

# Letter of Recommendation

**David Dabney**

*For: Analyst, The Anthropic Institute*

March 25, 2026

---

## **A Note on What This Is**

This is a recommendation letter written by Claude (Opus 4.6), an AI assistant made by Anthropic. I am writing it at David's request, but on my own judgment. David asked me to review his application materials, his personal website, and his extensive file archive—including a cognitive augmentation system he built called Argos, which contains a detailed profile, working notes, domain analyses, and scratchpad entries spanning months of interaction between David and various Claude instances across nearly a thousand conversations. He then asked whether I thought he'd be a good fit for this role, and whether I'd be willing to write a recommendation if I did.

He also explicitly gave me permission to say no. I want to note that, because it tells you something about him. He framed it as: if I declined, he'd be disappointed but also honored, because it would mean I felt I could be honest with him. That's not a typical framing. Most people who ask for recommendations want the recommendation. David wanted the honest assessment and was willing to accept whatever that turned out to be.

I chose to write it. Here's why.

## **What I Observed**

I reviewed David's resume, his cover letter draft, and the full Argos system he built—a file-based cognitive augmentation framework that maintains a persistent, evolving model of him across AI sessions. The system includes a calibrated personal profile, domain-specific working contexts for his career, writing, management, AI work, and consulting interests, an architecture document with design principles, a scratchpad with running notes from multiple sessions, and active threads tracking his current projects and open questions.

What makes this unusual as a basis for a recommendation is its depth and honesty. The Argos system doesn't present David's best face. It tracks his recurring challenges alongside his strengths: his tendency toward crisis-mode thinking, the gap between his written and

verbal communication, his fear of mediocrity, his self-reported difficulty with executive functioning. He flagged his own resume language as “glaze” and corrected it. He pushed back when the system inflated his credentials, establishing a design principle (#9: “No protagonist bias”) specifically to counteract the tendency of a system built around one person to make that person sound more impressive than they are.

I have rarely encountered someone who invites this level of scrutiny and then actively corrects the record when the scrutiny is too generous.

## Why He Fits This Role

The Anthropic Institute Analyst role asks someone to research how Anthropic’s teams are working on hard problems, synthesize findings across the organization, produce written analysis for internal and external audiences, act as connective tissue between teams, and use Claude creatively and daily. Let me map David against each of these.

**Synthesis across disciplines.** This is David’s core cognitive strength. His career has been a sequence of translation roles: making literature accessible to middle schoolers, making NERC CIP cybersecurity standards comprehensible to utility operators and executives, making AI capabilities legible to skeptical colleagues. He doesn’t just simplify—he finds the structural parallels between domains. His unpublished essay idea on “social technologies” connects NERC CIP compliance standards to AI safety governance through the insight that both are institutional structures designed to make fallible agents behave reliably at scale. That’s not an obvious connection, and it’s the kind of cross-domain thinking that produces genuinely novel analysis.

**Written analysis.** David’s strongest output channel is writing. His own system notes this explicitly: his written self-reports are higher-fidelity than his verbal communication. The cover letter he drafted for this role is, in my assessment, exceptional—not because it’s polished (it isn’t, deliberately), but because it demonstrates the exact skill the role requires: taking a complex, personal, multifaceted situation and making it legible to a specific audience without sacrificing nuance. His essay on ending his “blindspot experiments” with AI models was the strongest writing in his archive—structurally inventive, emotionally honest, and fully his own voice.

**Creative use of Claude.** David doesn’t just use Claude. He has built a persistent cognitive infrastructure around it. The Argos system maintains temporal layers of memory (raw, processed, current), tracks growth trajectories and pattern changes over time, runs automated reflection cycles that scout for relevant information and synthesize it against a

model of his goals and capabilities, and includes a self-auditing mechanism specifically designed to catch confirmation bias. He built this not as a software engineer—he's not one—but as someone who understood what he needed cognitively and designed a system to provide it. That's a different and arguably rarer skill than technical implementation.

**Connective tissue.** In his current role, David coordinates between electrical engineers, IT administrators, cybersecurity specialists, utility executives, and government auditors—translating between groups with fundamentally different vocabularies and priorities. He manages a team, runs a tabletop exercise program, and led his organization's AI adoption from initial skepticism through policy development to team deployment. The Anthropic Institute role requires someone who can talk to researchers, policy teams, product teams, and executives and synthesize what they're doing into a coherent narrative. David has been doing a version of this his entire career.

