CMSI 620T
DATABASE SYSTEMS GRADUATE TUTORIAL
https://dondi.lmu.build/fall2022/cmsi620

Fall 2022—Pereira 140
MW 1:45–3:25pm, 3 semester hours
Office Hours: MTW 4:30–6pm,
    or by appointment *(don't hesitate to ask!)*

John David N. Dionisio, Ph.D.
email: *dondi@lmu.edu*
Doolan 102; (310) 338-5782

Objectives and Outcomes

This course introduces current research directions in the computer science subfield of *databases*, which is concerned with the theory, design, and implementation of systems that manage large amounts of data. Long after the course concludes, my hope is that you will:

1. **Have a sense for what comprises academic research in the computer science subfield of database systems.**

2. **Have some familiarity with research questions that have been asked about certain database paradigms and models.**

3. **Be able to identify and possibly answer questions based on a given body of database research.**

In addition to the course-specific content, you are also expected to:

4. **Follow disciplinary best practices throughout the course.**

Prerequisites/Prior Background

In and of itself, this course expects some hands-on familiarity with existing database models and technologies. To supply or reinforce this background for students who do not have it yet, the course runs concurrently with CMSI 3520. As graduate students, you may choose how to engage this material for yourself such that you can credibly accomplish the objectives of this course.

Materials and Texts

Foundational content is to be provided by the materials in CMSI 3520. Content specific to this course will consist of research-oriented publications that will either be provided by the instructor or discovered by yourselves and your classmates throughout the semester.

Course Work and Grading

Your final grade will be based on the percentage of the points you get for the following deliverables against the total number of possible points:

Initial research survey	50
Research paper directed draft 1	50
Research paper directed draft 2	50
Research paper directed draft 3	50
Final database research paper	200
Total	400 points

Percentages ≥ 90% get an A– or better; ≥ 80% get a B– or better; ≥ 70% get a C– or better. I may nudge grades upward based on qualitative considerations such as degree of difficulty, effort, class participation, time constraints, and overall attitude toward the course.

Term Portfolio: Journey to a Research Paper

Your accumulated assignments for the semester comprise the *term portfolio*—the final, definitive artifact that demonstrates the course's outcomes. It is how you show whether you have, indeed, accomplished the objectives of this course.

For this course, the term portfolio is a sequence of artifacts leading to a final research paper on a database topic of your choice. We will start with an initial research survey, seeded by the so-called "Seattle database report" which comes out every five years. The most recent version was published just this August, 2022. A copy of the report has been uploaded to Brightspace.

Yes you'll want to read it, although you may not understand it fully at first. It is recommended that you revisit it throughout the semester, as your foundational knowledge and practical experience of databases grows via the concurrent CMSI 3520 course content.

The target scope of the final research paper is 2000–3000 words, not including references. An accompanying proof-of-concept project is recommended but not necessarily required, depending on your final topic.

We will build up to the final version throughout the semester via "directed" drafts—i.e., starting from the initial survey, you will get individualized feedback on how to proceed, and you will aim to fulfill those next steps in the next draft, totaling three (3) before the final version. Scores will be based on the degree to which feedback is addressed from one draft to the next.

Each research paper milestone is to be treated and presented like an assignment. Its number is its due date in *mmdd* format, and it is always due by 11:59:59.999pm of that date. Unless otherwise indicated, the documents are to be uploaded to Brightspace in their designated "folders."

Role of Undergraduate Content

Looking at research papers alone can be quite abstract and limited until you have logged direct, hands-on time with the kinds of systems that those research papers talk about. That is the role of the CMSI 3520 content. You are not obligated to do the undergraduate course work (i.e., you will not be graded for it), but working on it to some degree will be very beneficial especially if you have not had a lot of exposure to databases yet.

Workload Expectations

In line with the *LMU Credit Hour Policy*, the workload expectation for this three-credit-hour course is a minimum of $3 \times 3 = 9$ hours of work by an average student per week, including the time that we spend together in the classroom.

Attendance

Attendance at all CMSI 3520 sessions is expected, but not absolutely required. If you must miss class, it is your responsibility to notify me about this and keep up with the course.

The last day to add or drop a class without a grade of W is September 2. The withdrawal or credit/no-credit deadline is November 4.

Academic Honesty

Loyola Marymount University is a community dedicated to academic excellence, student-centered education, and the Jesuit and Marymount traditions. As such, the University expects all members of its community to act with honesty and integrity at all times, especially in their academic work. Academic honesty requires that all members of the LMU community act with integrity, respect their own intellectual and creative work as well as that of others, acknowledge sources consistently and completely, act honestly during exams and on assignments, and report results accurately. As an LMU Lion, by the Lion's Code, you are pledged to join the discourse of the academy with honesty of voice and integrity of scholarship.

