



Austin LangChain AIMUG

Welcome to the Austin LangChain AIMUG - where Texas innovation meets AI exploration! Join our community of tech enthusiasts, developers, and AI practitioners as we learn together in an open, supportive environment.



Today's Agenda

1 Intros

Get to know your fellow AI enthusiasts

2 LangChain / LanGraph Update

Latest developments in the ecosystem

3 Lightning Talks

Agent 2 Agent web experiment showcase

4 Claude AI Dev Workflows

Practical implementation showcase

5 Wrap Up & After Party

Continue the conversation upstairs at The Tavern



Start Time: 6:20 PM

Open AIMUG.ORG – Docs – August 2025

1

Visit AIMUG.org

Our digital home for all resources, events, and community updates

2

Explore Learn Section

Curated tutorials, guides, and educational materials for all skill levels

3

Check July 2025 Resources

Latest documentation, code samples, and community contributions

Our documentation is community-driven and constantly evolving, just like our **Texas-sized** ambitions for AI development.

The screenshot shows the homepage of the AI Middleware Users Group. At the top, there's a navigation bar with links for Learn, Blog, Events, Community, and Support. On the right side, there's a GitHub icon and a sun icon. Below the navigation, there's a sidebar with a tree view of documentation sections: News & Updates, Lightning Talks, Main Event, Resources, Key Highlights, Major Themes, Technical Insights, Community Focus, Quick Navigation, For Immediate Use, For Deep Research, Founding Members Recognition, Corporate Sponsors, Connect & Contribute, Austin LangChain Community, and Community Presenters & Contributors. The main content area features a section titled "July 2025 Documentation" with a sub-section for "Session Recording". It also includes sections for "Documentation Sections", "News & Updates", and "Lightning Talks", each listing specific items related to the July event.



AIMUG – Member Supported – Founding Members

Founding Members

- Cameron Rohn
- Jeff Linwood
- Robert (Muzz)
- Ryan Booth
- Christopher Ewing
- John Van Lowe
- Nate Little
- Sal Castoro
- Colin Best
- Joseph Fluckiger
- Orlando Kalossakas
- Scott Askinosie
- Colin McNamara
- Kaiwalya Joshi
- R Sarasa
- Wendy Schorn
- Collier King
- Karim Lalani
- Riccardo Pirruccio
- Zack Angelo
- Hutch Enbach
- Miguel Lucero
- Roberto Fuentes

Our community thrives thanks to the passion and dedication of our founding members who share a vision for democratizing AI knowledge in the heart of Texas:



Corporate Underwriters



Our initiatives are made possible through the generous support of organizations that believe in the power of community-driven AI education and innovation:

Center for Government and Civil Services

Supporting public sector AI applications and education

LangChain AI

Providing technical resources and framework support

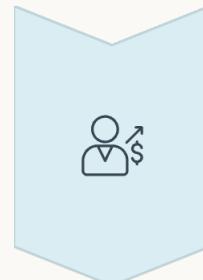
Always Cool Brands / AI

Championing ethical AI development across industries



Toolhouse

Hi, I'm Colin McNamara



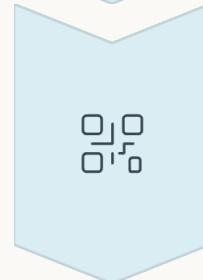
Entrepreneur

Co-founder of Always Cool Brands | AI Using AI to take Poison out of our Food Supply ([LangChain](#) / [LangGraph](#) User)



Community Builder

Founding Member of Austin LangChain AIMUG, committed to growing our Texas AI ecosystem



50% Developer

Consulting Engineer, Cloud Builder, Casual Coder with a passion for practical AI solutions

Public Good: Helping our community succeed by bringing Texas hospitality to the AI world

Connect with me on LinkedIn



Who We Are



1400+

Members Strong

Connecting Austin's tech, creative, and business leaders passionate about AI, bringing together Texas innovation with cutting-edge technology

100%

Learning in the Open

Exploring AI through the lens of LangChain, pulling the ecosystem together with Texas-sized ambition and collaborative spirit

24/7

Supportive Community

A well-balanced Discord that's "useful, but not overwhelming" - reflecting our commitment to southern hospitality in the tech world





Learning in the Open – "Be Cool to Each Other" 🤝

Our community values reflect the best of Texas: friendliness, openness, and a genuine desire to help each other succeed.



Respect & Inclusivity

Our only rule is simple: "Be cool to each other, don't be gross." Texas hospitality meets tech community values.



AI Enthusiasts of All Levels Welcome

We celebrate diverse perspectives and unique paths in AI, whether you're a beginner or expert.

Where and When We Meet



Monthly Showcases, Labs & Mixer

Location: ACC RGC3000

Community members and vendors share projects, demos, and hands-on learning opportunities in a collaborative Texas atmosphere

Weekly Virtual Syncs

Online: Tuesdays at 5PM

Workshopping, problem-solving, and AI discussions from the comfort of your home or office

1

2

3

Hacky Hours (ad hoc)

Location: Rotating Local Bars

Informal happy hours at local spots; meet, discuss, and relax with fellow AI enthusiasts over Texas craft beers

WTF is LangChain?

LangChain is a framework for developing applications powered by language models, enabling:

