

Pete Stergion

Enterprise AI Enablement & Solution Delivery at Cornell University | Claude Deployments · MCP Integrations · Workflow Automation · AI Governance & Cost Management

Ithaca, NY · github.com/pete-builds · stack.brooksnewmedia.com · linkedin.com/in/pete-builds

AI Tech Lead at Cornell University with 20 years of enterprise IT experience bridging the gap between complex technology and the people who use it. My current work spans two complementary roles: Technical Lead for Cornell's AI Innovation Hub and AI Tech Lead and Desktop Engineer for Cornell University Library. I work with departments across the university to assess needs, gather requirements, evaluate technical feasibility, and deliver practical AI solutions. My experience includes enterprise Claude deployments and staff on-boarding, MCP integrations, reusable Skills and GitHub plugin marketplaces, AI workflow automation, cost management, and security governance. I focus on integrating AI into existing enterprise environments in ways that are secure, maintainable, and useful to the people they serve.

SKILLS

- **Enterprise AI Platforms:** Claude Desktop, Claude Code, Cowork, Microsoft 365 Copilot · Claude Enterprise onboarding and deployment planning · Amazon Bedrock architecture evaluation · LiteLLM AI gateway administration
- **AI Architecture & Integrations:** Model Context Protocol (MCP), FastMCP, streamable HTTP, REST APIs · n8n workflow automation · Claude Skills, GitHub plugin marketplaces, system prompts, agent instructions
- **AI Enablement & Solution Design:** Business requirements gathering, technical feasibility assessment, solution evaluation, staff onboarding and training · Acceptance testing, technical documentation, stakeholder collaboration
- **AI Governance & Cost Management:** Data classification, secure deployment controls, access management, API key lifecycle · AI usage monitoring, token cost optimization, team budgets and chargeback · Human-in-the-loop workflows
- **Development & Automation:** Python, PowerShell, Bash, Docker, GitHub Actions, n8n · CI/CD, software packaging, supply-chain security (SBOM, provenance)
- **Enterprise IT & Security:** Microsoft Entra, Active Directory, Intune, Jamf Pro, SCCM · Vulnerability management, Tenable, endpoint security, enterprise software deployment

EXPERIENCE

Cornell University

Tech Lead, Cornell AI Innovation Hub

Jan 2026 to Present · Part time alongside the Library IT role · Ithaca, NY

- **AI Intake & Solution Design.** Work with university stakeholders to assess AI use cases, gather requirements, evaluate feasibility, and recommend technical approaches. Technical Lead for a Law Library resource discovery project, defining a Claude Skill-based solution with measurable acceptance criteria and validation controls before development. Lead a student engineering team toward a December 2026 demonstration.
- **Enterprise AI Architecture & Governance.** Member of a five-person team supporting Claude Desktop onboarding to Anthropic's Enterprise platform through Cornell's GenAI Premium Service. Collaborate with the AI Hub and central IT to establish technical requirements, deployment restrictions, and security controls aligned with university data risk classifications. Developed onboarding documentation and evaluated deployment paths involving Cornell's AI gateway, Amazon Bedrock, and direct Anthropic subscriptions.
- **Claude Desktop Deployment & Integration.** Built the proof of concept, credential management helpers, and Windows and MacOS deployment packages integrating Claude Desktop with Cornell's AI gateway for approved moderate-risk data use. Coordinated the macOS pilot handoff to central endpoint engineering for native Jamf deployment. The service progressed from a 20-user technical pilot to approximately 200 faculty and staff by September 2026. Created end-user documentation and IT support procedures.
- **MCP & API Integrations.** Developed Cornell's web-search MCP server using n8n and the university's AI gateway, authenticating individual users through their own API credentials rather than shared service keys. Coordinated gateway registration, validated end-to-end functionality in Claude Desktop, and documented the integration for operational handoff.
- **Reusable Claude Skills & Staff Enablement.** Helped onboard university staff to Claude tools and workflows. Contributed a reusable gateway-usage Skill to the AI Hub's public GitHub repository and developed its Claude Code plugin marketplace to simplify Skill distribution and reuse across teams.
- **AI Workflow Automation & Business Impact.** Co-led an AI solution for Cornell's Office of the Treasurer, developing a Python workflow that matches unidentified payments against three years of financial records using AI-assisted search and

analysis. Treasury recovered \$100,000 across five payments following the tool's initial recommendations. Incorporated human review, confidence scoring, and backtesting to evaluate accuracy and identify limitations. Optimized model usage for token efficiency.

