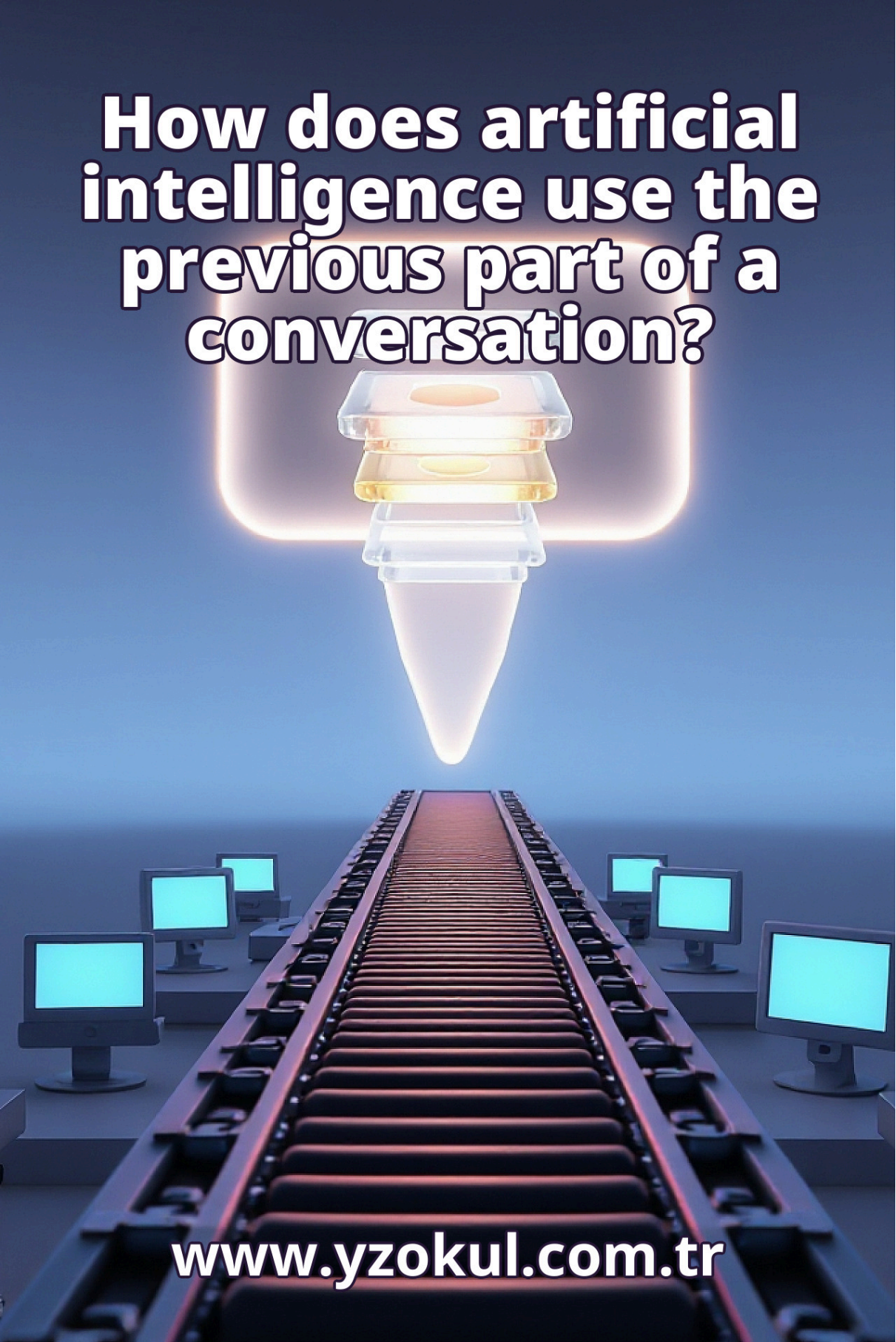
How does artificial intelligence use the previous part of a conversation?



www.yzokul.com.tr

How does artificial intelligence use the previous part of a conversation?



