

Sikha Pentyala

✉ sikha@uw.edu  sikhapentyala.github.io  Google Scholar  LinkedIn  ORCID

Summary

AI researcher with a focus on Responsible and Trustworthy AI, specializing in privacy-preserving techniques. My current research focus spans privacy in AI and synthetic data generation.

Appointments and Experience

Assistant Professor (Tenure-track) School of Engineering & Technology University of Washington Tacoma (UW Tacoma)	starts Sep 2026 <i>Tacoma, WA</i>
Research Scientist UW School of Nursing ◦ Development of AI Tutor and responsible AI micromodules	Jul 2026 – Sep 2026 <i>Seattle, WA</i>
Postdoctoral Researcher UW Data Science Postdoctoral Fellow, UW eScience Institute – University of Washington Tacoma ◦ Research on privacy-preserving synthetic data generation for biomedical data	Jul 2025 – Jul 2026 <i>Seattle, WA</i>
AI Research Summer Associate JPMorgan Chase ◦ Research on the interpretability of time series models	Jun – Aug 2024 <i>New York, NY</i>
AI Research Summer Associate JPMorgan Chase ◦ Research on the intersection of privacy and explainability in ML	Jun – Aug 2023 <i>New York, NY</i>
Pre-doctoral Instructor University of Washington Tacoma ◦ Served twice as primary instructor for Fundamentals of OOP	Dec 2022 – Jun 2023 <i>Tacoma, WA</i>
Research Intern Mila – Quebec AI Institute ◦ Research on privacy-preserving fair item rankings and on training ML models in a federated setting while preserving user privacy and achieving group fairness	Jul 2022 – Mar 2022 <i>Remote</i>
Research Assistant – Privacy-preserving machine learning University of Washington Tacoma	Mar 2022 onwards <i>Tacoma, WA</i>
Scientific Officer Nuclear Power Corporation of India Ltd.	Sep 2011 – Apr 2018 <i>Kudankulam, TN, India</i>

Education

Ph.D. in Computer Science and Systems University of Washington Tacoma (GPA: 4.0/4.0)	Spring 2022 – Spring 2025
M.S. in Computer Science and Systems University of Washington Tacoma (GPA: 4.0/4.0)	2019 – 2020
Bachelor of Computer Science Jawaharlal Nehru Technological University, Anantapur (Gold Medalist)	2005 – 2009

Awards

Major Fundings and Grants

- [2026 AI@UW's SEED-AI Grant](#)  for Advancing Responsible and Ethical AI Literacy for Healthcare Professionals and Personalized Education (co-PI)
- 2023 JP Morgan Chase Fellowship to support PhD research on synthetic data generation (PhD Fellowship)
- Carwein-Andrews Distinguished Fellowship for 2023-2024, UW Tacoma, School of Engineering and Technology to support PhD research on synthetic data generation (PhD Fellowship)

Competitions

- [1st prize](#)  in red teaming for bulk RNA-seq of CAMDA 2026 Health Privacy Challenge
- 1st and 2nd place respectively in the tracks for bulk RNA-seq and single-cell RNA-seq synthetic data generation of the [CAMDA 2025 Health Privacy Challenge](#) 
- [Winner of NIST CRC Challenge 2025](#)  - *Can you re-identify the CRC synthetic data samples (i.e., reconstruct targeted individuals' correct attribute values)?*
- [SNAKE Challenge](#) : won 1st place on membership attacks against synthetic data
- [U.S.-U.K. PETs Prize Challenges](#) : 2nd prize in Phase 2 of Track A
- [Winner of Track III of the iDASH2021 secure genome analysis competition](#) 

