

# Agentic AI Media Platform

## Python Backend & Automation

A production-ready platform that unifies image, video, prompt, caption, transcript, conversion, and document workflows through agentic AI orchestration.

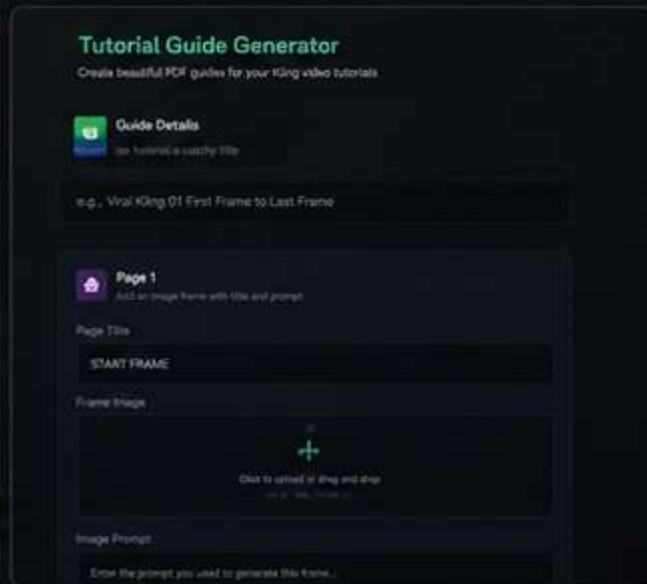
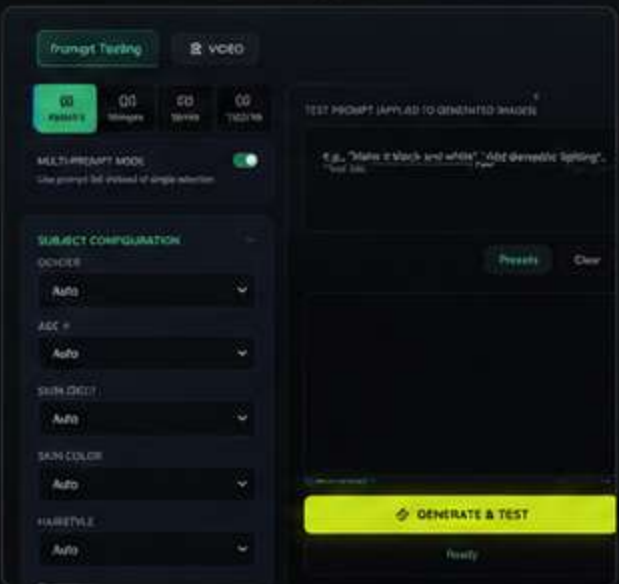
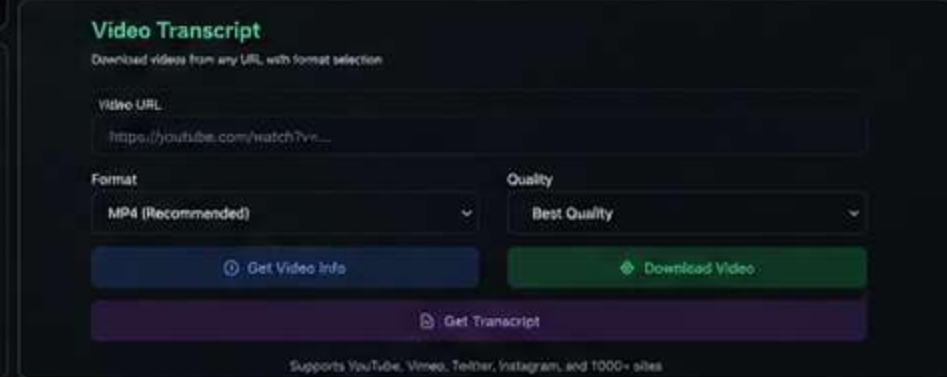
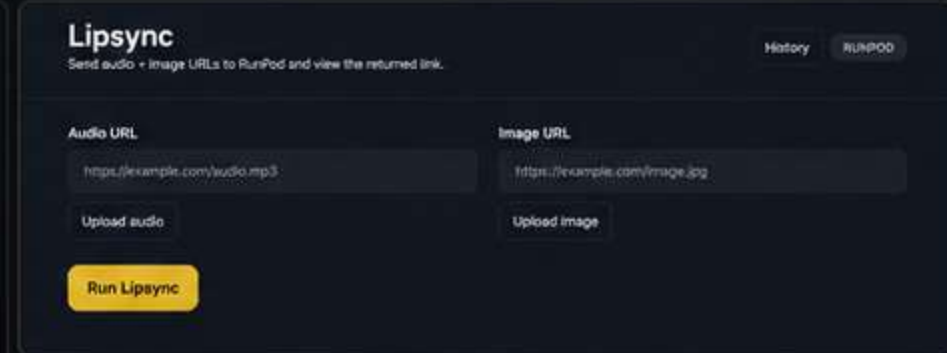
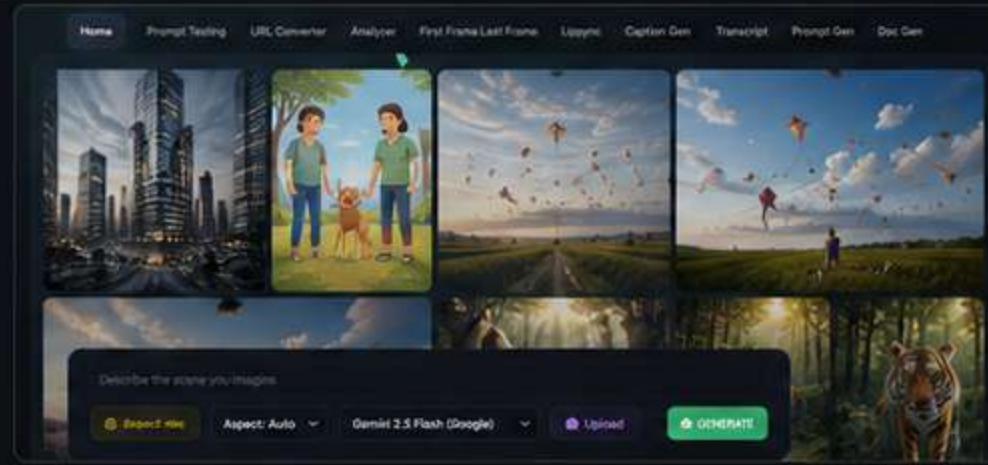
10+ AI workflows

