
ELLM

Cloud Models in eLLM

Cloud models in eLLM offer optional access to external AI services, enhancing local capabilities for complex tasks. They come with privacy controls and budget management, allowing organisations to balance cost, access, and data security.

AUTHOR

Paul Flanders

PUBLISHED

8 Jun 2026

DOCUMENT

InfoHub — Knowledge, news, and updates from EssingtonITS

Cloud Models in eLLM

Paul Flanders · 8 Jun 2026

What this product is for

On-premises local models keep everything in-house, but sometimes you want the extra capability of a large external model, for a difficult reasoning task, for example. Cloud models let eLLM call those external providers when your organisation chooses to allow it, while still keeping you in control of access, privacy, and cost.

Main features

- Connect external providers as an optional extra to your local models.
- They appear under **Cloud models** in the chat model picker.
- Each can carry a per-model instruction and have its cost tracked.
- Access can be restricted to specific groups.
- Governance controls add privacy redaction and a monthly spending budget.

How it works

From a user's point of view

In the chat model picker, cloud models appear in the Cloud models group. You can pick one when you want its capabilities, provided your account is allowed to use it. Automatic routing may also choose a cloud model when it is the best fit, though it prefers free local models where they suffice.

From an administrator's point of view

Administrators register cloud models on the Models page, marked as the Cloud tier. Cloud models have extra settings: a per-model instruction, input and output costs for spend tracking, an option to keep them out of automatic routing, a group restriction, and a flag for newer reasoning-style models that need a different request format. Privacy redaction and a monthly budget are set under cloud governance in system settings. See the Managing models and System settings articles.

Common tasks

- **Use a more capable model:** pick a cloud model from the picker, if permitted.
- **Limit a cloud model to certain teams:** restrict it to groups (administrators).
- **Keep an expensive model for deliberate use only:** opt it out of automatic routing.
- **Protect sensitive data:** enable privacy redaction in cloud governance.
- **Control spending:** set a monthly cloud budget.

Things to know

- Cloud models are optional. eLLM works fully without them, on local models alone.
- Using a cloud model sends content to an external provider, so privacy redaction and group restrictions matter.
- Cloud models can carry a cost; administrators set their prices for tracking and can cap monthly spend.
- Automatic routing prefers free local models, so cloud models are typically used when chosen or when they are clearly the best fit.
- Newer reasoning-style models may need a special setting to work correctly.

Troubleshooting

- **A cloud model isn't available to me:** it may be restricted to other groups. Ask your administrator.
- **A newly added cloud model errors:** the model name, access key, or the reasoning-model setting may need correcting; an administrator can check the Models page.
- **Cloud spending is higher than expected:** review the monthly budget and confirm which models are in routing.
- **I'm concerned about sensitive data:** ask your administrator to enable privacy redaction, or use a local model.

Frequently asked questions

What makes a model a "cloud" model?

It is provided by an external service rather than running on your own servers.

Does using a cloud model send my data outside?

Yes, to the external provider. Privacy redaction and group restrictions help manage this; a local model keeps everything in-house.

Will automatic routing use cloud models?

It can, but it prefers free local models. Administrators can also keep specific cloud models out of routing.

Are cloud models required?

No. They are an optional addition. eLLM runs fully on local models alone.

Summary

Cloud models add the capability of large external providers to eLLM as an optional extra. They appear under Cloud models in the picker and can be restricted to groups, priced for tracking, and governed with privacy redaction and a monthly budget. Because they send

content outside your organisation and can cost money, eLLM gives administrators clear controls, and local models remain the private, free default.

Need assistance navigating the complexities of eLLM? EssingtonITS offers expert guidance and tailored IT solutions to help you succeed. Visit EssingtonITS.co.uk for comprehensive support.