Academic dishonesty will be treated as an extremely serious matter, with serious consequences that can range from receiving no credit for assignments/tests to expulsion. It is never permissible to turn in any work that has been copied from another student or copied from any source (including the Internet) without properly acknowledging/citing the source. It is never permissible to work on an assignment, exam, quiz, or any project with another person unless your instructor has indicated so in the written instructions/guidelines. It is your responsibility to make sure that your work meets the standard of academic honesty set forth in the "Academic Honesty Policy" found at:

https://academics.lmu.edu/honesty

Special Accommodations

The Disability Support Services (DSS) Office offers resources to enable students with ADD/ADHD; physical, learning, and psychiatric disabilities; and those on the autism spectrum to achieve maximum independence while pursuing their educational goals. Staff specialists interact with all areas of the University to eliminate physical and attitudinal barriers. Students must provide documentation for their disability from an appropriate licensed professional. Services are offered to students who have established disabilities under state and federal laws. DSS personnel also advise students, faculty, and staff regarding disability issues. Students who need reasonable modifications, special assistance, academic accommodations or housing accommodations should direct their request to the DSS Office as soon as possible. All discussions will remain confidential. The DSS Office is located on the 2nd floor of Daum Hall and may be reached by email at *dsslmu@lmu.edu* or phone at (310) 338-4216. Please visit *http://www.lmu.edu/dss* for additional information.

Topics and Important Dates

The following topics are aligned with the topic sequence for CMSI 3520. As mentioned earlier, this is meant to be foundational, so that the ideas you encounter in the literature can be made concrete and real for you. The full coverage for this course consists of these foundational topics *along with* with articles and publications found through your literature review and search outside of class. Specifics may change as the course progresses. University dates (italicized) are less likely to change.

August	Overview; introduction to data sets and database application tiers; hands-on with a database fiddle
September	In-memory database code-along; file databases; database server setup, initialization, startup and shutdown, and general use; introduction to relational databases and SQL
September 2	*Last day to add or drop a class without a grade of W*
September 5	*Labor Day; no class*
October	Advanced SQL: aggregation, subqueries, JSON fields; document-centric databases
November	Graph databases; relational database theory: algebra and calculus; functional dependencies and normalization; integrity constraints; indexing and transactions
November 4	*Withdraw/credit/no-credit deadline*
November 23–25	*Thanksgiving; no class*
December	Database research paper work sessions; miscellaneous topics *(time permitting)*
December 12	*Final database research paper due*

You can view my class calendar and office hour schedule in any iCalendar-savvy client. Its subscription link can be found on the course web site (it's too long to provide in writing).

Tentative Nature of the Syllabus

This syllabus and its contents are subject to revision; students are responsible for any changes or modifications announced or distributed in class, emailed to students' LMU Lion accounts, or posted on LMU's course management system.

Course Evaluations

Student feedback provides valuable information for continued improvement. All students are expected to fairly and thoughtfully complete a course evaluation for this course. This semester, course evaluations will be administered online through the Blue™ evaluation system. You will receive an email notification at your Lion email address when the evaluation form is available. You may also access the evaluation form on Brightspace (*https://brightspace.lmu.edu*) dashboard during the evaluation period. Your responses will be anonymous and will not be linked to you in any way.

Course Outcomes

1 Have a sense for what comprises academic research in the computer science subfield of database systems.

"Knowing about databases" is the purview of the concurrent undergraduate content—that is the beginning. After learning what is known, we want to shift toward what is <u>not</u> known: that is where research is taking place.

2 Have some familiarity with research questions that have been asked about certain database paradigms and models.

This is the classic "standing on the shoulders of giants" idea: you will want to emerge from this course with an awareness of what has been tackled before, and how those have been answered. These answers take their most concrete and applicable forms by way of existing, operational database management systems, three of which will be introduced to you through the concurrent undergraduate content.

3 Be able to identify and possibly answer questions based on a given body of database research.

This is the culmination of the prior two objectives: when you are familiar with the areas that are investigated under the umbrella of database systems and know some questions that have been asked in the past, the hope is that you will then be able to do this on your own.

4 Follow disciplinary best practices throughout the course.

4a	*Communicate your ideas as clearly as possible.*	Scholarly discourse is still discourse—and thus, clarity of thought can only be conveyed through clarity of word.
4b	*Communicate your ideas as concisely as possible.*	Clarity frequently accompanies conciseness, but sometimes will feel at odds with it. Avoid embellishment and unnecessary detail; stay organized so that you retain sharp focus on what you need to get across at any given moment.
4c	*Polish matters.*	Well-executed work—minimal or no errors, consistent formatting, adherence to style manuals, technically and mechanically correct writing, comprehensible diagrams, to name a few—has an impact on how your readers or audience perceive your ideas. Polish conveys your capacity for detail and your work ethic, and when it is missing, it can negatively distract your readers or audience from what you truly want to convey.
		A concrete way toward polish is to proofread your work multiple times. Asking "another set of eyes" is also a good way to catch issues. Heed constructive feedback not only in terms of content but also in terms of execution—in this course, this is actually a major contributor to your final grade!
4d	*Use available resources and documentation to find required information.*	The need to look things up never goes away. Remember also that the course instructor counts as an "available resource," so this outcome includes asking questions and using office hours.
4e	*Use version control effectively.*	Although this might not apply to your work in this particular course, for the field in general it is good to know that, in addition to simply using version control correctly, effective use also involves appropriate time management, commit frequency, and descriptive commit messages.
4f	*Meet all designated deadlines.*	