



THE UNIVERSITY OF WINNIPEG

APPLIED COMPUTER SCIENCE

Course Number: GACS-7102
Course Name: Web and document databases
Course Webpage: <https://www.acs.uwinnipeg.ca/ychen2/WebDocument.html>

Instructor Information

Instructor: Dr. Yangjun Chen
E-mail: y.chen@uwinnipeg.ca
Office Hours: 4:00 pm - 5:00 pm on Mondays and Wednesdays
12:00 noon – 3:00 pm Friday

Class meeting time: Mondays/Wednesdays 11:30 - 12:45 pm location: 3D03

Important Dates

- | | |
|---|----------------------------|
| 1. First Class: | Monday, January 6, 2025 |
| 2. Reading Week (no classes): | February 16-22, 2025 |
| 3. Midterm Exam: | Monday, March 03, 2025 |
| 4. Final Withdrawal Date w/o academic penalty*: | Friday, March 14, 2025 |
| 5. Last Class: | Wednesday, April 2, 2025 |
| 6. Final Exam: | replaced by course project |
| 7. University closures: Louis Riel Day | Monday, February 17, 2025 |
| Good Friday | Friday, April 18, 2025 |

*A minimum of 20% of the work on which the final grade is based will be evaluated and available to the student before the voluntary withdrawal date.

Course Objectives / Learning Outcomes

In this course, students will gain a good understanding and knowledge of research issues associated with two types and databases. In particular, students will study basic theoretic issues of web and document databases: system architectures, XML data storage and data compression, data retrieval and twig matching, data stream system, as well as the search engine architecture. Another specific methodology related to Graph databases will also be discussed.

Evaluation Criteria

1. Assignments (24%)
 - 3 assignments, worth 8% each
 - Individual due dates will be posted on ion.uwinnipeg.ca/~ychen2/advancedAD.html
 - Assignments will be accepted up to 1 day late with a 25% penalty
2. Midterm Test (26%)
 - During the regular class time (see Important Dates)
4. Course project (50%)
 - Algorithm implementation in C++
 - Experiments
 - Report

Students should contact the instructor as soon as possible if extenuating circumstances require missing a test or examination. A medical certificate from a practicing physician may be required before any adjustments are considered.

Test / Exam Requirements

- Photo ID is required for the final exam.
- The use of computers, calculators, phones, or other electronic devices is not permitted during exams.
- Midterm and final exams are closed-book.

Students should contact the instructor as soon as possible if extenuating circumstances require missing a lab, assignment, test or examination. A medical certificate from a practicing physician may be required before any adjustments are considered.

Students with documented disabilities, temporary or chronic medical conditions, requiring academic accommodations for tests/exams (e.g., private space) or during lectures/laboratories (e.g., note-takers) are encouraged to contact Accessibility Services (AS) at 204-786-9771 or accessibilityservices@uwinnipeg.ca to discuss appropriate options. All information about a student's disability or medical condition remains confidential.
<https://www.uwinnipeg.ca/accessibility-services>.

Students may choose not to attend classes or write examinations on holy days of their religion, but they must notify their instructors at least two weeks in advance. Instructors will then provide opportunity for students to make up work examinations without penalty. A list of religious holidays can be found in the 2024-25 Undergraduate Academic Calendar online at <http://uwinnipeg.ca/academics/calendar/docs/important-notes.pdf>

Final Letter Grade Assignment

Historically, numerical percentages have been converted to letter grades using the following scale. However, instructors can deviate from these values based on pedagogical nuances of a particular class, and final grades are subject to approval by the Department Review Committee.

A+	90 – 100%	B+	75 – 79%	C	60 – 64%
A	85 – 89 %	B	70 – 74%	D	50 – 59%
A-	80 – 84%	C+	65 – 69%	F	below 50%

Required Text Book / Reading List

- Class Notes will be available on ion.uwinnipeg.ca/~ychen2/WebDocument.html
- Reference book:
 - 1) *Database Systems (the complete book)*, 2nd Ed. by Carcia-Molina, Ullman, Widom, Pearson Prentice Hall, 2009.
 - 2) *Fundamentals of Database Systems*, R. Elmasri and S. B. Navathe, 5th, 6th, 7th edition, Addison Wesley.
 - 3) *Introduction to Algorithms*, T.H. Corman, C.E. Leiserson, R.L. Rivest, and C. Stein, 2nd, 3rd edition.

Prerequisite Information* (This information can be found in the UW General calendar)

- Consent of the Department Graduate Program Committee Chair or Instructor.
- Make sure that you have the necessary prerequisites to take this course. If you have not successfully completed the above listed course(s), it is in your interest to drop the course.

Regulations, Policies, and Academic Integrity

Students are encouraged to familiarize themselves with the Academic Regulations and Policies found in the University Academic Calendar at:

<https://uwinnipeg.ca/academics/calendar/docs/regulationsandpolicies.pdf>

Particular attention should be given to subsections 8 (Student Discipline), 9 (Senate Appeals) and 10 (Grade Appeals).

