Revised: CET 07/2023, Viterbi 08/2023

USC Viterbi
School of Engineering

Course ID and Title: EE599: Foundations of Security, Privacy, and Trustworthiness in AI and Blockchain Systems

Units: 4

Term—Day—Time: Fall 2025 — TuTh — 10–11:50am

Location: KAP137

Instructor: Salman Avestimehr
Office: EEB 500B
Office Hours: TuTh, 9:00-10:00am
Contact Info: 213-740-7326. avestime@usc.edu.

Teaching Assistant:
Office: [TBD]
Office Hours: [TBD]
Contact Info: [TBD]

Catalogue Description

This course covers principles and algorithms for secure, private, and trustworthy computing, focusing on artificial intelligence (AI) and blockchain technologies.

Course Description

This course offers a comprehensive introduction to the principles and algorithms that enable secure, private, and trustworthy computing, with a focus on applications in artificial intelligence (AI) and blockchain technologies. The course is structured into five key sections:

1. **Security and Privacy Fundamentals**: Information-theoretic measures of security and privacy, along with foundational ciphers for secure communication.
2. **Cryptographic Protocols**: Essential concepts in key exchange and public-key encryption.
3. **Secure Machine Learning**: Techniques for secure and privacy-preserving machine learning, with emphasis on federated learning and differential privacy.
4. **Trustworthy AI**: Exploring trustworthiness in large language models (LLMs) and mechanisms for ensuring their reliability.
5. **Cryptocurrencies and Blockchain**: An in-depth look at Bitcoin, cryptocurrencies, and the cryptographic foundations of blockchain

Learning Objectives

By the end of this course, students will develop a foundational understanding of the key principles and algorithms that enable secure, private, and trustworthy computing and communication, particularly within the contexts of artificial intelligence (AI) and blockchain technologies. Upon completing the course, students will:

- Gain a deep understanding of **information-theoretic measures** for security and privacy, as well as foundational techniques such as the **one-time pad** and **information-theoretically secure communication**.

- Master the essentials of cryptographic techniques, including **stream ciphers**, **block ciphers**, **key exchange**, **public-key encryption**, and **cryptographic protocols** like **Diffie-Hellman**, **secure multi-party computation**, **Yao's garbled circuits**, **oblivious transfer**, **zero-knowledge proofs**, and more.
- Understand techniques for **secure and privacy-preserving machine learning**, with a focus on **federated learning** and **differential privacy**.
- Grasp the fundamentals of **large language models (LLMs)**, with emphasis on assessing and quantifying **uncertainty**, mitigating **hallucination**, and building **trustworthiness** in AI systems.
- Be introduced to **cryptocurrencies** and their cryptographic underpinnings, gaining a solid understanding of **Bitcoin**, its mining strategies, **proof-of-work** consensus mechanisms, and **proof-of-stake** protocols.

Prerequisite(s): EE503 (or equivalent)
Co-Requisite(s): none
Concurrent Enrollment: none
Recommended Preparation: none

Course Notes

Grading type: Letter
Papers and handouts that will be provided throughout the semester on Brightspace

Required Readings and Supplementary Materials

Papers and handouts that will be provided throughout the semester on Brightspace

Description of Assignments and How They Will Be Assessed

- There will be ~5 Homework assignments throughout the course.
- Late HW will not be accepted. A late assignment results in a zero grade. Please have your homework turned in by the beginning of lecture on the date that it is due.
- Homeworks will be assigned on Thursdays and collected the following week in class (on Thursdays at the beginning of the lecture)
- Show your work in your homework solution; the correct answer alone is worth only partial credit.
- Homework collaboration is encouraged. This is discussing problems and solution strategies with your classmates, the TA, and/or the instructor and is to be distinguished from copying solutions of others which is prohibited.
- For computer-based assignments no code can be shared or copied from the internet. The only exception is code provided to the entire class by the instructor or TA.

Exam Policy

- The course will have 1 midterm exam on Thursday March 13th , 2025 in class.

Final Project Policy

- Proposals will be due in March (exact date to be announced)
- Topics can be suggested by the students or taken from a list of suggested topics to be provided.
- The deliverables are:
    - A final report of approximately 5 pages.
- A presentation to be made on a projects-day event at the end of the semester. (details will be announced later)

Grading Breakdown

Assessment Tool (assignments)	% of Grade
Homework Assignments	25
Midterm Exam	35
Final Project	40
TOTAL	100

Grading Scale

Final grades will be assigned by a combination of student score distribution (curve) and the discretion of the instructor. Final grades are nonnegotiable.

Assignment Submission Policy

Homework assignments will be assigned on Thursdays and collected the following week in class (on Thursdays at the beginning of the lecture)

Course-Specific Policies

[Add any additional policies of which students should be aware: late work submissions, missed classes, use of technology in the classroom, etc. Course-specific policies differ from university policies in that they are set by each instructor or department/program.]

Attendance

Attendance in lectures is required.

