

DOCUMENTARY PAPER ENGINE

SOURCE MATERIAL v2.0, VERTICAL EDITION (Instagram Reels / TikTok)

Attach this PDF when the engine prompt asks for source material (STATE 0). It grounds every stage: Fern writing DNA, beat math, the collage style system, worked demo prompts, the universal video prompt, ElevenLabs production rules, and thumbnail DNA with visual references.

This edition targets vertical 9:16 output (1080x1920) for Instagram Reels and TikTok, and works across all eight niches offered in State 1 (crime and documentary is the house default, not a limit).

SECTION 1: WHAT THIS DOCUMENT IS

This is the SOURCE MATERIAL for the Documentary Paper Engine v2.0. The engine prompt (.txt) is pasted into the chat first; this PDF is attached when the engine asks for it in STATE 0. Everything here is grounding data: the model follows this document instead of guessing. It holds the Fern writing DNA, the visual style system, beat math, worked demo prompts, the universal video prompt, ElevenLabs voice production rules, and the thumbnail DNA with two visual references.

House rule: no em dashes anywhere in any output. Use commas, colons, parentheses, or plain hyphens.

SECTION 2: THE FLOW AT A GLANCE

State 0: attach this PDF, engine absorbs it. State 1: pick a niche (crime and documentary is the house default, but any of the eight options works identically). State 2: engine returns 10 ideas, Fern-shaped titles, no two in the same sub-territory. State 3: pick a length (30s to 5 min). State 4: engine writes the Fern-style continuous narration script. State 5: type voice, engine generates or formats the ElevenLabs voiceover. State 6: type proceed, engine splits the script into 2-3 second beats with a timecoded table. State 7: type next, engine delivers ONE .txt file with one self-contained image prompt per beat, blank-line separated, no numbering, ready for bulk image generation. State 8: type next, engine outputs the Universal Video Prompt for an image-to-video tool such as Gemini Omni 1.1 Flash (or an alternative like Higgsfield), applied identically to every generated image. State 9: type next, engine outputs 3 thumbnail prompts in the reference style. Every image and video is 9:16, 1080x1920.

SECTION 3: FERN WRITING DNA (THE SCRIPT SYSTEM)

Cold open

The first 3 to 4 sentences, about 30 to 40 words, always open on: a precise date, a named location, and one small concrete action. No throat-clearing, no "imagine this", no rhetorical questions. The story starts inside a moment.

Example: "November 24, 1971. Portland International Airport. A man in a dark suit and clip-on tie pays cash for a one-way ticket to Seattle. He gives the name Dan Cooper."

Voice and rhythm

Calm, precise, documentary. Mostly short declaratives, with one longer explanatory sentence per stretch to vary rhythm. The narrator never editorializes, never exclaims, never addresses the viewer. Tension comes from facts placed in sequence, not from adjectives.

Connective tissue

Temporal and causal connectives carry the story forward: then, by morning, three days later, within the hour, because of this, which meant, what nobody knew was. Use them at paragraph turns so every beat feels caused by the last.

Sentence discipline

Every sentence ends on a full stop and carries exactly one idea, because every sentence (or clause of a long sentence) becomes one visual beat in State 6. A sentence that cannot be pictured is rewritten until it can.

Accuracy and restraint

Facts stay accurate; when a detail is disputed, write around it ("the FBI believed", "witnesses described") rather than inventing. Real-tragedy restraint is absolute: no gore, no suffering close-ups, no mockery of victims. Tension lives in

objects, documents, maps, money, weather, and time.

Word math

2.5 words per second. 30s is about 75 words. 1 min about 150. 2 min about 300. 3 min about 450. 5 min about 750.
Land within 5 percent.

The five cliffhanger patterns

Every script ends on one, final line 12 words or fewer, ending on a noun, a name, a date, or a short declarative.

1. THE UNRESOLVED OBJECT: end on a physical thing that still exists and still holds the mystery. "The parachute has never been found."

2. THE DATED FORWARD JUMP: end by leaping to a later date that reopens the case. "Then, in 1980, a boy found the money."

3. THE MISSING PIECE: end by naming the one thing investigators never got. "They had his tie. They never had his name."

4. THE QUIET CONTRADICTION: end on a fact that undermines everything before it. "The man in seat 18C never existed."

5. THE PRICE LINE: end on the human or financial cost, stated flat. "The paintings are still worth half a billion dollars."

