

Paul Davidson

(Last updated: August 12th, 2026)

[LinkedIn](#) · [Personal Website](#)

Focus

My current observation of AI Safety is that the space is very fragmented and in an early state of maturity while evolving rapidly. Although I may go deep and narrow to initially prove myself, I think my most meaningful contributions will come from zooming out and supporting the larger bodies of work by connecting ideas, encouraging standardization, improving usability and supporting meta-analysis.

Improving automated evals is a natural place for me to start with projects, and hopefully future roles. [I wrote my initial assessment of the space](#) before the Technical AI Safety Course which includes ideas like implementing multiple judges, creating a scientific playbook and program, and democratizing then studying setups and outputs. There's also an emerging study of evaluating multi-agents which will further expand the space, and I find it really interesting as well.

My role in AI safety will serve as an amplifier by leaning into my 18 year career as a proven generalist, and transferring those skills, particularly from running an online experimentation platform and program.

Target roles

- Medium (20-30+ person) and growing org: Product Manager, (Technical) Program Manager, Research Operations - serve as an amplifier to engineer and researcher's iterative efforts.
- Founder - only on very specific meta-type initiatives, not on a research category or niche

Experience

Doubled experimentation volume 2 years in a row, reduced analysis time by 40%, leading to 500 concurrent cross-platform tests, and \$1 billion in estimated loss prevention from unsafe releases.

- **Online experimentation platform:** Spent 3.5 years titled as the product manager of a 20-person team that launched and evolved Expedia Group's experimentation platform to be more democratic, safer, easily understood, deliver trustworthy results and support meta-analysis.

- **Online experimentation program:** Those same 3.5 years I carried overlapping responsibilities of working within the 15,000 person org to promote shift of user roles, and best scientific practices.

Going 0-1 and leading other successful initiatives across a generalist career:

- **ML and AI systems:** Supported testing ML models at Expedia Group, and proposed AI-powered test analysis insights plus meta-analysis. Built and grew ni.com's first chat product. Developed a personal second brain project using Claude Code.
- **Data platforms:** Expedia's experimentation platform processed billions of events, petabytes of data into metrics and insights daily.
- **Documentation systems:** Led the initial phases of NI's transition to a system that could dynamically re-construct their documentation on the fly based on metadata and tagging.
- **Software development:** Supported 20-person engineering team. School and work projects in VB, C++, and LabVIEW. Light editing and debugging of SQL and Python. Developed a custom personal website and sub-site using Claude Code. Comfortable reading and debugging code.
- **Communication and strategy:** 10 years of running discovering, proposing strategies, cross-functionally aligning leadership stakeholders, measuring success, implementing and releasing projects.

Engagement with AI safety

- **AI Safety Courses:** BlueDot Impact AGI Strategy & Technical AI Safety (both completed), ARENA self-paced, focus on Python & PyTorch (August\September 2026)
- **Projects:** BlueDot Technical AI Safety August 2026 (applying), SPAR Fall 2026 (applied)
- **Coaching:** Successif (applied), BlueDot Impact (applied), 80,000 Hours (applied)
- **Programs:** Lateral Workshop (applied)
- **Community:** Berkeley EAGx (August 2026), BlueDot Evals Reading Club and Austin Meetup
- **Writing:** 1 [blog post](#) linkposted to [EA Forums](#), 11 others backlogged