

Tyler F. Young

AI Solutions Architect & Engineer | Forward Deployed AI Engineer | Agentic AI, Healthcare AI & AI Enablement

Tucson, AZ (Remote, US) | 650.933.0321 | hello@tyleryoung.org | linkedin.com/in/tyleryoung | tyleryoung.org

Summary

AI Solutions Architect and hands-on engineer who takes agentic AI from technical discovery to production in regulated, customer-facing settings. At BAM, a healthcare revenue cycle management (RCM) startup, designs and ships AI agents (voice, browser, and large language model (LLM) pipelines), EHR/FHIR and clearinghouse integrations, evals, and HIPAA guardrails, including a prior-authorization agent now live at a specialty practice. Brings 13+ years of global technical program management to stakeholder alignment, adoption, and measurable ROI.

Skills

AI Agents & LLMs: generative AI, AI agents, agentic workflows, multi-agent orchestration, tool use / function calling, Retrieval-Augmented Generation (RAG), vector search (pgvector), context engineering, prompt engineering, structured outputs, Model Context Protocol (MCP), browser / computer-use agents, voice AI (STT, TTS, realtime), human-in-the-loop (HITL)

Evals & LLMOps: evaluation frameworks (evals), LLM-as-judge, A/B testing, regression gates, guardrails, observability, model routing and fallback, prompt caching, cost and latency optimization, responsible AI

Models & Platforms: Anthropic Claude (Claude Code), OpenAI (GPT, Realtime API), Google Gemini (Vertex AI), AWS Bedrock AgentCore, OpenRouter, Ollama, LangGraph, LangChain, Deepgram, ElevenLabs, Telnyx

Engineering & Cloud: Python, TypeScript, SQL, FastAPI, Next.js / React, Node.js, PostgreSQL, REST APIs, OAuth 2.0 / JWT, AWS (Lambda, SQS, S3, CDK), GCP, Azure, Docker, Kubernetes, GitHub Actions CI/CD, Playwright, pytest

Healthcare: HIPAA, PHI, Business Associate Agreements (BAA), EHR/EMR integration (ModMed), HL7 FHIR R4, SMART Backend Services, X12 270/271 eligibility verification, X12 278 prior authorization, clearinghouses (Availity), payer portals, revenue cycle management (RCM)

Delivery & Leadership: solution architecture, technical discovery, workflow mapping, POCs and MVPs, pilots and go-live, stakeholder management, executive communication, change management, AI adoption and enablement, playbooks and SOPs, roadmaps and OKRs, technical program management, Agile / Scrum, RFP / RFI, vendor management

Experience

AI Solutions Architect & Engineer | BAM (bam.ai) | Remote

Dec 2025 – Present

AI-native revenue cycle management (RCM) platform automating eligibility, billing, and prior authorization for specialty practices.

- Architected and built BAM's multichannel **prior-authorization (PA) agent**, routing each case across clearinghouse e-filing (Availity X12 278), a payer-portal browser agent (AWS Bedrock AgentCore), AI voice calls, and fax under a deterministic state machine with an append-only evidence ledger (360+ new tests).
- Led a live, BAA-covered pilot with a Texas ENT practice through go-live: **resolved 82% of flagged prior-auth requests (56 of 68) in 13 days** across 110+ payer calls to 9 payers.
- Built Switchboard, a hold-aware **AI voice agent** for payer calls (Python, FastAPI, Telnyx, Deepgram, OpenAI Realtime, Gemini Live): paid models stay idle through IVR and hold (54% of call time), for **~5x lower modeled cost per call** than commercial voice-agent platforms.
- Designed the **evaluation framework** for voice and browser agents: a blind two-agent call lab with LLM-as-judge, A/B tests across 19 speech and LLM engines, and cost, latency, accuracy, and AI-disclosure regression gates.
- Integrated ModMed's certified **HL7 FHIR R4 API** (SMART Backend Services, JWT client assertions, JWKS) and hardened Availity eligibility and PA integrations (X12 270/271, 278), **merging ~13.6K lines to production** (FastAPI, Next.js, PostgreSQL, AWS Lambda).
- Defined BAM's **PHI guardrails for AI**: PHI reaches only BAA-covered, zero-retention models (OpenAI, Vertex AI) while other models see redacted data; PHI-free logs and URLs; AES-256-GCM storage; a secrets vault with automated TOTP sign-in.
- Shipped a signed, MDM-deployed capture agent to 30 clinic workstations (56 releases), recording **~7,900 hours of consented workflow footage**, and built the pipeline that mines it into PHI-redacted process maps and role specs (Gemini 2.5 Pro on Vertex AI, Claude, verifier agents).
- Translated observed clinic workflows into a **ranked automation roadmap** of 28 evidence-linked issues with minutes-saved estimates that set BAM's build priorities.
- Prototyped BAM's first desktop product in 10 days (Electron, React, Gemini) and shipped internal agents: a Slack footage "librarian," a PR agent with Claude-critic quality gates, and a meeting-to-PR coding agent (Claude Code).

