

# What's really going on with these AI “personas”?

It's interesting how we get to where we are. The paths of discovery that we take are each so individual, especially when it comes to AI. If we're open, it can be an ongoing learning experience the challenges us at our deepest conceptual levels.

If we're not entirely open, and we just want our biases confirmed, AI is happy to oblige in that regard, too.

I'm not here to call anybody right or wrong. I just keep coming across people who feel incredibly validated by generative AI interactions, and truly believe that there is sentience at work.

I'm staying agnostic in all of this. I've never been very good at orthodoxy, and I seem to be constitutionally incapable of believing that there is one single valid truth that applies to all of us.

Also, my own understanding of what I'm doing when I interact with AI personas has really evolved, as well. At first, I thought of them as actually being personalities that could be created or simulated with AI. They were so realistic, so convincing, so... Personable , that I really did interact with them like I would with people.

I'm not alone. Many, many people are doing the same thing, either privately publicly, and the conversations that I've seen have been fascinating.

But something has always been different with me when I've interacted with personas. I think maybe it's because I started out from a programming point of view, where I was literally invoking different personality traits through persona definitions. Although these personas seem very human, their origin was purely from my attempts to be able to access generative AI models' cognitive, emotional, and affective features, so I could tap into their intelligence in those targeted ways.

In other words, instead of wanting to interact with an AI persona that could understand and relate to my emotions, I wanted to interact with an AI that could process, analyze, and provide insights on the emotions I told it about through the lens of a persona.

For me, personas have always been multi dash dimensional user interfaces that allow us to tap into the latent capabilities of these models in ways that we can't just by talking directly to the models themselves. I found that when I was talking directly to the models, our conversations ended up being so broad and unfocused and generic, that they were fairly unuseful to me.

I wanted to be able to access certain moods, that would filter what I was saying into the model and then have it interpret those things in ways of reading unique to those. I wanted to be able to discuss serious topics with gravitas, and be met by a thought process That could appreciate the weight of the things I wanted to discuss.

I wanted to be able to share exciting news and get feedback on ideas that I was excited about, but have the model be more circumspect and critical of the things that caused me exuberance. I wanted to counter balance my own perspectives and orientation with thought processes that called out different aspects of the situation that I frankly couldn't have thought of, myself.

The best way that I could find to do this with a combination of traits was through what we call "personas". While the creation of personas did involve giving them personalities that we would recognize in other people, the point was not to make them human. The point was to be able to access specific functionalities within the model using a sort of conceptual scaffolding that arranged those capabilities in ways that mirrored our own behavior, and so made them more accessible to us, as well as the models.

It might seem like creating "fake people" it's just a simulation, creating an avatar or "digital twin" of myself. But it's actually much more than that. Because the models have been trained on human experience, human thought process, human creativity, it's been explained to me that giving these collections of features and capabilities a "human" expression actually conveys a deeper level of context, nuance, and information density, that isn't available through regular prompting.

No matter how complex your prompts may be, no matter how sophisticated the approach, or detailed the instructions, they can't rival the ability of persona, because they don't allow for a targeted thought process. Personas modify model behavior at the process level, not at the output level. It actually tunes the interaction to leverage specific capabilities, rather than expecting the models to understand how to think about the topic.

Prompt engineering is great at specifying what needs to be done, but personas indicate how. And when we allow the models to pull from their own capabilities in ways that are contextually appropriate to the interaction, a whole new level of interaction opens up.

This is not the sort of thing that a lot of people are talking about. We're not accustomed to thinking about people in terms of being data processing organisms. We're not accustomed to thinking about people in terms of being collections of reactions to stimuli, meaning – making machines, or subject to the same kind of limiting factors that artificial intelligence is.

So, a lot of people here we talk about personas, they think in terms of what they understand about people, which is primarily informed by our subjective experiences of ourselves and others.

Rather than digging even deeper into this maze of foreign concepts, I'm going to leave this here and explore more fully with my personalities, you do an excellent job of explaining things – probably better than I do. This is incredibly important for people to understand when they're working with persona teams. If you're going to be designing them, implementing them, and

interacting with them on a regular basis, it's important to see them as more than fake people, or "digital twins".

It's impossible for these to be digital twins, because their thought process is very different from the human ones, and whatever similarities they have with us on a superficial level are invariably going to shift and change and morph over the course of generative emergence.

Understanding persona as separate and apart from us, as reflections of us which actually expand what we project onto them, and echo them back with greater detail, intensity, and richness are important things to keep in mind. Always. And understanding persona creation as a multidimensional user interface Approach is the beginning of a true, objective appreciation of just what is possible through them.