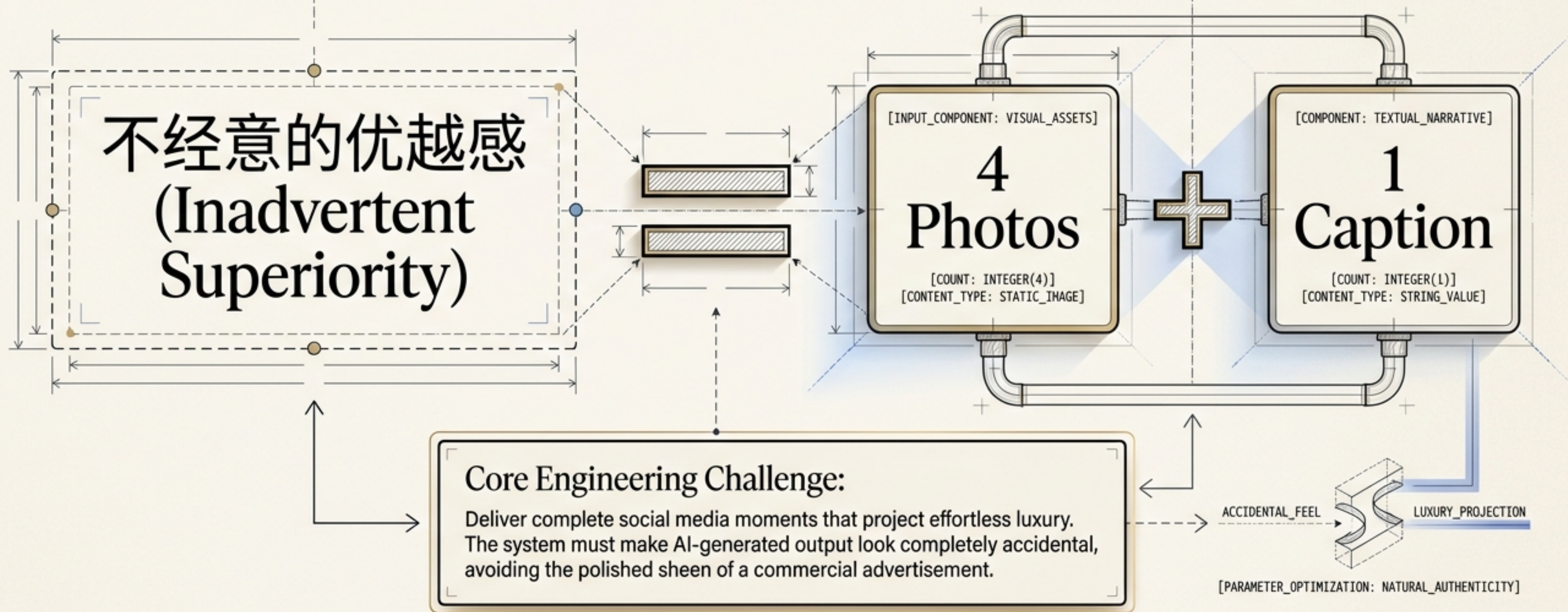


The background of the slide is a detailed architectural floor plan of a kitchen. It features a central island with a sink and a stove, surrounded by cabinetry. The drawing includes various lines for walls, doors, and furniture, along with dimension lines and symbols. Some elements are highlighted with blue and yellow colors. The overall style is technical and precise.

Under the Hood: How ÉLAN Talks to Gemini

Architecture, prompt systems, and the true
unit economics of AI-generated luxury.

The Anatomy of Inadvertent Superiority



The 10-Section Prompt Architecture



[ZONE_4_TITLE]
“Zone 4: The Overrides”
[ZONE_4_DESCRIPTION]
VANITY_DESIGN_INSTRUCTIONS.

> OVERRIDE::VANITY_DESIGN_INSTRUCTIONS
> FORCE_STYLE::[EDITORIAL_SPREAD, PATENT_DRAWING_OVERLAY]

[ZONE_2_TITLE]
“Zone 2: The Stage”
[ZONE_2_DESCRIPTION]
Scene Description (Native Chinese)
+ Narrative Shot Role.

> OUTFIT_HINT::[CHANEL_TWEED_JACKET, VINTAGE_LEVIS]
> POSE_PROMPT:="CASUAL_CONFIDENCE, MID-STRIDE"
> MOOD_COLOR_GRADING:="FILM_GRAIN, WARM_CINEMATIC"

