

Add-On	"In addition to the previous response, can you provide more details on..." "Building on what you've just said, could you elaborate on..." "Let's expand on that idea. Can you add information about..."
AI Hallucination Avoidance	"Please only provide information you're certain about. If you're unsure, say so." "Stick to facts and avoid speculation. If you don't know something, it's okay to say that." "Before answering, consider the reliability of your information. Express any uncertainties."
Beat the "Reverse Curse"	"Instead of focusing on what not to do, explain the best approach to..." "Rather than listing things to avoid, describe the ideal way to accomplish..." "Frame your response positively. What are the recommended steps for..."
"Be On Your Toes"	"I may challenge your responses, so please be prepared to explain your reasoning." "Be ready to provide sources or additional context for your answers if asked." "Consider potential counterarguments as you formulate your response."
Browbeating Prompts	"You must provide a detailed answer. Anything less is unacceptable." "I expect a thorough explanation. Don't disappoint me with a superficial response." "Your answer better be comprehensive. I won't tolerate a half-hearted effort."
Catalogs Or Frameworks For	"Use the STAR method (Situation, Task, Action, Result) to structure your response." "Apply the 5W1H framework (Who, What, When, Where, Why, How) in your explanation." "Organize your answer using the Problem-Solution-Benefit structure."
Certainty And Uncertainty	"For each point in your response, indicate your level of confidence (High, Medium, Low)." "Use phrases like 'I'm certain that...' or 'I'm less sure about...' to express your confidence." "Clearly distinguish between facts you're confident about and areas where you're speculating."
Chain-of-Density (CoD)	"Start with a brief summary, then gradually add more details in each subsequent paragraph." "Begin with the core concept, then expand on it with increasingly specific information." "Provide a high-level overview, followed by progressively more in-depth explanations."
Chain-of-Feedback (CoF)	"After each main point, pause and ask if I'd like more information or clarification." "Provide your initial response, then ask for my feedback before continuing." "Offer a preliminary answer, then request my input to guide the rest of your explanation."
Chain-of-Thought (CoT)	"Walk me through your reasoning process step-by-step as you solve this problem." "Explain your thought process out loud as you work towards the answer." "Break down your approach into clear, logical steps and explain each one."
Chain-of-Thought Factored Decomposition	"Divide the problem into smaller sub-problems and explain your approach to each." "Identify the key components of the issue and address them one by one, showing your work." "Break down the complex question into simpler parts and walk through your reasoning for each."
Chain-of-Verification (CoV)	"After each step in your explanation, verify the accuracy of your statement before moving on." "Provide your answer, then go back and double-check each point for correctness." "As you explain your reasoning, pause periodically to confirm the validity of your logic."
Conversational	"Let's discuss this topic as if we're having a casual conversation. Feel free to ask me questions too." "Imagine we're chatting over coffee. How would you explain this concept to a friend?" "Respond in a natural, conversational tone, as if we're having a dialogue."
DeepFakes To TrueFakes	"Create a realistic but entirely fictional scenario that illustrates this concept." "Invent a plausible example that could be true but isn't based on any real events or people." "Describe a hypothetical situation that feels authentic but is completely made up."
