

FORGE AI

*Motivation, Use Cases, Planned
Features*

AGENDA

Introduction and Motivation

Use Cases

Planned Features

Technologies Used



Introduction and Motivation

2020 marks the new **golden age** of AI — There's a model for almost every task:

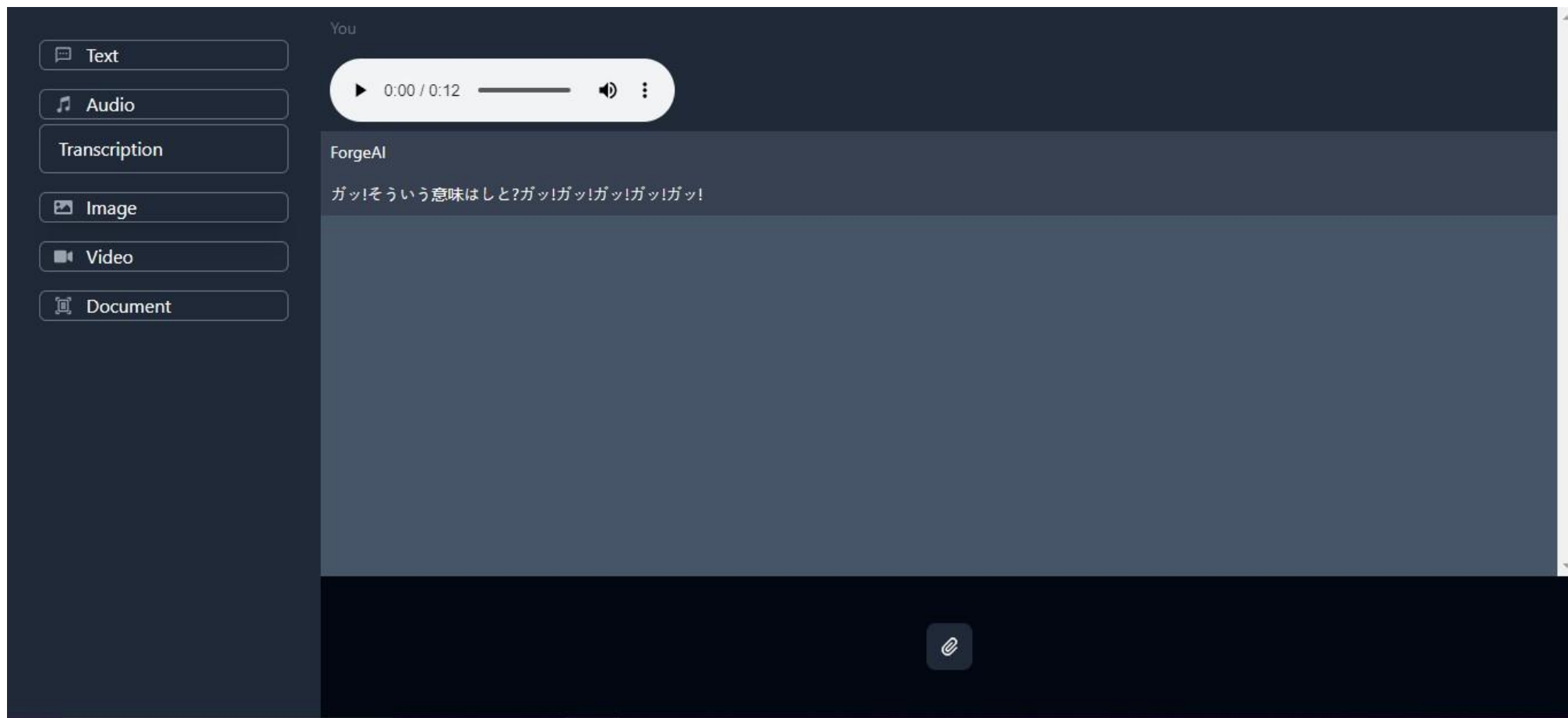
Pictures and Movies: DALLE, SORA

Text, Code Generation: Co-pilot, chatGPT, Gemini, LLaMa

Audio: Whisper, Pycomposer

Most of these models are offered in different websites, which is a bit inefficient.
How about instead having **a single platform** for all your AI needs?
Welcome to Forge AI!

Forge AI



Use Cases and Planned Features

As ForgeAI seeks to be all-in-one platform, we'll tackle the most important subsets:

Text2Text - Used for Question and Answering, code Generation

Text2Speech – For smart Agents, such as Alexa, or creative purposes (Videos)

Speech2Text – Transcribe audio in multiple languages;

Image/Video2Text – Caption or Summarize a sequence of images (e.g Movie Summarization);

-> **Problem:** Video2Text is one of the hardest problems to Tackle! One of the models (mPLUG-2) is not Available anymore, hence we approach it **differently**.

Technologies Used

Frontend: HTML, CSS, JS (React), Tailwind

Backend:

- Azure storage (blob to store images, text, videos + table to store user history);
- FastAPI to communicate with azure storage + load models

Tools: ML models (from Hugging Face and API's)

Extra: Azure Video Indexer for the **Video Captioning model**