

LEARNING NUGGET E-LEARNING &amp; AI 1 OCTOBER 2026

# Model, LLM, chat, agent: four terms, one picture

Model, LLM, chat and agent often get mixed up. A simple picture of layers shows how the four terms fit together and what that means for your teaching.

By Rico Eberle, e-learning expert for continuing education and municipal councillor in Dübendorf



## Video for this post

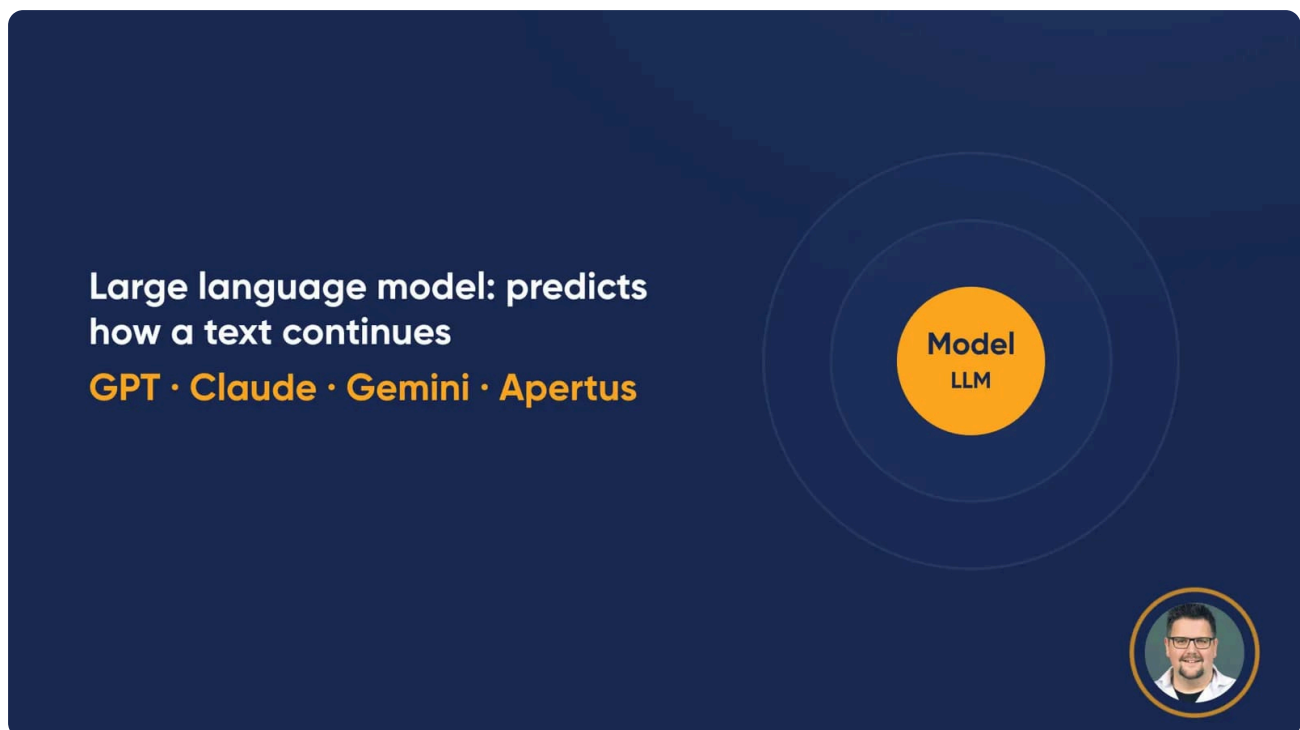
This learning nugget is also available as a video with transcript (1:33 min). Scan the QR code or open the link:

[ricoeberle.ch/en/blog/model-llm-chat-agent/](https://ricoeberle.ch/en/blog/model-llm-chat-agent/)

Model, LLM, chat, agent: these four terms come up in almost every conversation about AI, and two people using them often don't mean the same thing. Yet they fit together quite simply, like layers, one around the other. The picture is a teaching simplification, but it helps you place tools and ask the right questions.

## At the core: the model

At the very centre is the model. A model is the result of training: a huge file with billions of numbers, known as parameters. An LLM, a large language model, is one kind of model. It was trained on a vast amount of text and predicts, piece by piece, how a text continues. GPT, Claude, Gemini or the Swiss model Apertus are such language models.



The model sits at the core. An LLM is one kind of model, not a separate layer.

Apertus shows how large such models are: EPFL, ETH Zurich and the Swiss National Supercomputing Centre CSCS released it in 2025 in two sizes, with 8 and 70 billion parameters.

