# Nishant Vishwamitra

864-624-3133 | nishant.vishwamitra@utsa.edu | cse.buffalo.edu/∼nvishwam

## Career Objectives

To achieve information assurance for current and future information systems, continuing research and development, and education in the area of cyber security are strongly required. My career objective is to play a major role in this arena by pursuing research and development, and education programs in cyber security, Artificial Intelligence and privacy.

## Research Interests

- Artificial Intelligence (AI) Security
- Security and Privacy in Online Social Networks
- Cyberbullying and Online Hate Speech Detection, Explanation and Mitigation
- Access Control Models and Mechanisms

## Education

**University at Buffalo, SUNY** Buffalo, NY
*Ph.D. Candidate in Computer Science and Engineering* *August 2020 – June 2022*

- Advisor: Prof. Hongxin Hu
- Dissertation: Cyber-harassment Framework in Online Social Networks
- GPA: 4.00/4.00

**Clemson University** Clemson, SC
*Graduate Study in Computer Science* *August 2015 – May 2020*

- Advisor: Prof. Hongxin Hu
- GPA: 3.66/4.00

**Visvesvaraya Technological University** Bangalore, India
*Bachelor of Engineering in Electronics and Communication* *September 2007 – May 2011*

## Research Experience

- Adversarial Attacks on AI Systems
  - ∗ Proposed Multimodal Decoupling Attacks (MDA) framework to study the adversarial robustness of multimodal AI [19]
- Cyber harassment, Cyberbullying and Online Hate Speech
  - ∗ Proposed factors of COVID-19 related multimodal Hateful Memes and measurement analysis of the detection capability of multimodal models on COVID-19 related Hateful Memes. [20]
  - ∗ Proposed novel methodology for the discovery of image-based Cyberbullying related factors and designed multimodal AI for its detection. [3]
  - ∗ Designed novel explanation method of BERT (a transformer-based model) attention, and discovered novel COVID-19 related hate keywords in Twitter. [4, 6]
  - ∗ Explored cyberbullying defense using AI in mobile devices. [12]
- Privacy and Security in Online Social Networks
  - ∗ Conducted studies on how crowds validate ground truth data in crisis situations [15]
  - ∗ Conducted studies on the effect of AI granularity on trust, data mining concerns and AI aversion [16]
  - ∗ Designed a novel system for enabling automatic, content-based photo privacy management in a user-specific manner in Online Social Networks based on Collaborative Filtering AI [1]
  - ∗ Developed a taxonomy of obfuscation techniques for content-level photo privacy management in Online Social Networks. [7, 9]
  - ∗ Proposed a novel access control model for photo sharing in Online Social Networks based on protection of Personally Identifiable Information (PII) items. [11]

## Publications (* Denotes High School Students Advised by Me)

[1] **Vishwamitra, Nishant**, Yifang Li, Hongxin Hu, Kelly Caine, Long Cheng, Ziming Zhao, and Gail-Joon Ahn. “PrivacyRec: Automated content-based photo privacy recommendations”. *Proceedings of the 12th ACM Conference on Data and Application Security and Privacy (CODASPY) 2022.*

[2] Matthew Costello, Long Cheng, Feng Luo, Hongxin Hu, Song Liao, **Vishwamitra, Nishant**, Mingqi Li, and Ebuka Okpala. “COVID-19: A pandemic of anti-asian cyberhate”. *Journal of Hate Studies*, 17(1), 2021.

[3] **Vishwamitra, Nishant**, Hongxin Hu, Feng Luo, and Long Cheng. “Towards understanding and detecting cyberbullying in real-world images”. In *Proceedings of the 28th Annual **Network and Distributed System Security Symposium (NDSS) (Top conference in computer security. Known as a “Big 4” Security Conference, Acceptance rate: 15.2%)***, 2021.

[4] **Vishwamitra, Nishant**, Ruijia Hu*, Feng Luo, Long Cheng, Matthew Costello, and Yin Yang. “On analyzing covid-19-related hate speech using bert attention”. In *Proceedings of the 19th IEEE International Conference on Machine Learning and Applications (ICMLA)*, pages 669–676. IEEE, 2020.

[5] Ruijia Hu*, Wyatt Dorris*, **Vishwamitra, Nishant**, Feng Luo, and Matthew Costello. “On the impact of word representation in hate speech and offensive language detection and explanation”. In *Proceedings of the Tenth ACM Conference on Data and Application Security and Privacy (CODASPY)*, pages 171–173, 2020.

