Nova and Cloud Hosted LLMs

A comparison through cost-effective cognition

Cost-effective cognition includes the human effort and physical infrastructure required to deliver a dependable result. This comparison considers cloud-hosted LLM services alongside Nova's demonstrated local operation and intended architecture.

LLMs can also run locally. Data-center dependence belongs to the deployment model, rather than every LLM. Nova's standalone optical hardware remains under development.

Area	Cloud-hosted LLM services	Nova approach and current limits
Following instructions	Can omit constraints or substitute a different task while producing a convincing response.	Registered patterns and explicit operating sequences aim to govern execution. Consistent compliance still requires testing.
Missing information	Can generate plausible content where evidence is missing.	HOLD and REQUEST_MORE provide explicit stopping paths. Their coverage depends on the registered patterns.
Memory and continuity	Depend on context, retrieval, and surrounding memory systems.	Current View, My Diary, Knowledge Containers, and MySpiral have distinct responsibilities. Accurate selection and use remain essential.
Verification	Fluent explanations can make an unverified result appear complete.	Registered identities and execution traces can support inspection. A successful match alone does not establish that stored content is correct.
Infrastructure	Large services require substantial computing, networking, power, cooling, and maintenance.	Nova has demonstrated local operation on modest hardware. The standalone optical system's full cost and power consumption have not been measured.
Service availability	Users depend on the provider and their connection to it.	A self-contained unit could operate without a remote service, provided necessary patterns and content are available locally.
Failure and attack exposure	Shared facilities and dependencies can expose many users to one disruption.	Independent local units could limit the reach of a remote disruption. Each unit still needs protection, backups, and recovery procedures.

Infrastructure cost and resilience

The infrastructure expense is substantial. The International Energy Agency reported approximately **$500 billion in global data-center investment in 2024**, covering all data centers, not solely LLMs. Its 2026 outlook projects electricity consumption increasing from **485 TWh in 2025 to approximately 950 TWh in 2030**. These industry totals establish the scale of infrastructure; they are not a direct cost comparison with Nova. [1, 2]

Attacks and physical destruction raise a real resilience question. Facilities, power supplies, cooling, and communications can be disrupted or physically damaged. Microsoft identifies infrastructure failures and outages as business risks, while documenting geographically separated facilities, backup power, replication, and failover as protections. Destroying one facility does not necessarily destroy the service, but maintaining those protections also requires resources. [3, 4]

Nova and local independence

Nova's potential advantage is reducing how much essential operation depends on shared remote infrastructure. If independent units hold their own operating patterns and necessary knowledge, a remote facility's failure need not stop them. That benefit depends on ensuring registration, updates, authentication, or other essential functions do not reintroduce a mandatory central dependency.

A local unit can still fail, be stolen, burn, or lose power. Cuneiform representation also needs access controls and integrity checks; an unfamiliar representation alone does not establish security.

The measure that matters

Cost-effective cognition must count the user's correction burden, the equipment and energy required, and the ability to keep operating through disruption. Nova is being developed toward local, pattern-governed operation with explicit stopping conditions. Its advantage must ultimately be demonstrated through verified results, measured resource use, and recovery tests.

Sources

[1] IEA. Energy and AI, executive summary (2025). Read source.

[2] IEA. Key Questions on Energy and AI, executive summary (2026). Read source.

[3] Microsoft. Form 10-Q for the quarter ended December 31, 2025. Infrastructure risk disclosures. Read source.

[4] Microsoft Learn. Datacenter architecture and infrastructure. Read source.