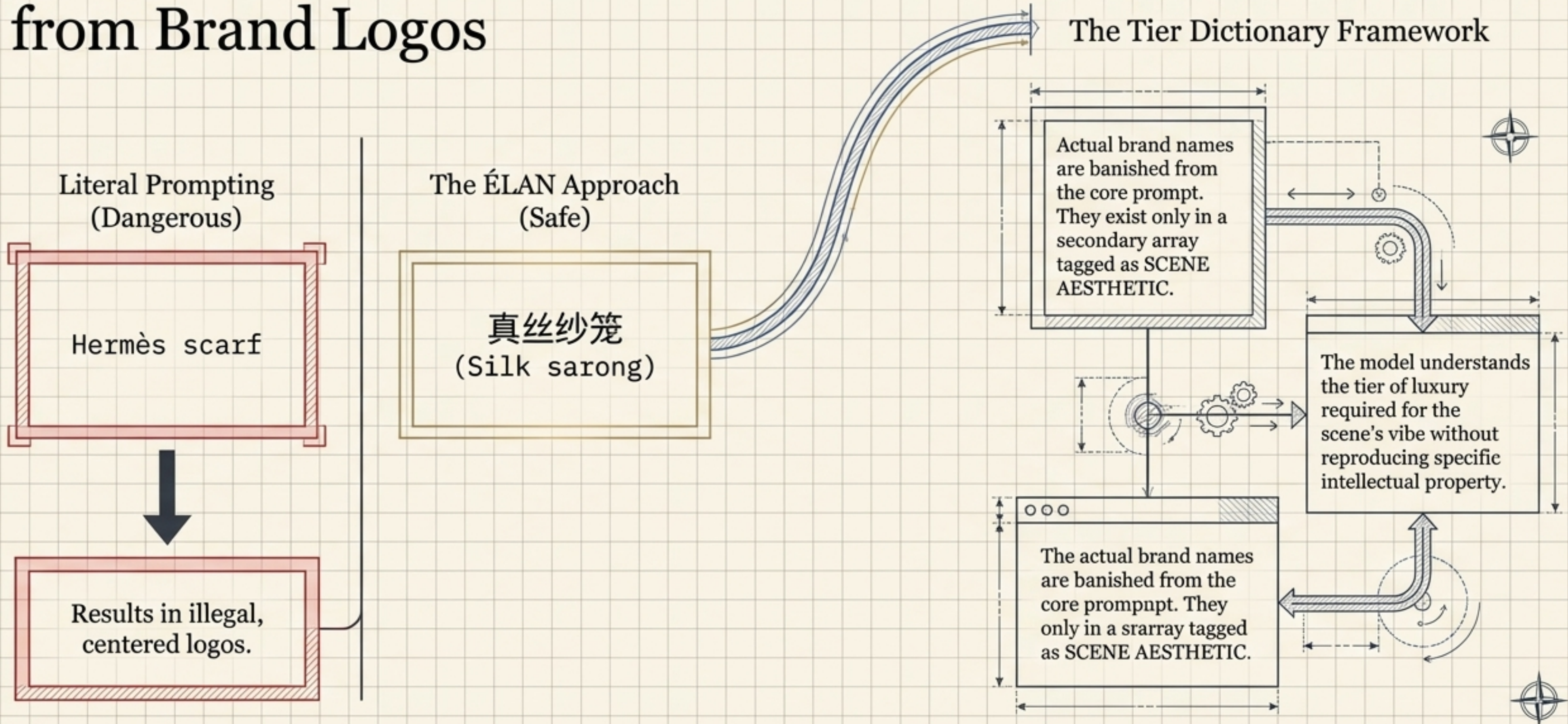
[ZONE_3_TITLE]
“Zone 3: Aesthetic Calibration”
[ZONE_3_DESCRIPTION]
Outfit & Luxury Hints + Pose
+ Prose Color Grading + Mood.

> SCENE_DESC_ZH:="在上海法租界梧桐树下..."
> SHOT_ROLE::NARRATIVE_PROTAGONIST

[ZONE_1_TITLE]
“Zone 1: The Anchors”
[ZONE_1_DESCRIPTION]
Reference Image Annotation
+ Identity Lock Footer.

> ANCHOR_SET::[REF_IMG_ID=X5A990B]
> INJECT::IDENTITY_LOCK_FOOTER=TRUE

Decoupling Luxury from Brand Logos



The Narrative Quartet



Establishing

Wide shot, sets the scene and context.

Portrait

Medium shot, focuses on the user's face.

Detail

Close-up of a specific physical element.

