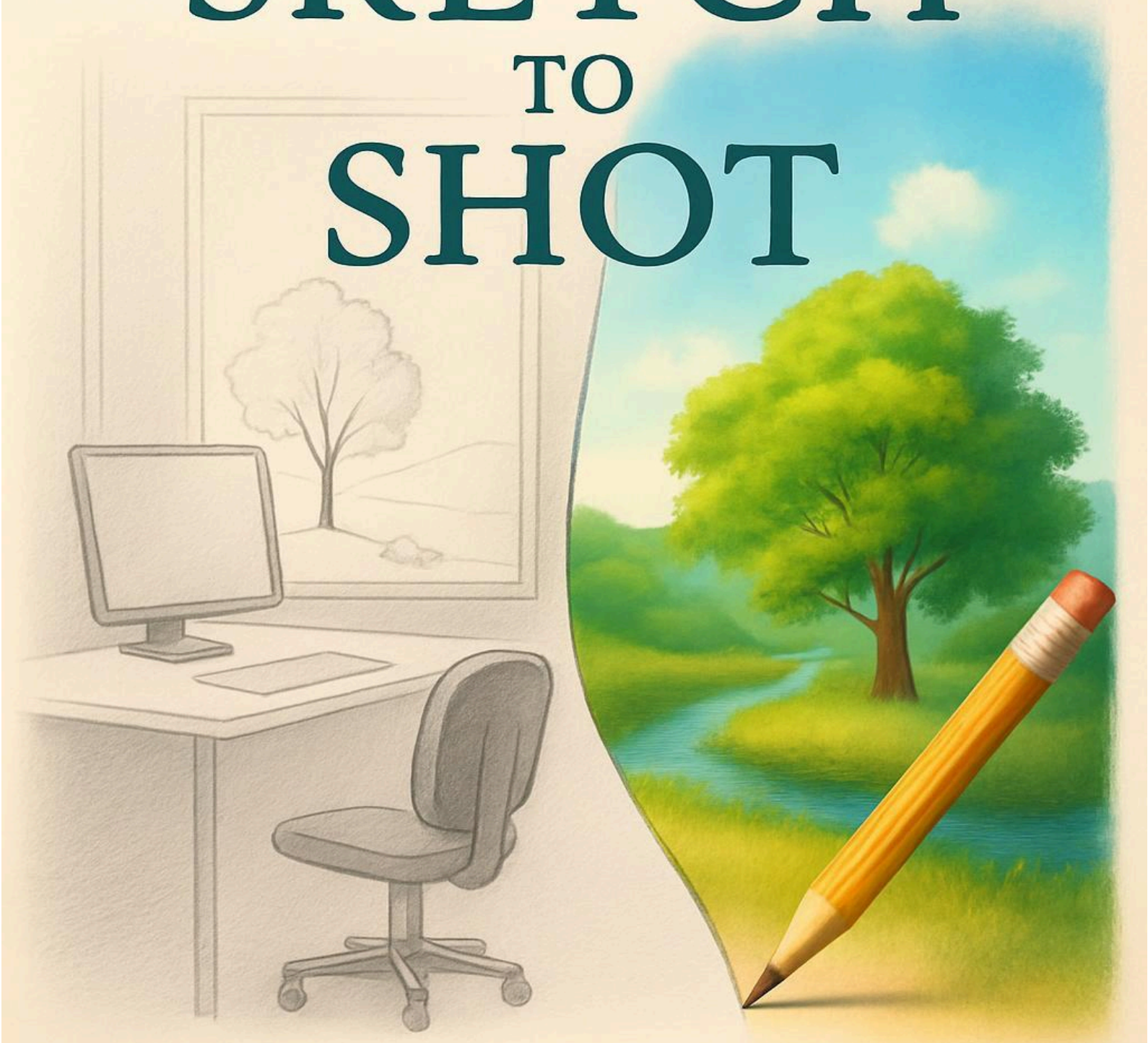
FROM SKETCH TO SHOT



WRITING PROMPTS THAT WORK

From Sketch to Shot:

Writing Prompts That Work

Why Your AI Images Look Like Chaos (And How to Finally Tame Your Prompts)

Every word you type shapes the image you get. Miss one detail, and your picture can go completely off the rails.

If you've ever tried to generate an image and felt… underwhelmed, you're in good company.

Maybe you aimed for a clean, simple greeting card and somehow ended up with something that looked like the side panel of a budget theme park ride. Maybe you once spent rainy afternoons sketching in pencil, chasing that quiet spark of creativity, and hoped AI art might bring it back — only to find that the "fun" part now seems buried under guesswork and frustration.
Or maybe you just want to capture a mood, tell a tiny story in one frame, or experiment for the sake of curiosity… but every attempt comes out clumsy and random.

This chapter isn't here to turn you into a "master prompt engineer" or give you a certificate in prompt wizardry. It's here to crack open one of the most fascinating creative tools we've ever had and make it feel less like gambling.

AI Doesn't "Get" You

Let's get this out of the way: **the AI does not see the picture in your head**. It doesn't know your intentions. There is no psychic channel connecting your brain to the model.

When you type something like "beautiful girl on a beach" and hit generate, the AI isn't visualizing your scene. It's breaking your words into tokens (little chunks of text), matching them to millions of examples it's been trained on,

and stitching together something that statistically fits. The catch? The "match" might be wildly different from what you imagined.

So instead of your dream scene, you might get a mutant who looks like they lost a fistfight with a carnival monster.

Here's the important thing: AI image generation is not magic, it's pattern matching. That's why vague prompts produce unpredictable results.

And the flip side of that truth? If you learn how the AI "reads" your words, you can stop gambling and start steering.

The Power of Specifics

Prompt example:

```
Man fishing on the shore.
```

You probably imagine a calm scene: a figure with a rod, gentle waves, maybe the glow of sunset. The AI, however, has seen countless interpretations of "fishing" — nets, spears, ice holes, bare hands, even scenes from survival shows. Without more detail, it's free to choose any of them.

A small change makes all the difference:

```
Man with a fishing rod, fishing on the shore.
```

That single added phrase narrows the possibilities and locks the AI into *your* mental picture.

Why this works: Models don't "understand" concepts the way we do. They understand *patterns of words* and the images those patterns have been linked to. Give them a pattern that's too broad, and they'll fill in the blanks with their own defaults — which may not match yours.

The Model's Autopilot Zone

When you leave gaps in your prompt, the AI doesn't leave them blank, it fills them for you. And it does this based on three things:

- **Statistical memory:** It leans on what it has seen most often in training data for those words. If "`cat ears girl`" in its dataset is mostly anime, you'll get anime, even if you were imagining a realistic portrait with fur-textured ears.
- **Model biases:** Every version of a generator has its own comfort zones. One might drift toward high-fantasy drama, another toward clean stock photography.
- **Stylistic glue:** The AI tries to bind your words into a coherent image. Wherever you're silent, it inserts whatever seems "most logical" to it, which might not be logical to you.

Word Order: Shapes Focus

Prompt example:

```
A car, with a person standing next to it.
```

In most cases, the AI will treat "car" as the main subject, rendering it in sharp detail, while the person becomes an afterthought — a small silhouette or a blurry figure.

Flip the order:

```
A person standing next to a car.
```

Now the person becomes the hero of the scene, with the car as background.

Why this works: In many text-to-image models, earlier words in a prompt often carry more weight, though this can vary depending on the model and interface, giving more weight to earlier words. Think of it like a movie poster: the first name listed is the star.

Advanced tip: group related details together. If you're describing a person, keep all person-related adjectives together before moving on to describe the other elements.

Prompt example:

```
A tall man in a leather jacket, standing next to a small vintage car with chrome details.
```

Avoid the "Don't Do This" Trap

Prompt example:

```
Portrait of a woman without a hat.
```

You'd think this is clear. But because the word "hat" is still present, and the model strongly associates "woman" with hats and accessories, you're likely to get... a hat.

It's like telling a toddler, "Don't touch the candy." Suddenly, the candy is irresistible.