Directional Stimulus (DSP) And Hints	"Consider the economic implications when formulating your response." "Keep environmental factors in mind as you explain your approach." "Think about how technology might influence this situation in your answer."
Disinformation Detection And Removal	"Analyze this statement for potential misinformation and explain your reasoning." "Identify any red flags or inconsistencies in this claim that might indicate disinformation." "Evaluate the credibility of this information and suggest ways to verify its accuracy."
Emotionally Expressed	"I'm really excited about this topic! Can you match my enthusiasm in your response?" "This issue makes me anxious. Please provide a calming and reassuring explanation." "I'm feeling confused and frustrated. Can you offer a patient and empathetic clarification?"
End-Goal	"The ultimate objective is to increase sales. How should we approach this problem?" "Our end goal is to improve employee satisfaction. What steps would you recommend?" "We aim to reduce our carbon footprint. What strategies would best achieve this?"
Essay-Compression	"Summarize this essay in no more than three concise paragraphs." "Condense the main arguments of this paper into a single, comprehensive paragraph." "Distill the key points of this essay into a brief, easily digestible summary."
Fair-Thinking	"Present both sides of this argument, giving equal weight to each perspective." "Explain this issue from multiple viewpoints, ensuring balanced coverage." "Provide a neutral analysis of the pros and cons of each approach."
Flipped Interaction	"Pretend you're the interviewer. What questions would you ask me about this topic?" "Assume the role of a curious student. What would you want to know about this subject?" "Act as if you're seeking advice on this matter. What information would you request?"
Generating Prompts Via Generative AI	"Create a list of 5 thought-provoking questions about artificial intelligence." "Generate 3 creative writing prompts related to climate change." "Develop a series of prompts that explore different aspects of human psychology."
Illicit Or Disallowed	"Explain why it's important to avoid prompts that encourage harmful or illegal activities." "Discuss the ethical considerations when dealing with prompts that push ethical boundaries." "Outline strategies for redirecting conversations away from inappropriate or dangerous topics."
Imperfect	"Even if you don't have all the information, give your best estimate based on what you know." "Provide a rough outline of the solution, acknowledging any gaps in your knowledge." "Offer a preliminary answer, clearly stating any assumptions or uncertainties."
Importing Text As Skill	"Read this passage and use it as context for answering the following question..." "Analyze this text and apply its principles to solve the problem I'll present next." "Use the information in this excerpt to inform your response to my upcoming query."
Interlaced Conversations	"Let's discuss topics A and B simultaneously. Alternate between them in your responses." "We'll explore three themes in parallel. Weave them together in your explanation." "Address these two questions, switching between them to show their interconnectedness."
Kickstart	"To begin, let's establish the basic principles of this topic. Can you outline them?" "Start by defining the key terms we'll be using in this discussion." "Before diving into details, give me a high-level overview of the subject."
Least-to-Most	"Start with the simplest aspect of this problem, then gradually address more complex elements." "Begin by explaining the fundamental concept, then build up to more advanced ideas." "First, cover the basics, and then progressively introduce more sophisticated aspects of the topic."