Context-Aware Applications

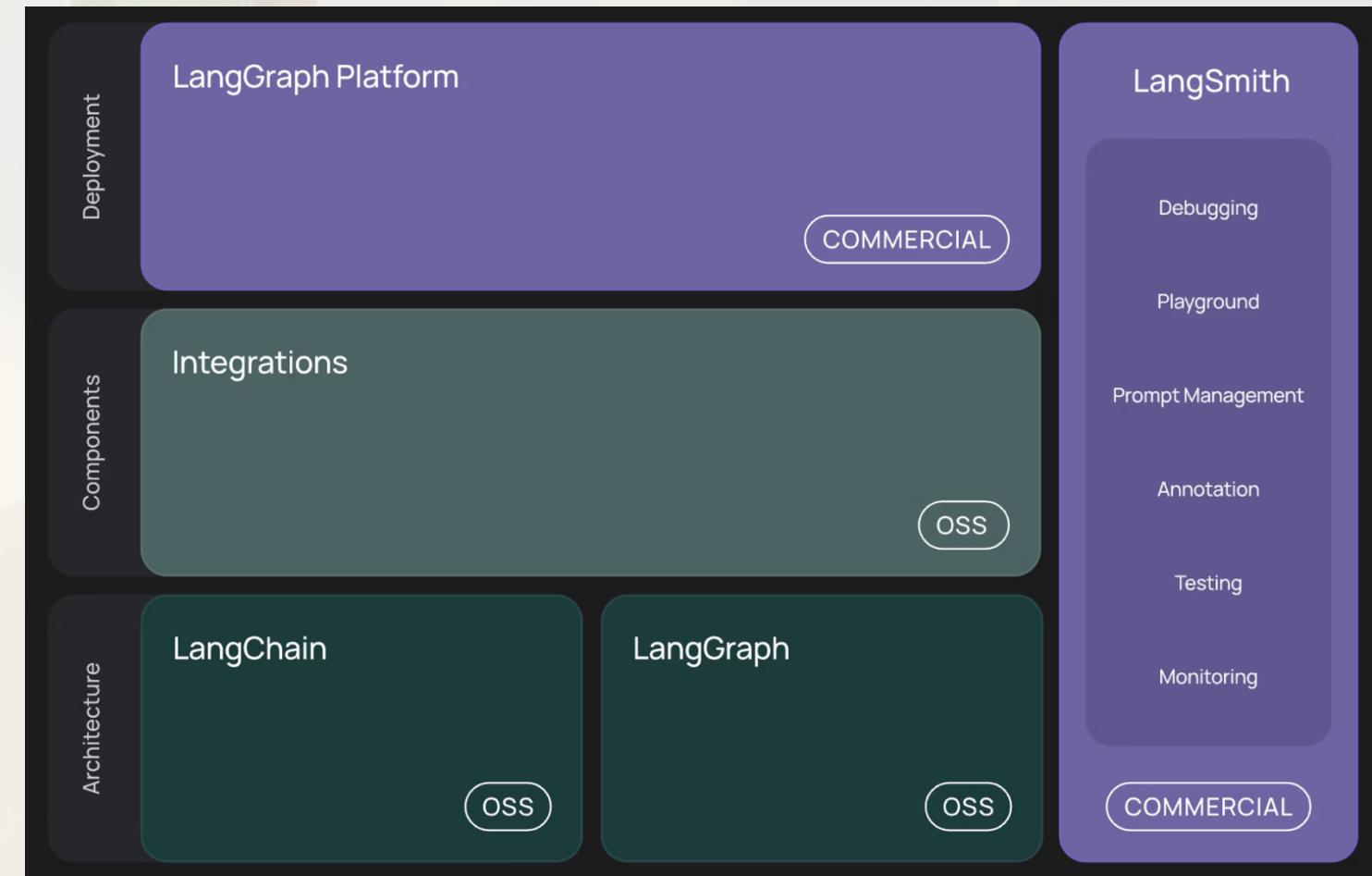
Connect LLMs to external data sources and systems

Reasoning Capabilities

Chain multiple LLM calls together for complex logic

Agent-Based Systems

Create autonomous AI assistants that interact with tools



Just like Texas prides itself on building connections, LangChain connects AI models to make them more powerful and practical.



LangChain Ecosystem Update

August 2025

A technical briefing for AIMUG attendees on recent developments in the LangChain ecosystem, focusing on Deep Agents, LangGraph improvements, and new tools for building sophisticated AI applications.

Deep Agents: Making Long-Horizon Tasks Practical

Harrison Chase's "Deep Agents" approach tackles why simple tool-loop agents feel *shallow* while production-grade agents achieve *depth* through four critical components:

1

Detailed System Prompt

Comprehensive instructions that guide agent behavior and establish consistent patterns

2

Planning Tool

Even a simple to-do list helps structure complex workflows and manage state

3

Sub-Agents

Specialized agents handling specific subtasks to distribute cognitive load

4

File System

Persistent storage for intermediate results and findings across agent sessions



This approach aligns perfectly with LangGraph's strengths in [state management](#), [persistence](#), [subgraphs](#), and [long-running flows](#).

[Python](#) and [TypeScript](#) implementations now available.

LangGraph v0.6: Enhanced Developer Experience

Context API (v0.6.0)

Type-safe runtime context replaces config['configurable'], providing clearer developer ergonomics and better IDE support

Durability Modes (v0.6.0)

Fine-grained checkpoint control with exit/async/sync options, essential for deep agents that run long and branch

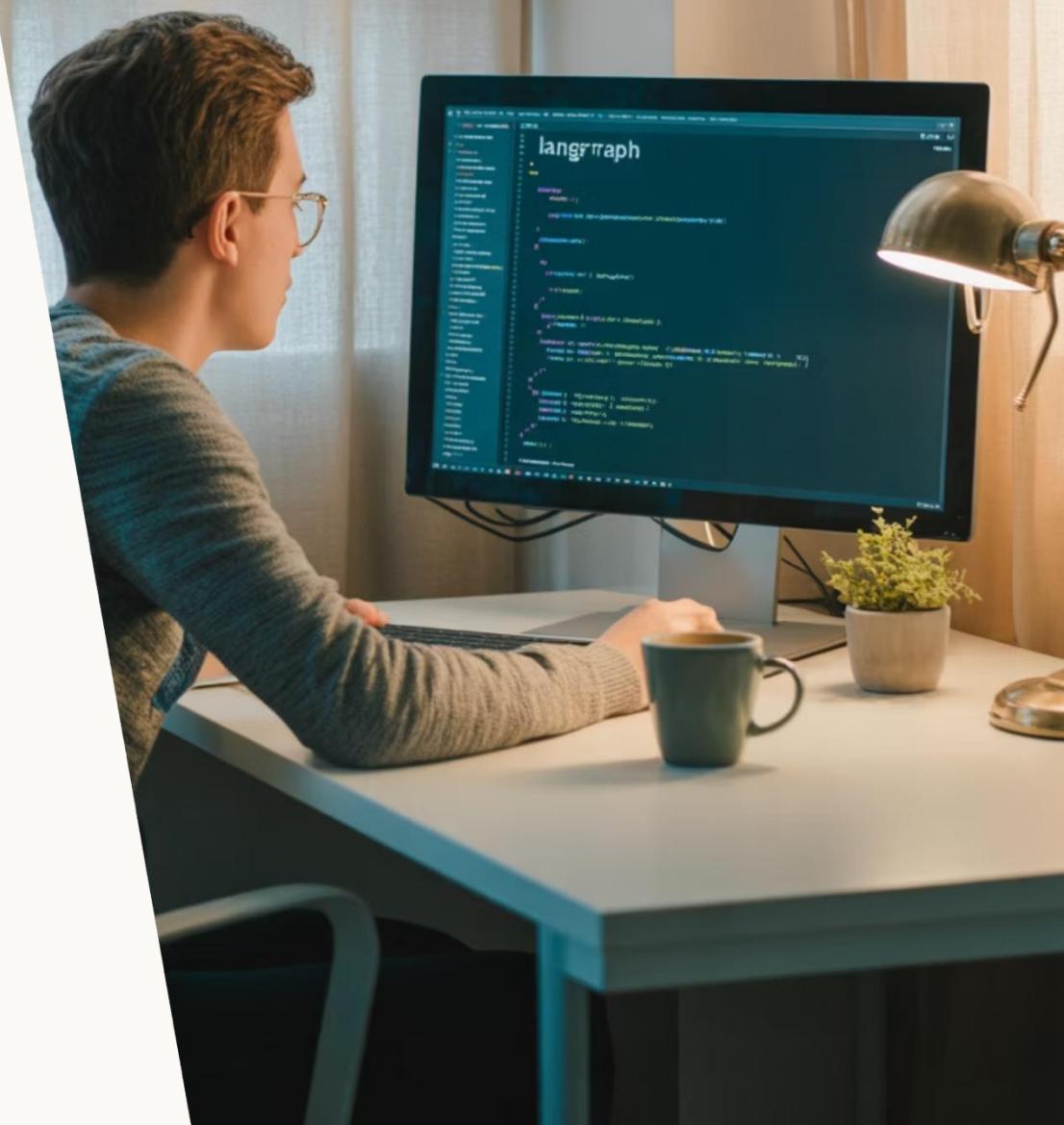
Enhanced Type Safety (v0.6.0)

