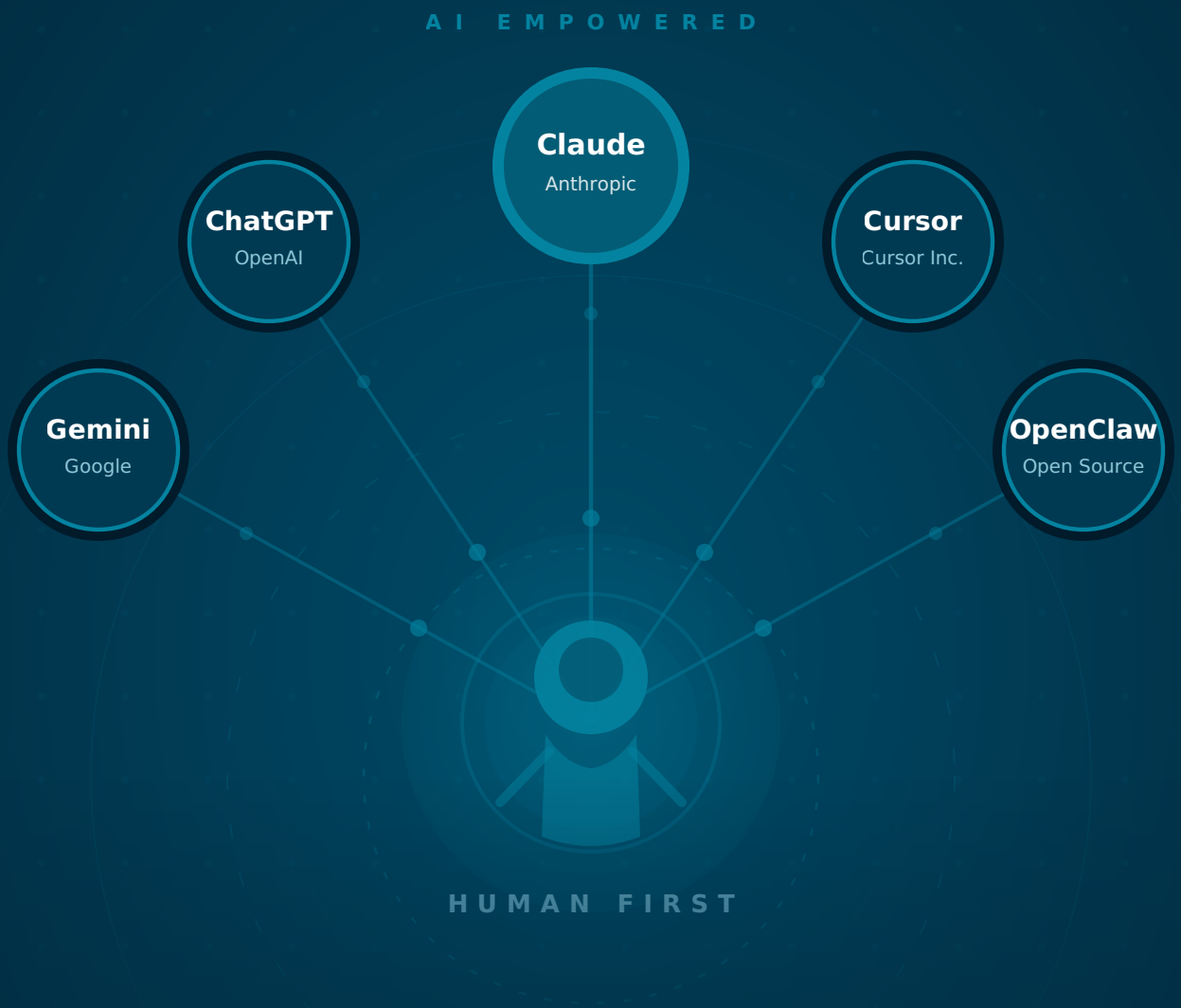

NAVIGATING THE

AI Tool Landscape

A ministry leader’s comparison of today’s top AI platforms — what they offer, where they excel, and what matters most for your mission.

