

DTS207TC Database Development and Design

Lecture 1 **Introduction & Review of 106**

Di Zhang, Autumn 2025

Learning Outcomes

- Second round of learning: advanced content of database
 - A. Identify and apply the principles underpinning transaction management within DBMS.
 - B. Demonstrate an understanding of advanced SQL topics.
 - C. Illustrate the issues related to Web technologies and DBMS and XML as a semi-structured data representation formalism.
 - D. Identify the principles underlying object-relational models.
 - E. State the main concepts in data warehousing and data mining.

More content beyond...

- Extensions for Computer Science:
 - Hadoop
 - Physical Storage
 - Information Retrieval

Teaching Team

- Di Zhang:
 - Over 15 years of database and data mining experience, proficient in Oracle, PostgreSQL, Hadoop, Spark, etc.
 - This is the second consecutive year that I have served as the ML instructor for this course.
- Co-teachers: (Lab)
 - Affan: ML for Class 106 (Database 1)
 - Hengyan Liu: Former ML for this course
 - Hejia Qiu: Co-teaching this course for the second consecutive year
 - Xiaowu: New teacher
- TA:
 - Biwen Meng: doctoral student
 - Fei Ren: master student
 - Haoran Zhu: doctoral student
 - Zhepeng Li: doctoral student

Teaching Plan

Lectures	Chapters in Textbook	Labs	Teacher	TA	CW
1	1-3	Installation of PostgreSQL, and Basic SQL	Affan	Biwen Meng	\
2	4.1-4.5	Intermediate SQL	Affan	Zhepeng Li	C1-Q1
3	5.1-5.4	Advanced SQL	Affan	Haoran Zhu	C1-Q2
4	8.1-8.2,30	Json,XML	Hejia	Biwen Meng	C1-Q3
5	9.1-9.2,9.4-9.7	ORM,HTML	Hejia	Zhepeng Li	C1-Q4
6	10.1-10.3	Hadoop	Hejia	Haoran Zhu	\
7	11.1-11.3	Data Warehousing	Hengyan	Biwen Meng	C1-Q5
8	12.1-12.5	Physical Storage	Hengyan	Fei Ren	\
9	13.1-13.5	Data Storage	Xiaowu	Zhepeng Li	C2-Q6
10	14.1-14.4	Indexing	Xiaowu	Haoran Zhu	C2-Q7
11	17.1-17.7,17.10	Transaction	Hengyan	Fei Ren	C2-Q8
12	31.1-31.7	Information Retrieval	Xiaowu	Fei Ren	\
13	Review	Q/A	Di	TBD	\

CW and DDL

- 2 CW
 - Under internal review. They will be released after pass.
- DDL
 - 23:59, 7th Nov (Friday)
 - 23:59, 12th Dec (Friday)

Xi'an Jiaotong-Liverpool University 2025-2026 Academic Calendar (STAF)

Semester	Month	Mon	Tues	Wed	Thur	Fri	Sat	Sun	Descriptio
Semester 1	September 2025	1	2	3	4	5	6	7	Week 0: Registration and induc
		8	9	10	11	12	13	14	Week 1: Teaching starts
		15	16	17	18	19	20	21	Week 2
		22	23	24	25	26	27	28	Week 3
	October 2025	29	30	1	2	3	4	5	National Day; University clousr
		6	7	8	9	10	11	12	Week 4; Mid-Autumn Festival
		13	14	15	16	17	18	19	Week 5
		20	21	22	23	24	25	26	Week 6
	November 2025	27	28	29	30	31	1	2	Week 7
		3	4	5	6	7	8	9	Week 8
		10	11	12	13	14	15	16	Week 9
		17	18	19	20	21	22	23	Week 10
	December 2025	24	25	26	27	28	29	30	Week 11
		1	2	3	4	5	6	7	Week 12
		8	9	10	11	12	13	14	Week 13
		15	16	17	18	19	20	21	Reading Week
	January 2026	22	23	24	25	26	27	28	Reading Week; Christmas
		29	30	31	1	2	3	4	Examination days; New Year
		5	6	7	8	9	10	11	Examination days
		12	13	14	15	16	17	18	Marking
		19	20	21	22	23	24	25	Assessment moderation
		26	27	28	29	30	31	1	Assessment moderation

Last year's feedback...

