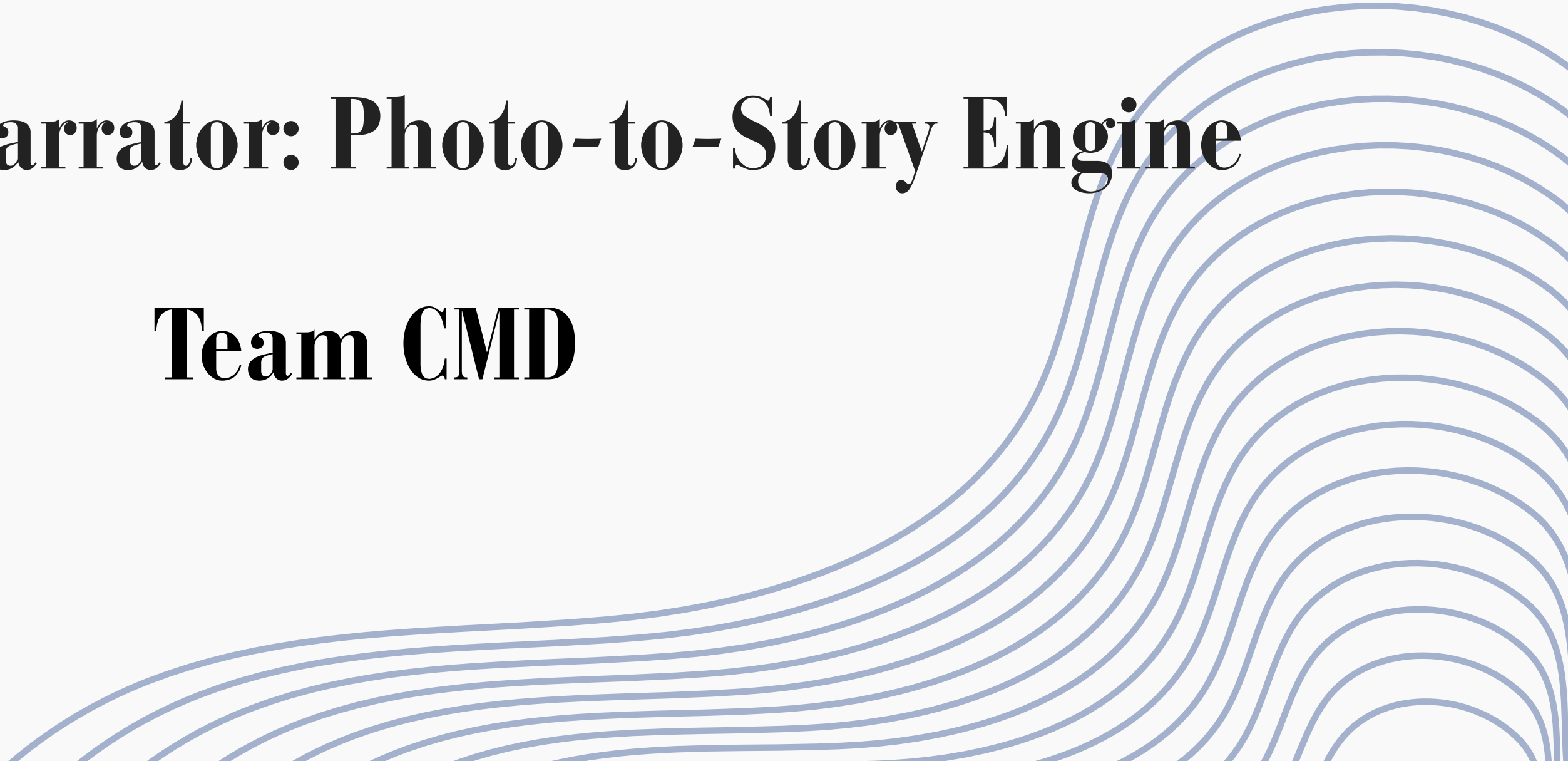




UDBHAV'25

The Visual Narrator: Photo-to-Story Engine

Team CMD



IDEA



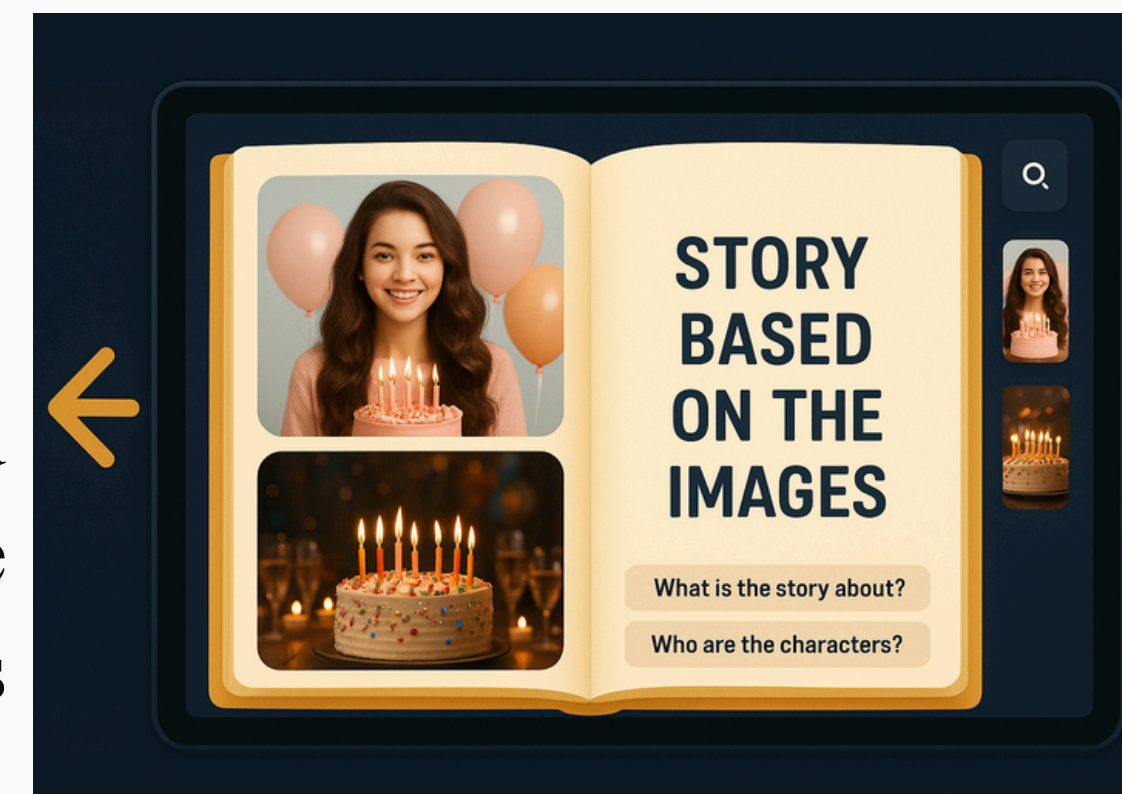
THE PROBLEM:

We are the last few generations whose childhood photos were in albums and teenage photos are in digital format. And the children now are completely transitioning to the digital medium. And we forget the story and memory behind that day as we grow old.

