

# Wenxin Jiang

*Ph.D. Student*  
*Elmore Family School of Electrical and Computer Engineering*  
*Purdue University*  
*West Lafayette, IN 47906*

*jiang784@purdue.edu*  
*765-409-1715*  
*<https://wenxin-jiang.github.io>*

## RESEARCH THEME

---

My research interests lie in *Software Engineering for Machine Learning*, especially the reusability, reproducibility, and trustworthiness of deep learning pre-trained models (PTMs).

## EDUCATION

---

<b>Ph.D, Electrical and Computer Engineering</b> , <i>GPA: 3.7/4.0</i> <i>Purdue University, West Lafayette, IN</i>	2020–present
<b>B.Sc. Applied Physics</b> , <i>GPA: 3.3/4.0</i> <i>Southeast University, Jiangsu, China</i>	2016–2020
<b>Study Abroad Program, Engineering Physics</b> , <i>GPA: 3.8/4.0</i> <i>University of California, Santa Barbara, CA</i>	2019

## PROFESSIONAL EXPERIENCE

---

<b>Graduate Research Assistant</b> <i>Purdue University — Advised by James C. Davis</i>	2021–present
<b>TensorFlow Model Developer</b> <i>Purdue University × Google</i>	2021–present

## REFEREED CONFERENCE PUBLICATIONS

---

- [1] **Jiang**, Synovic, Hyatt, Schorlemmer, Sethi, Lu, Thiruvathukal, and Davis. *An Empirical Study of Pre-Trained Model Reuse in the Hugging Face Deep Learning Model Registry*. Proceedings of the ACM/IEEE 45th International Conference on Software Engineering (**ICSE’23**).
- [2] **Jiang**, Synovic, Jajal, Schorlemmer, Tewari, Pareek, Thiruvathukal, and Davis. *PTMTorrent: A Dataset for Mining Open-source Pre-trained Model Packages*. Proceedings of the 20th Annual Conference on Mining Software Repositories — Data and Tool Showcase Track (**MSR-Data’23**).
- [3] Davis, Jajal, **Jiang**, Schorlemmer, N. Synovic, and G.K. Thiruvathukal. *Reusing Deep Learning Models Challenges and Directions in Software Engineering*. Proceedings of the IEEE John Vincent Atanasoff Symposium on Modern Computing (**JVA’23**).
- [4] Montes, Peerapatanapokin, Schultz, Guo, **Jiang**, and Davis. *Discrepancies among Pre-trained Deep Neural Networks: A New Threat to Model Zoo Reliability*. Proceedings of the 30th ACM Joint European Software Engineering Conference and Symposium on the Foundations of Software Engineering — Ideas, Visions, and Reflections track (**ESEC/FSE-IVR’22**).

## REFEREED WORKSHOPS, DEMONSTRATIONS, AND COMPETITIONS

---

- [1] **Jiang**, Synovic, Sethi, Indarapu, Hyatt, Schorlemmer, Thiruvathukal, and Davis. *An Empirical Study of Artifacts and Security Risks in the Pre-trained Model Supply Chain*. Proceedings of the 1st ACM Workshop on Software Supply Chain Offensive Research and Ecosystem Defenses (**SCORED’22**).

- [2] Synovic, Hyatt, Sethi, Thota, Shilpika, Miller, **Jiang**, Amobi, Pinderski, Laufer, Hayward, Kingensmith, Davis, and Thiruvathukal. *Snapshot Metrics Are Not Enough: Analyzing Software Repositories with Longitudinal Metrics*. Proceedings of the 37th IEEE/ACM International Conference on Automated Software Engineering — Demonstrations track (**ASE-Tool Demonstrations'22**).
- [3] Veselsky, West, Ahlgren, Thiruvathukal, Klingensmith, Goel, **Jiang**, Davis, Lee, and Kim. *Establishing trust in vehicle-to-vehicle coordination: a sensor fusion approach*. Proceedings of the 23rd Annual International Workshop on Mobile Computing Systems and Application (**HotMobile'22**).

## TECHNICAL REPORTS

---

- [1] **Jiang**, Cheung, Thiruvathukal, and Davis. *Exploring Naming Conventions (and Defects) of Pre-trained Deep Learning Models in Hugging Face and Other Model Hubs*. <https://arxiv.org/pdf/2310.01642>. 2023.
- [2] Jajal, **Jiang**, Tewari, Woo, Lu, Thiruvathukal, and Davis. *Analysis of Failures and Risks in Deep Learning Model Converters: A Case Study in the ONNX Ecosystem*. <https://arxiv.org/abs/2303.17708>. 2023.
- [3] **Jiang**, Banna, Vivek, Goel, Synovic, Klingensmith, Thiruvathukal, and Davis. *Challenges and Practices of Deep Learning Model Reengineering: A Case Study on Computer Vision*. <https://arxiv.org/abs/2303.07476>. 2023.
- [4] Banna, Chinnakotla, Yan, Vegesana, Vivek, Krishnappa, **Jiang**, Lu, Thiruvathukal, and Davis. *An Experience Report on Machine Learning Reproducibility: Guidance for Practitioners and TensorFlow Model Garden Contributors*. <https://arxiv.org/abs/2107.00821>. 2021.

## POSTERS

---

- [1] Schorlemmer, **Jiang**, and Davis. *Machine Learning Supply Chain Security*. 2023 Purdue CERIAS Symposium (**CERIAS'23**). *Award: Best Poster — 2nd-place*.
- [2] **Jiang**, Schorlemmer, and Davis. *Trustworthy Re-use of Pre-trained Neural Networks*. 2023 Purdue CERIAS Symposium (**CERIAS'23**).

## TEACHING ASSISTANT

---

ECE 595 – Advanced Software Engineering Purdue University	Spring 2022
--	-------------

## INVITED TALKS

---

An Empirical Study of Pre-Trained Model Reuse in the Hugging Face Deep Learning Model Registry Purdue University Programming Languages Group, Seminar	2023
Deep Learning Model Reengineering: An Exploratory Case Study on Computer Vision Purdue University Programming Languages Group, Seminar	2022

## AWARDS AND RECOGNITION

---

ACM SIGSOFT CAPS Travel Grant (ICSE'23)	2023
Purdue Graduate Student Government and the Graduate School Travel Grant (ICSE'23)	2023
ACM SIGSOFT CAPS Travel Grant (ESEC/FSE'22)	2022
Study Abroad Fellowship, Southeast University	2019

Second prize, Vision Guided Robot Competition, Southeast University	2019
Distinction Award, Southeast University	2018
Third prize, Structural Innovation Invitation Competition, Southeast University	2017

## ACTIVITIES AS A REFEREE

---

Sub-Reviewer: LCTES'23, ESEC/FSE'23, ASE'22	2023
---	------

## DEPARTMENTAL SERVICE

---

Director, Southeast University TV Station	2016-2019
Member, Southeast University Physics Student Council	2016-2017

## PROFESSIONAL MEMBERSHIPS

---

Student member, Association for Computing Machinery (ACM)
Student member, Institute of Electrical and Electronics Engineers (IEEE)