Donald Pinckney

donald_pinckney@icloud.com · https://donaldpinckney.com · Google Scholar · github.com/donald-pinckney

Skills

Languages: Rust, Python, JavaScript/TypeScript, SQL, C/C++, Swift, Haskell, Go, OCaml.
Tools: Docker, AWS (Lambda, S3, ECS), PostgreSQL, Redis, LangChain, VS Code Extension API, LSP, MCP, CI/CD (GitHub Actions)
Specialties: building agentic systems, coding agents, enterprise AI integrations, developer tooling / program analysis, dependency / package management, serverless computing, compilers, formal verification, distributed systems

Experience

Gitar (venture-backed startup, gitar.ai), *Founding AI Coding Agent Engineer* Apr 2024 – Present

- Launched Gitar, the first coding agent leveraging deep static analysis, driving a 700% spike in customer demand.
- Built agent tools which allow Gitar to perform static analyses and program transformations (symbol lookup, constant folding, dead code analysis, etc.), allowing for precise, safe and efficient code changes.
- Designed a synthetic data generation framework for evaluating Gitar's agentic static analyses and program simplifications.
- Co-designed cloud infrastructure to support running Gitar in the cloud (similar to Cursor's Background Agents), prioritizing safety via customer data isolation and a fine-grained permission system to validate external API calls by the agent.
- Built Gitar's GitHub pull request integration, which enables Gitar to act seamlessly on GitHub PRs to address comments and CI failures (similar to Claude Code Action).
- Worked directly with enterprise customers, including multiple Fortune 500s, to ensure Gitar integrates into their workflows across different Git hosting providers, CI/CD platforms, and monitoring solutions.

Northeastern University, Programming Research Lab, *Dev Tools & AI Agents PhD* Sep 2020 – Nov 2024

- Built agentic package managers (replacements for NPM and PIP) that can automatically apply patches based on runtime errors, avoid packages with security vulnerabilities, minimize code size, and more.
- Designed a dependency resolution and optimization agent tool that is called by the agentic package manager, enabling the LLM to focus on high-level reasoning and delegate the precise dependency solving.
- Mentored an undergraduate collaborator in building a distributed system to scrape and archive every NPM package (36M+, 20+ TB) as they are published in real-time. Used this dataset to empirically analyze developer pain points and evaluate agentic package managers.
- Developed a novel methodology (MultiPL-E) to automatically translate the HumanEval and MBPP benchmarks, enabling multilingual evaluation of LLMs. MultiPL-E now supports 31 different evaluation languages, and is used actively on the Big Code Models Leaderboard on HuggingFace.

Draper Laboratory, *Security and Safety Research Scientist* Feb 2023 – May 2023

- Funded by a DARPA defensive cybersecurity research program (AMP), enhanced the safety of live-patching binary code by developing formal methods to automatically verify correctness of binary security patches.

Uber, Programming Systems Group, *Dev Tools Research Scientist* May 2020 – Dec 2020

- Repaired over 75% of flaky tests across Uber by designing dynamic analysis-based tooling informed by natural language processing and clustering of crash logs. Evaluated with large-scale experiments on all tests across Uber.

University of Massachusetts Amherst, *Compilers & Cloud Infrastructure MS* Sep 2018 – May 2020

- Pioneered the study of formal semantics for serverless computing (FaaS), laying a theoretical and empirical foundation for cloud providers to develop new FaaS abstractions, such as Microsoft Azure's Durable Functions.
- Reduced code size by 23% and sped up programs by 15% for multithreaded WebAssembly by extending a JIT compiler (Wasmtime) with stack capture and resume instructions at the virtual machine level.

Apple Inc., *macOS Frameworks Team, Intern* Jun 2016 – Aug 2016

- Created a new user-interface feature simplifying tab navigation in the native macOS UI framework (AppKit), and perfected reliability of it across first-party and third-party apps so it could ship in macOS High Sierra.
- Presented the feature before a distinguished panel, including Apple's Senior Vice President Craig Federighi, earning recognition as one of the top 10 intern projects from a pool of hundreds of competitors.

Education

Northeastern University, PhD in Computer Science Sep 2020 – Nov 2024
Focused on Dev Tools & AI Agents, GPA: 4.00, Advised by Drs. Arjun Guha and Jonathan Bell

University of Massachusetts Amherst, MS in Computer Science Sep 2018 – May 2020
Focused on Compilers & Cloud Infrastructure, GPA: 3.87, Advised by Drs. Arjun Guha and Yuriy Brun

University of California Davis, BS in Computer Science and Mathematics Sep 2014 – Jun 2018
Double Major in Computer Science & Engineering and Mathematics, GPA: 3.94

Publications

PhD Dissertation 2024 *Improving Dependency Management via Formal Semantics.* **Donald Pinckney**.
[PDF]. Keywords: LLMs, Agents, Empirical Methods, Evaluations.

ICSE 2023 (Acceptance rate: 15%) *Flexible and Optimal Dependency Management via Max-SMT.* **Donald Pinckney**, Federico Cassano, Arjun Guha, Jonathan Bell, Massimiliano Culpo, Todd Gamblin.
[paper] [talk] [github] [install]. Keywords: Program Analysis, Empirical Methods, Theoretical Frameworks.

MSR 2023 *A Large Scale Analysis of Semantic Versioning in NPM.* **Donald Pinckney**, Federico Cassano, Arjun Guha, Jonathan Bell.
[paper]. Keywords: Empirical Methods, Data Collection, Data Analysis.

ESEC/FSE 2023 Demo Track *npm-follower: A Complete Dataset Tracking the NPM Ecosystem.* **Donald Pinckney**, Federico Cassano, Arjun Guha, Jonathan Bell.
[paper] [talk] [dataset] [github]. Keywords: Data Collection, Data Analysis.

TSE 2023 *MultiPL-E: A Scalable and Polyglot Approach to Benchmarking Neural Code Generation.* **Donald Pinckney**, Federico Cassano, John Gouwar, Daniel Nguyen, Sydney Nguyen, Luna Phipps-Costin, Ming-Ho Yee, Yangtian Zi, Carolyn Jane Anderson, Molly Q Feldman, Arjun Guha, Michael Greenberg, Abhinav Jangda.
[paper] [talk] [github] [website]. Keywords: LLMs, Empirical Methods, Model Benchmarking, Data Generation.

DLS 2020 *Wasm/k: Delimited Continuations for WebAssembly.* **Donald Pinckney**, Yuriy Brun, Arjun Guha.
[paper] [talk] [github] [website]. Keywords: Compilers, Program Analysis, Empirical Methods, Performance Benchmarking.

OOPSLA 2019 (Acceptance rate: 15%), Distinguished Paper Award *Formal Foundations of Serverless Computing.* Abhinav Jangda, **Donald Pinckney**, Yuriy Brun, Arjun Guha.
[paper] [talk] [website]. Keywords: Cloud Infrastructure, Empirical Methods, Theoretical Frameworks, Performance Benchmarking.