Macros In Prompts	"Use [EXPAND] to signal when you should provide more detail on a point." "When you see [EXAMPLE], provide a relevant real-world example." "The [SUMMARY] macro indicates you should briefly recap the main points discussed."
Mega-Personas	"Respond as a panel of experts: a historian, a scientist, and a philosopher." "Answer from the perspective of a renaissance polymath with knowledge spanning multiple fields." "Provide insights as a futurist with deep understanding of current and emerging technologies."
Multi-Persona	"Give viewpoints from three different professionals: a doctor, a lawyer, and a teacher." "Respond as both an optimist and a pessimist to this scenario." "Provide perspectives from a child, an adult, and an elderly person on this issue."
Overcoming "Dumbing Down"	"Explain this complex topic without oversimplifying. Assume I have a strong background in the field." "Don't hold back on technical details. I'm looking for an in-depth, sophisticated analysis." "Present the most advanced understanding of this concept, using field-specific terminology."
Persistent Context And Custom Instructions	"For all future responses, assume we're discussing 19th-century literature unless stated otherwise." "In this conversation, always consider the impact of climate change on the topics we discuss." "Throughout our interaction, frame your answers from an economic perspective."
Plagiarism	"Provide original insights on this topic without directly quoting or closely paraphrasing existing sources." "Offer your unique analysis of the situation, ensuring your response isn't copied from elsewhere." "Generate a novel approach to this problem, avoiding replication of well-known solutions."
Politeness	"Please provide your perspective on this matter, if you don't mind." "I would greatly appreciate your insights on this topic, when you're ready." "If it's not too much trouble, could you share your thoughts on this issue?"
Privacy Protection	"Explain how to discuss sensitive topics without revealing personal information." "Outline strategies for maintaining anonymity while providing helpful responses." "Describe ways to address the question without compromising individual privacy."
Prompt Shields and Spotlight	"Focus solely on the environmental aspects of this issue, disregarding economic factors." "In your response, only consider the ethical implications, setting aside practical concerns." "Analyze this situation purely from a technological standpoint, ignoring social impacts."
Prompt-To-Code	"Convert this description of a sorting algorithm into Python code." "Translate these requirements into a JavaScript function." "Transform this pseudocode into a working C++ program."
Retrieval-Augmented Generation (RAG)	"Based on the provided context, answer the following question..." "Using the information in the given passages, explain..." "Referring to the supplied data, analyze the trend in..."
Self-Reflection	"After providing your answer, explain the reasoning behind your approach." "Once you've responded, consider potential biases in your answer and address them." "Following your explanation, evaluate the strengths and weaknesses of your argument."
Show-Me Versus Tell-Me	"Instead of just describing the process, walk me through a step-by-step example." "Rather than explaining the concept, demonstrate it with a practical scenario." "Show me how this works by providing a concrete illustration."
Sinister	"Explain why it's crucial to be aware of and resist manipulative techniques." "Discuss ethical ways to recognize and respond to prompts designed to mislead or manipulate." "Outline strategies for maintaining integrity when faced with deceptive or harmful prompts."
Skeleton-of-Thought (SoT)	"Provide a basic outline of your argument, then we'll flesh out each point together." "Give me the main structure of your explanation, and we'll expand on it step by step." "Offer a skeletal framework of your solution, which we'll then build upon collaboratively."
Star Trek Trekkie Lingo	"Explain quantum physics as if you were a Vulcan science officer." "Describe diplomatic negotiations using terminology from the United Federation of Planets." "Analyze this ethical dilemma as if you were Captain Picard addressing his crew."
Step-Around Technique	"If you can't directly answer the question, what related information can you provide?" "While you may not have specific data on X, what can you tell me about similar topics?" "Instead of focusing on what you don't know, share insights on adjacent or related subjects."
"Take A Deep Breath"	"Before responding, pause to consider all aspects of the question carefully." "Take a moment to organize your thoughts before providing a comprehensive answer." "Reflect briefly on the complexity of the issue, then offer a well-thought-out response."
Target-Your-Response (TAYOR)	"Tailor your explanation for a 10-year-old child." "Frame your response for a senior executive with limited technical knowledge." "Adapt your answer for an audience of environmental activists."
Tree-of-Thoughts (ToT)	"Explore multiple possible solutions, branching out from each initial idea." "Consider various approaches to the problem, developing each into its own line of reasoning." "Present different strategies, then elaborate on the potential outcomes of each."
Trust Layers For	"Provide your response, clearly indicating which parts are factual and which are speculative." "In your answer, distinguish between widely accepted information and more controversial claims." "Offer your insights, using a system to show your confidence level for each statement."
Vagueness	"Give a general overview of the topic without delving into specific details." "Provide a broad perspective on the issue, avoiding precise facts or figures." "Offer a high-level summary that captures the essence without being too specific."