Academic Integrity for this Class

Unless otherwise noted, this course will follow the expectations for academic integrity as stated in the USC Student Handbook. The general USC guidelines on Academic Integrity and Course Content Distribution are provided later in this syllabus.

If found responsible for an academic violation, students may be assigned university outcomes, such as suspension or expulsion from the university, and grade penalties, such as an "F" grade on the assignment, exam, and/or in the course.

Please ask the instructor if you are unsure about what constitutes unauthorized assistance on an exam or assignment, or what information requires citation and/or attribution.

Class Recordings and Course Content Distribution: You may not record this class without the express permission of the instructor and all other students in the class. Distribution of any notes, recordings, exams, or other materials from a university class or lectures — other than for individual or class group study — is prohibited without the express permission of the instructor; violations will be considered an intentional act to facilitate or enable academic dishonesty and reported to the university.

Use of Generative AI in this Course

Generative AI is permitted but limited as follows: In this course, you are permitted to use artificial intelligence (AI)-powered programs to help you, but only on assignments that explicitly indicate a permitted use of AI. However:

- You should also be aware that AI text generation tools may present incorrect information, biased responses, and incomplete analyses; thus, their answers may not meet the standards of this course.
- To adhere to our university values, *you must cite any AI-generated material (e.g., text, images, and other content) included or referenced in your work and provide the prompts used to generate the*

content. Using an AI tool to generate content without proper attribution will be treated as plagiarism and reported to the Office of Academic Integrity.

Please review the instructions in each assignment for more details on how and when to use AI Generators for your submissions.

Course Evaluations

Course evaluation occurs at the end of the semester university-wide. It is an important review of students' experience in the class. The process and intent of the end-of-semester evaluation should be provided. In addition, a mid-semester evaluation is recommended practice for early course correction. You may choose to contact CET for support in creating a mid-semester evaluation.

Course Schedule

	Topics/Daily Activities	Readings/Preparation	Deliverable/ Due Dates
Week 1	Course introduction; Introduction to cyphers; Information measures: entropy and mutual information.	Handouts will be provided.	
Week 2	One-time pad; information-theoretic secure communication; stream cyphers	Handouts will be provided. Hw 1 will be assigned.	
Week 3	Block cyphers	Handouts will be provided. Hw 2 will be assigned.	Hw 1 is due.
Week 4	Basic key exchange and public-key encryption	Handouts will be provided. Hw 3 will be assigned.	Hw 2 is due.
Week 5	Modular-arithmetic; Diffie-Hellman public key encryption	Handouts will be provided. Hw 4 will be assigned.	Hw 3 is due.
Week 6	Two-party secure multi-party computing; Yao's garbled circuit	Handouts will be provided.	Hw 4 is due.
Week 7	Zero-knowledge proofs Verifiable computing	Handouts will be provided. Hw 5 will be assigned.	Hw 5 is due.
Week 8	Secure and privacy-preserving machine learning	Handouts will be provided.	Project proposal is due.
Week 9	Federated learning and differential privacy	Handouts will be provided.	
Week 10	Introduction to large language models (LLMs) and their safety concerns	Handouts will be provided.	
Week 11	Approaches and metrics for assessing uncertainty in LLMs	Handouts will be provided.	
Week 12	Safety guards for LLMs	Handouts will be provided.	
Week 13	Intro to cryptography & crypto currencies	Handouts will be provided.	
Week 14	Overview of consensus protocols	Handouts will be provided.	
Week 15	Bitcoin nuts and bolts	Handouts will be provided.	Project reports are due.

Academic Integrity

The University of Southern California is foremost a learning community committed to fostering successful scholars and researchers dedicated to the pursuit of knowledge and the transmission of ideas. Academic misconduct — which includes any act of dishonesty in the production or submission of academic work (either in draft or final form) — is in contrast to the university's mission to educate students through a broad array of academic, professional, and extracurricular programs.

This course will follow the expectations for academic integrity as stated in the USC Student Handbook. All students are expected to submit assignments that are their own original work and prepared specifically for this course and section in this academic term. You may not submit work written by others or "recycle" work prepared for other courses without obtaining written permission from the instructor(s). Students suspected of engaging in academic misconduct will be reported to the Office of Academic Integrity.

Other violations of academic misconduct include, but are not limited to, cheating, plagiarism, fabrication (e.g., falsifying data), knowingly assisting others in acts of academic dishonesty, and any act that gains or is intended to gain an unfair academic advantage.

Academic dishonesty has a far-reaching impact and is considered a serious offense against the university. Violations will result in a grade penalty, such as a failing grade on the assignment or in the course, and disciplinary action from the university itself, such as suspension or even expulsion.

For more information about academic integrity see the student handbook or the Office of Academic Integrity's website, and university policies on Research and Scholarship Misconduct.

Please ask your instructor if you are unsure what constitutes unauthorized assistance on an exam or assignment or what information requires citation and/or attribution.