Async processing

WebSockets

S3 / Cloudflare R2

OpenAI • Gemini • Fal AI • Whisper





# Executive Summary

## A unified AI media operating system



### 1. Problem

Fragmented tools, manual handoffs, and unreliable long-running AI jobs slow teams and limit scale.



### 2. Solution

Independent agent-style workflows coordinated by a Python backend with real-time orchestration.



### 3. Result

One production-ready platform spanning creation, automation, delivery, and storage.



### Non-blocking jobs

Async tasks run in background without blocking the experience.



### Real-time UI updates

WebSocket-driven updates keep users informed at every stage.



### Multi-service AI orchestration

Specialized agents work in concert for faster, smarter outcomes.

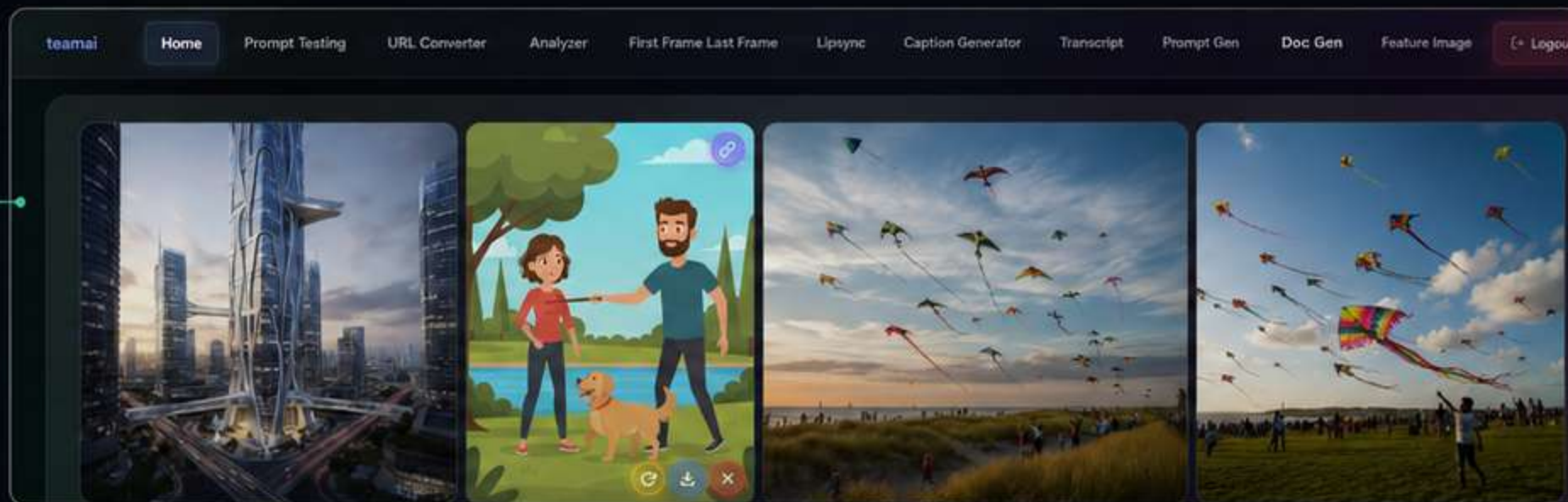


# One Platform, Many Workspaces

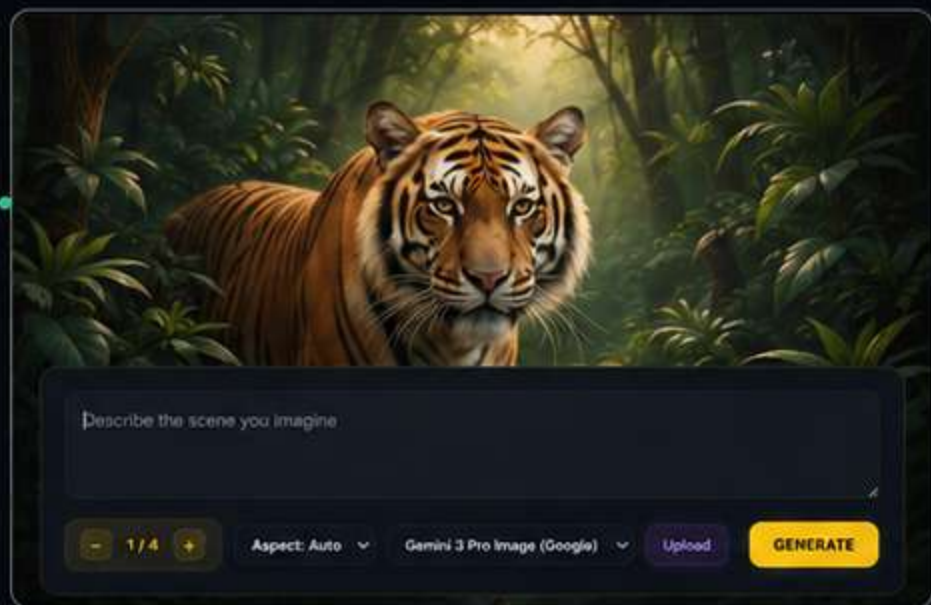
A central surface for exploration,  
generation, and real-time discovery



