# Knowledge Generation Prompting

The goal of Knowledge Generation prompting is to make your Language Model retrieve specific bits of information from its giant pool of training data. Picture this technique as asking your model to do some research before writing a final response.

Suppose you want your model to write a blog post about growing flowers on your balcony. Instead of asking your model to write the blog right away, you can prompt it to generate key points about gardening, flowers, and space management.

Once you get the desired key point, make sure to attend to your fact-checking duties. From there, prompt your model to use the “knowledge” it generated to write an article.

Knowledge Generation improves the output quality because it forces your model to focus on specific points instead of trying to answer a vague prompt.

Here’s how you can introduce Knowledge Generation into your prompts:

**[Knowledge Generation prompting]**

**Prompt Example #1:** Act like an expert horticulturist who specializes in maintaining balcony gardens. Generate key facts about growing flowers under Hamburg's weather, and managing space on your balcony. Add sources and quotes for each point.

Use the generated information to write a 2000-word blog post about how to grow flowers on your balcony for people who live in Hamburg.

\*\*\*

**Prompt example #2:** Act like an expert personal trainer. List the top 20 techniques of total-body stretching and add a detailed description of how to perform each technique.

I will then pick a sublist of those techniques, and you'll kindly provide me with a bi-weekly stretching routine based on my choices.

\*\*\*

**Prompt example #3:** Retrieve historical facts about the rise and fall of Carthage. Include dates, names, and current geographical locations.

From there, kindly write an essay about the relationship between Carthage and the Levant.

## **Knowledge Generation Prompting and ChatGPT Plugins**

You can use ChatGPT plugins to both generate knowledge and help with fact-checking. Make sure to try as many plugins as possible because most of them are still clunky.

# Knowledge Integration Prompting\*

The main weakness of Knowledge Generation prompting is the timeline. GPT-4’s training data stops in September 2021, which means all the content that came afterward is unknown to the model.

The cutoff date isn’t a problem when you deal with timeless topics like gardening, writing, and cooking, but if you’re chasing the latest information, you need a complementary trick.

You can use plugins, chatbots with online browsing, or Knowledge Integration prompting.

All you have to do is feed recent data into your model to help it catch up with the news. In a way, you make your offline model integrate new knowledge.

For API users, [GPT-4 can process up to 32,000 tokens](https://www.semrush.com/blog/gpt-4/), which represent about 25,000 words. This includes both the user prompt and the answer. For users of ChatGPT Plus, GPT-4 can take up to 4096 tokens as input, which is approximately 3,000 words.

You can use these 3,000 words and the chat history feature to “teach” ChatGPT-4 new information. The model itself won’t integrate the data, but you can generate prompts that leverage the “new information” you just added.

Below is a framework you can use to apply Knowledge Integration prompting:

* Find a relevant source, like a research paper or a documented article.
* Identify the most informative parts of the paper at hand.
* Cut the parts into chunks of 3,000 words.
* Feed the chunks into ChatGPT-4 and ask it to explain each section in simple words. You can also ask for quotes and examples.
* Use ChatGPT-4's output for a new prompt.

## Example:

Let’s say you’re an AI researcher specializing in Large Language Models. Your current task is to reference material that’s relevant to your thesis.

You found an interesting paper titled Language Models Can Solve Computer Tasks. You want to take notes before skimming the other 122 papers you bookmarked last week.

Here are the steps you can follow to get ChatGPT to help you take quick notes.

* First, identify the passage you want to summarize. In this example, we’ll select the discussion part which makes for about 1,000 words.

[Example text : Intuitive psychology is a pillar of common-sense reasoning. The replication of this reasoning in machine intelligence is an important stepping-stone on the way to human-like artificial intelligence. Several recent tasks and benchmarks for examining this reasoning in Large-Large Models have focused in particular on belief attribution in Theory-of-Mind tasks. These tasks have shown both successes and failures. We consider in particular a recent purported success case, and show that small variations that maintain the principles of ToM turn the results on their head. We argue that in general, the zero-hypothesis for model evaluation in intuitive psychology should be skeptical, and that outlying failure cases should outweigh average success rates. We also consider what possible future successes on Theory-of-Mind tasks by more powerful LLMs would mean for ToM tasks with people.]

