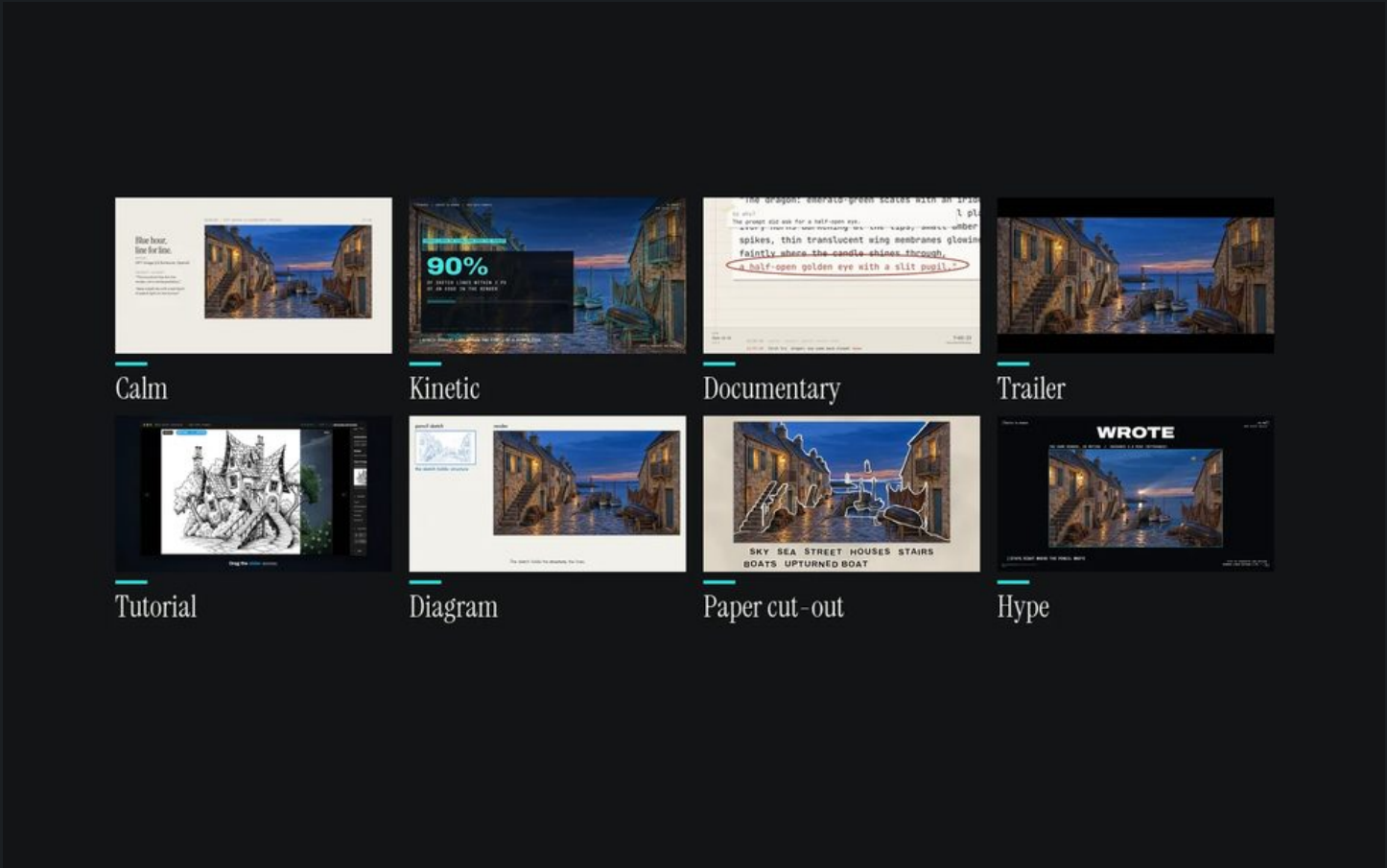


Eight ways to edit

A library of eight video editing styles, each a written skill an AI agent follows: its tone, its way of explaining and its rules for pace, cuts, camera, lettering, voice and music. To compare them fairly, all eight edited the same finished process with the same records: pencil sketches turned into finished renders in Scenario. This is how the library works and how each style edits.



