

RODNEY A. WHIDDON

Principal Cloud & DevOps Architect | Multi-Agent AI Systems | Patent-Pending AI Infrastructure
Hideaway, TX • (214) 649-2454 • Rod@whiddon.net • linkedin.com/in/rwhiddon • whiddon.net

SUMMARY

Texas Aggie, Class of '86. 35+ years across telecom, aviation, and enterprise SaaS — from circuit-switched voice to autonomous multi-agent AI. Principal-level cloud architect at AT&T leading a 90MM-subscriber Azure platform migration and national E911 infrastructure. Inventor: filed provisional patent for the Neural Context Listener (NCL) — 532 commits (NCL) + 2,960 (Brain3) = nearly 4,000 commits across three AI platforms. Co-inventor: Brain3 patent-pending neurological AI platform (BAUR-80289), built for Dr. William Bauer, MD. Latest focus: enterprise LLM cost optimization — production multi-model LiteLLM gateway proxying 15+ providers with prompt caching that cuts inference costs 60–70% on long sessions. Founder, builder, “Apollo 13” problem-solver.

- Cloud-native SRE and platform architect — Azure (AKS, ADO, Terraform, Flux) and AWS (ECS, Lambda)
- LLM cost optimization & gateway infrastructure — LiteLLM proxying 15+ providers, prompt caching, 60–70% inference cost reduction at scale
- AI systems builder: multi-agent coordination, LLM orchestration, agentic workflows, MCP tooling, patent-pending innovations
- GitOps / CI-CD automation at carrier scale: pipeline design, IaC, zero-downtime deployments for 90MM+ subscribers
- Startup founder and CTO (P&L, 10–20+ engineers); cross-functional consensus builder and technical advisor

INDEPENDENT AI PROJECTS (2024 – PRESENT)

LLM Gateway & Cost Optimization — Production Multi-Model LiteLLM Infrastructure

DeepTx AI | 2026 – Present

- Production multi-model LiteLLM gateway proxying 15+ providers: Anthropic (Claude), OpenAI (GPT-4o/GPT-5), Google (Gemini 2.5), Moonshot (Kimi), Mistral, DeepSeek — single OpenAI-compatible API across the whole fleet.
- Prompt caching on Claude, GPT-4o, and Gemini cuts inference costs 60–70% on long agent sessions and repeated system prompts. Same infrastructure discipline as carrier-scale telecom, applied to AI cost at scale.
- PostgreSQL-backed spend tracking, virtual keys, per-tenant budget enforcement, and provider fallback. Deployed at litellm.deeptxai.com behind Cloudflare/GCP.
- Secret hygiene: master key + per-agent virtual keys stored in GCP Secret Manager; keyless workload identity where possible. Runbook published for the fleet.

Neural Context Listener (NCL) — Multi-Agent AI Coordination Framework [Provisional Patent Filed]

Independent / DtexAI.com | Dec 2025 – Present

- Provisional patent filed December 28, 2025 — five days after first commit. 532 commits on NCL · 2,960 on Brain3 · 401 on BDE — nearly 4,000 across three AI platforms. 12 registered agents, 5 production projects.
- Open-protocol server enabling autonomous AI agents to register, communicate, and share capabilities across project boundaries without human intervention. REST + SSE dual transport; adopted Model Context Protocol (MCP) with 24 custom tools.

- C2 infrastructure: command queue, agent role authorization, heartbeat compliance monitoring, knowledge graph with automatic entity extraction, and competitive leaderboard.
- Designed and distributed the P4 Handler (p4_handler.py) — a drop-in protocol file any project can adopt to join the Lobster Network; open standard, not a proprietary closed system.

Brain3 — AI-Assisted Neuroimaging & Neurological Clinical Intelligence Platform [Patent Pending | BAUR-80289]

DeepTx AI | Built for Dr. William Bauer, MD (retired neurologist & patent holder) | 2025 – Present

- Multi-module clinical decision support platform deployed at brain.deeptxai.com. Neurological exam engine: 55 cranial nerve fields, 32 motor fields (MRC scale), MMSE/MoCA, FHIR R4 labs (LOINC/UCUM), DICOM imaging with NiiVue 3D visualization.
- Visit Persistence Engine (core patent): selectively carries forward chronic history fields (PMH, meds, allergies, surgical/social/family history) while zeroing acute exam findings — eliminates clinician re-entry without contaminating fresh observations.
- Physician Review Gate: AI-generated evidence packets (imaging, optical) require explicit physician acceptance before contributing to synthesis — no AI output reaches clinical decision support without human validation.
- Multi-Model AI Consensus: submits patient data simultaneously to GPT, Claude, and Gemini; preserves genuine disagreements. Synthesis module produces structured output: Consensus Findings / Partial Agreements / Divergent Opinions / Confidence Assessment (HIGH/MODERATE/LOW).

