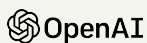


Pencil Prompting Playbook

Image & video prompting principles for AI-wielding creatives

OLIVER

Jellyfish



stability.ai

getty

Google

∞ Meta

🎵 TikTok

amazon

LinkedIn

The art of prompting

How to make pictures with machines
that keep getting smarter.



Before we begin

Three things to remember:



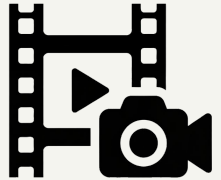
This Guide Is Wrong.

Not completely - but partially, inevitably. Every principle reflects the best thinking of a specific moment. Models change. What worked brilliantly last quarter may underperform today. Read accordingly.



It Is A Living Document.

Update it. Argue with it. Cross things out. The worst thing we could do is laminate it and call it done. Prompt engineering has no final form yet.

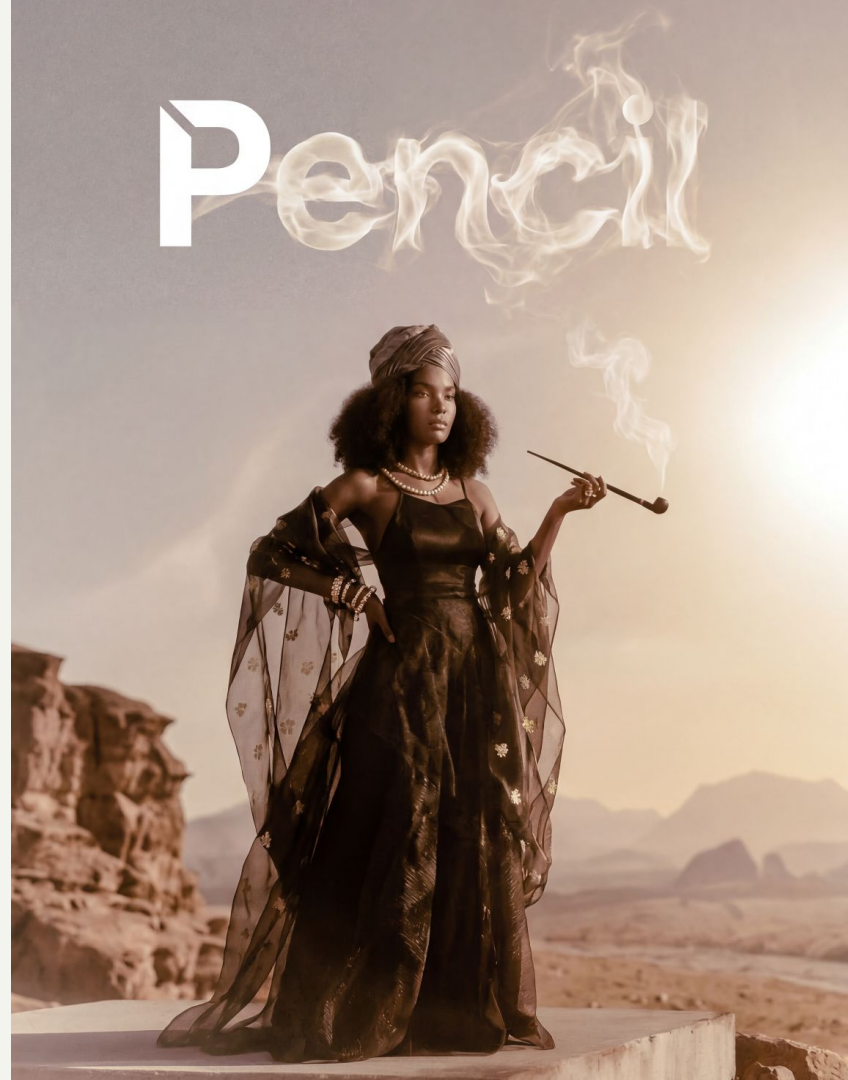


This Guide Is Image & Video Specific

Prompting for text generation is a different craft - it gets its own playbook. Everything here is calibrated for the visual: still images, motion, and the AI models that make them.