AI Tech Lead and Desktop Engineer, Cornell University Library

Oct 2023 to Present · AI Tech Lead title since Jun 2026 · Full-time · Ithaca, NY · Hybrid

- **AI Cost Management & Tokenomics.** Manage AI gateway access (LiteLLM), API credentials, model permissions, budgets, and per-team chargeback for Library IT, allocating usage costs to individual university accounts. Developed tools to improve spending visibility, including a Claude Skill for real-time budget tracking, a consent-based usage reporting tool, and a Claude Code status line displaying live token costs.
- **AI Intake & Technical Evaluation.** Assess departmental AI requests, evaluate feasibility and security requirements, and recommend appropriate technical solutions. Evaluated local open-source models for research workflows, established restricted testing requirements for EBSCO's MCP integration, and assessed the Library catalog's MCP server for context overhead and retrieval accuracy. Developed a validation Skill to identify misleading search results.
- **Enterprise AI Deployment & Security.** Lead the Library's Claude Code pilot with a \$10,000 annual budget, \$50 monthly per-user key limits, and 90-day credential rotation. Established a phased staff onboarding approach and authored a reusable secure deployment proposal using Jamf-managed configuration, enforced AI gateway routing, credential protections, and restrictions on unauthorized tool execution.
- **AI Governance & Risk Management.** Contributed Responsible AI guidelines addressing human oversight, transparency, data sovereignty, and risk-based controls. Serve as a Unit Security Liaison to Cornell's IT Security Office, coordinating vulnerability remediation, security compliance, and risk mitigation across Library IT systems.
- **Staff Training & AI Adoption.** Developed and delivered a Microsoft 365 Copilot Prompt Clinic for more than 76 Library staff, covering practical prompt engineering, responsible AI usage, and university data classification requirements. Help staff understand how to use approved AI tools effectively and safely within Cornell's enterprise environment.
- **Cross-Functional Project Management.** Coordinated Cornell University Library's Springshare MFA rollout, working with staff, system administrators, and central IT to implement authentication requirements and address shared-account security risks. Developed support documentation, communicated implementation deadlines, and guided staff through account transitions and exception procedures to ensure compliance and minimize service disruptions.

Desktop Engineer, Central IT · *Aug 2018 to Oct 2023 · Ithaca, NY*

Partnered with the IT Security Office on the campus-wide Secure Connect (Beyond Identity) passwordless rollout, and led development of the Windows 11 task sequence for Cornell's technical service providers. Worked with the campus technical support providers to make sure our endpoint infrastructure was up to date, reliable, and secure.

Buffalo State University

Client Systems Engineer · *Jul 2006 to Jul 2018 · Buffalo, NY*

Packaging and campus-wide software distribution lead on a three-person team, BitLocker key escrow and recovery, and macOS administration on Jamf.

Brooks New Media

Founder · *2009 to Present · Part-time · Ithaca, NY*

Web services for small businesses on a self-hosted production estate.

SELECTED WORK

mcp-unifi (24 stars). An MCP server that gives agents safe access to production UniFi network infrastructure: a read-only mode that hides every mutating tool, dry-run previews, rollback on partial failure, and a secret-scrubbed audit log. Signed images with a bill of materials and provenance, listed on the official MCP Registry.

open-model-arena. Blind side-by-side model comparison with an Elo leaderboard, so a new model release is judged on its answers, not its name.

writ-protocol. Agent architecture, authorization, governance, original technical design

EDUCATION

Medaille University, Bachelor of Business Administration

Cornell University: INFO 5940 AI Chatbots RAG AI Agents, Turning Point and Developing Facilitation Skills, five-day leadership and facilitation programs