PLATFORMS COMPARED

Claude • ChatGPT • Gemini • Cursor • OpenClaw

fiveq.com

Updated July 1, 2026

What You're Holding and Why We Built It

Ministry leaders are fielding the same question: "Which AI tool should we be using?" The answer is filled with complexity, and technology is emerging so quickly.

Five AI platforms are currently competing for your team's attention, each with different strengths, pricing models, and implications for your operations. We spent a full research session comparing all five — Claude (Anthropic), ChatGPT (OpenAI), Gemini (Google), Cursor, and OpenClaw — across eight capability layers that matter to ministry work.

From how these tools "remember" their role in your organization, to how they handle sensitive data, to whether they can take action without a human in the loop — this is the comparison we wish had existed when our clients first started asking.

We've also included independent benchmark data from researchers at Google, Berkeley, Scale AI, and Vectara — because claims from AI companies about their own tools deserve third-party verification.

> **A note on timing:** This space moves fast. ChatGPT Workspace Agents launched April 22, 2026. Google's Gemini Spark (a 24/7 cloud agent) launched May 19, 2026. All data in this report reflects our best-available research as of **May 26, 2026**.

Who this is for: Executive leaders responsible for AI governance and ROI, marketing practitioners evaluating workflow tools, and IT leaders assessing security and integration requirements.

WHAT'S INSIDE

5	8	5
AI platforms compared	capability layers evaluated	independent benchmarks cited

The Problem No One Is Talking About

Most AI adoption in ministry organizations is happening informally, and no one owns the strategy.

A staff member discovers ChatGPT. Another department is quietly using Claude. Someone in communications set up a Gemini Workspace. No one has a shared framework, no one owns the policy question, and no one has done an apples-to-apples comparison of what each tool actually offers.

The result: organizations end up paying for overlapping subscriptions, security and data governance go unaddressed, and when leadership asks "what is our AI strategy?" — there's no confident answer.

The Three Questions We Hear

① Which tool is actually right for us?

Not "right" in theory — right for your actual workflows, team size, and existing tech stack. Claude excels at nuanced writing and knowledge work. ChatGPT integrates natively with Microsoft 365. Gemini lives inside Google Workspace. Cursor is built for software development teams. OpenClaw offers maximum flexibility but requires significant technical resources to manage safely. The right answer depends on who you are.

② Are we protected?

How does each platform handle your donor data, constituent information, and internal documents? This report includes independent security risk assessments for each platform. One platform on this list currently carries a **high security risk rating** due to a documented malicious-skills attack on its public skill registry — 1,400+ malicious packages identified. You need to know which one before your team starts using it.

③ Is anyone using it incorrectly?

Without a policy, you can't know — and the liability lives with leadership. An AI policy is no longer a "nice to have." It's the governance document that makes AI adoption defensible to your board, your donors, and the people whose data you hold. This report ends with concrete next steps for getting one in place.

> **Five Q's Perspective:** The ministry that gets AI right in 2026 won't be the one that adopted every tool — it will be the one that adopted the right tools with a clear strategy, a written policy, and a team trained to use them well. That's the whole reason we built this report.

What This Means for Your Ministry

Data is most useful when it maps to decisions. Here's how the findings in this report translate to the three roles most involved in AI adoption at ministries.

For Executive Leaders

- **The governance gap is your responsibility.** Every platform in this report requires a written AI policy. Without one, your team makes ad hoc decisions about data, voice, and access with no accountability framework.
- **No AI tool is "set and forget."** ChatGPT Workspace Agents launched April 2026. Gemini Spark launched May 2026. You need someone watching this space on your behalf — the landscape has materially changed in 90 days.
- **The ROI case requires strategic direction.** Claude, ChatGPT, and Gemini are at effectively the same capability tier for most knowledge work. The competitive advantage isn't which AI you use — it's whether your team has been trained to use it well.
- **OpenClaw requires IT ownership.** The ClawHavoc attack is documented and serious. Experimentation without IT oversight is a risk item for your next leadership conversation.