You ask how artificial intelligence uses the earlier part of a conversation. The short answer: the chat system shows the earlier messages to the model again, and the model reads them together with your new question. This happens each time you send a message. For example, if you first ask about cats and then say tell me more, the system includes your first question so the model knows what more refers to.



Each message you type is just text sent to the AI. The AI's reply is also just text. Both your messages and the AI's replies are kept in the chat window. For example, when you type What is a dog? and the AI answers, that question and answer stay visible in the chat. The chat system stores this text so it can be used later.



The model itself has no eyes or ears and no live memory. It only works with the text it is given at that moment. It does not see your screen or hear your voice. For example, if you say Hello, the model receives only the text Hello and maybe earlier messages if the system sends them. It cannot know anything else unless it is in that text.



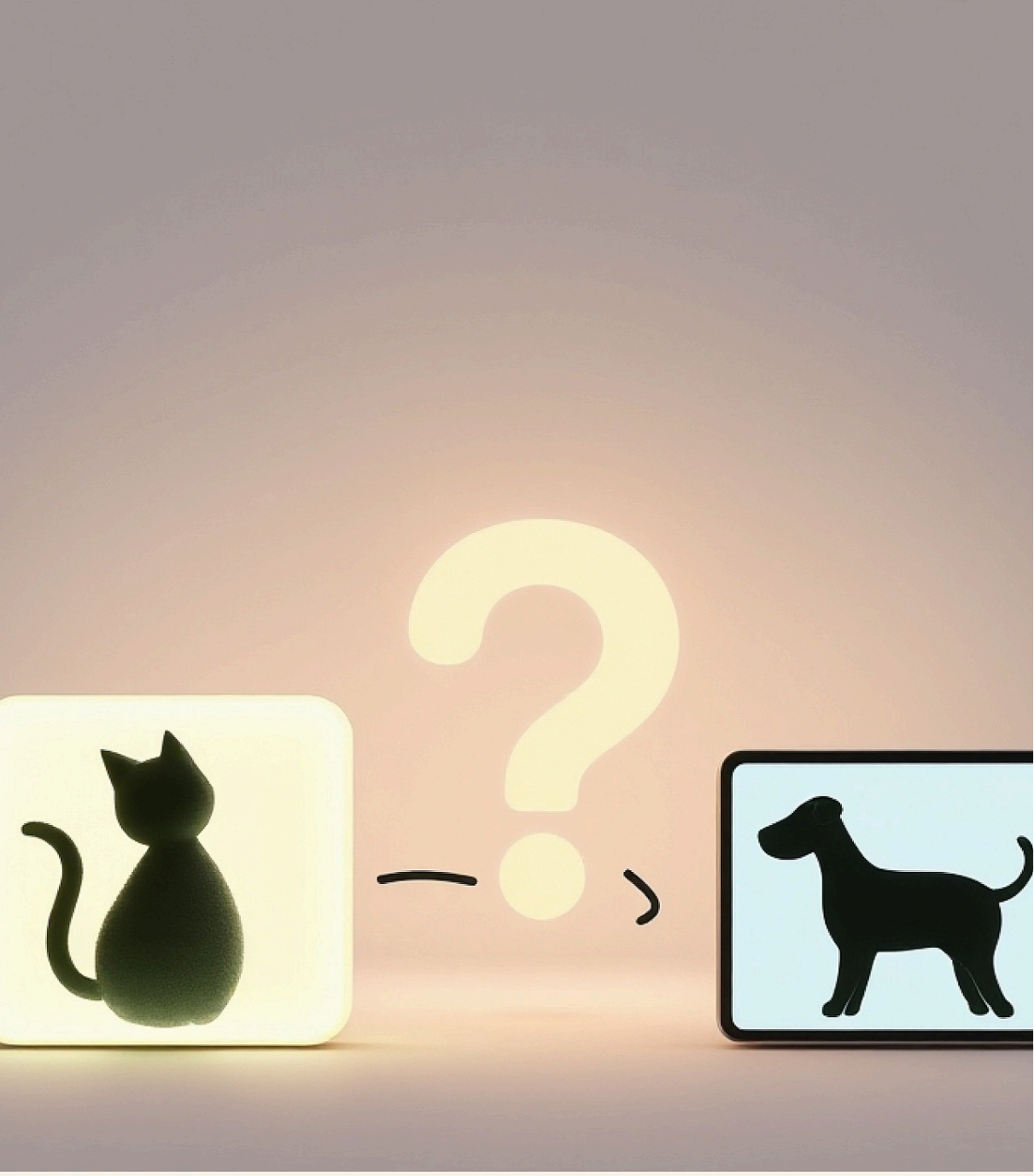
The chat system takes the previous messages and sends them to the model again with each new question. The model reads this whole block of text as its input. For example, if you asked about dogs and then ask How big are they?, the system sends your first question, the AI's answer, and your new question together. The model then sees all of it.



