

Prompt cheatsheet: video, image and audio

Values measured from 8,835 real prompts, most used first. Fragments are drop-in clauses. Write the aspects in the order shown and put the exclusions last.

Video

Order: duration > aspect > medium > shotSize > angle > movement > lens > dof > lighting > grade > motion > mood > audioBed > pacing > negatives		
ASPECT	VALUES	FRAGMENT
Duration	6 seconds, 8 seconds, 10 seconds, 15 seconds, 30 seconds, 60 seconds	15-second cinematic clip.
Aspect ratio	16:9, 9:16, 1:1	Framed 16:9 for landscape playback.
Shot size	close-up, wide shot, medium shot, extreme close-up, establishing shot, two-shot, macro shot, over-the-shoulder, closeup	Extreme close-up on the texture, filling the frame.
Camera angle	overhead, low-angle, eye level, three-quarter, top-down, profile, high-angle	Low-angle looking up, subject towering over the lens.
Camera movement	push-in, tracking shot, handheld, orbit, locked-off, slow zoom, truck, slow dolly, parallax, pull-back, slow pan, whip pan, crane up, dolly in	Slow push-in toward the subject, ending tight.
Lens and focal length	35mm, macro lens, anamorphic, 100mm, 50mm, 85mm, 28mm, wide-angle lens, 32mm, 50 mm, fisheye, telephoto, 40mm, 35 mm	Shot on an 85mm lens, compressed and flattering.
Depth of field	shallow depth of field, bokeh, rack focus, f/1.4, f/1.8, f/2.0, soft focus, f/2.8	Shallow depth of field, background dissolving into bokeh.
Lighting	soft lighting, rim lighting, rim light, silhouette, volumetric lighting, golden hour, volumetric light, soft light, side lighting, chiaroscuro, key light, low-key, practical lighting, backlit	Soft lighting from a large source, shadows barely there.
Colour and grade	vibrant, pastel, film grain, black-and-white, natural color, high contrast, desaturated, crushed blacks, natural colour, low contrast	Vibrant saturated grade, colours pushed but not clipping.
Medium and style	cinematic, photorealistic, anime, documentary, 3d render, unreal engine, stop-motion, 35mm film, watercolor, vector, photoreal, claymation, line art, octane	Cinematic, shot on 35mm film.
Subject motion	reveal, float, drift, slow motion, glides, rotates, time-lapse, ripples, unfolds, hovers, cascade	The lid glides open on hidden runners, slow and even.
Mood	playful, energetic, whimsical, calm, dramatic, minimal, elegant, uplifting, moody, ominous, ethereal, luxurious, serene, nostalgic	Calm and unhurried throughout.
Audio bed	percussion, bells, ukulele, xylophone, piano, drums, bass, marimba, strings, synth, guitar, choir, handclaps, pads	Minimal deep electronic ambience with soft mechanical clicks.
Pacing and cuts		One continuous take, no cuts.
Resolution / quality tier	480p, 720p, 1080p, 4k, basic, high, ultra, pro, std, standard, bitrate_mode standard high	--resolution 4k
First/last frame anchoring	--start-image, --end-image, start frame, end frame, start = end, one-shot, loop, end-frame reveal, generic image/reference role , not a start-frame role	`--start-image` anchors the first frame. Prompt describes motion.
Ad format / mode (branded video)	ugc, ugc_how_to, ugc_unboxing, product_showcase, product_review, tv_spot, wild_card, ugc_virtual_try_on, virtual_try_on, hook, setting, ad reference	Looks like a real person filmed on phone
Edit instruction (video editing workflows)	make the jacket red, sketch, timestamp	--prompt "make the jacket red"

Rule out. Text and branding: text, logos, subtitles, watermark, watermarks, captions, on-screen text, text on screen; Anatomy and faces: extra limbs, distorted faces, extra characters, hands, extra fingers, extra animals; Content and safety: violence, scary imagery, scary elements, weapons; Audio: music, narration; Other: dialogue, copyrighted characters, talking, cuts, horror, humans, clothes, flickering

Image

Order: aspect > medium > shotSize > angle > composition > lens > dof > lighting > grade > material > background > mood > text > negatives

