CS6526: PROJECT IN TRUSTWORTHY AI

Effective Term
Semester B 2025/26

Part I Course Overview

Course Title
Project in Trustworthy AI

Subject Code
CS - Computer Science
Course Number
6526

Academic Unit
Computer Science (CS)

College/School
College of Computing (CC)

Course Duration
Two Semesters

Credit Units
0-6

Level
P5, P6 - Postgraduate Degree

Medium of Instruction
English

Medium of Assessment
English

Prerequisites
Students should have completed at least 12 credit units

Precursors
Nil

Equivalent Courses
Nil

Exclusive Courses
CS6522 Project in Autonomous Driving or
CS6523 Internship in Autonomous Driving or
CS6524 Project in Generative AI or
CS6525 Internship in Generative AI or
CS6527 Internship in Trustworthy AI or
CS6528 Internship in Artificial Intelligence or
CS6529 Project in Artificial Intelligence or

CS6539 Internship in Autonomous Driving or
CS6540 Internship in Generative AI or
CS6541 Internship in Trustworthy AI or
CS6542 Internship in Artificial Intelligence or

Part II Course Details

Abstract

This project-based course focuses on the design, development, and evaluation of AI systems that prioritize trustworthiness through fairness, transparency, accountability, and robustness. Students will work in teams to tackle real-world problems, applying ethical frameworks, technical tools, and governance strategies to ensure AI systems align with societal values. The course emphasizes hands-on implementation, critical analysis, and stakeholder communication.

Course Intended Learning Outcomes (CILOs)

	CILOs	Weighting (if app.)	DEC-A1	DEC-A2	DEC-A3
1	Identify a challenging trustworthy AI problem, analyze the problem in detail; and propose innovative solutions through computing means.		x		
2	Provide a proof-of-concept for the solution by designing and developing a working system or application.			x	x
3	Implement and evaluate the developed system or application to match the initial system requirements.				x
4	Document and report the system design process, study, implementation and evaluation findings using different communication media.				x

A1: Attitude
Develop an attitude of discovery/innovation/creativity, as demonstrated by students possessing a strong sense of curiosity, asking questions actively, challenging assumptions or engaging in inquiry together with teachers.

A2: Ability
Develop the ability/skill needed to discover/innovate/create, as demonstrated by students possessing critical thinking skills to assess ideas, acquiring research skills, synthesizing knowledge across disciplines or applying academic knowledge to real-life problems.

A3: Accomplishments
Demonstrate accomplishment of discovery/innovation/creativity through producing /constructing creative works/new artefacts, effective solutions to real-life problems or new processes.

Learning and Teaching Activities (LTAs)

	LTAs	Brief Description	CILO No.	Hours/week (if applicable)
1	Project planning	Students will identify the problem for investigation and draft a project plan with appropriate milestones.	1	

2	Project proposal	Students will analyze the problem identified and research on existing and/or related solutions. Then, in consultation with their supervisors, they will propose their own designs and solutions.	2	
3	Project implementation and evaluation	Students will implement the proposed solutions and validate their designs by testing and evaluating the completed solution.	3	
4	Project documentation	Students will document and explain their work in regular progress reports and a final report. At the end, they are required to present their projects in the form of oral presentation and demonstration.	4	

Additional Information for LTAs

Teaching pattern:

Suggested lecture/tutorial/laboratory mix: 8 hours individual consultation per semester.

The course is designed to guide students in proposing and managing their own projects in Trustworthy AI. Each student will find an academic staff to supervise the project on a one to one basis.

The role of the supervisor is to closely monitor the project progress with project meetings regularly, in order to give advice to the student, to establish criteria for assessment, and to advise on possible solutions and potential problems at an early stage. In particular, the supervisor is expected to encourage the student to explore innovative approaches and alert the student to the possibility of alternative and novel solutions to problems encountered.

Assessment Tasks / Activities (ATs)

	ATs	CILO No.	Weighting (%)	Remarks ("-" for nil entry)	Allow Use of GenAI?
1	Project management and individual development of the student	1	10	For assessment of technical merit, report, and presentation, the project committee assigns two examiners, including the supervisor. The Supervisor is required to give detailed grading reports on all aspects of assessment. The Assessor will evaluate the CILOs 2-4 of the project. The Course Leader will review all projects, moderate consistency across a wide range of projects, and, where necessary, resolve discrepancies between grading of the Assessor and the Supervisor, drawing on the expertise of domain experts as needed.	No