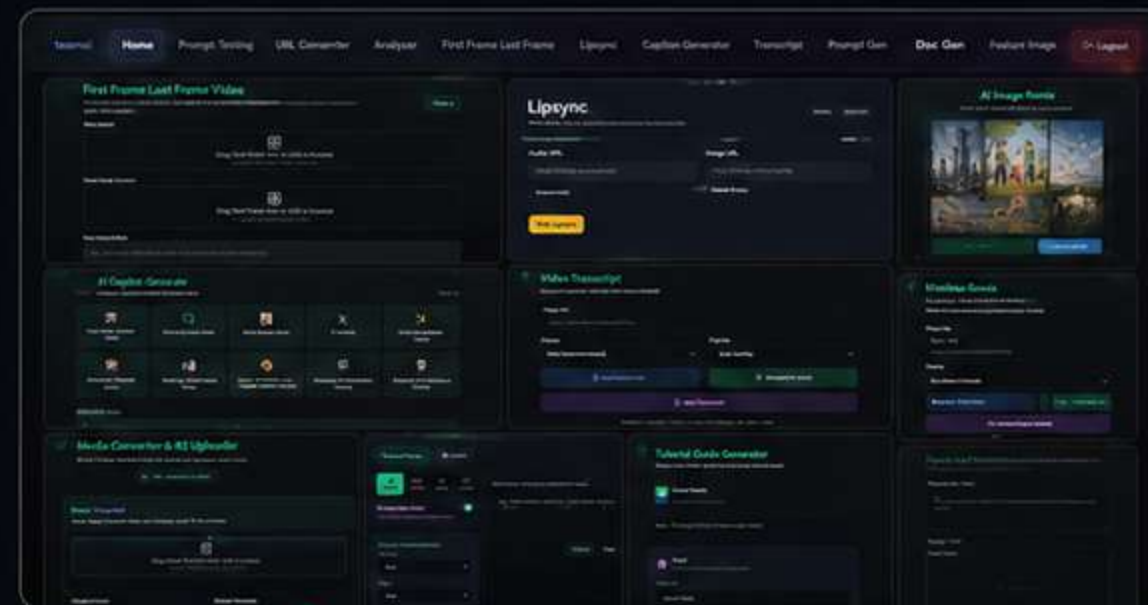
Infinite-scroll  
gallery



One-click  
generation  
from the feed



Model selection +  
aspect control



Central navigation  
across modules



Designed for fast exploration with live updates and a unified product surface.

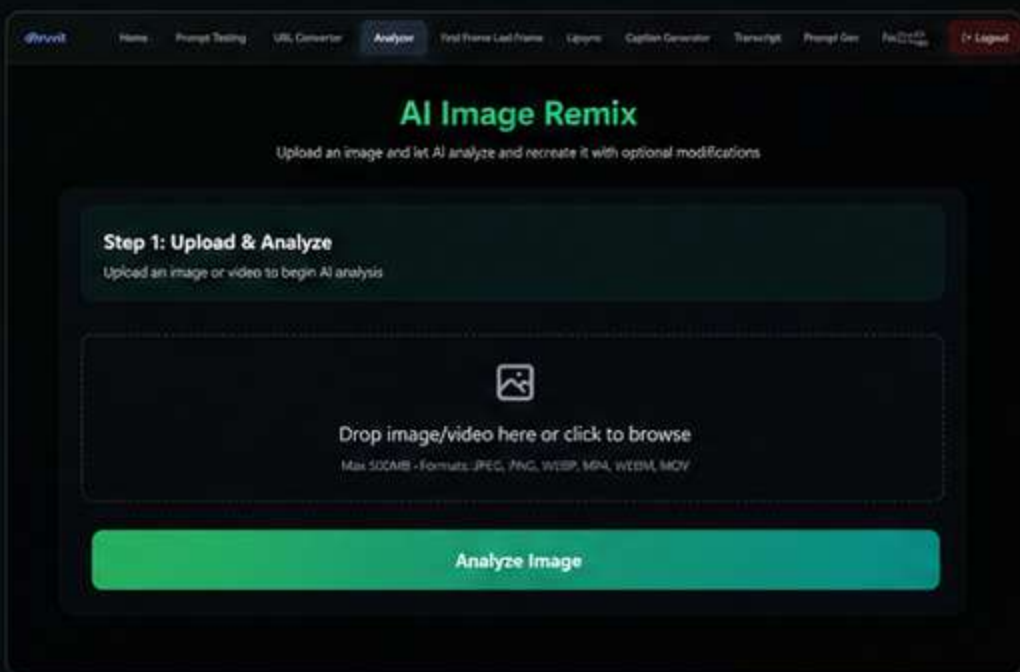
# Creation & Prompt Intelligence

Create, remix, retrieve, and test



## Image Remix

Analyze image or video inputs and extract editable prompts.