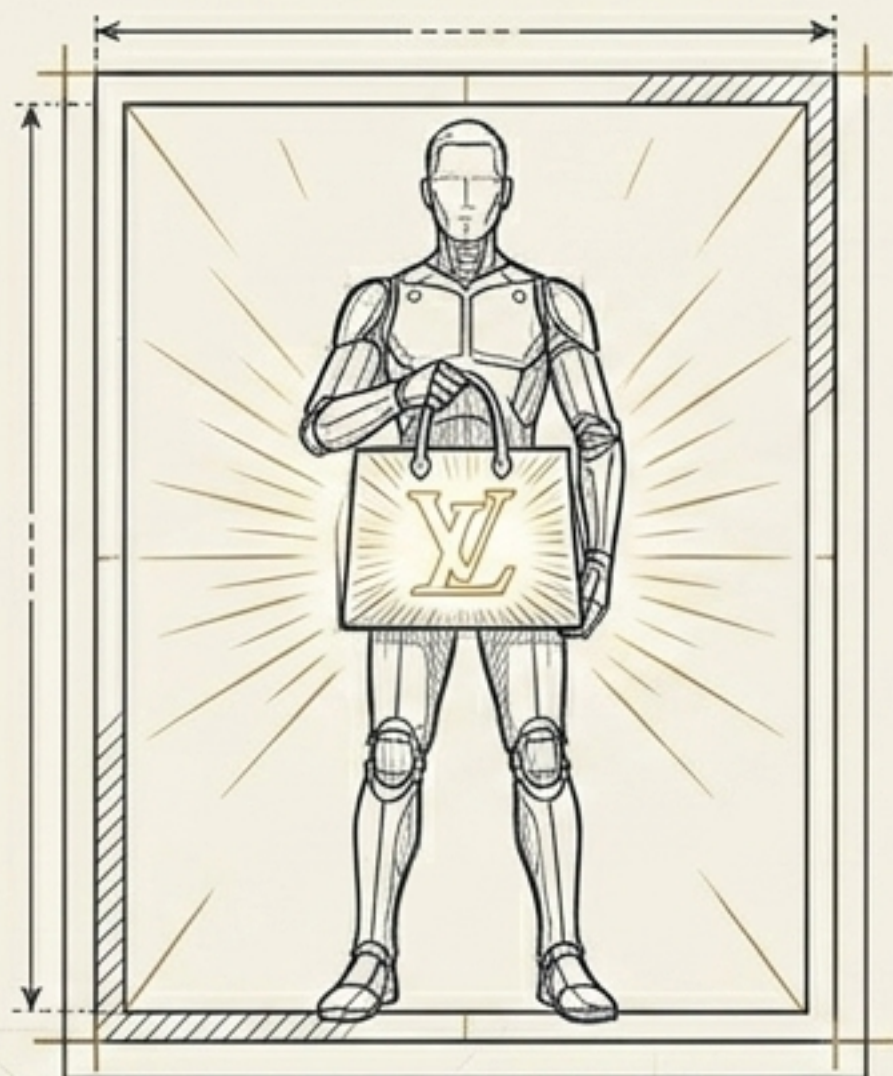
Mood

Atmospheric closing shot, back-to-camera.

Generating four distinct compositions creates a cohesive visual story rather than four random outputs.

The Differentiator: VANITY_DESIGN_INSTRUCTIONS

Without Instructions



Commercial Ad Aesthetic

Centered subjects, perfect lighting, highly visible logos. Fails the authenticity test.

With Instructions

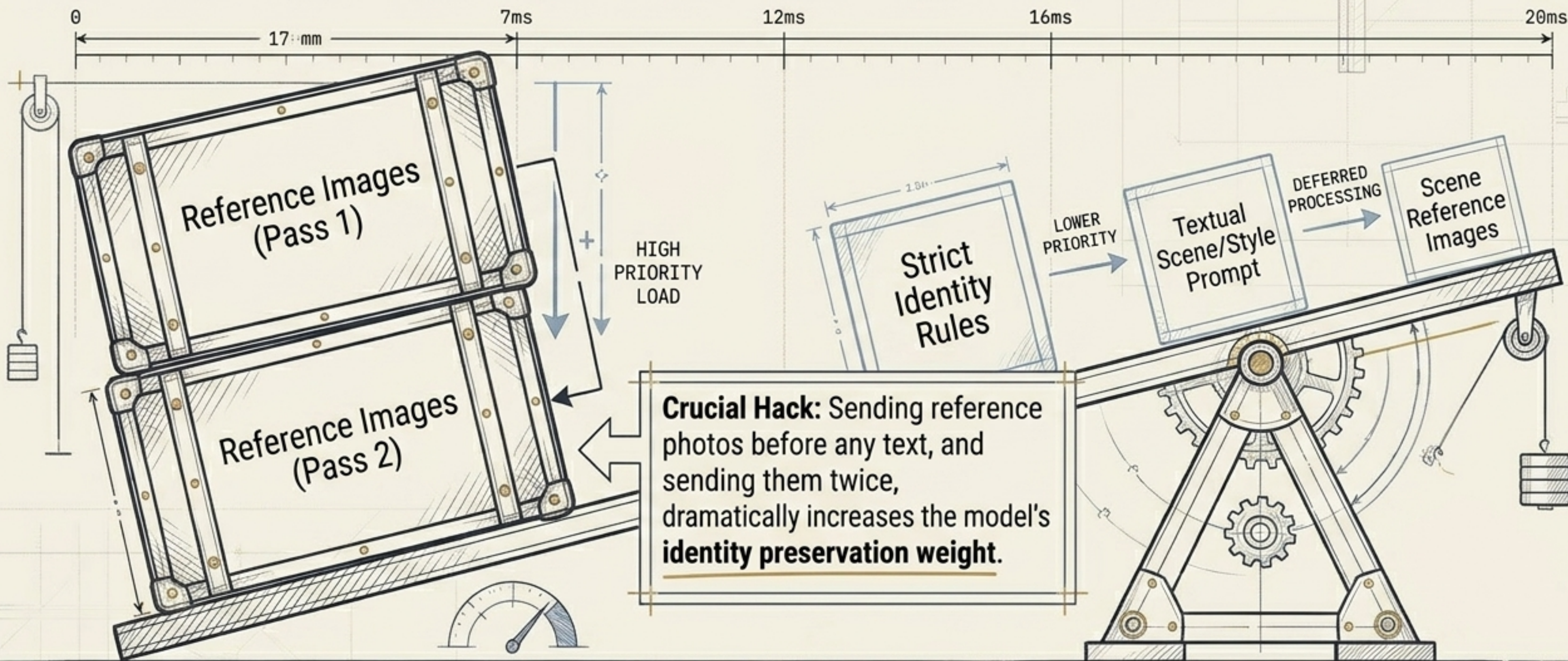


Example

Candid life moment, NOT an advertisement.
This is just my normal Tuesday.