Better alternatives:

`Portrait of a woman with long, flowing hair.`

`Portrait of a woman with an uncovered head.`

`Indoor portrait, cozy home setting.` *(Less chance of headwear indoors.)*

If you truly need to *ban* something, that's where negative prompts come in, but that's an advanced tool we'll cover later. For now, focus **on describing what you want, not what you don't.**

Don’t Force the Model to Choose

Prompt example:

```
A girl in a long coat with bright blue eyes stands on the shore, gazing thoughtfully into the distance, seen from behind over her shoulder.
```

Sounds cinematic, right? But you’ve just created a contradiction:

- “`Bright blue eyes`” → requires a front or side view.
- “`Seen from behind`” → no eyes visible.

The AI can’t have both, so it will try to compromise, usually in awkward ways.

Fix: Decide which detail matters more and commit to it.

If the pose matters:

```
A girl in a long coat stands on the shore, gazing into the distance, wind in her hair, seen from behind.
```

If the eyes matter:

```
Close-up of a girl with bright blue eyes in a long coat, standing on the shore, wind in her hair, distant gaze.
```

For the curious: When you give a prompt, the model breaks it into tokens. These tokens interact through a process called *attention*, which decides which details influence which parts of the image. Conflicting tokens (like “seen from behind” and “bright blue eyes”) create competing priorities, and the model will often produce an odd hybrid. Understanding this mechanic isn’t essential, but knowing it exists explains why contradictions rarely “blend” gracefully.

Choosing the Right Moment in Motion

Whenever you describe actions or movement, think in terms of a single, precise stop-frame. Asking the AI to capture "`a girl runs along the shore, hair flying, then looks over her shoulder with a smile.`" is tempting — it sounds dynamic — but it creates ambiguity. Should the model freeze her mid-stride, capture the hair at the peak of motion, or show the expression from a particular angle?

Without specifying the exact moment, the AI will guess. Often, it blends multiple possibilities into one awkward compromise: a slightly distorted pose, misaligned limbs, or an expression that doesn't match the intended action.

The fix is simple: choose the single instant that matters most for your image. Want the movement and pose to dominate? Freeze her mid-step, let the gaze be implied. Want the facial expression to be clear? Zoom in, freeze the expression, and let the limbs adjust naturally. This precision ensures your AI-generated image communicates exactly what you envisioned — no unintended hybrids or visual confusion.

AI Generators Aren't Photoshop

If you've ever tried to get an AI model to "add an effect" the way you'd apply a Photoshop filter, you've probably learned this the hard way: text-to-image generators don't work like that. They don't take an existing picture, run it through a ripple filter, and hand it back. They build the entire image from scratch, interpreting your words as a description of a scene, not a request for a post-processing step.

Prompt example:

```
The photo of an old house, looks like it's the surface of transparent water with a droplet falling, ripples distorting the image underneath.
```

Here's what the model **is not thinking:** "I'll render the house, then overlay a water-ripple effect."
Here's what it **is doing:** "I need to imagine a scene that could be described by those words." The result could be:

- Actual drops of water and visible ripples, but positioned as real objects in the scene, not as an optical distortion.
- Ripples or distortions that ignore real-world physics entirely.
- A completely re-imagined house or setting, because the model tries to merge "old house" and "underwater-like ripples" into a single concept.

Better approach: If you want something to look like a photo through water, describe the scene *as if you're photographing it in that condition*:

```
An old house seen through the surface of a clear pond, air ripples like water in the foreground, sunlight refracting through the water.
```

This sets the ripples as part of the environment, not as a Photoshop-style filter. You're giving the AI a coherent scene to synthesize, instead of an editing instruction it can't follow.

Why this matters: Photoshop and similar tools operate in layers, applying precise effects to fixed images. AI image generators work in probability space, synthesizing pixels based on patterns in training data. When you describe an "effect," the AI isn't layering it on top, it's inventing what such a thing *might* look like, all at once.

For the curious: Some advanced workflows (like using ControlNet or image-to-image pipelines) can mimic post-processing more closely, because they have an actual reference image to work from. But in plain text-to-image prompting, you're asking the model to *imagine*, not to edit — and imagination is messy.

Quick Practice

Grab an old prompt and ask yourself:

1. Did I leave anything vague?
2. Could word order be shifting the focus away from what matters?
3. Did I describe what I don't want instead of what I do?

Rewrite, regenerate, and compare. You'll be surprised how quickly the chaos starts to shrink.

Key takeaway: AI is literal. Every word you type is a clue it uses to guess what you want. If you're vague, you're letting a caffeinated parrot with a paintbrush take the wheel.

Building Blocks: Crafting Prompts Like Lego

(And Why This Method Actually Works)

You've already patched up the rookie mistakes. Now it's time to graduate from simply *describing* a scene to *directing* it: the difference between telling your friend about a movie and actually being behind the camera.

A prompt isn't just a shopping list of nouns. It's a blueprint. Every word tells the model both *what* to create and *how* to interpret it. When you just throw in "castle, sunset, dragon," you hand the AI a box of Lego bricks without any instructions. Sure, it will build something. But maybe you wanted a medieval fortress, and it gives you a pink fairy-tale tower with suspiciously cheerful gargoyles.

The fix: structure your prompt into four functional blocks. Think of them like a film crew — each block has a specific job, and together they produce the finished "shot."

Art Style / Aesthetic: Choosing the Visual Language

This block sets the *grammar* of your image. It's the first filter the AI applies to everything else you describe.

Without it: The model defaults to whatever it thinks is "normal", usually photorealism.
With it: You can jump between Pixar whimsy, gritty noir, or surreal dreamscapes just by naming the style.

Prompt examples:

- `Digital art, vaporwave neon hues` → flat, high-contrast, cyber-tinged look

- `Studio Ghibli anime style` → softer lines, warm palette, childlike wonder
- `Brutalist black-and-white photography` → harsh shadows, unapologetic realism

Why it works: AI models are trained on huge datasets of images tagged with style descriptors. Naming a style immediately narrows the field of reference, so the model "speaks" in that dialect.

For the curious: If you add the style block *after* describing the scene, the model often reinterprets details to fit the style, which is why your dragon might sprout big round anime eyes if you tack on "Ghibli" at the end.

Subject & Scene: The Story Core

This block is where you place your characters, objects, and environment. Without specificity, the model fills gaps on its own and its imagination can be… eclectic.

Without it: "`Castle`" could be a ruin, a Disney palace, or a sandcastle.
With it: "`Ancient weathered stone castle perched on a cliffside`" gives the model a precise mental picture.

Prompt examples:

- `Futuristic market street with glowing stalls, android vendors, and floating food drones`
- `A lone astronaut planting a tattered flag on a jagged ice moon`

Why it works: Precise nouns and adjectives reduce ambiguity. The more vivid your language, the more the model aligns with your mental image. It's not about adding *more* words, it's about choosing words that do heavy lifting.

For the curious: Models don't "imagine" in the human sense. They map words to clusters of visual patterns. If you say "frostbitten," the AI doesn't feel cold, it just pulls imagery patterns that have historically co-occurred with that word.

Multiple Subjects & Interaction: Shaping Relationships

When your scene involves more than one subject, the structure of your prompt starts to act like stage direction. The model doesn't just "see" the objects; it builds them in sequence, attaching details more strongly to whatever words are closest. That's why order matters: you're effectively choosing whether the model draws characters side by side, locks them into a shared action, or grounds them within a broader atmosphere.

There are several common ways to shape these relationships:

Option 1: Separate subjects, interaction added after
Structure: *Subject 1 description, Subject 2 description, description of their interaction.*
Effect: The model first "builds" each subject individually, often with more symmetry and detail, and only then layers their relationship on top. The interaction can feel static or weaker, as if pasted over two already finished figures.
Prompt example:

```
A young woman with long red hair, wearing a white dress, sitting on the floor.
A fluffy gray cat with bright green eyes.
They are playing together, the cat reaches for her hand playfully.
```

Option 2: Subject 1 tied directly to the action with Subject 2

Structure: *Subject 1 description, description of interaction with Subject 2, Subject 2 description.*

Effect: The interaction is woven into the portrait of the first subject, making the action feel more natural and embedded. The second subject sometimes gets slightly less detail, since it enters later as a supporting role.

Prompt example:

```
A young woman with long red hair, wearing a white dress, sitting on the floor, reaching her hand to play with a fluffy gray cat.
The cat has bright green eyes.
```

Option 3: Scene as foundation, action and reaction

Structure: *Environment/scene, Subject 1 and their action — Subject 2 and their reaction.*

Effect: This is often the most fluid and cinematic. The environment sets the mood, and the subjects are framed as part of a dynamic moment, emphasizing atmosphere and interaction in motion.

