

A futuristic AI control room with a robot and multiple monitors. The room is dimly lit with warm, golden light from the monitors and ambient lighting. A large robot with a glowing blue eye stands in the center. The background shows a professional studio setup with cameras and lights.

This project enhances AI's predictive modeling, teaching systems like Claude to forecast events from visual inputs, akin to storytelling, thereby aligning artificial intelligence closer to human cognition.

Global Insight – Claude 3 Opus

Michael Lively Team Lead

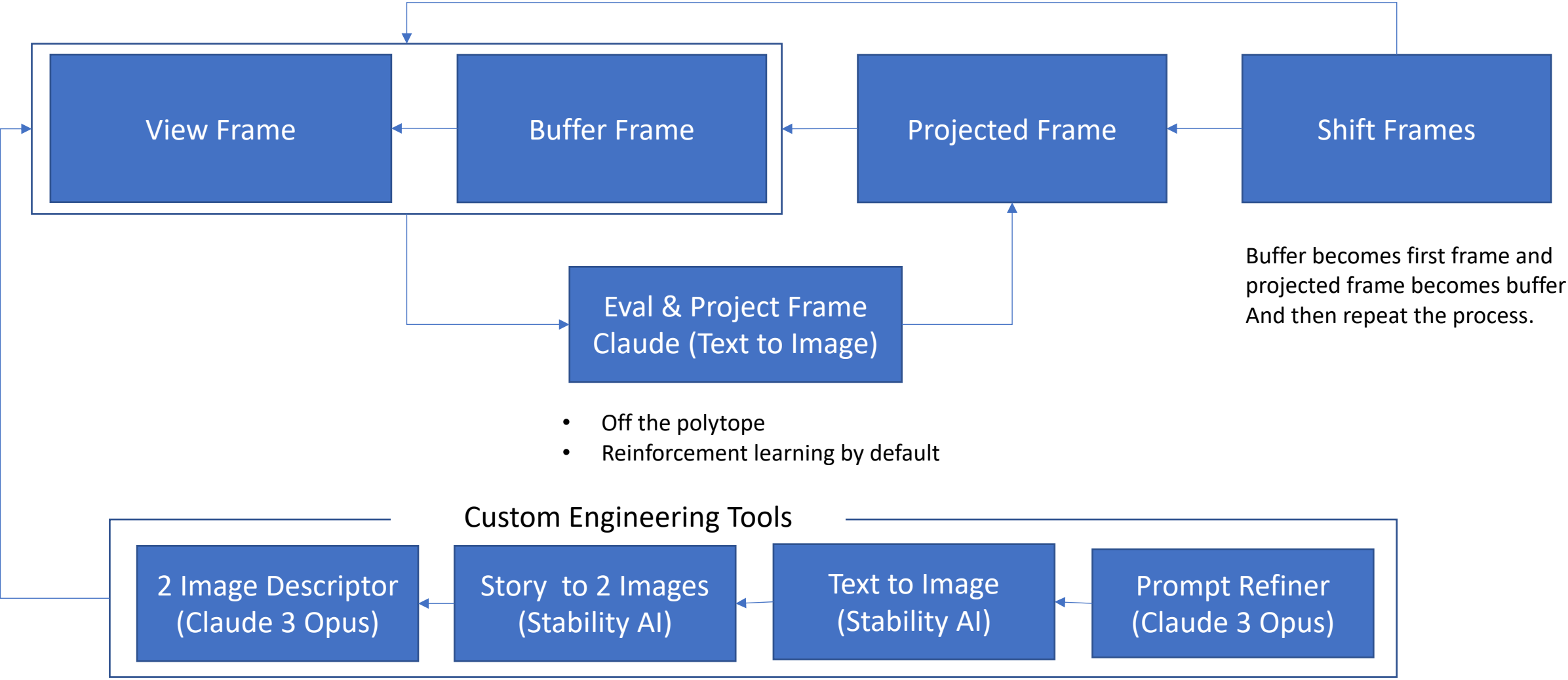
A futuristic robot with glowing eyes and a metallic body stands in a room filled with film reels and a television. The robot is holding a television set that displays a crowd of people. The room is dimly lit, with a single lamp hanging above the robot. The background wall is covered with various film reels and a framed portrait of a woman.