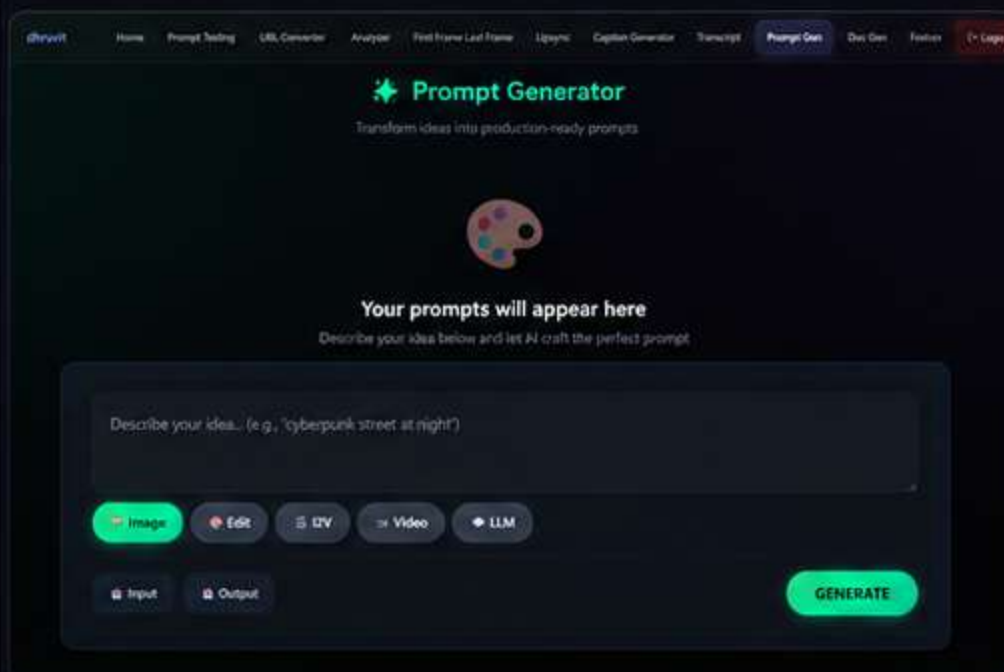


- ✔ Prompt extraction from media
- ✔ Model-specific prompt generation
- ✔ Preset-based experimentation
- ✔ Multi-prompt testing



## Prompt Generator

Retrieve proven patterns from a vector database and tailor outputs to specific AI models.

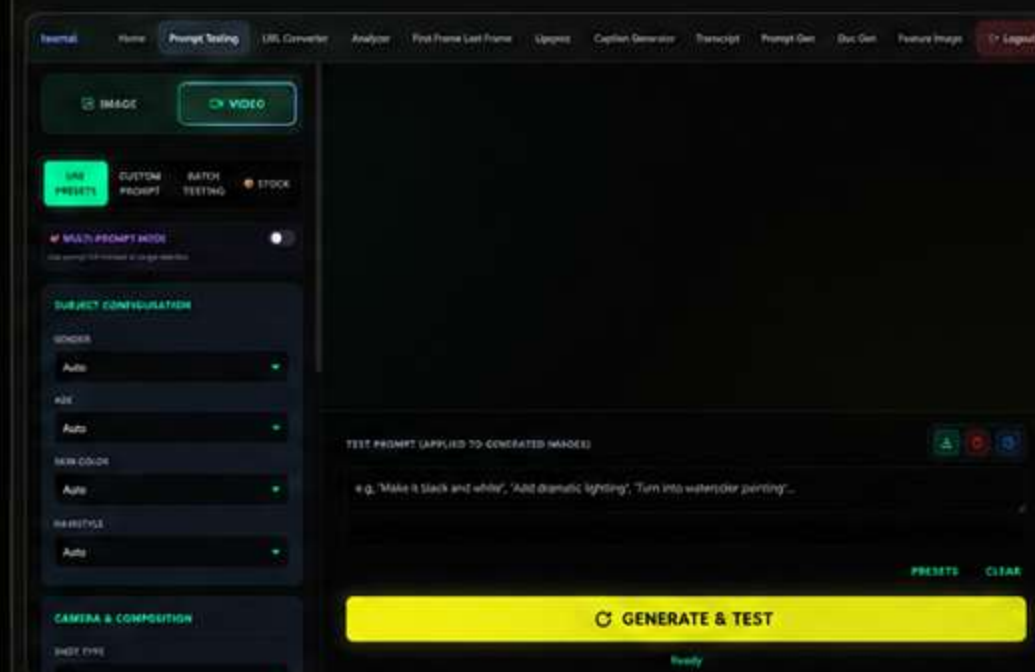


- ✔ Prompt extraction from media
- ✔ Model-specific prompt generation
- ✔ Preset-based experimentation
- ✔ Multi-prompt testing



## Prompt Testing Lab

Test raw prompts across models with presets, seeds, aspect ratios, and safety controls.



