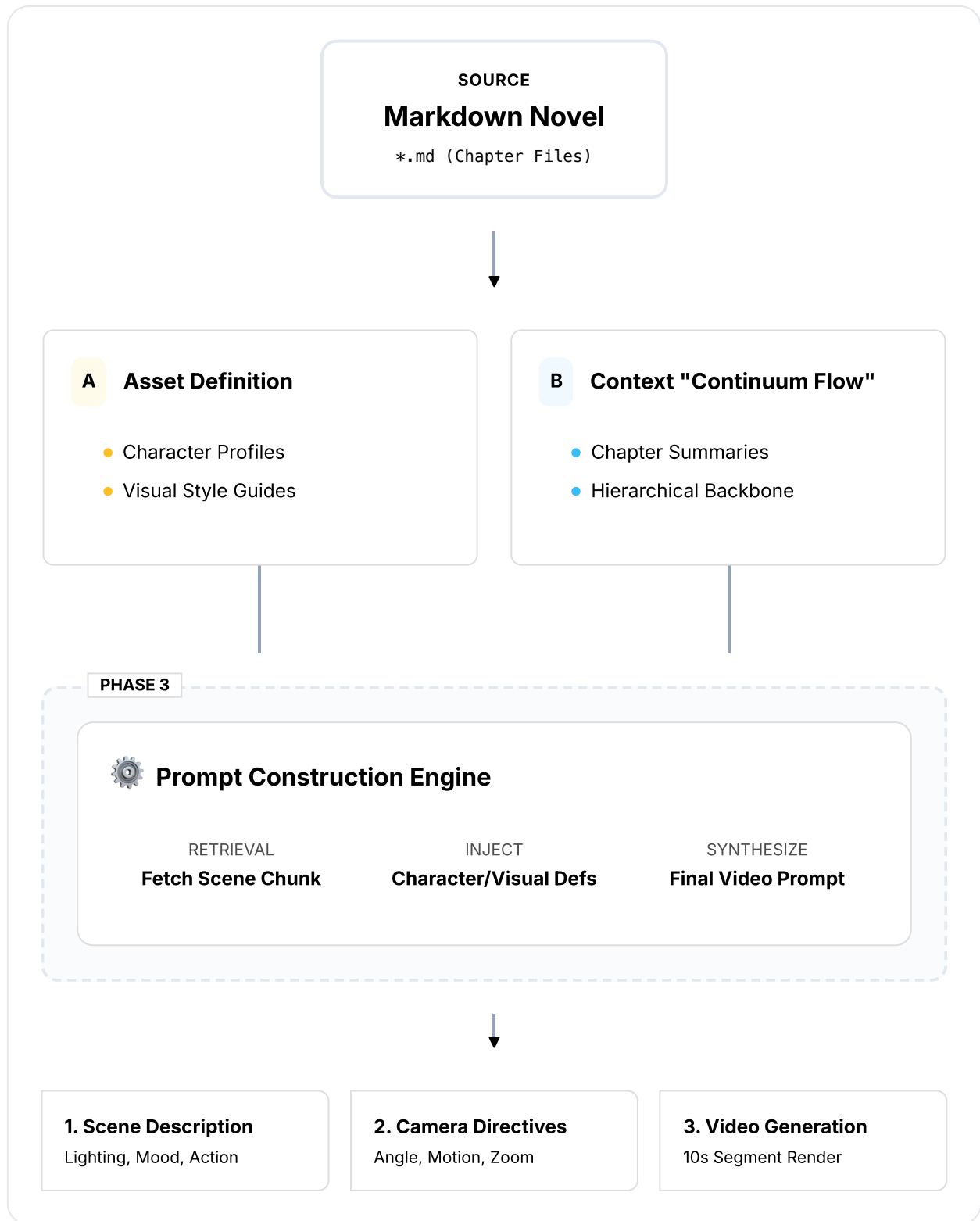# System Architecture

Core Application Architecture & Context Management Flow

The **Novel Video Generator** architecture is designed for high-fidelity narrative consistency. Unlike standard video generation pipelines that operate shot-by-shot, our system utilizes a multi-layered approach that separates static world-building from dynamic narrative progression. By caching character identities and environmental constraints in a "Global State," we ensure that AI hallucinations are minimized during the final render phase.

Application Data Pipeline

The following diagram visualizes the end-to-end flow from raw Markdown input to the final 10-second video segments. The architecture is split into three primary phases: Static Asset Definition, Dynamic Context Management (The "Continuum Flow" Agent), and the Prompt Construction Engine.

```
                    SOURCE
                Markdown Novel
                *.md (Chapter Files)
```

A **Asset Definition**

- Character Profiles
- Visual Style Guides

B **Context "Continuum Flow"**

- Chapter Summaries
- Hierarchical Backbone

PHASE 3

⚙️ **Prompt Construction Engine**

RETRIEVAL	INJECT	SYNTHESIZE
Fetch Scene Chunk	**Character/Visual Defs**	**Final Video Prompt**

1. Scene Description	**2. Camera Directives**	**3. Video Generation**
Lighting, Mood, Action	Angle, Motion, Zoom	10s Segment Render