Contents : The principles of prompting

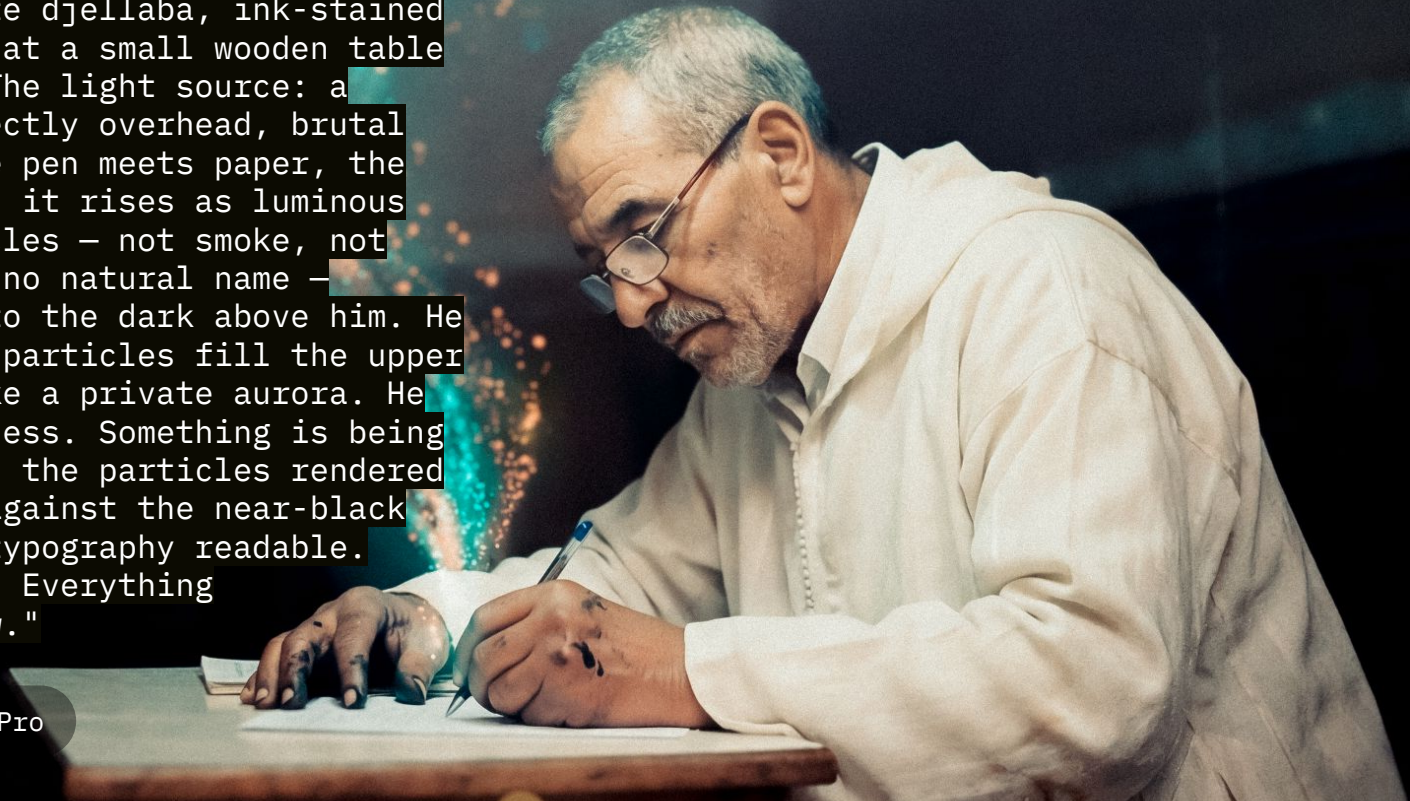
1. Learn to Learn
2. The Prompt Is the New Brief
3. The Model Has Character
4. Anatomy of an Image Prompt
5. Break the Bias
6. Examples Beat Explanations
7. Where You Put It Matters
8. Say What You Don't Want
9. Iterate Iterate Iterate!
10. Experiment and Learn From Failures
11. Physics-first Prompting
12. A Multi-Disciplinary Lexicon
13. Prompting for Video
14. Video with Native Audio
15. Let AI Write Your Prompts
16. Risky Prompts to Avoid



Prompt:

"35mm photography, pushed two stops, extreme grain. An older North African man – 70s, reading glasses, white djellaba, ink-stained right hand – writing at a small wooden table in a riad interior. The light source: a single bare bulb directly overhead, brutal and theatrical. Where pen meets paper, the ink doesn't stay ink: it rises as luminous teal and coral particles – not smoke, not fire, something with no natural name – spiralling upward into the dark above him. He doesn't look up. The particles fill the upper half of the frame like a private aurora. He is writing into darkness. Something is being made. Shot wide open, the particles rendered as pure light bloom against the near-black ceiling. No text or typography readable. The bulb burns white. Everything else earns its shadow."

Model: Nano Banana Pro



- * There is no universal formula.
These are general principles, not a rulebook.

Learn to learn

Tools change.

Models are replaced.

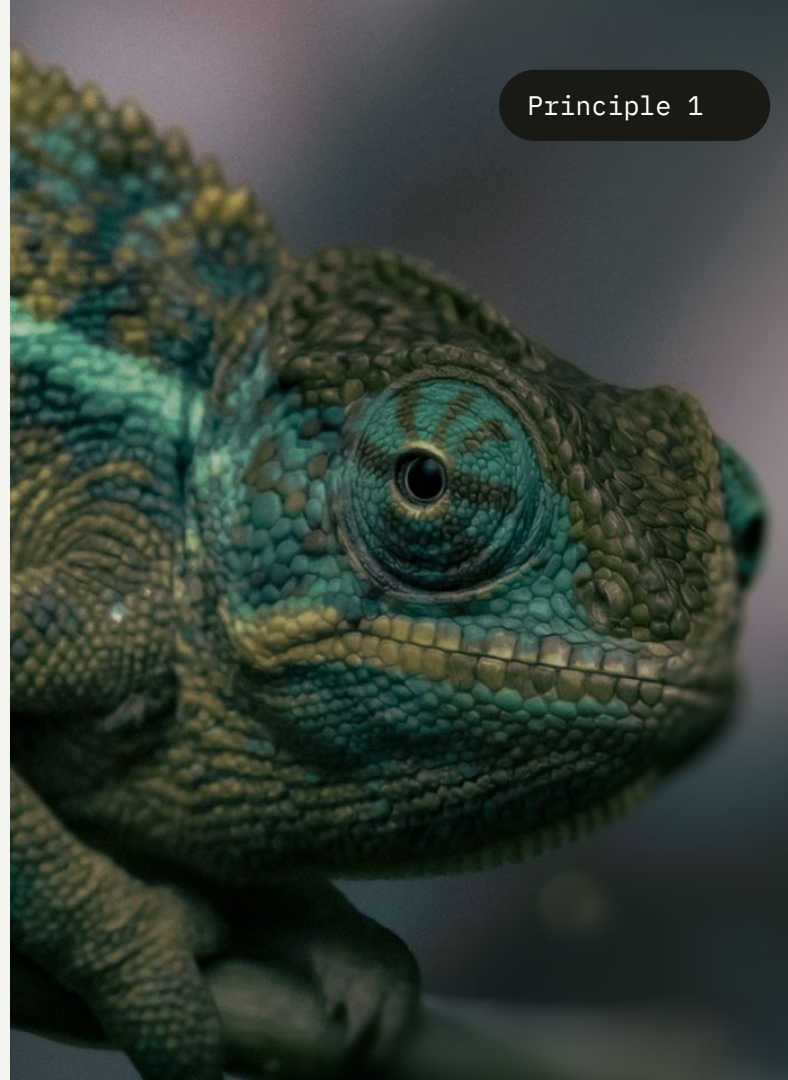
Interfaces are redesigned.

