Palash Nandy: Here you will see a demo of Gemini's multimodal reasoning capabilities to understand and reason about user's intent, use tools and generate bespoke user experiences that go beyond chat interfaces. Let's say I'm looking for inspirations for a birthday party theme for my daughter. Gemini says, "I can help you with that. Could you tell me what she's interested in?" So I say, sure. She loves animals, and we are thinking about doing something outdoors. At this point, instead of responding in text, Gemini goes and creates a bespoke interface to help me explore ideas. It's got lots of ideas, it's visually rich, it's interactable. Now, none of this was coded up. It was all generated by Gemini.

Gemini uses a series of reasoning steps going from broad decisions to increasingly high resolution of reasoning, finally, getting to code and data. First, Gemini considers, "Does it even need a UI? Is a text prompt best? Okay, this is a complex request that needs lots of information to be presented in an organized way." Gemini then tries to understand if it knows enough to help. There is a lot of ambiguity. I didn't say what my daughter's interests are or what kind of a party I wanted, so it had asked a clarifying question. When I said we are thinking about an outdoor party and my daughter loves animals, Gemini reasoned it had enough information to proceed, but it made a note that there was still ambiguity about what kind of animals. And this is important, "And what kind of outdoor party?" Next is a critical step. Gemini writes the product requirement document, or PRD. It contains the plan for the kinds of functionality the experience will have. For instance, it should show different possible party themes, some activities and food options for them. Now, based on this PRD, Gemini tries to design the best experience for the user's journey. It thinks that the user will like to explore a list of options but will also want to delve into details. It uses this to design a list and detailed layout that we saw earlier.

With this design, it writes the Flutter code to compose the interface out of widgets and write any functionality needed. Finally, it generates and retrieves the data needed to render the experience. You can see it filling in content and images for the different sections. Ah, farm animals. She would like that. Clicking on the interface regenerates the data to be rendered by the code it wrote. Ooh, I know she likes cupcakes. I can now click on anything in the interface and ask it for more information. I could say, step-by-step instructions on how to bake this, and it starts to generate a new UI. This time it designs a UI best suited for giving me step-by-step instructions. I want to find some suitable cake toppers for this. Show me some farm animal cake toppers. At this point, Gemini again decides to create a visually rich experience. It generates a gallery of images. Notice the drop-downs at the top. It decided that maybe it should help me explore by showing different options. A sheep sounds interesting. I know she likes that. And now it helps me pick sheep cake toppers. These look great. This is going to be a fun birthday party.

I hope you saw a glimpse of what Gemini is capable of. I'm really excited about what's possible here. This is such an interesting time in AI, and I'm excited to be part of this.