

Piyush Vyas

Forward-Deployed Architect | Agentic AI | Creator of Browser Agent Protocol

812-391-4102 • work@piyushvyas.com • piyushvyas.com • LinkedIn • GitHub

EXECUTIVE PROFILE

Hands-on technical leader with **5.5+ years** building and shipping production AI systems at enterprise scale, with deep experience across **applied research, ML engineering, product development, solutions architecture, and go-to-market**. Led **30+ GenAI engagements** for Fortune 500 and startups, delivering **40 to 70% cost reductions** and contributing to **8-figure pipeline**. Personally architect and code production systems while leading teams and driving **multi-million dollar ARR**. Built and open-sourced **Browser Agent Protocol (BAP)**, an open standard for AI agents to control web browsers with native MCP integration, and **skill-tools**, the first complete toolchain for the Agent Skills ecosystem, enabling parsing, validation, scoring, routing, and generation of SKILL.md files with zero LLM dependencies. Published researcher with **Best Paper and Outstanding Paper awards**. Operate fluidly across engineering, product, sales, and executive leadership. Build zero-to-one and scale what works.

VALUE DELIVERED

30+

Deployments Shipped

8-Fig

Pipeline Influenced

40–70%

Cost Reductions

8 Wk

Avg. Time-to-Value

PROFESSIONAL EXPERIENCE

Deep Learning Architect / Delivery Lead

May 2024 to Present

Amazon Web Services, Generative AI Innovation & Delivery

Scope: End-to-end technical and business ownership of GenAI engagements across AMER/EMEA, from discovery through production deployment.

- Led **30+ GenAI production deployments** for Fortune 500 and startups, averaging **8-week time-to-value** from discovery to launch and influencing **8-figure pipeline**
- Architected and coded agentic AI, RAG, IDP, conversational AI, and browser automation systems. Developed **15+ reference architectures** and production code samples adopted network-wide
- Delivered **35+ executive workshops** for C-suite and engineering leaders. Orchestrated cross-functional teams of 4 to 6 including consultants, FDEs, and applied scientists
- Drove **40 to 50% operational cost savings** for customers. Accelerated partner GTM through reusable assets and enablement programs
- Partnered with product teams, SI partners, and ISV alliances to shape roadmaps and co-develop joint solutions

Senior Machine Learning Engineer

Apr 2023 to May 2024

Numerade Labs, EdTech Startup (Series A)

Scope: Technical lead for applied AI with full ownership from ideation to production. Direct report to VP of Engineering.

- Owned AI product roadmap and technical strategy across all ML initiatives
- Built AI video generation system from scratch (image gen, animation, TTS, avatars) and deployed LLMs via vLLM achieving **10x throughput** and **60% memory reduction**
- Built and managed team of 4 ML engineers and 2 interns. Promoted 1 to senior role within 12 months
- Led B2B adoption across **50+ academic partners** achieving **90%+ feature adoption**, directly contributing to Series A growth metrics
- Delivered **70% cost reduction**, **30% adoption increase**, and **45% engagement lift**

Senior Machine Learning Scientist

Oct 2021 to Apr 2023

Spectrum Labs, Trust & Safety Startup (Acquired)

Scope: Founding scientist for speech AI capability. Zero-to-one ownership from research through production at scale. Direct report to Head of Data Science.

- Built proprietary LLMs from scratch including fine-tuning, evaluation, quantization, and deployment. Created PyTorch toolkit adopted by entire DS team
- Led team of data scientists and ML engineers. Established research-to-production processes and mentored junior engineers into technical leads
- Scaled system to **12+ languages** processing **billions of daily messages**, achieving **40% latency reduction** at 98% accuracy
- Partnered with CTO and GTM on product strategy and B2B customer engagements influencing **multi-million dollar ARR**

Research Engineer

Aug 2021 to Oct 2021

Comcast Applied AI

- Built production ASR system (wav2vec 2.0) achieving **30% WER reduction**, serving **millions of users**. Published at **IEEE ICASSP 2022**

Research Engineer

May 2020 to Aug 2021

Indiana University Bloomington

- Led research in speech translation and assessment, publishing at INTERSPEECH and IEEE ICASSP with **2 best paper awards**

SKILLS

Agentic AI	Multi-Agent Orchestration, MCP Servers, Browser Automation, Protocol Design, RAG, Eval Frameworks, Guardrails
Engineering	Python, TypeScript, PyTorch, LangChain, Strands, Playwright, vLLM, FastAPI, Docker
Cloud & Infra	AWS (Bedrock, SageMaker, Lambda), OpenAI API, Anthropic API, Vector Databases
Open Source	Browser Agent Protocol (JSON-RPC/WS, dual TS/Python SDKs) · Skill Tools (BM25 routing, plugin arch)
Delivery	Solutions Architecture, Customer-Embedded Deployment, Product Strategy, Executive Workshops, Sales Engineering, Technical Discovery

PUBLICATIONS

P. Vyas et al., “Optimally Encoding Inductive Biases into the Transformer Improves E2E Speech Translation,” *INTERSPEECH 2021*, **Best Student Paper Award**

Z. Zhang, P. Vyas et al., “End-To-End Non-Intrusive Model for Subjective and Objective Real-World Speech Assessment,” *IEEE ICASSP 2021*, **Outstanding Student Paper Award**

R. Tang, K. Kumar, J. Xin, P. Vyas et al., “Temporal Early Exiting for Streaming Speech Commands Recognition,” *IEEE ICASSP 2022*

EDUCATION

Indiana University Bloomington , M.S. Intelligent Systems Engineering	2020
Devi Ahilya Vishwavidyalaya , Post Graduate Diploma Computer Applications	2018
Devi Ahilya Vishwavidyalaya , Bachelor of Commerce	2017