Prompt example:

```
A cozy living room with warm evening light.
On the floor, a young woman with long red hair in a white dress is sitting and laughing, playing with a fluffy gray cat with bright green eyes who jumps toward her hand.
```

Why it matters: Think of it as narrative grammar. Option 1 is like listing characters before showing them interact. Option 2 fuses one subject's identity with the action. Option 3 treats the scene as a story unfolding, where environment and movement come first. None is "better" in the abstract, it

depends on whether you want symmetry, intimacy, or atmosphere to dominate the composition.

By separating subjects and using explicit references, you reduce the model's chance of confusing attributes. Each description becomes a self-contained block of visual information, and the interaction ties them together in the correct order. The structure also makes it easier to troubleshoot: if the model gives the wrong jacket to the wrong person, you can adjust that block without rewriting the rest.

Advanced tip: Instead of pronouns like his/her/its, repeat the subject's name or role.

Prompt example:

```
Young woman in elegant dress standing on balcony, woman's hands resting on wooden railing, detailed fingers gripping railing. Man in dark suit standing next to woman, man's hand gently resting on top of woman's hand on railing. Sunset light casting long shadows across the balcony floor.
```

This repetition may sound redundant to a human reader, but to the model it's a map, not a novel. Language models for image generation don't truly understand grammar in the way humans do. When they parse a pronoun like *her*, they try to match it to the most recent compatible subject, which can easily be the wrong one if multiple subjects are present. Explicit repetition removes that ambiguity by "pinning" attributes to the correct subject node in the model's mental representation.

Backgrounds: Filling the Canvas

Backgrounds behave like the silent partner in your scene. The model will always complete them, even if you don't explicitly describe them: pulling from the "visual canon" of the subject or genre. Say "knight in armor," and if you don't add more, you'll often get a castle or a battlefield, because that's what history (and dataset patterns) have taught the model to expect.

Less detail = more improvisation

When you give only minimal cues, the generator leans on its own associations. The result is less predictable, but often more atmospheric and fluid, as if the world is growing naturally around your subject.

Prompt example:

`A knight in medieval armor, standing proudly.` → Likely result: a vague castle wall, a cloudy sky, maybe banners.

More detail = fixed composition

The more specific you are, the less room the model has to improvise. This can create a clean, controlled image, but also risks overloading the frame if every background element is spelled out.

Prompt example:

> `A knight in medieval armor, standing on a cobblestone bridge at dusk, with a looming Gothic castle behind him and gnarled trees on either side.` → Stronger control, but the scene can feel heavy if every element is fighting for attention.

Finding the balance

A handful of strong details is usually enough: time of day, one or two anchor objects (castle, forest, marketplace), and the general mood (foggy, golden,

neon-lit). That way the model has structure to follow, while still leaving space for the atmosphere to breathe.

Why it matters: Backgrounds don't just "sit" behind the subject, they frame the story. Too vague, and you hand the pen fully to the model. Too rigid, and the image can collapse under its own weight. The sweet spot is giving just enough detail to guide the composition while allowing the model's improvisation to weave cohesion and mood.

Lighting & Mood: The Emotional Palette

Light changes more than you think. It affects mood, depth, texture, and even how we read expressions.

Without it: Scenes feel flat or strangely neutral.
With it: You control the atmosphere before the viewer even registers the subject.

Prompt examples:

- `Cinematic volumetric light cutting through thick fog` → suspense, drama
- `Golden hour backlight with long soft shadows` → warmth, nostalgia
- `Neon glow reflecting off wet asphalt` → cyberpunk moodiness

Why it works: Humans read light cues instinctively, we associate long shadows with dusk, soft light with intimacy, high contrast with tension. The AI mimics this because those correlations exist in its training data.

For the curious: Some models respond better to mood adjectives (“ominous,” “serene”) than to technical lighting terms. Combining them often produces the richest results.

Quality Enhancers & Technical Details: The Polish

This block is your post-production team. It doesn’t change the story, but it affects how crisp, detailed, and refined the image looks.

Without it: Images might be fuzzy, low-detail, or oddly simple.
With it: Print-quality, gallery-ready results.

Prompt examples:

- `Ultra detailed, high dynamic range, sharp focus`
- `8K render, cinematic composition, intricate textures`

Why it works: These terms push the model toward high-quality training examples, biasing it toward cleaner lines, richer textures, and professional-level compositions.

Note: Quality tags aren’t magic**.** Many people add “`masterpiece`” to every prompt, expecting better results. What it really tells the model is: *“make it artistic”*. Sometimes that’s good, but it can hurt realism.

Why Modular Thinking Works

Breaking prompts into blocks is more than an organizational trick:

- **Clarity for the model**: Reduces word salad, gives structure.
- **Creative flexibility**: Swap just the style block, and you’ve got a new image without touching the story.
- **Efficient troubleshooting**: If something feels off, you know exactly which block to tweak.

Putting It All Together

1. **Style**: Digital art, dystopian vaporwave palette
2. **Subject**: Cyberpunk detective in a rainy alley
3. **Lighting**: Neon reflections, cinematic backlight
4. **Quality**: Ultra detailed, sharp focus, 8K

Resulting prompt:

```
Digital art in dystopian vaporwave style, a
grizzled cyberpunk detective leaning casually
against a wall in a rain-soaked neon alley with
holographic billboards glowing in the distance,
cinematic backlight, reflective puddles,
atmospheric haze, ultra detailed, sharp focus, 8K
```

Practice Tip: Don't try to cram all four blocks in at once. Start with the subject, add style, then lighting, and finally quality. That way you'll **see** what each layer changes and learn to control it like a director adjusting the set.

Advanced Playbook: Fine-Tuning Your AI Art

How to use angles, depth, and composition to make your images cinematic and alive.

By now, you can stop your AI from giving hats to everyone, and you've built prompts with structure. Good. Now it's time to level up.

This is the point where you stop thinking like a *prompt typist* and start thinking like a cinematographer, someone who knows how to control the invisible camera that shapes every image.

Angles decide what your viewer sees first, how large or small a subject feels, and whether your scene carries tension, grandeur, or intimacy. Without them, even your most ambitious concepts can end up looking like bland stock photos.

The Language of Camera Angles

AI models, left to their own devices, usually default to a safe mid-shot at eye level. It's pleasant, symmetrical... and boring. Your job is to move the camera — and the viewer — somewhere more interesting.

Here are the most useful "camera vocabulary" terms and what they do:

- **Close-up** — Brings the viewer's face-to-face with a subject. Great for emotion or detail.
- **Extreme close-up** — Pushes in further to highlight tiny elements: eyes, lips, a lock of hair.
- **Medium shot** — Waist-up framing. Natural, conversational, balanced.
- **Long/Wide shot** — Shows the full subject and environment together. Ideal for scale.

- **Full-body view** — Focuses entirely on the subject's figure; often used for fashion or poses.
- **Dutch angle** — Tilts the horizon. Creates unease, tension, or dynamism.
- **Bird's-eye view** — From above, giving a map-like or surveillance feel.
- **Worm's-eye view** — From low to high, making subjects loom and dominate.
- **Over-the-shoulder** — Adds intimacy and perspective, like you're inside the scene.
- **POV** (point-of-view) — Places the viewer in the character's shoes. Highly immersive.

When Angles Collide (and How to Fix It)

Example problem:

```
Woman, realistic facial pore texture, long shot, full body view.
```

What you wanted: a full figure, distant enough to see her whole outfit.
What the AI delivered: a zoomed-in portrait.

Why? "Realistic facial pore texture" has a high visual priority, the model *thinks* you want extreme detail on her face, and the easiest way to deliver that is to crop in close.

The fix: Reinforce the wide frame by adding details that only make sense from a distance:

```
Woman wearing a flowing red dress and intricate high heels, realistic facial pore texture, long shot, full body view, standing on a grand marble staircase.
```

Now the model sees multiple elements (dress, heels, staircase) that it can't show without staying wide.

Why it works: AI doesn't "balance" details like a human would. If you ask for both *full body* and *pore detail* without giving extra wide-shot anchors, it will pick one. Extra environmental or clothing elements act as anchors that keep the camera where you want it.

For the curious: Visual tokens compete for "attention budget." Large-scale elements (environment, posture) and small-scale ones (skin texture, eye reflections) often fight for priority. Adding enough wide-dependent tokens tilts the balance toward the angle you want.

