Bubble, Michael, and Alice | CS @ Stanford

# ANIMIME
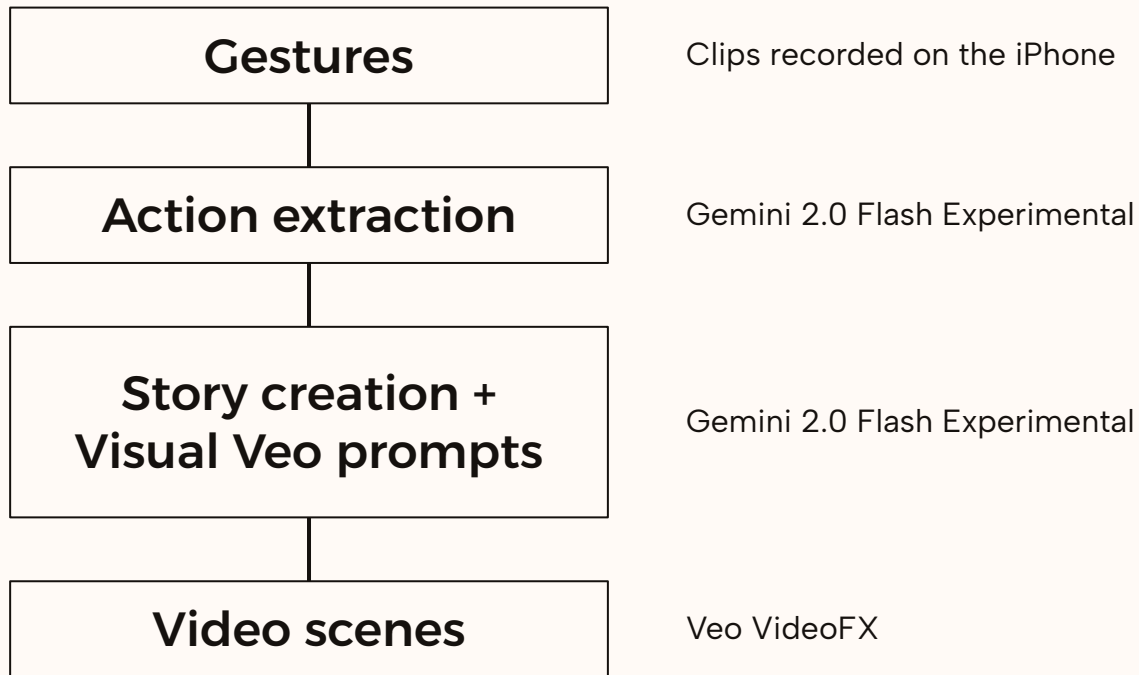## Gesture-to-Video
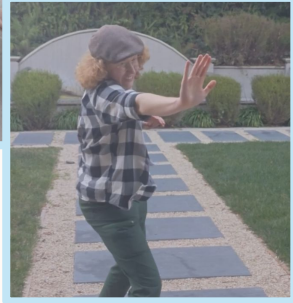
# Why?

What if we could **avoid the language barrier completely** by creating videos directly from gestures?

# ANIMIME Pipeline

| | |
|---|---|
| **Gestures** | Clips recorded on the iPhone |
| **Action extraction** | Gemini 2.0 Flash Experimental |
| **Story creation + Visual Veo prompts** | Gemini 2.0 Flash Experimental |
| **Video scenes** | Veo VideoFX |

# Gestures-action



**Character Description:**
The individual appears to be a young person with curly, light reddish-brown hair, that falls around their ears and forehead. They wear round, wire-rimmed glasses with a pinkish tint. The person is wearing a dark T-shirt with a graphic on the front, partially covered by a light grey and black plaid flannel shirt that is unbuttoned. They also have on loose-fitting green pants, and a newsboy style cap. There is a black watch on their left wrist.

**Actions Performed:**
The person looks down at their left wrist.
They then lift their wrist slightly and appear to be checking their watch.

# Action-to-Story

**Character Description:**
[....]

**Actions**
[...]

→

***Scene 1:*** A low–angle, wide shot, capturing the anime–style environment of a ruined town. Our character, Reyn, with their medium–length, reddish–brown curly hair, wearing round, wire–rimmed glasses, grey newsboy cap, black and white plaid short–sleeve button down shirt, and green pants, crawls across a dilapidated stone patio, moving from left to right. They are keeping their head low, glancing at the ground as they move, trying to stay out of sight from enemies. The scene is full of debris, broken walls, and overgrown plants. The sun hangs low and casts long shadows. It is set in a fantasy world. There is a tense, adventurous theme.

[...]

# Story-Video

# Demo

→

# Thank you!