ASPECT	VALUES	FRAGMENT
Aspect ratio	16:9, 9:16, 1:1, 4:5, 21:9, 3:2, 4:3	Square 1:1.
Composition and framing	close-up, wide shot, medium shot, extreme close-up, two-shot, closeup, over-the-shoulder, establishing shot, macro shot, full shot	Centred on a single vertical axis, symmetrical.
Shot size	close-up, wide shot, medium shot, extreme close-up, two-shot, closeup, over-the-shoulder, establishing shot, macro shot, full shot	Extreme close-up, texture filling the frame.
Camera angle	three-quarter, profile, eye-level, low angle, overhead, high angle, top-down, bird's-eye, dutch angle	Three-quarter view, turned slightly off axis.
Lens and focal length	35mm, 85mm, 50mm, 110 mm, 16mm, anamorphic, 24mm, 70mm, 48mm, fisheye, 600mm, 450mm, 40mm, 297 mm	85mm portrait lens, natural compression.
Depth of field	shallow depth of field, bokeh, f/1.4, f/7.1, soft focus, f/2.8, deep focus, f/2.0, f/4, f/2, f/1.8, f/5.6	Shallow depth of field at f/1.8, background soft.
Lighting	silhouette, rim light, soft light, soft lighting, backlit, key light, golden hour, high-key, rim lighting, backlight, softbox, low-key, volumetric light, bounce light	Large softbox front left, soft shadows falling right.
Colour and palette	vibrant, pastel, film grain, black-and-white, desaturated, high contrast, natural color, natural colour, monochrome, low contrast, sepia	Vibrant, fully saturated, clean whites.
Medium and style	cinematic, photorealistic, documentary, anime, 3d render, line art, watercolor, vector, 16mm, 35mm film, photoreal, kodak vision3, isometric, flat illustration	Photorealistic editorial photograph.
Materials and finish	glossy, matte, ceramic, concrete, leather, velvet, satin, brass, linen, frosted, marble, chrome, walnut, oak	Brushed metal with a satin finish, no mirror reflections.
Background		Seamless mid-grey studio backdrop, no horizon line.
Mood	elegant, playful, minimal, dramatic, calm, energetic, whimsical, moody, luxurious, nostalgic, uplifting, ethereal, cosy, gritty	Elegant and restrained.
Text handling		No text, no lettering, no numbers anywhere in the image.
Subject + setting + style (the opening clause)	a red fox curled in a snowy pine forest, golden hour, cinematic, subject + action, concrete and physical	a red fox curled in a snowy pine forest, golden hour, cinematic
Typography named in the prompt (when text is generated)	Anton, Anton SC, Bebas Neue, Oswald, Archivo Black, Montserrat, Poppins, Roboto Condensed, Inter, Barlow Condensed, Playfair Display, Cormorant Garamond, DM Serif Display, Fraunces	The prompt must state: Exact literal copy; Display/body font family names and which text uses each
Identity lock / subject fidelity	IDENTITY LOCK, photographic identity match, same bone structure, eye shape, nose, lips, jawline, skin tone, hairline and hair texture, Do not beautify, average, or restyle the face, IMAGE REFERENCES manifest	CHARACTER N: the person from attached face reference #K , IDENTITY LOCK: reproduce this exact person with a photographic identity match , same bone structure, eye shape, nose, lips, jawline, skin tone, hairline and hair texture. Do not beautify, average, or restyle the face. Expression: <emotion phrase>.

