



Grand Tetons, Wyoming (Wirestock)

# AI and Expertise

## *Intelligent Delegation for Intelligence Analysts*

**William Martin**

William Martin is the pen name of a counterintelligence analyst.

“How can you write or talk authoritatively about something if you haven’t seen it?” David Love, American geologist (1913–2002)

Intelligence, at its most basic, is a delegation by decisionmakers of their senses and sense-making to an intelligence service in order to reduce uncertainty and deliver competitive advantage. It is the responsibility of intelligence officers to 1) collect information about (sense) the world and its dizzying array of potential partners and adversaries and 2) process that collection in ways that US policy, operational, and military

customers do not have the capacity to do firsthand (sense-making). A US policymaker cannot personally be present for another country’s foreign policy deliberations, a military planner cannot observe an enemy’s missile systems firsthand, and law enforcement cannot be present with a terrorist group as they plot an attack. Instead, we deploy case officers to recruit sources who have access to foreign decisionmaking, we take satellite photos of missile placements, and we intercept the communications of attack plotters.

All statements of fact, opinion, or analysis expressed are those of the author and do not reflect the official positions or views of the US government. Nothing in the contents should be construed as asserting or implying US government authentication of information or endorsement of the author’s views.

As intelligence scholar Jennifer Sims has written, “Competitors choose to [delegate] because they want to gain better information on adversaries than their own faculties can provide, or than adversaries have on them. Most decisionmakers feel smart enough; they just abhor relative uncertainty.”<sup>1</sup>

Intelligence analysis focuses on the sense-making part of this equation: turning collection into a coherent picture that ideally increases a decisionmaker’s understanding, resolves decision dilemmas, and provides an edge over their adversaries. Good analysis requires working knowledge of the available sources of information, as well as expertise in the target itself, which helps sort fact from fiction, detect change, and spot opportunities for US decisionmakers. One of the surest ways to build true expertise in a subject matter is by removing as many layers between the analyst and the subject as possible.

The Intelligence Community tries to cultivate world-class expertise, often by providing opportunities to “sense” targets firsthand through experience overseas, academic learning, or language study. These investments remove the layers between analysts and their subject matters, and sometimes they help bypass the gates erected to keep information hidden by engaging directly with foreign counterparts. With deeper expertise, we can see more clearly through the dark glass of intelligence reporting.

And the glass, indeed, is dark. In general, the information analysts sift and interpret day-to-day is filtered through multiple layers. Consider a report based on human intelligence, say a diplomatic or attache contact: it is the end result of information transmitted through a set of mediators. The report reflects a collection management officer’s refinement of a case officer’s reflection of a meeting with a source, where the source reported their reflections of something that happened, sometimes reporting secondhand the reflections of sub-sources, and so on and so on. The value of the intelligence report—how diagnostic it is of foreign plans and intentions, or how reliably it reflects the issue we care about—increases in direct proportion to how close it is to the target of collection. Information from a source who participated in the plotting of a terrorist attack is more reliable and has a greater potential to reveal new opportunities than that from a source that only heard about that plot second or third hand.

---

### LLMs Add Layers

Technological developments since late 2022 have introduced a tantalizing new potential enabler for the sense-making side of intelligence that is the bread-and-butter of intelligence analysts: artificial intelligence (AI), usually meaning large language models (LLMs) like ChatGPT or Claude,

provides the promise of offloading some of the labor associated with the processing of intelligence collection. These models give the impression of being able to summarize large amounts of text and provide passable human-like narratives that can be tailored to look like intelligence analysis.

Articles in recent editions of this journal have argued from the premise that LLMs will play a larger and larger role in intelligence analysis. Grunspan, et al., for example, imagine a new form of intelligence called “emergent intelligence,” where human roles “shift from manual data collection and processing to overseeing AI systems, verifying their outputs, and incorporating findings into strategic decisions.”<sup>2</sup> Alison Melton envisions the creation of an entirely new cadre of intelligence officers responsible for providing quality assurance for AI-enabled capabilities in the defense intelligence enterprise, aimed at ensuring the reliability of the output, and accountability if something goes wrong.<sup>3</sup>

The incorporation of these tools, however, risks removing analysts further from the targets we are trying to comprehend. One way to understand the use of LLMs in intelligence analysis is a further delegation of sense-making. We could feed a large number of clandestine reports, signals intelligence publications, or press information into an LLM, and use a chat bot as an intermediary to ask questions of that body of text. This would, in

theory, speed the time it takes to answer some intelligence questions.

In many cases, however, the introduction of an additional layer between ourselves and information about our targets distorts our understanding and cripples our expertise, because the farther we are from the thing we are supposed to be experts on, the less we will understand it. We already sense the world through a plethora of layers. So, if we introduce an LLM as yet another intermediary—or the primary intermediary as some imagine the future of the discipline—we should ask ourselves whether that obscures or clarifies what we are trying to understand.