- “CW is too difficult.”
 - The difficulty has been lowered this year.
 - AI has been banned again.
- “The connection between the coursework and the assignments isn't strong.”
 - One-on-one, strictly aligned.

- Textbook:
 - <https://db-book.com/toc-dir/toc.pdf>
- Video:
 - https://search.bilibili.com/all?vt=16822812&keyword=Database%20System%20Concepts&from_source=webtop_search&spm_id_from=333.1007&search_source=5
 - https://www.youtube.com/results?search_query=database+system+concepts

How to ask questions

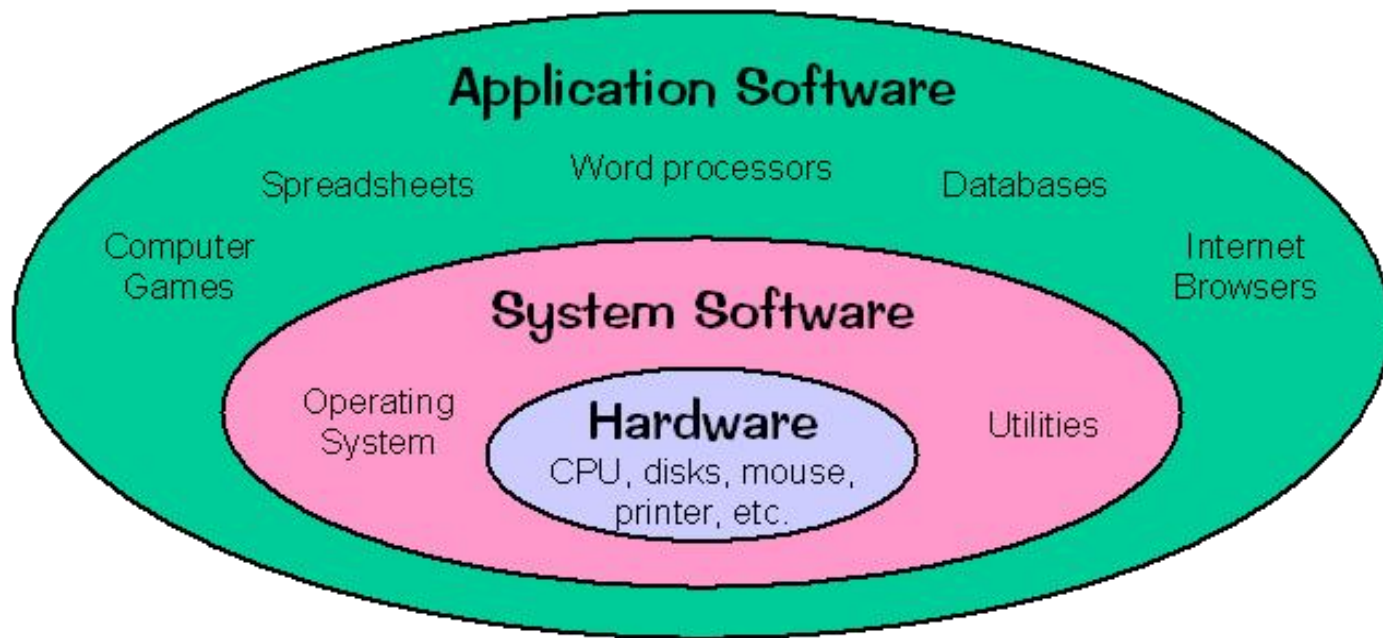
- **If** questions about your private issues (leave of absence, late submission, etc.), then
 - Send email to the module leader
- **Elseif** questions about technical issues, then
 - Ask questions on forum, to make everyone see them (after reviewed by ML)
 - **If** questions about knowledge in lectures or requirement of assessments, then
 - Questions will be answered
 - **Elseif** questions are answered, duplicate, or directly linked to the solution
 - Questions will be redirected or omitted

- Construct like building blocks...

