SAIS '27: USENIX Conference on Secure Agentic-AI Systems

June 7–8, 2027, Santa Clara, CA, USA

Sponsored by USENIX, the Advanced Computing Systems Association

The **USENIX Conference on Secure Agentic-AI Systems (SAIS '27)** will take place on June 7–8, 2027, in Santa Clara, CA, USA.

Important Dates

- Paper submissions due: **Thursday, February 4, 2027, 23:59 AoE**
- Notification to authors: **Tuesday, March 16, 2027**
- Shepherded approval due: **Tuesday, March 30, 2027**
- Final papers due: **Tuesday, April 20, 2027, 23:59 AoE**

Overview

The USENIX Conference on Secure Agentic-AI Systems (SAIS) [pronounced /sɛz/] is a new conference dedicated to the security and privacy of AI-powered autonomous and agentic systems. SAIS brings together researchers, practitioners, developers, and users from around the world to explore cutting-edge ideas and principled approaches to designing, verifying, and deploying secure and privacy-preserving AI agents in the real world.

Topics of Interest

USENIX SAIS welcomes contributions on all aspects of safety, security, and privacy of agentic AI-based systems. SAIS is a systems-security venue, and as such all contributions must be relevant to the broader systems security community. Topics of interest include, but are not limited to:

- Adversarial Control of Agent Behavior
  Attacks exploiting untrusted inputs or tool outputs.
  - Prompt Injections
  - Jailbreaks and Completion Manipulations
  - Backdoor and Trigger-Based Manipulation
- Privacy and Information Leakage
  Risks related to data exfiltration and inference.
  - Secret Exfiltration via Crafted Queries
  - PII Leakage and Training Data Extraction
  - Cross-Context Inference Attacks
- Secure Integration with Tools
  Security of external environments and plugins.
  - Supply-Chain and Tool Exploitation
  - Capability Confinement and Least-Privilege Design
  - Formal Guarantees of Safety for Interfaces
- Defense Mechanisms
  Architectural and policy-level defenses.
  - Training-Free Defenses (e.g., Trajectory Re-execution)
  - Architectural Hardening (e.g., Secure Memory Design)
  - Policy and Planning Guards
- Evaluation & Threat Modeling
  Standardizing threats and benchmarks.
  - Taxonomies and Shared Threat Models
  - Red-Teaming Pipelines
  - Metrics and Reproducibility
- Containment of Cyber-Capable Agents
  Controlling the threat of autonomous hacker agents.
  - Sandboxing and Policy Systems to Contain Highly Capable Agents
  - Firewalls to Mediate Inter-Agent Communications
- Negative Results
  Understanding whether progress is real.
  - Results "Proving a Negative" by Identifying Unworthy Lines of Research
  - Lessons Learned from Unsuccessful Experimental Methodologies
  - Results Debunking Earlier Findings by Describing Precisely How and Why Proposed Techniques Fail Where Earlier Successes Were Claimed

Submission Policies and Instructions

Submissions must be made through the USENIX SAIS '27 submission system linked from the Call for Papers web page. Submitted papers should describe original, scientifically sound work produced by the authors. Submissions must follow the USENIX Conference Submission Policy (www.usenix.org/conferences/author-resources/submissions-policy).

Formatting Requirements & Page Limit

Submissions must be in PDF format and typeset using USENIX's LaTeX template and style files (www.usenix.org/conferences/author-resources/paper-templates). Papers must use U.S. letter-sized pages, a two-column format, 10-point Times Roman type with 12-point leading (single-spaced), and a 7" × 9" text block. Please ensure that your submission remains fully legible when printed or rendered in grayscale.

The main body of the submission must not exceed 13 pages, excluding references and appendices. The complete paper, including references and appendices, must not exceed 30 pages. The main body must be self-contained, and reviewers are not required to read the appendices.

Attempts to circumvent the page limit by reducing whitespace (e.g., using negative vspace, savetrees, or titlesec, or removing author blocks) or by altering the default formatting (e.g., changing fonts) are prohibited and may result in rejection.

Rev. 11/04/26

Submission Instructions and Procedures

By the submission deadline, all authors must (1) create or update their HotCRP profiles to include their ORCID and (2) confirm the submission terms in HotCRP. If a co-author is unavailable and cannot fulfill these requirements, please contact the Program Co-Chairs and explain the circumstances. Except in such exceptional cases, papers for which any author has not completed these requirements will be desk rejected.

Submissions must not substantially overlap with work that has already been published or is simultaneously under review by a journal or a conference or workshop with proceedings. Duplicate submission, once detected, will result in desk rejection and may lead to restrictions on submissions to future SAIS conferences. To identify potential duplicate submissions, the Program Committee may share relevant information with the chairs of other conferences and workshops or with journal editors.