Recognitions, Other Fundings and Miscellaneous

- 2024-2025 School of Engineering & Technology's Student Justice Equity Diversity & Inclusion Award, UW Tacoma
- 2024-2025 School of Engineering & Technology's Outstanding Graduate Research Award, UW Tacoma
- 2023-2024 School of Engineering & Technology's Outstanding Student Leadership Award, UW Tacoma
- 2023 UPE Executive Council Award by UPE International Honor Society for the Computing and Information Disciplines
- 2023 Andrew and Julie Fry Innovation Award, UW Tacoma
- 2022-2023 School of Engineering & Technology's Student Justice Equity Diversity & Inclusion Award, UW Tacoma
- 2022-2023 School of Engineering & Technology's Outstanding Graduate Research Award, UW Tacoma
- WiML travel funding award (WiML@NeurIPS 2022), PPAI travel funding award (PPAI@AAAI 2024), ICML travel funding award (ICML 2024), UWT CTF awards (2022,2023,2024), Graduate Student Conference Presentation Award (2022, 2024), AAAI NSF travel award (2025)
- Carwein-Andrews Endowment award for 2022-2023, UW Tacoma, School of Engineering and Technology
- 2021-2022 School of Engineering & Technology's Outstanding Graduate Research Award, UW Tacoma
- Gold Medalist at Jawaharlal Nehru Technological University, Anantapur, ECM, 2009

Publications

Proceedings

- M. Kumar, Q. Lou, P. Barreto, M. De Cock, **S. Pentyala**. *FHAIM: Fully Homomorphic AIM For Private Synthetic Data Generation* [↗](#), 43rd International Conference on Machine Learning (ICML), 2026 (acceptance rate: 26.6%)
- **S. Pentyala**, A. Babu, J. Ahn, M. De Cock. *Federated Synthetic Data Generation for Hepatology Research: A UNOS based Evaluation* [↗](#), 17th ACM Conference on Bioinformatics, Computational Biology, and Health Informatics (ACM BCB), 2026 (acceptance rate: <20%)
- **S. Pentyala**, Z. Pan, P.J. McKeever, J. Banerjee, L. Foschini, M. De Cock. *Synthetic Germline VCF Generation for Rare Diseases: Case Study in NF1* [↗](#), 17th ACM Conference on Bioinformatics, Computational Biology, and Health Informatics (ACM BCB), 2026
- P. Garg, S. Iribarren, **S. Pentyala**, Y. Rodriguez, P. Carmiol-Rodriguez, A. Vidrio, C. Kwanin, J. Sprecher, J. Roberti. *Designing Structured Conversational Support for Tuberculosis Treatment Adherence and Patient Coping* [↗](#), 10th Workshop on Computational Linguistics and Clinical Psychology (CLPsych), 2026
- D. Filienko, M. De Cock, **S. Pentyala**. Federated Generation of Synthetic RNA-seq Data, Third IEEE Secure and Trustworthy Machine Learning Conference 2025, 3rd RECOMB Satellite Conference on Biomedical Data Privacy, 2026
- S. Golob, **S. Pentyala**, A. Maratkhan, M. De Cock. *Privacy Vulnerabilities in Marginal-based Synthetic Data* [↗](#), Third IEEE Secure and Trustworthy Machine Learning Conference 2025, (acceptance rate: 29%)
- **S. Pentyala**, N. Neophytou, A. Nascimento, M. De Cock, G. Farnadi *Privacy-Preserving Group Fairness in Cross-Device Federated Learning* [↗](#), Proceedings of Machine Learning Research (PMLR) 2024 (acceptance rate: 16%)
- **S. Pentyala**, M. Pereira, M. De Cock. *CaPS: Collaborative and Private Synthetic Data Generation from Distributed Sources* [↗](#), 41st International Conference on Machine Learning (ICML), 2024 (acceptance rate: 27.9%)
- S. Golob, **S. Pentyala**, R. Dowsley, B. David, M. Larangeira, M. De Cock, A. Nascimento. *A Decentralized Information Marketplace Preserving Input and Output Privacy* [↗](#), Proceedings of the 2nd Data Economy Workshop (DEC23) - SIGMOD 2023 Workshop, 2023
- Jia Ao Sun, **S. Pentyala**, M. De Cock and G. Farnadi. *Privacy-Preserving Fair Item Ranking* [↗](#), 45th European Conference on Information Retrieval 2023 (acceptance rate: 29%)
- **S. Pentyala**, R. Dowsley and M. De Cock *Privacy-Preserving Video Classification with Convolutional Neural Networks* [↗](#), 38th International Conference on Machine Learning (ICML), 2021 (acceptance rate: 21.46%)
- F. Tabet and **S. Pentyala**, B. Patel, et al. *OSMRRunner: A System for Exploring and Fixing OSM Connectivity* [↗](#), 22nd IEEE International Conference on Mobile Data Management (MDM), 2021