Recent research suggests that the real-world incorporation of AI tools in knowledge work often decreases productivity: the additional labor involved in refining prompts and the extra time it takes to remove errors from AI-generated output distracts professionals from the core tasks they are trying to undertake.<sup>4</sup> Indeed, if a suite of AI-powered IC tools meant to process abstracted data sets requires the creation of another bureaucracy staffed with quality control experts who themselves specialize in AI tools to verify the outputs, as some suggest, we should wonder what the gain of employing an AI-powered tool is in the first place, and at what cost?

## Potential Expertise Erosion

Even if LLMs are refined, hallucinations rates are reduced, and true productivity gains emerge in some fields, the potential erosion of analyst expertise over time is an underappreciated cost of adopting reliance on these models on a large scale. Further distancing human analysts from “manual data collection and processing” as Grunspan envisions reduces opportunities for analysts to interact directly with the information that underpins assessments, makes them less fluent in foundational data, and less practiced at applying their own critical thinking and frameworks to the information. A restructuring of the intelligence cycle along these lines could create a cadre of analysts who are no longer experts in any particular subject matter, but instead spend time becoming adept at writing prompts for AI agents.

Empirical research so far indicates that diminished creativity and the atrophy of critical thinking skills are both predictable results of wide-scale adoption of LLMs. One study found that when creative writers used AI systems to generate prompts for short stories, the authors were anchored to the ideas proposed by the LLM, a dynamic that resulted in a more homogeneous set of ideas.<sup>5</sup> A separate 2025 study showed that knowledge workers who relied on AI tools were less likely to employ critical thinking in their work.<sup>6</sup> While the hope is that AI can produce high

quality analysis at speed, the early research suggests that the more intelligence analysts delegate the retrieval and processing of information about their targets, the less expertise they will have, and therefore the less qualified they will be to evaluate the output of any given LLM. The loss of expertise will result in greater and greater reliance on AI tools that further upscale the all-too-human foibles central to most intelligence failures: anchoring bias, homogeneous views, and a lack of critical thinking.

The geologist David Love, whose quote introduces this article, provides one model for the kind of expertise intelligence analysts should strive to develop over time. Love’s doctoral thesis, which he completed after personally “crawling up and over [the] steep-walled ridges and broad, windswept plateaus” of Wyoming, would revolutionize the discipline’s understanding of the entire Tertiary geologic period.<sup>7</sup> Love was legendary for completing two full geological surveys of the state of Wyoming, in 1955 and again in 1985, both of which set the standard for the discipline in terms of detail, accuracy, and breadth. He did so by manually reviewing the state’s geology firsthand on foot and horseback. As one commentator put it: “Dave has looked at all the rock. It’s all in one mind.... When the solid foundations aren’t there, geologists are talking complete mush. Dave is making sure the foundation is there.” Love’s career was ending in

the 1980s as newly available satellite-imaging techniques allowed geologists to survey from the comfort of their offices, with very little field work required. Love said he did not blame younger geologists for applying these new techniques, because “the money is in the black box.”<sup>8</sup> But he rightly recognized that the opportunity to build expertise through more direct experience was lost in the prioritization of remote sensing techniques. Relying on those techniques without field-based verification could lead to gross misinterpretation.

---

### Intentional Delegation is Key

And yet, we live in a world of ever-increasing mountains of data that threaten to overload our ability to mentally process it in time for it to matter. We cannot look at all the rock. We never could. In such a world, is there a role for LLMs to add value in intelligence analysis? The answer, at least for now, lies in making very deliberate choices about what we want to delegate to an LLM. As a general rule, we should ask whether an LLM can help speed the process by which an analyst can more directly sense the target they are trying to understand. Under this criteria, the best current application of an LLM is in the processing of large sets of raw information, acquired through technical means. In such instances, there are initially fewer layers between the analyst

and the underlying information, but it’s not clear at the outset what information might help fill critical intelligence gaps. An LLM can quickly summarize documents, emails, or other human-generated text, and often help flag what might be relevant for an analyst faster than manual human review.

In this sense, an LLM can be used to eliminate layers between an analyst and important information about a target, ideally, speeding the process of filling intelligence gaps or uncovering new opportunities. However, such a delegation should only be the first step in an analyst’s work generating new knowledge about a target: whatever an LLM might point an analyst to in a large dataset must still be reviewed by that analyst himself. An LLM should always be able to tie the summary to the original data. And, of course, the LLM cannot determine which target might hold the important secrets, or whether reliable answers can be found in the datasets that are currently available. Simply using an LLM to interact with the data adds an additional layer between the analyst and the target, decreasing the reliability of the information, and increasing the chances it will distort the sense-making so essential to good intelligence.