Course Content Distribution and Synchronous Session Recordings Policies

USC has policies that prohibit recording and distribution of any synchronous and asynchronous course content outside of the learning environment.

Recording a university class without the express permission of the instructor and announcement to the class, or unless conducted pursuant to an Office of Student Accessibility Services (OSAS) accommodation. Recording can inhibit free discussion in the future, and thus infringe on the academic freedom of other students as well as the instructor. (Living our Unifying Values: The USC Student Handbook, page 13).

Distribution or use of notes, recordings, exams, or other intellectual property, based on university classes or lectures without the express permission of the instructor for purposes other than individual or group study. This includes but is not limited to providing materials for distribution by services publishing course materials. This restriction on unauthorized use also applies to all information, which had been distributed to students or in any way had been displayed for use in relation to the class, whether obtained in class, via email, on the internet, or via any other media. Distributing course material without the instructor's permission will be presumed to be an intentional act to facilitate or enable academic dishonestly and is strictly prohibited. (Living our Unifying Values: The USC Student Handbook, page 13).

Statement on University Academic and Support Systems

Students and Disability Accommodations:
USC welcomes students with disabilities into all of the University's educational programs. The Office of Student Accessibility Services (OSAS) is responsible for the determination of appropriate accommodations for students who encounter disability-related barriers. Once a student has completed the OSAS process (registration, initial appointment, and submitted documentation) and accommodations are determined to be reasonable and appropriate, a Letter of Accommodation (LOA) will be available to generate for each course. The LOA must be given to each course instructor by the student and followed up with a discussion. This should be done as early in the semester as possible as accommodations are not retroactive. More information can be found at osas.usc.edu. You may contact OSAS at (213) 740-0776 or via email at osasfrontdesk@usc.edu.

Student Financial Aid and Satisfactory Academic Progress:
To be eligible for certain kinds of financial aid, students are required to maintain Satisfactory Academic Progress (SAP) toward their degree objectives. Visit the Financial Aid Office webpage for undergraduate- and graduate-level SAP eligibility requirements and the appeals process.

Support Systems:
Counseling and Mental Health - (213) 740-9355 – 24/7 on call
Free and confidential mental health treatment for students, including short-term psychotherapy, group counseling, stress fitness workshops, and crisis intervention.

988 Suicide and Crisis Lifeline - 988 for both calls and text messages – 24/7 on call
The 988 Suicide and Crisis Lifeline (formerly known as the National Suicide Prevention Lifeline) provides free and confidential emotional support to people in suicidal crisis or emotional distress 24 hours a day, 7 days a week, across the United States. The Lifeline consists of a national network of over 200 local crisis centers, combining custom local care and resources with national standards and best practices. The new, shorter phone number makes it easier for people to remember and access mental health crisis services (though the previous 1 (800) 273-8255 number will continue to function indefinitely) and represents a continued commitment to those in crisis.

Relationship and Sexual Violence Prevention Services (RSVP) - (213) 740-9355(WELL) – 24/7 on call
Free and confidential therapy services, workshops, and training for situations related to gender- and power-based harm (including sexual assault, intimate partner violence, and stalking).

Office for Equity, Equal Opportunity, and Title IX (EEO-TIX) - (213) 740-5086
Information about how to get help or help someone affected by harassment or discrimination, rights of protected classes, reporting options, and additional resources for students, faculty, staff, visitors, and applicants.

Reporting Incidents of Bias or Harassment - (213) 740-2500
Avenue to report incidents of bias, hate crimes, and microaggressions to the Office for Equity, Equal Opportunity, and Title for appropriate investigation, supportive measures, and response.

The Office of Student Accessibility Services (OSAS) - (213) 740-0776
OSAS ensures equal access for students with disabilities through providing academic accommodations and auxiliary aids in accordance with federal laws and university policy.

USC Campus Support and Intervention - (213) 740-0411
Assists students and families in resolving complex personal, financial, and academic issues adversely affecting their success as a student.

Diversity, Equity and Inclusion - (213) 740-2101
Information on events, programs and training, the Provost's Diversity and Inclusion Council, Diversity Liaisons for each academic school, chronology, participation, and various resources for students.

USC Emergency - UPC: (213) 740-4321, HSC: (323) 442-1000 – 24/7 on call
Emergency assistance and avenue to report a crime. Latest updates regarding safety, including ways in which instruction will be continued if an officially declared emergency makes travel to campus infeasible.

USC Department of Public Safety - UPC: (213) 740-6000, HSC: (323) 442-1200 – 24/7 on call
Non-emergency assistance or information.

Office of the Ombuds - (213) 821-9556 (UPC) / (323-442-0382 (HSC)
A safe and confidential place to share your USC-related issues with a University Ombuds who will work with you to explore options or paths to manage your concern.

Occupational Therapy Faculty Practice - (323) 442-2850 or otfp@med.usc.edu
Confidential Lifestyle Redesign services for USC students to support health-promoting habits and routines that enhance quality of life and academic performance.