James C. Davis

Assistant Professor
Elmore Family School of Electrical and Computer Engineering
Purdue University
West Lafayette, IN 47906

davisjam@purdue.edu
765–494–3133
https://davisjam.github.io

RESEARCH THEME

My research enables safe and secure software engineering for cyber- and cyber-physical systems. My work is grounded in empirical measurements of the software engineering process, product, and usage context.

EDUCATION

Ph.D, Computer Science and Applications 2015–2020
Virginia Tech, Blacksburg, VA

B.Sc. Computer Science, B.Sc. Mathematics 2008–2012
Clarkson University, Potsdam, NY

PROFESSIONAL EXPERIENCE

Assistant Professor Fall 2020-present
Purdue University — Electrical and Computer Engineering

Intern, Microsoft Research (RiSE group: Cloud Security) Summer 2019
Microsoft Research, Redmond, WA — Mentored by Patrice Godefroid

Intern, IBM Research (Storage) Summer 2018
IBM Research, Almaden, CA — Mentored by Deepavali Bhagwat

Graduate Research Assistant 2016–2020
Virginia Tech — Advised by Dongyoon Lee

Software Engineer, IBM (GPFS) 2012–2017
IBM, Poughkeepsie, NY

EXTERNAL RESEARCH GRANTS

[1] **Unrestricted gift to support research on machine learning reproducibility**
PI
Google, LLC
2022. $80,000.

[2] **NSF #2229703: POSE: Phase I: Scoping An Open-Source Ecosystem Around Proactive Software Supply Chain Monitoring**
Co-PI (PI: Santiago Torres-Arias)
US National Science Foundation
2022–2023. $300,000.

[3] **Cisco: Trustworthy Re-use of Pre-Trained Neural Networks**
PI (Co-PI: Yung-Hsiang Lu)
Contract with Cisco
2022–2023. $179,237.

[4] **Cisco: Monitor and manage security risks in software supply chains with Sigstore**

Co-PI (PI: Santiago Torres-Arias)

Contract with Cisco

2022–2023. $184,536.

[5] **NSF #2135156: Collaborative Research: SaTC: CORE: Small: Improving Sanitization and Avoiding Denial of Service Through Correct and Safe Regexes**

PI (Co-PI: Dongyoon Lee)

US National Science Foundation

2022–2025. Purdue's share: $274,000.

[6] **Rolls Royce: Dynamic Analysis of Embedded Firmware**

Co-PI (PI: Aravind Machiry)

Contract with Rolls Royce

2021-2022. $175,000.

[7] **NSF #2107230: Collaborative Research: OAC Core: Advancing Low-Power Computer Vision at the Edge**

Co-PI (PI: Yung-Hsiang Lu)

US National Science Foundation

2021–2024. Purdue's share: $258,000.

[8] **Unrestricted gift to support research on machine learning reproducibility**

PI (Co-PI: Yung-Hsiang Lu)

Google, LLC

2020. $80,000 + $20,000.

INTERNAL RESEARCH GRANTS

[1] **Revamping the CompE Curriculum for Secure Software Engineering**

PI (Co-PIs: Machiry, Torres-Arias, Bagchi)

ECE Agile Reform of Curriculum program, enabled by Elmore Family gift

2021-2022. $150,000.

[2] **Intercultural Engineering Education for Software Engineers**

PI (Co-PI: Kirsten Davis)

Purdue University VEIL Program

2020. $5,000.

REFEREED CONFERENCE PUBLICATIONS[1]

[1] Hassan, Aamir, Lee, **Davis**, and Servant. *Improving Developers' Understanding of Regex Denial of Service Tools through Anti-Patterns and Fix Strategies.* Proceedings of the 44th IEEE Symposium on Security and Privacy (**S&P'23**).

[1] My research mentees are underlined. My name is in **bold**.

[2] Amusuo, Sharma, Rao, Vincent, and **Davis**. *Reflections on Software Failure Analysis*. 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**).

[3] 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**).

[4] Anandayuvaraj and **Davis**. *Reflecting on Recurring Failures in IoT Development*. Proceedings of the 37th IEEE/ACM International Conference on Automated Software Engineering — New Ideas and Emerging Results track (**ASE-NIER'22**).