The model evaluates the old messages and the new question as one piece of text. From this it works out what the new question means. It does not treat them separately. For example, if the old text is about weather and the new question is Will it rain?, the model connects it to weather and understands that it means the weather. This helps it give a relevant answer.



You ask about a topic, then say tell it more briefly. Because the earlier text is included, the model knows what it refers to. For example, you ask What is photosynthesis? and get a long answer. Then you say Tell it more briefly. The model sees the earlier explanation and shortens that same topic. Without the earlier text, it would not know what to shorten.



Small words such as it, that, or the second one only make sense with the earlier text. The resent messages give the model that missing meaning. For example, if you say I have a cat and a dog. The second one is bigger. The model needs the earlier sentence to know that the second one means the dog. The chat system provides that earlier sentence.



The system may send all previous messages or only a selected portion of them. It chooses what fits and what seems useful. For example, if the conversation is very long, it might send only the last few messages. This selection helps the model focus on the most relevant recent text. The choice depends on the system's design and limits.



The context window is the limited amount of text the model can take in at once. Old messages take up room in this window. For example, if the window holds 4000 words, a long conversation may fill it up. The model can only use text that fits inside this window. When the window is full, older text may be left out.



When the conversation gets long, the oldest text may be dropped or shortened. Then the model can no longer refer to it. For example, after many messages, the first question might be removed to make space. If you then ask a question about that first topic, the model may not have the information. It may give a less accurate answer or ask for clarification.



