Shlok Gilda

✉ shlokgilda@ufl.edu   shlokgilda   shlokgilda   shlokgilda

Education

2021 – 2026 **Ph.D. Computer Science, University of Florida**.
Research Interests: *Open-Source Software Supply Chain Security; Misinformation Analysis; Natural Language Processing; Artificial Intelligence*
Thesis Title: *How Communication Dynamics Shape Vulnerability Management in Open-Source Software*
Advisor: *Dr. Bonnie Dorr*
GPA: *3.96/4.0*

2021 – 2022 **M.Sc. Computer Science, University of Florida**.
Advisor: *Dr. Daniela Oliveira*
GPA: *3.96/4.0*

2014 – 2018 **B.E. Computer Engineering, University of Pune**.
Thesis Title: *User Privacy in Consumer IAM.*
Advisor: *Dr. Geetanjali Kale*
GPA: *3.56/4.0*

Research Publications

Conference Proceedings

1. **S. Gilda**, K. Martiny, J. Ho, L. Tinnel, G. Denker, and B. J. Dorr, "Navigating the Blue Nowhere: A Framework for Mapping Validated Adversarial Trajectories," in *ICDM Workshop 2025*, Presented, IEEE, 2025. URL: https://ascend-data.sri.com/docs/publications/ascend-2025-GTA.pdf.

2. Q. Yang, T. Christensen, **S. Gilda**, J. Fernandes, D. Oliveira, R. Wilson, and D. Woodard, "Are Fact-Checking Tools Helpful? An Exploration of the Usability of Google Fact Check," in *5th EAI International Conference on Data and Information in Online Systems*, 2024. DOI: https://doi.org/10.1007/978-3-031-97352-9_7.

3. L. Giovanini, **S. Gilda**, M. Silva, F. Ceschin, P. Shrestha, C. Brant, J. Fernandes, C. S. Silva, A. Grégio, and D. Oliveira, "People Still Care About Facts: Twitter Users Engage More with Factual Discourse than Misinformation," in *Security and Privacy in Social Networks and Big Data*, **Luiz Giovanini and Shlok Gilda are co-first authors. Best Paper Award.**, Singapore: Springer Nature Singapore, 2023, pp. 3–22, ISBN: 978-981-99-5177-2. DOI: https://doi.org/10.1007/978-981-99-5177-2_1.

4. **S. Gilda**, T. Jain, and A. Dhalla, "None Shall Pass: A Blockchain-Based Federated Identity Management System," in *Inventive Computation and Information Technologies*, Singapore: Springer Nature Singapore, 2022, pp. 329–352, ISBN: 978-981-19-7402-1. DOI: https://doi.org/10.1007/978-981-19-7402-1_24.

5. **S. Gilda**, L. Giovanini, M. Silva, and D. Oliveira, "Predicting Different Types of Subtle Toxicity in Unhealthy Online Conversations," 12th International Conference on Emerging Ubiquitous Systems and Pervasive Networks / 11th International Conference on Current and Future Trends of Information and Communication Technologies in Healthcare, vol. 198, 2021, pp. 360–366. DOI: https://doi.org/10.1016/j.procs.2021.12.254.

6. **S. Gilda** and M. Mehrotra, "Blockchain for Student Data Privacy and Consent," in *2018 International Conference on Computer Communication and Informatics (ICCCI)*, 2018, pp. 1–5. DOI: 10.1109/ICCCI.2018.8441445.

7. **S. Gilda**, “Source Code Classification using Neural Networks,” in *2017 14th International Joint Conference on Computer Science and Software Engineering (JCSSE)*, 2017, pp. 1–6. doi: 10.1109/JCSSE.2017.8025917.

8. **S. Gilda**, H. Zafar, C. Soni, and K. Waghurdekar, “Smart Music Player Integrating Facial Emotion Recognition and Music Mood Recommendation,” in *2017 International Conference on Wireless Communications, Signal Processing and Networking (WiSPNET)*, 2017, pp. 154–158. doi: 10.1109/WiSPNET.2017.8299738.

US Patents

1. P. Gokhale, **S. Gilda**, S. Malik, S. H. Rizvi, and R. Poulose, “Identity Attribute Confidence Scoring while Certifying Authorization Claims.” url: https://uspto.report/patent/app/20200322342.

Under Review

1. S. Gilda and **S. Gilda**, *Principled Design for Epistemic Accountability in AI-Assisted Engineering*, in *ICLR 2026 Workshop Trustworthy AI*, **Sankalp Gida and Shlok Gilda are co-first authors.**, 2026. doi: https://doi.org/10.48550/arXiv.2601.21116.

