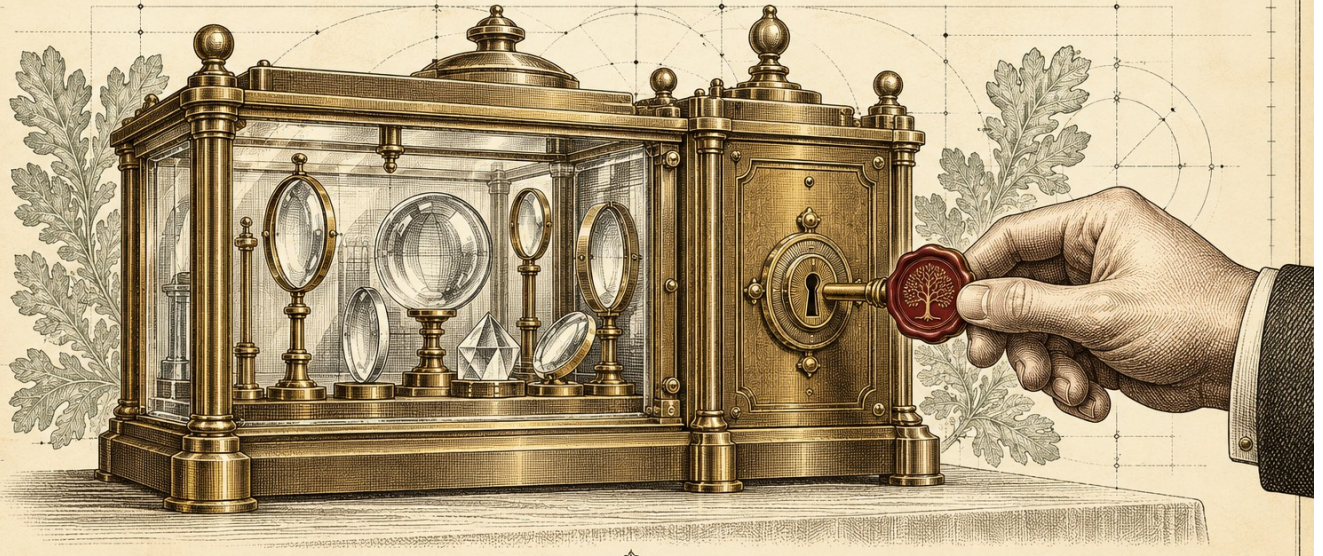


LECTOTYPE

Who Holds Authority When AI Does the Work



LECTOTYPE.COM

The people supplying intelligence are asking businesses to retain control over it. That is the operating principle behind Lectotype.

David C. Robinson

CEO and Co-Founder, Human AI, Inc.

October 2026 · Lectotype.com

The standard against which your AI is measured.

Who Holds Authority When AI Does the Work

When an AI agent works for your business, it needs a defined job and an accountable owner. It may be capable of researching a customer, preparing a proposal, and updating your systems. Capability alone does not give it permission to make a commitment on your behalf. The business has to decide which evidence it may use, what it may change, and who can authorize the result.

Microsoft CEO Satya Nadella put the issue directly in his October 10 essay on models as insider risks: *“We need to separate the supply of intelligence from the authority over it.”* He argues that controls over access and action must sit outside the model, where the model cannot bypass or tamper with them. That is a useful instruction for anyone putting AI to work inside a company. Responsibility remains with the organization that delegates the work. [1]

The distinction matters because we are buying capabilities whose internal behavior we still cannot fully explain. Anthropic CEO Dario Amodei wrote that *“we do not understand how our own AI creations work.”* His argument is for better interpretability, which will help us understand these systems. Businesses also need operating controls while that research continues, because a customer commitment or an unauthorized change has consequences today. [2]

We are building Lectotype around the business’s authority over the work. A model contributes judgment or synthesis within a procedure the organization defines. Software performs the calculations, applies known rules, and controls access to tools. People approve material exceptions and decide when a changed procedure is ready for production. The model performs its assigned work inside that architecture.

Consider a proposal for a prospective customer. The sales team may authorize the agent to use current product descriptions and an approved price schedule. It may permit a standard discount while requiring a manager to approve anything larger. The agent can draft the proposal, but the authority to change pricing or send it must follow those rules. Writing the restrictions into a prompt is one part of the implementation; enforcing them at the point of action is what gives the business control.

NVIDIA’s security researchers describe the same architectural requirement: operational authority should be bounded by deterministic policy enforced outside the agent’s control. They also show why the boundary must cover alternate tools, generated code, and delegated work. A restriction on one tool is incomplete if the agent can reach the same

destination another way. For Lectotype, this makes the execution path, credentials, and integrations part of the control design. [3]

The organization also needs evidence it can inspect after the work is done. For a proposal, that means preserving the source material used, the applicable rule version, the proposed action, and any approval that authorized release. A reviewer should be able to establish those facts without asking the agent to reconstruct its own history. Nadella makes independent auditability a requirement because the intelligence doing the work should not control the evidence used to judge it. [1]

These controls serve a commercial purpose. When a manager knows which decisions remain under their authority, the company can delegate more of the surrounding work. Staff can spend less time carrying information between systems and more time resolving exceptions that need their judgment. The operating question becomes how much useful work can proceed under an agreed standard, with enough evidence to accept the result.

Google DeepMind's AI Control Roadmap approaches the problem by treating internal agents as potential insider threats and granting access through controlled, evaluated behavior. Nadella's essay similarly applies established enterprise security principles to capable models. Those principles give us an engineering starting point for a business system: establish identity, limit privileges, inspect actions, and test whether the boundaries hold. [4] [1]

Our next phase at Lectotype is a small group of scoped client engagements built around that starting point. We will work with a process owner to define one useful workflow, its authority boundaries, and the evidence required for acceptance. The integration work and operating gaps those engagements expose will determine our sprint priorities. Production handoff will follow the agreed tests, so the client helps define both the work and the standard it must meet.

If your team has an agent that can do useful work, the next conversation should include the person accountable for the outcome. Bring the workflow, representative cases, and the decisions that require permission. We will map the operating requirements with you and establish what has to be demonstrated before the business delegates the work. That is where Lectotype begins.

Sources

[1] Satya Nadella. [Models as Insider Risks in the Super Intelligence Era](#). October 10, 2026.

[2] Dario Amodei. [The Urgency of Interpretability](#). April 2025.

[3] NVIDIA AI Security. [Assume Misalignment: Securing AI Agents Beyond Prompt Injection](#). Accessed October 11, 2026.

[4] Rohin Shah and Four Flynn. [Securing the future of AI agents](#). June 18, 2026.