

System Technical Architecture & Generation Pipeline

A comprehensive technical breakdown of Hatao AI's autonomous visual intelligence platform. How the system transforms natural language intent into layered vector DOM layouts, professional print typography, isolated raster graphics, and real-time motion clips.

100%
DOM LAYER
SEPARATION

300 DPI
PRINT PRECISION
OUTPUT

< 150ms
REALTIME SSE
STREAM

10+ AI
SPECIALIZED MODELS

5 Formats
PNG / PDF / PPTX /
HTML / MP4

THE FUNDAMENTAL PARADIGM SHIFT: RASTER PIXELS VS. VECTOR DOM

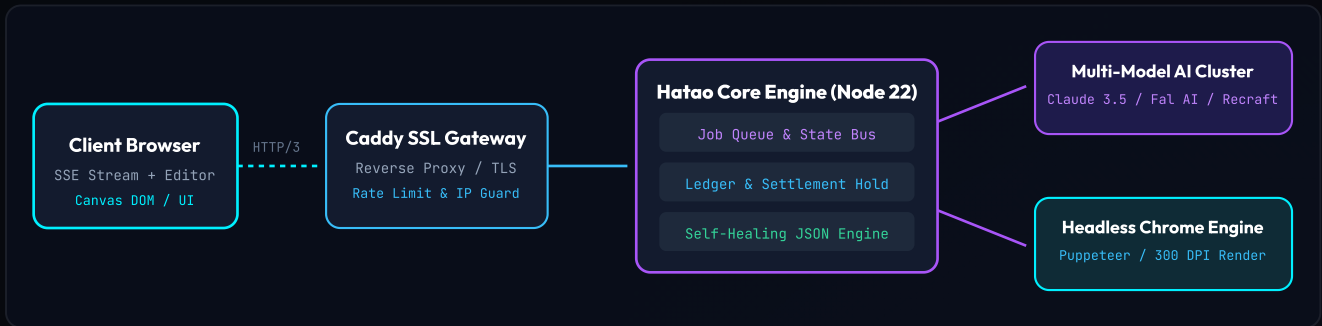
✗ Legacy Generative AI (Midjourney, DALL-E, Ideogram)

- Flat Pixel Baking:** Text, badges, and background are fused into uneditable RGB bitmaps.
- Typo Destruction:** A single spelling mistake requires a complete regeneration roll.
- Zero Brand Control:** Inability to adjust specific hex colors, reposition copy, or swap logos.
- Low Resolution Traps:** Upscalers blur fine hairline strokes and geometric rules.
- Static Deliverable:** No route to editable HTML, PowerPoint, or dynamic motion layers.

⚡ Hatao AI: Autonomous Semantic DOM Synthesis

- Separated Vector Layers:** Typography lives in native CSS/DOM text nodes; graphics are isolated assets.
- Instant WYSIWYG Editing:** Click any headline or body block to edit wording, font, size, or alignment.
- Deterministic Brand Kits:** Seamless one-click injection of corporate fonts, hex palettes, and slogans.
- True Print Resolution:** Native vector fonts render razor-sharp at 300 DPI up to 2400px canvas bounds.
- Omni-Format Export:** High-res PNG, vector PDF, editable PowerPoint PPTX, native HTML, and video clips.

HIGH-LEVEL TOPOLOGY & DATA FLOW



The 7-Stage Autonomous Visual Pipeline

From prompt ingestion to final multi-format bundle delivery, every creation request traverses seven strictly isolated, idempotent processing stages designed for sub-second failure detection and zero-loss recovery.

01 Ingestion, IP Origin Guard & Dual-Ledger Hold SAFETY & ACCOUNTING

The request enters through `POST /api/designs`. Network origin is normalized and hashed via HMAC-SHA256 salted keys to enforce the strict **1-free-trial per origin policy**. If authenticated or paying, credits are calculated via `models.estimate()` and placed into a non-overdrawing atomic hold inside the account ledger before any upstream API token is consumed.