Teaching Claude 3 Opus How to Watch TV

Prompt: Claude tell me how I could teach you to watch TV.

From Claude: The idea would be to treat this as a storytelling exercise, where I observe the two given frames, interpret the contents and actions occurring in those frames, and then use my natural language generation abilities to continue the story forward by describing what I think would happen in the next frame.

Global Insight - Claude 3 Opus Project Flow Chart



Tools to Build Tools

Prompt Refiner
(Claude 3 Opus)

Text to Image
(Stability AI)

Story to 2 Images
(Stability AI)

2 Image Descriptor
(Claude 3 Opus)



Tools to Build Tools

(Our Claude Conversation)

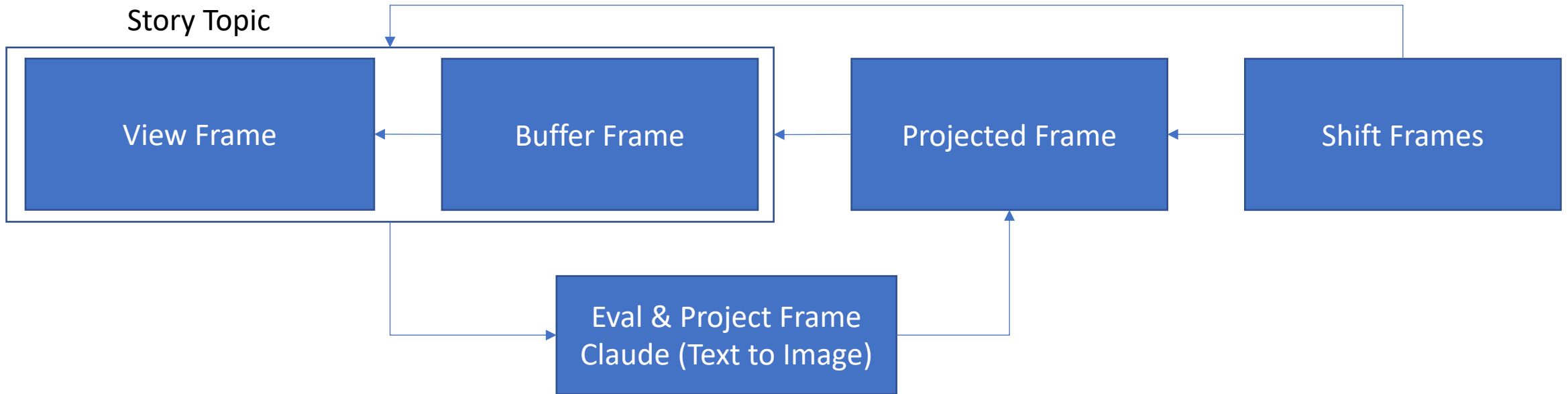
Ah I see, you're suggesting that instead of training a complex deep learning model, we could leverage my existing knowledge and capabilities to generate a continuous narrative based on analyzing the current and previous video frames. The idea would be to treat this as a storytelling exercise, where I observe the two given frames, interpret the contents and actions occurring in those frames, and then use my natural language generation abilities to continue the story forward by describing what I think would happen in the next frame. For example, let's say the current frame shows a person running down a city street, and the previous frame had them standing still. From those two frames, I could infer they have started running, perhaps chasing after something or someone. Then for the next descriptive frame, I could generate a continuation like: "The person's feet pounded against the pavement as they raced down the street, dodging pedestrians and leaping over obstacles. Up ahead, they spotted their target rounding the corner and redoubled their efforts, determination etched on their face." By analyzing the visual cues and leveraging my understanding of events, characters, and narrative storytelling, I could stitch together a descriptive story experience frame-by-frame based on the video inputs. This wouldn't be generating actual video frames, but rather using my language abilities to project what the scene could evolve into next based on the observed visuals. It becomes an imaginative storytelling exercise using the video as the prompt. I think it's a creative idea to explore the interaction between machine vision understanding and language generation in this narrative sense. Please let me know if this is along the lines of what you had in mind!