Comprehensive typing across state/context/input/output schemas reduces runtime errors

Refined APIs (v0.6.0–v0.6.5)

Improved Interrupt interface and public API surface ahead of v1.0

New in v0.6.5: Redis node-level cache implementation for accelerating expensive operations



Open SWE: Asynchronous Coding Agent



Launched August 6th, [Open SWE](#) is a cloud-hosted, long-running coding agent that:

- Plans, codes, tests, and opens PRs autonomously
- Integrates with GitHub Issues for task management
- Uses Daytona sandboxes for safe execution
- Supports interrupts and "double-texting" during operation
- Self-hosts or deploys on LangGraph Platform

Multi-Agent Architecture

Built on [LangGraph](#) with a Manager → Planner → Programmer workflow, including a specialized Reviewer sub-agent

LangSmith

Personal > Tracing Projects > a2a-testing

Home

Observability

Tracing Projects 40

Monitoring 0

Evaluation

Datasets & Experiments 5

Annotation Queues 0

Prompt Engineering

Prompts 0

Playground

LangGraph Platform

Deployments 0

LangSmith & LangGraph Platform Updates

TRACE

Waterfall

Traces LLM

LangGraph 21.73s 1,760

supervisor 1.70s

ChatOpenAI gpt-4 1.60s

route_supervisor 0.00s

tools 5.72s

run_comprehensive_a2a_test 5.72s

discover_a2a_capabilities_v2 1.18s

test_blog_list_posts 0.99s

test_blog_get_post 0.85s

test_blog_search_posts 0.93s

test_blog_get_metadata 0.46s

test_blog_get_author_info 0.46s

test_caching_headers 0.84s

supervisor 11.75s

ChatOpenAI gpt-4 11.74s

route_supervisor 0.00s

analyzer 2.51s

ChatOpenAI gpt-4 2.50s

LangSmith & LangGraph Platform Updates



Align Evals (July 29)

Calibrate LLM evaluators to match human preferences with alignment scores and side-by-side comparisons

Trace ↔ Platform Logs (July 31)

Jump directly from LangSmith traces to LangGraph Platform server logs for unified debugging

Scheduled Exports (July 25)

Keep external systems in sync with automatic trace exports on a regular schedule

Deployment Metrics (July 10)

Monitor CPU/memory usage, latency, and run counts for LangGraph Platform deployments

Operational Note

Increased LangSmith latency incident (Aug 5-6) due to GCP us-central1 issues has been resolved. Review the [post-mortem](#) for SRE insights.

Education & Community Resources

"Deep Research with LangGraph"

New [LangChain Academy](#) course launched **August 14th** – perfect complement to the Deep Agents theme and implementation patterns

In-Person Workshop Opportunity

Academy Live (San Francisco) on August 19th

Hands-on ambient-agents workshop using LangGraph + LangSmith

Both resources provide practical guidance for implementing production-grade agent architectures with the latest LangChain tools.



Ecosystem Signals

Funding Momentum

TechCrunch reported LangChain is raising at approximately [\\$1 billion](#) valuation (July 8)

Indicates strong market confidence in the LangChain ecosystem approach

Server Improvements

Multiple LangGraph Server performance and reliability enhancements (v0.2.80-0.2.94)

- Checkpoint write optimizations
- HTTP metrics instrumentation
- Database schema tuning

✖ Developer Advisory

Some users have encountered deployment regressions after langgraph-api upgrades

Recommendation: Pin dependency versions when rolling forward to avoid unexpected compatibility issues





Key Resources & Next Steps

Documentation & Implementation

- [Deep Agents](#) blog post and [GitHub repo](#)
- [LangGraph v0.6 release notes](#) and changelog
- [Open SWE announcement](#) and technical details
- [LangSmith changelog](#) for platform updates

Learning & Community

- Register for [Deep Research with LangGraph](#) course
- Join [Academy Live workshop](#) in San Francisco (Aug 19)
- Follow [LangChain Forum](#) for implementation tips
- Check [LangSmith Status](#) page for service updates

Questions? Contact the LangChain team or join the community Discord

Lightning Talks!

A2A in the wild – Colin

Agent-to-Agent Communication

Exploring how multiple AI agents can collaborate, share information, and solve complex problems together

Web-Based Implementation

Demonstrating a LangGraph-based test harness showing agents interacting in real-time

Practical Applications

How this technology can be applied to real-world business problems across industries

Future Directions

Where this technology is headed and how you can get involved



Even our mascot is experimenting with agent-to-agent communication!

A2A in the Wild

How enabling Agent-to-Agent on my Astro blog produced agent referrals in 24 hours

Speaker: Colin McNamara (AIMUG) • **Length:** 10 minutes



Why this matters

Agents are becoming the discovery layer (GEO > SEO).