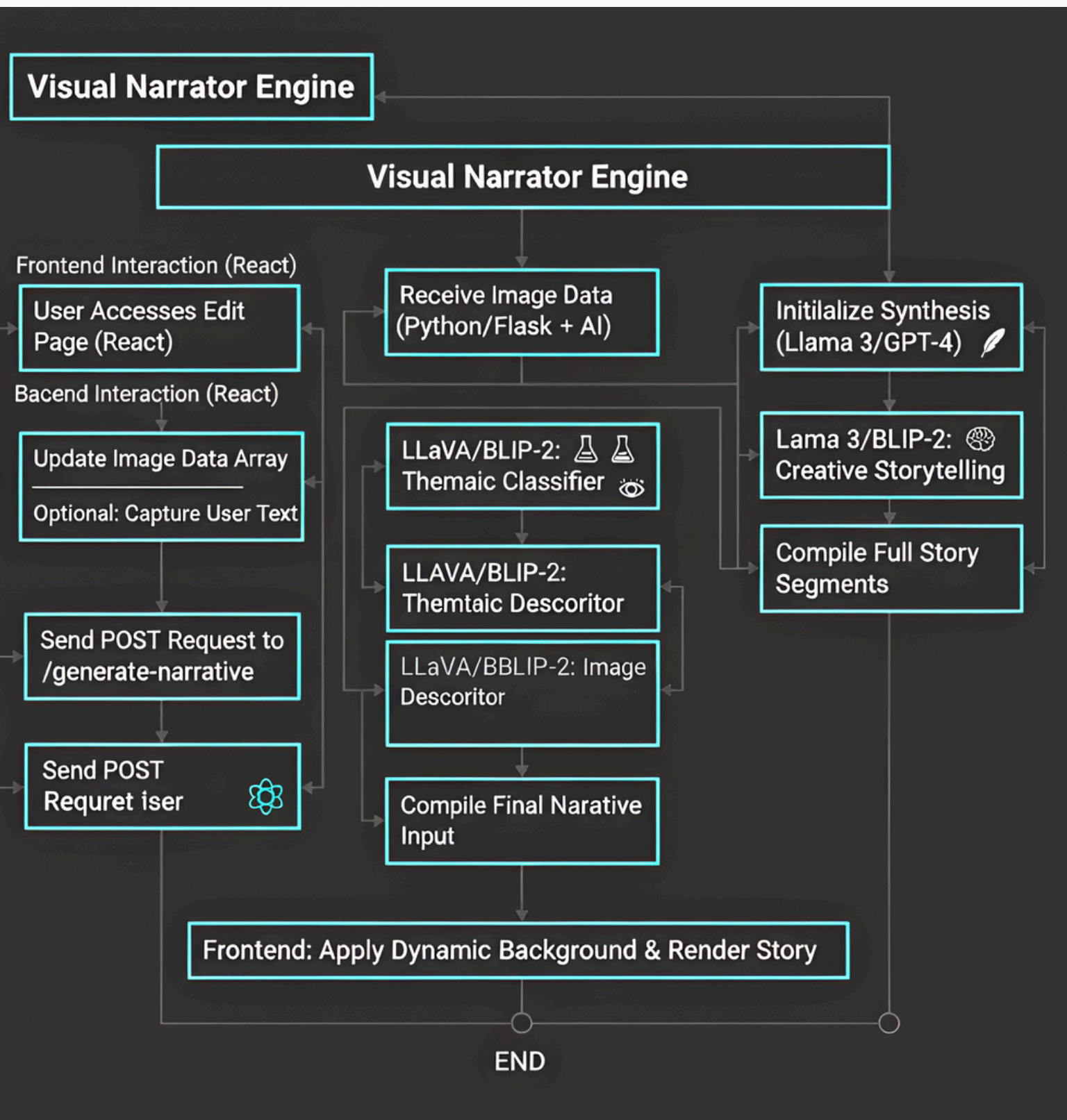
OUR SOLUTION:

So we create an digital album/storybook which you can take anywhere you want, which keeps all the stories we once lived and forgot. And the forgotten memories can be rewritten. In a customizable, editable albums you can carry anywhere.

It creates a personal event chronicle, ready to playback whenever nostalgia strikes. It can be used to create albums, a diary, which lives forever with personal notes



TECHNICAL DETAILS



FRONTEND: 

React :Builds dynamic interfaces, instantly updates UI, and manages user images dynamically

BACKEND: 

Flask: A lightweight Python backend connecting frontend images to AI models and Neon database.

AI USED:

BLIP-2: Converts input images into descriptive text captions

LLama 3: Weaves BLIP-2's image descriptions into a cohesive and engaging story

LLaVA: Understands complex visual details and context.

GPT-4: Writes highly immersive and human-like story

DATABASE:

NEON-It is a serverless Postgres platform. Used to save the albums created by user



BUSINESS SCOPE AND IMPACT

Business scope

- The albums can be edited and made into a physical album format which can be further printed and sent to the buyer
- We can create a 3D model of the moment and send it to the user



Impact

- Relive the old memories with a new story
- Can help kids learn by turning their surroundings into immersive stories
- It creates a digital diary that transports users back to the moment, offering a fully immersive experience whenever they choose to revisit it

CHALLENGES AND FUTURE SCOPE

CHALLENGES:-

- 1:Some user expect very cinematic or perfect stories, which AI may not always deliver.
- 2:Ensuring the story follows logic flow based on single image by avoiding un-realistic storytelling.
- 3:Handling the photos which are blurry, low-light, or having complex scenes.

FUTURE SCOPE :-

- 1:Allow user to choose the tone (funny, emotional, dramatic, cinematic, etc..).
- 2:Convert generated stories in AI-narrated audio with different voice of celebrities/loved-ones.
- 3:Implement polylingual narrative generation.
- 4:Build a mobile app with low-data version for wider accessibility.