2. K. Yamoah, G. Agyapong, N. Parekh, D. Brinkley, C. Jayaweera, **S. Gilda**, B. J. Dorr, E. Dorley, and K. Scroggins, *An Elicitation-Matrix Approach to Pragmatic Context Modeling in Low-Resource Machine Translation: The Case of Akuapem Twi*, in *The 39th International Florida AI Research Society (FLAIRS) Conference*, 2026.

In-Progress

1. **S. Gilda**, M. Botacin, and B. Dorr, *Temporal Evolution of Security Concerns in OSS: Investigating the Role of Contributor Characteristics and Behaviors.*

2. **S. Gilda** and B. Dorr, *Developing a Communication-Based Health Score for OSS Projects: Insights and Recommendations.*

Invited Talks

2025

- **Predicting OSS Vulnerabilities Through Communication Analysis: A Work in Progress**, OpenSSF Community Day North America, Colorado.
- **Communication-Driven OSS Security**, Invited Talk for NLP Applications Course, University of Florida, Florida.

Employment History

Jan. 2021 – · · · ·

Graduate Research Assistant, University of Florida.

* Advancing research at the intersection of Natural Language Processing and Cybersecurity under the supervision of Dr. Bonnie Dorr. My work centers on two key areas: developing a thesis that analyzes how communication dynamics influence open-source software security, and developing a neuro-symbolic framework for an IARPA-funded project to construct validated, temporally-aware knowledge graphs from multi-modal Cyber Threat Intelligence (CTI). Previously advised by Dr. Daniela Oliveira.

* Developing a thesis titled, *"How Communication Dynamics Shape Vulnerability Management in Open-Source Software"*, which integrates longitudinal analysis of open-source repositories to study how communication dynamics (e.g., sentiment, toxicity, topics, stances, and outrage) influence their security posture. As part of this work, designing and implementing the ***FORCE: Framework for Open-source Risk and Community Evaluation*** to analyze temporal vulnerability evolution in open-source repositories and correlate it with vulnerability dynamics.

* Spearheading the development of the Cyber Behavior Pattern Extractor (CBPE), a neuro-symbolic framework designed to address factual unreliability in Large Language Models (LLMs) when processing CTI reports. This work introduces a novel, two-stage automated validation loop that verifies syntactic and semantic correctness against source material, eliminating the need for a pre-existing trusted knowledge base. The pipeline formalizes multi-modal CTI data (text and images) into Concrete Syntax Trees (CSTs) to build a validated, temporally-aware knowledge graph for modeling adversarial behavior.

Jan. 2024 – May 2024

AI Resident, SandboxAQ.

* Led the development of a machine learning pipeline to classify cryptographic strengths, achieving 73% accuracy and an AUC of 0.79 on a dataset of over 300, 000 encrypted files across 8 cryptographic algorithms.

* Integrated 7 novel randomness features through advanced statistical analyses, enhancing feature extraction and significantly boosting model performance.

* Managed the project end-to-end, from dataset curation and experimental design to presenting findings to stakeholders, demonstrating the potential of machine learning in cryptographic security assessments.

Jun. 2023 – Aug. 2023

Research Intern, Accenture Security Labs.

* Led a data science initiative at Accenture, analyzing 100, 000+ commits and 500 users across 20 OSS repositories using TensorFlow, Neo4J, and Python to identify malicious developers.

* Engineered a Python-based data pipeline for Git/GitHub metadata, employing graph-based models and clustering algorithms (K-means, DBSCAN) for enhanced data analysis and community detection.

* Formulated and validated a machine learning ruleset for user classification, presenting key cybersecurity insights to senior leadership, demonstrating potential industry applications.

Employment History (continued)

Apr. 2020 – Dec. 2020

Software Engineer, Moxie.xyz.
* Successfully enhanced Moxie's user sign-up and onboarding experience by integrating OAuth 2.0 with Facebook and Instagram, streamlining access and increasing user engagement.
* Achieved a remarkable 99.9% data availability at Moxie by managing extensive user data with Apache Cassandra, ensuring robust data handling capabilities for thousands of daily user interactions.
* Revolutionized media processing on the Moxie platform by developing advanced video recording and compression features using FFMPEG, achieving a 40% increase in efficiency and significantly improving user experience.

Jun. 2019 – Apr. 2020

Software Engineer, Pepo.com.
* Boosted user engagement at Pepo by 35% by developing a personalized feed algorithm that delivered tailored content, significantly enhancing user satisfaction and platform stickiness.
* Enhanced the user onboarding experience by streamlining sign-up and authentication processes through seamless OAuth 2.0 integration with major social platforms, facilitating easier access and increased user growth.
* Elevated app responsiveness and user interaction at Pepo by implementing WebSockets, leading to a 25% improvement in real-time communication efficiency, enriching the user experience.
* Leveraged Apache Cassandra for robust data storage solutions and integrated Google Firebase Cloud Messaging (FCM) for precise in-app and push notifications, driving user engagement and improving key platform metrics.