Anonymous Submission for Double-Blind Review

Papers must be prepared for anonymous review. Author names and affiliations must not appear in the paper, and authors must not otherwise reveal their identities in the submission. References to the authors' prior work should be written in the third person, as if the work had been conducted by others. A blinded reference should be used only in the unusual case where referring to it in the third person would likely reveal the authors' identities. For submitted artifacts, neither links included in the paper nor the linked websites may reveal the authors' identities or track reviewers who access them. Papers that are not properly anonymized may be desk rejected.

Although submissions must be anonymous, authors may give talks about their work, post preprints online, disclose security vulnerabilities to vendors or the public, or otherwise discuss their work during the review process. Authors should, however, avoid timing publicity or public announcements to coincide with the review period, or engaging in other behavior that could reasonably be perceived as an attempt to reveal their identities to reviewers.

Open Science Policy

All submissions must include an Open Science Appendix describing the artifacts and how to access them or an explicit description of why artifacts cannot be provided. To preserve double-blind review, artifacts must be shared through anonymous links with no tracking mechanisms. Authors are solely responsible for ensuring no identifying information is exposed (e.g., usernames, organization names, commit history). For source code repositories, we recommend https://anonymous.4open.science, using the conference ID "SAIS27" to apply the review cycle's expiration deadline. Check any expiration times carefully: files must remain accessible for the entire review period, until the final approval deadline.

Anonymous URLs should be included in the paper's Open Science Appendix. After the submission deadline, URLs are final, but a 3-day grace period is provided to anonymize and upload artifacts. During this period, the paper itself cannot be modified. Once the grace period ends, artifacts are considered frozen and must not be updated.

We expect all submitted artifacts to be available publicly on paper acceptance; however, we acknowledge that some artifacts cannot be shared publicly, e.g., due to licensing restrictions, adversarial risk, or harm to study subjects. Artifacts help the Program Committee to review the paper; therefore, if at all possible, authors should submit the artifacts and note in the Open Science Appendix those which will not be made public. If submission of artifacts is not possible (e.g., due to NDA or legal constraints), the Open Science Appendix must explain such omissions explicitly.

For all submitted artifacts that can be made public, paper acceptance is conditional on their public availability. Camera-ready paper versions will include an Open Science Appendix that includes a non-anonymous, stable artifact link.

Ethical Considerations

Some submissions raise ethical considerations that should be explicitly addressed. In such cases, authors should include a clearly labeled section in the appendix (e.g., "Ethics" or "Ethical Considerations") discussing the ethical implications of their work.

For example, submissions involving human subjects or raising ethical or legal concerns relevant to the SAIS community should state whether an appropriate ethics review (e.g., IRB review or approval) was conducted and explain how the relevant ethical and legal concerns were addressed. Submissions reporting potentially high-impact vulnerabilities should describe in detail the steps the authors have taken, or plan to take, to mitigate the associated risks, such as responsibly disclosing the vulnerabilities to affected vendors.

Authors of submissions that raise ethical concerns but do not provide an adequate discussion may be asked to provide additional information. The Program Committee reserves the right to reject a submission if the authors do not provide sufficient evidence that relevant ethical or legal concerns have been appropriately addressed.

Use of AI

Submitted work should respect the integrity of scientific inquiry and the time and effort of reviewers. The use of AI tools to assist in preparing submissions is permitted. However, human authors remain fully responsible for all submitted content and results, including their accuracy and integrity.

By submitting to USENIX SAIS '27, authors affirm that they have carefully read and take responsibility for all AI-generated content included in the submission as if they had produced it themselves, including text, experimental data, and references.

Violations of this policy that result in fabricated or hallucinated content, nonexistent or incorrect references, misattributed authorship, invented claims, falsified results, or poor-quality exposition may result in rejection.

Poor-quality exposition can arise with or without the use of AI tools. However, excessive reliance on AI tools can exacerbate the problem by reducing the time authors spend engaging closely with their own writing: organizing ideas into a logical structure, deciding what should be included or omitted, and working through different formulations to express technical ideas as clearly and precisely as possible. For authors who are still developing their technical writing skills, extensive reliance on AI tools can also deprive them of valuable opportunities to practice and improve those skills. Examples of poor-quality exposition include:

- Excessive use of vague or grandiose language that obscures rather than clarifies the technical contents.
- Adjectives, nouns, or phrases that sound impressive but convey little substantive meaning.
- Overly complicated sentence structures with clauses that can be removed without loss of meaning.

- Repeated explanations of the same idea in multiple places.
- Excessive or unnecessary use of contrastive constructions and transitions.
- Inconsistent terminology, such as using different terms for the same concept without a clear reason.
- Unnecessary formalizations, notation, or formulas that do not improve precision, clarity, or understanding.