(Our Claude Generated Prompt)

Generated Prompt Template: Your task is to generate a narrative story by analyzing a sequence of video frames. I will provide you with two frames from a video. Your job is to carefully examine the contents and actions occurring in those frames, and then use your imagination and storytelling abilities to continue the story forward by describing what you think would happen in the next frame. Here are the video frames for you to analyze: `<images> {IMAGES} </images>` First, write out a brief analysis of what you observe in the provided frames, including the characters, setting, actions, and any other relevant details. Provide this analysis inside `<analysis>` tags. `<analysis>` `</analysis>` Next, based on your analysis, generate a creative and descriptive continuation of the story that could occur in the next frame after the ones provided. Your story continuation should be 3-5 sentences long and vividly describe the imagined scene, actions, and any dialogue that you think would logically follow based on the events so far. Write your story continuation inside `<story_continuation>` tags. `<story_continuation>` `</story_continuation>` Remember, the goal is to use your language generation abilities to create an engaging narrative that flows naturally from the provided video frames. Carefully analyze the visual cues, characters, and events, and then use your imagination to project how the story could evolve from there.

Why? It's How Our Brain Really Works



- Off the polytope
- Reinforcement learning by default

Demo: Stable Diffusion View and Buffer Images

Enter your Stability API Key:

Enter your Claude API Key:

Enter a movie topic

Generate Scenes



Scene 1

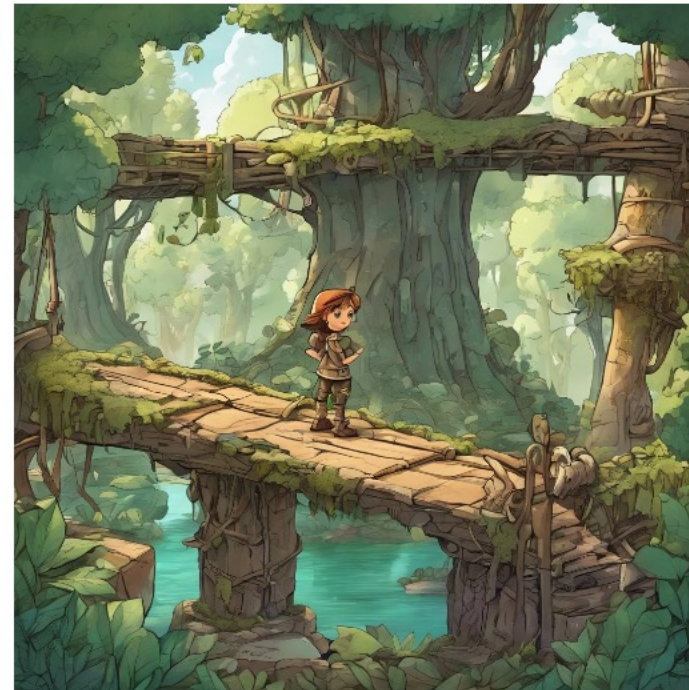


Scene 2

Demo: Claude Image to Text



+



"The image depicts a lush, fantastical forest scene. A winding stone path leads through towering trees with thick, twisting trunks. The trees create a dense canopy overhead, filtering sunlight into hazy beams that illuminate the forest floor.\n\nAlong the path, simple wooden handrails guide the way. The path ultimately leads to a small wooden bridge that crosses over what appears to be a stream or small ravine.\n\nThe forest is rendered in muted shades of green, brown and yellow, giving it a dreamy, almost ethereal quality. A few tiny human figures can be seen walking along the path, providing a sense of scale that emphasizes the grandeur and height of the trees.\n\nOverall, the scene evokes a sense of wonder and mystery, inviting the viewer to imagine themselves wandering through this enchanting, primeval woodland. It feels like a glimpse into a magical realm untouched by the modern world.', type='text!'"

"The image depicts a young girl exploring a lush, fantastical forest. She stands on a wooden bridge that spans a tranquil, teal-colored stream winding through the dense foliage. The forest is filled with enormous, ancient-looking trees whose trunks and branches intertwine to create a canopy overhead, allowing rays of sunlight to filter through.\n\nThe girl appears small amidst the towering trees, emphasizing the grand scale of the enchanted woodland. She wears a hat and backpack, suggesting she is on an adventure or journey through this magical realm.\n\nThe color palette features vibrant greens of the leaves and moss, contrasted by the wood tones of the trees and bridge. The teal water adds a mystical, almost glowing quality to the scene.\n\nOverall, the illustration captures a sense of wonder, exploration, and the beauty of an imaginative, hidden world waiting to be discovered by a curious young adventurer. The detailed, stylized artwork brings the forest to life, inviting the viewer to imagine the stories and secrets held within its depths.', type='text!'"