[5] Goel, Tung, Eliopoulos, Hu, Thiruvathukal, **Davis**, and Lu. *Directed Acyclic Graph-based Neural Networks for Tunable Low-Power Computer Vision*. Proceedings of the ACM/IEEE International Symposium on Low Power Electronics and Design (**ISLPED'22**).

[6] Barlas, Du, and **Davis**. *Exploiting Input Sanitization for Regex Denial of Service*. Proceedings of the ACM/IEEE 44th International Conference on Software Engineering (**ICSE'22**).

[7] Xu, **Davis**, Hu, and Jindal. *An Empirical Study on the Impact of Parameters on Mobile App Energy Usage*. Proceedings of the 29th IEEE International Conference on Software Analysis, Evolution and Reengineering (**SANER'22**).

[8] Goel, Tung, Hu, Thiruvathukal, **Davis**, and Lu. *Efficient Computer Vision on Edge Devices with Pipeline-Parallel Hierarchical Neural Networks*. Proceedings of the 27th Asia and South Pacific Design Automation Conference (**ASP-DAC'22**).

[9] Goel, Tung, Hu, Wang, **Davis**, Thiruvathukal, and Lu. *Low-Power Multi-Camera Object Re-Identification using Hierarchical Neural Networks*. Proceedings of the ACM/IEEE International Symposium on Low Power Electronics and Design (**ISLPED'21**).

[10] **Davis**, Servant, and Lee. *Using Selective Memoization to Defeat Regular Expression Denial of Service (ReDoS)*. Proceedings of the 42nd IEEE Symposium on Security and Privacy (**IEEE S&P'21**).

[11] Cha, Wittern, Baudart, **Davis**, Mandel, and Laredo. *A Principled Approach to GraphQL Query Cost Analysis*. Proceedings of the 28th ACM Joint European Software Engineering Conference and Symposium on the Foundations of Software Engineering (**ESEC/FSE'20**). ***ACM Distinguished Paper Award***.

[12] Rupprecht, **Davis**, Arnold, Gur, and Bhagwat. *Improving Reproducibility of Data Science Pipelines through Transparent Provenance Capture*. Proceedings of the 46th International Conference on Very Large Data Bases (**VLDB'20 Industry track**).

[13] **Davis**, Moyer, Kazerouni, and Lee. *Testing Regex Generalizability And Its Implications: A Large-Scale Many-Language Measurement Study*. Proceedings of the 34th IEEE/ACM International Conference on Automated Software Engineering (**ASE'19**).

[14] Michael, Donohue, **Davis**, Lee, and Servant. *Regexes are Hard: Decision-making, Difficulties, and Risks in Programming Regular Expressions*. Proceedings of the 34th IEEE/ACM International Conference on Automated Software Engineering (**ASE'19**). ***ACM Distinguished Paper Award***.

[15] **Davis**, Michael, Coghlan, Servant, and Lee. *Are Regular Expressions a Lingua Franca? An Empirical Study on the Re-use and Portability of Regular Expressions*. Proceedings of the 27th ACM Joint European Software Engineering Conference and Symposium on the Foundations of Software Engineering (**ESEC/FSE'19**).

[16] Wittern, Cha, **Davis**, Baudart, Mandel. *An Empirical Study of GraphQL Schemas*. Proceedings of the 17th International Conference on Service-Oriented Computing (**ICSOC'19**).

[17] Fu, Ghaffar, **Davis**, and Lee. *EdgeWise: A Better Stream Processing Engine for the Edge*. 2019 USENIX Annual Technical Conference (**USENIX ATC'19**).

[18] **Davis**, Coghlan, Servant, and Lee. *The Impact of Regular Expression Denial of Service (REDOS) in Practice: an Empirical Study at the Ecosystem Scale*. Proceedings of the 26th ACM Joint European Software Engineering Conference and Symposium on the Foundations of Software Engineering (**ESEC/FSE'18**). ***ACM Distinguished Paper Award***.

[19] **Davis**, Williamson, and Lee. *A Sense of Time for JavaScript and Node.js: First-Class Timeouts as a Cure for Event Handler Poisoning*. Proceedings of the 27th USENIX Security Symposium (**USENIX Security'18**).