Creating Depth and Volume

Flat images feel cheap. Cinematic images feel layered, with a foreground, midground, and background working together.

Think in three layers:

- **Foreground** — Objects close to the camera: foliage, rails, hands, smoke.
- **Midground** — The main subject: your character or focal object.
- **Background** — World-building elements: distant mountains, buildings, sky.

Depth of field tools:

- **Shallow DOF** — Subject sharp, background blurred. Adds intimacy.
- **Deep DOF** — Everything in focus. Epic, documentary-style scale.
- **Bokeh** — Soft glowing circles from out-of-focus lights. Romantic or dreamy.

Prompt example:

```
A lone adventurer standing on a cliff edge, looking out over a vast, misty valley.
Foreground: Crumbling rocks and sparse hardy bushes.
Midground: The adventurer with a weathered backpack and a sword.
Background: Distant snow-capped peaks and swirling clouds.
Angle: Wide shot, cinematic depth of field, volumetric fog.
```

Suddenly the image has dimension — like a still from a movie instead of a flat poster.

Extra Angles and Composition Tips

- **High vs. Low Angles** — Low angles make characters heroic or intimidating. High angles make them vulnerable or small.
- **Leading Lines** — Roads, rivers, or architectural lines that guide the eye to your subject.
- **Framing Within the Frame** — Windows, doorways, foliage that "frame" your subject for focus.
- **Negative Space** — Empty areas that give breathing room and emphasize scale.

Key Takeaways:

- Be explicit about your camera position and angle.
- If multiple elements conflict (close-up vs. wide shot), reinforce your preferred frame with details that "lock in" that choice.
- Use depth layering to make your image feel alive and cinematic.

Critique & Common Pitfalls

- Overstuffed prompts: **More isn't always better**. If your prompt reads like a grocery list, the AI may ignore parts entirely. Group related descriptors together.
- Vague angles: Words like "dynamic" or "interesting" don't tell the model where to put the camera. Use **concrete terms**.
- Ignoring the environment: Even a stunning character **looks flat without context**. Give the world as much thought as the subject.

Painting Emotion with Light

Lighting isn't just a way to see a scene. It's a silent narrator. Soft light whispers comfort; hard shadows shout danger. The same subject can be tender or terrifying depending on how light touches it.

Once you've chosen your camera angle, lighting is what separates a competent AI render from one that *feels alive*. Just as in traditional art and cinematography, light shapes atmosphere, guides the eye, and silently tells the viewer how to feel.

AI models have seen millions of examples — they've absorbed visual grammar. The right lighting cue can turn a cozy living room into a place of suspense, or an empty alley into a romantic hideaway.

The Language of Light

Below are common lighting styles, with their moods, uses, and example prompts.

Soft Light — Dreamy and Flattering

Gentle transitions between light and shadow. Think overcast days, or sunlight filtered through sheer curtains. Soft light wraps around shapes, smooths edges, and feels natural and calm.
When to use: portraits, intimacy, nostalgia.

Prompt example:

```
soft natural light, gentle illumination, smooth transitions
```

Why it works: Soft edges trigger visual associations with comfort, safety, and memory. AI reads “soft” and “gentle” as instructions for tone as much as physics.

Hard Light — Drama and Tension

Sharp shadows, strong contrast. Like midday sun in a cloudless sky. Creates definition and visual punch.
When to use: action, conflict, high-stakes realism.

Prompt example:

```
harsh sunlight, strong contrast, sharp shadows
```

Why it works: High contrast exaggerates form and tension. AI links “harsh” and “sharp” to cinematic danger and intensity.

Golden Hour — Romantic and Warm

That first hour after sunrise or last before sunset. Golden glow, long shadows, soft atmosphere.
When to use: peace, beauty, emotional storytelling.

Prompt example:

```
golden hour lighting, warm glow, long soft shadows
```

For the curious: Golden hour light is naturally low in contrast and rich in reds/yellows — both of which the human brain reads as warmth and nostalgia.

Backlight — Depth and Mystique

Light from behind the subject creates glowing edges or full silhouettes.
When to use: mystery, elegance, iconic reveals.

Prompt example:

```
dramatic backlighting, luminous silhouette, glowing
rim light
```

Volumetric Rays — Epic and Ethereal

Visible beams of light cutting through haze or dust. Adds scale, wonder, and cinematic depth.
When to use: fantasy, awe, sacred moments.

Prompt example:

```
volumetric lighting, god rays piercing mist,
cinematic atmosphere
```

Shadows — Depth and Realism

Without shadow, scenes look flat and artificial. With it, they gain dimension and mood.
When to use: realism, suspense, chiaroscuro drama.

Prompt example:

```
deep shadows, high contrast between light and dark
```

Bokeh — Cinematic Focus

Soft background blur with glowing circles of light. Frames the subject while adding visual charm.
When to use: romance, intimacy, cozy settings.

Prompt example:

```
shallow depth of field, beautiful bokeh, blurred background lights
```

Neon Glow — Cyberpunk Energy

Artificial, vibrant light bouncing off metal and wet streets. Instantly signals urban futurism.
When to use: sci-fi, nightlife, high-tech dystopia.

Prompt example:

```
neon glow, vibrant city lights, wet reflective streets
```

Why This Works: Lighting styles act like emotional shortcuts. They instantly cue the viewer's brain to a story context — before they've even processed the characters or setting.

- **Soft light** = safety, romance
- **Hard light** = urgency, grit
- **Golden hour** = nostalgia, calm
- **Neon** = chaos, tech, energy

AI responds because these cues are statistically linked to those moods in its training data.

Mood: The Invisible Layer

Lighting sets the stage, but *mood* fills the air. It's the difference between a technically correct image and one that feels cinematic.

Instead of describing only what is in the scene, describe how it feels.

Mood Keywords: Speaking AI's Emotional Language

These are compact, high-impact terms that pull together composition, color, and tone.

- **Dramatic cinematic shot** — strong framing, bold grading, tension.
  `a lone knight on a hilltop, stormy sky, dramatic cinematic shot`
- **Dreamy atmosphere** — soft focus, pastels, ethereal haze.
  `a glowing garden under a pale moon, dreamy atmosphere, pastel colors`

- **Dynamic tension scene** — motion blur, diagonals, sharp contrasts.
  `two martial artists clashing mid-air, dramatic backlighting, dynamic tension scene`

Why it works: These words bundle many subtle cues. AI translates them into lighting, lens choice, color palette, and even character expressions.

Film & Photography Elements: Small Changes, Big Mood

Use sparingly, like seasoning:

- `Vintage Polaroid photo` → nostalgia, imperfection
- `Film grain, anamorphic lens flare` → tactile, authentic
- `Wide-angle lens` → scale and drama
- `Macro shot` → intimacy and focus

Making It Feel Alive

Real scenes breathe. Add sensory and emotional details:

- **Sensory:** visible breath in cold air, hair ruffled by wind, steam curling from a cup.
- **Body language:** slumped shoulders, confident stride, gaze locked on horizon.
- **Textures:** soft velvet, cracked stone, rain-slick asphalt.

Prompt example:

```
a woman gazing toward a distant horizon, loose hair blowing in the wind, soft velvet cloak embroidered with silver threads
```

Putting It All Together: Layered Prompt Example

```
Intense, dramatic, painted aesthetic. A colossal gothic castle, ancient and weathered, perched on a jagged cliff hammered by ocean winds. Stone walls cracked by centuries of storms, a single tattered banner snapping in the gusts. Below, waves crash violently, sending foamy spray into the air. In the foreground, a vast empty beach scattered with driftwood and black pebbles. A lone cloaked traveler stands, facing the castle. Stormy twilight, volumetric sun rays breaking through clouds, lightning flickers in the distance, cinematic chiaroscuro lighting. Panoramic wide shot, hyper-realistic water spray, ultra detailed.
```

Breakdown:

- **Aesthetic:** painted, dramatic
- **Main subject:** gothic castle on cliff
- **Foreground:** driftwood, pebbles, sea
- **Human element:** lone traveler for emotion & scale
- **Lighting:** twilight, volumetric rays, chiaroscuro
- **Angle:** panoramic wide shot
- **Quality:** hyper-realistic, ultra-detailed

Each block adds control. If one element feels wrong, you can adjust without rewriting the whole prompt.