02 Dynamic Canvas Sizing & Creative Art Direction SPATIAL INTENT

The Art Director LLM evaluates user intent to select the optimal physical aspect ratio, pixel boundaries (700px – 2400px), and visual aesthetic direction (editorial, luxury, brutalist, playful, or technical). Format is locked as poster, flyer, billboard, or social card with explicit spatial rationale.

03 Composition Reference Synthesis VISUAL RHYTHM

A prompt enhancer generates an atmospheric composition reference via Fal AI (Recraft V3 / Flux 1.1). *Crucial architectural distinction:* this reference image is stored solely under `reference/` for layout guidance and spatial hierarchy analysis; it is **never** sliced or shipped directly into client assets.

04 Integrated Art Direction & Layer Planning TOPOLOGY DECISION

Claude 3.5 Sonnet / Opus 5 performs deep multimodal reasoning across the reference image and user brief. It outputs a comprehensive plan selecting one of five distinct asset topologies: `code-native`, `slot-matrix`, `continuous-scene`, `cutout-stack`, or `layered-collage`, alongside font stacks, exact live wording, and hex color roles.

05 Autonomous Asset Synthesis & Transparency Masking ASSET ENGINE

If the chosen topology requires visual elements, prompts are programmatically injected with strict negative constraints (No lettering, words, logos, or UI). Assets are synthesized in parallel via Fal AI with transparency channels enabled, then imported into the project directory via isolated sandbox scripts.

06 Code-Native Semantic Layout Synthesis DOM GENERATION

The coding engine synthesizes a self-contained, standards-compliant `index.html` file with encapsulated CSS styles. All textual elements are written into distinct `<div data-layer-id="...">` container nodes with CSS flex/grid layout rules, high-contrast typography, and explicit bounding boxes.

07 Headless Chrome Preflight, Font Stack & Critique Verification QUALITY CONTRACT

Puppeteer executes a headless Chromium instance (with `--no-sandbox` isolation under Linux root) to inspect the rendered design. It verifies font availability against installed typography stacks, tests for container collisions, takes high-res screenshots, runs a visual critique pass with the vision model, and delivers the finalized bundle.

⚡ Resilient Self-Healing LLM Extraction Engine PRODUCTION HARDENED

Large creative design plans often exceed 10,000 tokens. Minor LLM syntax anomalies—such as unescaped literal newlines in multi-line markdown fields, trailing commas, or markdown code-fence wrappers—are automatically parsed and repaired via our multi-pass byte scanner before parsing, eliminating pipeline crashes and preventing stranded jobs.

Multi-Model AI Orchestration & The In-Browser Studio

Hatao AI does not rely on a single monolithic provider. We orchestrate specialized foundational models across reasoning, graphics, and video motion, coordinating their strengths through unified task adapters.

FOUNDATION MODEL CATALOG & TIER ALLOCATION

Task Category	Primary Engine	Provider	Role & Specialized Capability	Credits / Unit
Layout Reasoning	Claude 3.5 Sonnet / Opus 5	Anthropic	Spatial hierarchy, typographic pairings, semantic DOM layout synthesis	4 - 8 cr
Visual Verification	Claude 3.5 Sonnet (Vision)	Anthropic	Preflight layout critique, contrast analysis, container overlap detection	4 cr
Vector Graphic Art	Recraft V3 (SVG & Raster)	Fal AI	Clean graphic artwork, isolated icons, vector illustrations with alpha	8 cr
Photorealistic Imaging	Flux 1.1 Pro / Dev	Fal AI	Photographic depth, complex textures, composition reference baselines	6 - 10 cr
Cinematic Motion	Kling 2.5 Turbo Pro	Fal AI	Dramatic camera motion, fluid 3D transformations of finished designs	35 - 70 cr
Synchronized Video	Wan 2.5 (Audio Motion)	Fal AI	HD video clips (720p/1080p) with synchronized cinematic audio soundtrack	25 - 50 cr
Dynamic Camera Pan	Luma Dream Machine	Fal AI	Immersive 3D push, orbit, and zoom camera physics around graphic assets	30 cr

