

Part 1: Foundational Systems

1.1. Asset Management & Metadata Database

Database: Implement a database (e.g., Firestore, DynamoDB) to store asset metadata.

Metadata Schema: Each asset must include:

s3_path: Full S3 path.

asset_type: video, audio, image, or narration.

category: User-defined category.

tags: An optional list of text tags.

duration: Asset length in seconds (auto-calculated on upload).

Asset Upload UI: Create a dashboard UI to:

Upload files to S3.

Assign a category.

Add one or more tags.

Asset Linking Logic:

Tagged Assets: Can only be paired with assets sharing the exact same tag.

Untagged Assets: Can be paired with any footage for the job.

1.2. Automated Narration Engine

UI Tool: Build a dashboard tool to upload a structured text file (e.g., CSV).

Input File Format: The file must have columns for text, campaign_part, and optional tags.

Backend Workflow: Create a script that:

Reads the uploaded file.

Generates audio via the ElevenLabs API for new text.

Saves the MP3 to S3.

Creates a new asset entry in the database with complete metadata.

Part 2: Job Configuration UI & API

All settings must be sent as a single JSON "Job Instruction Set" to the worker.

2.1. Main Job Setup

Source Video Selection: Field to select one or more source videos.

Number of Variations: Input for how many unique videos to create per source.

2.2. Dynamic Structure Definition

Number of Parts: UI must ask for the number of parts for the custom video segment.

Mode Selection:

If 1-Part, show simple processing options.

If Multi-Part, generate a configuration section for each part.

2.3. Configuration Options

For 1-Part Mode:

Background Audio: Allow selection of one or more audio categories.

Overlay: Checkbox to enable a random image/logo overlay.

On-Screen Text:

Add text to the video only if asset tags match text file tags.

Provide 10 random font and text size options.

Uniqueness Engine:

Checkbox to enable.

If enabled, show options for randomizing contrast, color, audio, etc.

For Multi-Part Mode (Per-Part Settings):

Asset Selection: For each part, allow selection of one or more asset categories.

Footage Stitching:

Specify the number of video clips to use per part.

Set a min/max duration for each clip.

Narration: Allow selection of narration assets.

Soundtrack:

Select a background music category.

Specify which parts the soundtrack spans (e.g., Part 1 to Part 2).

Text Overlay:

Checkbox to enable.

Text input field.

Use a predefined list of fonts, selected randomly.

Image/Logo Overlay: Checkbox to enable.

Final Global Settings:

Horizontal Flip: Master checkbox to flip the initial source video only.

Uniqueness Engine: Master checkbox to enable final randomization of contrast, brightness, speed, and pitch.

Part 3: Worker Engine Logic

3.1. Core Worker Process

Parse Instructions: Read the JSON job configuration from the SQS message.

Variation Loop: Loop for the specified "Number of Variations."

Assemble "Shot List": Inside the loop, query the database for a new, random set of assets matching the job rules.

Generate Video:

If 1-Part Mode: Apply selected overlay, audio, and uniqueness effects to the source video.

If Multi-Part Mode:

Determine Durations: Video segment length is determined by narration audio duration.

Generate Parts: Create each part as a temporary video file.

Stitch Footage: Stitch video clips together with hard cuts, looping or trimming to match narration duration.

Mix Audio: Combine narration and background music.

Combine Parts: Use FFmpeg's concat filter to join all temporary part files.

Final Stitch: Combine the original source video with the newly generated multi-part video.

Apply Uniqueness Engine (If Enabled): Perform a final FFmpeg pass to apply randomized changes (contrast, speed, pitch, etc.).

Upload: Upload the final video to S3.

Repeat Loop: Continue until all variations are created.