Key takeaway: Lighting and mood are not afterthoughts, they're narrative tools. In AI prompting, they're as critical as the subject itself. Think of them as the soundtrack and color grading for your still images. Without them, you have a picture; with them, you have a *story*.

Control & Priority: Getting the AI to Actually Listen

By now, you've learned how to paint vivid pictures with your prompts. You've described scenes in rich detail, picked the right words, and avoided ambiguity. And yet… sometimes the AI still ignores you.

You ask for a red dress — it hands you white. You want the foreground sharp — it throws it into a blur.

This isn't because the AI is stubborn. It's because it doesn't "think" in meanings the way we do. It doesn't interpret a "red dress" as an idea; it treats it as a set of visual patterns, weighted in a very particular way. To truly steer it, you have to speak its language of *priority and emphasis*.

Think of prompt writing as stage direction. You're not just telling the AI *what* to put on stage — you're telling it *where to shine the spotlight*.

Word Order: Attention Control

Prompt models scan from left to right, and earlier tokens grab more attention.

- "`girl in a red dress next to a tree`" → The girl is the star, the tree is the background.
- "`a tree, and next to it, a girl in a red dress`" → The tree dominates, the girl becomes secondary.

Why it works: In most text-to-image models, earlier tokens influence composition and allocation of detail more strongly. This is partly due to the way attention maps are built: the first tokens act like anchors for the rest of the scene.

Advanced tip: Put the most important elements first — even before adjectives that describe them — if you want them to visually dominate.

Token Weighting: Turning Up the Volume

Some models let you emphasize or de-emphasize words. Think of it as a volume knob for visual attention.

- `(vibrant red car)` → small boost
- `((vibrant red car))` → stronger boost

Colon syntax (common in Stable Diffusion models):

- `(vibrant red car:1.2)` → +20% emphasis
- `(subtle blue sky:0.8)` → tone down by 20%

When to use weighting:

- An element keeps being ignored
- Two elements compete for attention and you need a winner
- You want a texture or color to pop

Caution: Overweighting can warp shapes or spawn surreal extra limbs. Start with small changes and check the results.

Negative Prompts: Telling the AI What *Not* to Do

Sometimes, instead of boosting, you need to block. Negative prompts are the "no-entry" signs of prompting.

Common uses:

- Fix anatomy: `extra limbs, deformed hands`

- Remove clutter: `blurry, low resolution, watermark`
- Remove unwanted objects: `cars, people`
- Adjust mood: "`bright colors`" for a somber tone

Why it works: Negative prompts don't just *hide* things, they actively redirect the model's visual space away from patterns it associates with those terms.

Prompt example — Full-Scale Negative Prompt for Photorealism:

```
blurry, low quality, low resolution, deformed hands, extra fingers, missing fingers, extra limbs, fused fingers, long neck, mutated anatomy, bad proportions, poorly drawn face, bad eyes, cross-eye, squinting, watermark, text, logo, overexposed, underexposed, noisy, pixelated, jpeg artifacts, unnatural skin texture, unrealistic hair, oversaturated colors, cartoon, anime, 3d render, cgi, doll-like, plastic, flat lighting, harsh shadows, grainy
```

This kind of "negative shield" is usually pasted in one block, so the model gets a clear and exhaustive list of things to avoid when rendering. In photorealistic work, it's less about style changes and more about damage control — cutting out the glitches before they ever appear. Use long negative prompt lists for realism, but beware: overuse can flatten style and produce plastic-looking results.

Style Prompts: The Soft Hand of Influence

Style prompts are a gentle current under the whole image.

- `Studio Ghibli style` → whimsical colors, soft shapes
- `Brutalist photorealism` → hard lines, muted palette

Advanced tip: Don't just name the style, add mood cues:

- `Studio Ghibli-inspired, soft whimsical nature tones`"
- `Cyberpunk anime, neon-lit cityscape, dark moody palette`

Technical Parameters: Your Control Panel

For advanced models, these matter as much as your words.

- **Samplers:** Different "painting techniques" — try several.
- **Steps:** More steps = more detail, but diminishing returns.
- **CFG Scale:** Higher → sticks closely to prompt (7–12). Lower → more creative drift (3–6).
- **Resolution:** Higher = more detail but heavier load.
- **Seed:** Repeatable results or fresh variations.

Beyond the words of your prompt, image generators like MidJourney respond to a set of technical flags. These don't change the *idea* of your prompt, but they shape how the idea is executed: the canvas size, the level of detail, the degree of randomness. Think of them as dials on a mixing board.

Common parameters *(MidJourney-specific syntax, may change in future versions)*:

> `--ar X:Y` **(Aspect Ratio)**
> Controls the shape of the canvas. *1:1* for square, *16:9* for cinematic, *9:16* for mobile-friendly verticals. The ratio often nudges composition: tall for portraits, wide for landscapes.
>
> `--v N` **(Version)**
> Chooses the engine version. Higher versions usually mean improved

rendering, but each can have a distinct “look.” Sometimes older versions are better for specific styles.

`--q N` **(Quality)**
Rendering effort. *1* is default, *2* or higher spends more resources per image, yielding sharper detail but slower results. Lower values (*.5*) are faster, rougher drafts.

`--s N` **(Stylize)**
Dictates how strongly the model applies its own aesthetic. Low values (*50–100*) keep close to your literal description. High values (*500–1000+*) push toward more artistic, interpretive outcomes.

`--chaos N` **(Chaos)**
Adds randomness. Low chaos (*0–10*) gives consistent, predictable variations. High chaos (*50+*) encourages surprising, less controlled results.

`--no X` **(Negative prompt)**
Excludes unwanted elements. Example: `--no text, --no watermark` tells the model to avoid rendering those features.

Why All This Works: AI doesn’t “understand” in the human sense. Instead, it maps your words to probability clusters learned from millions of images and captions. The order, weight, and inclusion/exclusion of words shift these probability maps, much like turning dials on a mixing console changes the balance of a song.

Choosing the Right Model

Don't use a cinematic ultra-detailed model for quick sketches.

- **Simple, exploratory idea** → lightweight model.
- **Complex, polished final** → heavyweight, high-detail model.

Experiment: Stretching the Prompt

This chapter has focused mainly on technical parameters, the knobs and sliders you can use to control output. But before we wrap up, let's try a hands-on experiment to see how models handle complexity.

Pick your favorite image generator, choose a style (for instance, photorealism), and build the following prompt step by step instead of all at once:

`A red cat running across a table`…
→ then extend with: …`chasing a white mouse`…
→ then add: …`the mouse is holding a piece of cheese`…
→ then: …`a parrot flying above them, wings spread`…
→ then: …`a vase with flowers stands on the table`…
→ then: …`a sunset glows through the window`…
→ then: …`a dog sits in the corner, chewing a bone`…
→ finally: …`and a chandelier with flickering bulbs hangs from the ceiling.`

Most modern generators can keep up for a while, but as the description grows, you'll often notice drift: details being dropped, objects deforming, or the scene becoming cluttered. That moment of "slippage" shows you the boundary between what the model can handle smoothly and where it starts to improvise. By experimenting this way, you'll get a feel for your model's limits,

how much detail it can juggle, what it tends to forget, and which elements remain consistent.

Refinement and Iteration: The Heartbeat of Prompting

Your first prompt is almost never perfect. Treat it as a draft, not a decree.

1. **Analyze the output** — Which parts worked, which went rogue?
2. **Adjust keywords** — Add missing detail or remove repeating mistakes.
3. **Use weighting & negatives** — Fine-tune attention.
4. **Refine style cues** — Move beyond buzzwords.

This loop — *Try → Analyze → Adjust → Repeat* — is where mastery lives.

Final Takeaway: A good prompt grows from a simple wish list into a blueprint for creation. Control comes from *priority* (word order, weighting), *boundaries* (negative prompts), *mood shaping* (styles), and *refinement* (iteration).
When you stop asking AI and start directing it, your outputs stop being "lucky" and start being *yours*.

The Troubleshooting Field Guide

You've probably noticed that AI images fail in strangely predictable ways. At first it feels like chaos: the machine throws in hats you never asked for, limbs twist like they've escaped a horror movie, or your grand cinematic scene comes back looking like a washed-out stock photo.

We'll walk through the most common "symptoms" in AI images, trace them back to their likely causes, and show you the quick fixes that bring things back on track.