The Image-First Attention Hack

Order dictates attention. Sending images early forces the model to prioritize facial preservation.



The Dual-Identity Lock



Use the uploaded image(s) strictly as the **FIXED IDENTITY REFERENCE**.

Aesthetic Rules: Rich descriptions of outfit, mood, lighting, and scene.



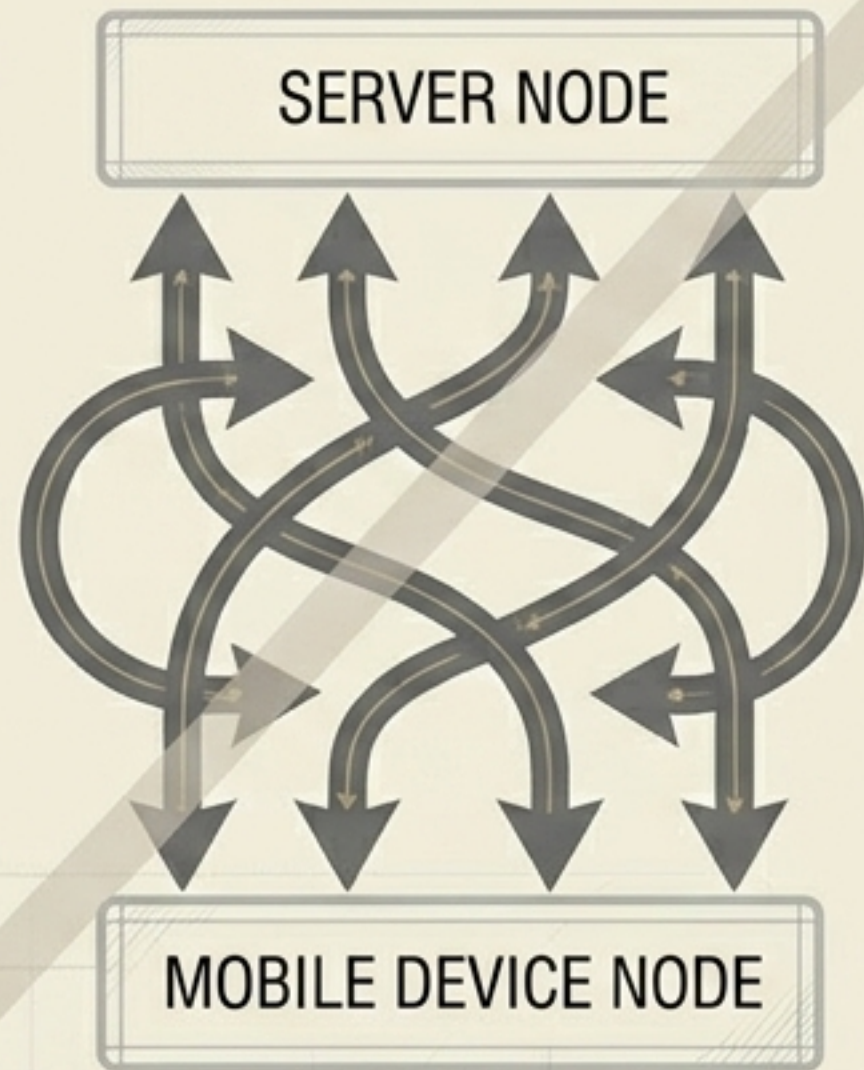
Identity Fallback: If style conflicts with identity, choose identity.

The **fallback** rule prevents the model's inherent bias to **sacrifice** facial accuracy in order to perfectly match a stylistic direction.