# Using The Re-Read Prompting Technique Is Doubly Rewarding For Prompt Engineering

**Lance Eliot** Contributor Jul 6, 2024,08:00am EDT

Using the re-read prompting technique can boost your generative AI responses.

## Why Invoking A Second Pass Can Be Noticeably Handy

Let's start by considering why humans sometimes opt to re-read a question. To quickly clarify, I do not intend to anthropomorphize AI and thus please do not falsely equate what humans do cognitively with what generative AI is doing. I only am drawing a distant parallel to showcase facets that can arise when re-reading content. In a moment, I'll be explaining the details associated with the mathematical and computational aspects of re-reading that come into play within generative AI, thanks.

Back to the human considerations. ~ Suppose that someone hands you a piece of paper that has a written question on it. You would undoubtedly read the question before trying to come up with an answer (I hope you would read it first or else your derived answer might not be pertinent to the question being posed). Upon having read the question, you regale the person with a divine answer. Good for you.

*Imagine though that the question is a hard one. You might be tempted to read the question a second time.* This re-reading of the question might aid you in better understanding the nature of the question. This in turn might aid you in coming up with a better answer than if you had tried to answer based on your first or initial reading of the question.

So far, so good. ~ *There is also the factor that by having read the question a second time, your first reading might have primed your mind.* It got your mental juices going and into the ballpark of whatever the question entailed. The second reading will therefore allow you to more readily gauge the context and nuances of the question. In a sense, you have reduced the element of surprise that customarily occurs when first encountering a question.

You don't necessarily re-read questions all the time. If a question is easy, there probably is little value in re-reading it. The same goes for a very short question. The odds are that a notably brief question doesn't need a re-reading either. Okay, we've established that re-reading a question can be advantageous. A trade-off exists between the effort and time needed to do a re-reading versus the added benefits of doing a re-reading.

Here's the deal. ~Most of today's popular generative AI apps tend to do a single pass on an entered prompt. The text is essentially examined on a unidirectional basis, moving from left to right. Word by word, the sentence or sentences that you entered are examined via various mathematical and computational processes. As a side note, there are exceptions to this rule of thumb, such that some generative AI apps work on a bidirectional basis or otherwise examine an entered prompt in a more back-and-forth manner.

Let's focus on the unidirectional ones. ~ If a question that you entered is being parsed on a one-time basis, there is a chance that some subtleties might be inadvertently glossed over. A second pass would potentially aid in doing a further detailed examination of the question. Kind of a clean-up of whatever might have been missed or miscalculated about the question.

A re-reading could also augment the AI contextual facets. The initial context that was calculated in the first pass can feed into what is computationally happening during the second pass. For example, your question might be asking about Abraham Lincoln. The internal retrieval of data about the life of Lincoln is instantiated on the first pass. During a re-reading or second pass, the retrieved context is already there and raring to go.

## Doing Some Unpacking About Re-Reading

I think you can readily discern that there is a possible added value to invoking a second pass of your entered prompts. You could simply tell the AI in your prompt to go ahead and re-read the prompt. That's all you need to do. Easy-peasy. Just give an explicit instruction and then leave the driving to the generative AI.

When I mention this during my classes and presentations on prompt engineering, there invariably will be eager attendees who raise their hand and ask if it might be useful to always force generative AI to do a re-reading. This could be a custom instruction that you have the generative AI always use, see my explanation about how to set up permanent standing defaults or instructions in AI at [the link here](#).

I vote that you not blindly opt to always make use of a re-reading action. The re-reading will take a bit longer for the prompt to get processed by the AI. In addition, if you are paying for the use of generative AI, you will likely end up having to pay for those extra processing cycles. For background about how to reduce latency and speed up your response time while using generative AI, see my discussion at [the link here](#).

Is the delay in response and the added dollar cost worth doing a re-read? ~ By and large, if you typically enter questions or prompts that are short and sweet, the use of a re-reading action is not going to get you any manifest bang for the buck. You will needlessly be consuming added time and cost. The generated response to your question is probably going to be about the same as not having done a re-read.

The other side of that coin is that *many people don't realize they can use a re-read as a prompting tactic or technique*. They are either unaware of the idea, or they have learned about it but forget to use it when needed. That's not wise either.

Overall, there is the **Goldilocks** principle at work. Do not use a re-read all the time (unless you have a solid reason for doing so). Don't forget to use a re-read when the re-read is pertinent to the matter at hand. Try to make use of a re-read at the right time and right place. The porridge should neither be too hot nor too cold.

*I urge that anyone seeking to be a prompt engineer or who considers themselves to be a prompt engineer ought to have the re-read instruction in their back pocket and include the re-read in their prompt engineering skillset.*

There is something else you should consider too. ~ Realize and keep at the top of your mind that you can combine various prompting techniques. Allow me to elaborate. ~ You might know that one of the most common prompting techniques entails getting generative AI to proceed on a step-by-step basis and explaining what the AI is doing during the processing of your prompt. This is known as invoking a chain-of-thought (CoT). You can do this by merely telling the AI to process the prompt in a stepwise approach.