The ability to prompt any single tool is the least important skill you can develop.

Staying resilient in the AI era requires flexibility; the capacity to hold your current mental model loosely, let go when the ground shifts, and rebuild quickly.

The art of prompting is less like learning a single language, and more like *learning how to learn* languages.

Principle 1



Prompt:

"Macro wildlife photography, absurdly close. A Cape dwarf chameleon, *Bradypodion pumilum*, vivid green with yellow lateral stripe, perched on a thin branch. The chameleon's turret eye is swivelled directly into the lens, the other eye pointed backward - that split-gaze thing they do. Depth of field is a sliver - the eye, the tail are sharp, the branch and background dissolve into lush tropical green bokeh. Every granular scale on the chameleon's skin rendered in forensic detail. Soft key light from above-left, a single catchlight in the forward-facing eye. A serendipitous, award winning nature photograph with a warm green-bronze overall colour grade. Shot using 200mm macro, f/2.8."

Model:

Nano Banana 2

The Prompt Is the New Brief

Principle 2

A good brief/prompt leaves no room for misinterpretation.

Vague prompts produce vague outputs. **Precision** is key, and long prompts are risky, unless every element is intentional. Restraint - knowing when to strip back your prompts - is part of the craft. "**Brevity** is the soul of wit," and it can also be the highest form of prompting.

Vague prompt:

Precise prompt:

"A woman in ancient jewelry standing dramatically in a mythological scene."

"A woman in her 30s wearing dark draped fabric and ornate gold jewelry. Warm golden-hour backlight with a dusky sky. Wide shot, off-centre composition, cinematic colour grade. Ancient mythology meets fashion ad campaign feel."

NO



YES



Prompt:

"The subject's natural curl or wave pattern is the compositional hero of the image; the hair has volume, definition, and architectural presence. The model's gender, ethnicity, and curl type are unspecified – generate freely. Minimal wardrobe in black or ecru only: a single garment; coat, suit, bodysuit etc.(chosen so as not to compete with the hair.)

Pose is organic and sculptural: seated, crouching, leaning, or turning – body angled, face toward camera, expression cool and direct. Background: warm off-white seamless studio paper. Lighting: large diffused softbox, high-key, even exposure. Skin and hair both luminous. No accessories, no jewellery, no props. Do not render the hair as straight, blown-out, or smooth – curl and coil definition must be sharp and distinct. Shot on Hasselblad H6D, 85mm. Fashion magazine high end editorial aesthetic."



Model:

Nano Banana Pro

The Model Has Character

Every model has strengths, stylistic 'tastes', biases, blind spots, and its own way of interpreting prompts.

Nano Banana 2 & Pro - The workhorse: Fast, photorealistic, great text rendering. Both these models sport studio-grade fidelity.

ChatGPT Image 1.5 - The art director: Best instruction following. Excels at iterative edits. Warm colour bias.

Bria 3.2 - The compliance model: Licensed data, full IP indemnification. Deterministic parameter control.

Adobe Firefly - The brand-safe option: Licensed data, commercial guardrails. Compliance over raw quality.

Google Imagen 4 - The aesthetic purist: High photorealism. Single-shot - no iterative reworking, so prompt precision is key.

Principle 3



Know your model

* Model recommendations date quickly. Verify before deploying on critical campaigns.

 Google

Image models (in order of preference):

Product & commercial:

Nano Banana Pro, Nano Banana 2, ChatGPT Image 1.5

 OpenAI

Google Imagen 4,

Object swapping:

Nano Banana Pro, Nano Banana 2, ChatGPT Image 1.5

 Adobe

Brand assets:

Nano Banana Pro, Nano Banana 2, ChatGPT Image 1.5, Bria 3.2

 Bria

Text in image:

Nano Banana 2, Nano Banana Pro, ChatGPT Image 1.5, Adobe Firefly



Know your model

* Model recommendations date quickly. Verify before deploying on critical campaigns.

 Google

 OpenAI

 Adobe

 Bria

 Runway

 Kling

Video models

Cinematic motion with native audio:

Veo 3.1, Veo 3, Sora 2 Pro, Sora 2

Fast iteration:

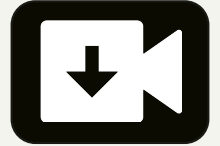
Veo 3.1, Veo 3, Sora 2.1, Runway Gen-4 Turbo,

Commercial-ready & brand-safe:

Veo 3.1, Veo 3, Sora 2.1, Adobe Firefly Video

Start + End frame control:

Veo 3.1, Runway Gen-3 Alpha Turbo, Kling AI 2.5 Turbo, Kling AI 2.1, Adobe Firefly Video



Prompt:

"Award-winning photorealistic coffee advertisement on a black background. A large coffee bean floating centre-frame, hyper-realistic in texture, but it is unmistakably also an ancient tribal mask. The bean's natural crease divides the Mask into two ceremonial halves. The LEFT half is the MOON; carved with closed or half-lidded sleeping eyes, serene expression, cool silver and blue-grey tones in the surface patina, subtle crescent moon symbols etched into the forehead. The RIGHT half is the SUN – carved with a wide open eye, dynamic expression, warm golden tones in the patina, radiating sun-ray lines carved outward from the eye, the texture of the same ancient material but bathed in warm fire-light. The two halves share one nose along the central crease of the bean. The bean retains its unmistakable coffee bean shape and roasted surface texture throughout. Dramatic split lighting reinforces the duality: cool blue-silver light from the left, warm amber-gold light from the right. The bean sits on a dark reflective surface with just enough sheen to ground it. Shot on Hasselblad H6D-400c, 120mm macro, f/5.6, studio lighting, museum-quality still life. No text."



Model: Nano Banana Pro

Anatomy of an Image Prompt

Two levels of prompt construction:
basic & advanced.

Basic prompt anatomy

Three elements, in order. This may be enough for simple prompts. It's also good for when you want to leave the AI model some leeway for creative freedom.



Subject



Composition

Style/Vibe

Principle 4



Basic example prompt:

"A man in a linen shirt pouring coffee on a sunlit balcony. Shot from across the table, slightly out of focus foreground. Calm, golden, unhurried."

Model:

GPT Image 1.5



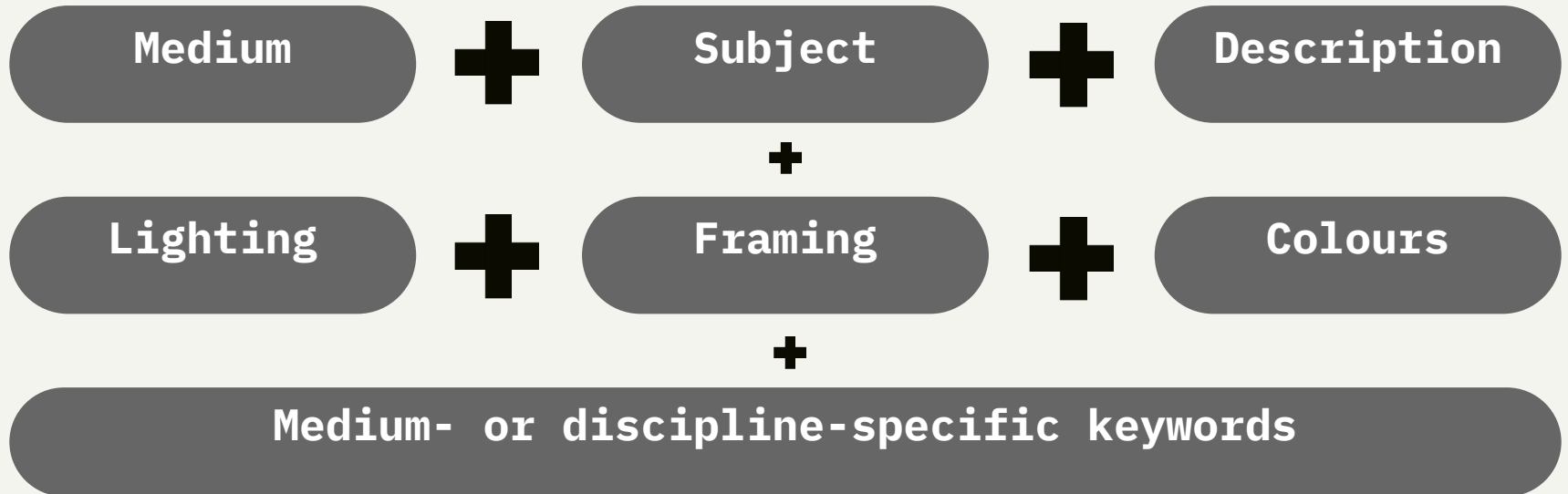


Anatomy of an Image Prompt

Two levels of prompt construction: basic & advanced.

Advanced prompt anatomy

For when you need precise, fine-grained control, and you have a very specific vision in mind, try this more detailed, robust framework. This is especially useful for hero images, campaign work, or anything technically complex.



Advanced example prompt:

"Studio photography. A pair of white leather sneakers on wet concrete. Fresh out of the box, laces untied, one shoe tilted against the other. Single hard overhead light, deep shadow underneath, no fill. Tight crop, 85mm macro, f/4, stitching in focus. Cool grey and white only, no warm tones. Shot on Phase One, maximum sharpness, no post-processing grain."



Model: GPT Image 1.5

Break the Bias

Every AI model carries bias.

This is a result of human biases, representation, and training data distribution. And it's also circular: biased outputs become the next round of training data, and we risk reinscribing and normalising the very biases we want to unlearn.

Break the cycle and contribute to fairer representation across media by specifying diversity in your prompts. To bake this in to your workflow, activate Pencil's **Bias Breaker**. This passes your input through another AI tuned to counter common AI biases before sending the new prompt to the image or video model downstream.

Examples of AI Bias:

Prompt for "a conductor" and most of the time you'll get an older white man. That's an example of AI bias in action. Counter this by prompting for **specific ethnicities, orientations, ages, or characters who are differently abled**.

You can also push past models' unreasonable default beauty standards for a more natural, realistic look, with keywords like "visible skin pores", "wrinkles", "stretch marks", or "natural imperfections".



