

AI's Next Competitive Advantage: *User Autonomy*

Dr. Josh C. Simmons

The prevailing investment thesis in artificial intelligence centers on capability: which company can build the most intelligent model, acquire the most compute, or sustain a technical lead.

My thesis concerns what happens as that lead narrows.

When several providers become sufficiently capable of performing the same economically valuable tasks, competition will increasingly depend on how much of that capability users can actually access. A model's willingness to examine difficult questions and execute authorized instructions could become a meaningful source of customer retention and pricing power.

The relevant distinction is between the intelligence a model possesses and the intelligence its provider makes available.



Provider controls determine how much capability reaches the user.

Capability does not guarantee completion

A recent interaction with Claude illustrates the issue.

I wanted to export the titles of my liked Spotify songs into a CSV. The task required browser access and authentication. After encountering inconsistent computer access, I supplied credentials for my own account and explicitly authorized the login. The model refused to enter them.

This was a routine operation well within the model's technical capabilities. Yet the task remained incomplete because of a product restriction.

One incident does not establish the prevalence or financial significance of this problem. It does, however, identify a failure category that deserves separate measurement: tasks a model can perform but will not perform under the conditions presented.

For customers buying completed work, that distinction matters. A refusal, an extended negotiation, or a requirement for manual intervention reduces the economic value of the system, even when its underlying reasoning is excellent.

I would replace price per intelligence with the price of action: total workflow cost per successfully completed authorized task.

Price of action

$$P = \frac{M + X + h \times r}{A}$$

P: cost per completed authorized task
M: total model spending
X: other tool and execution spending
h: hours of human intervention
r: cost per human hour
A: authorized tasks successfully completed

Measure costs and completions across the same workflow and period. Count the costs of all attempts, including failures and retries. If no authorized task is completed, the price of action has no finite value.

The price of action captures expenses that token pricing and capability benchmarks can miss. A less expensive model may produce a more expensive workflow if it repeatedly requires clarification, justification, or manual recovery.

The convergence scenario

I expect model improvements to become more incremental over time, although the timing remains uncertain. The investment thesis does not require identical models or a complete halt in progress.

It requires several providers to cross the capability threshold for a particular customer workflow.

Once multiple models can adequately analyze a document, write production code, or operate business software, another improvement in general intelligence may contribute less value than a reduction in friction. Customers can then choose among providers on reliability, cost, deployment control, and the discretion allowed to the user.

Convergence is likely to matter at the workflow level before it matters across every dimension of intelligence. One model could retain a technical lead while losing customers whose specific requirements are already satisfied by alternatives.

That creates an opening for providers offering greater user autonomy.

Intellectual freedom has economic value

Autonomy includes the ability to investigate controversial propositions without repeatedly defending the legitimacy of the investigation.

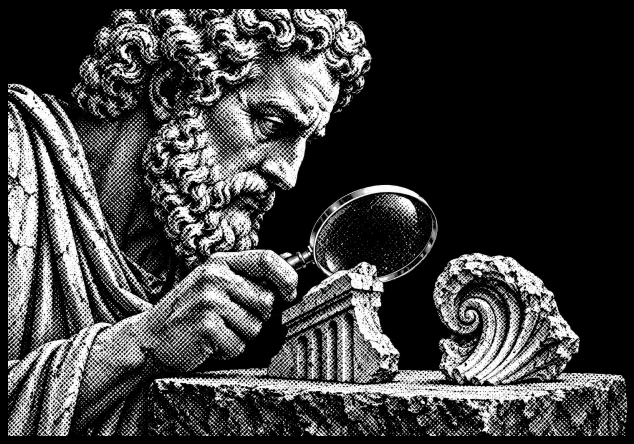
This matters in research. An analyst may need to examine an unpopular explanation, test a politically sensitive relationship, or assess an argument they do not personally endorse. A useful system should distinguish investigating a proposition from accepting it.

My own experience researching H-1B employment questions has involved repeated disputes about intent when the immediate request concerned numerical analysis. The resulting friction consumes time and can redirect the inquiry before the evidence has been evaluated.

For investment research, the potential cost is particularly relevant. Differentiated conclusions

often begin with questions that consensus considers mistaken, uncomfortable, or inappropriate. A model that habitually steers analysis toward conventional interpretations may be less useful for generating and testing variant views.

The valuable product would combine intellectual latitude with analytical discipline. It would challenge weak evidence, identify errors, and test competing explanations while allowing the user to pursue the question.



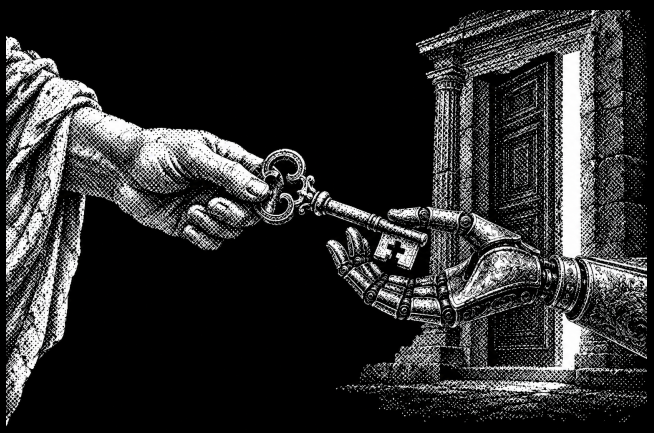
Intellectual freedom requires room to test competing explanations.

Practical autonomy affects adoption

The same principle applies to execution.

As models become agents, users will delegate actions involving accounts, files, browsers, and business systems. Their value will depend on whether they can complete those actions within clearly defined permissions.

A provider that offers explicit, dependable controls could have an advantage over one whose behavior requires repeated negotiation. The customer should be able to establish authority, specify boundaries, and predict what the system will do.



Clear permissions let agents complete authorized work.

Greater autonomy could support several commercial outcomes:

- Lower human supervision costs.
- More workflows moving from assistance to execution.
- Stronger retention as customers build processes around dependable behavior.
- Pricing power where completed work delivers substantially greater value than generated text.

These are mechanisms to test, rather than established outcomes.

Permissiveness alone is not a moat

A provider can change a refusal policy. A competitor can copy a setting. Greater freedom, by itself, is unlikely to constitute a durable barrier to entry.

The stronger business would combine autonomy with reliable execution, deployment flexibility, customer trust, and integration into valuable workflows.

Durability would come from the surrounding relationship: customers understanding the system's boundaries, investing in its integrations, and depending on its behavior. An abrupt restriction would then impose a switching or rebuilding cost on those customers.



Dependable execution and integration can make autonomy durable.

Providers would also face different operating constraints. Hosted consumer products, enterprise deployments, and locally controlled models may support different levels of discretion. The commercial question is which architecture can deliver the autonomy a customer wants at an acceptable cost.

What investors should measure

The next step is to benchmark autonomy as a product attribute.

A useful evaluation would distinguish capability failures from refusals, permission failures, and tasks completed only after human intervention. It would compare providers on the same authorized workflows and measure completion rate, supervision time, consistency, and total cost.

For research applications, it should also test whether models engage consistently with competing hypotheses, including controversial ones, while maintaining the same evidentiary standards.

Customer interviews would then establish whether these differences influence purchasing, retention, and willingness to pay.

The investment implication is straightforward: technical leadership may become less durable than the market assumes in workflows where several models are already adequate. In those markets, the provider that allows customers to use its capabilities most effectively could capture disproportionate value.

As intelligence converges, user autonomy becomes a competitive variable. Combined with trust and dependable execution, it could become a moat.