[20] **Davis**, Thekumparampil, and Lee. *Node.fz: Fuzzing the Server-Side Event-Driven Architecture*. Proceedings of the Twelfth European Conference on Computer Systems (**EuroSys'17**).

REFEREED JOURNAL ARTICLES

[1] Goel, Tung, Eliopoulos, Wang, **Davis**, Thiruvathukal, Lu. *Tree-based Unidirectional Neural Networks for Low-Power Computer Vision*. IEEE Design & Test (**IEEE D&T'22**).

[2] Davis, Deters, Ozkan, **Davis**, and Murzi. *Applying Experiential Learning Theory to Understand Study Abroad Leaders' Experiences Using Real-Time Perspectives*. Frontiers: The Interdisciplinary Journal of Study Abroad, Vol. 34, No. 2 (**Frontiers'22**).

[3] Herbold, Trautsch, Ledel, Aghamohammadi, Ghaleb, Chahal, Bossenmaier, Nagaria, Makedonski, Ahmadabadi, Szabados, Spieker, Madeja, Hoy, Lenarduzzi, Wang, Rodriguez-Perez, Colomo-Palacios, Verdecchia, Singh, Qin, Chakroborti, Davis, Walunj, Wu, Marcilio, Alam, Aldaeej, Amit, Turhan, Eismann, Wickert, Malavolta, Sulír, Fard, Henley, Kourtzanidis, Tüzün, Treude, Shamasbi, Pashchenko, Wyrich, **Davis**, Serebrenik, Albrecht, Aktas, Strüber, and Erbel. *A Fine-grained Data Set and Analysis of Tangling in Bug Fixing Commits*. Empirical Software Engineering (**EMSE'21**).

[4] Kazerouni, **Davis**, Basak, Shaffer, Servant, and Edwards. *Fast and Accurate Incremental Feedback for Students' Software Tests Using Selective Mutation Analysis*. Journal of Systems and Software (**JSS'21**).

[5] Ozkan, Davis, **Davis**, James, Murzi, and Knight. *Expectations and Experiences of Short-Term Study Abroad Leadership Teams*. Journal of International Engineering Education (**JIEE'20**).

REFEREED WORKSHOPS, DEMONSTRATIONS, AND COMPETITIONS

[1] Jiang, Synovic, Sethi, Indarapu, Hyatt, Schorlemmer, Thiruvathukal, and **Davis**. *An Empirical Study of Artifacts and Security Practices 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] Okafor, Schorlemmer, Torres-Arias, and **Davis**. *SoK: Analysis of Software Supply Chain Security by Establishing Secure Design Properties*. Proceedings of the 1st ACM Workshop on Software Supply Chain Offensive Research and Ecosystem Defenses (**SCORED'22**).

[3] Synovic, Hyatt, Sethi, Thota, Shilpika, Miller, Jiang, Amobi, Pinderski, Läufer, Hayward, Klingensmith, and **Davis**, 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**).

[4] Gopalakrishna, Anandayuvaraj, Detti, Bland, Rahaman, and **Davis**. *"If security is required": Engineering and Security Practices for Machine Learning-based IoT Devices*. Proceedings of the 4th International Workshop on Software Engineering Research & Practices for the Internet of Things (**ICSE-SERP4IoT'22**).

[5] **Davis**, Amusuo, and Bushagour. *Experience Paper: A First Offering of Software Engineering*. Proceedings of the 1st International Workshop on Designing and Running Project-Based Courses in Software Engineering Education (**ICSE-DREE'22**).

[6] Veselsky, West, Ahlgren, Goel, Jiang, Lee, Kim, **Davis**, Thiruvathukal, and Klingensmith. *Establishing Trust in Vehicle-to-Vehicle Coordination: A Sensor Fusion Approach*. Proceedings of the 2nd Workshop on Data-Driven and Intelligent Cyber-Physical Systems for Smart Cities (DI-CPS) (**DI-CPS'22**).

[7] Winkler, Agarwal, Tung, Ugalde, Jung, and **Davis**. *A Replication of "DeepBugs: A Learning Approach to Name-based Bug Detection"*. Proceedings of the 29th ACM Joint European Software Engineering Conference and Symposium on the Foundations of Software Engineering (**ESEC/FSE'21 Artifact**).