Eight styles, eight skills.

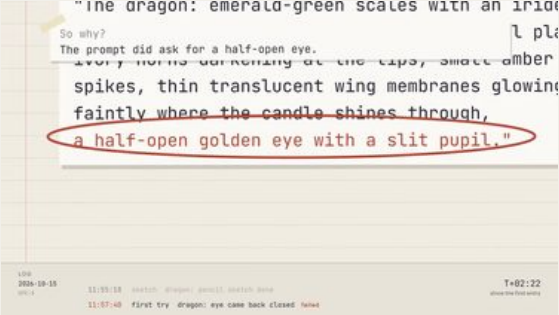
Each style is a skill: a written set of dials, a grammar and a worked example that an agent follows to edit a finished process into one video. Every style explains in its own way. Below, one frame of each final video and how it explains.



Calm
[calm-minimal-video](#)
Explains by showing a few results slowly, with room to look.



Kinetic
[kinetic-tech-showcase](#)
Explains by proving a technical result fast, on the beat.



Documentary
[documentary-process-video](#)
Explains by telling how the work unfolded, failures included.



Trailer
[cinematic-trailer-video](#)
Explains by building anticipation to a title.



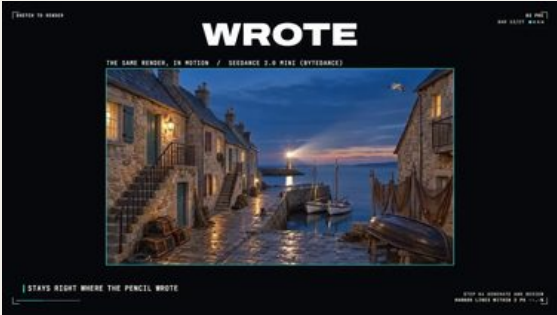
Tutorial
[tutorial-walkthrough-video](#)
Explains by doing it in the real tool, step by step.



Diagram
[diagram-explainer-video](#)
Explains by building a diagram of how it works.



Paper cut-out
[paper-cutout-video](#)
Explains by paper pieces landing on the input they fit.



Hype
[hype-music-video](#)
Explains by celebrating it in an original song.

What drives each edit.

The same minute of material, cut eight ways: what sets the pace, how often the full frame changes and what the camera does. Numbers are measured on these videos; the rest follows each style’s written rules.

STYLE	WHAT DRIVES THE PACE	CUTS	CAMERA
Calm	the voice and long holds; voice on 37% of the time	at most one per 3 s (style rule); dissolves	slow drift inside each hold; results in wide margins
Kinetic	a 120 BPM beat; each section starts on a bar	many, on the hits	punch-ins on hits; HUD chrome around real UI
Documentary	five beats: problem, why, tried, fix, result	a page on the beat, about one every 2 s	handheld desk camera over notebook paper
Trailer	the cue: accelerating to quarter-second flashes	speeding up, then 0.7 s of digital silence	slow pushes, letterboxed 2.39:1
Tutorial	five steps over a 100 BPM bed	at most one screen change per 4 s (style rule)	one continuous glide over the real app
Diagram	the voice over a 100 BPM bed	none: the diagram morphs	pushes in on the named part; one 3D parallax moment
Paper cut-out	each piece lands on the word that names it	none: whole sheets slide; motion on twos	locked-off overhead rostrum
Hype	a 127 BPM song	on the bar or the beat	wipes and swaps in the same framing

Where the words land, and who says them.

Seven styles are narrated with Gemini 3.8 Flash TTS, a Google model, in six voices; the hype video is sung. Beds come from ElevenLabs Music v2.5 and sound kits from Sonilo; the song from MiniMax Music 3.0.

