UNSW Course Outline

COMP9321 Data Services Engineering - 2024

Published on the 14 Feb 2024

General Course Information

Course Code : COMP9321
Year : 2024
Term : Term 1
Teaching Period : T1
Is a multi-term course? : No
Faculty : Faculty of Engineering
Academic Unit : School of Computer Science and Engineering
Delivery Mode : Online
Delivery Format : Standard
Delivery Location : Kensington
Campus : Sydney
Study Level : Undergraduate, Postgraduate
Units of Credit : 6

Useful Links

Handbook Class Timetable

Course Details & Outcomes

Course Description

This course aims to introduce the student to core concepts and practical skills for engineering the data in Web-service-oriented data-driven applications. Specifically, the course aims to expose students to basic infrastructure for building data services on the Web, including techniques to

access and ingest data in internal/external sources, develop software services to curate (e.g. extract, transform, correct, aggregate the data), develop services to apply various analytics and develop services to visualize the data to communicate effectively using data. The course uses the Python programming language as the practical basis for its modules. However, the concepts taught are universal and can be applied to any other web development language/framework.

Course Aims

Software engineering has advanced rapidly in recent years. The knowledge-, service-, and cloud-based economy in parallel with the continuous improvement in connectivity, storage and data processing capabilities allow access to a data deluge from sensors, social-media, news, user-generated, government and private data sources. Accordingly, in a modern data-oriented landscape, data-driven applications may need to deal with a collection of datasets - from relational to NoSQL - that holds a vast amount of data gathered from various private/open data islands. Therefore, well-engineered service-oriented functionalities are critical for ingesting, organizing and querying the growing volume of data in data-driven application.

This course aims to introduce the student to core concepts and practical skills for engineering the data in service-oriented data-driven applications. Specifically, the course aims to answer these questions:

- How to develop services to access and ingest data in internal/external sources of the data?
- How to develop services to use Databases (from Relational to NoSQL) as a Service for persisting user information?
- How to develop services to Curate (e.g. Extract, Transform, Correct, Aggregate, and Merge/Split) the data?
- How to develop services to apply analytics (e.g. by leveraging Machine Learning and Natural Language Processing techniques) to the curated data?
- How to develop services to visualize the data to communicate effectively with data.

Course Learning Outcomes

Course Learning Outcomes
CLO1 : Describe the main components of a data-oriented Web-based application
CLO2 : Read and ingest data from a variety of common sources
CLO3 : Apply techniques Extract, Transform, Correct, Aggregate, and Merge/Split to prepare and publish data
CLO4 : Identify and apply appropriate analytics and visualization techniques on prepared data
CLO5 : Design and develop a non-trivial web based data services application from scratch
CLO6 : Identify common security threats to a Web application
CLO7 : Evaluate the performance of a web based data services and identify bottlenecks

Course Learning Outcomes	Assessment Item
CLO1 : Describe the main components of a data-oriented Web-based application	• Assigment 2 • Final Exam
CLO2 : Read and ingest data from a variety of common sources	• Assignment 1
CLO3 : Apply techniques Extract, Transform, Correct, Aggregate, and Merge/Split to prepare and publish data	• In-class Quizzes • Assignment 1 • Assigment 2
CLO4 : Identify and apply appropriate analytics and visualization techniques on prepared data	• In-class Quizzes • Final Exam • Assignment 1 • Assigment 2
CLO5 : Design and develop a non-trivial web based data services application from scratch	• Final Exam • Assignment 1 • Assigment 2
CLO6 : Identify common security threats to a Web application	• Final Exam • Assigment 2
CLO7 : Evaluate the performance of a web based data services and identify bottlenecks	• In-class Quizzes • Final Exam

Learning and Teaching Technologies

Moodle - Learning Management System | Microsoft Teams | WebCMS

Learning and Teaching in this course

- Lectures: introduce concepts, show examples (Online lecture for 2024T1)
- Self Guided Lab Activities: introduce technology required for the assignments
- Consultation Labs every day of the week
- Online Quizzes: revision of the concepts introduced in Lectures and Lab Work
- Assignments: solve significant problems
- Final Exam: Validate the stdudents understanding of the course material over all

Teaching Rationale

This course is taught the way it is because we believe in structured learning, also learning by doing. We provide timely feedback for learning via small, step-by-step individual assignments that gradually build up knowledge and practical skills.