Review and Publication Process

All submissions will be judged on relevance, originality, correctness, and quality of exposition. Accepted papers will be published in the proceedings of USENIX SAIS '27.

Conflicts of Interest

At submission time, authors must identify any conflicts of interest (COIs) with Program Committee members and specify the nature of each conflict. COIs include advisor-advisee relationships, both current and former; current affiliation with the same institution; professional collaborations within the past two years, regardless of whether they resulted in a publication or funding; and close personal relationships. Program Committee members who have a conflict with a submission will be excluded from its evaluation.

Authors must not declare conflicts for the purpose of excluding Program Committee members with whom no genuine conflict exists. Doing so may constitute grounds for rejection. The Program Co-Chairs reserve the right to request additional information or justification for any declared conflict. Authors who have concerns about whether their submissions will receive fair treatment should instead contact the Program Co-Chairs and provide a clear explanation of the circumstances and any special consideration they are requesting.

Final versions of accepted papers must acknowledge all sources of funding. Authors should also disclose any affiliations, financial or other interests, or other circumstances that may be relevant to readers in interpreting the work and its implications. For examples of relevant disclosures, authors may wish to consult the 2026 IEEE S&P Financial Conflicts Policy (www.ieee-security.org/TC/SP2026/financial-con.html).

Confidentiality of Submissions

The Program Committee and external reviewers are required to treat all submissions, including associated artifacts, as confidential. However, the Program Co-Chairs or designated committee members may share submissions with the chairs of other conferences for the limited purpose of identifying duplicate or concurrent submissions. Submissions accompanied by nondisclosure agreements will not be considered.

Accepted submissions will remain confidential until they are published on the USENIX SAIS '27 website. Rejected submissions will continue to be treated as confidential indefinitely.

Desk Rejection

The Program Co-Chairs reserve the right to desk reject submissions when they determine that further review is not warranted. Reasons may include, but are not limited to, serious editorial or presentation problems, lack of fit with the conference scope, insufficient technical quality, fabricated or incorrect references, or failure to comply with the CFP.

Unlike many other major security conferences, SAIS uses a single-round review process. To avoid unnecessarily burdening reviewers, a submission may be desk rejected within approximately one week of submission if the Program Co-Chairs determine that it has substantial technical or presentation deficiencies or is unlikely to be competitive for acceptance. Submissions that proceed beyond this initial screening will undergo the full review process.

Paper Availability

As part of USENIX's open-access policy, all papers will be available online before the conference. If your accepted paper should not be published prior to the event, please notify production@usenix.org after you submit your final accepted paper. USENIX also allows authors to retain ownership of the copyright in their works, requesting only that USENIX be granted the right to be the first publisher of that work. See our sample consent form (www.usenix.org/sites/default/files/consent_author_proceedings.pdf) for the complete terms of publication.

By submitting a paper, you agree that, if the paper is accepted, at least one of the authors will register to attend the conference at full price (i.e., not the student rate) and to present the paper; USENIX members at the Sustainer level and higher may apply their membership discounts to their registrations. If an author plans to present more than one paper, one full-price registration will still be required for each paper. If the conference registration fee will pose a hardship for the presenter of an accepted paper, please contact conference@usenix.org.

Alternative arrangements can be made if there are justifiable difficulties in travel and will be allowed only on a case-by-case basis with permission from the Program Co-Chairs.

Authorship

USENIX expects authors to adhere to high ethical standards and to ensure transparency and integrity in the publication of their work. Clear and consistent authorship practices support USENIX's core mission and are essential to maintaining confidence in the research record. See the USENIX Authorship Policy (www.usenix.org/conferences/authorship-policy) for complete information, including criteria for authorship, details about pseudonymous authorship, and the difference between contributors and authors.

Contact Information

Specific questions about submissions may be sent to the Program Co-Chairs at sais27chairs@usenix.org. The chairs will respond to questions about the submission process if contacted at least a week before the submission deadline.

Further questions? Contact the Program Co-Chairs, sais27chairs@usenix.org, or the USENIX office, submissionspolicy@usenix.org.

Conference Organizers

General Chair

Mihai Christodorescu, *Google*

Program Co-Chairs

Earlence Fernandes, *Gray Swan AI and University of California, San Diego*
Ninghui Li, *Perplexity AI and Purdue University*

Program Committee

TBA

Steering Committee

Kamalika Chaudhuri, *Google DeepMind*
Mihai Christodorescu, *Google*
Anupam Datta, *Snowflake*
Earlence Fernandes, *Gray Swan AI and University of California, San Diego*
Casey Henderson-Ross, *USENIX Association*
Somesh Jha, *Google*
Edward Suh, *NVIDIA*
Florian Tramèr, *ETH Zurich*
David Wagner, *University of California, Berkeley*