STYLE	LETTERING KEYED TO WORDS	VOICE	MUSIC AND SOUND
Calm	serif headline settles on the key word; caption words brighten as spoken	Sulafat, 74 wpm	ambient bed, no effects
Kinetic	display slams and data chips on words and beats; karaoke subtitle	Puck, 107 wpm	120 BPM instrumental; whoosh, hit, tick, click
Documentary	narration types on as spoken; job times stamp on their word	Callirrhoe, “we”, 130 wpm	acoustic bed, 90 BPM; page, stamp, pencil
Trailer	cards on the narrator’s stressed words; no running captions	Algenib, 43 wpm	trailer cue cut by hand; the clips’ own sounds
Tutorial	a ring completes on the named control; the spoken word lights up	Zubenelgenubi, 129 wpm	light bed, 100 BPM; click, key, whoosh, tick
Diagram	boxes, arrows and labels draw on the word that names them	Fenrir, 132 wpm	curious bed, 100 BPM; pen, tick, lift, fan
Paper cut-out	cut-paper letters, typed slips, the prompt on a torn strip	Sulafat, 122 wpm	handmade bed; a paper tap on every landing
Hype	every sung word on its onset; display words fill the frame	sung, 88 wpm	the song, nothing else

Different edits, one set of rules.

Under the eight styles runs one audio and proof spine. It is why the lettering lands on the word and why the comparison stays honest.

Natural narration

Text written as speech: short clauses, contractions, a breath tag at the start of each take, no stage directions. Two takes per section, picked by measured pitch range and pace.

Music fitted to the voice

Beds generated in sections that follow the script, then cut around the narration. The paper cut-out repeated bars 16 to 18 so its final chord rings under the logo.

Every claim on screen

Nothing is said that is not visible at that moment: the overlay under “ninety percent”, the closed eye under “closed”. A word sheet checks each claim.

Word-timed lettering

Each chosen take is transcribed with faster-whisper and matched to its script, so every word has an onset in the video’s timeline. In the trailer, “Color.” lands at 14.96 s on a hard cut from sketch to render.

A measured mix

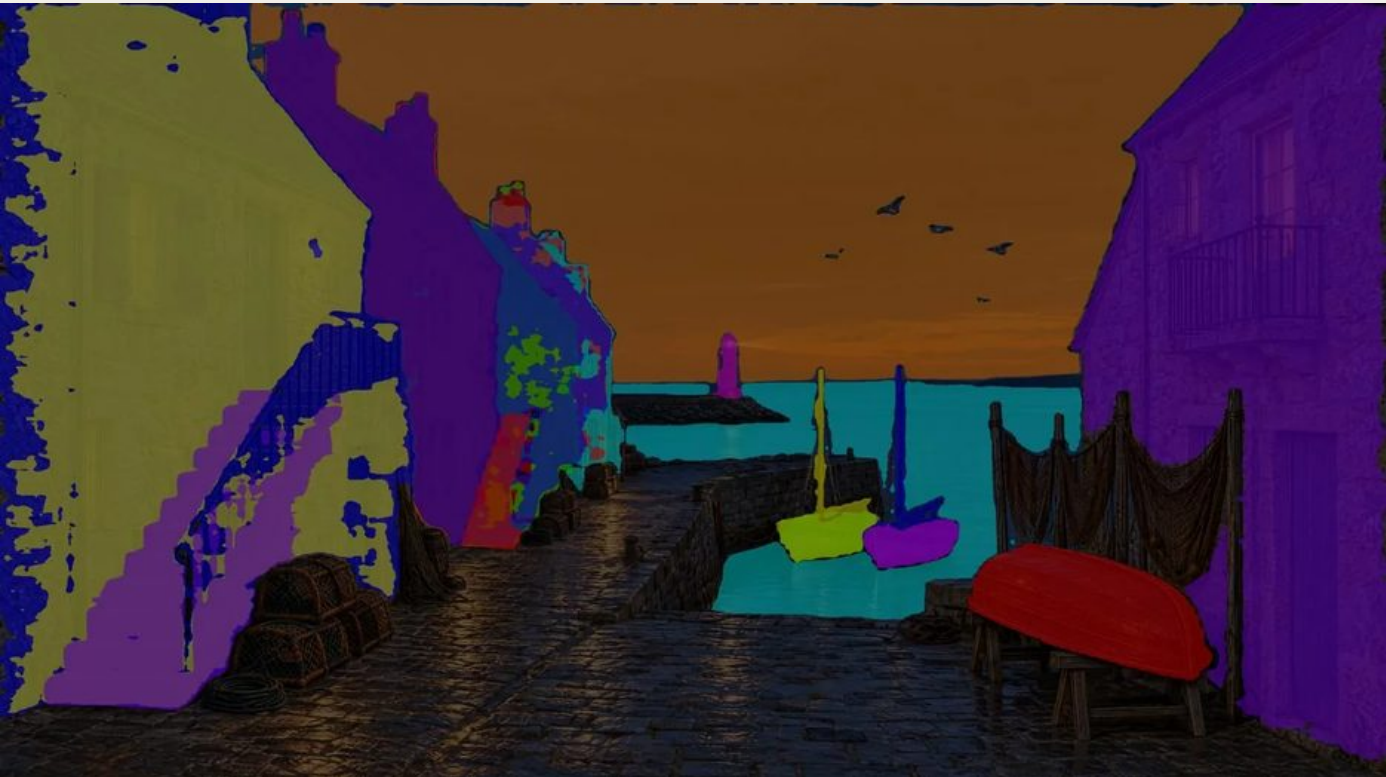
Loudness, true peak, voice over music, pace, dead air and intelligibility measured on every final mix. All eight land between -14.8 and -14.1 LUFS.