Jun. 2018 – Aug. 2020

Software Engineer, Ost.com.
* Enabled secure and efficient blockchain transactions on the OST Platform by developing a REST API with NodeJS and Ruby on Rails, seamlessly integrating Ethereum blockchain to support over 1, 000 transactions/second.
* Enhanced the platform's security and scalability by implementing peer-to-peer (P2P) technologies and data encryption, ensuring the safe handling of thousands of consumer-app tokenization transactions.
* Achieved exceptional system throughput of over 500 transactions per second by adeptly utilizing technologies such as RabbitMQ, Memcached, Redis, ElasticSearch, and AWS DynamoDB, facilitating robust multi-chain support and high-performance operations.
* Significantly improved platform scalability and user experience by innovating with database sharding and smart contract-based user account recovery methods, leading to a 40% increase in overall system performance.

Employment History (continued)

Jun. 2017 – Jun. 2018

Research Intern, IBM India Software Labs.
* Played a pivotal role at IBM in co-developing a Hyperledger Fabric-based IAM system, incorporating zero-knowledge authentication and advanced cryptographic schemes like ECC and HMAC-SHA512, substantially enhancing the security of user identity verification processes.
* Elevated data security and user sovereignty by implementing cutting-edge access control measures, including split-key cryptography and proxy re-encryption, enabling secure and authorized data access by identity authorities without compromising user control.
* Streamlined the process of secure identity claims transfer and efficient blockchain data retrieval by integrating and customizing OpenID Connect within the Websphere Liberty Server, enhancing system interoperability and user convenience.
* Co-authored a US patent for an innovative method of calculating identity attribute trust scores, making a significant contribution to the project's intellectual property and setting a new standard in identity verification technology.

Teaching Experience

Spring 2026	**Instructor on Record**, CAP 4641 Natural Language Processing, University of Florida.
Fall 2025	**Lead Teaching Assistant**, CAI 6307 Natural Language Processing, University of Florida.
Spring 2025	**Teaching Assistant**, CAI 6307 Natural Language Processing, University of Florida.
Fall 2024	**Teaching Assistant**, CAP 4641 Natural Language Processing, University of Florida.
Summer 2024	**Teaching Assistant**, COP 3530 Data Structures and Algorithms, University of Florida.

Service

2026	**Reviewer**, LREC.
2025	**Reviewer**, IEEE Transactions on Privacy.
	Reviewer, IEEE Transactions on Dependable and Secure Computing.
	Reviewer, COLING.
2024	**Reviewer**, IEEE Access.
	Reviewer, LREC-COLING.
	Artifact Evaluation Program Committee, Usenix Security.
	Program Committee, Eighth Workshop on Online Abuse and Harms (WOAH).
	Program Committee, Computing Conference.
2023	**Program Committee**, Seventh Workshop on Online Abuse and Harms (WOAH).
	Program Committee, Usenix SOUPS Posters.
	Program Committee, Computing Conference.
2022	**Student Volunteer**, ACM CSCW.
	Program Committee, Sixth Workshop on Online Abuse and Harms (WOAH).
	Program Committee, Usenix SOUPS Posters.
	Reviewer, IEEE Open Journal of the Computer Society.
	Program Committee, 2nd International Conference on Emerging Trends and Innovations in ICT.

Service (continued)

2019 **Reviewer**, IEEE Access.

Skills

Coding	Python, JavaScript, Node.JS, C, C++, SQL.
Databases	MySQL, DynamoDB, Neo4J, Cassandra.
ML Frameworks	PyTorch, Tensorflow, Scikit-Learn, spaCy.

Miscellaneous Experience

Awards and Achievements

2025 **Travel Grant**, Linux Foundation Open Source Summit 2025.

2023 **Best Paper Award**, SocialSec 2023.

2022 **Student Conferenceship**, ACSAC 2022.

2021 **Student Travel Grant**, IEEE S&P 2021.

Student Grant, Usenix Enigma 2021.

Certification

2018 **Deep Learning Specialization**. Awarded by Coursera.org.

Sequence Models. Awarded by Coursera.org.

Convolutional Neural Networks. Awarded by Coursera.org.

Improving Deep Neural Networks: Hyperparameter Tuning, Regularization and Optimization. Awarded by Coursera.org.

Neural Networks and Deep Learning. Awarded by Coursera.org.

Structuring Machine Learning Projects. Awarded by Coursera.org.

References

Available on Request