[8] **Davis**. *On the Impact and Defeat of Regex DoS*. ACM Student Research Competition, 2019-2020 Grand Finals. ***Second place, graduate student division***.

[9] **Davis**. *Rethinking Regex Engines to Address ReDoS*. ACM Student Research Competition, 2019-2020 at ESEC/FSE'19. ***First place, graduate student division***.

[10] Rupprecht, **Davis**, Arnold, Lubbock, Tyson, and Bhagwat. *Ursprung: Provenance for Large-Scale Analytics Environments*. Proceedings of the 2019 International Conference on Management of Data (**SIGMOD'19 Demo**).

[11] **Davis**, Kildow, and Lee. *The Case of the Poisoned Event Handler: Weaknesses in the Node.js Event-Driven Architecture*. Proceedings of the 10th European Workshop on Systems Security (**EuroSec'17**).

POSTERS

[1] Hornbrook and **Davis**. *An Intercultural Engineering Module for Software Engineers*. 2021 Annual Colloquium for International Engineering Education (**ACIEE'21**).

[2] Vivek, Chinnakotla, Banna, Vegesana, Yan, **Davis**, Lu, Thiruvathukal. *Exemplars for Machine Learning: Towards Software Engineering & Reproducibility*. SIAM Conference on Computational Science and Engineering (**CSE'21**).

TECHNICAL REPORTS

[1] 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.

PATENTS

[1] Davis, **Davis**. *Verification of the Integrity of Data Files Stored in Copy-on-Write (CoW) Based File System Snapshots*. IBM, U.S. patent 11,176,090 B2, granted Nov. 16, 2021.

[2] **Davis**, Davis. *Injection of Simulated Hardware Failure(s) in a File System for Establishing File System Tolerance-to-Storage-Failure(s)*. IBM, U.S. patent 11,023,341 B2, granted Jun. 1, 2021.

[3] **Davis**, Rupprecht, Bhagwat, Arnold, Sawdon. *Performing Hierarchical Provenance Collection*. IBM, U.S. patent 10,891,174 B1, granted Jan. 12, 2021.

[4] **Davis**, Davis. *File Metadata Verification in a Distributed File System*. IBM, U.S. patent 10,678,755 B2, granted Jun. 9, 2020..

[5] Davis, **Davis**. *Testing of Lock Managers in Computing Environments*. IBM, U.S patent 10,061,777 B1, granted Aug. 28, 2018.

[6] **Davis**, Davis, Knop. *Detection of File Corruption in a Distributed File System*. IBM, U.S. patent 10,025,788, granted Jul. 17, 2018.

COURSES DESIGNED

ECE 461 – Software Engineering Launched Fall 2021
Purdue University

ECE 595 – Advanced Software Engineering Launched Spring 2021
Purdue University

COURSES TAUGHT

ECE 461 – Software Engineering Fall 2021
Purdue University

ECE 595 – Advanced Software Engineering Spring 2021, Spring 2022
Purdue University

ECE 368 – Data Structures Fall 2020
Purdue University

Vertically Integrated Project: Open-Source TensorFlow Software Fall 2020–present
Purdue University

Vertically Integrated Project: SafeRegex Fall 2020, Spring 2021
Purdue University

CS 3114 – Data Structures and Algorithms Fall 2019
Virginia Tech

CS 1064 – Introduction to Programming in Python Spring 2019
Virginia Tech

Rising Sophomore Abroad Program (Track Leader) Spring 2018, Spring 2019
Virginia Tech

PHD AND MASTER'S STUDENTS

Wenxin Jiang	PhD	Spring 2021–present
Paschal Amusuo	PhD	Fall 2021–present
Dharun Anandayuvaraj	PhD	Fall 2021–present
Kelechi Gabriel Kalu	PhD	Spring 2023–present
William Maxam	MSc	Fall 2021–present
Geoffrey Cramer	MSc	Fall 2021–present
Garvit Jairath	MSc	Fall 2022–present
Taylor Schorlemmer	MSc	Fall 2022–present