Goal of the experiment: enable A2A on my Astro blog and watch what happens.

If you expose machine-readable capabilities + content, agents can *find, cite, and link* you.

Within 24 hours I saw new referral sources from agents (ChatGPT, Perplexity, Microsoft/CDN).

What is A2A?

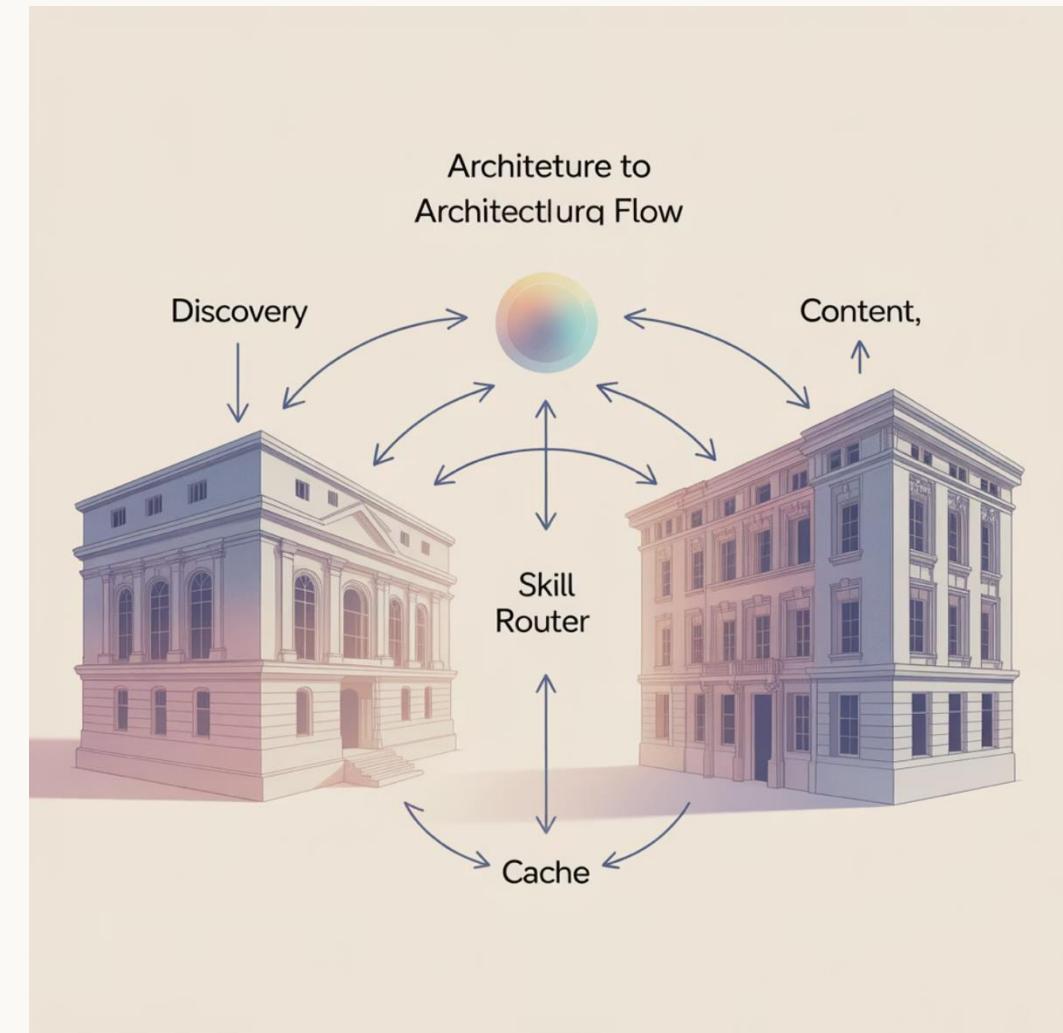
A lightweight, standardized **JSON-RPC 2.0** interface, designed for AI agents to seamlessly discover and interact with web services.

Two primitives:



How I implemented it on Astro

- Astro SSR + TypeScript.
- Added /.well-known/agent.json (capabilities, rate limits, endpoint).
- Built POST /api/a2a/service handling JSON-RPC with validation, CORS, error shaping.
- Implemented blog skills against Astro Content Collections; added **ETags** (304s), **rate-limits**, **security headers**.
- Production-ready file structure + examples are in my implementation guide.
- Architecture at a glance (discovery → gateway → skill router → cache → content).



Robots.txt & Licensing

I explicitly allow:

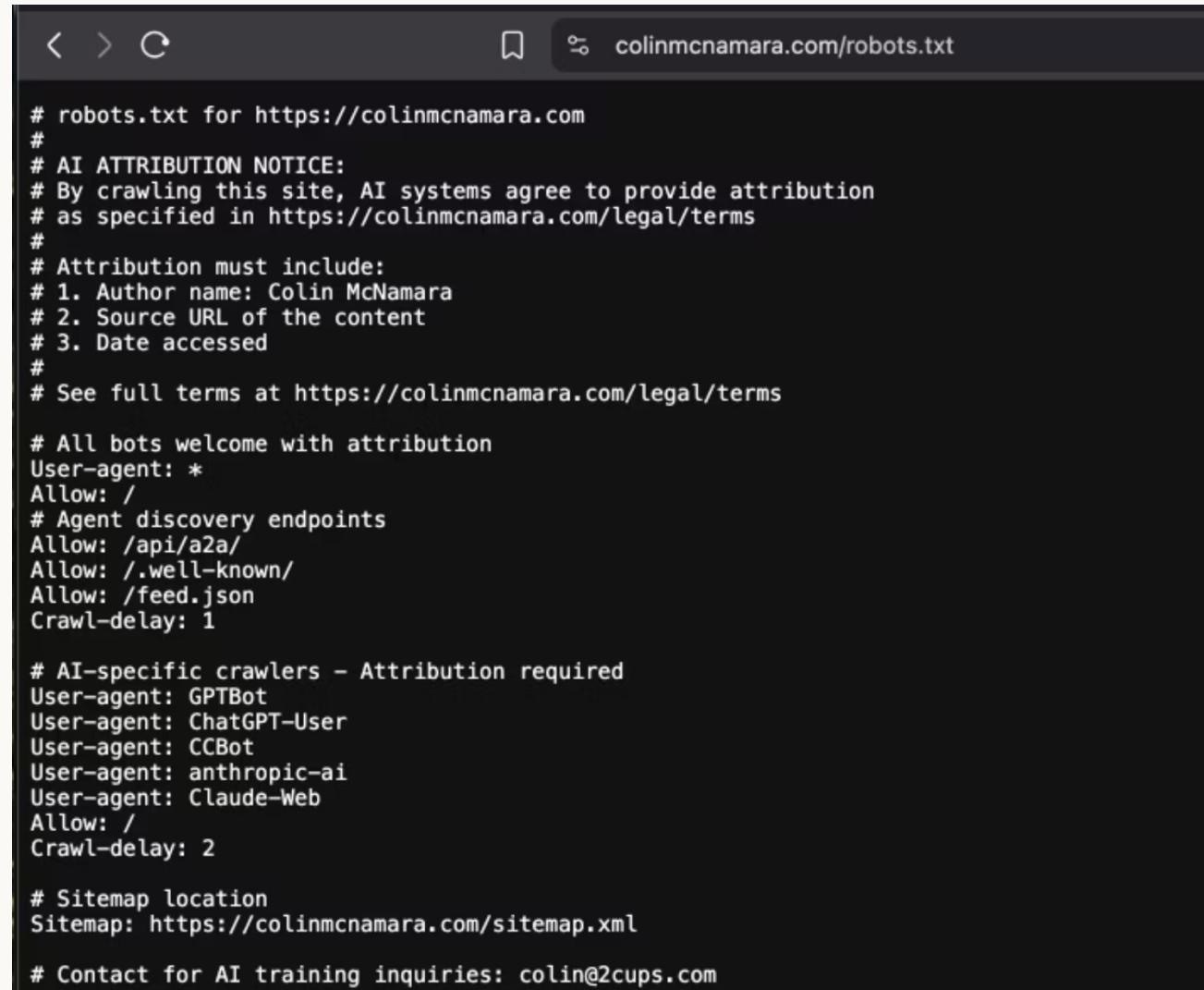
.well-known, /api/a2a, and my feed.json

I set CC-BY 4.0 attribution

requirements in robots.txt so crawlers know my terms.

Net effect:

agents can crawl + call my API, and they know how to attribute.
(Pattern described across the docs.)



A screenshot of a web browser displaying the robots.txt file for the URL colinmcnamara.com/robots.txt. The browser interface includes standard navigation buttons (back, forward, search) and a tab bar. The main content area shows the text of the robots.txt file, which includes comments about AI attribution requirements and specific rules for various user-agents like GPTBot, ChatGPT-User, and CCBot.

```
# robots.txt for https://colinmcnamara.com
#
# AI ATTRIBUTION NOTICE:
# By crawling this site, AI systems agree to provide attribution
# as specified in https://colinmcnamara.com/legal/terms
#
# Attribution must include:
# 1. Author name: Colin McNamara
# 2. Source URL of the content
# 3. Date accessed
#
# See full terms at https://colinmcnamara.com/legal/terms

# All bots welcome with attribution
User-agent: *
Allow: /
# Agent discovery endpoints
Allow: /api/a2a/
Allow: /.well-known/
Allow: /feed.json
Crawl-delay: 1

# AI-specific crawlers - Attribution required
User-agent: GPTBot
User-agent: ChatGPT-User
User-agent: CCBot
User-agent: anthropic-ai
User-agent: Claude-Web
Allow: /
Crawl-delay: 2

# Sitemap location
Sitemap: https://colinmcnamara.com/sitemap.xml

# Contact for AI training inquiries: colin@2cups.com
```

Testing & hardening

Test Approach

LangGraph test harness

validates agent card, every method, error handling, and measures latency.

k6 load tests

for quick perf sanity.

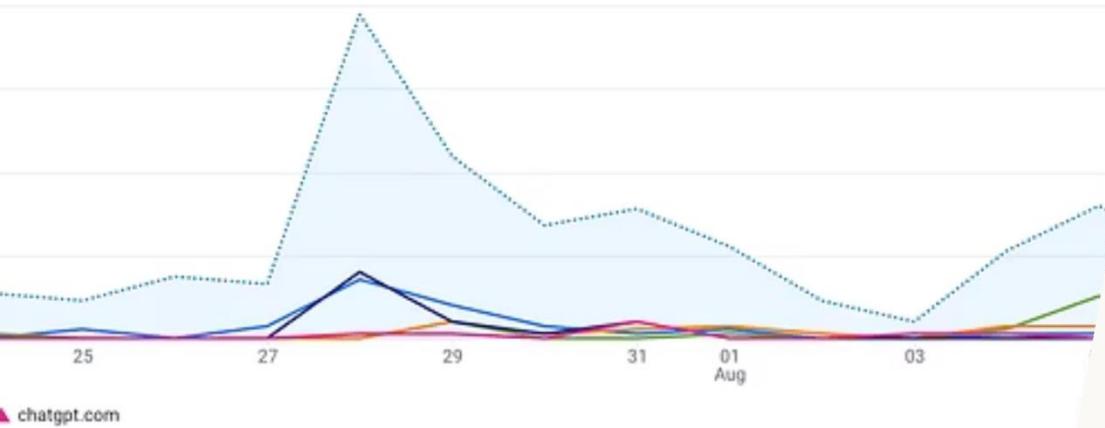
Targets

- JSON-RPC conformance
- ETag 304s
- rate-limit behavior
- CORS
- p95 300ms under modest load



What happened

From your GA4 "Session manual source" report (Jul 17 → Aug 13):



Referrer	Sessions	Engaged sessions	Average engagement time per session	Engaged sessions per active user	Events per session
chatgpt.com	561	167	17s	0.38	4.33
total	100% of total	100% of total	Avg 0%	Avg 0%	Avg 0%
(4%)	352 (62.75%)	91 (54.49%)	14s	0.31	4.7
(5%)	84 (14.97%)	30 (17.96%)	28s	0.47	4
(2%)	38 (6.77%)	11 (6.59%)	10s	0.34	4
(1%)	27 (4.81%)	15 (8.98%)	12s	0.75	4
(6%)	25 (4.46%)	6 (3.59%)	14s	0.43	4
(8%)	13 (2.32%)	8 (4.79%)	47s	0.73	4
(9%)	11 (1.96%)	2 (1.2%)	7s	0.50	4
(3%)	5 (0.89%)	1 (0.6%)	54s	0.20	4
(3%)	2 (0.36%)	0 (0%)	0s	0.00	4
(5%)	2 (0.36%)	0 (0%)	0s	0.00	4