Pre-processing and Definition Phase

Before a single frame of video is generated, the system must establish the "Ground Truth" of the story world. In a traditional film production, this is the pre-production phase: casting, costume design, and location scouting. In Continuum Flow, this is an automated agentic workflow that builds a static
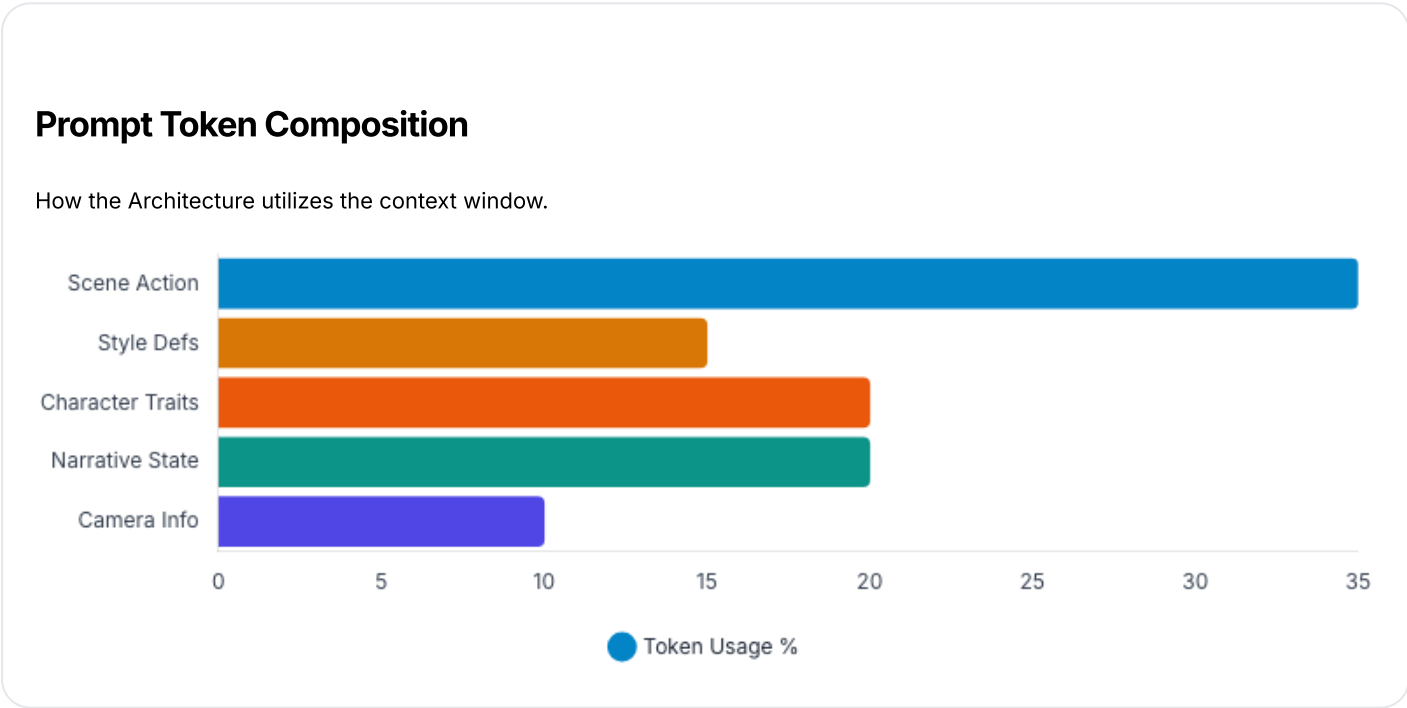
reference database. This phase is critical because generative video models (unlike text models) require explicit visual instructions for every frame to prevent hallucination or morphing of character identities.

Character Profile Definition

The first agentic workflow triggers the **Character Definition Agent**. This agent scans the entire corpus (all Markdown files) to identify unique entities. It then synthesizes comprehensive profiles for each character.

Visual Attribute Locking

To maintain consistency in video generation, character descriptions must be translated into immutable visual prompts. The agent generates a "Character Reference Sheet" for each entity, defined in a rigid JSON schema. This schema acts as the "Source of Truth" for all subsequent generation steps.

Prompt Token Composition

How the Architecture utilizes the context window.

Attribute Category	Data Points Captured	Purpose in Video Generation
Physicality	Height, body type (e.g., ectomorph), skin texture, eye shape, hair hex code.	Ensures the silhouette and basic appearance remain constant across varied camera angles.
Costume	Primary Outfit, Secondary Outfit, Accessories (e.g., "Silver Locket").	Prevents the model from "hallucinating" different clothes in every shot.
Identity Anchors	Scars, tattoos, distinct hairstyles, specific props (e.g., "glowing staff").	These are high-weight tokens injected into every prompt to force model attention on unique identifiers.
Style LoRA	Reference to specific Low-Rank Adaptation models or embeddings.	Links the text profile to a specific visual model trained on the character's likeness.

Psychological and Narrative Roles

Beyond visuals, the profiles include "Behavioral Tensors"—descriptions of how a character moves and reacts. A "nervous" character requires video instructions for "jittery camera movement" or "fidgeting hands," while a "stoic" character requires "static framing" and "minimal micro-expressions." These behavioral traits are encoded as metadata that influences the camera direction in later phases.

Scene Description and Environmental Modeling

The system creates a **Global Location Registry**. Similar to character profiling, the **Environment Agent** scans the text to identify recurring locations.

- **Lighting and Mood**: For each location, the agent defines the baseline lighting (e.g., "volumetric god rays," "cyberpunk neon," "dim candlelight") and atmospheric mood.

- **Spatial Geometry**: To ensure characters move consistently through space, the agent estimates the geometry of key sets (e.g., "The kitchen island is to the left of the fridge").

The Pre-processing Workflow

1. **Entity Extraction**: An NLP entity extraction model runs over the full text.

2. **Cluster Analysis**: Mentions of "John," "Jonathan," and "The Detective" are clustered to verify they refer to the same entity.

3. **Profile Synthesis**: An LLM aggregates all descriptors into a unified profile.

4. **Conflict Resolution**: A specialized **Conflict Agent** flags inconsistencies for review.

Processing Pipeline Load

Execution density across the delivery lifecycle.