[6] Wyatt Dorris*, Ruijia Hu*, **Vishwamitra, Nishant**, Feng Luo, and Matthew Costello. “Towards automatic detection and explanation of hate speech and offensive language”. In *Proceedings of the Sixth International Workshop on Security and Privacy Analytics (IWSPA)*, pages 23–29, 2020.

[7] Yifang Li, **Vishwamitra, Nishant**, Hongxin Hu, and Kelly Caine. “Towards a taxonomy of content sensitivity and sharing preferences for photos”. In *Proceedings of the 2020 Conference on Human Factors in Computing Systems (**CHI) (Top conference in HCI. Acceptance rate: 24.3%)***, pages 1–14, 2020.

[8] Xiang Zhang, **Vishwamitra, Nishant**, Hongxin Hu, and Feng Luo. “CrescendoNet: A new deep convolutional neural network with ensemble behavior”. In *Proceedings of the 17th IEEE International Conference on Machine Learning and Applications (ICMLA)*, pages 311–317. IEEE, 2018.

[9] Yifang Li, **Vishwamitra, Nishant**, Bart P Knijnenburg, Hongxin Hu, and Kelly Caine. “Effectiveness and users’ experience of obfuscation as a privacy-enhancing technology for sharing photos”. *Proceedings of the 20th ACM Conference on Computer-Supported Cooperative Work and Social Computing (**CSCW) (Top conference in HCI)***, 1, 2017.

[10] Yifang Li, **Vishwamitra, Nishant**, Hongxin Hu, Bart P Knijnenburg, and Kelly Caine. Effectiveness and users experience of face blurring as a privacy protection for sharing photos via online social networks. In *Proceedings of the Human Factors and Ergonomics Society (HFES) Annual Meeting*, volume 61, pages 803–807. SAGE Publications Sage CA: Los Angeles, CA, 2017.

[11] **Vishwamitra, Nishant**, Yifang Li, Kevin Wang, Hongxin Hu, Kelly Caine, and Gail-Joon Ahn. “Towards PII-based multiparty access control for photo sharing in online social networks”. In *Proceedings of the 22nd ACM on Symposium on Access Control Models and Technologies (SACMAT)*, pages 155–166, 2017.

[12] **Vishwamitra, Nishant**, Xiang Zhang, Jonathan Tong*, Hongxin Hu, Feng Luo, Robin Kowalski, and Joseph Mazer. “MCDefender: Toward effective cyberbullying defense in mobile online social networks”. In *Proceedings of the 3rd ACM on International Workshop on Security And Privacy Analytics (IWSPA)*, pages 37–42, 2017.

[13] Yifang Li, **Nishant Vishwamitra**, Bart P. Knijnenburg, Hongxin Hu, and Kelly Caine. Blur vs. block: Investigating the effectiveness of privacy-enhancing obfuscation for images. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, pages 39–47, 2017.

[14] Xiang Zhang, Jonathan Tong*, **Vishwamitra, Nishant**, Elizabeth Whittaker, Joseph P Mazer, Robin Kowalski, Hongxin Hu, Feng Luo, Jamie Macbeth, and Edward Dillon. "Cyberbullying detection with a pronunciation based convolutional neural network". In *2016 15th IEEE international conference on machine learning and applications (ICMLA)*, pages 740–745. IEEE, 2016.

## Under Revision

[15] Pienta D. and **Vishwamitra N.** and Berente N. and Thatcher J.and S. Somanchi. "Why Do Crowds Validate False Data: Systematic Errors in Validating Crowdsourced Ground Truth During a Crisis". ***Management Information Systems Quarterly, (MISQ)*** **(ABS 4* Level, JCR Impact Factor 4.373, Financial Times 50) Status: 1$^{st}$ Round Invited Revision.**

## Selected Working Papers

[16] Pienta D., **Vishwamitra N.**, Thatcher J., A. Johnston, and H Hu. "Tell Us Your Secrets: Trust, Data Exfiltration, and Aversion in the Design of Protective Artificial Intelligence". Target: *Management Information Systems Quarterly (MISQ)* Status: Final Iterations.

[17] Nirmalee I. Raddatz, Dan Pienta, **Vishwamitra, Nishant**, and Jason Thatcher. "AI Transparency and Interpretability Study". Target: *Management Science Special Issue on The Human-Algorithm Connection* Status: Working.

[18] **Vishwamitra, Nishant**, Nirmalee I. Raddatz, Dan Pienta, and Jason Thatcher. "AI Fairness Study". Target: *Management Information Systems Quarterly (MISQ) Special Issue on Digital Technologies and Social Justice* Status: Final Iterations.

