**ENTERPRISE TECHNOLOGY INTEGRATION (ETI)**

**ETI 435: Enterprise Analytics**

3 Credits

Analytics and big data, enabling analytics through information technology, ROI in analytics, leveraging proprietary data for analytical advantage, analytics on the web, analytics of online engagement, applying analytics at production scale, predictive analytics in the cloud, analytical technology and the business user, using analytics for improved organization performance, organizing analysts, engaging analytical talent, analytics governance, and building a global analytical capability, and analytics case studies in healthcare, manufacturing, HR, financial services, etc.

**Enforced Prerequisite at Enrollment:** IST 210 and ETI 461

**ETI 461: Database Management and Administration**

3 Credits

A study of advanced topics in database management systems that are fundamental to effective administration of modern enterprise information systems. The objective of the course is to enable a student to comprehend a range of issues in modern database management and administration. The students will learn advanced SQL, database system development lifecycle topic that include: database planning, requirements and design, database selection and application design, prototyping, implementation, testing operational and maintenance; database performance tuning concepts, monitoring the system for improved performance, and DBMS performance tuning; database transaction management covering transactions and the ACID properties, concurrency control techniques, and database recovery management; query processing and optimization techniques via query decomposition and optimization options; introduction to distributed processing and distributed database concepts, components and characteristics of DDBMS, and distributed database design; web connectivity technologies and XML; introduction to Business intelligence and data warehouses; introduction to Big data, NOSQL and cloud databases; and database security and database administration.

**Enforced Prerequisite at Enrollment:** IST 220 and IST 242 and ETI 461

**ETI 463: Distributed Database Management Systems**

3 Credits

Today, most enterprise databases are no longer a centralized data store that is accessed by thousands of users from multiple locations which may be globally situated. These databases are typically web-based and distributed across multiple sites for availability, low latency and better reliability. This course exclusively focuses on the design and system issues related to such distributed database systems. An initial review of relational DBMS is required in the first week of the course. Students will learn the architectural options and design issues and choices for DDBMSs. Design considerations include fragmentation alternatives (vertical or horizontal), fragment allocation and the data directory. Database integration covers at schema matching, integration, and mapping. Data cleaning is also studied under database integration. Processing distributed queries is challenging, and this topic is studied next by first trying to understand the query processing problem followed by: the objectives for query processing, characterization of query processors and layers of query processing. Query decomposition and localization of distributed data is then studied. The next issue is the problem of optimizing distributed queries using various techniques such as centralized query optimization, join ordering in distributed queries and distributed query optimization using dynamic, static. Semi-join and hybrid approaches. The ACID properties of transactions are studied and different types of transactions, and this is followed by distributed concurrency control using techniques such as locking, timestamps, and optimistic concurrency control algorithms. Deadlock management which is a problem in concurrency control schemes is also studied. Distributed reliability to address failures in DDBMSs is addressed by studying local reliability protocols, distributed reliability protocols, dealing with site failures and network partitioning. Data replication is an important aspect of reliability and various replication update management strategies and replication protocols are studied. Modern databases are web-based and this topic is considered next. Topics studied include: web searches, web querying and distributed XML processing. Many databases have moved to the cloud and cloud data management covers: cloud deployment models, service models SQL data services, and so on.

**Enforced Prerequisite at Enrollment:** IST 210 and ETI 461