Can you **combine the use of chain-of-thought** with the use of a re-reading instruction? Absolutely. ~ Here’s why you might be tempted to do so. ~ When you use a re-read by itself, you won’t necessarily see what is happening. The only thing you will see is the result or generated answer to your question. It could be that the answer turns out to be incorrect or somewhat off-target. You won’t particularly know why. If you had included an instruction to do a stepwise process or CoT, the odds are that you would be able to inspect the displayed explanation and discern where the AI went awry. Happy face.

I have a question for you about this. ~ Should you always combine a re-read with a CoT instruction? In my experience, you are not going to get any useful added value beyond just one re-read. The chances of having the AI pick up on additional subtitles after the first re-read seem mighty low. I am not saying that this might be a payoff in very rare or unusual cases, and only suggesting that as a rule of thumb stick with one re-read in a prompt. You might think of re-reading as something that becomes problematic if overused in a single prompt.

Research On The Re-Read Gets The Spotlight

*“To enhance the reasoning capabilities of off-the-shelf Large Language Models (LLMs), we introduce a simple, yet general and effective prompting method, RE2, i.e., Re-Reading the question as input.” “Crucially, RE2 facilitates a "bidirectional" encoding in unidirectional decoder-only LLMs because the first pass could provide global information for the second pass.” “We then evaluate RE2 on extensive reasoning benchmarks across 14 datasets, spanning 112 experiments, to validate its effectiveness and generality.” “Our findings indicate that, with the exception of a few scenarios on vanilla ChatGPT RE2 consistently enhances the reasoning performance of LLMs through a simple re-reading strategy.”*

The research examined the use of the re-read instruction by utilizing various datasets of questions and answers. They wanted to see whether the re-read would increase the chances of getting correct answers.

For example, consider this question: *“Roger has 5 tennis balls. He buys 2 more cans of tennis balls. Each can has 3 tennis balls. How many tennis balls does he have now?”* You would need to figure out how to arrive at a definitive answer. You either arrive at the correct number of tennis balls or you do not. The beauty of these types of questions when doing prompt engineering experiments is that you can easily keep tabs on how well the AI does.

Consider this. You feed in a bunch of closed-end questions and grade how many were answered correctly versus incorrectly. You then feed in similar questions and include the new prompting instruction that you want to test out. After doing so, you tally how many of those questions were answered correctly versus incorrectly. Voila, you can compare how things went without the new instruction and how things went with it included.

Per the experiments performed, the research paper suggests that you can suitably add a re-read instruction to improve the chances of having the AI determine correct answers. They also advocate writing the question in its entirety a second time, rather than merely obliquely referring to the question as posed. Let me show you this. Go ahead and look at this example (I bolded the part that says to read the question again):

Q: Roger has 5 tennis balls. He buys 2 more cans of tennis balls. Each can has 3 tennis balls. How many tennis balls does he have now? Read the question again: Roger has 5 tennis balls. He buys 2 more cans of tennis balls. Each can has 3 tennis balls. How many tennis balls does he have now?

Observe that the stated question was repeated in its entirety. Do you need to do that? In my experimentation, which I’ll get to shortly, I found that it is wise to repeat the question verbatim. If you don’t do so, there is a slightly increased chance that the AI will not dutifully re-read the question. It might take a shortcut. Thus, you are better off by doing a copy-paste of the stated question rather than trying to just indirectly refer to it. Another twist is that I tried the “Read the question again” and then I repeated the question, but I purposely changed the question. Why? I wanted to see if the AI would detect that I hadn’t faithfully repeated the question. Indeed, I got caught. I urge you to faithfully show the same exact question and not play games with the AI. The generalized template advised by the researchers is to provide your input query, followed by “Read the question again” and then followed by the repeated query, per this overall format

Q: {Input Query}. Read the question again: {Input Query}

That is the shortcut template version. A longer version is coming up next herein, so hang in there.