[19] **Vishwamitra, Nishant**, Hongxin Hu, Long Cheng, Feng Luo, and Matthew Costello. "Robustness of multimodal learning in an adversarial setting". Target: *Proceedings of the International Joint Conference on Artificial Intelligence (IJCAI) 2022* Status: Final Iterations.

[20] Zheyuan Ma, **Vishwamitra, Nishant**, Anoosha Seelam, and Hongxin Hu. "On analysis and detection of covid-19 related hateful memes". Target: *Proceedings of the International AAAI Conference on Web and Social Media (CSCW) 2022* Status: Final Iterations.

[21] **Vishwamitra, Nishant**, Keyan Guo, and Hongxin Hu. "On understadning and mitigating new waves of online hateful content". Target: *Proceedings of the USENIX Security Symposium 2022* Status: Final Iterations.

## Poster Presentations

[22] **Vishwamitra, Nishant**. "AutoPri: Collaborative photo privacy in online social networks". *AI for IndustryConference, CUiCAR, Greenville SC*, 2018.

## Leadership/Supervision Experience

### DEFCON Self-Driving Capture the Flag (CTF) 2021 — July 2021 – August 2021

*University at Buffalo, SUNY* — *Buffalo, NY*

- Led a team of four Ph.D. students and one graduate student in the DEFCON Autodriving CTF challenge. Overall, over 100 teams participated in this event all over the world.
- Our team **finished 5$^{\mathbf{th}}$** in this challenge overall, and **1$^{\mathbf{st}}$** in USA.
- We **finished 1$^{\mathbf{st}}$** all over the world in the Lane Detection challenge, a task fully handled by me.

### AI Cybersecurity Labs Development — June 2021 – Present

*University at Buffalo, SUNY* — *Buffalo, NY*

- Leading the design of the NSF funded AI Cybersecurity Labs, a suite of labs to train students in AI security issues.
- The labs are currently used as part of the course curriculum in Clemson University and North Carolina A&T University.
- Multiple talks were given at the GenCyber 2021 event hosted in North Carolina A&T University about the labs.

- Link: https://colab.research.google.com/drive/1dmwP_N7fkqZIbPVInb2XBgF4NnMUX1-p?usp=sharing

**UB Cybersecurity Research Lab Paper Presentations** August 2020 – Present

*University at Buffalo, SUNY* *Buffalo, NY*

- Led the initiative of weekly paper presentations, where each student in the lab presents and discusses research papers published in top cybersecurity and AI conferences.
- An important outcome of this initiative is our paper reading group list, a repository of all the research papers and talk videos.

**Co-supervision of Incoming Students** Dec. 2019 – Present

*University at Buffalo, SUNY* *Buffalo, NY*

- As part of my Ph.D. studies, I co-supervised multiple incoming students.
- Led Zheyuan Ma (UB Ph.D, 2021) and Anoosha Seelam (UB MS, 2020) in the analysis and measurement of COVID-19 related Hateful Memes.
- Mentored Rui Cao (Clemson University MS, 2020) in design and development of a robustness analysis framework for Multimodal AI.
- Mentored Roger Hu (Duke University BS, 2021) and Wyatt Dorris (Clemson University BS, 2021), **two high school students** from the D.W. Daniel High School, Clemson, in two research papers published at ICMLA 2021 and IWSPA 2020 on explanation and detection of hateful tweets.

**Host of Graduate Student Roundtable Meetings** Dec. 2019 – Present

*University at Buffalo, SUNY* *Buffalo, NY*

- I regularly host the graduate student roundtable meetings, held as part of the CSE department faculty hiring process.
- These meetings are crucial for the incoming faculty members to know about various research activities in the department, and also for graduate students to know about the research objectives of incoming faculty.

## Grants

**Why do crowds validate false data? Systematic errors in validating crowdsourced ground truth during a crisis** Waco, TX

*Sponsoring Agency: Baylor University Research Committee ONE-URC Program* *June 2021 – May, 2022*

- Status: Granted
- Award Amount: $4,500.00
- This project focuses on studying why crowds validate fake information online using eye-tracking studies.
- I am the **Co-PI** of the grant.

## Grant Proposal Contribution (Participated in Writing and Research Development)

**Cyber-Hostility and COVID-19** Buffalo, NY

*Sponsoring Agency: NSF* *June 2020 – May 2022*

- Status: Granted
- Award Amount: $199,996.00
- In this project, we examine the new wave of cyber-hostility encountered during the COVID-19 pandemic, by studying data collected from major social media websites.
- The sections that discussed our **expertise in the area of hate speech detection, project objectives, and workflow** were contributed by me.

