

# James C. Davis

*School of Electrical and Computer Engineering  
Purdue University  
West Lafayette, IN 47906*

*davisjam@purdue.edu  
765-494-3133  
<https://davisjam.github.io>*

## RESEARCH INTERESTS

---

My research improves the quality of computing systems, grounded in empirical measurements of how and why they are used. My projects have flavors of software engineering, cybersecurity, and systems.

## PROFESSIONAL EXPERIENCE

---

<b>Assistant Professor</b> <i>Purdue University — Electrical and Computer Engineering</i>	Fall 2020-present
<b>Intern, Microsoft Research (RiSE group: Cloud Security)</b> <i>Microsoft Research, Redmond, WA — Mentored by Patrice Godefroid</i>	Summer 2019
<b>Intern, IBM Research (Storage)</b> <i>IBM Research, Almaden, CA — Mentored by Deepavali Bhagwat</i>	Summer 2018
<b>Graduate Research Assistant</b> <i>Virginia Tech — Advised by Dongyoon Lee</i>	2016-2020
<b>Software Engineer, IBM (GPFS)</b> <i>IBM, Poughkeepsie, NY</i>	2012-2015, Summer 2016, Summer 2017

## EDUCATION

---

<b>Ph.D, Computer Science and Applications</b> <i>Virginia Tech, Blacksburg, VA</i>	2015-2020
<b>B.Sc. Computer Science, B.Sc. Mathematics</b> <i>Clarkson University, Potsdam, NY</i>	2008-2012

## RESEARCH GRANTS

---

- [1] **NSF #2107230: OAC Core: Advancing Low-Power Computer Vision at the Edge**  
Co-Principal Investigator (PI: Yung-Hsiang Lu)  
*US National Science Foundation*  
2021-2024. Purdue's share: \$250,000.
- [2] **Unrestricted gift to support research on machine learning reproducibility**  
Principal Investigator (Co-PI: Yung-Hsiang Lu)  
*Google, LLC*  
2020. \$80,000.
- [3] **Unrestricted gift to support research on machine learning reproducibility**  
Principal Investigator (Co-PI: Yung-Hsiang Lu)  
*Google, LLC*  
2020. \$20,000.

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

Principal Investigator (Co-PI: Kirsten Davis)

*Purdue University VEIL Program*

*2020. \$5,000.*

**REFEREED CONFERENCE PUBLICATIONS**

---

- [1] Goel, Tung, Hu, Wang, **Davis**, Thiruvathukal, 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**).
- [2] **Davis**, Servant, 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**).
- [3] Cha, Wittern, Baudart, **Davis**, Mandel, 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*.
- [4] Rupperecht, **Davis**, Arnold, Gur, 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**).
- [5] **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**).
- [6] 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*.
- [7] **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**).
- [8] Wittern, Cha, **Davis**, Baudart, Mandel. *An Empirical Study of GraphQL Schemas*. Proceedings of the 17th International Conference on Service-Oriented Computing (**ICSOC'19**).
- [9] Fu, Ghaffar, **Davis**, and Lee. *EdgeWise: A Better Stream Processing Engine for the Edge*. 2019 USENIX Annual Technical Conference (**USENIX ATC'19**).
- [10] **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*.
- [11] **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**).
- [12] **Davis**, Thekumparampil, and Lee. *Node.fz: Fuzzing the Server-Side Event-Driven Architecture*. Proceedings of the Twelfth European Conference on Computer Systems (**EuroSys'17**).

## JOURNAL ARTICLES

---

- [1] Kazerouni, **Davis**, Basak, Shaffer, Servant, Edwards. *Fast and Accurate Incremental Feedback for Students' Software Tests Using Selective Mutation Analysis*. Journal of Systems and Software (**JSS'21**).
- [2] Ozkan, Davis, **Davis**, James, Murzi, Knight. *Expectations and Experiences of Short-Term Study Abroad Leadership Teams*. Journal of International Engineering Education (**JIEE'20**).

## REFEREED SHORT PAPERS

---

- [1] Winkler, Agarwal, Tung, Ugalde, Jung, **Davis**. *A Partial 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**).
- [2] **Davis**. *On the Impact and Defeat of Regex DoS*. **ACM SRC – Grand Finals. Second place, graduate student division.**
- [3] **Davis**. *Rethinking Regex Engines to Address ReDoS*. **ACM SRC – ESEC/FSE'19. First place, graduate student division.**
- [4] 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**).
- [5] **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'20**).