Journals

- S. Golob, **S. Pentyala**, M. De Cock. *SoK: Reconstruction Attacks on Synthetic Tabular Data (Insights from Winning the NIST CRC)*, 26th Privacy Enhancing Technologies Symposium (PoPETs), 2026 (acceptance rate: 4%)
- R.J.M. Maia, D. Ray, **S. Pentyala**, R. Dowsley, M. De Cock, A. Nascimento and R. Jacobi.

2023. *An End-to-End Framework for Private DGA Detection as a Service* [🔗](#), PLOS ONE 2024 (acceptance rate: 20%)

- J. Vos, **S. Pentyala**, S. Golob, R. Maia, D. Kelly, C. Martins, Z. Erkin, M. De Cock, A. Nascimento. *Privacy-Preserving Membership Queries for Federated Anomaly Detection* [🔗](#), Proceedings on Privacy Enhancing Technologies (PoPETS), 2024 (acceptance rate: 20%)

Workshops and Abstracts

- **S. Pentyala**, Z. Pan, L. Foschini, M. De Cock, J. Banerjee. *Auditing Patient Privacy Risk in Synthetic Rare Disease Germline Data*, 3rd RECOMB Satellite Conference on Biomedical Data Privacy, 2026
- **S. Pentyala**, P. McKeever, J. Banerjee, L. Foschini, M. De Cock. *Privacy-Preserving Genomics Data Sharing on the NAIRR*, NAIRR Annual Meeting, 2026
- D. Filienko, E.C. Neto, **S. Pentyala**, J. Banerjee, L. Foschini, M. De Cock. From Graphical Models to Foundation Models: Synthetic Bulk RNA-seq Data Generation. 34th Conference on Intelligent Systems for Molecular Biology & 25th European Conference on Computational Biology (ISMB/ECCB), 2026
- C. Jarell, D. Filienko, J. Kim, E. Szebenyi, **S. Pentyala**, S. Golob, L. M. De Cock. Does Synthetic Bulk RNA-seq Data Protect Donors? Privacy Auditing through Membership Inference Attacks. 34th Conference on Intelligent Systems for Molecular Biology & 25th European Conference on Computational Biology (ISMB/ECCB), 2026
- A. Babu, J. Ahn, M. De Cock, **S. Pentyala**. Federated Synthetic Data Generation to Accelerate Hepatology Research. ACM Cascadia Women in Computing Celebration, 2025
- P. McKeever, D. Filienko, S. Golob, S. Menzies, **S. Pentyala**, J. Banerjee, L. Foschini, M. De Cock. Comparison of Single Cell RNA Synthetic Data Generators: a CAMDA Health Challenge Analysis. 33rd Conference on Intelligent Systems for Molecular Biology & 24th European Conference on Computational Biology (ISMB/ECCB), 2025
- S. Menzies, **S. Pentyala**, D. Filienko, S. Golob, J. Banerjee, L. Foschini, M. De Cock. Synthetic Data Generation for bulk RNA-seq Data: a CAMDA Health Challenge Analysis. 33rd Conference on Intelligent Systems for Molecular Biology & 24th European Conference on Computational Biology (ISMB/ECCB), 2025
- **S. Pentyala**, G. Sitaraman, T. Claar, M. De Cock. *End to End Collaborative Synthetic Data Generation*, PPAI-2025 (AAAI-25 Workshop on Privacy-Preserving Artificial Intelligence), 8 pages
- **S. Pentyala**, S. Menzies, M. De Cock. *Privacy-Preserving Synthetic Data Generation on the NAIRR*, NAIRR Pilot Inaugural Annual Meeting, 2025
- T. Claar, S. Golob, **S. Pentyala**, G. Sitaraman, M. De Cock, J. Banerjee, L. Foschini. *Securely Generating Synthetic Genomic Data from Distributed Data Silos*, 11th International Workshop on Genome Privacy and Security (GenoPri'24), 2024
- S. Golob, **S. Pentyala**, A. Maratkhan, M. De Cock. *High Epsilon Synthetic Data Vulnerabilities in MST and PrivBayes*, PPAI-2024 (AAAI-24 Workshop on Privacy-Preserving Artificial Intelligence), 8 pages
- **S. Pentyala**, S. Sharma, S. Kariyappa, F. Leuce and D. Magazzeni. *Privacy-Preserving Algorithmic Recourse*, 3rd International Workshop on Explainable AI in Finance, ICAIF 2023, 8 pages
- M. Pereira, **S. Pentyala**, A. Nascimento, R. T. de Sousa Jr., M. De Cock *Secure Multi-party Computation for Synthetic Data Generation from Distributed Data*, SyntheticData4ML