**Collaborative Research: EAGER: SaTC-EDU: Learning Platform and Education Curriculum for Artificial Intelligence-Driven Socially-Relevant Cybersecurity** Buffalo, NY

*Sponsoring Agency: NSF* *June 2021 – May 2023*

- Status: Granted
- Award Amount: $130,000.00
- In this project, we will develop curricular modules and hands-on labs to educate both CS and non-CS students on AI-driven cyberharassment detection, related attacks against AI models, and social issues in AI models for cyberharassment detection.
- Overall, the six labs have been designed by me. The sections that **discussed the six labs** in this project were contributed by me.

**Collaborative Research: Elements: Integrative Cyberinfrastructure**

**for Enhancing and Accelerating Online Abuse Research** Buffalo, NY
*Sponsoring Agency: NSF* *December 2021, Under review*

- Status: Applied
- Award Amount: $300,000.00
- This project will develop the first *scalable*, *sustainable*, *customizable*, *extendable*, *portable*, and *user-friendly* Integrative Cyberinfrastructure for Online Abuse Research (ICOAR), which fills a much needed gap and advances research capability for researchers in both CISE and SBE communities to apply advanced ML methods for online abuse research.
- Major challenger, technical overview, application layer and multiple sample projects contributed by me.

**Defending Against Cyberbullying in Instagram Images** Buffalo, NY
*Sponsoring Agency: Facebook* *July 2021, Under review*

- Status: Applied
- Award Amount: $50,000.00
- This project focuses on studying the factors of Instagram-specific image-based cyberbullying and developing AI-based techniques for its detection.
- All sections of the grant application were contributed by me.

## Teaching Experience

**Assistant Professor** Aug. 2022 – Present
*The University of Texas at San Antonio* *San Antonio, TX*

- Teacher for the following courses
  - IS-3423 Network Security, Fall 2022
- My responsibilities included
  - Instructor in-charge for lectures and hands-on labs in all formats

**Graduate Teaching Assistant (TA)** Jan. 2017 – Jan. 2020
*Clemson University* *Clemson, SC*

- Teaching assistant for the following courses
  - CPSC-8580 Security in Emerging Systems, Fall 2020
  - CPSC-8580 Security in Emerging Systems, Fall 2019
  - CPSC-8580 Security in Emerging Systems, Fall 2018
  - CPSC-6200 Computer Security Principles, Fall 2018
  - CPSC-4200 Computer Security Principles, Fall 2018
  - CPSC-8580 Security in Emerging Systems, Spring 2018
  - CPSC-6200 Computer Security Principles, Fall 2017
  - CPSC-4200 Computer Security Principles, Fall 2017
  - CPSC-8570 Network Technologies Security, Spring 2017
- My TA responsibilities included
  - Instructor in-charge for cyber security labs
  - Designing and grading hands-on cyber security labs
  - Answer students' questions about lab tasks, troubleshoot student issues, and help students' setup lab environment
  - Grading assignments, quizzes, exams and other submissions
  - Filling-in for the instructor for face-to-face course lectures
- The cyber security labs designed for courses are listed below
  - Content-level photo privacy management (CPSC 8580, CPSC 6200, CPSC 4200)
  - Hacking deep learning models (CPSC 8580)

## Additional Experience

**System Engineer** Jun. 2011 – Aug. 2015
*Infosys Technologies Ltd.* *Bangalore, India*

- Designed a web application for a leading banking company.
- The application is used by bankers for book-keeping of on-going projects.
- Worked with a large team from multiple locations.

## Collaborators

- Computer Science
  - Ziming Zhao, Assistant Professor at SUNY Buffalo (AAAI [19])
  - Feng Luo, Professor at Clemson University (NDSS [3], AAAI [19], ICMLA [4], CODASPY [5])
  - Long Cheng, Assistant Professor at Clemson University (NDSS [3], AAAI [19], ICMLA [4], CODASPY [5])
- HCI and Social Science
  - Kelly Caine, Associate Professor at Clemson University (CHI [7], CSCW [9], SACMAT [11])
  - Matthew Costello, Assistant Professor at Clemson University (ICMLA [3], CODASPY [5])
- Information Systems
  - Daniel Pienta, Assistant Professor at Baylor University (MISQ [15, 16, 17])
  - Jason Thatcher, Professor at Temple University (MISQ [15, 16, 17])
  - Nicholas Berente, Associate Professor at University of Notre Dame (MISQ [15])
  - Allen Johnston, Professor at University of Alabama (MISQ [16])
  - Nirmalee Raddatz, Assistant Professor at The University of Memphis (MISQ Special Issue on Digital Technologies and Social Justice [17])
  - Sriram Somanchi, Assistant Professor at University of Notre Dame (MISQ [15])

