I320D: Database Design

Spring 2022, UT Austin iSchool

Sam Oh, Visiting Professor at UT Austin, School of Information
Ph.D. in Information Transfer, Syracuse University 1995

Instructor Bio

Sam Oh is *an **Affiliate Professor at University** of **Texas at Austin and** University* of **Washington** (UW) **iSchool**. *His expertise* includes **Data modeling, Metadata** *and* **Ontology design, Data analytics,** *and* **Knowledge management**. He has consulted many companies and government sectors in Korea. He is a ***Past Chair** and Current Ambassador of* **iSchools, Inc**. He is the new Executive Director *of* **Dublin Core Metadata Initiative (DCMI)**. He is the Chair of TC46/SC4 (Technical Interoperability), also chaired both TC46/SC9 (Identification & Description) for 6 years and ISO/IEC JTC1 SC34 (Document Description *and* Processing Languages) for 9 years. He taught classes at the iSchools such as **Sungkyunkwan University, Syracuse, Pittsburgh,** *and* **University of North Carolina** *at* **Chapel Hill**.

Instructor	*Professor Sam Oh* *Email:samoh@g.skku.edu*
Textbook	*Rob, P. and Coronel, C. Database Systems: Design, Implementation and Management. 13th Edition. Thompson Course Technology.*
Course Objectives	*To learn how to develop sound database designs by data modeling methodology.* *To gain competent understanding of relational model and normalization process* *To understand structured query language (SQL), which is widely used for defining and manipulating the relational databases.* *To gain competent understanding of MySQL for database implementation.* *To gain competent understanding of a data cleaning tool.* *To gain competent understanding of SPARQL, which is widely used to extract what you need from big data sets.* *To learn how to extract data from Wikidata using SPARQL* *To learn how to do data analyses using Knime*
Teaching Methods	*The class will be conducted virtually and there will be formal lectures, student presentations, and cyber lab hours.*
Course Outline	*Entity-Relationship Model* *SQL: Structured Query Language* *Relational Model* *Normalization* *Database Implementation using MySQL* *Data Cleaning and analyses using Knime* *SPARQL and Wikidata*
References	*Fidel, Raya. Database Design for Information Retrieval: A Conceptual Approach. John Wiley and Son, Inc. ISBN: 0-471-82786-X.* *Elmasri, R. and Navathe, S. Fundamentals of Database Systems. Benjamin and Cummings. The Latest Edition.* *Jack Park and Sam Hunting. XML Topic Maps. Addison-Wesley.*
Grading	*Grading scale: The final grade will be given using Plus and Minus.* *Grading criteria: Class Participation (10%), ER Modeling (20%), SQL (10%), Normalization (10%), Data cleaning (10%), Database Implementation (20%), and Data Analytics and Visualization (20%),* *There will be no exams in this class, but only assignments.*

Database Design: Class Schedules
Class Time: Tue, Thu 5:00 – 6:30 PM Austin (Tue, Thu 3:00-4:30 PM PST | Wed, Fri 7:00-8:30 AM KST)

Session	Date	Topic	Reading	Assignment Due
M1	1/18	Course Overview Technology Trends		
M2.1	1/20	Entity-Relationship Model (ER) 1	Chapter 4	
M2.2	1/25	ER Model 2	Chapter 1	
M2.3	1/27	ER Model 3 ER-Diagraming Exercises: #1 and #2	Chapter 2	
M2.4	2/1	ER Model 4 Checking A1 Answer (ERD) ER-Diagraming Exercises: #3	Chapter 5	A1: Products and Parts DB (ERD) 5:00 PM Austin Time
M2.5	2/3	ER Model 5 ER-Diagraming Exercises: #4 and #5		
M3.1	2/8	Relational Model 1 Checking A2 Answer (ERD)	Chapter 3	A2: Book Printing DB (ERD) 5:00 PM Austin Time
M3.2	2/10	Relational Model 2 ER-Diagraming Exercises: #6		
M4.1	2/15	Normalization 1 Checking A3 Answer (ERD)	Chapter 6	A3: Shoes Store DB (ERD) 5:00 PM Austin Time
M4.2	2/17	Normalization 2 ER-Diagraming Exercises: #7		-
M4.3	2/22	Normalization 3 Normalization Exercises: #1-#3 Checking A4 Answer (ERD)		A4: Student Grade DB (ERD) 5:00 PM Austin Time
M4.4	2/24	Normalization 4 Normalization Exercises: #4-#6 ER-Diagraming Exercises: #8		
M5.1	3/1	Structured Query Language (SQL) 1 ER-Diagraming Exercises: #9	Chapter 7	
M5.2	3/3	Structured Query Language (SQL) 2 Checking A5 Answer (Normalization)		A5: Normalization 5:00 PM Austin Time
M5.3	3/8	Structured Query Language (SQL) 3 ER-Diagraming Exercises: #10	Chapter 8	
M5.4	3/10	Structured Query Language (SQL) 4 ER-Diagraming Exercises: #11		
	3/15	Spring break (No class)		
	3/17	Spring break (No class)		
M6.1	3/22	Database Implementation 1 (MySQL) Checking A6 Answer (SQL)		A6: SQL 5:00 PM Austin Time

M6.2	3/24	Database Implementation 2 (MySQL) ER-Diagraming Exercises: #12		
M6.3	3/29	Database Implementation 3 (MySQL) (LAB) Checking A7 Answer (ERD)		A7: Video Store DB (ERD) 5:00 PM Austin Time
M7	3/31	Wiki Family (Wikidata, MediaWiki, Wikibase,DBpedia) 1		
M8.1	4/5	SPARQL 1		A9: Database Implementation 5:00 PM Austin Time
M8.2	4/7	SPARQL 2		
M9.1	4/12	Data Cleaning (Knime) 1		
M9.2	4/14	Data Cleaning (Knime) 2		
M9.3	4/19	Data Cleaning (Knime) 3		
M9.4	4/21	Data Cleaning (Knime) 4 ER-Diagraming Exercises: #13		
M10.1	4/26	Data Analytics using Knime 1		A10: Data Cleaning 5:00 PM Austin Time
M10.2	4/28	Data Analytics using Knime 2 ER-Diagraming Exercises: #14		
M10.3	5/3	Data Analytics using Knime 3 (LAB)		
M11	5/5	Presenting Final Projects		A11: Data Analytics and Visualization 5:00 PM Austin Time

May 9-10, 15   *Monday-Tuesday, Sunday: No-class days.*

May 11-14, 16-17   *Wednesday-Saturday, Monday-Tuesday: Spring semester final examinations except in the School of Law.*