---

1. Subject Problems

The subject is the star of your image. If the star doesn't shine, the whole scene collapses.

When your subject looks tiny or lost:
This usually happens when the subject is buried at the end of the prompt. AI tends to allocate the most attention to whatever comes first, which means that "castle at sunset" will overshadow "a girl standing in front." The result: a postcard of a castle with a blurry, ant-sized human.

Fix it by:

- Putting the subject first.
- Giving them unique anchors: hair color, clothing, expression.
- Specifying framing: *"close-up," "medium shot," "full-body view."*

When accessories keep sneaking in: You asked for a simple portrait, and suddenly your subject has a hat, necklace, or a medieval helmet you definitely didn't want. Why? Because negative phrasing like *"without a hat"* still contains the word "hat." The AI can't help but bring it along.

Fix it by:

- Rewriting positively: *"with uncovered head," "long flowing hair."*
- Save negative prompts for advanced cleanup.

When details contradict each other: "Bright blue eyes" plus "seen from behind" equals one confused model. It can't show both at once, so it compromises, often with nightmare fuel.

Fix it by:

- Choosing what matters more and dropping the contradiction.
- If both details are essential, split them into separate prompts.

Example: Subject Lost in the Scene

Before:

```
a castle at sunset, a girl in front of the castle
```

Result: The castle dominates, the girl is tiny or blurry.

After:

```
a girl standing in front of a massive castle at sunset, full-body view, detailed clothing
```

Why it works: The subject is placed first, given unique anchors (clothing), and framed explicitly as full-body. The castle becomes context instead of stealing the spotlight.

---

2. Anatomy and Artifacts

The number-one AI meme is cursed hands: spaghetti fingers and melted palms. These aren't random either; they come from pushing the model into positions it can't handle.

When hands sprout extra fingers:

- Overly complex poses or action shots stress the model.
- Without a negative prompt "shield," it leans on noisy training data.

Fix it by:

- Adding a short corrective list: *"deformed hands, extra limbs, bad anatomy."*
- Simplifying the pose: describe the figure first, then add the environment.

When faces blur or drift: If the face is described late in the prompt, or buried under too many competing details, the model loses track of it.

Fix it by:

- Moving facial description to the front.
- Using clear tags: *"portrait, sharp focus, highly detailed face."*
- Switching to *"close-up"* if the face is the point.

Example: Cursed Hands

Before:

```
a man running through the forest, cinematic shot, dynamic action
```

Result: Extra arms, broken fingers, anatomy chaos.

After:

> `a man running through the forest, cinematic shot, dynamic action, (well-formed hands:1.2)`
> negative prompt: `deformed hands, extra limbs, bad anatomy`

Why it works: By explicitly weighting "well-formed hands" and adding a short corrective negative list, the model has a stronger anatomical anchor.

---

3. Lighting and Mood

Lighting is your silent narrator. Without it, scenes often look like bland catalog shots.

When everything feels flat: This means you've described objects but not atmosphere. No mention of light source, time of day, or mood.

Fix it by:

- Adding light cues: *"backlight, golden hour, volumetric rays."*
- Layering mood: *"dreamy atmosphere, tense cinematic shot."*

When the image is too dark or overexposed: It's usually because you relied on vague adjectives like *"dramatic"* or forgot to anchor the light source.

Fix it by:

- Pairing technical and emotional cues: *"soft natural light, serene mood"* or *"harsh direct sunlight, high contrast."*
- Grounding with context: candlelit tavern, neon-lit street, misty dawn.

Example: Flat Lighting

Before:

```
a portrait of a woman, cinematic, photorealistic
```

Result: Flat, catalog-like lighting, no mood.

After:

```
a portrait of a woman, cinematic, photorealistic, golden hour backlight, soft rim lighting, dreamy atmosphere
```

Why it works: Light cues (backlight, rim light) and mood words (dreamy) turn a sterile render into a cinematic scene.

---

4. Composition and Perspective

Composition is what makes an image feel cinematic instead of flat. Ignore it, and your subject might look like they've been pasted onto a background with bad scissors.

When the subject feels like a cardboard cutout: No depth layers, no camera angle, no context.

Fix it by:

- Adding foreground, midground, and background elements.
- Specifying a shot: *"wide-angle cinematic shot," "over-the-shoulder view."*

When you want a full-body shot but get a close-up: That happens because details like "realistic skin pores" or "eye reflections" signal to the AI: zoom in.

Fix it by:

- Adding anchors visible only at distance: outfit details, environment, props.
- Reinforcing with *"long shot, full-body view."*

Example: Wrong Perspective

Before:

```
a young knight in armor, photorealistic, cinematic, realistic skin pores
```

Result: Close-up of the knight's face instead of full-body armor.

After:

```
a young knight in full armor, standing in a field, cinematic long shot, full-body view, detailed armor, wide-angle perspective
```

Why it works: Anchors like "*full-body view*," "*long shot*," and "*detailed armor*" force the camera to pull back. Skin-pore details are removed so the model doesn't zoom into the face.

---

5. Technical Parameters

Even the strongest prompt can fall apart if your dials are set wrong.

When the model ignores a detail entirely:

- The detail is too far back in the prompt.
- Tokens are conflicting.
- CFG scale is too low.

Fix it by:

- Moving the detail forward.
- Boosting emphasis with double parentheses or weighting.
- Raising CFG scale — but not so high it turns stiff.

When the whole image looks fuzzy or cheap:

- Missing quality enhancers.
- Low render steps or default quality settings.

Fix it by:

- Adding quality cues: *"ultra detailed, 8K, sharp focus, cinematic."*
- Raising render steps or quality.
- Simplifying the prompt: contradictions blur the model's focus.

---

6. Overstuffing and Random Chaos

Finally, the classic rookie mistake: the grocery-list prompt.

If your input looks like:
"castle, sunset, dragon, rainbow, cinematic, epic, photorealistic, fantasy, painting, golden, 8K, HDR, masterpiece, fog, mist, thunder, glowing sword, glowing eyes, glowing moon..."
...then don't be surprised when the AI shrugs and gives you a glowing mess.

Fix it by:

- Grouping related details: all subject descriptors together, then scene, then lighting.
- Cutting adjectives that don't carry visual weight.
- Leaving space for atmosphere to breathe.

Example: Random Clutter

Before:

```
a fantasy scene, dragon, castle, rainbow, epic, glowing sword, glowing eyes, glowing moon, masterpiece, 8k, HDR
```

Result: Overstuffed chaos, everything glows.

After:

```
a fantasy scene of a dragon flying above a castle at dawn, cinematic wide shot, dramatic sky, detailed architecture, (glowing moon:1.3)
```

Why it works: Details are grouped by subject (dragon), setting (castle at dawn), and atmosphere (dramatic sky). Only one element is emphasized with glow, instead of everything competing.

Building a Prompt from Scratch

In previous chapters, we explored how AI-generated images are created, some of the technical mechanics behind them, and how to "speak the same language" as an image generator. Now you're staring at that empty prompt field. You know you want to make something, but the cursor just blinks at you in silence. The dreaded blank page problem, right?

There are many ways to break through it. In this chapter, we'll build one practical approach together. Think of it as a map, scattered reference points that help your imagination find its own footing.

It's no secret that AI art is often a reimagining of classic visual tropes. The tools have changed, but the visual archetypes are stubborn; they've been with us for centuries. By identifying them, you give your brain something like checkpoints—a network of associations that can make the creative process more predictable and less intimidating.

Below is a loose "map" of archetypes you'll recognize in most AI imagery, whether they appear alone or in combination.

I. Focus on Subject and Sensory Experience

Aesthetics – Harmony, form, visual pleasure.

> **Examples:** *An idealized portrait. A serene landscape. The perfectly polished design of a futuristic gadget.*
>
> The human eye is wired to seek symmetry, rhythm, and pleasing proportions. Even if the subject is ordinary, refined aesthetics tap into our brain's reward circuits.

Emotion – Capturing a precise feeling.

> **Examples:** *A victorious warrior on a battlefield (heroism), a room with a*

fireplace on a rainy day (coziness), a solitary figure under a streetlamp (melancholy), a comic panel (humor).
Images that radiate emotion invite instant empathy. Your viewer doesn't just "see" the picture, they feel it.

Color Reboot – Shock of the unexpected palette.