## Media and News

- My paper on image-based cyberbullying accepted at NDSS [3] was reported on ***SUNY Buffalo CSE News***. Article Link: https://engineering.buffalo.edu/computer-science-engineering/news-and-events/news.host.html/content/shared/engineering/home/articles/news-articles/2021/cybersecurity-research-showcased-at-network-and-distributed-system-security-symposium.detail.html
- My work on content-level photo privacy protection [11] was reported in ***Clemson IDEAS magazine***. Article Link: https://cecas.clemson.edu/ideas/archives/spring20/. Video Link: https://www.youtube.com/watch?v=n2Hnvc8xcX4
- My paper on online hate speech detection and understanding [6] with two high school students has been reported in ***Clemson Newsstand***. Article Link: https://news.clemson.edu/artificial-intelligence-could-help-stem-the-tide-of-online-hate-speech-clemson-university-researchers-say/
- My work on image-based cyberbullying was shared by the NDSS Symposium on ***Twitter***. Link: https://twitter.com/NDSSSymposium/status/1364666537559298048?s=20

## Invited Talks

- **Great Lakes Security Day 2021.** I am invited to talk on visual cyberbullying threat and defenses at Great Lakes Security Day (GLSD) 2021, on November $12^{th}$2021. GLSD brings together premier practitioners, researchers, students, and funding partners in security, to share latest advances, debate roadmaps and future directions, create new collaborations, and seek new opportunities in cybersecurity, in and around Western and Upstate New York.
- **GenCyber 2021 at North Carolina A&T University.** I was invited to talk on AI-related cybersecurity challenges and cyberbullying defense at the GenCyber 2021 event at North Carolina A&T University, during two sessions held on July $21^{st}$2021, and July $28^{th}$2021. The GenCyber 2021 was attended by **60+ high school students** and teachers in a 2-week summer camp centered on cyber security education.
- **UpBeat at SUNY Buffalo, CSE Department.** I was invited to talk on visual cyberbullying defense at UpBeat, organized by the CSE department at SUNY Buffalo on October $1^{st}$2021. The UpBeat is a weekly event attended by faculty and graduate students to learn about ongoing research projects in the department.

## Technical Program Committee

Conference Committee

- Annual Computer Security Applications Conference (ACSAC) Artifacts Evaluation Committee, 2021
- IEEE International Workshop Big Data Security and Services (BigDataService), 2018–2020

Poster Committee

- Poster Program of ACM Conference on Data and Application Security and Privacy (CODASPY), 2018–2020

## Conference Paper Reviewer

- The Fourteenth International Conference on Advances in Computer-Human Interactions (ACHI), 2021
- International Conference On Design Science Research In Information Systems And Technology (DESRIST), 2021
- International Workshop on Automotive and Autonomous Vehicle Security (AutoSec), 2021
- The Web Conference (WWW), 2021
- ACM Conference on Computer and Communications Security (CCS), 2018–2021
- ACM Symposium on Information, Computer and Communications Security (AsiaCCS), 2018–2021
- Annual Computer Security Applications Conference (ACSAC), 2018–2021
- ACM Conference on Data and Application Security and Privacy (CODASPY), 2018–2020
- ACM Symposium on Access Control Models and Technologies (SACMAT), 2018–2020
- IEEE Conference on Communications and Network Security (CNS), 2018–2021
- IEEE International Conference on Computer Communications and Networks (ICCCN), 2018–2020
- ACM SIGCOMM Workshop on Security in Softwarized Networks: Prospects and Challenges (SecSoN), 2018
- ACM International Workshop on Security in Software Defined Networks & Network Function Virtualization (SDN-NFV Security), 2017–2020
- International Conference on Information and Communications Security (ICICS), 2018–2020
- International Conference on Privacy, Security and Trust (PST), 2018–2020
- IEEE International Conference on Cloud Networking (CloudNet), 2018–2020
- IEEE International Conference on Smart City Innovations (SCI), 2018–2021-0

## Journal Paper Reviewer

- Transactions on Dependable and Secure Computing (TDSC), 2017–2020
- Transactions on Information Forensics  Security (TIFS), 2018–2020

## Honors And Awards

- Talford family fellowship for cyber security research, 2018

## Technical Skills

**Languages**: Python, C/C++, Latex, JavaScript, HTML/CSS, R
**Frameworks**: PyTorch, Tensorflow
**Developer Tools**: Git, Vim, Eclipse
**Libraries**: Pandas, NumPy, Matplotlib
**MOOC**: Deep Learning by deeplearning.ai on Coursera (January 15, 2019)