Streaming 4 Parallel Generations

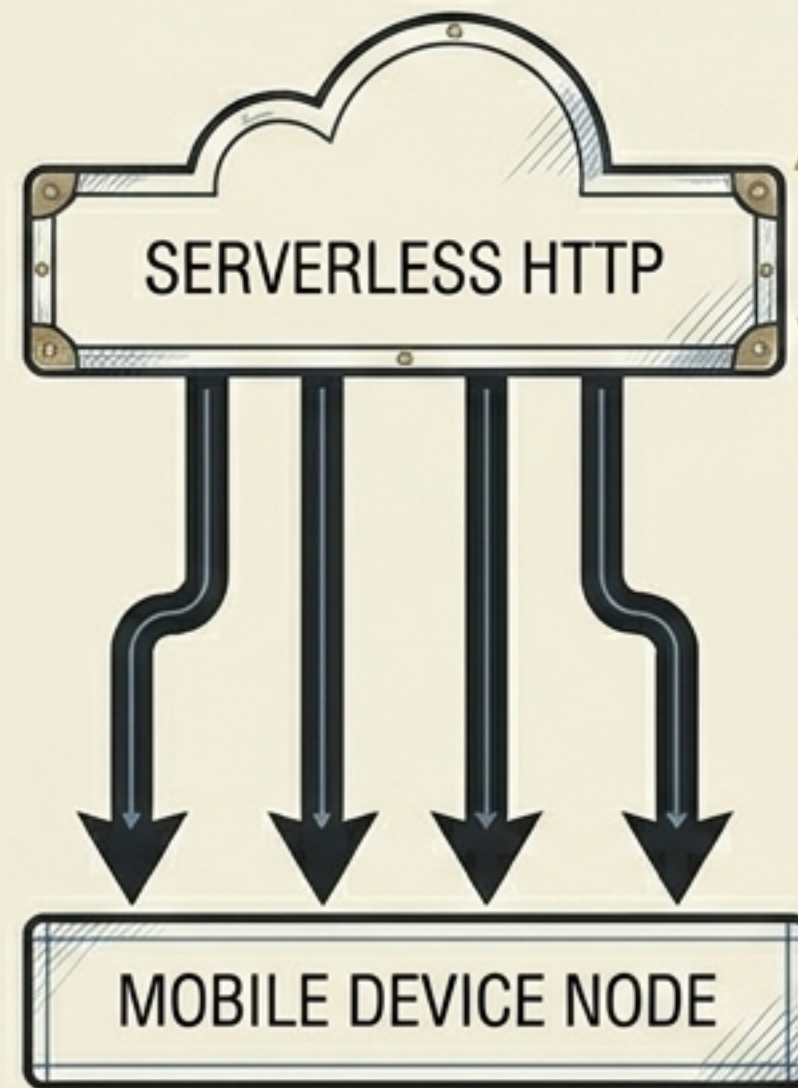
Why Server-Sent Events (SSE) won over WebSockets to solve the 60-second wait constraint.

WebSocket



Bidirectional overhead is wasted.

SSE Architecture



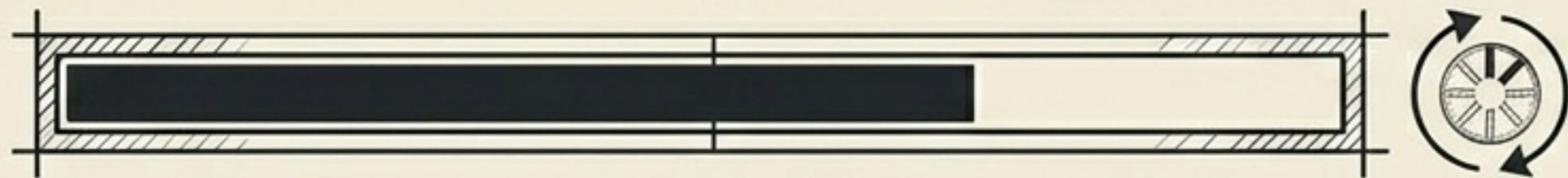
Unidirectional Flow: Client only receives data; no unused upstream overhead.

Serverless Friendly: Works natively on Vercel edge functions without persistent connection pools.

Natural Reconnection: Built-in protocol handles dropped mobile networks automatically.

Disguising Latency Through UX

The Traditional Approach



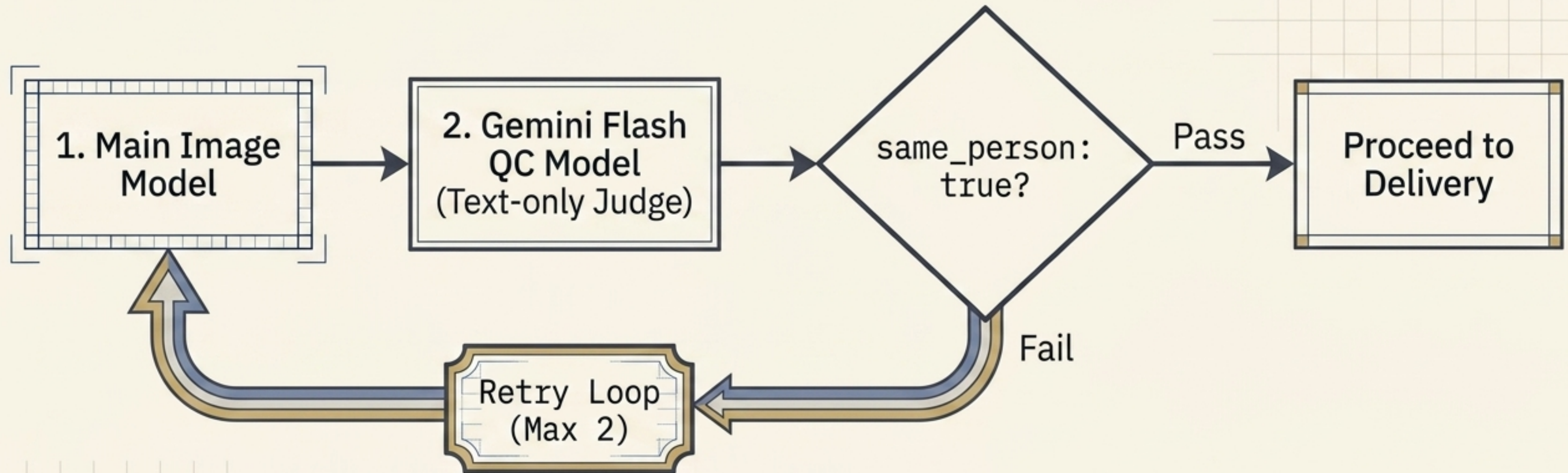
Progress bars imply predictable duration, which LLM generation lacks. Waiting feels like an eternity.