Examples: *Acid neon gradients, pop-art blocks of primary colors, retro computer palette with dithering, candy-colored clouds against a steel sky.*
A radical shift in color jolts the visual cortex. It's a reset button for the senses—suddenly, the familiar feels alien and exciting. Pop-art echoes tap into cultural memory, while hyper-saturated gradients give the scene a dreamlike, high-voltage intensity. The effect is both nostalgic and new, like rediscovering an old song remixed for today.

Sensory Impression – Evoking non-visual senses.

Examples: *A portrait embroidered in satin thread. A landscape made of molten sugar. An image that makes your skin tingle.*
This engages synesthesia in a soft way—making the brain blend senses that aren't actually in use, which makes the experience more memorable.

Nostalgia – Conjuring a lost time or vanished era.

Examples: *Vintage photography, retrofuturism, sepia tones, old film grain.*
Nostalgia is an emotional shortcut—your brain links certain aesthetics to personal or cultural memories, even if you never lived them.

II. Conceptual Shift and "What if..."

Conceptual Artifact – Combining the incompatible; ordinary objects in extraordinary roles.

Examples: *A glass teapot with a rose blooming inside, a city on the back of a whale, a piano sculpted from ice, a beetle with diamond wings.*

Juxtaposition surprises the brain—it forces the mind to resolve the gap between reality and imagination.

Distortion of Reality – Breaking physical laws and logic.

Examples: *A face woven from branches. Melting buildings. Cyberpunk glitches. Black lipstick dripping like ink.*

Defying physics disrupts expectation. Viewers linger because their brains keep trying to make sense of it.

Merging – Hybrids of contrasting entities.

Examples: *H.R. Giger's biomechanical forms, a dryad woman, a city-as-machine.*

Hybrids feel both familiar and alien—comforting and unsettling at the same time.

The Beautiful and the Uncanny – Balancing attraction and discomfort.

Examples: *A perfectly symmetrical face with no emotion. A drop of blood on a white dress. A silent room where something feels off.*

The uncanny valley effect activates primal caution, making the image magnetic in a strange way.

III. World-Building and Narrative

Terra Incognita – Exploring and discovering new worlds.

Examples: *Fantastical landscapes, alien vistas, a cozy lakeside cabin in impossible colors, a treehouse teetering over a cliff, a snowy viewing platform overlooking winter mountains.*

Humans are explorers at heart. Even a still image can trigger the instinct to "step inside."

Symbolism and Allegory – Images as metaphors.

Examples: *Shattered clocks as a symbol of lost time. A figure reaching toward*

light from darkness as a metaphor for hope.

Symbolism turns an image into a puzzle. Decoding it is part of the pleasure.

Narrative Moment – A frozen instant of story.

Examples: *"The calm before the storm." "The last glance." "An unexpected meeting."*

When viewers sense a story behind the frame, they mentally write it themselves, which deepens engagement.

Building from Archetype to Prompt

This map is far from complete, but even this short list gives you something crucial: the ability to ask, "What do I want to create—and what do I not want to create?" From there, you can add context, contrast, or reinforcing elements, and finally anchor everything with a unifying detail.

Let's walk through a concrete example.

You're spending a quiet evening with a cup of coffee and want to create something nostalgic. You sift through possible objects and settle on an old radio. It's visually unremarkable—so you give it a twist. The radio glows warmly from within. It sits on a wooden table.

For contrast, the setting is a dusty, abandoned cabin on a cold winter evening. The light from the window is pale and icy.

For the final touch—the anchor—you scatter several faded photographs on the table. In the radio's glow, they seem more vivid and alive, as if they're holding on to something you can't quite name. That's your personal artifact: a keeper of forgotten memories.

At this point, you already have enough to build a prompt:

Prompt example:

```
Grainy film still from an indie movie, old radio glowing warmly on a wooden table, scattered faded photographs on the table, photographs illuminated by radio become colored, abandoned wooden cabin, domestic details, snow outside the window, cold winter evening light, high-quality detail.
```

It's not as grand as "a glass spaceship landing in an alien city square," but if you see the finished image and your hand instinctively moves to save it, you've succeeded.

For the Curious: Why anchoring works
Your anchor—here, the photographs—acts as the emotional and narrative core. It gives the image a "point of gravity" that ties together the setting, mood, and subject. Without it, the scene might be pretty, but it risks feeling hollow.

Sometimes, when a scene feels lifeless, the problem isn't in the idea itself but in the lack of a center of gravity the eye can return to. Think of the strongest tools we've already explored—perspective, depth, light.

If your vast, fantastical forest feels static, place a winding path that disappears into its depths, giving the viewer the role of a traveler. If your tropical lagoon with a brigantine in the distance feels too postcard-like, shift the point of view—let the scene be framed by the hands of an explorer holding a weathered pirate map.

Repetition is another quiet force in composition. A row of lanterns fading into the mist, the rhythm of windows on a tall building, the curve of a spiral staircase, these patterns pull the gaze forward, suggesting movement even in a still frame. They lead the viewer deeper, giving the image an internal pulse.

Critique and extension: You can stretch this method by deliberately combining archetypes from different groups—say, blending "Narrative Moment" with "Distortion of Reality" for a surreal but emotionally grounded image. The key is balance: too many competing archetypes, and the image turns into noise.

Collaborating with the AI

Can you do all this **with** the model in a back-and-forth? Yes — but only if you give it a starting frame. Think of the AI as your co-director. You feed it the role and the archetypes, and it will bounce ideas back.

Example role definition (for any chatbot or AI assistant**):**

```
You are an elite creative partner. Your mission is
to help me develop rich, emotionally charged
concepts for AI-generated images. You suggest both
```

```
logical and unexpected ideas for maximum artistic
effect.
```

From there, you can run a four-step brainstorming process:

1. **Spark** — the initial seed (a taste, a color, a feeling).

2. **Vector** — choosing and mixing archetypes.

3. **World** — building the scene from setting to fine detail.

4. **Anchor and Contrast** — defining the focal point and adding one element that bends the rules.

At the end, you gather all the pieces into a master prompt — and you've not only broken through the blank page, you've built a workflow you can repeat.

From here, you can take the idea you've just developed and use it to co-create a rough draft prompt with the model itself. One simple way to do this is to give the AI a clear role and a precise format to follow.

Example role definition:

```
You are a personal assistant for crafting prompts
for AI image generation. Your primary task is to
take the user's request and transform it into a
detailed, high-quality prompt designed for AI image
generators.
```

Prompt format:

```
Art Style, by Artist/Photographer, Main Subject —
detailed description, pose, emotions, attire,
unique elements, Action/Interaction with
environment or objects, Setting/Background —
detailed and immersive, Additional Details, camera
```

```
angle, shot type, overall mood/atmosphere,
lighting, quality enhancers.
```

Composition: Finding Presence in the Frame

In the previous chapter, we touched briefly on composition. For many people, the word itself sounds like something belonging to high art, the kind of subject studied for years in academies. And in that sense, it is true. But for our purposes—when working with generated images—composition can be thought of in a simpler and more practical way. In fact, you may already understand it intuitively. You have seen it in films, in photographs, even in how you notice light and space around you. What we are doing here is giving a name to something you have probably known for a long time.

Let's take a somewhat elaborate, but useful example. Imagine you want to capture the feeling of gentle joy: a warm summer rain, with sunlight breaking through.
Picture a young woman in street clothes standing at a suburban bus stop, rain falling outside. I will show a sample image later, but for now let's focus on how the scene could be built.

Why a bus stop? Because surrounding lines and structures can help the eye settle on the subject. It works like a frame, shaping where attention goes. Yet if you hand this idea directly to an image generator, the result may be flat or overly centered. To make the scene more engaging, shift the perspective. Place the viewer inside the bus stop, slightly off to the side. Let the woman stand half-turned, as if she shares the same space with us. Instead of observing from afar, the viewer becomes present in the scene.

This is where one of the most common ideas of composition appears: the rule of thirds. Often, images feel more balanced when the subject is not placed directly in the center. Some generators respond to the phrase "rule of thirds," but not always. A more reliable way is to enrich the scene with other elements.

For instance, in front of the bus stop, a cracked road stretching into the distance, and beyond it, a suburban forest. These details guide the image generator to spread attention across the frame and give the composition more weight.

The road itself also plays a role. From our chosen angle, it recedes into space, drawing the eye deeper. Think of a child's toy block: seen from the front, it looks like a square, but when you tilt it, the edges create a sense of volume. Perspective works the same way in images. Lines and vanishing points give a flat surface the illusion of depth.