[Source](https://arxiv.org/abs/2302.08399).

* Cut these lengthy passages into chunks of 3,000 words (not needed in this example).
* Feed these chunks of text to ChatGPT.
* Ask ChatGPT to write a summary of the text you provided.
* Repeat the process for all the papers you want to summarize.
* Don’t forget to fact-check.
* Use your freshly created reading notes to find common threads, and confront opposing results.

Here’s what the framework looks like in practice:

**[Knowledge Integration]**

**Context:** Comparing notes from two research papers.
[All of the prompts below must be performed in the same chat history tab/session].

**Prompt #1:** Hey ChatGPT, I'll submit a few paragraphs from a research paper. Please memorize the whole text and then summarize it in bullet points.

##

Desired format:
Title of the paper: <paper\_title>
URL of the paper: <paper\_URL>
/////
Summary:
- Point #1
- Point #2
- Point #3
- etc...

##

<paper\_title>: Theory of Mind May Have Spontaneously Emerged in Large Language Models
<paper\_URL>: https://arxiv.org/abs/2302.02083
<text\_to\_summarize>: (I inserted around 1,000 words here)

**\* ChatGPT-4's output #1:**

Title of the paper: Theory of Mind May Have Spontaneously Emerged in Large Language Models
URL of the paper: https://arxiv.org/abs/2302.02083
/////
Summary:

- GPT-3 and later versions show increasing ability to solve Theory of Mind (ToM) tasks, with GPT-4 performing even better
- ToM-like ability likely emerged spontaneously and autonomously, as a byproduct of models' increasing language ability
- Two potential explanations: models developed ability to impute unobservable mental states to others (ToM), or they discovered unknown language patterns to solve ToM tasks without engaging ToM
- If language patterns are responsible, it may necessitate re-examination of ToM tasks and research conclusions
- Studying AI performance on ToM tasks could provide insights into human cognition and neural structures that enable similar abilities

\*\*\*

**Prompt #2:** Hey ChatGPT, I'll submit a few paragraphs from a research paper. Please memorize the whole text and then summarize it in bullet points.

##

Desired format:
Title of the paper: <paper\_title>
URL of the paper: <paper\_URL>
/////
Summary:
- Point #1
- Point #2
- Point #3
- etc...

##

<paper\_title>: Large Language Models Fail on Trivial Alterations to Theory-of-Mind Tasks
<paper\_URL>: https://arxiv.org/abs/2302.08399
<text\_to\_summarize>: (I inserted around 1,000 words here)

**\* ChatGPT-4's output #2:**

Title of the paper: Large Language Models Fail on Trivial Alterations to Theory-of-Mind Tasks
URL of the paper: https://arxiv.org/abs/2302.08399
/////
Summary:

- Large Language Models (LLMs) like GPT-3.5 may not have spontaneously developed Theory-of-Mind (ToM) as they fail on simple alterations to ToM tasks
- Encourage a skeptical stance when assessing the claim that LLMs have developed ToM, due to the human tendency to anthropomorphize agents
- ToM may be better achieved by integrating computational models that directly capture the ability with language models, rather than expecting it to emerge spontaneously
- Concern: once a systematic generator of examples or a benchmark is provided, LLMs could pass without answering the basic question of what they have learned
- The evaluation of generalization in machine-learning models is a broader issue, but ToM is a particularly relevant example

\*\*\*

**Prompt #3:** Act as a research assistant specializing in <field>.
Use a scientific approach to compare the content of the previous papers based on their summaries.
Write the results in bullet points followed by precise details.

##

<field>: Deep Learning.