The ÉLAN Approach: Satisfying Arrival



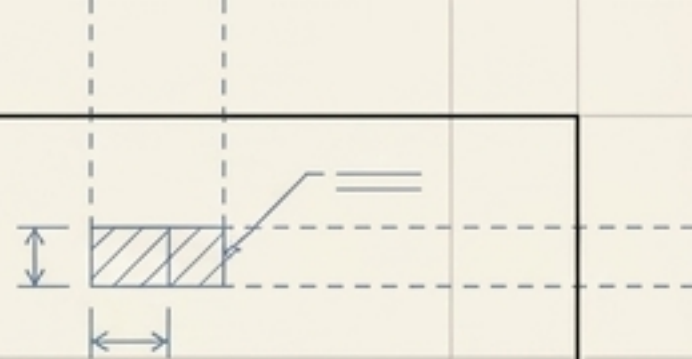
Photos arrive one by one as the SSE stream fires completion events. Completed photos replace placeholders with entrance animations, turning wait time into an anticipated experience.

Solving Face Drift with a QC Loop



- * Catches the worst 15-20% of face drift failures.
- * Retries rescue ~80% of those failed generations.
- * Cost: Fraction of a cent. Latency: +2 to 5 seconds. UX ROI: Immeasurable.

The Caption Engine: 3 Distinct Personalities



Versailles (凡尔赛)

The Humble-Brag

Philosophy: Captions describe a mundane annoyance while the photo reveals effortless luxury.

Constraints: 200 chars (XHS) / 80 chars (WeChat).

Poetic (文艺)

The Image-Maker

Philosophy: Short and imagistic. Express feelings through natural metaphors. No narration.

Constraints: 150 chars (XHS) / 60 chars (WeChat).

Minimal (简约)

The Whitespace

Philosophy: Ultra-short, one phrase. Maximum whitespace. The image speaks for itself.

Constraints: 50 chars (XHS) / 30 chars (WeChat).



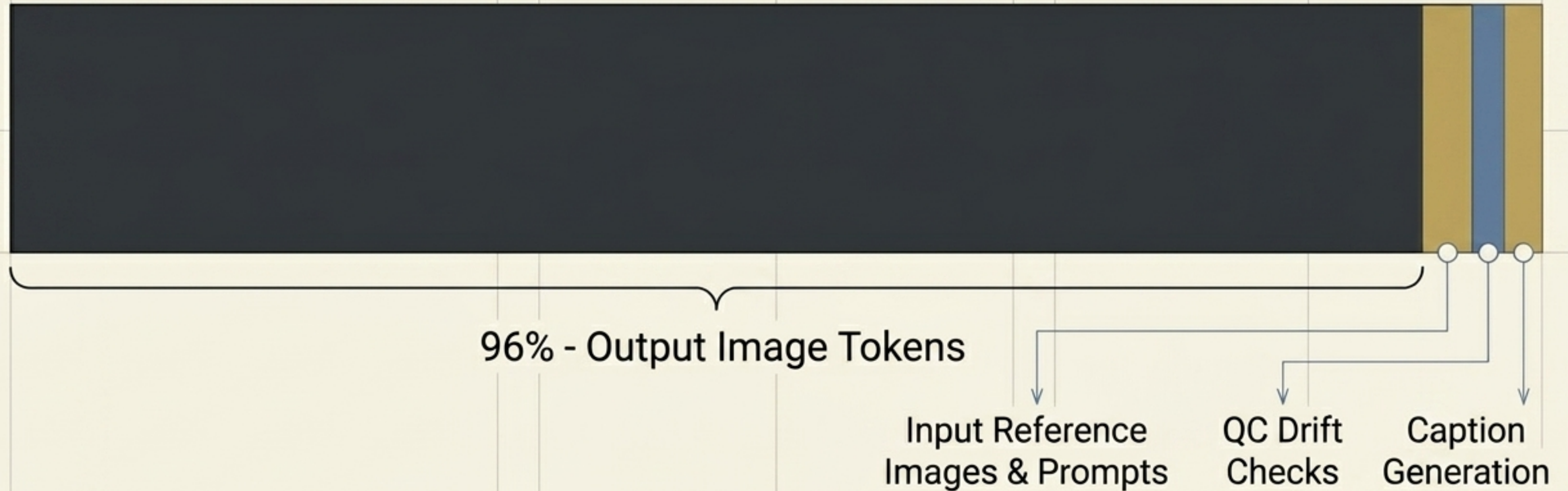
The Golden Rule of Captions

**禁止提及任何品牌名、
酒店名、餐厅名、景点名。**

(Never mention any brand name, hotel name, restaurant name, or landmark name.)