Assessments

Assessment Structure

Assessment Item	Weight	Relevant Dates
In-class Quizzes	10%	Due Date: Throughout the term.
Assignment 1	10%	Start Date: Week 3 Due Date: Week 5
Assigment 2	20%	Start Date: Week 5 Due Date: Week 7
Final Exam	60%	Start Date: Exam Period Due Date: Exam Period

Assessment Details

In-class Quizzes

Course Learning Outcomes

- CLO3 : Apply techniques Extract, Transform, Correct, Aggregate, and Merge/Split to prepare and publish data
- CLO4 : Identify and apply appropriate analytics and visualization techniques on prepared data
- CLO7 : Evaluate the performance of a web based data services and identify bottlenecks

Detailed Assessment Description

5 Quizzes over all

Assignment 1

Course Learning Outcomes

- CLO2 : Read and ingest data from a variety of common sources
- CLO3 : Apply techniques Extract, Transform, Correct, Aggregate, and Merge/Split to prepare and publish data
- CLO4 : Identify and apply appropriate analytics and visualization techniques on prepared data
- CLO5 : Design and develop a non-trivial web based data services application from scratch

Detailed Assessment Description

Data ingestion, manipulation and visualization

Assigment 2

Course Learning Outcomes

- CLO1 : Describe the main components of a data-oriented Web-based application
- CLO3 : Apply techniques Extract, Transform, Correct, Aggregate, and Merge/Split to prepare and publish data

- CLO4 : Identify and apply appropriate analytics and visualization techniques on prepared data
- CLO5 : Design and develop a non-trivial web based data services application from scratch
- CLO6 : Identify common security threats to a Web application

Detailed Assessment Description

Data publication as a RESTful service API

Final Exam

Course Learning Outcomes

- CLO1 : Describe the main components of a data-oriented Web-based application
- CLO4 : Identify and apply appropriate analytics and visualization techniques on prepared data
- CLO5 : Design and develop a non-trivial web based data services application from scratch
- CLO6 : Identify common security threats to a Web application
- CLO7 : Evaluate the performance of a web based data services and identify bottlenecks

Assessment Length

2 hours + reading time

Assignment submission Turnitin type

Not Applicable

General Assessment Information

Grading Basis

Standard

Course Schedule

Teaching Week/Module	Activity Type	Content
Week 1 : 12 February - 18 February	Lecture	Course overview and Introdcution to Data Services Engineering
	Laboratory	No Consultation Lab sessions. student should start by the Setup Python, Flask, NumPy, Pandas.
Week 2 : 19 February - 25 February	Lecture	Data Access and ingestion
	Assessment	Quiz1
	Laboratory	Accessing NoSQL DB, API data sourced, CSV files, text files.
Week 3 : 26 February - 3 March	Lecture	Data Cleansing and Manipulation
	Laboratory	Cleansing data with Python Pandas
	Assessment	Assignment1 specification Released
	Assessment	Quiz2
Week 4 : 4 March - 10 March	Lecture	Data Visualization
	Laboratory	Using matplotlib library for charts and plots
	Assessment	Quiz3
Week 5 : 11 March - 17 March	Lecture	Building a Data service (part1)
	Assessment	Assignment1 is Due Assignment2 Specification is released
Week 6 : 18 March - 24 March	Other	Flexibility Week, no activties planned
Week 7 : 25 March - 31 March	Lecture	Building a Data service (part2)
Week 8 : 1 April - 7 April	Lecture	Data Analytics Applied Techniques and Tools part1
	Assessment	Quiz4
	Assessment	Assignment2 is Due Assignment3 Specification released
Week 9 : 8 April - 14 April	Lecture	Data Analytics Applied Techniques and Tools part2
Week 10 : 15 April - 21 April	Lecture	Final wrap-up
	Assessment	Quiz5 Assignment3 is Due

Attendance Requirements

Students are strongly encouraged to attend all classes and review lecture recordings.

General Schedule Information

Note: The Course Schedule might change according to the progress of the topics and feedback throughout the course. Although the quizzes are planned as indicated, but the actual number may vary depending on the progress of the weekly topics.

Course Resources

Prescribed Resources

Although there is no official textbook for the course, but here are some of useful books:

- Python for Data Analysis, Wes McKinney
- RESTful Web Clients: Enabling Reuse Through Hypermedia, By Mike Amundsen

- Mastering Machine Learning with Scikit-Learn, Second Edition. Gavin Hackeling

Recommended Resources

recommended resources will be shared for each courses and access to useful resources will be greanted via the access to DataCamp website

Course Evaluation and Development

This course is evaluated each session using the myExperience system.

Changes made in response to MyExperience feedback

In the spirit of continuous improvement, the following adjustments were made to address students' comments in the MyExperience Survey from the last offering:

- Relying on Three additional Tutors to ensure fast response on the Course Forum
- Creating Assignment specs workgroup, for developing the Assignments' specifications. This will help making sure everything is clear, checked, double checked and triple checked.
- Establishing stronger line of communication between the teaching team through weekly stand-ups and frequent check-ins. This will minimize misunderstandings and increase consistency.
- Include more in-lecture exercises.
- Adjusting the strategy for Bonus marks

Staff Details

Position	Name	Email	Location	Phone	Availability	Equitable Learning Services Contact	Primary Contact
Lecturer	Morty Al-Banna					Yes	Yes
Administrator	Mehmet Rad					No	No
Tutor	Dylan Sanusi-Goh					No	No
	Tim Arney					No	No

Other Useful Information

Academic Information

I. Special consideration and supplementary assessment

If you have experienced an illness or misadventure beyond your control that will interfere with your assessment performance, you are eligible to apply for Special Consideration prior to, or

within 3 working days of, submitting an assessment or sitting an exam.