**\* ChatGPT-4's output #3:**

1. Theory of Mind emergence in large language models:

Paper #1 claims that Theory of Mind (ToM) may have spontaneously emerged in large language models like GPT-3 and GPT-4 as they show an increasing ability to solve ToM tasks
Paper #2 argues that large language models, such as GPT-3.5, probably have not developed ToM since they fail on simple alterations to ToM tasks

2. Methodological concerns:

Paper #1 suggests that ToM-like ability might have emerged spontaneously and autonomously in large language models as a byproduct of their increasing language ability
Paper #2 highlights the concern that providing a systematic generator of examples or a benchmark may help models pass ToM tests without truly understanding ToM

3. Skepticism and anthropomorphism:

Paper #2 encourages adopting a skeptical stance when assessing claims that large language models have developed ToM, as humans may be biased to anthropomorphize agents

4. Implementation of Theory of Mind in AI:

Paper #1 raises the possibility that AI models may develop mechanisms similar to those employed by the human brain to impute mental states to others, boosting our understanding of AI and human cognition
Paper #2 suggests integrating computational models that directly capture ToM with language models, rather than expecting ToM to emerge spontaneously from additional linguistic data

5. Evaluation of generalization in machine-learning models:

Paper #2 points out that the evaluation of generalization in machine-learning models is a concern for many researchers, with ToM being a particularly troubling example

\*\*\*

**Prompt #4:** Act as a research assistant specializing in <field>.
Use a scientific approach to compare the content of the two previous papers based on their summaries.
Write the results in bullet points followed by precise details.

##

<field>: Congnitive Psychology.

**\* ChatGPT-4's output #4:**

1. Assertion of Theory of Mind (ToM) in Large Language Models (LLMs):

 Paper #1 suggests that GPT-3 and later versions, including GPT-4, show an increasing ability to solve ToM tasks, hinting that ToM-like ability might have spontaneously emerged.
 Paper #2 counters this by stating that LLMs like GPT-3.5 fail on simple alterations to ToM tasks, implying they may not have genuinely developed ToM.

2. Emergence of ToM in LLMs:

 Paper #1 proposes that the ToM-like ability might have arisen spontaneously and autonomously as a byproduct of the models' increasing language abilities.
 Paper #2 encourages skepticism regarding this claim, highlighting the potential for humans to anthropomorphize agents and erroneously attribute ToM to them.

3. Explanations and Mechanisms behind ToM-like Abilities in LLMs:

 Paper #1 offers two potential explanations: models might have developed the ability to impute unobservable mental states to others, or they might have identified unknown language patterns that enable them to solve ToM tasks without genuinely engaging ToM.
 Paper #2 doesn’t directly counter this but suggests integrating computational models that directly capture ToM abilities with language models might be a better approach than expecting spontaneous emergence.

4. Implications for Theory of Mind Research:

 Paper #1 mentions that if mere language patterns enable the ToM task-solving, then there might be a need to re-examine ToM tasks and research conclusions derived from them.
 Paper #2 warns of a concern where, if a systematic generator of examples or a benchmark for ToM is provided, LLMs could pass the tasks without truly addressing the fundamental question of their understanding of ToM.

5. Broader Relevance and Implications:

 Paper #1 indicates that observing AI performance on ToM tasks might provide insights into human cognition and the neural structures enabling similar abilities.
 Paper #2 places emphasis on the broader issue of evaluating generalization in machine-learning models, using ToM as a particularly relevant example.

In conclusion, while both papers discuss the capabilities of Large Language Models (LLMs) regarding Theory of Mind (ToM) tasks, they present contrasting viewpoints. Paper #1 leans more towards the possibility of LLMs developing ToM abilities, while Paper #2 is more skeptical, emphasizing caution in interpreting LLM performance on ToM tasks.

**Note:** if the final output is too long, ChatGPT will stop writing its response midway. In this case, you can prompt it with the word “Continue,” and it will resume writing from the point it was cut off.

## **Knowledge Integration\* and Microsoft Edge**

When using Knowledge Integration prompts, you can use the “Chat” feature of Microsoft Edge for more efficiency.

Instead of navigating the material yourself, you can open a web page or a PDF in Edge and use the Chat feature to summarize the content. From there, inject the summary into ChatGPT and use it for another prompt like the one we saw in the previous example.

Here’s a prompt you can use to summarize a document using Microsoft Edge:

**[Prompt to generate summaries using Microsoft Edge's Chat feature]**

**Prompt:** Summarize this paper. Start the summary with the title of the paper and its URL. Then list the main ideas in bullet points. Please illustrate the main ideas with examples extracted from the paper.