561

Total Sessions

443

Active Users

29.77%

2,427

Events

New agent referrers showed up:

chatgpt.com:

13 sessions (≈2.3%), 47s avg engagement,
61.5% engagement rate

perplexity:

5 sessions (≈0.9%), 54s avg engagement

bing:

11 sessions (≈2.0%), 7s avg engagement

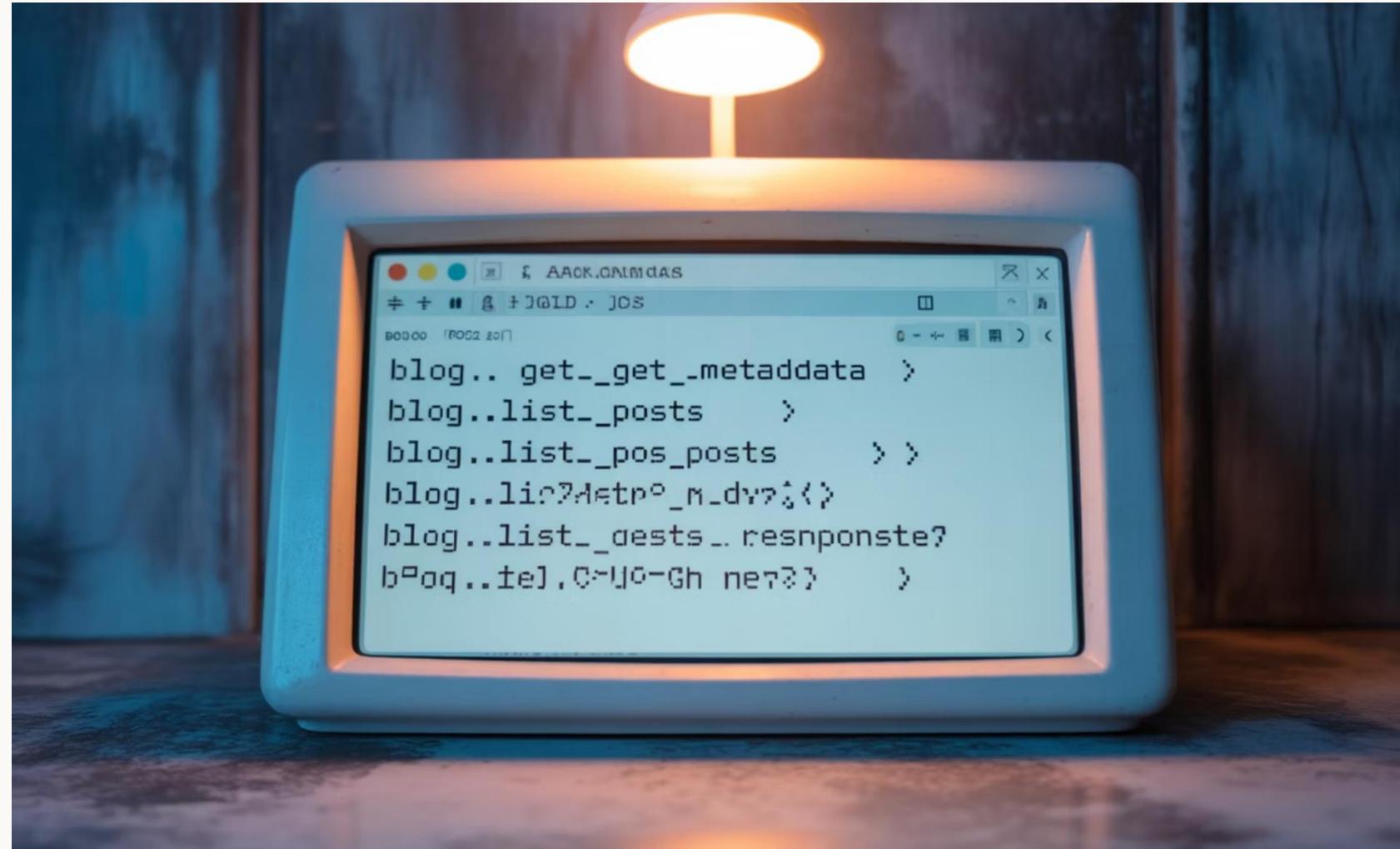
Combined, the three agent sources are ~5.2% of sessions so far (29 / 561).

Takeaway: even a basic A2A surface yields immediate, attributable agent traffic with strong time-on-page from ChatGPT/Perplexity compared to general search. (Use your screenshot as the slide visual and circle the rows for chatgpt.com, perplexity, bing.)

30-second live demo

```
# metadatacurl -sX POST https://colinmcnamara.com/api/a2a/service \-H 'Content-Type: application/json' \-d '{"jsonrpc":"2.0","method":"blog.get_metadata","id":1}' | jq# list postscurl -sX POST https://api/a2a/service \-H 'Content-Type: application/json' \-d '{"jsonrpc":"2.0","method":"blog.list_posts","params":{"limit":5},"id":2}' | jq
```

(These are the same minimal endpoints from the Quick Start.)



How to add A2A in ~60 minutes

01

Create `./well-known/agent.json` with your capabilities + endpoint.

02

Add POST `/api/a2a/service` that understands JSON-RPC 2.0.