## 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**, Rupprecht, Bhagwat, Arnold, Sawdon. *Performing Hierarchical Provenance Collection*. IBM, U.S. patent US10,891,174B1, granted Jan. 12, 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 application 20200264961. U.S. patent pending.
- [3] Davis, **Davis**. *Verification of the integrity of data files stored in copy-on-write (CoW) based file system snapshots*. IBM, U.S. patent application 20200242075. U.S. patent pending.
- [4] **Davis**, Davis. *File metadata verification in a distributed file system*. IBM, U.S. patent 10,678,755B2, 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

---

<b>ECE 595 — Advanced Software Engineering</b> <i>Purdue University</i>	Spring 2021
<b>ECE 30861 — Software Engineering</b> <i>Purdue University</i>	Fall 2021

## COURSES TAUGHT

---

<b>ECE 30861 — Software Engineering</b> <i>Purdue University</i>	Fall 2021
<b>ECE 595 — Advanced Software Engineering</b> <i>Purdue University</i>	Spring 2021
<b>ECE 368 — Data Structures</b> <i>Purdue University</i>	Fall 2020
<b>VIP: Open-Source TensorFlow Software</b> <i>Purdue University</i>	Fall 2020, Spring 2021
<b>VIP: SafeRegex</b> <i>Purdue University</i>	Fall 2020, Spring 2021
<b>CS 3114 — Data Structures and Algorithms</b> <i>Virginia Tech</i>	Fall 2019
<b>CS 1064 — Introduction to Programming in Python</b> <i>Virginia Tech</i>	Spring 2019
<b>Rising Sophomore Abroad Program (Track Leader)</b> <i>Virginia Tech</i>	Spring 2018, Spring 2019

## MASTER'S AND PHD THESIS STUDENTS CURRENTLY SUPERVISED

---

Wenxin Jiang	PhD	Spring 2021-present
--------------	-----	---------------------

## INVITED TALKS

---

<b>Regexes Awry: Characterizing and Defeating Regex-based Denial of Service</b> <i>Clemson University CS department colloquium</i>	2020
<b>Regex-based Denial of Service</b> <i>Clarkson University CS department colloquium</i>	2020
<b>Regexes are Hard: Qualitative and Quantitative Perspectives</b> <i>NC State CS department colloquium</i>	2019
<b>The Dangers of Copy/Pasting Code</b>	2019

Episode of the Podcast “The Secure Developer”: <https://tinyurl.com/DavisResearchPodcast>

<b>Regexes in the Wild</b> <i>Virginia Tech department seminar</i>	2019
<b>Academic Perspectives on Node.js</b> <i>Node.js Collaborator Summit, Vancouver</i>	2018
<b>International Engineering</b> <i>Rising Sophomore Abroad Program, Virginia Tech</i>	Annual, 2015-2019

## AWARDS AND RECOGNITION

---

VIP Outstanding Team Mentor Award, Purdue TensorFlow Team	2021
ACM Distinguished Paper Award, ESEC/FSE 2020	2020
Outstanding Graduate Student Service Award, CS@VT	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
ACM Distinguished Paper Award, ESEC/FSE 2018	2018
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
IBM Poughkeepsie’s New hire of the month	2014
Frederica Clarkson Award	2012
Clarkson University’s Outstanding Senior (x2): Mathematics, Computer science	2012
Clarkson University Phalanx Commendable Leadership	2011

## ACTIVITIES AS A REFEREE

---

Member, ASE Program Committee	2021
Member, ICSE Demonstrations Track	ICSE 2021
Member, ESEC/FSE Artifact Evaluation Committee	2020, 2021
Reviewer, ACM Transactions on Software Engineering (TSE)	2020-present
Reviewer, Springer Empirical Software Engineering (EMSE)	2020-present
Member, CGO Artifact Evaluation Committee	CGO 2019

## DEPARTMENTAL SERVICE

---

Committee member, Purdue ECE Undergraduate Curriculum Committee	2020-present
President, Virginia Tech CS Graduate Student Council	2018-2019
Organizer, Virginia Tech Systems Reading Group	2017-2020

## SHORT COURSES AND WORKSHOPS ATTENDED

---

Effective College Teaching (Brent & Felder)	2020
Intercultural Pedagogy Grant Training Program, Purdue CILMAR	2020

## PROFESSIONAL MEMBERSHIPS

---

Member, Association for Computing Machinery
---