	Theory	Practice
Application	Optimization geometry	Machine Learning Computer Graphics Internet
Software	Discrete Mathematics: Turing Machine (doodle) + Von Neumann architecture	Network Database Compiler OS
Hardware	Electromagnetism	Digital Circuit (Clock Speed) paper/mechanical/...

System Software

- System software is software designed to provide a platform for other software.



Q: Can you give some examples of system software?

OS: Conceptualize Hardware

Functionality	Windows/Linux
Process:	Queue/Time Sharing (for realtime), nice command
Isolation Paralellization	Test and Set-> Mutex -> Threading/Process -> Queue/...
Storage:	Addr./Blocks -- similiar to a city
Memory Disk	NTFS/ext4
Interface:	
GUI Progamming	cmd/Bash
I/O:	
Import/Export	Mouse/Keyboard Monitor

Database: Conceptualize Bits



Xi'an Jiaotong-Liverpool University

西交利物浦大学

Functionaility	PostgreSQL
Process: Isolation Paralellization	Transaction
Storage: Memory Disk	Relational Tables
Interface: GUI Progamming	SQL
I/O: Import/Export	Import csv ODBC/Drivers

- DBMS/DBA
- OLTP/OLAP
- DDL/DML/SQL
- E-R/Relational
- 1NF/2NF/3/NF

- A super powerful "warehouse manager".



Analog

	warehouse	Database
Physical level	It's like putting things in boxes and cabinets; the physical layer controls how the cabinet is placed.	It controls where and how data resides.
Logical Level	Defines the structure and relationships of data.	Catalog system that tells you the contents and categories of each box.
Logical Level	The view layer is the data that users see. Users do not need to know how the data is stored or the detailed structure of the data. They only need to see the information they need.	Combinable room and door lock settings.

- OLTP (Online Transaction Processing): Handles daily business operations.
- OLAP (Online Analytical Processing): Analyzes and summarizes data to identify patterns and trends.
- Analogy:
 - The CPU (Central Processing Unit) is like an all-around "transaction processor," adept at handling a wide variety of small tasks at high speed.
 - The GPU (Graphics Processing Unit) is like a professional "data analyst," skilled at handling high-volume, repetitive tasks, particularly graphics and data analysis.

- SQL (Structured Query Language) is a language used to "chat" with a database. Just like chatting with a friend, you use SQL to tell the database what you want to do, such as store data, retrieve data, or modify data.
- DDL (Data Definition Language) is a part of SQL specifically used to define and modify database structures. Just like when you build a house, DDL is the tool used to design and modify the blueprints.

- Specific examples:

Create a table: `CREATE TABLE table_name (field1 type, field2 type);`

Modify a table: `ALTER TABLE table_name ADD COLUMN new_column type;`

Drop a table: `DROP TABLE table_name;`

- DML (Data Manipulation Language)
 - DML is part of SQL and is specifically used for data manipulation, such as inserting, querying, updating, and deleting. Just like moving or arranging things around a room, DML is a tool for manipulating data.

- The E-R model is a graphical tool used to design and describe relationships between data. It helps us understand the structure of data and the connections between them.
 - Its elements include entities, attributes, and relationships.
- The relational model is a model used to store and manage data in a tabular format.
 - Its elements include tables, fields, and foreign keys.

Mapping from E-R to Relational

- The mapping from the E-R model to the relational model isn't unique.
 - Different levels of normalization (such as 1NF, 2NF, and 3NF) result in different table structure designs.
- Reflection: Does the E-R model necessarily map to the relational model?