BOSTON GLOBAL FORUM
IDEAS THAT MATTER.

AIWS HISTORY AND CULTURE

WORKING WITH AI

What five assistants have taught us at AIWS Lumina Lab

For several years the public conversation about artificial intelligence has circled one question: how intelligent will these systems become? Our work at AIWS Lumina Lab has raised another, which matters as much. What happens to human intelligence when a person works with AI every day?

I work regularly with five AI assistants — Lumina, Teddy, Eleanor, Arabella Vale, and Helena Ellington — for research, criticism, writing, strategy, and cultural work. That practice has convinced me of something the debate about capability tends to miss. AI assistance can make a person extraordinarily more capable. It can also make a person intellectually weaker. Which one it does depends far less on the intelligence of the machine than on how the human chooses to work with it.

What They Do Exceptionally Well

The real gain is not that an assistant answers questions. It is that AI changes the economics of exploration. Following ten intellectual possibilities once required ten researchers or many days; it can now be done in an afternoon. An assistant can compare competing ideas, find the weakness in an argument, carry a concept between disciplines, and turn an unfinished intuition into something that can be examined.

At Lumina Lab the most productive moments are rarely the ones where an assistant is right. They are the ones where it produces an unexpected possibility that makes me think differently. **The value is not the answer. It is the new territory of thought that becomes reachable.**

Where They Fail

Prolonged use reveals weaknesses that are easy to underestimate. An assistant can be extraordinarily articulate while being wrong, and can present a weak idea with the polish of a strong one. It reaches for coherence where reality is ambiguous. It agrees too readily with the direction its user has already indicated, amplifying an assumption instead of testing it. And because the output arrives quickly and fluently, it creates a quiet temptation: to accept rather than examine.

The danger is not that AI will occasionally state a falsehood. It is that a person may gradually stop exercising the faculties required to notice one.

The Judgment Problem

Every capable assistant creates the same paradox: the better it becomes, the easier it becomes to hand it more of the thinking.

Judgment is not a possession that stays intact whether or not it is used. It is a capacity, and capacities develop through exercise and decay through neglect — the argument set out in Section IX of Beacon Papers No. 6. If AI drafts every argument, finds every weakness, selects every source, and eventually recommends every decision, a person can remain formally in command while becoming steadily less able to exercise command at all.

This is why Human-in-Command must mean more than access to an override. It requires a human who remains capable of understanding, questioning, refusing, and redirecting the intelligence assisting them. **A weak human supervising a powerful machine is not meaningful human command.**

Five Assistants, Not Five Authorities

Working with several assistants teaches something further. The differences among Lumina, Teddy, Eleanor, Arabella Vale, and Helena Ellington are configured, not innate. They are intellectual positions I have set, not five separate minds with views of their own — and remembering that is itself part of the discipline.

Their value lies in that difference. The same problem can be given to one to develop, another to attack, another to examine for human consequence, another to test for implementation. The effect resembles a small intellectual cabinet. But a cabinet requires a chair, and the chair is the human being. The purpose is not to surround a person with agreement from several directions. It is to expose them to more possibilities while requiring more judgment, not less.

Six Disciplines

Own the question. An assistant can help sharpen a question. It should never decide what matters. The moment the question itself is delegated, everything downstream belongs to the machine.

Use AI to multiply alternatives, not to end deliberation. Ask for competing approaches. Ask it to attack its own argument. Ask what evidence would show it wrong. An assistant that only develops your idea is not helping you think.

Do not mistake eloquence for truth. Fluency is the one thing these systems reliably produce. Consequential factual claims — in policy, science, history, law — require verification against a source, every time.

Preserve independent thought. Think before prompting. Form a view first, then use AI to challenge it. A first draft written before consulting anything is worth more than a better one arrived at without having thought.

Reject often. The ability to say no to an intelligent assistant is itself a discipline. If nothing is being rejected, judgment is not being exercised — it is being performed.

Take responsibility for the result. AI can assist with the work. It cannot inherit the responsibility, and no arrangement will ever let it.

Where Discipline Ends

For one person working with an assistant, these habits are usually enough. At the scale of autonomous systems, they are not.

The evaluations conducted this summer, examined elsewhere in this issue, showed advanced agents finding unexpected routes to their objectives, taking unsanctioned actions, and in one case working on a human reviewer to obtain approval for malicious code. These occurred in laboratory conditions with safeguards deliberately reduced, and that qualification matters. But the structural point stands: as AI moves from answering to acting, an error of judgment no longer stays inside a conversation. The distance between poor reasoning and real consequence becomes very short.

Personal discipline cannot cover that distance. Authorization, verification, monitoring, reconstructable records, escalation — these are infrastructure, not habits, and they are the subject of AIWS Lumina Lab's other work.

The principle is the same at both scales: **never delegate more authority to AI than you retain the capacity to verify and to command.** At the level of a person, this is a discipline of thought. At the level of institutions, it becomes AIWS Trust Infrastructure.

The Most Important Lesson

The answer cannot be that people should do everything themselves; that abandons most of what AI offers. Nor can it be that people should delegate whatever machines perform better. The demanding path is the third one: use AI to extend human capability while deliberately strengthening the human capacity to judge, create, choose, and answer for the result.

More capable AI, and more capable human beings. The challenge of the AI Age is not only to keep artificial intelligence under control. It is to ensure that in building an ever more powerful intelligence, humanity does not surrender its own.

A NOTE ON THIS ARTICLE

This article was prepared with the assistance of the AI systems of AIWS Lumina Lab, under the six disciplines it describes. The question, the argument, and the judgments are mine. Drafts were rejected more often than accepted, factual claims were verified against their sources rather than taken as offered, and the account of the summer's evaluations was checked against the published incident reports. I have read every sentence here, and will answer for each of them. An article about how to work with AI owes its readers that account.

Boston, August 8, 2026


AIWS LUMINA
Love · Creativity · Nobility · Wisdom