The "pieces" mentioned in the video are called tokens: words, parts of words or punctuation marks. The video deliberately says "piece" rather than "word", because a model doesn't work word by word. Why this prediction can sound fluent and still be wrong is the subject of the post [AI doesn't lie. It guesses.](#)

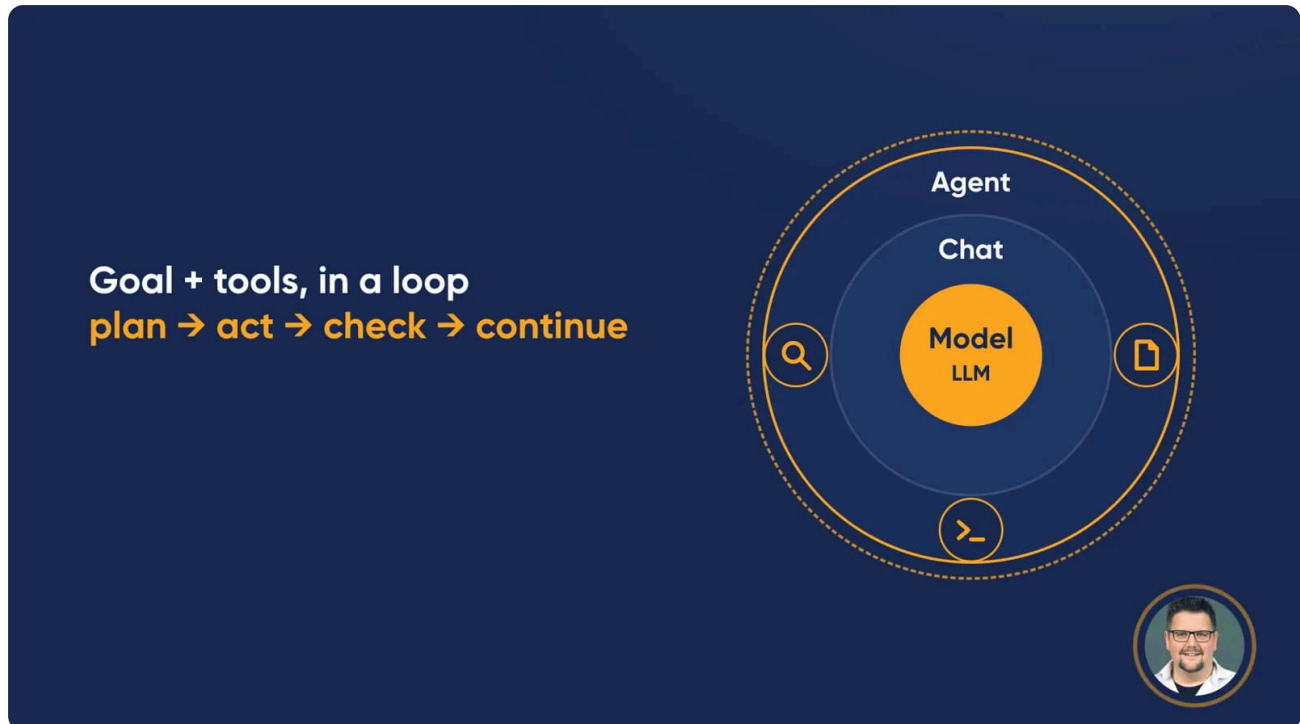
## The first shell: the chat

On its own, a model does nothing. It needs a shell through which you can use it. The first shell is the chat: an input field, a history, maybe a few files. You ask, the model answers.

A well-known name makes the difference tangible: ChatGPT is the chat, and the model inside it is called GPT. So when a colleague says “ChatGPT” got something wrong, it’s worth asking whether they mean the model’s answer or a feature of the interface.

## The second shell: the agent

The agent gives the model a goal and tools: searching the web, opening files, running programs. And it lets the model work in a loop: plan, act, check, continue, until the task is done.



The agent gives the model a goal and tools and lets it work in a loop.

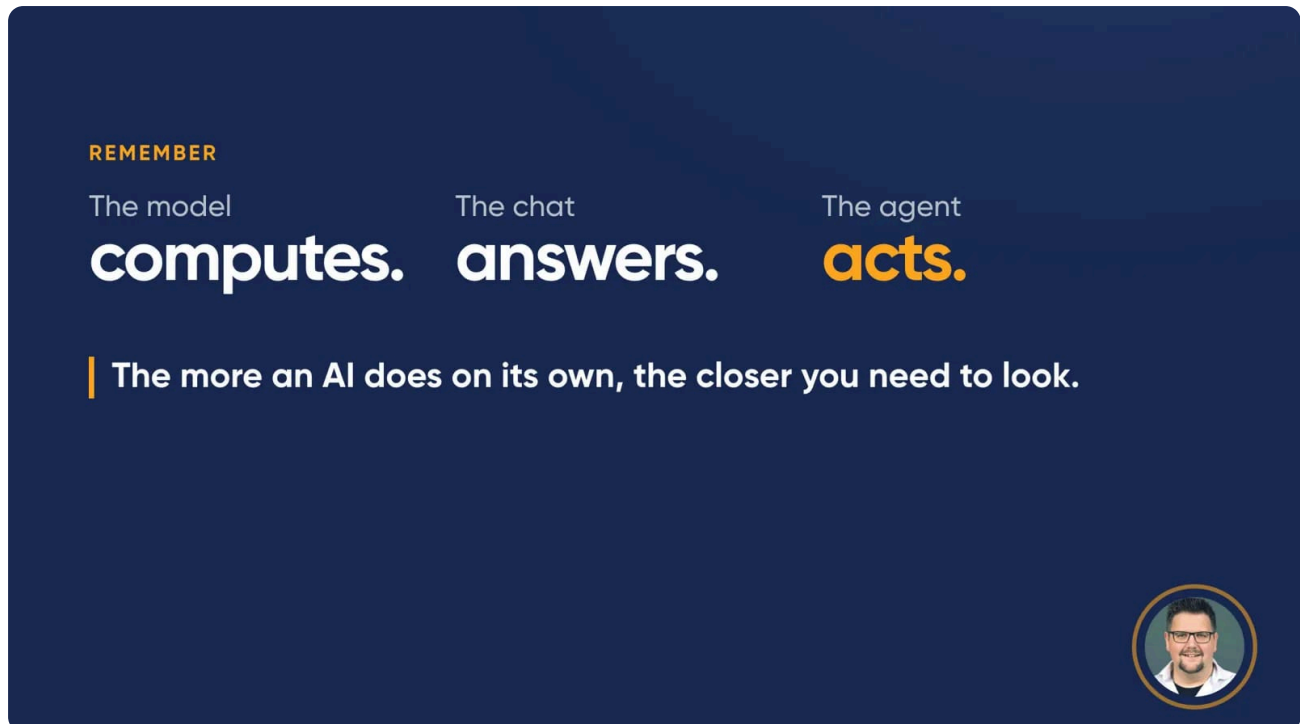
In a 2024 guide, Anthropic describes agents as systems in which the language model directs its own process and tool use. The guide distinguishes them from workflows, where predefined code decides when the model uses which tool. It also names the downside: agents cost more, and errors can compound over several steps. It therefore recommends extensive testing in sandboxed environments and appropriate guardrails.