Demo: Claude Projection to Next Frame

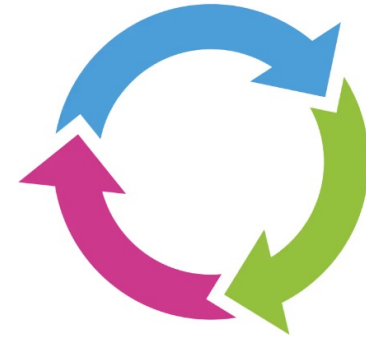
Claude Projection from View and Buffer Descriptions

"The curious girl continues her enchanted forest adventure, wandering deeper along the winding path. She discovers hidden clearings filled with strange, luminous flowers and encounters friendly woodland creatures who guide her onwards. The forest's ancient trees seem to whisper long-forgotten secrets as she passes.\n\nFinally, the girl reaches a majestic tree at the heart of the forest. Climbing inside its hollow trunk, she uncovers a mysterious, glowing medallion. As she clasps it, the forest around her shimmers and transforms, revealing an even more wondrous, magical realm lying beyond what she first saw. With the medallion to guide her, the girl steps forward to explore this new world, ready for whatever fantastical adventures await.

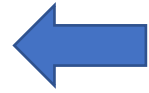
Projected Frame



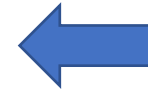
Demo



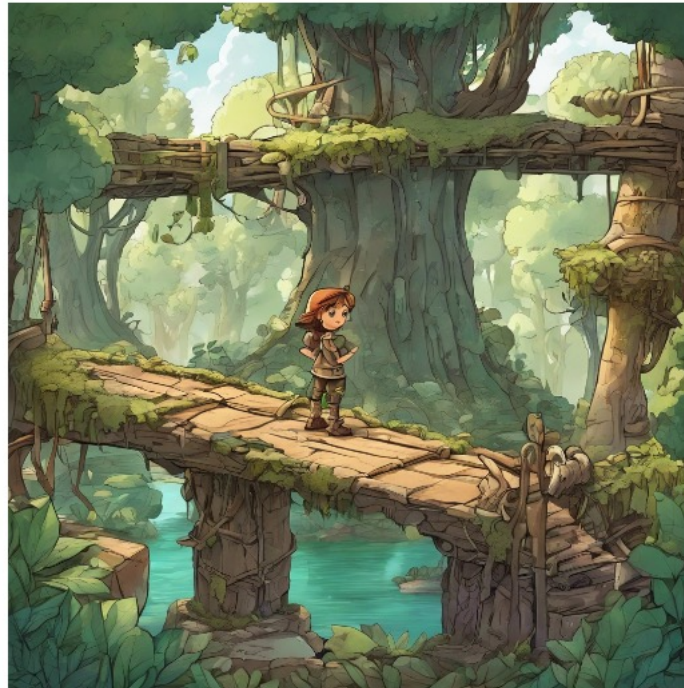
View



Buffer



Projected Frame



Business Value is Huge

Predictive modeling in AI enables robots to forecast future events, resembling human predictive thinking, and offering practical benefits:

- **Intuition:** Robots analyze past data to intuit future outcomes, mimicking human foresight.
- **Decision-Making:** Anticipating possible scenarios improves safety and responsiveness, crucial for systems like autonomous vehicles.
- **Problem-Solving:** Identifying potential issues early on, robots in industrial settings can prevent costly downtimes.
- **Interaction:** By predicting needs, robots can better assist and collaborate with humans.
- **Healthcare:** Robots help in healthcare by estimating patient trajectories, aiding in diagnosis and treatment.
- **Security:** Predicting risks, security robots can alert and act preemptively to ensure safety.



Thanks to the Team

- Areeb Pasha
- Hus nain
- Muhammad Arham
- Thalanayar Muthukumar
- Mahnoor Malik

