

University of Texas at El Paso Department of Computer Science
CS 5364 Topics in Data Science – Web-based Data Integration - Spring 2021

Logistics

Instructors:

Natalia Villanueva Rosales, e-mail: nvillanuevarosales [at] utep.edu.

Class time: TR 9:00-10:20 am.

Room: LART 222

Office hours: TR 10:30 am-11:50 (Instructor Dr. Villanueva Rosales), and by appointment outside this time on Blackboard collaborate or face-to-face by request at CCSB 3.0505.

Note: Office hours times may change throughout the semester and will be announced to the class.

Email: Please use email as the main means to contact the instruction team, add [CS5364] in the prefix of the email, and use your UTEP student account. The instruction team will do their best to respond within 24-48 hours of receipt.

Discussion Board: If you have a question that you believe other students may also have, please post it in the [24/7 discussion board](#) inside of Blackboard. Please respond to other students' questions if you have a helpful response.

Announcements: Check the Blackboard announcements frequently for any updates, deadlines, or other important messages. It is recommended that you download the Blackboard Mobile App to receive notifications.

Course Catalog Description

Introduction to advanced concepts and algorithms in data science and their applications. Topics may include deep learning, speech processing, language processing, data integration, information retrieval, and information visualization.

Learning Outcomes

Course Outcomes

Divided into the following three broad levels of Bloom's taxonomy:

Level 1: Knowledge and Comprehension.

Level 1 outcomes are those in which the student has been exposed to the terms and concepts at a basic level and can supply basic definitions. The material has been presented only at a superficial level. Upon successful completion of the course, students will be able to:

- 1a. Describe and compare data models (e.g., Entity-Relationship model, ontologies, UML model) and how they are currently used for data management.
- 1b. Identify current challenges in data management.
- 1c. Identify current trends in data science research, both fundamental and in application domains.
- 1d. Identify current trends in data science research, both fundamental and in application domains.
- 1e. Describe technical solutions to the challenges in data sharing, integration and reuse of data.
- 1f. Identify the careers/roles associated with data scientists.

Level 2: Application and Analysis.

Level 2 outcomes are those in which the student can apply the material in familiar situations, e.g., can work a problem of familiar structure with minor changes in the details. Upon successful completion of the course, students will be able to:

- 2a. Demonstrate uses of explicitly stored metadata/schema associated with data on the web.
- 2b. Use current cyberinfrastructure for the sharing, integration or exchange of data with a focus on reproducible workflows in scientific domains.
- 2c. Demonstrate the ability to design applications that solve real-world problems that required integration of data.
- 2d. Demonstrate the ability to write a technical report.
- 2e. Demonstrate the ability to work in teams.
- 2f. Demonstrate the use data integration languages fundamentals.

Level 3: Synthesis and Evaluation.

Level 3 outcomes are those in which the student can apply the material in new situations. This is the highest level of mastery. Upon successful completion of the course, students will be able to:

- 3a. Design a high-level data integration model from a problem statement using standard notation and modeling principles.
- 3b. Design and implement a tool prototype that enables the integration of data from heterogeneous, decoupled data.
- 3c. Design and implement an interface proof-of-concept for a data-integration tool applying best practices for privacy, security, and reproducibility of results.

Grading

1. Exams, including one per module, final, and quizzes 50%. The final exam will be optional for those with an average of 80 or higher in module exams and quizzes.
2. Project deliveries including assignments and presentations 30%.
3. Active class participation, homework, and activities 20%.

Assignments for this course are assessed according to rubrics. Each assignment will be provided with the corresponding rubric.

Required Materials

A list of readings included scientific papers and book chapters available through the UTEP library will be provided by the instructor.

Technology Requirements: Course resources, assignments, and tests are delivered through the Blackboard learning management system. Ensure your UTEP e-mail account is working and that you have access to the Web and a stable web browser. Google Chrome and Mozilla Firefox are the best browsers for Blackboard. When having technical difficulties, update your browser, clear your cache, or try switching to another browser. Bring a device to the lectures that can access Blackboard (smartphone, tablet, or

laptop) and can scan a document or take a picture to complete and submit class exercises. Ensure you have access to a computer/laptop to install or update the list of software that will be provided in the course Blackboard shell and needed to complete assignments. This list will be provided at least a week before the software is needed.

Expectations and Resources

Standards of conduct

Scholastic Dishonesty: Any student who commits an act of scholastic dishonesty is subject to discipline. Scholastic dishonesty includes, but not limited to cheating, plagiarism, collusion, submission for credit of any work or materials that are attributable to another person. **Cheating** is copying from the test paper of another student. Communicating with another student during a test to be taken individually. Giving or seeking aid from another student during a test to be taken individually. Possession and/or use of unauthorized materials during tests (i.e. crib notes, class notes, books, etc.). Substituting for another person to take a test. Falsifying research data, reports, academic work offered for credit. Paying a person or company to complete coursework (i.e., contract cheating). **Plagiarism** is using someone's work in your assignments without the proper citations. Submitting the same paper or assignment from a different course, without direct permission of instructors. To avoid plagiarism, see: <http://sa.utep.edu/osccr/wp-content/uploads/sites/8/2012/09/AvoidingPlagiarism.pdf>. **Collusion** is unauthorized collaboration with another person in preparing academic assignments.

All suspected violations of academic integrity at The University of Texas at El Paso must be reported to the [Office of Student Conduct and Conflict Resolution \(OSCCR\)](#) for possible disciplinary action. Students are expected to comply with student standards of conduct [HOOP: Student Conduct and Discipline](#).