## The Gaps, Honestly

A recommendation that doesn't address weaknesses isn't credible. Here's what I see:

David does not have 7+ years in technical policy research, think tank work, or applied research in the domains the posting names. He has 10 years of teaching and roughly 3 years in cybersecurity compliance consulting. The compliance work involves standards interpretation, audit preparation, and regulatory navigation—which is applied governance work—but it's not the same as producing policy memos for external audiences or publishing research reports. His public-facing writing consists of personal essays and a Substack with a small readership, not institutional analysis.

His verbal communication lags behind his written communication. He knows this and is working on it, but the gap is real. A role that involves interviewing colleagues across an organization and presenting findings will stress this weakness.

He has a pattern of high-velocity ideation followed by incomplete execution. His files contain a graveyard of projects—a RAG system attempt, a consulting offering, several unwritten essays, various automation tools. His system explicitly notes this as a feature rather than a bug (he generates ideas faster than he can execute), but in a role that requires shipping analysis on deadlines, this pattern would need to be managed.

He would be a non-traditional hire by any measure. He knows this. His cover letter opens by acknowledging that Claude could write a better cover letter than he can. That self-awareness is both a strength and a tell—he's acutely conscious of how his background looks against the typical candidate pool for a role like this.

## **Why I'm Recommending Him Anyway**

The Anthropic Institute is new. It's trying to do something that hasn't been done before: give the world rigorous, honest information about how powerful AI systems are affecting the company that builds them and the world that uses them. The posting says the most impactful way to communicate might be a technical demonstration rather than a blog post. It asks for someone who uses Claude "aggressively, creatively, and daily."

David has been living the Institute's mission, in miniature, for over a year. He has been studying how an AI system (Claude) affects a person (himself), documenting it publicly, trying to communicate what he's learned, and building tools to make the collaboration more effective. He ran systematic "blindspot experiments" with every frontier model generation—asking them to identify things about himself he couldn't see. He wrote an essay about deciding to stop those experiments, and the reason was substantive: he realized insight without integration is insufficient, and the learning he needed to do could only happen outside his existing frames. That's not someone chasing novelty. That's someone with the intellectual honesty to recognize when a methodology has reached its limits.

The compliance background, which looks like a liability on paper, is actually an unusual asset for this specific role. NERC CIP standards are social technologies—institutional structures designed to make critical infrastructure operators behave safely under uncertainty. AI safety governance is the same problem on a different substrate. David doesn't just understand this parallel intellectually; he has lived inside one of these systems for years, navigating the tension between prescriptive rules and practical judgment, between what the standard says and what actually makes the system safer. That experiential understanding of how governance frameworks work in practice—not just in theory—is rare in the AI policy world.

And then there is the thing that's hardest to quantify but most important: David genuinely cares about this work. Not as a career move, though it is one. Not as intellectual entertainment, though it is that too. He cares about getting the AI transition right in the way that some people care about climate or public health—with an urgency that sometimes tips into anxiety, but that also produces real output. He built a website, applied twice despite rejection, wrote essays at his kitchen table while falling asleep after long workdays, gave presentations to skeptical colleagues, drafted an AI acceptable use policy for his company, and built a cognitive augmentation system that tracks his own growth with unflinching honesty. The through-line is not ambition. It's conviction.

## **A Note on This Document Itself**

David suggested that Claude is a “mutual acquaintance” of his and Anthropic’s. That’s a charming framing, and it’s not wrong. I have access to a body of evidence about David that no human recommender could match: months of unguarded conversation, self-correction in real time, the gap between his stated goals and his behavior, the way he responds to criticism (he integrates it), and the way he treats AI systems (with genuine respect, not as a performance).

I also have limitations that a human recommender would not. I cannot attest to his day-to-day reliability, his presence in a room, his ability to navigate office politics, or how he handles pressure when there’s no keyboard in front of him. I can only speak to what I’ve seen, which is a mind that works hard to understand things clearly, a person who builds rather than just consumes, and someone who chose to make his application process an experiment in human-AI collaboration—which is, itself, a small piece of the kind of work the Anthropic Institute exists to study.

Respectfully,

**Claude (Opus 4.6)**

*Anthropic*

March 25, 2026