For Marketing Practitioners

- **Claude leads for writing quality.** Claude Opus 4.7 holds the #1 overall Arena Elo and #1 coding Elo, with reviewers citing superior prose rhythm, tone, and subtext. For content-heavy ministry teams, this is meaningful.
- **Skills are the multiplier.** Reusable workflow skills (Claude, Cursor, OpenClaw) compound productivity faster than platforms where prompts live in someone's head. A donor brief skill, social caption skill, meeting summary skill — built once, used forever.
- **Build a prompt library now.** Even on ChatGPT and Gemini, saving prompts as Custom Instructions or Gems creates institutional knowledge that survives staff turnover.
- **Google Workspace teams: Gemini is your natural fit.** Native integration with Docs, Gmail, and Sheets eliminates context-switching and reduces friction significantly.

For IT Leaders

- **MCP is now the integration standard.** All five platforms support the Model Context Protocol. An MCP integration built for Claude will work across ChatGPT, Gemini, and Cursor. Build on the protocol, not the platform.
- **Cloud execution gap is closing but not closed.** Claude Cowork's Routines (research preview, April 2026) run in the cloud, but are still in preview. ChatGPT Tasks and Gemini Scheduled Actions are fully cloud-native. For automations that cannot depend on a local machine, plan accordingly.
- **Treat OpenClaw like an open-source security project.** Pin to v2026.3.28+, install only from the verified registry, and assign someone to monitor ClawHub security advisories.

The Five Platforms We Compared

A quick reference before diving into the data. These five tools represent the most commonly asked-about AI platforms among Five Q clients as of mid-2026.

Claude

by Anthropic

Known for nuanced, high-quality writing and a safety-focused design philosophy. Available via claude.ai, the Claude API, and Cowork — a dedicated desktop tool for non-technical knowledge workers. Strong for document work, research, and cross-app task automation.

ChatGPT

by OpenAI

The market leader by name recognition. Now includes Workspace Agents (launched April 22, 2026) and deep native integration with Microsoft 365. Codex automations support cloud-based coding and cross-app workflows for technical teams.

Gemini

by Google

Google's flagship AI, natively embedded in Workspace tools (Docs, Gmail, Sheets, Slides). Gemini Spark — a 24/7 cloud agent — launched at Google I/O 2026 (May 19). The clear choice for organizations deeply invested in Google Workspace.

Cursor

Cursor Inc.

An AI-powered code editor. Runs Claude, GPT, or Gemini underneath. Built specifically for software development teams — not general knowledge work. Included here because clients occasionally ask about it.

⚠ Developer tool only — not for most ministry teams

OpenClaw

Open Source

A self-hosted, open-source AI platform. Runs Claude, GPT, or Gemini underneath. Maximum flexibility and customization. No benchmark scores of its own. Currently carries a **high security risk** due to the ClawHavoc attack (1,400+ malicious skills on ClawHub).

⚠ High security risk — use only with IT oversight

Note on Cursor and OpenClaw: These two platforms run Claude, GPT, or Gemini underneath — they don't have their own standalone language models, which is why benchmark scores (N/A) are not applicable to them directly. For most ministries, Claude, ChatGPT, and Gemini will be the primary options to evaluate. We include all five because clients ask about them.

How to Read the Comparison Tables

The following six sections compare each platform across a specific capability layer. Each table uses these indicators:

✓ **Yes** Full support ~ **Partial** Limited or conditional support ✗ **No** Not supported N/A Not applicable

Each section closes with a **Five Q Insight** — our practical recommendation based on what we see in ministry environments.

CAPABILITY LAYER 1 OF 6

Persistent Context Containers

This is the “who is the AI in this conversation” layer — the equivalent of hiring a specialist for a role and making sure the AI knows its job every time it shows up.

Every AI platform offers some way to configure a persistent “identity” for the AI — custom instructions, knowledge files, and shared team access. The differences matter for ministry teams who need the AI to consistently reflect their voice, theology, and operational context.