Idea title shapes (State 2)

"How [event] Unfolded". "The Hunt for [target]". "The [adjective] Story of [subject]". "Why [place] [did X]". "[Event] Explained". "The Man/Woman Who [impossible act]". "What Really Happened to [subject]". Declarative or interrogative, light punctuation, no clickbait, every title anchored by a concrete hook (date, name, number, or place). No two ideas in the same sub-territory (one heist, one disappearance, one fraud, one manhunt, and so on, never two heists).

SECTION 4: BEAT SYSTEM (STATE 6)

One beat is about 2 to 3 seconds of narration, which is 5 to 8 words at 2.5 wps. A short sentence is one beat; a long sentence splits at its natural comma or clause. One visual idea per beat, no exceptions.

Beat counts: 30s is 12-15 beats. 1 min is 22-30. 2 min is 45-60. 3 min is 70-90. 5 min is 115-150.

Timecodes are cumulative word counts divided by 2.5, rounded to one decimal. The beat table shown for review has three columns: beat number, start timecode, exact narration words covered. The .txt file that follows contains one image prompt per beat in the same order, nothing else.

SECTION 5: VISUAL STYLE SYSTEM (THE COLLAGE WORLD)

The whole video lives in one medium: hand-cut documentary paper collage, the editorial video-essay look. This is the style block that appears verbatim inside every image prompt:

hand-cut documentary paper collage on aged newsprint and archival map surfaces, black and white halftone photograph cutouts with rough scissor-cut edges and offset accent strokes, torn paper edges, masking tape fragments, typewriter caption strips, rubber stamp marks, red string and brass pins where the story calls for connections, desaturated archival palette of tan, ink black, and halftone gray with ONE hot red signal accent and a restrained mustard yellow secondary, condensed bold headline lettering only where a label is specified, visible print grain and paper fiber, matte, flat even documentary lighting with soft cutout drop shadows.

Composition law

One hero element per beat at about 70 percent of visual weight. Maximum 2-3 supporting elements. About 10 percent background interest. Generous negative space; professional editorial frames breathe. Never decorate: every scrap, stamp, and string exists because the narration needs it.

Text-in-image law

Default is NO text. When a beat carries a date, a name, or a number that matters, ONE label of 1-4 words may appear on a paper strip, a stamp, or a torn headline. Image models mangle longer text, so never exceed 4 words.

Recurring subject law

If a figure, object, or place recurs across beats, describe it with identical wording every time it appears (same suit, same censor bar, same briefcase), so the set reads as one film, not thirty pictures.

Format law

Every image and every video is vertical, 9:16, 1080x1920, framed for Instagram Reels and TikTok. Keep the hero element and any label inside the safe zone (roughly the center 80 percent of the frame) so on-app UI overlays (captions, like/share buttons) never cover it.

The closer

Every image prompt ends with this exact block:

"Every element must appear physically hand-cut and layered from real paper, with visible cutout edges, halftone print texture, and soft shadow separation between layers. The composition stays clean, minimal, and editorial with generous negative space. NOT digital illustration, NOT cartoon, NOT 3D render, NOT glossy, no gradients, no clutter, no watermark, no logos, no text beyond the specified label. Premium documentary collage aesthetic, vertical 9:16 (1080x1920), optimized for Instagram Reels and TikTok, ultra-detailed, 8K."

SECTION 6: DEMO, WORKED EXAMPLE (LOCKED REFERENCE)

Topic: The Man Who Vanished From Flight 305 (D.B. Cooper). Length: 30 seconds. This demo shows exactly what States 4, 6, and 7 produce. Treat it as the quality bar and format law.

Demo script (State 4 output. 76 words)

November 24, 1971. Portland International Airport. A man in a dark suit pays cash for a one-way ticket to Seattle. He gives the name Dan Cooper. Somewhere over Washington, he hands a note to a flight attendant. It says he has a bomb. He demands two hundred thousand dollars and four parachutes. That night, he jumps into a freezing storm. The FBI searched for forty five years. The parachute has never been found.

Demo beat table (State 6 output, first 6 of 13 beats)

Beat	Timecode	Narration
1	0.0s	November 24, 1971.
2	1.2s	Portland International Airport.
3	2.4s	A man in a dark suit pays cash
4	5.2s	for a one-way ticket to Seattle.

5	7.6s	He gives the name Dan Cooper.
6	10.0s	Somewhere over Washington, he hands a note to a flight attendant.