THE INTERACTIVE VISUAL STUDIO EDITOR ARCHITECTURE

Interactive DOM Canvas Engine

Unlike canvas-based raster tools that draw non-selectable pixels on a single bitmap, Hatao AI's editor mounts a live HTML/CSS document directly inside a responsive sandboxed iframe with bidirectional postMessage bindings:

- data-layer-id tagging on every visual component.
- Sub-pixel draggable position handles and resize anchors.
- Live inline text editing with contenteditable synchronization.
- Layer hierarchy reordering (z-index plane management).

Deterministic Brand Kit Automation

Corporate visual identities are stored as structured JSON profiles attached to the user account. During generation and editing, brand assets are dynamically injected into CSS custom properties:

```
:root { --brand-primary: #00f0ff; --brand-secondary: #a855f7; --brand-accent: #38bdf8; --brand-font: 'Outfit', sans-serif; --brand-logo: url('/data/brand-kits/logo.svg'); }
```

MOTION SYNTHESIS: FROM STATIC GRAPHIC TO CINEMATIC VIDEO

Once a design is finalized, Hatao AI offers single-click motion generation (POST /api/videos). The backend extracts the verified high-resolution poster PNG as the keyframe anchor (first frame), validates aspect ratio and duration constraints, and dispatches the job to Fal AI's Kling or Wan video clusters. The resulting 60 FPS video clip is rendered and attached to the project package.

Omni-Format Export Engine & Production Infrastructure

A single creation session in Hatao AI produces an entire ecosystem of synchronized marketing deliverables—print, web, presentation decks, and video motion—all generated from the same unified design source.

THE 5 UNIFIED DELIVERABLES

300 DPI Print-Ready PNG

RASTERIZATION

Rendered via headless Chromium GPU pipeline at the precise declared pixel dimensions (e.g. 1600×2400). Validated via binary IHDR chunk dimension checks to guarantee zero layout clipping or scale drift.

Native Vector PDF

HIGH PRECISION

Direct headless PDF synthesis with embedded web fonts and CMYK-compatible color boundaries. Retains vector clarity for commercial print production and high-volume billboard runs.

Editable PowerPoint Deck (.PPTX)

ENTERPRISE

Python-driven layout translation (via python-pptx). Maps DOM text nodes to native Microsoft PowerPoint editable text boxes and shapes, enabling executive presentations in corporate environments.

Cinematic Motion Clip (.MP4)

SOCIAL & ADS

H.264 video rendering with synchronized audio. Ready for immediate deployment across Instagram Reels, TikTok, YouTube Shorts, and programmatic digital ad displays.

PRODUCTION INFRASTRUCTURE & SECURITY SAFEGUARDS

Deployment Topology

Dedicated Cloud Droplet (206.189.138.142). Node.js 22 LTS under systemd supervisor (hataoai.service). Caddy HTTP/3 reverse proxy with automated Let's Encrypt TLS renewal.

Anti-Abuse & Origin Protection

Salted HMAC-SHA256 origin fingerprinting prevents trial evasion via cookie clearing. Same-Origin verification headers block CSRF attacks on all mutating endpoints.

Atomic Payment Settlement

Idempotent payment webhooks (Cashfree & Stripe). Account holds guarantee zero provider overruns; unused reserved credits are automatically released on job completion.

SYSTEM HEALTH & VERIFICATION SUMMARY

Production Diagnostic Audit: 26/26 Tests Verified (100% Pass Rate)

Verified live end-to-end design generation (adc08c31-dba2-4106-ba6c-0112cb1e7b06), real-time SSE bus, credit settlement, headless Chrome sandbox bypass, and full multi-format asset downloads on production droplet.

ALL SYSTEMS OPERATIONAL