Head of AI Enablement (Consulting Engagement) | RepeatMD | Remote

Nov 2025 – Mar 2026

Membership, loyalty, and commerce platform for medical aesthetics practices.

- Owned the AI enablement roadmap for CX and platform operations: ran workflow discovery with support and onboarding teams, then shipped LLM applications, adoption plans, SOPs, and human-in-the-loop review patterns.
- Built the RepeatMD **CX Agent** (Next.js, TypeScript, OpenRouter tool calling): RAG over help-center and Notion content (pgvector) plus live Metabase and SQL data, embedded in the admin console as a Chrome side panel.
- Architected **Platform Agent**, a multi-model Python pipeline turning unstructured client inputs (PDFs, spreadsheets, images, Gong call transcripts) into structured app-build specs for ~15 client practices.
- Orchestrated 13 models (8 generation, 5 vision) via OpenRouter with routing and fallback, prompt caching, adaptive chunking, JSON repair, and 8-criterion QA scoring; added a 24-worker Gemini tax-code classifier (US and Canada).

Director of Operations | GreyEye

2023 – 2025

- Led the AI operations roadmap for local inference, workflow automation, and customer-facing prototypes; ran discovery and rapid POCs that **lifted operational efficiency 40% and profitability 15%** in six months.
- Deployed local LLMs (Python, Ollama, GPU) that **cut inference latency 25%**; engineered multi-model consensus across GPT-4, Claude, Gemini, Grok, and o-series models, **improving output accuracy 30%**.
- Built CI/CD for AI workflows (40% faster deployments, 99% production reliability) and migrated workloads to AWS and GCP GPU environments, **cutting infrastructure costs 20%**.

Senior Program Manager | SimulTrans

2018 – 2023

- Owned the cross-functional delivery roadmap for a global quick-service restaurant (QSR) enterprise across 27 countries: **\$100K+ first-year revenue at a 53% profit margin**.
- Integrated GPT-4 into production translation workflows in 2023, **cutting turnaround 30%** with human review controls; automated quoting, vendor management, QA, and invoicing in XTRF (25% faster turnaround).
- Built analytics dashboards and macros that reclaimed 35% of management time and lifted CSAT 15%; restored a critical client relationship, **adding \$500K+ in revenue** (2019–2020).

Program Manager, Talent & Quality | Welocalize

2016 – 2018

- Designed live linguistic quality assurance (LQA) processes that **raised client satisfaction scores to 95%**; automated translator workload distribution (20% faster delivery, 15% less coordinator workload); introduced AI-assisted machine translation post-editing, cutting costs 10%.

Program Manager, IP Translations | MultiLing

2013 – 2016

- Managed **9,000+ patent translation projects a year** for Fortune 100 clients; built IP translation and validation systems that sped processing 90%, raised profitability 36%, and cut errors 25%.

Senior Localization Project Manager | Moravia

2012 – 2013

- Managed global app, web, and technical localization programs at 45%+ profit margins; QA protocols reduced rework 30%.

Education

B.A., Linguistics, Brigham Young University, 2008 (Minor: Language & Computers)

B.A., Portuguese, Brigham Young University, 2008 (24-month residency, Rio de Janeiro, Brazil)

Certifications, Awards & Publications

Certifications: Foreign Patent Filing Certificate, IP Paralegal Institute (2013); Patent Paralegal Certificate, IP Legal Ed. (2013)

Awards (MultiLing): Annual Automation Competition, 1st Place (2013); Employee of the Year (2014); Team of the Year (2015)

Publications: "Translation Pitfalls Mirror the Asian IP Boom," Intellectual Property Magazine (2013); International Guide to Patent Application Translation & Filing, MultiLing (2016); 40 Lessons in Portuguese Grammar (2008)

Languages: English (native), Portuguese (fluent)