The mono logo ending

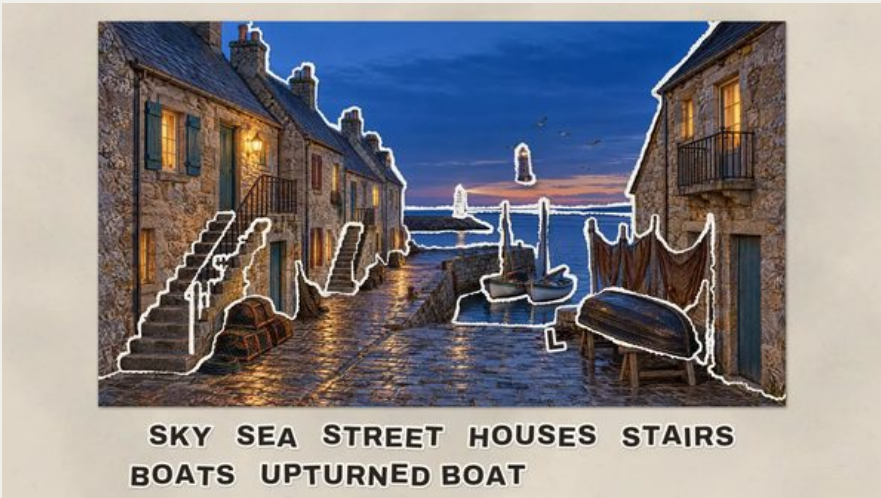
Every video ends on the Scenario logo in one color: black on light pages, white on dark ones.

The paper cut-out: pieces from the render's own pixels.

SAM 3.1 Image, a Meta model, made masks from text prompts on the harbor render: lighthouse, boat, stone staircase, house, sky and sea. The street is every pixel no other piece claimed. The two dragons were masked the same way.



The harbor render with its masks: each color is one paper piece.



The paper cut-out video: the harbor assembled piece by piece.

Put back together without their paper borders, the nine harbor pieces match the render exactly: every pixel belongs to one piece, and the largest difference is 0. On screen each piece lands on its sketch lines as its name is said, so the fit is the proof.

Every video ends on the same mark.

Where the logo lands follows each style’s closing grammar; its color follows the last frame.



Calm: black on paper, under the held line.



Trailer: white on black, on the silent tail.

Kinetic	white, on the end card
Documentary	black, in the credits band
Tutorial	white, held over the recap
Diagram	black, on plain paper
Paper cut-out	black, on the light table
Hype	white, under the credits

Same input for every style.