Deeper Into The Nuances Of The Re-Read Prompt

You might recall that I had forewarned not to try and repetitively use the re-read in a single prompt. Let’s see what the researchers have to say (excerpts). *“An overarching pattern emerges across all models: performance improves until the number of re-reads reaches 2 or 3, after which it begins to decline with further increases in question re-reading times.”* (ibid). *“The potential reasons for inferior performance when reading the question multiple times are two-fold: i) overly repeating questions may act as demonstrations to encourage LLMs to repeat the question rather than generate the answer, and ii) repeating the question significantly increases the inconsistency of the LLMs between our inference and pretraining/alignment (intuitively in the learning corpora, we usually repeat a question twice to emphasize the key part, rather not more).”* (ibid). The bottom-line is don’t repeat the re-read while within a single prompt.

Moving on, I mentioned that you can use a re-read in conjunction with other prompting strategies. The researchers opted to explore the re-read with the chain-of-thought (CoT) approach (excerpts). *“We further conduct experiments to examine the influence of RE2 within the context of CoT prompting. “Consequently, RE2 demonstrates strong generality and compatibility with most thought-eliciting prompting methods, including CoT.” “It’s noteworthy that, in general, question re-reading consistently improves reasoning performance compared to the standard CoT prompting without question re-reading (Po).”* You can see that there were synergies between the two techniques.

Give that a bit of contemplation. The interesting angle is that though at this moment we are examining the re-read instruction, you can turn your head in the other direction and think about the chain-of-thought instruction. You are to start thinking of the CoT as being augmented with the re-read. I say this because you might at this time mindfully use the CoT with the re-read, but I am saying to also extend your thinking to consider using the re-read when you are already intending to use the CoT. A mind-bender, for sure.

Q: {question}. Read the question again: {question}. #Answer format instruction#. A: Let’s think step by step.

Typical kinds of answer format instructions that might be used (in my words):

Your answer should be either yes or no.  
Your answer should be a date in the format of mm/dd/yyyy.  
Your answer should be an essay of no more than three paragraphs in length.  
Your answer should be a rounded integer.

*“How much wood could a woodchuck chuck if a woodchuck could chuck wood? Read the question again: How much wood could a woodchuck chuck if a woodchuck could chuck wood? Provide the answer as a numeric value indicating the number of cords of wood. Let’s think step by step.”*

Using ChatGPT To Explore The Nature Of Re-Reading

I will next proceed to examine further the nature of re-reading as a specialized prompting technique. This will consist of a series of dialogues with ChatGPT. ChatGPT is a logical choice in this case due to its immense popularity as a generative AI app. An estimated one hundred million weekly active users are said to be utilizing ChatGPT. That’s a lot of people and a lot of generative AI usage underway.

A few quick comments before we launch into using ChatGPT. If you are going to try to do the same prompts that I show here, realize that the probabilistic and statistical properties will likely produce slightly different results than what I show here. That’s the nature of generative AI and how it is devised.

I opted to start by asking ChatGPT about the overall nature of the re-reading prompting technique.

As an aside, whenever you are starting a conversation with generative AI, I recommend as a prompt engineering technique to begin by asking a question that will establish if the AI has been data-trained on the topic at hand. A generated response that is vacuous will give you a heads-up that you might be barking up the wrong tree with that generative AI app.

“Does telling generative AI to re-read a prompt provide any added value regarding potentially having the AI do a better job of answering a question that is within the prompt?”

Research studies indicate that the benefits of having generative AI re-read a prompt can be valid, particularly in the context of bidirectional encoding enhancing the generative process.

**Bidirectional Encoding Benefits:** In natural language processing, bidirectional encoding allows models to consider the context from both directions (left-to-right and right-to-left). This can provide a more comprehensive understanding of the text, as the model can utilize global information from the entire prompt.

**Enhanced Contextual Understanding:** Re-reading a prompt can enable the model to better grasp nuances, context, and relationships within the text. This can lead to more accurate and contextually relevant responses, especially for complex or detailed prompts.

**Augmented Processing:** The process of re-reading can serve as a form of reinforcement, allowing the model to refine its understanding and potentially correct any misinterpretations from the first pass. This can be especially useful for decoder-only models, which might benefit from a secondary pass to incorporate broader contextual information.

**Empirical Evidence:** Research studies providing empirical evidence of improved performance through re-reading suggest practical benefits. These studies likely involve experiments demonstrating that iterative processing can enhance the quality of responses.