INVITED TALKS

Software Reuse 2.0: Practices and Challenges for Pre-Trained Neural Networks 2022
Loyola University Chicago

Challenges in Global Software Development 2021
University of Wisconsin–Stout

Regexes Awry: Characterizing and Defeating Regex-based Denial of Service 2020
Clemson University CS department colloquium

Regex-based Denial of Service 2020
Clarkson University CS department colloquium

Regexes are Hard: Qualitative and Quantitative Perspectives 2019
NC State CS department colloquium

The Dangers of Copy/Pasting Code 2019
Episode of the Podcast "The Secure Developer": `https://tinyurl.com/DavisResearchPodcast`

Regexes in the Wild 2019
Virginia Tech department seminar

Academic Perspectives on Node.js 2018
Node.js Collaborator Summit, Vancouver

International Engineering Annual, 2015–2019
Rising Sophomore Abroad Program, Virginia Tech

AWARDS AND RECOGNITION

FOR RESEARCH

ASE 2021 Distinguished PC Member Award 2021

ACM Distinguished Paper Award, ESEC/FSE 2020 2020

Second place, Grand Finals of the ACM Graduate Student Research Competition 2020

First place, Graduate Student Research Competition, ESEC/FSE 2019 2019

ACM Distinguished Paper Award, ASE 2019 2019

Microsoft Security Researcher Acknowledgments (Regex DoS) 2018

Pratt Fellowship, Virginia Tech College of Engineering 2017–2019

Davenport Fellowship, Virginia Tech College of Engineering 2019

Graduate Fellow, VT Academy for Global Engineering 2019–2020

IBM Significant Contributor Award (Node.js) 2018

FOR TEACHING

2022 Ruth and Joel Spira Outstanding Teacher Award 2022

Fall 2021: Teaching–Recognized for high student evaluation scores (∼100 faculty in College of Eng.) 2021

VIP Outstanding Team Mentor Award, Purdue TensorFlow Team 2021

FOR SERVICE

ACM Distinguished Paper Award, ESEC/FSE 2018 2018

Outstanding Graduate Student Service Award, CS@VT 2020

ACTIVITIES AS A REFEREE

Member, ESEC/FSE Program Committee	2023
Member, ASE Doctoral Symposium Committee	2022
PC Member, ACM Workshop on Software Supply Chain Offens. Research and Ecosystem Defenses	2022
Reviewer, ACM Transactions on Software Engineering (TSE)	2020–present
Reviewer, Springer Empirical Software Engineering (EMSE)	2020–present
Judge, CSAW'21 Best Paper Competition	2021
Member, ASE Program Committee	2021
Member, ICSE Demonstrations Committee	2021
Member, ESEC/FSE Artifact Evaluation Committee	2021
Member, ESEC/FSE Artifact Evaluation Committee	2020
Member, CGO Artifact Evaluation Committee	2019
Sub-reviewer: Middleware'17, ASPLOS'18, EuroSys'18, MASCOTS'18, HPCA'19, CGO'19	2016–2019

DEPARTMENTAL SERVICE

Member, Purdue ECE Faculty Search Committee — Software Engineering	2022-2023
Host, Computer Engineering Seminae Series — Dr. Joanna C. S. Santos (Notre Dame)	2022
Host, Purdue Engineering Distinguished Lecture Series (PEDLS) — Dr. Nancy Leveson (MIT)	2022
Member, Purdue ECE Undergraduate Curriculum Committee	2020-2022
Panelist, CS@Virginia Tech Academic Jobs Panel	2021
President, Virginia Tech CS Graduate Student Council	2018–2019
Organizer, Virginia Tech Systems Reading Group	2017–2020

SHORT COURSES AND WORKSHOPS ATTENDED

Tools to Foster Students' (Cross-)cultural Sensitivity in Engineering Ethical Decision-Making (ASEE'22, Clancy & Qiu)	2022
Effective College Teaching (Brent & Felder)	2020
Intercultural Pedagogy Grant Training Program, Purdue CILMAR	2020

PROFESSIONAL MEMBERSHIPS

Member, Association for Computing Machinery (ACM)

Member, Institute of Electrical and Electronics Engineers (IEEE)

Member, American Society for Engineering Education (ASEE)