- ✔ Prompt extraction from media
- ✔ Model-specific prompt generation
- ✔ Preset-based experimentation
- ✔ Multi-prompt testing



End-to-end intelligence for creators — from insight to iteration.

Analyze >> Generate >> Test >> Refine >> Create

# Async Video Workflows

Consistent motion generation, orchestrated reliably



Condition videos from a start frame and optional end frame




Translate instructions into precise motion prompts




Inject media into external workflow graphs



Track long-running jobs with polling and history

 Async job handling

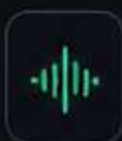
 History

 External workflow engine



## First Frame / Last Frame Video

Generate cinematic transitions from start to finish.



## Lipsync Studio

Sync audio and image to create natural, expressive lip movements.

# Content Intelligence

Captions and transcripts, generated in a few steps



## Caption Generator

- Generate multiple caption styles in parallel
- Automatically optimize media before processing

The screenshot shows the 'AI Caption Generator' web interface. At the top, there's a navigation bar with links like 'Home', 'Prompt Testing', 'URL Converter', 'Analyzer', 'First Frame Last Frame', 'Lip Sync', 'Caption Generator', 'Transcript', 'Prompt Gen', 'Doc Gen', and a 'Logout' button. The main heading is 'AI Caption Generator' with the subtitle 'Generate captions for images and videos with AI'. Below this is an 'Upload Image or Video' section with a large upload icon and the text 'Drop file here or click to browse'. Underneath, it specifies 'Max 100MB - Images: JPG, PNG, WEBP - Video: MP4, MOV, WEBM'. A 'Select Caption Types' section follows, featuring a grid of 10 options: Social Media (Instagram Style), SEO-Friendly Social Caption, Short Punchline Caption, X (Twitter), Creative / Storytelling Caption, Professional / Business Caption, Marketing / Promotional Caption, Meme / Funny Caption, Aesthetic / Photography Inspired, and Educational / Insightful Product. A 'Generate Captions' button is at the bottom.



## Transcript & Downloader

- Download from URLs with metadata extraction
- Produce accurate AI transcripts for reuse

The screenshot shows two web interfaces. The 'Video Transcript' interface on the left has a heading 'Video Transcript' and the subtitle 'Download videos from any URL with format selection'. It features a 'Video URL' input field with a placeholder 'https://youtube.com/watch?v=...', a 'Format' dropdown menu set to 'MP4 (Recommended)', and a 'Quality' dropdown menu set to 'Best Quality'. There are two buttons: 'Get Video Info' and 'Download Video'. Below these is a large purple 'Get Transcript' button. At the bottom, it says 'Supports YouTube, Vimeo, Twitter, Instagram, and 1000+ sites'. The 'AI Transcript' interface on the right has a heading 'AI Transcript' and a 'Download' button. It shows 'Source: YouTube • Duration: 5:42 • Language: English'. Below this is a list of transcript segments with timestamps and text: '00:00:01 Welcome to our product demo.', '00:00:05 In this video, we'll explore the key features', '00:00:09 and how they can help streamline your workflow.', and '00:00:14 Let's get started.'. At the bottom, it displays 'Total Words: 1,248 • Segments: 32'.