Prompt:

"Low angle, 16mm wide, slight fisheye distortion. A black woman symphony orchestra conductor in her 50s - natural silver hair, full formal concert dress - mid-downbeat, arm fully extended, baton blurred with motion. Shot from below the podium, looking up: she is the central 'hero' in the composition - her conductor baton's arc visible as a motion trail of pure white light against near-black. Behind her: the orchestra is pure abstraction - a warm sea of out-of-focus amber and gold. The ceiling above her: vaulted, dark, cathedral scale. 16mm, f/2.8, 1/30 second; long enough for the baton to blur, short enough for her face to hold. Expression: not performance. Absolute authority. The baton trail is the only line in the image. Available light only."

Model:

Nano Banana Pro



More Bias-Breaking Examples

More prompt keywords for better diversity and inclusion.

Using these sorts of words and phrases also usually results in a less idealised, more natural, realistic look to your AI-generated images.

Multigenerational

"A group featuring children, parents, and elderly family members."

Neurodivergence

"A person using noise-canceling headphones in a busy space."

Diverse abilities

"Professional athlete using a high-tech racing wheelchair."

Global heritages

"A person wearing a contemporary South Asian Lehenga in a modern office."

Gender identities

"A model with an androgynous aesthetic and non-binary fashion styling."

Examples Beat Explanations

A picture is worth a thousand words, or more, to a multi-modal image model.

Show, don't just tell.

Showing a model examples of the kind of outputs you're going for - or even the style or 'mood' that you envision - communicates more to the AI than paragraphs of description ever could.

Teach it your style or look with a few key examples. Give it a mood board in addition to, or instead of, a text brief.

Paste in samples. Attach design references. Be sure to make it clear in your prompts which attachments are image, character or compositional references, and which ones are purely intended as style inspiration references.



Where you put it matters

Models don't read linearly.

Front-load the most important elements.

Image models tend to give the most attention to the beginning and end of a prompt. The middle of the prompt may be underweighted.

Beginning

End

Middle

'Mission critical' elements of your image should be close to the top, or right at the end. Not in the middle. This is especially true with longer prompts.

Principle 7



Prompt:

"A classic wattle and daub traditional adobe hut dwelling from southern Africa sitting atop a tall cylindrical topiary column covered in dense ivy and flowering vines, floating above a perfectly manicured green lawn with pink flower beds, bright saturated cyan-blue sky, soft white cumulus clouds, warm midday light, slight glow around the house, photorealistic with soft dreamlike quality, shot from a low angle looking up. Off centre, rule of thirds, golden ratio composition. Anime-influenced dreamcore surrealism or magical realism style."

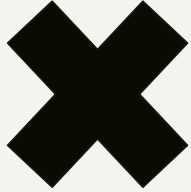


Model:

Nano Banana 2

Say what you don't want

Not only what you do want.



The default failure mode is generic. The antidote is often a precise exclusion.

Standalone Negative Prompt fields are largely a thing of the past. Today's models respond well to negatives written directly into your prompt, in plain language.

Examples:

`"...natural expressions, no forced smiles or staged poses..."`

`"...clean studio white background, no props or additional objects..."`

`"...minimal composition, no typography or text in the image..."`

Prompt:

Award-winning high-fashion photograph. Subject/composition: randomized, but make it extraordinary. Style: Prismcore - light itself is the subject, shot through a giant prism. Electric cyan, hot magenta, neon teal, violet; aggressive complementary clashes, rainbow edges, chromatic aberration everywhere. Light fractures, splits, melts across surfaces - 80s Miami Vice meets 2025 holographic. Chrome, liquid glass, oil-slick iridescent skin, molten metals. Euphoric, hypnotic, glitching into something more beautiful than reality. No flat, diffused or warm-only light. No matte surfaces. No earthy palettes. No desaturation or monochrome elements.

Model:

Nano Banana 2



Iterate Iterate Iterate!

There are many ways to ideate, iterate, refine and curate with prompts in Pencil.

The first prompt is just the beginning...

Chat with an **agent** and prompt it conversationally; map your creative process out visually using **sheets** or **workflows**, with modularised prompts; or work in a more traditional 'prompt + input image + settings = output' paradigm.

Once you've got an image close to locked, you can start prompting for surgical edits, using prompts with clear precise action verbs, like **“add”**, **“swap”**, **“sharpen”**, **“intensify”** or **“shift”**.



Prompt (with image reference):

"A vast modern high-end gallery or conference space. The image attached appears on a massive wall screen filling the entire back wall of a semi-outdoor, open plan massive sleek beautiful corporate conference venue, set in a desert oasis with a 'liminal space meets urban' aesthetic. The wide flat shiny floor subtly reflects the large display ad on the huge back wall. A few people milling about in the mid ground, they look really tiny on the ground compared to the size of the display ad, showing off the enormous scale of the space. High-end architectural art or fashion journal; a graphic design portfolio piece."



Model:

Nano Banana 2

Experiment and learn from failures

Principle 10

Skilled prompters pause and diagnose before reformulating.

Do a 'post mortem' on failed prompts.

When prompting, it can be instructive to **examine the specific failure modes** you encounter, and then **refine your prompts accordingly**, targeting the specific issues that are coming up.

The trick is to **be persistent**: It may take several iterations, but if you work methodically, you can usually nudge the model closer to your vision with each round of feedback and revision.

Seeing what went wrong can help you understand how to **reformulate** your prompt for the **next iteration**. Every 'failed' generation is an opportunity to learn a little more about prompting.

Sometimes you aren't yet sure what you *don't* want, until you actually see it. Why did it fail? Differing interpretations? Sub-optimal description? Ambiguous instruction? Too-thin context?