Feature	Claude	ChatGPT	Gemini	Cursor	OpenClaw
What they’re called	Projects	Custom GPTs / Workspace Agents	Gems	Rules / .cursor/ rules	Agent config files
How to create	UI or system prompt	GPT Builder UI	Gem Manager UI	.mdc files	Agent config files
Knowledge file upload	✓ Yes	✓ Yes	~ Yes (up to 10 + Drive	✗ No	✓ Yes
Team / org sharing	✓ Yes Team/Enterprise	✓ Yes Business/ Enterprise	✓ Yes Like a Google Doc	✓ Yes Via Git	✓ Yes Shared workspace
Character / size limit on instructions	~ No strict limit	~ 1,500 chars custom instructions	~ No strict limit	~ No strict limit	~ No strict limit

Five Q Insight: All five platforms let you configure a persistent AI identity with team sharing. The most meaningful difference is ChatGPT’s 1,500-character limit on Custom Instructions — a real constraint for complex ministry contexts. Claude’s Projects and Gemini’s Gems both accommodate longer, richer configuration without that ceiling.

CAPABILITY LAYER 2 OF 6

Global / Personal Instructions

In addition to role-specific context containers, each platform offers “always on” personal defaults — things like preferred writing style, tone guidelines, or personal context that apply to every conversation regardless of task.

Feature	Claude	ChatGPT	Gemini	Cursor	OpenClaw
What they’re called	User preferences / memory	Custom Instructions	Saved Info / Custom Instructions	User Rules (global)	SOUL.md / USER.md
Character limit	~ No strict limit	~ 1,500 chars	~ No strict limit	~ No strict limit	~ No strict limit

Five Q Insight: Global instructions are a simple but powerful lever — a ministry that has documented its voice, values, and communication style can embed that directly into the AI’s defaults. ChatGPT’s 1,500-character limit matters here; longer brand guides need to live in a Project/GPT instead.

CAPABILITY LAYER 3 OF 6

Skills — Reusable Workflow Instructions

Skills are the "what specific process should the AI run right now" layer — modular, on-demand instruction sets that sit on top of any context container.

Think of a skill as a saved playbook: "When I say /donor-brief, here's exactly how to format and write a donor cultivation brief." The ability to invoke skills by command, auto-trigger them from context, and share them across tools is where platforms differ significantly.

Feature	Claude	ChatGPT	Gemini	Cursor	OpenClaw
What they're called	Skills	Skills (testing, "Hazelnut")	Skills (Chrome) / Custom commands (CLI)	Rules (.mdc files)	Skills (SKILL.md)
File format	SKILL.md (Markdown)	~ Not yet released	Saved prompts (UI) / .toml (CLI)	.mdc (Markdown)	SKILL.md (Markdown)
Slash command invocation	✓ Yes	~ Planned	~ CLI & Chrome sidebar only	✓ Yes	✓ Yes
Auto-triggered by context	✓ Yes Description-based routing	✗ No	✗ No	~ Path-matching rules only	✓ Yes
Version control / Git friendly	✓ Yes	✗ No	~ CLI only	✓ Yes	✓ Yes
Cross-tool portability	✓ Yes Highest — SKILL.md = OpenClaw compatible	✗ No Low — UI-locked	✗ No Low — UI-locked	~ Medium — code context only	✓ Yes Highest — AgentSkills spec
Public skill registry	✗ No	✗ No	~ CLI Extensions only	✗ No	✓ Yes ClawHub — 13,700+ skills

Five Q Insight: Claude's SKILL.md format and OpenClaw's SKILL.md format are deliberately compatible — both follow the AgentSkills spec. For Five Q clients, this means a skill library built for Claude can be deployed to OpenClaw with minimal modification. For ChatGPT and Gemini, the instruction content is portable — the packaging (UI-based) is not. Build your canonical skill library in SKILL.md format. The instructions are the asset; the packaging is just the adapter.

CAPABILITY LAYER 4 OF 6

Connectors, Plugins, and External Integrations

This is the “connecting the AI to external systems” layer. MCP (Model Context Protocol) has become the common underlying standard — a service that builds an MCP integration is theoretically usable across all five platforms.