2	Technical merit of the proposed solution, including the degree of innovation in the proposed design or solution	2, 3	50	For assessment of technical merit, report, and presentation, the project committee assigns two examiners, including the supervisor. The Supervisor is required to give detailed grading reports on all aspects of assessment. The Assessor will evaluate the CILOs 2-4 of the project. The Course Leader will review all projects, moderate consistency across a wide range of projects, and, where necessary, resolve discrepancies between grading of the Assessor and the Supervisor, drawing on the expertise of domain experts as needed.	No

3	Standard of final documentation	4	30	Allowed for writing polishing. For assessment of technical merit, report, and presentation, the project committee assigns two examiners, including the supervisor. The Supervisor is required to give detailed grading reports on all aspects of assessment. The Assessor will evaluate the CILOs 2-4 of the project. The Course Leader will review all projects, moderate consistency across a wide range of projects, and, where necessary, resolve discrepancies between grading of the Assessor and the Supervisor, drawing on the expertise of domain experts as needed.	Yes

4	Standard of oral presentation	4	10	For assessment of technical merit, report, and presentation, the project committee assigns two examiners, including the supervisor. The Supervisor is required to give detailed grading reports on all aspects of assessment. The Assessor will evaluate the CILOs 2-4 of the project. The Course Leader will review all projects, moderate consistency across a wide range of projects, and, where necessary, resolve discrepancies between grading of the Assessor and the Supervisor, drawing on the expertise of domain experts as needed.	Yes

Continuous Assessment (%)

100

Additional Information for ATs

Dissertation-type Course:
This course falls under the academic regulation for dissertation-type courses (AR12.6). The course assessed through 100% coursework.
Each student is assigned a supervisor from the academic staff for individual consultation.
The normal duration of the course is two semesters, after which the dissertation must be submitted.
The maximum duration of the course is four semesters, after which no further extension must be permitted.
Dissertation-type courses may NOT be repeated.

Assessment Rubrics (AR)

Assessment Task

1. Project planning

Criterion

1.1 ABILITY to IDENTIFY problems for investigations.
1.2 ABILITY to PLAN a project schedule with appropriate milestones, and MAINTAIN the project schedule.

Excellent
(A+, A, A-) High

Good
(B+, B, B-) Significant

Fair
(C+, C, C-) Moderate

Marginal
(D) Basic

Failure
(F) Not even reaching marginal levels

Assessment Task
2. Project proposal

Criterion
2.1 ABILITY to ANALYZE a problem.
2.2 ABILITY to EVALUATE, COMPARE, and CONTRAST existing solutions.
2.3 ABILITY to DESIGN and INNOVATE new solutions.

Excellent
(A+, A, A-) High

Good
(B+, B, B-) Significant

Fair
(C+, C, C-) Moderate

Marginal
(D) Basic

Failure
(F) Not even reaching marginal levels

Assessment Task
3. Project implementation and evaluation

Criterion
3.1 ABILITY to IMPLEMENT the proposed solution.
3.2 ABILITY to VALIDATE and TEST the implemented solution.
3.3 ABILITY to EVALUATE and INTERPRET results from the design, and COMPARE with existing solutions.

Excellent
(A+, A, A-) High

Good
(B+, B, B-) Significant

Fair

(C+, C, C-) Moderate

Marginal

(D) Basic

Failure

(F) Not even reaching marginal levels

Assessment Task

4. Project documentation

Criterion

4.1 ABILITY to DOCUMENT the progress of the project in interim reports.
4.2 ABILITY to DOCUMENT the OUTCOMES of the project in a final report.
4.3 ABILITY to DEMONSTRATE project outcomes in an oral presentation.

Excellent

(A+, A, A-) High

Good

(B+, B, B-) Significant

Fair

(C+, C, C-) Moderate

Marginal

(D) Basic

Failure

(F) Not even reaching marginal levels

Part III Other Information

Keyword Syllabus

The Project has no fixed formal syllabus. Each student will be required to undertake an individual piece of work, which is related to the trustworthy AI. The topic area of the dissertation will be chosen so that the aims of the Project can be achieved. Criteria for topic choice include: (i) compatibility with a subject area of Trustworthy AI; (ii) availability of a qualified supervisor; (iii) appropriate academic level; (iv) availability of necessary specialized resources. Following topic areas will be included. 1) Foundations of Trustworthy AI: Ethics, principles, and regulations. 2) Bias Detection and Mitigation (e.g., adversarial debiasing, fairness metrics). 3) Explainability Methods (e.g., LIME, SHAP, counterfactual explanations). 4) Robustness and Security (e.g., adversarial attacks, model testing). 5) Accountability Frameworks (e.g., model cards, impact assessments). 6) Stakeholder Engagement and Human-Centered Design.

Reading List

Compulsory Readings

	Title
1	N/A

Additional Readings

	Title
1	N/A