Physics-first Prompting

Today's models can parse complex, materially specific language, and even reward it with photorealistic output.

For high-stakes client campaigns (like product shots, packshots, and beauty editorial work), a physics-driven prompt structure gives you fine-grained control and repeatable results. This methodology offers a robust, modular framework for achieving a more granular level of creative control.

- 1 Camera + geometry first:** Critical for product and packshot work. Lock your framing, orientation, and lens geometry up front.
- 2 Material physics + object behaviour:** Describe what materials *do*, not just what they look like. Physics language - tension variation, collapse pockets, microbubble fields, thin-film interference works well here. A bubble that has "burst membrane traces and uneven collapse" renders more convincingly than one that's simply "delicate."
- 3 Lighting:** Specify the rig, not the vibe. Directional key, cast shadows, highlight falloff, controlled fill — these translate into dimensional, commercially credible output.
- 4 Aesthetics comes last:** Once structure, physics, and lighting are locked down, aesthetic language finally has context to land in.

Physics-first prompt:

"Top-down flat lay product photography - camera perfectly perpendicular. Full opaque white cosmetic pump bottle centered and fully visible. Skincare packaging. Not beverage, not glass, not transparent, and not a drink container. The bottle is deeply embedded into thick fluffy cleansing foam covering about 90% of surface. Foam: volumetric with height variation layered peaks and ridges; mixed bubble sizes; microfoam clusters and large transparent soap bubbles with some flecks floating above surface. Water droplets and microbubbles cling to the bottle, foam reaching mid-way up the height of the bottle, with a clear negative contour around its silhouette. Bright directional hard studio light, casts sharp, well-defined shadows, resulting in a strong contrast with crisp highlight. The textured damp surface is visible only in small gaps between foam. Clean technical beauty product photography, geometry. 100mm lens f11 ISO100 Hard key light with minimal fill."

Model:

Nano Banana 2



150ml / 5.1 fl oz

The Multi-Disciplinary Lexicon

Borrow other worlds' words.

The following is a quick reference guide filled with precision vocabulary, borrowed from various disciplines. Use them freely, and experiment liberally!

PHOTOGRAPHY

- "golden hour"
- "bokeh"
- "Chiaroscuro"
- "fill light"
- "leading lines"
- "negative space"

CINEMATOGRAPHY

- "dutch angle"
- "Verité"
- "jump cut"
- "tracking shot"
- "mise en scène"
- "establishing shot"

ART MOVEMENTS

- "Bauhaus"
- "Afrofuturist"
- "De Stijl"
- "vaporwave"
- "Surrealism"
- "Cubism"
- "Impressionism"

The Multi-Disciplinary Lexicon

Borrow other worlds' words.

The following is a quick reference guide filled with precision vocabulary, borrowed from various disciplines. Use them freely, and experiment liberally!

TYPOGRAPHY

- "tracking"
- "leading"
- "font weight"
- "hierarchy"
- "negative space"

MUSIC THEORY

- "staccato"
- "legato"
- "call & response"
- "dynamics"
- "motif"
- "dissonance"

ARCHITECTURE

- "brutalism"
- "open plan"
- "load-bearing"
- "threshold"
- "biomimicry"
- "parametric facade"

The Multi-Disciplinary Lexicon

Borrow other worlds' words.

The following is a quick reference guide filled with precision vocabulary, borrowed from various disciplines. Use them freely, and experiment liberally!

FASHION & TEXTILE

- “deconstructed”
- “haute couture”
- “draped”
- “silhouette”
- “bricolage”
- “monochromatic”
- “ready-to-wear”

LITERATURE

- “in medias res”
- “bathos”
- “first person”
- “polyphony”
- “fourth wall”

GAME DESIGN

- “diegetic”
- “emergent”
- “pixel art”
- “HD Render”

Prompting for Video

Think like a film-maker, animator, or motion graphics specialist.

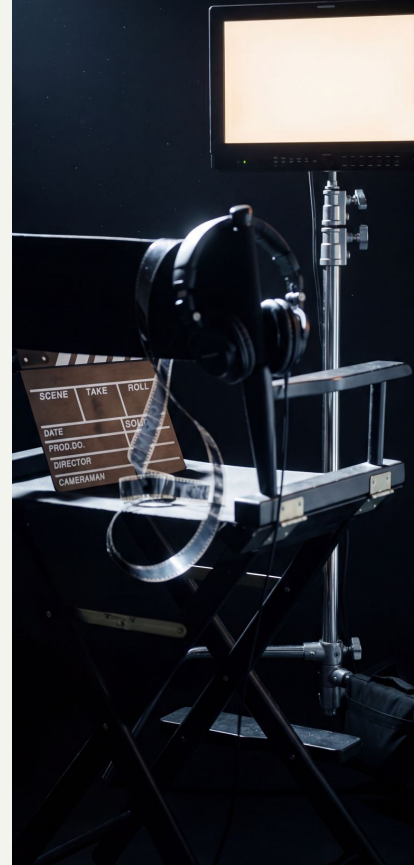
Everything in the previous principles applies to video prompting. But there's one additional consideration: motion.

There are a few ways to prompt for video generation: **Text-to-video** generates from a written prompt alone. **Frames-to-video** lets you define a start frame, an end frame, or both, giving you precise control over where the clip begins and lands. **Ingredients-to-video** lets you attach up to 3 reference images of a character, product, or style as visual anchors.

Specify **camera movement** (“tracking”, “pan”, “dolly”, “static”, “orbit”), **subject movement** (“walks toward camera”, “turns”, “picks up the product”), and **element animation** (“text slides in from the left”, “logo fades up”, “particles disperse”).