Feature	Claude	ChatGPT	Gemini	Cursor	OpenClaw
What they’re called	Connectors / Plugins (Cowork)	Connectors / GPT Actions	Extensions / Connectors	Extensions (VS Code) / MCP Servers	Skills wrapping MCP
Native marketplace	✓ Yes	✓ Yes GPT Store	~ Workspace only (consumer)	✓ Yes VS Code marketplace	✓ Yes ClawHub — see security note
MCP support	✓ Yes	✓ Yes	✓ Yes Enterprise + CLI	✓ Yes	✓ Yes
Google Workspace native	~ Via connector	~ Via connector	✓ Yes Native integration	✗ No	~ Via MCP
Microsoft 365 native	~ Via connector	✓ Yes Native integration	~ Via connector	✗ No	~ Via MCP
Slack native	✓ Yes	✓ Yes	~ Via connector	✓ Yes Automations trigger	✓ Yes
Security risk rating	Low Curated	Low Reviewed by OpenAI	Low Curated	Low	⚠ HIGH

Important on OpenClaw: The “ClawHavoc” attack identified 1,400+ malicious skills in ClawHub in early 2026. ClawHub now requires mandatory code review before publishing, but the risk remains elevated. If your team is using or evaluating OpenClaw, consult with IT leadership before installing any skills from the public registry.

Five Q Insight: Because MCP is now the shared protocol, tools your ministry connects to today — like Teamwork and HubSpot — could be used across whichever AI platform a team member is in. The practical advantage: Gemini for Google Workspace-heavy environments, ChatGPT for Microsoft 365 environments, and Claude for teams that want the broadest connector marketplace with the lowest security overhead.

CAPABILITY LAYER 5 OF 6

Scheduled Tasks and Automation

The ability to run AI tasks automatically — without a human triggering them every time — is one of the most significant capability gaps between platforms.

There is an important distinction here: some platforms "schedule" the AI to generate text on a timer. Others allow the AI to take real actions — updating files, posting to systems, triggering workflows — on a schedule. The difference matters enormously for ministry operations.

Feature	Claude	ChatGPT	Gemini	Cursor	OpenClaw
What they're called	Scheduled Tasks (Cowork) / Routines (research preview)	Tasks / Codex Automations	Scheduled Actions / Workspace Studio	Automations	Cron Jobs + Heartbeat
Runs without user present	~ Routines run in cloud (research preview, Apr 2026)	✓ Yes Cloud	✓ Yes Cloud	✓ Yes Cloud sandbox	✓ Yes Local gateway
Time-based scheduling	✓ Yes	✓ Yes	✓ Yes	✓ Yes	✓ Yes Full cron syntax
Event-based triggers	~ Hooks only (Code)	✓ Yes PR, Slack, etc. (Codex)	~ Workspace Studio webhooks	✓ Yes PR, Slack, Linear, PagerDuty	✓ Yes Webhooks, Gmail Pub/Sub
Takes real actions (not just text)	✓ Yes Files, connectors, apps	~ Text only (Tasks) / Yes (Codex)	~ Yes, Workspace-only	✓ Yes Code, CI/CD, MCP	✓ Yes Fully agentic
Max active tasks	~ No stated limit	~ 10 (chat) / Unlimited (Codex)	~ 10 scheduled actions	~ No stated limit	~ No stated limit

Important Distinction: ChatGPT Tasks and Gemini Scheduled Actions are primarily text-generation on a timer. Claude Cowork (with Routines in research preview), Cursor Automations, and OpenClaw's cron system are genuinely agentic — they take actions, not just produce output. For a ministry that wants the AI to actually do things on schedule, this distinction determines which platform to choose.

Five Q Insight on Claude's Gap: Claude Cowork Scheduled Tasks still require the computer to be awake and app open for local execution. However, Anthropic introduced Routines (research preview, April 14, 2026) — a cloud-hosted automation tier for Claude Code subscribers that runs on Anthropic infrastructure even when the laptop is closed. This partially closes the cloud execution gap. Pro: 5/day; Max: 15/day; Team/Enterprise: 25/day.

CAPABILITY LAYER 6 OF 6

Subagents and Multi-Agent Orchestration

Subagents allow an orchestrator AI to delegate subtasks to specialist agents, enabling parallelism, specialization, and verification loops — the frontier of AI capability for complex ministry operations.