ASPECT	VALUES	FRAGMENT
Aesthetic / seasonal preset	Clean girl, Cottagecore, Quiet luxury, Dark academia, Y2K, Christmas, Valentine's, Halloween, Black Friday, None, Clean studio, Lifestyle, Conceptual, With a model	Christmas version, quiet-luxury aesthetic
Product-photo genre / mode	product_shot, lifestyle_scene, closeup_product_with_person, moodboard_pin, hero_banner, social_carousel, ad_creative_pack, virtual_model_tryout, conceptual_product, restyle, levitating, floating, splash, frozen motion	bottle of cold-brew on a sunlit kitchen counter, IG feed
Logo / vector mark specification	lettermark_monogram, pictorial, abstract, mascot, emblem, one dominant unified silhouette, negative-space device, motif treatment, unexpected locked color-role pairing, symmetry or intentional asymmetry, internal detail limit, small-size scalability, centered isolated presentation, minimal anchor points	Flat vector design, clean lines, no shadows, no texture, no text. Clean editable vector paths, SVG-friendly, minimal anchor points.
Logo placement on a mockup surface	Target surface, Position, Scale, Orientation, Color, horizontally centered, upper third, optical center aligned to panel, clear-space margin, align to <panel edge/seam/baseline>, follow surface perspective	[PLACEMENT LOCK] Target surface: <exact object panel/face/material>. Position: <exact alignment and location, e.g. horizontally centered, upper third, optical center aligned to panel>. Scale: logo occupies <specific proportion> of the target surface while keeping <specific clear-space margin>. Orientation: align to <panel edge/seam/baseline>; follow surface perspective without changing logo proportions. Color: use the supplied <full-color/black/white> variant exactly. State why it contrasts correctly with the material/background.
Seamless / tileable behaviour	seamless tileable game texture tile, uniform pattern density, perfectly seamless edges that wrap horizontally and vertically, no border, no vignette, no single focal object, photographed-flat appearance, seamless repeating pattern	seamless tileable game texture tile of <description>, uniform pattern density,
Key colour for transparency	bright magenta #FF00FF, bright green #00FF00, bright blue #0000FF	on a solid uniform bright <KEY COLOR> background, crisp edges, no drop shadow outside the element
Reference-image role and what it controls	style donors only, Image0, Image1, Image2, official logo, approved base scene, style-reference, representative-visual, grade/energy only	Make an Animated Explainer. Take only the visual render style and color grading of the input image(s); mix the styles if there is more than one image. Never use the characters, inscriptions, etc. from the input image(s) unless the instructions below ask you to. Use only the render style, and follow the user's instructions below:
Design-board direction vocabulary (web/layout imagery)	Pristine Light (paper/cream/off-white, dark ink), Deep Dark (charcoal/graphite), Bold Studio Solid (oxblood, royal blue, forest, vermilion, emerald fields), Quiet Premium Neutral (bone, sand, taupe, stone, smoke), technical grid/dot field, solid with soft ambient depth, full-bleed cinematic imagery, tactile paper/material texture, cinematic centered minimalist, asymmetric split, floating polaroid scatter, inline typography behemoth, editorial offset, massive image-first with restrained text	"website design mockup, desktop landing page section, [SECTION ROLE], [theme paradigm + exact palette words], [typography character] typography, [hero architecture / composition anchor], [background model], [narrative spine motif], professional layout, clear hierarchy and spacing, award-winning web design" , plus: "no watermark, no browser chrome".
Reference photos for identity training	front, 3/4 left, 3/4 right, slight up/down, indoor, outdoor, soft, harsh, neutral, smiling, talking, head shot, head-and-shoulders, full body	Higher variety = better identity capture.
Image-to-3D input framing	t-pose, a-pose, Full body in frame, nothing cut off, Clean/plain background, Limbs visually separated from the torso, One character, no props overlapping the silhouette, lowpoly, three-quarter isometric view	Full body in frame, nothing cut off.

Rule out. Text and branding: text, logos, captions, watermark, watermarks, logo, labels; Anatomy and faces: hands, arms, extra people, extra characters; Frame artefacts: frame, border, mockup; Content and safety: scary elements, weapons prominently displayed; Style and clutter: people, props; Other: overlays, slogan, letters, sunglasses, dramatic pose, helmet, police department names, cinematic color grading

Audio

Order: duration > role > mood > key > tempo > instrumentation > arrangement > vocals > ending > negatives

ASPECT	VALUES	FRAGMENT
Duration	15 seconds, 30 seconds, 60 seconds, 90 seconds, 120 seconds	90-second bed.
Role		Instrumental music bed to sit under narration.
Tempo	70 BPM, 90 BPM, 100 BPM, 116 BPM, 128 BPM	Buoyant mid-tempo around 116 BPM.
Key and tonality	major key, minor key, modal, pentatonic	Bright major key.
Instrumentation	percussion, ukulele, piano, bass, handclaps, marimba, bells, strings, guitar, xylophone, drums, harp, synthesizer, pads	Marimba, ukulele, xylophone, light percussion and soft handclaps.
Mood	warm, playful, cheerful, buoyant, wholesome, ambient, uplifting, driving, triumphant, minimal, dreamy, suspenseful	Cheerful, wholesome and playful.
Arrangement and dynamics	airy, midrange, headroom, no abrupt hits, narration-friendly, smooth transitions, gentle dynamics	Airy and narration-friendly, gentle dynamics, no abrupt hits.
Vocals		No vocals, no lyrics, no spoken words.
Ending		End on a clean major-key resolution, no fade.
Voice and spoken line	voice_type preset element, voice_id, speech_rate, one voice per speaking entity, exact line, performance direction	The gate is open. Move!
Narration writing (spoken script)	20-24 words, about 8-9 seconds, under roughly 9.5 seconds, plain spoken text only, no timecodes, emotion cues, parentheticals, or stage directions, Spell numbers out, hook through payoff, ~150 words per minute	Write one plain line per clip, sized for about 8-9 seconds and normally 20-24 words. Keep every take under roughly 9.5 seconds.

Rule out. Text and branding: spoken words, words, other words, harsh textures, ad-libs that obscure the words; Content and safety: scary sounds; Audio: vocals, lyrics, sound effects, speech, music, spoken narration, voices, melody; Other: medical sounds, abrupt peaks, dark, dramatic tension, heavy drums, long instrumental passages, sung parts, dialogue