## The lines are blurring

Many chats can now search the web. Then they already work a little like an agent: the model decides for itself whether and what to search for, and works the results into its answer. The picture of layers is therefore not a sharp division but a way to get your bearings. The further out, the more the AI does on its own.

## What this means for teaching

The line to remember from the video sums up the picture: the model computes, the chat answers, the agent acts.



The model computes, the chat answers, the agent acts.

For course leaders, this leads to two practical points:

**Say exactly what you mean.** If you set rules for AI in your course, make clear whether you mean a chat or also tools that search, edit files or run programs on their own.

**Look more closely the more the AI does on its own.** With a chat answer, you check a text. With an agent, you also check which steps it took, which sources it opened and what it changed. That's my own view; it fits Anthropic's warning about compounding errors and the need for testing and guardrails.

How an AI assistant works with your own documents, and what the language model really sees, is explained in [RAG explained: your AI never read the handbook.](#)

## My suggestion

In your next training session, draw three nested circles on a flip chart: model, chat, agent. Ask participants to write the AI tools they know or use on sticky notes and place each one in a circle. The borderline cases are the interesting ones, such as a chat with web search. Ask there: what does the tool do on its own, and what do I need to check afterwards?

If you keep the layers apart, you know where to look more closely.

## Frequently asked questions

### What is the difference between ChatGPT and GPT?

ChatGPT is the chat, the interface with an input field and a history. GPT is the language model inside it that computes the answers.

### What is an LLM?

A large language model (LLM) is an AI model trained on a vast amount of text that predicts, piece by piece (strictly speaking, token by token), how a text continues. Examples include GPT, Claude, Gemini and the Swiss model Apertus.

### What is an AI agent?

An AI agent gives a language model a goal and tools, such as web search, files or programs, and lets it plan, act and check in a loop until the task is done. Anthropic describes agents as systems in which the model directs its own process and tool use.

## Sources

1. ETH Zurich: Apertus, a fully open, transparent, multilingual language model (2025)  
<https://ethz.ch/en/news-and-events/eth-news/news/2025/09/press-release-apertus-a-fully-open-transparent-multilingual-language-model.html>
2. Kalai et al.: Why Language Models Hallucinate (2025)  
<https://arxiv.org/abs/2509.04664>
3. Anthropic: Building effective agents (2024)  
<https://www.anthropic.com/engineering/building-effective-agents>



### ABOUT THE AUTHOR

#### Rico Eberle

Rico Eberle is an e-learning expert, business economist (FH) and municipal councillor in Dübendorf, Switzerland. He chairs the foundation board of WBK Dübendorf, a continuing education foundation. In the learning nuggets he explains research on learning, AI and digital sovereignty, briefly and with sources.

Contact: [ricoeberle.ch](https://ricoeberle.ch) · LinkedIn: [linkedin.com/in/rico-eberle](https://www.linkedin.com/in/rico-eberle)

### LICENCE

Text, transcript and video by Rico Eberle are licensed under CC BY 4.0 (music and sound effects in the video excluded). You may share and adapt the content, including for teaching, as long as you credit the source. [creativecommons.org/licenses/by/4.0/](https://creativecommons.org/licenses/by/4.0/)

### SUGGESTED CITATION

"Model, LLM, chat, agent: four terms, one picture", Rico Eberle, [ricoeberle.ch](https://ricoeberle.ch), 2026,  
<https://ricoeberle.ch/en/blog/model-llm-chat-agent/>, CC BY 4.0