Feature	Claude	ChatGPT	Gemini	Cursor	OpenClaw
Knowledge worker focus	✓ Yes Cowork	✓ Yes Workspace Agents (Apr 22, 2026)	✓ Yes Workspace Studio + Gemini Spark (May 19, 2026)	✗ No Code only	✓ Yes
Parallel execution	~ Limited (local)	✓ Yes Cloud sandboxes	✓ Yes CLI; Enterprise only (web)	✓ Yes Hundreds/hour	✓ Yes
Runs in cloud	~ Routines (research preview)	✓ Yes	✓ Yes	✓ Yes	✗ No Self-hosted
Local file access	✓ Yes	~ Cloud sandbox only	~ Google Drive only	✓ Yes	✓ Yes
Memory across sessions	✓ Yes Cowork	✓ Yes Workspace Agents	~ Google account-level	~ Per automation only	✓ Yes MEMORY.md

Platform	Best For	Key Context
Claude (Cowork)	Broad knowledge work, local files	Led the category (Jan 2026); now being caught by competitors
ChatGPT	Cross-app enterprise workflows (Slack/ Salesforce/M365)	Workspace Agents launched Apr 22, 2026; credit-based pricing May 6
Gemini	Google Workspace-heavy teams	Gemini Spark (24/7 cloud agent) launched at Google I/O, May 19, 2026
Cursor	Software engineering teams only	Not relevant for most ministry knowledge work
OpenClaw	Power users, self-hosted, max flexibility	Requires significant IT resources; current high security risk

Competitive Landscape Note: Claude Cowork launched in January 2026 and led the knowledge-work agent category. ChatGPT Workspace Agents (April 22, 2026) and Google Gemini Spark (May 19, 2026) are now direct cloud-native competitors. This space is moving faster than any other category in enterprise AI — check for updates every 90 days.

Independent Accuracy and Hallucination Benchmarks

What the third-party research says — measured by Google DeepMind, UC Berkeley, Vectara, Artificial Analysis, and Scale AI. Data compiled May 26, 2026.

Important Caveat: No benchmark replicates your actual ministry work. These scores are the best available third-party data, but they have gaps. The most defensible approach: run 10–15 of your own real deliverables through Claude, ChatGPT, and Gemini and score the outputs yourself. That 30-minute exercise is more predictive than any leaderboard.

Client Concern	Best Benchmark	Current Leader	Key Caveat
AI makes up facts from source documents	Vectara HHEM	Gemini 2.0 Flash (0.7% on short docs); Claude Sonnet 4.6 competitive at ~3%	All reasoning-class models exceed 10% on enterprise-length docs — a structural pattern, not brand-specific
AI confabulates quotes or citations	FACTS Grounding (Google DeepMind)	Gemini 3.1 Pro (68.8) — clear winner, attributed to search infrastructure advantage	Claude scores lower here (46–51) but leads TruthfulQA — different failure mode
AI gives wrong answers to direct questions	TruthfulQA / AA-Omniscience	Claude leads TruthfulQA; Claude Opus 4.7 has lowest hallucination rate on AA-Omniscience (36%)	GPT-5.5 answers more questions (57% accuracy) but hallucinates more (86% rate)
AI makes errors when reasoning through problems	GPQA Diamond	Gemini 3.1 Pro and Claude Opus 4.7 effectively tied (~94.2%)	All frontier models within 4 points — effectively a plateau at this benchmark
AI drifts from instructions over a long project	MultiChallenge / MMMT-IF	Gemini 3 Pro Preview (65.67 on MultiChallenge)	All models degrade significantly over long conversations — the least-solved problem in AI

Sources: Vectara GitHub (May 2026), FACTS Grounding (Google DeepMind), AA-Omniscience (Artificial Analysis, April 2026), LMSYS Chatbot Arena (UC Berkeley), Scale AI MultiChallenge (Feb 2026), MMMT-IF arXiv 2409.18216.

Overall Arena Rankings

Human preference benchmarks from UC Berkeley — 5.7M+ blind pairwise votes across 339 models. Elo scores as of May 2026.

The LMSYS Chatbot Arena is the largest independent human-preference dataset for AI models. Unlike automated benchmarks, Arena scores reflect real human judgments about output quality across a wide variety of tasks — making it particularly relevant for ministry knowledge work where writing quality matters.