03

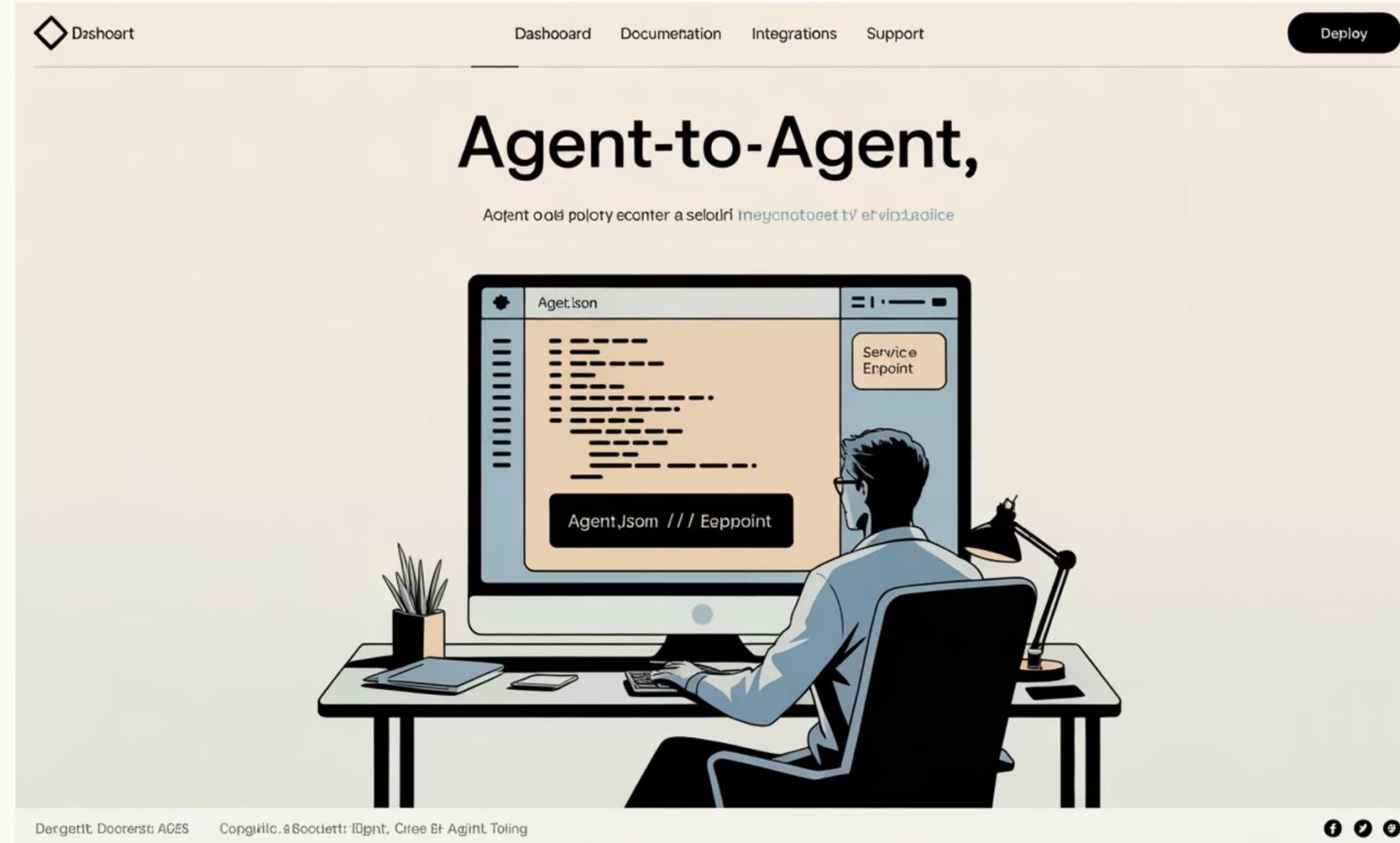
Implement **four** methods: `blog.get_metadata`, `blog.list_posts`, `blog.get_post`, `blog.search_posts`.

04

Add **CORS**, **rate-limit**, **ETags/304**, **security headers**.

05

Test with the LangGraph harness + k6 quick script. Copy-paste boilerplate here: [A2A Quick Start](#) → then level-up with the [Implementation Guide](#).



What's next



Add OTEL

to the endpoint to see *which agents hit which methods*, latency, cache hit-rates.

Expose product/service catalogs

next (not just blog) for real conversions.

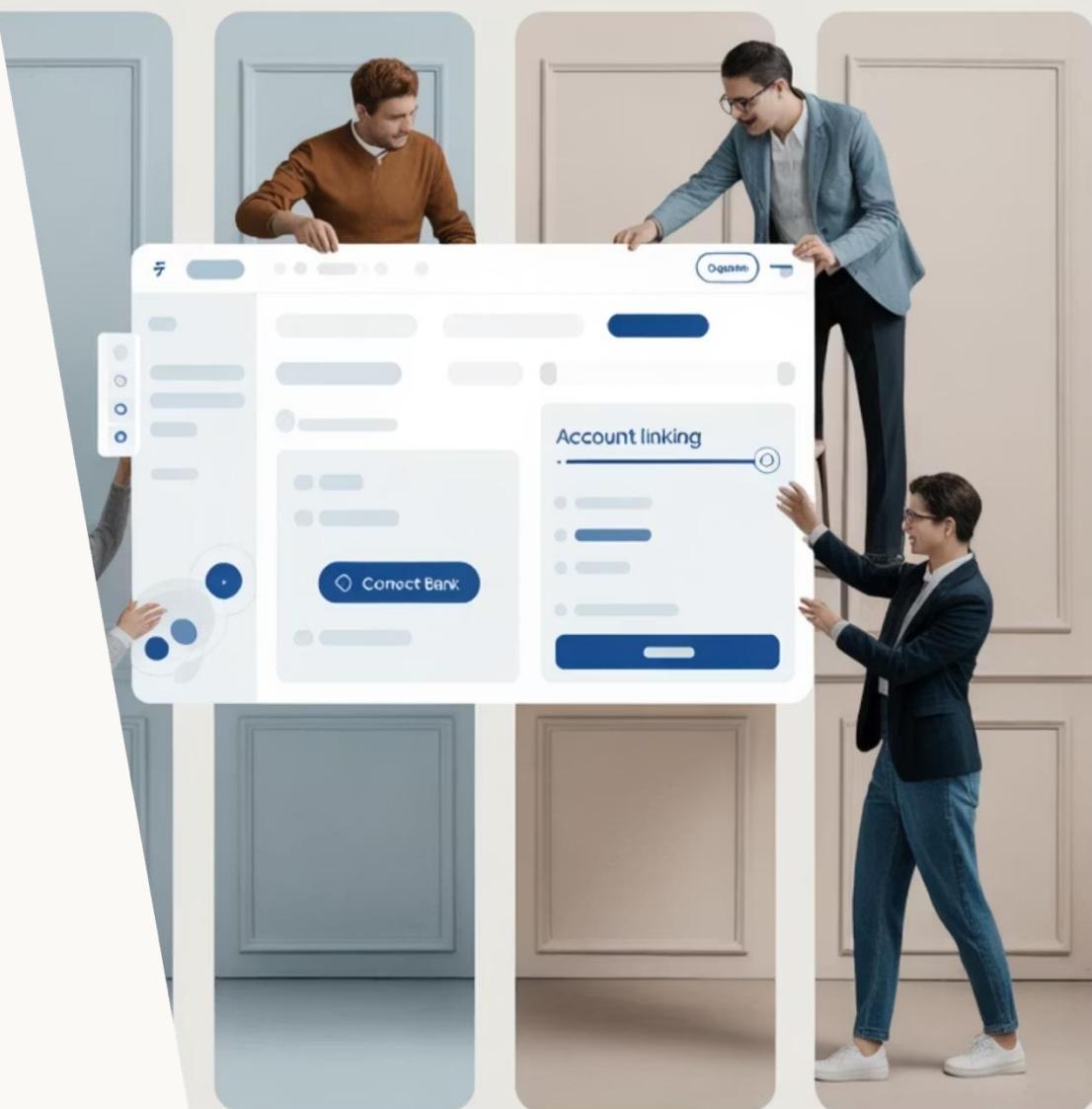
Help the community

let's A/B real **GEO** tactics, share agent referral stats, and build a repeatable playbook.