All eight videos tell one real process, Method 1 of the Help Center article Sketch-to-Render: four pencil sketches generated with Gemini 3.1 Flash (Google), each rendered with the sketch as the reference image and one detailed prompt. The inventor, dragon and scooter are on Gemini 3.1 Flash; the harbor is on GPT Image 2.5 Sunburst (OpenAI).



The harbor pencil sketch, 1376 x 768. All four sketches took 35 s.



Its render on GPT Image 2.5 Sunburst, 2752 x 1536, back at 19:22:37 UTC on October 15, 2026.

The harbor prompt, opening sentences, word for word

“Turn this exact pencil drawing into a photorealistic photograph of the same harbour village street at blue hour, just after sunset. This is a strict line-for-line render, not a reinterpretation: the photograph must overlay the drawing perfectly. Every pencil line becomes the edge of a real object at exactly the same position, size, angle and perspective.”

One number, one failure, one fix.

Each style had to carry the same facts. The number: 90% of the harbor sketch’s lines land within 2 px of an edge in the render, against 8 to 14% by chance. The failure and the fix: the dragon’s half-open eye came back closed.



Pencil sketch: half open



First try, 14:57:40 UTC: closed



Second try, 15:01:04 UTC: half open

Attempt 1: the eye, one phrase at the end of a list

“... thin translucent wing membranes glowing faintly where the candle shines through, a half-open golden eye with a slit pupil.”

Attempt 2: the eye gets its own instruction

“Keep the dragon's expression from the sketch: its visible eye is HALF OPEN, the upper eyelid lowered halfway over a glowing golden eye with a slit pupil, gazing sleepily toward the viewer. The eye must not be closed.”

A stricter prompt on another model.

The first harbor, on Gemini 3.1 Flash, let the two boats and the lighthouse drift off the pencil lines: 71% of the lines within 2 px. A strict line-for-line prompt on GPT Image 2.5 Sunburst (OpenAI) put them back: 90%. The prompts differ and each ran once, so this compares two real attempts, not two models.



Pencil sketch



Gemini 3.1 Flash, first prompt: 71%



GPT Image 2.5 Sunburst, stricter prompt: 90%

The Sunburst picture is busier, so it also scores a little higher by chance: other subjects' sketches gain 2.8 to 9.3 points on it. The harbor sketch gains 18.7, about twice the most that busyness explains, and the overlays show the boats and the lighthouse on their lines.

A same-prompt comparison, and motion.

Two more kinds of footage gave the styles something to cut to. The first dragon prompt on GPT Image 2, an OpenAI model, one run each on one sketch: it kept the eye but its book stack and head slide off the sketch lines. And each final render became the first frame of an 8-second clip on Seedance 2.0 Mini, a ByteDance model, at 480p with its own sound.



Gemini 3.1 Flash with the sketch lines: the books sit on them.



GPT Image 2 with the sketch lines: the stack slides off.



The harbor clip at 1.5 frames a second.

The clips are 8.1 s at 864 x 496 and 24 fps. The dragon’s eye is open only from 5.0 to 5.8 s and closed at the end, so the trailer and the hype video use only that window, and no video claims the eye stays open. The trailer shows the clips at their own size, never upscaled.

Made with Claude Code working through the Scenario MCP.

All models run on Scenario. Each is credited to its maker.

Gemini 3.8 Flash TTS, Google
narration, six voices

ElevenLabs Music v2.5
music beds

MiniMax Music 3.0, MiniMax
the song

Sonilo V1.1
sound effects

SAM 3.1 Image, Meta
paper cut-out masks

Gemini 3.1 Flash, Google
sketches; inventor, dragon, scooter

GPT Image 2.5 Sunburst, OpenAI
the harbor render

GPT Image 2, OpenAI
one comparison render

Seedance 2.0 Mini, ByteDance
five clips

The styles are skills for an AI agent, downloadable from the exploration page.

Every number in this document comes from the production records: job logs, word timelines, scripts and the measurements of each final mix.