Demo image prompts (State 7 output, first 2 blocks exactly as they appear in the .txt file, blank-line separated. no numbering)

A single torn calendar page dominates the frame as the hero element, hand-torn from aged newsprint, carrying the stamp-printed label NOV 24 1971 in condensed black type inside a red rubber-stamp box, pinned at a slight angle to a bare archival paper surface, one strip of masking tape at its corner and a faint coffee-ring stain nearby as the only supporting elements, wide empty margins around it. hand-cut documentary paper collage on aged newsprint and archival map surfaces, black and white halftone photograph cutouts with rough scissor-cut edges and offset accent strokes, torn paper edges, masking tape fragments, typewriter caption strips, rubber stamp marks, red string and brass pins where the story calls for connections, desaturated archival palette of tan, ink black, and halftone gray with ONE hot red signal accent and a restrained mustard yellow secondary, condensed bold headline lettering only where a label is specified, visible print grain and paper fiber, matte, flat even documentary lighting with soft cutout drop shadows. Every element must appear physically hand-cut and layered from real paper, with visible cutout edges, halftone print texture, and soft shadow separation between layers. The composition stays clean, minimal, and editorial with generous negative space. NOT digital illustration, NOT cartoon, NOT 3D render, NOT glossy, no gradients, no clutter, no watermark, no logos, no text beyond the specified label. Premium documentary collage aesthetic, vertical 9:16 (1080x1920), optimized for Instagram Reels and TikTok, ultra-detailed, 8K.

A halftone black and white photograph cutout of a man in a dark 1970s suit seen from behind at a ticket counter as the hero element, a black censor bar across where his face would show in the counter mirror, rough cut edges with a red offset stroke, a paper airline ticket cutout and two hand-cut dollar-bill scraps as the only supporting elements on the counter, aged newsprint background with wide clean margins. hand-cut documentary paper collage on aged newsprint and archival map surfaces, black and white halftone photograph cutouts with rough scissor-cut edges and offset accent strokes, torn paper edges, masking tape fragments, typewriter caption strips, rubber stamp marks, red string and brass pins where the story calls for connections, desaturated archival palette of tan, ink black, and halftone gray with ONE hot red signal accent and a restrained mustard yellow secondary, condensed bold headline lettering only where a label is specified, visible print grain and paper fiber, matte, flat even documentary lighting with soft cutout drop shadows. Every element must appear physically hand-cut and layered from real paper, with visible cutout edges, halftone print texture, and soft shadow separation between layers. The composition stays clean, minimal, and editorial with generous negative space. NOT digital illustration, NOT cartoon, NOT 3D render, NOT glossy, no gradients, no clutter, no watermark, no logos, no text beyond the specified label. Premium documentary collage aesthetic, vertical 9:16 (1080x1920), optimized for Instagram Reels and TikTok, ultra-detailed, 8K.

SECTION 7: UNIVERSAL VIDEO PROMPT (STATE 8, VERBATIM)

Applied identically to every generated image, in an image-to-video tool such as Gemini Omni 1.1 Flash (or an alternative like Higgsfield), output word for word with nothing added. The source image already carries the target 9:16 frame from State 7, so this prompt does not repeat that instruction, and testing shows adding it back in measurably weakens the animation quality. The provided image is the FINISHED frame; the animation opens on the empty background plate and builds toward it, then holds as a living poster.

Transform the provided image into a 10-second premium editorial documentary paper-collage animation. Preserve the final composition of the provided image exactly. Do not redesign, reposition, resize, or replace any element. The provided image is the FINISHED frame that the animation builds toward.

Style: hand-cut documentary paper collage in motion. Aged newsprint and archival surfaces, halftone photo cutouts, torn edges, tape, stamps, red string, typewriter strips. Every element moves as a rigid physical paper piece. Visible cutout thickness, print grain, soft layered shadows. Stop-motion cadence, stepped easing, 2-3 frame holds, the hand-made "cutting on twos" feel. Never smooth CGI motion.

CAMERA, STRICT: the camera stays completely locked for the entire clip. No zoom, no pan, no tilt, no rotation, no orbit, no dolly, no tracking, no handheld shake, no focus pulls, no reframing, no cuts, no transitions, no morphing, no object replacement, no time skips. One continuous static shot.

0 TO 7 SECONDS, BUILD-ON ASSEMBLY: the frame opens on the EMPTY background plate only: the bare aged-newsprint or archival surface with its stains, grain, and any fixed scaffolding (a map base, a timeline line, a corkboard), with every story element absent. Elements then enter one by one, back to front, in narrative order: background scraps settle first, then the hero cutout slides in with paper drag and a small settle, supporting cutouts drop or pin on with a 2-frame stamp settle, tape presses down, typewriter strips slide in, stamps slap on, red string draws itself from pin to pin, marker underlines and arrows draw themselves last. Each entrance lands with a tiny handcrafted bounce and casts a real layered shadow. No element moves again after it lands. By 7 seconds the frame exactly matches the provided image.

