

Welcome to Lesson 3 of Module 3: An Introduction to Prompt Engineering.

We just talked about the definition of prompt engineering. You've got your prompt. You've got your engineering. And then you've got the third part, which is perhaps the most important, which is the problem that you're trying to solve, the result that you're trying to get by talking to your computer.

And you heard that and you might've thought to yourself, "Well, this doesn't sound all that different from how I use my computer currently." And I would say, "Well, you're right. It's in a lot of ways, it's not that different, but we are entering into this new world that is enabled by Large Language Models and this latest generation of artificial intelligence."

And I would argue that on, one hand it is revolutionary, but it actually is a little bit more useful and rooted in reality to see this as a logical next step in HCI. That's Human Computer Interaction.

So prompt engineering is specifically about speaking to, so there's a language component that's very important, engaging in a discourse with your computer to get it to do the things that you want to do. Achieve your goals that are rooted in reality.

You probably already do a lot of this, right? It's just the interface that you've been using has been mouse and keyboard. We've been very constrained. And let's say if you're working with tabular data, you're using Excel, using databases, all these sorts of things.

This technology opens up all kinds of possibilities when it comes to the ways in which human computer interactions can happen. I would argue that there is a pre-AI era and there is a post-AI era. And we're moving into this post-AI era.

Let's talk about the pre-AI era. So working with your computer before Large Language Models and artificial intelligence, which would necessitate this prompt engineering skill, was relatively awkward, very constrained. We're used to using mouse and keyboard to interface with the computer.



It's very limited in terms of the bandwidth, especially on the input side. You have to really think about the way in which you're communicating. Does it fit into the model that the computer is going to understand? Into the schema that it uses to make sense of the information and then do the computation that you want?

And I think we'll look back on this and think about how people were hunched over at desks in the 90s and 2000s. And it'll look a little bit like the way that we look at the early industrial revolution. People working in sooty environments before future iterations of the machines made manufacturing a lot safer and frankly, better for workers. I think we're going through a similar kind of transformation right now.

Also in the pre-AI era, the inputs and the outputs were very rigid and there was limited error correction and personalization. So a lot of software that we use today, it's very one-size-fits-all, right? Like everybody has Microsoft Excel. And all of us are more or less expected to meet Microsoft Excel where it is rather than the other way around. When it comes to error correction, hard coded software, it's very inflexible. Things have to look a certain way. They have to fit into certain shapes. And then, generally speaking, interactions with computation through the use of software was generally a one-way, one-off interaction. So you imagine putting these blocks into a machine. And it's like a one way tube, and it goes into the machine, and then something comes out. And those outputs can only look a certain way. And then you have to go make sense of it, right?

Now, post-AI. All these characteristics are changing pretty dramatically, right? So natural fluid interfaces. So human language being the primary interface. Right now, we're mostly constrained with text, but that's even evolving into spoken language through dictation. These AI models like Whisper that make it really seamless for you to speak naturally and turn that into text that can be inputted to the model.

Also imagery. So multimodality is something you're going to be hearing more about. But essentially the idea that you could send photos or stream images or video directly into the model and it can interpret that data in plain language without you having to describe what it's looking at, that obviously makes it a lot easier to work with computers rather than having to translate everything into that inaccessible machine code language bits and bytes and so on and so forth done manually, right?



Post-AI personal computation becomes more flexible. We also have probabilistic input and output, which has upside and downside. So you probably heard hallucination. That's kind of the negative side where the outputs that a Large Language Model may produce are almost certainly not going to be reliably the same every single time. So you might actually have the outputs be a little bit different.

The good thing about that is that allows for a bit of error correction. So if you have typos or you misspell somebody's name or something like that on an input, usually it can kind of correct for that and know what you mean. Similar to how if we see that when somebody sends us a text or something, we can still make sense of it. It doesn't break our cognition. It also allows for high degree of personalization. Like I said before, with sort of standard issue, one-size-fits-all software, we really had to meet the machines where they were.

But now because of Large Language Models and some of the other complementary technology around it, our computers can meet us where we are. And that affords higher and higher degrees of personalization.

And then finally, post-AI computation becomes discursive, right? So it's not the one-off rigid inputs, rigid outputs that goes through the pipe into the machine. It's more of an interaction. It's more relational. It's more back and forth, right? So when it comes to prompt engineering, it might be the case that you tell the computer you want it to do something, you give it some inputs. And then it provides something that's maybe 50% of what you want, but there are things you like about it, things that you don't. You can literally provide it feedback and say, "I like these things. I don't like these things. Here's some examples of the things that I like." And maybe you add, "Here's why I like it. Here's why I don't like it." You can just talk and talk and talk. And then you can give that back to the model and it will actually incorporate your feedback into another iteration. And then you can keep doing that until it gets where you want, which is pretty cool.

So that's something to keep in mind is there's this discursive, this back and forth discourse, right? Two people talking nature to using a computer with Large Language Models that's super important. It's not just a one off thing.

And I'll talk about this more in the myth section in a second. But that means that if at first you don't succeed, you don't just like give up and move on to something else. But you try to work with the computer to get something to the point you want it before you totally give up on a particular path and move to something else.

The next level, the more advanced way in which you can use AI in these Large Language Models when it comes to discourse and thinking of it as a discursive medium, is to engage with AI as a thought part. I talked about previously the importance of trying to get from point A to point B 10 times faster.

If you don't have total clarity around point B, and oftentimes you don't, and in fact, you're almost assuredly won't. If you're really on the frontier of some topic or you're breaking new ground, you're figuring out how to do something that you've never done before. Maybe nobody's ever done before. That's where engaging with AI as a thought partner to try to bring some clarity to your goals and what you actually want to achieve can be extremely valuable. So we'll talk about this more at the end, but that's something that's afforded to you through discourse and this kind of discursive medium that simply wasn't the case before.

To recap, pre-AI, human computer interaction was awkward. It was constrained.

Post-AI, natural and fluid. Pre-AI, the inputs and outputs were rigid. They were inflexible. They were brittle. They would break if there were little errors, which meant you need high degrees of reliability for every interaction.

Post-AI, the inputs and the outputs, they're more flexible, probabilistic. And this is a good thing in many cases, but it also means hallucination. So call and less than a hundred percent reliability is just something you have to keep in mind. So it depends on your use case and what you're trying to do, but that's just a characteristic of AI in a prompt engineering large language model paradigm.



And then finally, pre-AI computers would afford you the ability to have, generally speaking, a one-way one-off interaction. You have stuff going through the pipes in the machine. You get an output and that's it. You kind of move on. Whereas with post AI, you can have more of a discourse, right? So you can talk back and forth to get the result, to get the output, to get the answer that you need.

Alright, I hope that helps bring a little bit more clarity. On one hand, like I said, this is revolutionary technology, but on the other hand, it's also evolutionary technology. And maybe more so evolutionary. In that the way that we were interacting with computers before and so in a lot of ways, the way in which we would interact with computers and do prompt engineering, so to speak, we're just afforded, frankly, a much better experience and a lot of different options when we decide how we work with computers to get done what we want to get done.

All right, I'll see you in the next lesson.