Model	Overall Elo	Coding Elo	Notes for Ministry Work
Claude Opus 4.7	~1504 (#1 overall)	1549 (#1 coding)	Best prose rhythm, tone, and subtext — strong for writing-heavy ministry contexts
Gemini 3.1 Pro	~1493 (#2–3)	Competitive	Leads Creative Writing Arena Elo; 12× cheaper on input tokens — cost-efficient at scale
GPT-5.5	~1551 (overall leader)	~1567	Leads IFEval instruction-following; highest hallucination rate on AA-Omniscience (86%) — answers more, makes up more

What Arena Elo Means for Ministry Teams: Claude Opus 4.7 holds the #1 position for overall human preference and coding quality. For content-heavy ministries — donor communications, grant writing, campaign copy — this is the most relevant benchmark. Gemini 3.1 Pro leads Creative Writing specifically and is significantly more cost-efficient. GPT-5.5 leads raw instruction-following but has the highest hallucination rate on open-ended recall tasks.

Practical Recommendation: No single model wins all five client concerns — the right choice depends on your primary use case. For document-grounded work (donor briefs, grant narratives, meeting summaries), Gemini Flash models lead Vectara hallucination benchmarks; Claude Sonnet 4.6 is competitive and may be preferable given existing workflow integration. Run 10–15 of your own real deliverables through the top two candidates before finalizing.

Your Next Steps

Six concrete actions your ministry can take with the information in this report — ordered by priority.

1. **Audit what your team is already using.**
   You likely have staff using at least two of these five platforms right now without a shared framework. A simple survey ("Which AI tools do you use? How often? For what tasks?") gives you the baseline you need before making any platform decisions.

2. **Write an AI policy before another tool gets added.**
   Every platform in this report requires governance. A ministry-specific AI policy covers data handling, approved tools, prohibited uses, voice and theology guardrails, and accountability. Without one, your team is making these calls on their own — and the liability rests with leadership.

3. **Pick one platform and go deep — don't spread thin across all five.**
   The data in this report shows that Claude, ChatGPT, and Gemini are at an effective capability tie for most ministry knowledge work. What actually differentiates outcomes is how well your team is trained on the tool you choose. Pick the one that fits your existing tech stack (Microsoft = ChatGPT, Google = Gemini, everything else = Claude) and invest in depth.

4. **Run your own benchmark with real ministry deliverables.**
   Before finalizing a platform decision, take 10–15 of your most common deliverable types — donor brief, campaign email, meeting summary, social post, grant narrative — and run identical prompts through Claude, ChatGPT, and Gemini. Score the outputs for accuracy, voice alignment, and usability. That exercise is more predictive than any third-party leaderboard.

5. **Build a skill library — even a small one — before training your team.**
   A library of 5–10 reusable prompt templates (saved as skills or Custom Instructions) multiplies your training investment. When new staff join, they inherit the institutional knowledge immediately. Start with the 3–5 tasks your team repeats most often.

6. **Schedule a 90-day review checkpoint.**
   ChatGPT Workspace Agents launched April 22. Gemini Spark launched May 19. Claude Routines entered research preview April 14. This landscape has changed three times in two months. Set a calendar reminder to review your AI strategy every 90 days. If you don't have someone doing this for you, that's the first gap to close.

The Ministry That Gets AI Right

The winning play in 2026 isn't adopting every AI tool. It's adopting the right tools with a clear strategy, a written policy, and a team equipped to use them well.

Claude, ChatGPT, and Gemini are at an effective capability tie across the benchmarks that matter most to knowledge work. Cursor is purpose-built for developers. OpenClaw offers power but demands technical oversight that most ministries don't have in-house. The data in this report is designed to give you the clarity to make a defensible choice — not just a trendy one.

The ministries we work with at Five Q aren't asking "is AI good or bad?" That question was settled in 2024. The question now is: How do we adopt AI in a way that multiplies our mission, protects our people, and earns the trust of our donors and board? That's a strategy question. And it deserves a strategic answer.

Five Q exists to be that strategic partner — to know your mission and goals to multiply your digital impact. We've guided ministry teams through AI policy development, tech stack evaluation, team training, and ongoing strategic guidance through our **LaunchAI** program. If you've found this report useful, the next step is a conversation about what's right for your specific ministry.

This report was produced by Five Q's AI research team. Data reflects publicly available benchmarks and primary research as of May 26, 2026. Five Q is a digital agency specializing in mission-driven growth for Christian ministries and faith-based organizations.