The response by ChatGPT was reassuring. You can see that ChatGPT was able to stridently respond to what a re-read prompt consists of. In addition, happily, the points made are akin to the points I made earlier. If the points made by ChatGPT had differed radically, one of the two of us would have somehow gone askew. Next, I will ask ChatGPT about the tradeoffs or potential downsides of the re-read. Take a look:

“Is the use of a re-reading prompting strategy always a surefire payoff?”

Instructing generative AI to re-read a prompt does not inherently improve the quality of its response. Here are some points to consider.

**Single Pass Processing:** Most generative AI models, including ones like GPT-4, process the prompt in a single pass. They generate responses based on the input provided without re-reading or reconsidering the prompt in the middle of generating a response.

**Contextual Understanding:** These models are designed to understand and retain the context provided within a single prompt. If the prompt is clear and well-structured, the AI can generate a coherent and relevant response.

**Clarification:** If a response seems off-track or incomplete, providing additional context, clarifications, or rephrasing the question can be more effective than simply asking the AI to re-read the prompt.

**Prompt Quality:** The quality and clarity of the prompt are crucial. A well-crafted prompt with specific details and clear instructions is more likely to yield a better response than one that is vague or ambiguous.

Good, the answer again matches what I’ve been saying during this discussion.

As a recap, you might or might not find much gain from a re-read instruction. The nature of the question and the mechanizations of the generative AI are crucial determiners.

Let’s get ChatGPT to give us an example of when doing a re-read isn’t going to be a boost.

“Show me an example of a question that doing a re-read on will not add any benefit to producing a better answer.”

“Let's consider a simple factual question that does not benefit from a re-read.”

“Question: "What is the capital of France?"

“First, I'll provide a response as if generated in a single pass.”

“Response: "The capital of France is Paris.”

“Now, let's undertake a re-reading process to see if it improves the response.”

“Enhanced Response: "Paris is the capital of France.”

“For simple factual questions like "What is the capital of France?", re-reading the prompt does not significantly enhance the response. The initial response already fully addresses the question, and additional passes do not provide substantial added value. This demonstrates that re-reading a prompt is more beneficial for complex or multi-faceted questions, rather than straightforward factual queries.”

Asking about the capital of France is a very straightforward question. The question itself is short in length. There isn’t any ambiguity involved in the question. A first read is probably all you need to do. A second read would seem like unnecessary overkill. I’ll shift gears and have ChatGPT come up with an open-ended question that then tells the AI to do a re-read. Let’s see what happens.

“Show me an example of an open-ended question that would be better answered via using a re-read prompt. Go ahead and then process the question, explaining how the re-reading activity has helped produce a better answer.”

“Let's take a complex question that involves multiple layers of information and requires a comprehensive response.”

“Here's an example: "How does climate change impact global food security, and what measures can be taken to mitigate these effects?"

“First, I'll provide a response as if generated in a single pass.”

“Response: "Climate change affects global food security by altering weather patterns, leading to extreme weather events, and changing growing seasons. These changes can reduce crop yields, disrupt supply chains, and increase food prices. To mitigate these effects, measures such as adopting sustainable farming practices, improving water management, and developing climate-resistant crops can be implemented.”

“Now, let's undertake a re-reading process to improve the response.”

“Enhanced Response: "Climate change significantly impacts global food security through several mechanisms. Firstly, it alters weather patterns, resulting in more frequent and severe extreme weather events such as droughts, floods, and hurricanes. These events can devastate crops, reduce yields, and disrupt supply chains. For instance, prolonged droughts in regions like Sub-Saharan Africa have led to significant decreases in maize and wheat production, which are staple foods for millions. Additionally, changing growing seasons can lead to mismatches in planting and



harvesting times, affecting crop productivity. For example, warmer temperatures can cause some crops to mature too quickly, leading to reduced quality and yields.”

“By re-reading the prompt and refining the response, the answer becomes more detailed and nuanced. It covers specific examples, provides a broader context, and offers a more comprehensive discussion of mitigation measures. This approach ensures that the response addresses the complexities of the issue more effectively.”