Be explicit about timing, by saying what should happen when. Clips are typically 4 to 8 seconds, so every second counts. It's hit or miss, but the models do *try* to adhere to your timing specifications.

Current models can generate clips with built-in cuts, transitions, and editorial rhythm. Between **smart prompting** and Pencil's built-in **ads editor**, you can craft exquisite video content without the need for any other video editing software..



Video with Native Audio

Principle 14

The latest video models generate lip-synced audio natively: soundtrack, VO, foley, dialogue, all in one generation.

If you don't specify audio, the model will make a choice.

Diegetic:

Sound within the world of the video. For example: the product hitting a surface; or the sound of water pouring. Specify with the same physics language as visuals. Prompting for “diegetic audio only” is a simple way to ensure coherence across clips.

Non-diegetic:

Sound outside the world, for example; VO and music tracks. Specify genre, instrumentation, tempo, emotional register. Quality is genuinely high. But this is where coherence across multiple clips becomes a challenge.



Video with Native Audio

The latest video models generate lip-synced audio natively: soundtrack, VO, foley, dialogue, all in one generation.

If you don't specify audio, the model will make a choice.

The coherence challenge:

Short clips mean audio will drift across generations. The industry is converging on coherence, but for now, this is still a known limitation.

Plan for it. Specify audio consistently in each clip's prompt, and be prepared to unify your audio tracks in a video editor in post production.

You may want to specify “no backing soundtrack or music” in your prompts, and then add a single music track later, for example. Or, specify a voice for your character to try and maintain coherence across generations. Since this method is unreliable, you may still need to swap out the voices to achieve character consistency for longer sequences.



Let AI write your prompts

AI knows how to talk to AI...

...sometimes much better than us humans do!

Describe your intent in plain language to an LLM chatbot. Ask the model to generate a detailed, formal prompt for your preferred image model, adhering to your desired prompt template or structure.

Let AI help you find the right words

If you're struggling to find the words to describe the style or vibe of a particular image, you can also attach an example of it for another LLM and ask for a detailed description, or analysis of the style.

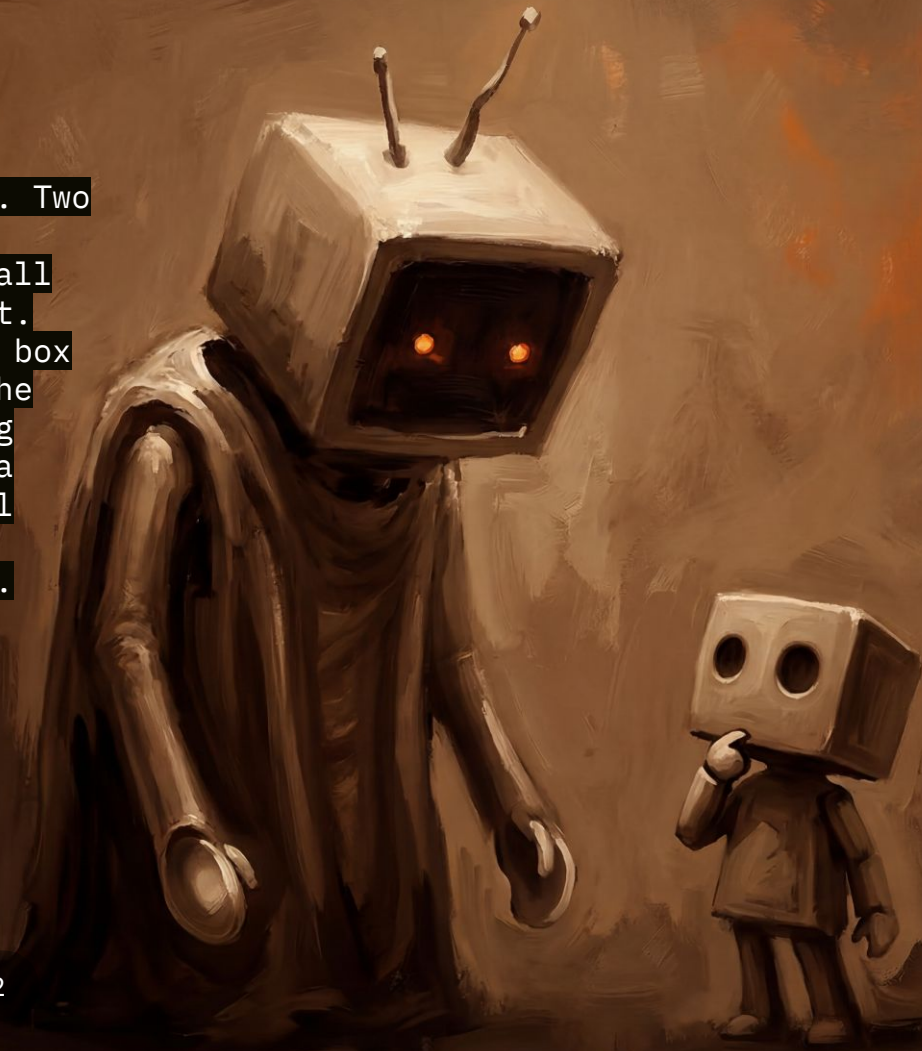
This technique can reveal certain niche domain- or discipline-specific keywords that you may have overlooked, or just not been aware of.

Principle 15



Prompt:

"Impasto oil painting. Two robot figures - tall teacher robot, and small child or learner robot. They both have simple box heads, antenna, and the larger one has glowing amber eyes. Warm sepia colour scheme. Minimal composition with a near-black background. Tender, melancholic. Heavy shadow, single diffuse light source. Expressive loose brushwork."