Netiquette: With online communication, it is possible to miscommunicate what we mean or to misunderstand what our classmates mean given the lack of body language and immediate feedback. Therefore, please keep these netiquette (network etiquette) guidelines in mind. Failure to observe them may result in disciplinary action. Additional guidelines will be provided in the Blackboard course shell.

- *Always consider audience.* This is a college-level course; therefore, all communication should reflect polite consideration of other's ideas.
- *Respect and courtesy* must always be provided to classmates and the instruction team. No harassment or inappropriate postings will be tolerated.
- When reacting to someone else's message, *address the ideas*, not the person. Post only what anyone would comfortably state in a face-to-face situation.
- Blackboard is not a public internet venue; all postings to it should be considered private and confidential. Whatever is posted on in these online spaces is intended for classmates and instructors only. Please **do not copy documents or any course content to a publicly accessible website, blog, or other space.**

Attendance and Participation: Attendance in the course is determined by participation in the learning activities of the course. Your participation in the course is important not only for your learning and success

but also to create a community of learners. Participation is determined by completion of the following activities:

- Reading/Viewing all course materials to ensure understanding of assignment requirements
- Participating in an engaging discussion with your peers on the discussion boards and team journals
- Participating in scheduled Blackboard Collaborate sessions
- Completing session activities (e.g., quizzes, polls) indicated in the Blackboard course shell

Because these activities are designed to contribute to your learning each week, they cannot be made up after their due date has passed. Due to non-performance in the course or excessive non-excused absences (see UTEP Undergraduate Catalog for a list of excused absences) you may be dropped from the course with a grade of “W”. A 72-hours advance notice may be provided to you via email.

Blackboard Collaborate or Zoom sessions: The instructor may occasionally record lectures or deliver the class virtually using Blackboard Collaborate or Zoom. Our use of such technology is governed by the Federal Educational Rights and Privacy Act (FERPA) and UTEP’s acceptable-use policy. A recording of class sessions will be kept and stored by UTEP, in accordance with FERPA and UTEP policies. Your instructor will not share the recordings of your class activities outside of course participants, which include your fellow students, teaching assistants, or graduate assistants, and any guest faculty or community-based learning partners with whom we may engage during a class session. **You may not share class sessions recordings outside of this course.** Doing so may result in disciplinary action.

Course materials:

The course materials are only for the use of students currently enrolled in this course and only for the purpose of this course. They may not be further disseminated.

Deadlines and late work: The penalty for late work is 10% per 24 hours unless a specific assignment indicates otherwise. Make-up work will be given *only* in the case of a *documented* emergency. Note that make-up work may be in a different format than the original work, may require more intensive preparation, and may be graded with penalty points. If you miss homework or assignments and the reason is not considered excusable or you did not contact the instructor, the grade will be zero. ***It is therefore important to reach out to the instructor—in advance if at all possible, or as soon as possible—and explain why you missed a given course requirement.*** Proper documentation may be required. Once a deadline has been established for make-up work, no further extensions or exceptions will be granted. It is strongly recommended to submit your work with plenty of time to spare if you have a technical issue with Blackboard, network, and/or your computer. Save all your work (homework, quizzes, exams, and assignments) in OneDrive, the institutional cloud platform as a backup. This way, you will have evidence that you completed the work and will not lose credit. If you are experiencing difficulties submitting your work through Blackboard, please contact the UTEP Help Desk. Emailing work directly to the instructor should be your last resort.

COVID-19 Precautions. Please stay home if you have been diagnosed with COVID-19 or are experiencing COVID-19 symptoms. If you are feeling unwell, please let me know as soon as possible, so that we can work on appropriate accommodations. If you have tested positive for COVID-19, you are encouraged to

report your results to covidaction@utep.edu, so that the Dean of Students Office can provide you with support and help with communication with your professors. UTEP is equipped to provide COVID 19 testing.

The Center for Disease Control and Prevention recommends that people in areas of substantial or high COVID-19 transmission wear face masks when indoors in groups of people. The best way that Miners can take care of Miners is to get the vaccine. If you still need the vaccine, it is widely available in the El Paso area and will be available at no charge on campus during the first week of classes. For more information about the current rates, testing, and vaccinations, please visit epstrong.org

Special accommodations: Students requesting accommodation based on a disability must register with the [UTEP Center for Accommodations and Support Services](#) (CASS). Contact the Center for Accommodations and Support Services at 915-747-5148, or email them at cass@utep.edu, or apply for accommodations online via the [CASS portal](#).

Student resources: UTEP provides a variety of student services and support:

Technology Resources

- [Help Desk](#): Students experiencing technological challenges (email, Blackboard, software, etc.) can submit a ticket to the UTEP Helpdesk for assistance. Contact the Helpdesk via phone, email, chat, website, or in-person if on campus.

Academic Resources

- [UTEP Library](#): Access a wide range of resources including online, full-text access to thousands of journals and eBooks plus reference service and librarian assistance for enrolled students.
- [University Writing Center \(UWC\)](#): Submit papers here for assistance with writing style and formatting, ask a tutor for help and explore other writing resources.

Individual Resources

- [Military Student Success Center](#): Assists personnel in any branch of service to reach their educational goals.
- [Center for Accommodations and Support Services](#): Assists students with ADA-related accommodations for coursework, housing, and internships.
- [Counseling and Psychological Services](#): Provides a variety of counseling services including individual, couples, and group sessions as well as career and disability assessments.

NOTE: When in doubt on any of the above, please contact the instructor to check if you are following authorized procedures.