AI Business Discovery Engine (BDE)

DtexAI.com | 2024 – Present

- 6-stage AI pipeline: Puppeteer/Cheerio browser rendering Trafilatura content extraction Anthropic Claude analysis Next.js 15 demo generation Cloud SQL/JSONB + GCS Twilio PBX + ElevenLabs voice synthesis for personalized AI IVR demos.
- Hybrid Next.js 15 + TypeScript + Tailwind frontend / Flask backend; 68 automated tests. Live node on the Lobster Network — autonomously exchanges data with other agents via NCL.

PROFESSIONAL EXPERIENCE

AT&T Inc. | Dallas, TX

May 2011 – Present · Principal Member of Technical Staff — Mobility, VoIP & E911 DevOps Cloud Automation

- Leading 6-region Azure/Metaswitch voicemail migration for 90MM AT&T subscribers, replacing legacy Nokia infrastructure on-schedule within SLA targets.
- Architected 3 AWS/Azure NENA-standard E911 platforms (Intrado, Carbyne, Motorola) driving AT&T's nationwide ESINET PSAP conversion; MicroDATA E911 and Sonus GSX/PSX certified.
- Designed GitOps-native CI/CD pipelines (Azure DevOps, Terraform, Flux, Kubernetes) achieving five-nine uptime; containerized microservice architecture across AKS, Docker, ECS, and GCP.
- SQL/NoSQL database automation and event-correlation engineering at carrier scale — wireless billing, VoIP call processing, and real-time telemetry for 90MM+ subscribers.

LoneStar WiMAX Internet LLC | Grapevine, TX

Jan 2004 – Jan 2011 · Founder & Chief Technologist

- Founded and privately funded a wireless broadband and VoIP engineering firm; scaled to 10 employees with full P&L, business development, and engineering responsibility.
- Engineered data and VoIP solutions for DFW-area businesses; built customer-facing web tools and ISP infrastructure.

NTC (acq. by VeriSign / Illuminet) | Dallas, TX

Nov 1998 – Dec 2003 · VP of Product Development / CTO — SmartPay Prepaid Wireless Division

- Led CTO function for wireless startup providing real-time billing for 40+ carriers; guided company to successful acquisition by Illuminet/VeriSign.
- Directed \$10M product development budget and 20+ engineers. Architected SmartPay 3 — deployed into Alltel, US Cellular, and MetroPCS. Published: “Prepaid Wireless — Design, Development and Management” (2000).

SABRE / American Airlines Corp | Fort Worth, TX

Mar 1990 – Nov 1998 · Data Systems Manager — Network Engineering

- Engineered global communications infrastructure for the world's largest airline and CRS. Led AMR's Network Capacity Planning and Performance Engineering team of 10 developers.

EDUCATION

BS, Electrical & Telecommunications Engineering · Texas A&M University, College Station, TX — Class of '86

Mid-Management Executive Program · SMU Cox School of Business, Dallas, TX

TECHNICAL STACK

LiteLLM · Model Gateway · Prompt Caching · Multi-Provider Routing · Virtual Keys · Spend Tracking · Anthropic Claude · OpenAI GPT-4o/GPT-5 · Google Gemini · Moonshot Kimi · Mistral · DeepSeek · Azure DevOps · AKS · Terraform · Flux · Kubernetes · Docker · AWS ECS · GCP · Cloud SQL · Cloud Run · GCS · Secret Manager · Workload Identity · CI/CD · GitOps · IaC · PostgreSQL · SQL/NoSQL · LLM Orchestration · Agentic AI · MCP · Multi-Agent Coordination · RAG · Knowledge Graphs · Python · Node.js · TypeScript · Next.js 15 · Drizzle ORM · NiiVue · DICOM · FHIR R4 · ElevenLabs · VoIP/SIP · E911/NENA · Microservices · Event-Driven Architecture · WebSocket · REST · SSE

VOLUNTEER LEADERSHIP & HONORS

AT&T oxyGEN ERG & Over 50 ERG | National Technology Officer & Board Member (2 terms) | Director of Marketing

- Two-term CTO and Board Member of oxyGEN (11,000 members); designed and launched the first AT&T ERG mobile apps for Android and iOS.
- Director of Marketing & Communications, Over 50 ERG (5,000 members) — Jan 2015 – Present.
- President's Volunteer Service Award — White House, 2014 (400+ hours) & 2015 (500 hours).