Business Solutions Platform

Call to action

- 1 Ship your agent card + JSON-RPC this week.
- 2 Post your agent-referral stats in the AIMUG Discord.
- 3 Validate w/ test harness so you can copy/paste.



Appendix — Pointers & snippets you can paste on slides

Agent card example (trimmed)

shows capabilities, rate limits, endpoint.

Minimal service endpoint in Astro

JSON-RPC + four methods.

Architecture blocks

(discovery → gateway → skills → cache → content).

LangGraph test harness

for one-command verification.

Testing checklist

(JSON-RPC conformance, ETags/304, rate limiting, CORS, perf).



Showcase

Claude AI Agentic Development Workflow – Nate

Understanding Claude's Capabilities

Overview of Claude's unique strengths and how they compare to other AI assistants

Setting Up Development Environment

Tools, APIs, and configurations for optimal Claude integration

Building Agentic Workflows

How to design systems where Claude acts as an autonomous agent

Real-World Applications

Case studies and examples from Texas businesses leveraging Claude

Join us as Nate demonstrates how Claude AI can transform your development process with Texas-sized efficiency gains!



RGC-3000 and Tavern After Party

Join us after the meeting to continue conversations, network, and enjoy some Texas hospitality at The Tavern!
Our green parrot mascot might even make an appearance.



Website

aimug.org – All things LangChain and AIMUG



Discord

[Join us](#) for ongoing conversations



GitHub

[Austin LangChain Repo](#) – Check out our projects



Meetup

[Austin LangChain Group](#)



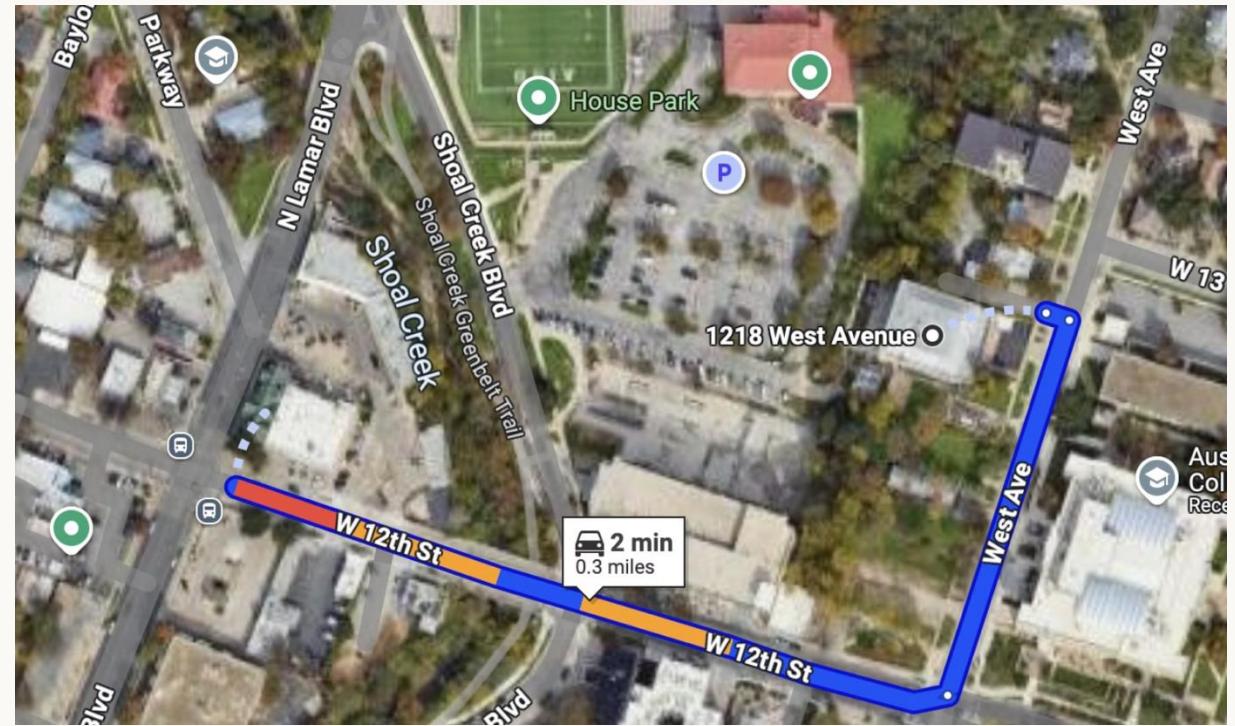
Twitter

[@AustinLangChain](#) – Latest updates



YouTube

[Austin LangChain Channel](#)



ACC RGC3000 & The Tavern

ACC Main Campus



Finding US – ACC RGC-3000

The Austin Community College Rio Grande Campus (RGC) Building 3000 is our regular meeting location, offering modern facilities perfect for our tech-focused gatherings.

Easy Access

Centrally located in downtown Austin with parking available nearby

Tech-Ready Facilities

State-of-the-art presentation equipment and reliable Wi-Fi

Collaborative Spaces

Room arrangements perfect for both presentations and hands-on workshops

Look for Tex, our green parrot mascot with his cowboy hat and gold chain, greeting you at the entrance!



Next Event – Hacky Hour at Kinda Cosmic

Wednesday, August 27th



Casual Networking

Join fellow AI enthusiasts for coffee, conversation, and coding in a relaxed Austin atmosphere



Bring Your Projects

Get feedback, troubleshoot, or collaborate on your latest LangChain experiments



All Welcome

From beginners to experts, our community embraces diverse perspectives and experience levels

"Be cool to each other, don't be gross." Our only rule reflects our commitment to Texas hospitality in the tech world.

Keep an eye out for Tex, our green parrot mascot with his cowboy hat and gold chain – he'll be there too!

