Title: Data and Information Management

Catalog Description

This course provides the students with an introduction to the core concepts in data and information management. It is centered around the core skills of identifying organizational information requirements, modeling them using conceptual data modeling techniques, converting the conceptual data models into relational data models and verifying its structural characteristics with normalization techniques, and implementing and utilizing a relational database using an industrial-strength database management system. The course will also include coverage of basic database administration tasks. In addition to developing database applications, the course helps the students understand how large-scale packaged systems are highly dependent on the use of DBMSs. Building on the transactional database understanding, the course also provides an introduction to data and information management technologies that provide decision support capabilities under the broad business intelligence umbrella.

Learning Objectives

Students will be able to:

- Understand the role of databases and database management systems in managing organizational data and information.
- Understand the historical development of database management systems and logical data models.
- Understand the role of information requirements specification processes in the broader systems analysis & design context.
- Understand the basic approaches to data modeling techniques and be able to provide a comparison between these techniques (i.e. object-oriented data modeling, generic data modeling, semantic data modeling, etc.)
- Use at least one conceptual data modeling technique (such as entity-relationship modeling) to capture the information requirements for an enterprise domain.
- Link to each other the results of data/information modeling and process modeling.
- Design high-quality relational databases.
- Understand the purpose and principles of normalizing a relational database structure and to design a relational database so that it is at least in 3NF.
- Implement a relational database design using an industrial-strength database management system, including the principles of data type selection and indexing.
- Use the data definition, data manipulation, and data control language components of SQL in the context of one widely use implementation of the language.
- Perform simple database administration tasks.
- Understand the concept of database transaction and apply it appropriately to an application context.
- Understand the basic mechanisms for accessing relational databases from various types of application development environments.
- Understand the role of databases and database management systems in the context of enterprise systems.

- Understand the difference between on-line transaction processing (OLTP) and on-line analytic processing (OLAP), and the relationship between these concepts and business intelligence, data warehousing and data mining.
- Create a simple data warehouse ("data mart").
- Understand how structured, semi-structured, and unstructured data are all essential elements of enterprise information and knowledge management. In this context, the students will learn the principles of enterprise search.

Topics

- Basic File Processing Concepts
- Database Approach
- Types of Database Management Systems
- Conceptual Data Model
    - Entity-relationship model
    - Object-oriented data model
    - Specific modeling grammars
    - Generic data modeling (i.e. natural language)
    - Semantic data modeling
- Logical Data Model
    - Hierarchical data model
    - Network data model
    - Relational data model
        - Relations and relational structures
        - Relational database design
            - Mapping conceptual schema to a relational schema
            - Normalization
            - Anomalies
- Physical Data Model
    - Indexing
    - Data types
- Database Languages
    - SQL: DDL, DML, and DCL
    - Using XML to represent data
- Data and Database Administration
- Data Views
    - Virtual views (i.e. declaring views, query views, etc.)
    - Modifying views (i.e. Triggers)
- Transaction Processing
- Using a database management system from an application development environment
- Use of database management systems in an enterprise system context
- Business intelligence
    - On-line analytic processing
    - Data warehousing
    - Data mining
    - Enterprise search

Discussion

- The course still has a strong focus on traditional data management: conceptual data modeling (using ER modeling as the primary technique), logical data modeling using the relational data model (including ER – relational conversion and normalization), and physical database implementation and manipulation using SQL.
- It is essential that the information requirements specification processes are firmly linked to the organizational SA&D processes and that students understand the role of conceptual data modeling as an integral part of the process of making sense of the domain.
- The focus on the physical data model and the DBA-level work on database implementation has been reduced to give more time on improved understanding of the role of databases in the enterprise application context and various business intelligence topics, including enterprise search. Still, the students should understand the basic nature of the DBA tasks and be able to make intelligent decisions regarding DBMS choice and the acquisition of DBA resources.
- It is critically important that the students will fully understand how dependent various large-scale packaged systems (including ERP systems) are on relational databases and how strongly success in maintaining them and in supporting their use in organizations is dependent on understanding data structures and data manipulation with SQL.
- The course should provide a practical understanding of how relational databases are used to support web-based applications.
- Coverage of relational algebra and calculus can be added to the course material to provide an in depth understanding of relational databases.