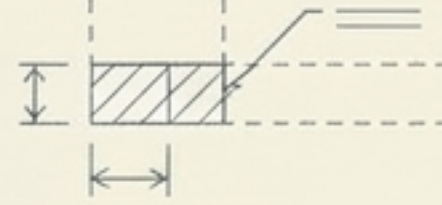
If the caption text explains the luxury, the illusion of effortless is immediately destroyed. The luxury must live exclusively in the unspoken context of the photo.

The 96% Reality of Unit Economics

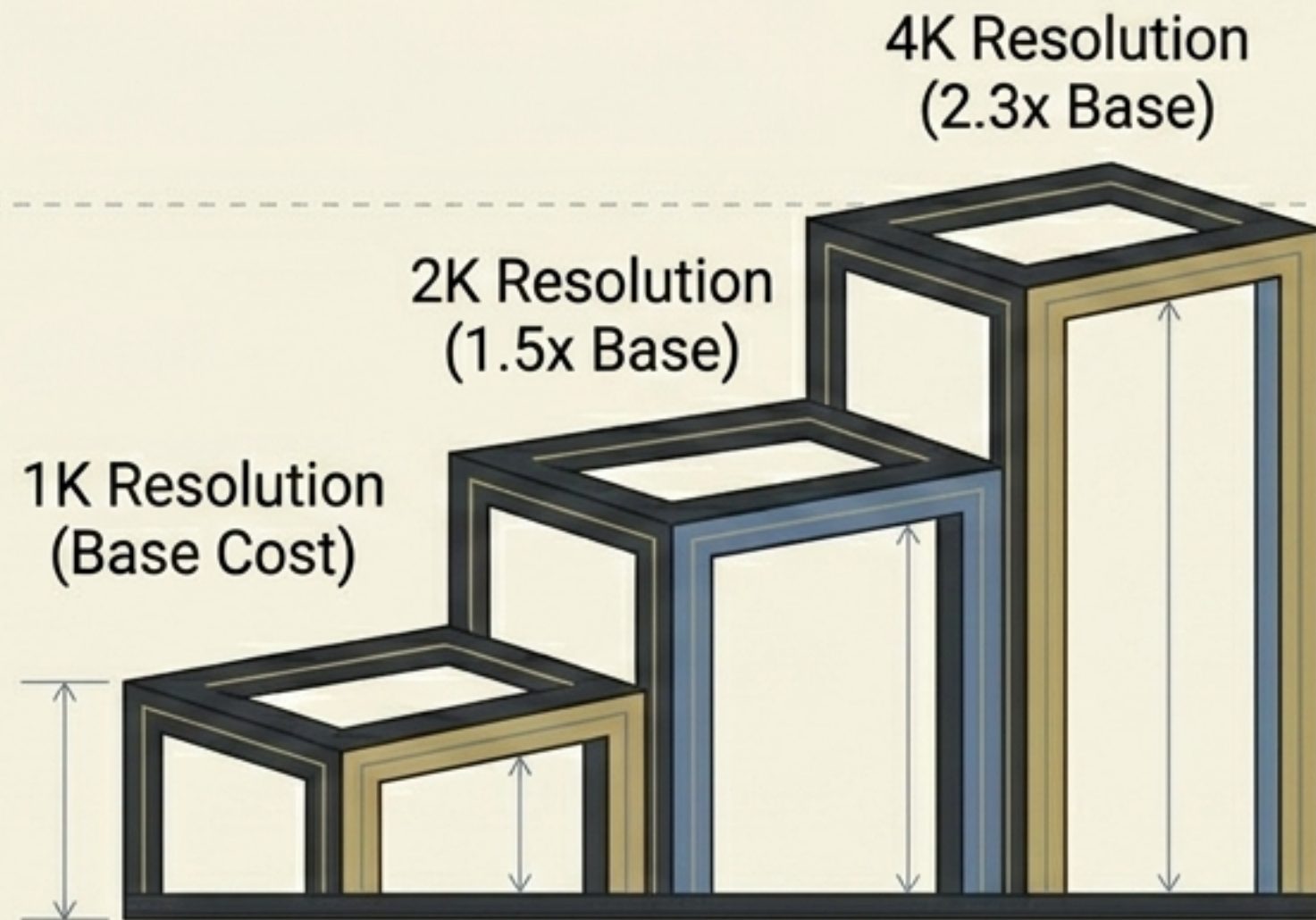


Generating images is roughly 120x more expensive per token than text. This mathematical reality killed the unlimited beta plan and forced a strict credit system to avoid bankruptcy.

The Hidden Multipliers: Resolution and Retries



The Baseline Math

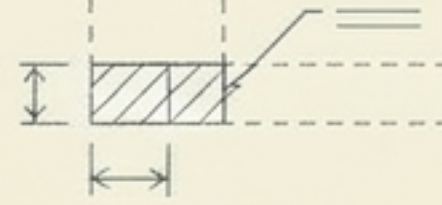


The QC Wildcard



The retry loop guarantees quality but shatters predictable economics. If a photo fails QC and retries twice, that single shot costs 300% of the base rate.

Harsh Realities & System Constraints



01. Prompt Tension



Rich scene descriptions inherently compete for attention with reference images, slightly degrading facial accuracy.