Please note that UNSW has a Fit to Sit rule, which means that if you sit an exam, you are declaring yourself fit enough to do so and cannot later apply for Special Consideration.

For details of applying for Special Consideration and conditions for the award of supplementary assessment, please see the information on UNSW's Special Consideration page.

II. Administrative matters and links

All students are expected to read and be familiar with UNSW guidelines and polices. In particular, students should be familiar with the following:

- Attendance
- UNSW Email Address
- Special Consideration
- Exams
- Approved Calculators
- Academic Honesty and Plagiarism
- Equitable Learning Services

III. Equity and diversity

Those students who have a disability that requires some adjustment in their teaching or learning environment are encouraged to discuss their study needs with the course convener prior to, or at the commencement of, their course, or with the Equity Officer (Disability) in the Equitable Learning Services. Issues to be discussed may include access to materials, signers or note-takers, the provision of services and additional exam and assessment arrangements. Early notification is essential to enable any necessary adjustments to be made.

IV. Professional Outcomes and Program Design

Students are able to review the relevant professional outcomes and program designs for their streams by going to the following link: https://www.unsw.edu.au/engineering/student-life/student-resources/program-design.

Note: This course outline sets out the description of classes at the date the Course Outline is published. The nature of classes may change during the Term after the Course Outline is published. Moodle or your primary learning management system (LMS) should be consulted for the up-to-date class descriptions. If there is any inconsistency in the description of activities between the

University timetable and the Course Outline/Moodle/LMS, the description in the Course Outline/ Moodle/LMS applies.

Academic Honesty and Plagarism

UNSW has an ongoing commitment to fostering a culture of learning informed by academic integrity. All UNSW students have a responsibility to adhere to this principle of academic integrity. Plagiarism undermines academic integrity and is not tolerated at UNSW. *Plagiarism at UNSW is defined as using the words or ideas of others and passing them off as your own.*

Plagiarism is a type of intellectual theft. It can take many forms, from deliberate cheating to accidentally copying from a source without acknowledgement. UNSW has produced a website with a wealth of resources to support students to understand and avoid plagiarism, visit: student.unsw.edu.au/plagiarism. The Learning Centre assists students with understanding academic integrity and how not to plagiarise. They also hold workshops and can help students one-on-one.

You are also reminded that careful time management is an important part of study and one of the identified causes of plagiarism is poor time management. Students should allow sufficient time for research, drafting and the proper referencing of sources in preparing all assessment tasks.

Repeated plagiarism (even in first year), plagiarism after first year, or serious instances, may also be investigated under the Student Misconduct Procedures. The penalties under the procedures can include a reduction in marks, failing a course or for the most serious matters (like plagiarism in an honours thesis or contract cheating) even suspension from the university. The Student Misconduct Procedures are available here:

www.gs.unsw.edu.au/policy/documents/studentmisconductprocedures.pdf

Submission of Assessment Tasks

Work submitted late without an approved extension by the course coordinator or delegated authority is subject to a late penalty of five percent (5%) of the maximum mark possible for that assessment item, per calendar day.

The late penalty is applied per calendar day (including weekends and public holidays) that the assessment is overdue. There is no pro-rata of the late penalty for submissions made part way

through a day. This is for all assessments where a penalty applies.

Work submitted after five days (120 hours) will not be accepted and a mark of zero will be awarded for that assessment item.

For some assessment items, a late penalty may not be appropriate. These will be clearly indicated in the course outline, and such assessments will receive a mark of zero if not completed by the specified date. Examples include:

- Weekly online tests or laboratory work worth a small proportion of the subject mark;
- Exams, peer feedback and team evaluation surveys;
- Online quizzes where answers are released to students on completion;
- Professional assessment tasks, where the intention is to create an authentic assessment that has an absolute submission date; and,
- Pass/Fail assessment tasks.

Faculty-specific Information

Engineering Student Support Services – The Nucleus - enrolment, progression checks, clash requests, course issues or program-related queries

Engineering Industrial Training – Industrial training questions

UNSW Study Abroad – study abroad student enquiries (for inbound students)

UNSW Exchange – student exchange enquiries (for inbound students)

UNSW Future Students – potential student enquiries e.g. admissions, fees, programs, credit transfer

Phone

(+61 2) 9385 8500 – Nucleus Student Hub

(+61 2) 9385 7661 – Engineering Industrial Training

(+61 2) 9385 3179 – UNSW Study Abroad and UNSW Exchange (for inbound students)

School Contact Information

CSE Help! - on the Ground Floor of K17

- For assistance with coursework assessments.

The Nucleus Student Hub - https://nucleus.unsw.edu.au/en/contact-us

- Course enrolment queries.

Grievance Officer - grievance-officer@cse.unsw.edu.au

- If the course convenor gives an inadequate response to a query or when the course convenor does not respond to a query about assessment.

Student Reps - stureps@cse.unsw.edu.au

- If some aspect of a course needs urgent improvement. (e.g. Nobody responding to forum queries, cannot understand the lecturer)