Workshop@NeurIPS 2022, 6 pages

- **S. Pentyala**, D. Melanson, M. De Cock and G. Farnadi. *PrivFair: a Library for Privacy-Preserving Fairness Auditing*, PPAI-2022 (AAAI-22 Workshop on Privacy-Preserving Artificial Intelligence), 8 pages
- **S. Pentyala**, N. Neophytou, A. Nascimento, M. De Cock, G. Farnadi . *PrivFairFL: Privacy-Preserving Group Fairness in Federated Learning*, Montreal AI Symposium (MAIS) 2022
- **S. Pentyala**, N. Neophytou, A. Nascimento, M. De Cock, G. Farnadi . *Towards private and fair federated learning*, WiML@NeurIPS 2022 (abstract/poster in Women in Machine Learning Workshop at NeurIPS, 2022)
- **S. Pentyala**, M. De Cock and R. Dowsley *Privacy-Preserving Video Classification*, WiCV 2021 (extended abstract/poster in Women in Computer Vision Workshop at CVPR, 2021)

Patents

- System and method for generating recourse paths with privacy guarantees (US Patent App. 18/517,211)
- System and method for low power multi-horizon time forecasting with interpretability (US Patent App. 2026/0220657)

Papers on Arxiv

- H. Ozturk, T. Afonja, J. Jalko, R. Binkyte, P. Rodriguez-Mier, S. Lobentanzer, A. Wicks, J. Kreuer, S. Ouaari, N. Pfeifer, S. Menzies, **S. Pentyala**, D. Filienko, S. Golob, P. McKeever, J. Banerjee, L. Foschini, M. De Cock, J. Saez-Rodriguez, M. Fritz, O. Stegle, A. Honkela. *Towards Useful and Private Synthetic Omics: Community Benchmarking of Generative Models for Transcriptomics Data*
- S. Golob, P. McKeever, **S. Pentyala**, M. De Cock, J. Peck. *Privacy Vulnerabilities in Synthetic Single-Cell RNA-Sequence Data.*
- **S. Pentyala**, D. Railsback, R. Maia, R. Dowsley, D. Melanson, A. Nascimento, M. De Cock. *Training Differentially Private Models with Secure Multiparty Computation.*

Invited Talks

-
- *Auditing Patient Privacy Risk in Synthetic Rare Disease Germline Data.* RECOMB Satellite Conference on Biomedical Data Privacy, 2026
 - *Illuminating Dark Data: Privacy-Preserving AI in Distributed Data Silos.* University of Central Florida, April 2026
 - *Privacy-Preserving Genomics Data Sharing on the NAIRR.* NAIRR Annual Meeting, Arlington, VA, March 2026
 - *Privacy-Preserving Synthetic Data Generation for NF1.* Children's Tumor Foundation Data Workshop / NF Conference, June 2025
 - *Privacy-Preserving Synthetic Data Generation across Data Silos.* NAIRR Pilot Inaugural Annual Meeting, Arlington, VA, March 2025
 - *Privacy-Preserving AI across Data Silos.* RAISE 2024, University of Washington Seattle, 2024
 - *Synthetic Data from Distributed Silos.* Women in Data Science (WiDS), UW Tacoma, 2024
 - *Private and Fair Federated Learning.* University of Washington Bothell, 2023
 - *Winning Solution to the U.S. PETs Prize Challenge.* Women in Data Science (WiDS), UW Tacoma, 2023