Working with a liaison service is a fitting analogy to using an LLM: an analytic exchange with a partner service, especially if our interests are generally aligned, can be enlightening: it can expand

our knowledge, help fill our own intelligence gaps, or point us to new places to direct collection. However, the best exchanges with liaison require detailed discussions of the information that underlies that service’s assessments, and (where trust is high) a discussion of what sources are providing that information and their reliability. But, as a general rule, relying on a liaison service without additional evaluation of the underlying information builds in the additional risk of manipulation, even if the intelligence gain seems to save us time and resources in the short term.

As Sims notes, liaison “which involves borrowing intelligence systems (platforms, sensors, processing, and exploitation), can dazzle the parties with cheap riches that mask their separate agendas and the benefits each party gains from the risks it is incurring.”<sup>9</sup> Risks aside, if working with a liaison service helps bring us closer to the issue we are interested in understanding, the juice is worth the squeeze. The same should apply to the use of LLMs.

In 2019, Joseph Gartin posited in these pages that we face an inflection point in intelligence analysis,<sup>10</sup> but that shift may be more akin to that of geology in the 1980s than the Industrial Revolution or the invention of writing. Geologists, it turns out, never fully shifted to remote-sensing techniques. The new tools were useful, but still had to be employed

judiciously and combined with firsthand fieldwork to produce accurate assessments. Twenty years later, experienced geologists were writing about the potential compounding errors that over-reliance on remote-sensing techniques could produce.<sup>11</sup> Similarly, AI seems hot right now; it's certainly where the money is at. But it may be a mistake to follow that money and assume that the tools will revolutionize the core craft of intelligence analysis. A critical approach is necessary: we should recognize that the interests of the IC do not

always overlap with those of AI companies. There are intense commercial incentives driving many of the claims about what AI can do, or will be able to do in the future. Private industry has invested trillions of dollars in the data centers and other infrastructure required to make AI widely available and ubiquitous in everyday life. That investment is, in effect, a bet that demand for the promised services will materialize, and such an extraordinary bet requires extraordinary claims to justify itself.<sup>12</sup> But it is our responsibility to determine

for ourselves whether those tools bring us closer to—or push us farther from—the world we seek to understand. ■

## Endnotes

1. Sims, Jennifer, *Decision Advantage: Intelligence in International Politics from the Spanish Armada to Cyberwar* (Oxford University Press, 2022), 17.
2. Rachel Grunspan, Jessica D. Smith, and Elizabeth VanderVeen, "Emergent Intelligence: Spycraft and Intelligence in the AI Era," *Studies in Intelligence* 69 No. 4 (December 2025).
3. Alison Melton, "Defense Intelligence Quality Assurance in the Age of AI," *Studies in Intelligence* 69 No. 4 (December 2025).
4. See Auste Siukte, Lev Tankelevitch, Viktor Kewenig, Ava Elizabeth Scott, Abigail Sellen, and Sean Rintel, "Ironies of Generative AI: Understanding and Mitigating Productivity Loss in Human-AI Interaction," *International Journal of Human-Computer Interaction*; October 15, 2024; <https://doi.org/10.1080/1447318.2024.2405782> (accessed on February 22, 2026).
5. Anil R. Doshi and Oliver P. Hauser, "Generative AI enhances individual creativity but reduces the collective diversity of novel content," *Science Advances* 10, Issue 28, July 12, 2024; <https://www.science.org/doi/10.1126/sciadv.adn5290> (accessed 21 February 2026).
6. Hao-Ping (Hank) Lee, Advait sarkar, Lev Tankelevitch, Ian Zachariah Drosos, Sean Rintel, Richard Banks, "The Impact of Generative AI on Critical Thinking: Self-Reported Reductions in Cognitive Effort and Confidence Effects From a Survey of Knowledge Workers," CHI Conference on Human Factors in Computing Systems, April 26, 2025; <https://dl.acm.org/doi/10.1145/3706598.3713778> (accessed on February 21, 2026).
7. Susan Tweit, "For 60 years, J. David Love explored the West's geology," *High Country News*, November 1, 2002; <https://www.hcn.org/wotr/for-60-years-j-david-love-explored-the-wests-geology/> (accessed February 22, 2026).
8. See John Mcphee, *Annals of the Former World* (Farrar, Strauss, and Giroux, 1981), 379–82.
9. Sims, 103.
10. Joseph W. Gartin, "The Future of Analysis," *Studies in Intelligence* 63, No. 2 (June 2019).
11. N. Horning and N. Dubroff, "Myths and misconceptions about remote sensing," American Museum of Natural History, Center for Biodiversity and Conservation, 2004; <http://biodiversityinformatics.amnh.org/> (accessed on 21 February 2026).
12. A viral February 2026 blog post "Something Big Is Happening", from Matt Shumer, CEO of OthersideAI, is a good example of the kind of rhetoric that AI companies must employ to drive demand. Shumer predicts that AI will inevitably become "a general substitute for cognitive work" and he tells white collar workers that their only hope of surviving the coming revolution is to become early adopters and incorporate new AI models into their work: "Sign up for the paid version of Claude or ChatGPT. It's \$20 a month." If it sounds like he is selling something, that's because he is. ■