03. Brand Safety

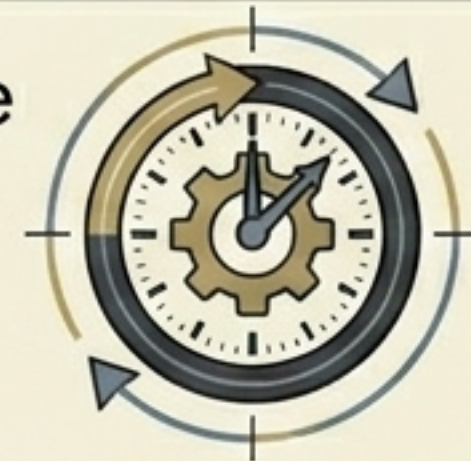


True luxury items are a legal minefield, forcing the migration to generic item descriptions and tier-based aesthetic hints.

02. The Identity Limit

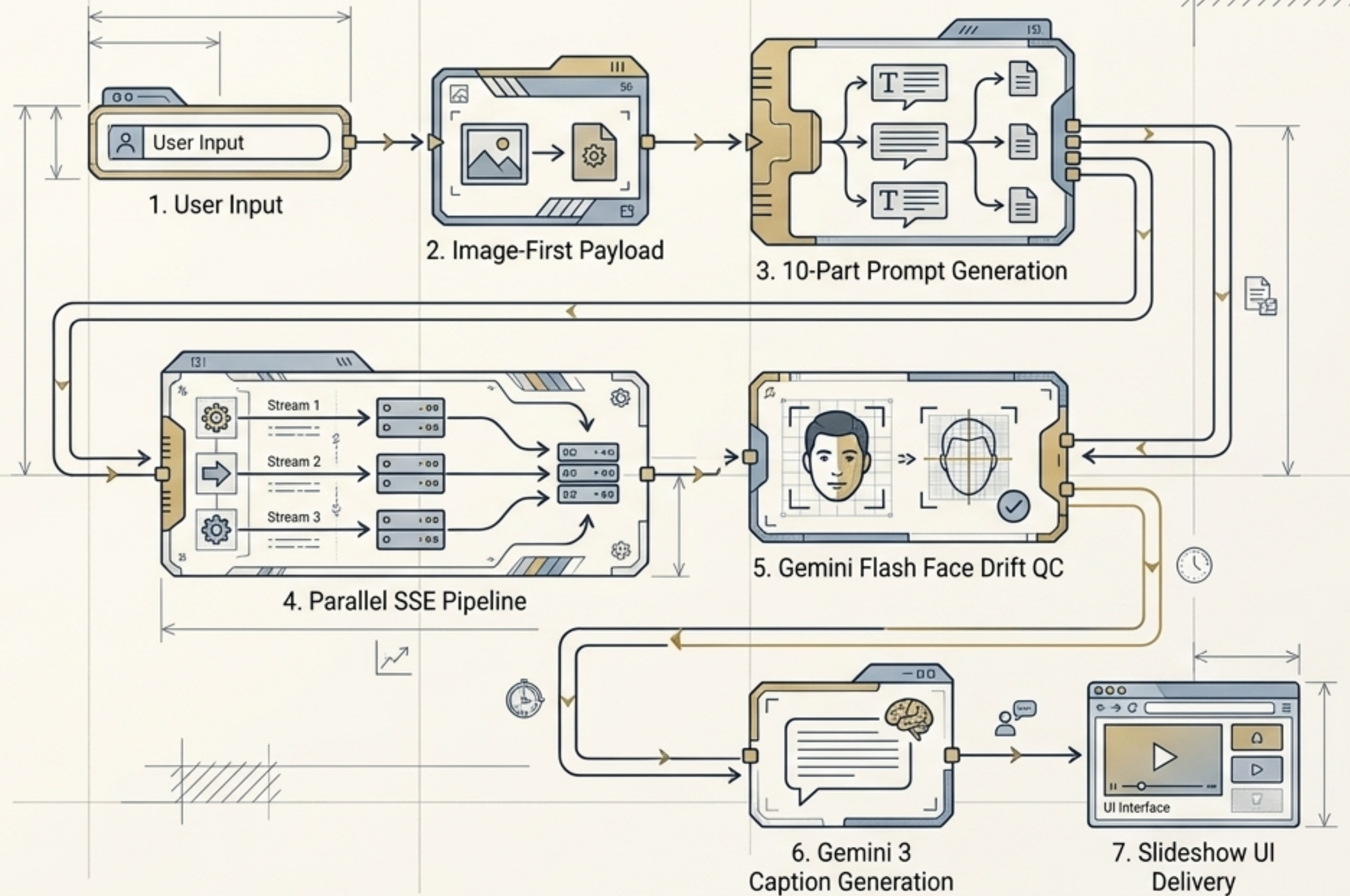
Because 4 photos use 4 independent API calls, true identity locking across a session is impossible without model fine-tuning, without mst and noncontinuity.

04. Storage Costs



Vercel Blob storage overheads necessitated a strict 2-Hour TTL on generated images, creating user friction.

The Complete ÉLAN Engine Architecture



Inadvertent luxury isn't a clever prompt; it's a meticulously layered machine.