- *Privacy-Preserving Synthetic Data Generation*. SyntheticData4ML Workshop, NeurIPS, 2022
- *Recorded talks*: ICAIF 2023, AFCEP at NeurIPS 2022, PPAI at AAAI 2022, WiCV at CVPR 2021

Professional Service

Program Committees and Reviewing

- Program Committee member, AAAI 2027
- *Conferences*: Reviewed for NeurIPS 2026, ICML 2026, ACL 2026, ARR May 2026, AAAI 2026, ACM CWIC 2025, SyntheticData4ML at NeurIPS 2022 and 2023
- *Sub-reviewer*: Reviewed for CIKM 2025, ECAI 2024 and 2025, NeurIPS 2022
- *Journals*: Reviewed for Journal of Artificial Intelligence Research (2025, 2026); ACM Transactions on Knowledge Discovery from Data (2026); International Journal of Information Management (2024); IEEE Transactions on Neural Networks and Learning Systems (2022–2024); IEEE Transactions on Dependable and Secure Computing (2022)

Organizing, Outreach, and Community

- Tutorial facilitator, [AI in Practice Summer Institute](#) 
- Student poster organizing team, ACM Celebration of Cascadia Women in Computing, 2025
- Co-organizer, Women in Data Science table at the Sisterhood Event, School of Engineering & Technology, UW Tacoma
- Invited contributor, NIST [Privacy-Preserving Federated Learning Blog Series](#) 
- Invited participant, virtual workshop on Advancing Privacy and Fairness, May 2023
- Roundtable lead on privacy, Algorithmic Fairness through the Lens of Causality and Privacy (AFCEP) Workshop, NeurIPS 2022
- Volunteer, ICML 2024; Women in Machine Learning at NeurIPS 2022
- Member, Women in Data Science

Mentoring & Teaching

Mentoring

- Mentored 8+ students (more than 1 Ph.D., 4+ M.S., 4+ B.S.) at UW Tacoma, UW Bothell, and the University of Central Florida
- Mentees are co-authors on peer-reviewed publications (ICML 2026, ACM BCB 2026, CAMDA 2025, GenoPri 2024, NAIRR Abstracts 2024); 2 were awarded Mary Gates Research Scholarships
- Research mentor, Mary Gates Endowment for Students, Spring 2025
- Committee member, capstone projects, School of Engineering & Technology, UW Tacoma, 2025 and 2026

Teaching

- **Guest lecturer and TA**, TCSS 555 (Machine Learning), UW Tacoma, Fall 2025
- **TA/grader**, TCSS 555 (Machine Learning), UW Tacoma, Fall 2024
- **TA/grader**, TCSS 343 (Design and Analysis of Algorithms), UW Tacoma, Fall 2024
- **Instructor of record**, TCSS 143 (Fundamentals of Object-Oriented Programming), UW Tacoma, Spring 2023

- **Instructor of record**, TCSS 143 (Fundamentals of Object-Oriented Programming), UW Tacoma, Winter 2023

Skills

- Research Areas: Privacy Enhancing Technologies (Federated Learning, Differential Privacy, Secure MultiParty Computation), Fairness in AI, Explainable AI
- Frameworks: MP-SPDZ, FairLearn, Flower, Darts
- Tools and Technologies: Keras, Sklearn, Pandas, TensorFlow, Pytorch, OpenCV, Docker, Git (Github), Jupyter
- Programming Languages: Java, Python
- Databases: MS-SQL, ORACLE DB, MySQL
- Cloud Technologies: Amazon Web Services, Azure Containers, NAIRR
- Servers: Apache-Tomcat, WAMP/LAMP, MS AD-DC and Exchange, BlueCoat