

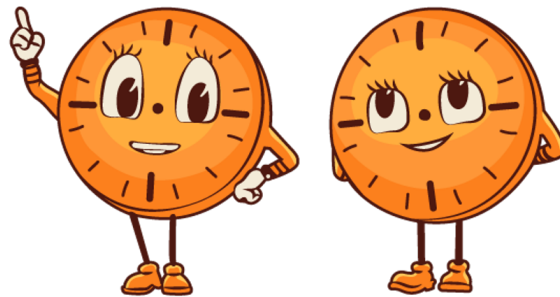


## Giving ChatGPT a virtual body

Picture this: ChatGPT, not just as a text-based assistant, but embodied as a virtual companion you can see and interact with in a mixed or virtual reality environment. This smart agent can engage with you, help you navigate complex scenarios, and even play games, all by understanding the world around you and making intelligent decisions.

We aim to create an innovative pipeline that harnesses the power of Large Language Models (LLMs) to bring virtual humans or objects to life. These entities will interact seamlessly with the world, making decisions about how to move, talk, and engage in various activities.

To achieve this goal, we plan to utilize Head-Mounted Displays (HMDs) equipped with sensors that can collect comprehensive data about the surroundings in the virtual world and also capture detailed information about the user, such as facial expressions and eye movements. This rich dataset will provide



clear inputs about both the environment and the user's interactions. Next, we will employ Multimodal Long-Context Vision-Language Models (VLMs) to process this data, enabling the virtual agent to understand the context and make informed decisions. This method will allow the agent to act accordingly, creating a responsive and engaging output that enhances the user's experience.

In this project, you will have the opportunity to collaborate with [Magic Leap](#).

### **Requirements:**

Knowledge in Deep Learning, or solid background in Machine Learning.  
Experience with PyTorch and Unity is an advantage.

**Interested? Please contact us for more details!**

### **Contact**

- Ard Kastrati: [kard@ethz.ch](mailto:kard@ethz.ch), ETZ G61.3
- Dushan Vasilevski: [dvasilevski@magicleap.com](mailto:dvasilevski@magicleap.com), Magic Leap