

Matt Welsh

<https://www.mdw.la>
<https://www.linkedin.com/in/welsh-matt/>

Email: mdw@mdw.la
Based in Seattle, WA

Professional Summary

I am a technology leader with 22 years of experience as a startup founder, engineering lead, software engineer, people manager, and professor. I've been the co-founder of Fixie.ai, head of engineering for OctoAI, led teams at Google and Apple, and was a tenured Professor of Computer Science at Harvard. I've built large systems supporting more than a billion users, deployed wireless sensors on active volcanoes, and published more than 70 scientific papers. My technical background is in machine learning systems, distributed systems, networks, mobile computing, and embedded systems. I love working on the bleeding edge of technology and developing products to improve the lives of users all over the world. I enjoy both leading teams and writing code.

Experience

Distinguished Engineer
Aryn.ai

Seattle, WA
March 2022 - March 2024

Aryn is building an unstructured data warehouse that leverages the power of Large Language Models to process and query complex documents. I am leading our engineering efforts to build an AI-powered query engine for complex, unstructured data.

Co-founder and Chief Architect
Fixie.ai

Seattle, WA
March 2022 - March 2024

I co-founded Fixie.ai to build an enterprise AI development platform to enable companies to leverage the power of Large Language Models in their own applications. We raised \$17M in VC funding from Redpoint, Madrona, Zetta, and SignalFire. As the original CEO and then Chief Architect, I was responsible for all of the day-to-day operations of the company from its inception, and led our first two rounds of fundraising. I designed and implemented the bulk of Fixie's original platform, and owned the core service logic, cloud infrastructure, LLM integrations, and developer SDKs. I also contributed substantially to the web front-end. I led our hiring efforts and grew the team to 10 people.

In early 2024, the company pivoted into a very different space and I left to explore other interests.

SVP of Engineering
OctoAI

Seattle, WA
April 2020 - March 2022

OctoAI is developing technology to optimize machine learning models for deployment. As the head of engineering, I grew the engineering org from about 10 people to more than 70 engineers over two years. At the time, OctoAI was developing a cloud-based service that optimized and packaged ML models, powered by Apache TVM, an open-source compiler framework for ML. I took care of all aspects of the engineering organization, from setting overall technical direction, recruiting world-class engineers, mentoring and onboarding, and building an inclusive company culture. I also contributed substantially to OctoAI's software stack, developing code in Rust and Python to power OctoAI's cloud platform. OctoAI was acquired by NVidia in 2024.

Software Development Engineer (T6)
Apple

Seattle, WA
January 2020 - April 2020

I joined Apple as the technical lead and manager of the former Xnor Platform team, as part of the AI/ML organization based in Seattle. We developed new capabilities to support fast model training and inference on Apple hardware, leveraging the special capabilities of Apple's CPU, GPU, and Apple Neural Engine chips.

Principal Engineer
Xnor.ai

Seattle, WA
March 2019 - January 2020

I was head of the Platform and Systems teams at Xnor.ai, a Seattle-based startup that developed fast, power-efficient AI capabilities for edge devices, such as low-power CPUs. Xnor's technology made it possible to run, for example, 30 fps real-time object and person detection and tracking on an embedded device such as a Raspberry Pi. I was both an individual contributor, designing and building major pieces of our systems, as well as a manager of two teams comprising 12 engineers. Our projects included:

- Xnor's CPU-efficient inference engine, which compresses and optimizes ML models to run on low-power devices.
- Xnor's AI2GO platform (<https://ai2go.xnor.ai>), a free SDK making it possible to embed fast AI models with just a few lines of code.
- Xnor's infrastructure for model training, testing, benchmarking, continuous integration, and code health.
- Specialized capabilities and integration support for Xnor's largest customer engagements.

Xnor was acquired by Apple in January 2020.

Principal Engineer and Engineering Director
Google, Inc.

Seattle, WA
June 2010 - March 2019

I led the Chrome Mobile team in Seattle and Kirkland, with a focus on optimizing the web experience for users in emerging markets such as India, Indonesia, and sub-Saharan Africa. I started the team and grew it to more than 40 engineers. As a result of our team's work, grew Chrome to more than a billion users in emerging markets. Our work spanned server-side and client-side technologies. Some of my team's projects include:

- Flywheel, a proxy service for compressing mobile web pages, in use by more than 650M users of Chrome, driving tens of billions of queries per day.
- Client-side optimizations to simplify page loading and rendering, using server hinting to tailor the optimizations for individual web pages.
- Adding capabilities to Chrome to enable downloading and prefetching of web content to enable users with intermittent connectivity to browse while offline.
- New browser features to help users discover new web content and more effectively multitask on small screens.

Professor of Computer Science
Harvard University

Cambridge, MA
July 2003 - July 2011

Led a research team of more than a dozen graduate students, postdocs, undergrads, and research staff focused on exploring the technology ramifications of tiny, embedded wireless sensors, with applications such as geological monitoring and healthcare.

- We developed new operating systems, network protocols, and programming models for networks of embedded, 8-bit microcontrollers with low-power radios.
- We evaluated our work through deep engagements with domain scientists, deploying sensor networks on two active volcanoes in Ecuador, on street lights throughout the city of Cambridge, and in lab tests for patients being treated for Parkinson's Disease.
- Taught undergraduate and graduate courses on Operating Systems, Computer Architecture, Distributed Systems, and Wireless Networking.
- Awarded tenure and promoted to full professor in 2010.

Senior Researcher
Intel Research

Berkeley, CA
August 2002 - July 2003

Following my PhD work at Berkeley, I spent a year at Intel Research, developing a new programming language and compiler chain for embedded wireless sensors, called NesC. This work was published in PLDI 2003, won the "most influential paper" award 10 years later, and has been cited more than 2500 times.

Skills

- I have extensive experience in machine learning systems, operating systems, networks, wireless, mobile computing, and distributed systems, and have published papers in and served on program committees for numerous conferences in these areas.
- Long track record of driving successful technology projects, ranging from academic research efforts to multi-team projects supporting more than a billion users.
- Strong full-stack and infrastructure software development expertise. I'm fluent in Python, Go, Rust, TypeScript, JavaScript, C++, Java, Swift, and SQL, with extensive experience developing with frameworks such as React, Next.js, Django, Actix, Google Cloud, AWS, Kubernetes, iOS, and Android. I still write tons of code.
- Seasoned organizational leader with experience starting and growing companies; building, scaling, and reorganizing teams; mentoring and career development; managing performance and balancing engineering investments.
- Strong communication skills as an author of numerous scientific articles and books, as well as experience as a teacher, technical speaker, and mentor.
- I am passionate about leveraging technology to address humanitarian needs and improve the lives of people.

Full academic CV with publication list is available at <https://www.mdw.la/mattwelsh-cv.pdf>

Education

Ph.D., Computer Science

UC Berkeley, December 2002

Thesis title: *An Architecture for Highly Concurrent, Well-Conditioned Internet Services*

Advisors: Prof. David Culler and Prof. Eric Brewer

Masters of Science, Computer Science

UC Berkeley, December 1999

Bachelor of Science, Computer Science

Cornell University, May 1996