Avoiding Academic Misconduct: Academic dishonesty is a very serious offense and will be dealt in accordance with the University's policies.

Detailed information can be found at the following:

- Academic Misconduct Policy and Procedures:
<https://www.uwinnipeg.ca/policies/docs/policies/academic-misconduct-policy.pdf> and
<https://www.uwinnipeg.ca/policies/docs/procedures/academic-misconduct-procedures.pdf>
- About Academic Integrity and Misconduct, Resources and FAQs:
<https://library.uwinnipeg.ca/use-the-library/help-with-research/academic-integrity.html>

Uploading essays and other assignments to essay vendor or trader sites (filesharing sites that are known providers of essays for use by others who submit them to instructors as their own

work) involves “aiding and abetting” plagiarism. Students who do this can be charged with Academic Misconduct.

Academic Integrity and AI Text-generating Tools: Students must follow principles of academic integrity (e.g., honesty, respect, fairness, and responsibility) in their use of material obtained through AI text-generating tools (e.g., ChatGPT, Bing, Notion AI, copilot). Use of AI Tools is prohibited in this course: students may face an allegation of academic misconduct if using them to do assignments, paper presentations, reports, and/or final project.

Non-academic misconduct: Students are expected to conduct themselves in a respectful manner on campus and in the learning environment irrespective of platform being used. Behaviour, communication, or acts that are inconsistent with a number of UW policies could be considered “non-academic” misconduct. More detailed information can be found here:

- Respectful Working and Learning Environment Policy
<https://www.uwinnipeg.ca/respect/respect-policy.html>,
- Acceptable Use of Information Technology Policy
- <https://www.uwinnipeg.ca/policies/docs/policies/acceptable-use-of-information-technology-policy.pdf>
- Non-Academic Misconduct Policy and Procedures:
<https://www.uwinnipeg.ca/policies/docs/policies/student-non-academic-misconduct-policy.pdf> and <https://www.uwinnipeg.ca/policies/docs/procedures/student-non-academic-misconduct-procedures.pdf>

Copyright and Intellectual Property: Course materials are the property of the instructor who developed them. Examples of such materials are course outlines, assignment descriptions, lecture notes, test questions, and presentation slides—irrespective of format. Students who upload these materials to filesharing sites, or in any other way share these materials with others outside the class without prior permission of the instructor/presenter, are in violation of copyright law and University policy. Students must also seek prior permission of the instructor/presenter before, for example, photographing, recording, or taking screenshots of slides, presentations, lectures, and notes on the board. Students found to be in violation of an instructor’s intellectual property rights could face serious consequences pursuant to the Academic Misconduct or Non-Academic Misconduct Policy; such consequences could possibly involve legal sanction under the Copyright Policy:

<https://copyright.uwinnipeg.ca/basics/copyright-policy.html>

Privacy

Students have rights in relation of the collecting of personal data the University of Winnipeg

- Student Privacy: <https://www.uwinnipeg.ca/privacy/admissions-privacy-notice.html>
- Zoom Privacy: <https://www.uwinnipeg.ca/privacy/zoom-privacy-notice.html>

Class Cancellation, Correspondence with Students and Withdrawing from Course

When it is necessary to cancel a class due to exceptional circumstances, the course instructor will make every effort to inform students via UWinnipeg email and Nexus.

Students are reminded that they have a responsibility to regularly check their UWinnipeg e-mail addresses to ensure timely receipt of correspondence from the University and/or the course instructor.

Please let the course instructor know if you plan on withdrawing from the course. Note that withdrawing before the VW date does not necessarily result in a fee refund.

Topics to be covered (tentative)

1. Database Basics
2. Relational data model and relational algebra
3. SQL language and JDBC
4. Normalization theory and lossless join
5. Static hashing and dynamical hashing
6. Index techniques
7. Semi-structured data model
8. Programming language for XML
9. Database systems and internet
10. Evaluation of tree pattern queries
11. Evaluation of reachability queries in graph databases
12. Data mining: finding most popular package
13. Graph databases

Lecture notes

Projects (sample topics):

- (1) Implementing different strategies for query evaluation in XML document databases
 - (2) Implementing Hopcroft-Karp algorithm for maximum bipartite matching
- More projects will be announced.

Guidance to project reports:

1. Introduction (including the problem description, motivation – its significance and application in the computer engineering and industry)
2. Related work (describe some important techniques related to the problem to be addressed)
3. Main thrust (detailed description of the method, formal algorithm, analysis of computational complexities: time and space overhead)
4. Future work (discussion on the possible improvements, or possible extension)
5. Experiments (main data structures used for implementation, description of the data used for tests, test results: charts, histogram, or tables)
6. References