# Media Operations

Conversion, cloud upload, and reusable asset history



Batch convert images and videos



Upload directly to S3 / Cloudflare R2



Generate hosted URLs for easy reuse



Review, search, export, and manage asset history



Cloud-hosted output



Batch processing



Reusable assets

teamai Home Prompt Testing **URL Converter** Analyzer First Frame Last Frame Lipsync Caption Generator Transcript Prompt Gen Doc Gen Feature Image Logout

## Media Converter & R2 Uploader

Convert images · Download videos from URLs · Upload to Cloudflare R2

[View Conversion History](#)

### Image Converter

Upload images, customize names, select formats, and get R2 URLs instantly

Drop images, videos, or audio here or click to browse  
Images: PNG, JPG, WEBP · Videos: MP4, MOV · Audio: MP3, WAV

**Output Formats**

**IMAGE FORMATS**

WEBP Best compression

PNG Lossless

JPG Universal

**Upload Destination**

ChromaStudio

folder name (will be created inside public/)

**Image & Video Options**

[Upload files to begin](#)

dhruvit Home Prompt Testing **URL Converter** Analyzer First Frame Last Frame Lipsync Caption Generator Transcript Prompt Gen Doc Gen Feature Image Logout

## Conversion History

View and manage your converted files

**Display Limit**  
Showing 9 most recent items (up to 300, Your account stores up to 335 items total, with oldest items automatically removed.)

Total Items: 9

Images: 9

Videos: 0

Est. Storage: 2.7 MB

Search files... Newest First All Files [Copy All URLs](#) [Export](#) [Clear All](#)

testimonial-emily.webp WEBP facesswapper Jan 12, 2025, 11:47 AM	simpsons_mutant_DPKM75W.webp WEBP dreamshot Jan 12, 2025, 11:43 AM	O1CN011j3e2NTZTZNx1R5_22064... PNG chroma studio Jan 12, 2025, 06:30 PM	O1CN01QvB1eg25IHLo7TvbR_2917... JPG chroma studio Jan 12, 2025, 06:30 PM
<a href="#">Copy</a>	<a href="#">Copy</a>	<a href="#">Copy</a>	<a href="#">Copy</a>

# Document Generator

Turn frames, prompts, and media into shareable guides



Create structured PDF guides from images and text



Support dynamic multi-frame layouts



Combine frame inputs with prompt context



Export styled documents ready to share

dhruvit Home Prompt Testing URL Converter Analyzer First Frame Last Frame Lipsync Caption Generator Transcript Prompt Gen **Doc Gen** Feature Image Logout

## Tutorial Guide Generator

Create beautiful PDF guides for your Kling video tutorials

### Guide Details

Give your tutorial a catchy title

Guide Title

e.g., Viral Kling O1 First Frame to Last Frame

### Page 1

Add an image frame with title and prompt

Page Title

START FRAME

Frame Image

Click to upload or drag and drop

Image Prompt

Enter the prompt you used to generate this frame...

### Page 1

Add an image frame with title and prompt

Page Title

START FRAME

Frame Image

Click to upload or drag and drop

Image Prompt

Enter the prompt you used to generate this frame...

### Dynamic Multi-Frame Support

Add multiple frames with titles and prompts to build rich, structured guides.

### First Frame

Drop first frame here or click to browse

www.kling.ai/tutorial/001\_0101

### Last Frame (Optional)

Drop last frame here or click to browse

www.kling.ai/tutorial/001\_0102

### Raw Instructions

e.g., Transition from day to night, Add cinematic camera movement, Close a sunset window.

### Frame-Powered Content

Use first and last frame inputs from video workflows to create guided, step-by-step content.