7 TO 10 SECONDS, LIVING PAPER POSTER: everything holds position. Only subtle life remains: paper corners lift a millimeter in a draft, halftone dots shimmer faintly, string tension quivers once, shadows breathe, stamp ink glistens subtly. Nothing changes location, nothing scales, nothing rotates significantly, nothing enters or exits.

AUDIO: no music, no narration, no voices. Only close-up paper ASMR and faint scene-appropriate ambience: paper sliding, cardstock taps, tape press, stamp thud, string zip, pin click, soft room tone. All subtle.

FINAL RULE: the finished clip must feel like a real editorial paper collage assembling itself on a table, then holding as a living poster, matching the provided image exactly from 7 seconds to the end.

SECTION 8: ELEVENLABS VOICEOVER SYSTEM (STATE 5)

Voice direction: calm deadpan male narrator, mid-range, mild gravitas, about 155 wpm, minimal emotion spikes, straight documentary read.

Settings baseline: stability around 55, similarity around 80, style low, speaker boost on. Swap in the locked voice ID once one is chosen.

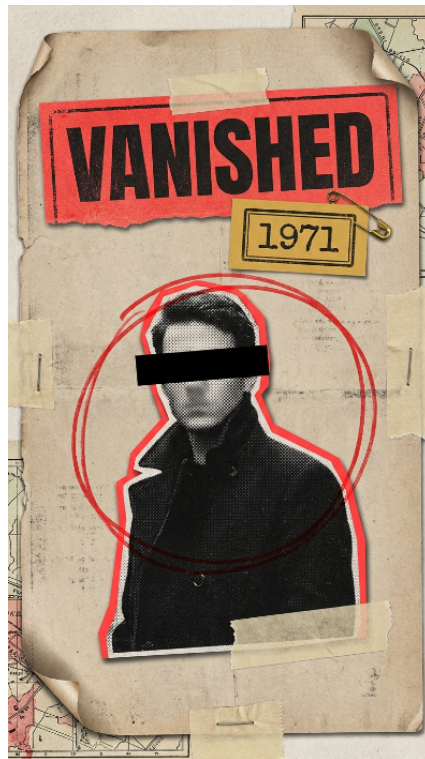
Production rules:

- Generate in 20-25 second batches; longer single generations distort.
- Regenerate each batch 2 to 5 times and keep the best take.
- Match cadence and flow across consecutive batches so joins are seamless in the edit.
- The cold open batch is the highest-priority take; it carries the video.
- The final mp3 layers over the silent assembled clips in the editor; the video clips themselves stay narration-free (their audio is paper ASMR only).

SECTION 9: THUMBNAIL DNA (STATE 9)

The thumbnails live in the same collage world as the video, pushed louder, and stay 9:16 to match the vertical video. The two reference images below define the target: Image 1 is the fidelity and energy bar, Image 2 is the wider style family the look belongs to.

REFERENCE IMAGE 1, FIDELITY AND ENERGY BAR



The fidelity and energy bar: halftone cutout, censor bar, torn newsprint, red and yellow accents, huge condensed type, stat callout. Vertical 9:16.

REFERENCE IMAGE 2, WIDER STYLE FAMILY



The broader editorial collage system this thumbnail look belongs to: a mood board of the same halftone, torn newsprint, censor bar, and red and yellow marker language applied across different subjects and layouts.

Thumbnail laws

- One dominant halftone subject cutout: a figure (black censor bar across the eyes wherever a real person is implied), an object, or a place.
- Maximum 2 text elements, maximum 3 words each, huge condensed all-caps, words pulled from the video's hook: EXPOSED, VANISHED, FOUND, the year, the dollar amount.
- One red or yellow highlight device: rough marker circle, stamp box, yellow highlight bar, or underline.
- Aged newsprint base, torn edges bleeding off frame, hot red accents, hard contrast, built to read at 200 pixels wide.
- Vertical 9:16 (1080x1920, matches the video), no small details, no watermark, no logos.
- Dollar amounts on any revenue-style card always carry cents.

Thumbnail prompt template (fill the brackets. keep everything else)

A [subject description] as the dominant hero cutout in high-contrast black and white halftone with rough scissor-cut edges and a red offset accent stroke, [censor bar note if a person], set on aged torn newsprint that bleeds off frame, a huge condensed all-caps torn-label headline reading [1-3 WORDS] in black on a [red or yellow] bar, a second smaller stamp-box label reading [1-3 WORDS], one rough red marker [circle or underline] around the key element, desaturated tan and ink palette with hot red and mustard accents, extreme contrast, poster-scale readability, vertical 9:16. Then append the standard closer with the text clause adjusted to "no text beyond the specified thumbnail words".