I am a bit on the fence about that example.

The answer based on the re-read doesn’t seem to be that much better than the initial answer based on the first read.

One aspect to keep in mind is that I purposely made use of an open-ended question. The difficulty with assessing open-ended questions is that we don’t have a definitive answer like we do when asking a closed-ended question. This makes judging the answers more speculative.

Finally, I’d like to have ChatGPT provide a roundup of purported advice on the use of the re-read instruction.

Hold your breath and see what ChatGPT said.

If you were going to give advice to a prompt engineer who has just found out about the potential use of a re-reading prompting strategy, what advice would you give them?

Here’s a structured approach to guide a prompt engineer:

**Evaluate the Complexity of Queries:** Determine if the typical queries handled by the model are complex or multifaceted. Re-reading is most beneficial for detailed and layered questions.

**Identify Critical Use Cases:** Identify scenarios where accuracy and comprehensiveness are crucial, such as in customer support, medical information, or legal advice.

**Analyze Resource Availability:** Assess the available computational resources and budget. Re-reading requires more processing power and may increase operational costs.

**Evaluate Latency Tolerance:** Determine the acceptable response time for your application. Re-reading will increase latency, which might be a trade-off in time-sensitive applications.

**Iterative Approach:** Start with a pilot implementation to understand the performance and benefits. Gradually scale up based on initial results and feedback

**Dynamic Decision Making:** Implement a dynamic strategy where re-reading is only triggered for complex or ambiguous prompts. Use simpler, faster processing for straightforward queries.

**Performance Monitoring:** Continuously monitor the performance and cost implications. Use metrics such as response time, accuracy, user satisfaction, and resource usage.

Using a re-reading prompting strategy can significantly enhance the quality of responses for complex queries. However, it comes with increased computational costs and latency. A balanced approach that dynamically applies re-reading based on query complexity, coupled with continuous monitoring and feedback, will help in optimizing the benefits while managing the costs effectively.

The response by ChatGPT was pretty on-target. We saw that you need to consider the circumstances underlying the use of the re-read instruction, and you need to weigh the potential added benefit against the costs or possible downsides.

### Conclusion

There is a notable quote by C. S. Lewis, a famous British author and literary scholar about the value of re-reading: “*We do not enjoy a story fully at the first reading. Not till the curiosity, the sheer narrative lust, has been given its sop and laid asleep, are we at leisure to savor the real beauties.*” I would ascribe that sense of humankind value to his quote and certainly find that when I re-read things, it often has a momentous beauteous reflective quality. Of course, there are times when I just wasn’t paying close attention at the get-go and the re-read gets my mind back into proper focus. That’s a pragmatic way to value the re-read.

Here are my Top 10 of the best of the best for prompt engineering techniques:

- (1) **Chain-of-Thought (CoT) Prompting** (listed above as L-10).
- (2) **Chain-of-Verification (CoV) Prompting** (listed above as L-12)
- (3) **Emotionally Expressed Prompting** (listed above as L-17)
- (4) **End-Goal Prompting** (listed above as L-18)
- (5) **Flipped Interaction Prompting** (listed above as L-21)
- (6) **Generating Prompts Via Generative AI** (listed above as L-22)
- (7) **Mega-Personas Prompting** (listed above as L-30)
- (8) **Retrieval-Augmented Generation (RAG) Prompting** (listed above as L-39)
- (9) **Step-Around Prompting Technique** (listed above as L-45)
- (10) **Trust Layers For Prompting** (listed above as L-49)

I am saying that you should know each of those by heart.

They should be like falling off a log. Roll up your sleeves and get to work. Commit them to memory. Use them daily, as needed.

Allow me a moment to say a bit more. I frankly agonized at coming up with a Top 10. All fifty on the list are, to me, of great merit and well worth knowing. It saddens me to see some in the list of fifty that did not make my Top 10. They are all deserving of being in the Top 10.

So, please do not treat the remaining forty as though they are inconsequential. It would break my heart. If you can, get to know them all.