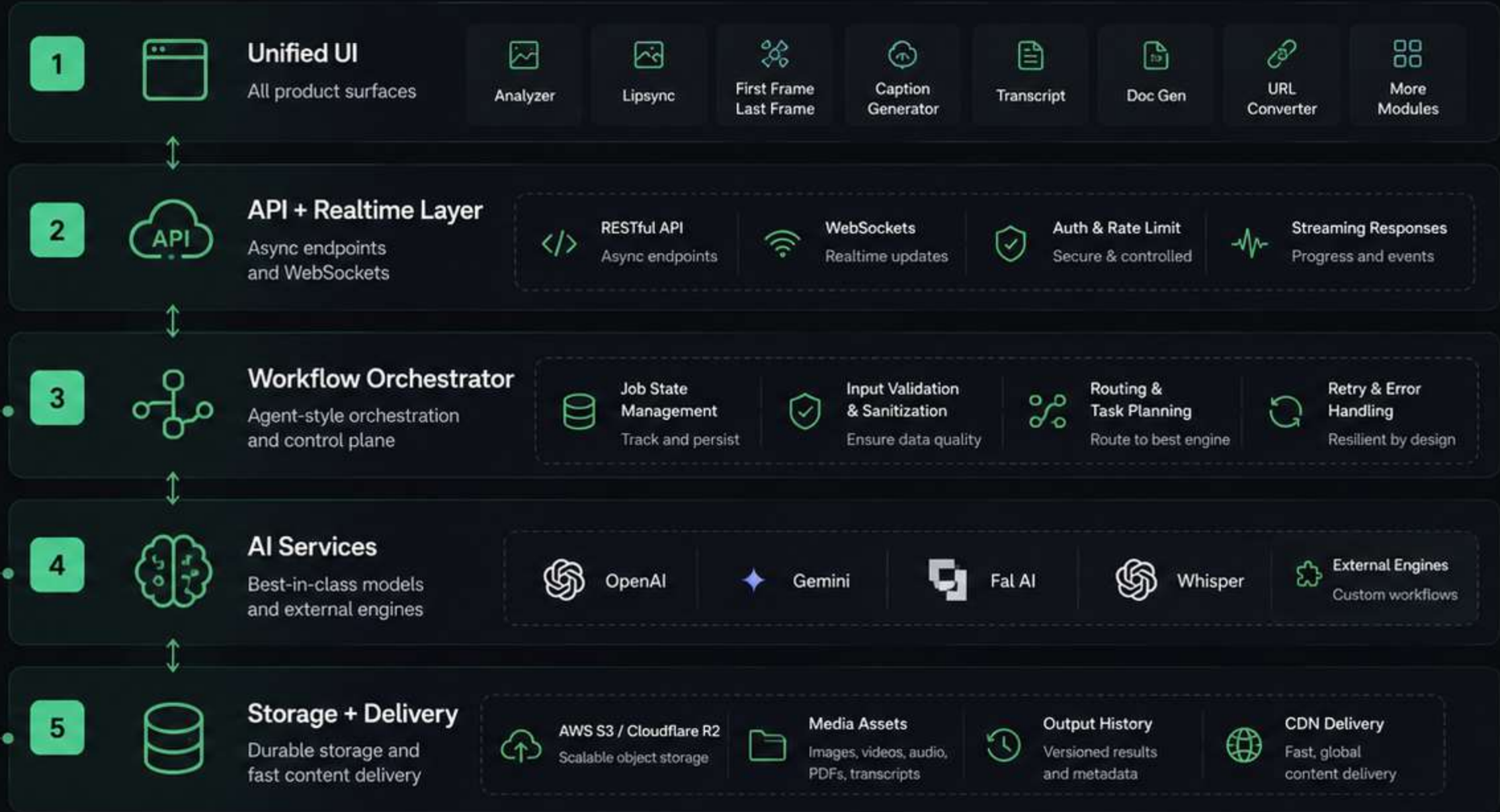
# Technical Architecture

Python backend, modular routing, and agent-style orchestration

**Non-blocking execution**  
Async processing for high throughput and responsiveness

**Clear separation of concerns**  
Decoupled layers for maintainability, scalability, and flexibility

**Production-ready reliability**  
Retries, validation, observability, and resilient infrastructure



# Role, Outcome & Value

Built end-to-end as a solo developer



## 1. My Role

System architecture · Python backend · AI workflows  
External service integrations · End-to-end product delivery



## 2. Outcome

Unified more than 10 AI-driven workflows in one production-ready platform.



## 3. Business Value

- Reduced manual effort across content and media operations
- Reliable long-running job execution with real-time updates
- Faster experimentation and iteration with unified tools
- Reusable, cloud-hosted outputs for any use case

✦ 10+ workflows

⚡ Real-time updates

🔄 Async jobs

👤 Solo-built



A scalable foundation for **AI-native media operations**.

## Agentic AI Media Platform

Python Backend & Automation

A production-ready platform that unifies image, video, prompt, caption, transcript, conversion, and document workflows through agentic AI orchestration.

✦ 10+ AI workflows    ⚡ Async processing    🔄 WebSockets  
☁️ S3 / Cloudflare R2    🤖 OpenAI · Gemini · Fal AI · Whisper