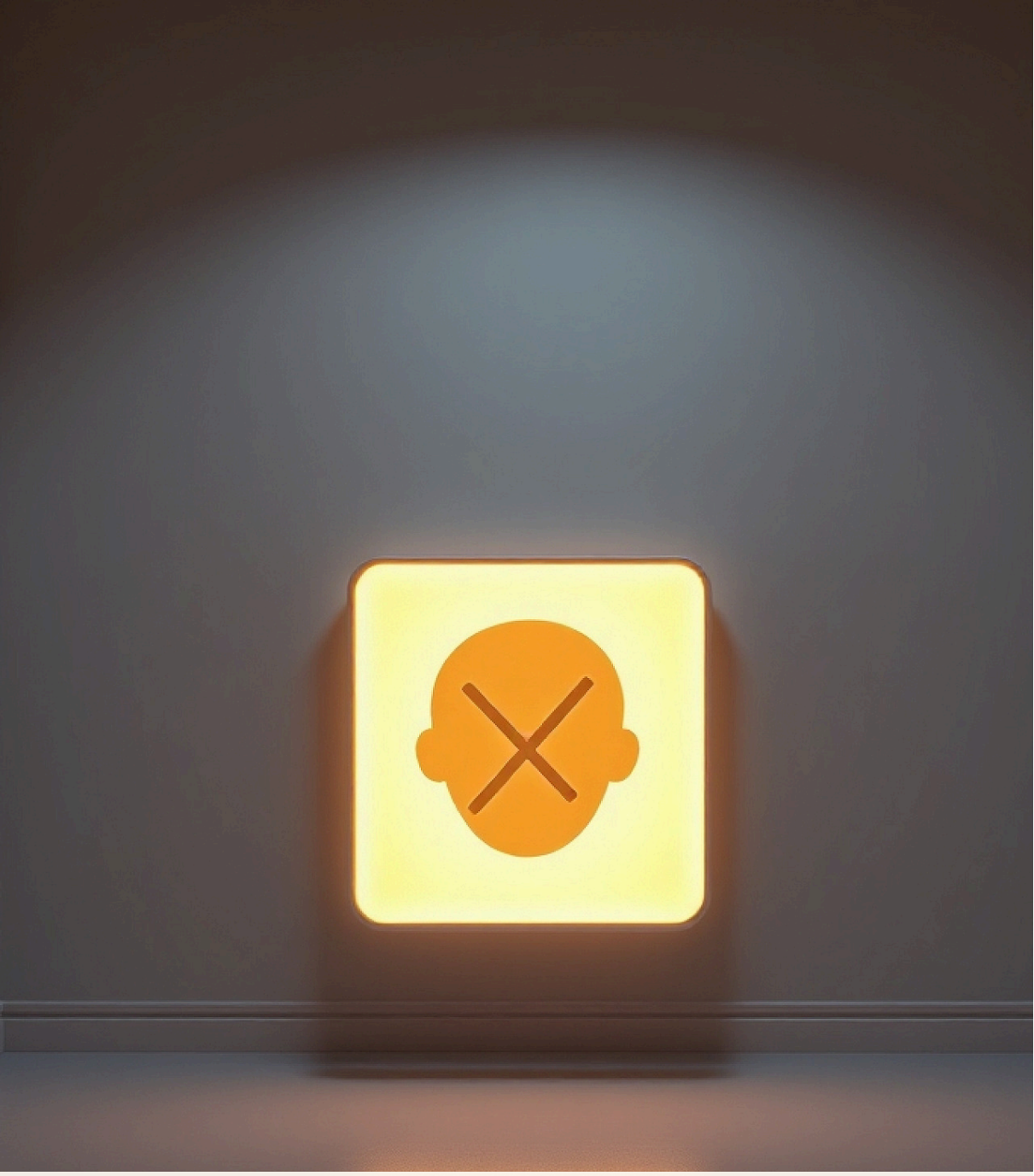
This is not remembering like a person does. The model does not keep anything between calls. The chat system keeps the text and hands it back each time. For example, if you close the chat and start a new one, the model does not recall the old conversation unless the system sends it again. Each call to the model is independent.



Some applications write a short summary of the earlier conversation. This summary is then sent instead of the full old messages. For example, after a long chat about a trip, the system might create a summary like The user plans a trip to Italy and asked about flights. Then it sends this summary with new questions. This saves space in the context window.



Some applications save notes or records of earlier chats and look them up later. These notes are added to the text the model receives. For example, a customer service bot might save a note that you asked about returns last week. When you chat again, it adds that note to the conversation. The model can then refer to it.



A common mistake is to think the model continuously remembers all your conversations. It does not. It only sees what the system sends in that moment. For example, if you tell the AI your name in one chat and then start a new chat, it will not know your name unless the system sends that information again. The model has no memory of its own.



Another mistake is to think the AI forgets on purpose. When it loses track, it is usually because old text fell outside the context window, not because it chose to forget. For example, if the conversation is very long and the earliest messages are dropped, the model cannot refer to them. It is not being stubborn; it simply does not have that text.



The chat system resends earlier messages or a portion of them. The model reads them with the new question. Limits, summaries, and records shape what it can use. For example, if you ask a follow-up question, the system includes relevant previous text so the model can understand. Remember that the model itself does not remember; the system provides the context each time.