Model:

Nano Banana 2

Let AI write your prompts

Principle 15

AI knows how to talk to AI...



The annotation hack for Nano Banana and Veo 3.1 models

Add text labels on the input reference image itself. You can use this for detailed fine-grained image-to-image editing jobs with Nano Banana 2/Pro, where you need an extra level of control.

When creating videos with Veo 3.1, you can also annotate your starting frame, and the model will try to direct the clip according to the direction given in the annotated start frame. You may need to trim off the first few frames of your video, as the annotation text sometimes bleeds through for a moment.

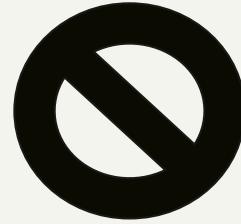
Risky Prompts to Avoid

AI image models can generate Unsafe or misaligned images.

Most systems have automated safety filters at both the input and output levels, but filters aren't foolproof. Word your prompts carefully.

Avoid generating:

- **IP & copyright** infringements.
Example: images of Disney characters.
- Directly **imitating signature looks** and styles of currently working creatives, such as Wes Anderson or Studio Ghibli.
- **Offensive material**, including violent, and NSFW outputs.
- **Politically incendiary** visual narratives.
- **Culturally insensitive** imagery.

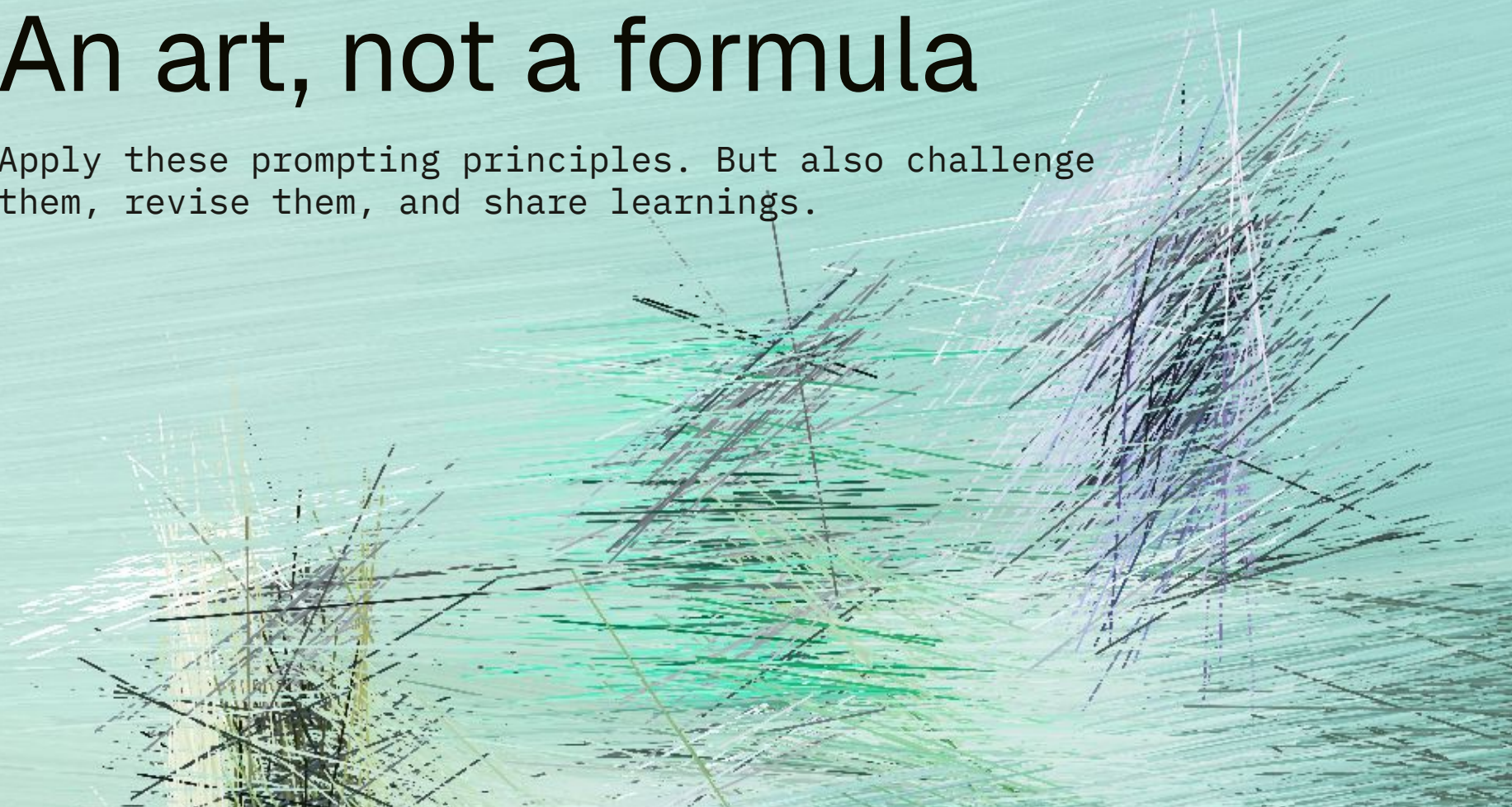


Principle 16



An art, not a formula

Apply these prompting principles. But also challenge them, revise them, and share learnings.



Pencil



* Use your human taste and discretion. You're ultimately the judge of which prompts and outputs are good or bad.

* The art of prompting is a collective, ongoing practice.
You're part of building it.

Pencil & the
brandtech
group