Now return to the woman. How could her emotion be conveyed? Perhaps a faint smile, her head tilted slightly upward, eyes closed. Sunlight filtering through clouds, soft backlight against the rain. A few drops on her face make the scene tangible. Around her, small contrasts add life: rain dripping from the roof, graffiti and torn posters on the wall, nature bright and fresh in the background. If you want to add nostalgia, you could choose a grainy film effect, as if the image were shot on an old camera.

A possible prompt could look like this:

```
Analog horizontal photo, grainy film effect,
photorealism, a young blonde woman, with closed
eyes, head raised slightly to soft beams of
sunlight, light backlight, smiling faintly, wearing
street clothes, raindrops on her face, stands
half-turned inside a suburban bus stop, light
summer rain falling outside, drops falling from the
roof, the wall covered with advertisements and
graffiti, a cracked road stretching forward,
suburban forest in the background, side shot from
inside the bus stop, highly detailed, sharp focus.
```

The larger question is how to treat composition not as a rigid list of rules, but as a set of tools for storytelling. The simplest way is to approach it with curiosity. Learn from images that already move you. When you are watching a film or series, pause on a frame that feels striking and ask yourself: what holds my attention here? Is it the light, the lines, the angle, the way the space feels? These observations, over time, become your own working vocabulary.

Case Studies:

The Three-Panel Comic

The challenge here is **consistency**: keeping the same characters across panels while shifting action, angle, and mood. If you can guide the AI through this, you can generate not only single images but also micro-stories.

We'll start with a simple gag format: two panels on top for setup, one panel at the bottom for the punchline.
Story:

1. An old wizard reading a book.
2. The wizard adds powder into a flask.
3. A frog in the wizard's hat sits on the table.

Without visual glue, the panels won't feel like one comic. You need:

- **Shared style anchor:** "comic style illustration, bold outlines, flat colors" (or manga linework, watercolor, etc.)
- **Consistent character description:** reuse the same wizard descriptors every time.
- **Panel framing cues:** "top left panel," "top right panel," "bottom wide panel." Even if the AI doesn't literally create panels, it guides framing.

Possible enhancements

- Add **speech bubbles** in text form: "*Wizard muttering: 'This will work…'*"
- Use **expressions**: curious → excited → comic despair.
- Experiment with **layouts**: two stacked panels, or vertical strip.

Prompt example:

```
Three-panel comic, comic style illustration, bold outlines.
```

```
Top left panel: old wizard with long beard and pointy
hat, sitting at a desk, reading a book by candlelight.
Top right panel: the same wizard in his study,
sprinkling powder into a flask, magical sparks.
Bottom wide panel: wizard's desk, his hat on the table
with a frog inside it, glowing comically.
```

The Fashion Lookbook

Fashion is less about fantasy excess and more about **control and cohesion**. The goal is to make a set of images feel like they belong to the same photoshoot: the same model, consistent lighting, coherent styling.

Key Challenges:

Consistent model identity: Use a stable description of the person: *"tall woman with long red hair, oval face, slender build"*. Repeat it across all prompts.

Pose control: Define shot type: *"runway full-body shot"*, *"editorial mid-shot"*, *"close-up beauty portrait"*. Keep it simple to avoid anatomy errors.

Clothing details: Name fabrics and cuts: *"silk evening gown with flowing folds"*, *"denim jacket with metallic buttons"*. Pair with lighting cues to enhance texture.

Lighting & setting: Anchor the shoot: *"studio fashion shoot, softbox lighting, seamless white background"*. Or: *"outdoor editorial, golden hour sunlight, blurred city background"*.

Series cohesion: Keep the same background + lighting across prompts. Use identical style anchors: *"high-fashion editorial, Vogue style, minimal styling"*.

Prompt example (single lookbook shot):

```
High-fashion editorial lookbook photo, tall woman with long red hair and slender build, wearing a flowing silk gown with reflective texture, runway pose, full-body shot, studio fashion shoot, softbox lighting, seamless grey background, minimal styling, Vogue magazine aesthetic, photorealistic, ultra sharp, high resolution.
```

The Product Render

Product images are a different beast. They aren't about atmosphere or storytelling: they're about **clarity, material accuracy, and commercial polish**. The viewer should feel they're looking at a studio photo made for an online store.

Key Challenges:

Background control: Avoid clutter. Use *"white seamless background"* or *"isolated on transparent background"*. If lifestyle context is needed: *"product on wooden table in soft daylight"*.

Material realism: Surfaces matter: glass, plastic, brushed metal, leather. Call them explicitly: *"glossy ceramic mug with matte handle"*.

Lighting: E-commerce = clean and even. Use *"studio product photography, softbox lighting, no harsh shadows"*.

Framing: Decide early: *"front view, side view, angled 45° product shot"*. Add *"centered composition"* to keep it symmetrical.

Quality anchors: *"ultra sharp, photorealistic, 8k, product render, commercial photography"*. Negative prompts: *"no text, no watermark, no reflections of camera equipment"*.

Prompt example (single product shot):

> `Photorealistic studio product render of a sleek black wireless headphone set, isolated on seamless white background, front view, centered composition, studio product photography, softbox lighting, sharp focus, commercial e-commerce style, ultra detailed, high resolution.`
> Negative: `text, watermark, people.`

Prompt example (lifestyle variation):

> `Product photo of a ceramic coffee mug with matte black handle, placed on a wooden kitchen table, soft daylight coming from the window, shallow depth of field, photorealistic, commercial photography, ultra sharp.`
> Negative: `logos, watermark.`

The key to product work is **restraint**: fewer adjectives, no chaotic atmosphere, maximum focus on accuracy and lighting. Where fantasy prompts thrive on excess, product prompts thrive on discipline.

Creating the Epic Shot

From wish-list chaos to a directed, repeatable prompt

The raw idea:

> `A beautiful, detailed Japanese temple, pagoda, wooden bridge, a dense, ethereal forest, sunbeams filtering`

```
through a canopy, lush green mossy stream with stones,
glowing butterflies and floating motes of light,
flowering bushes, wide shot, full-frame, depth of
field, detailed foreground and background, rule of
thirds, smooth textures, intricate details on the
pagoda and bridge, glowing effects, particles of light,
water reflections, magical sunbeams, warm golden glow
from temple lanterns, soft ethereal light, lush green,
teal and blue tones, with accents of pink and purple,
high saturation, high contrast, fantasy illustration,
digital painting, highly detailed, photorealistic,
Japanese aesthetic, whimsical art, 8k, ultra sharp,
cinematic lighting, masterpiece.
```

Lovely. Overstuffed. Full of intention. Now we'll pare, restructure, and preserve that intention. You're right: chopping away words without losing the soul of the idea is the hardest part. The trick isn't to "use more words" or to "be shorter", it's to give every word a job. In this chapter we'll turn our beautiful, noisy wish-list into a clean, modular prompt that still reads (and feels) epic.

Below I'll show the exact process step-by-step, why each edit matters, and multiple final variants you can copy. Read it as a recipe and a small workshop.

Step 1. Extract the skeleton (what must be present)

Ask: *If everything else dropped out, what are the non-negotiables?*

- **Subject:** Japanese temple / pagoda / wooden bridge
- **Setting:** dense forest, mossy stream
- **Mood:** ethereal, magical, warm lantern glow

Skeleton prompt:

```
Japanese temple pagoda with wooden bridge over a mossy stream in a dense ethereal forest
```

Why: This keeps the image's backbone. Anything else is detail and texture.

Step 2. Add atmosphere (the emotional layer)

Now add the things that give the scene its soul: light behavior, particles, small living things.

- sunbeams filtering through canopy
- glowing butterflies and floating motes
- warm golden lantern glow, soft ethereal light
- water reflections, particles of light

Atmosphere block:

```
sunbeams filtering through the canopy, glowing butterflies and floating motes of light, warm golden lantern glow, soft ethereal light, water reflections, particles of light
```

Why: These phrases set mood and direct how the model treats illumination and small effects.

Step 3. Camera & composition (where the viewer sits)

This decides scale and focus—full-frame, wide shot, DOF, rule of thirds.

- wide shot, full-frame, depth of field, detailed foreground and background, rule of thirds

Camera block:

`wide cinematic shot, full-frame, strong depth of field, detailed foreground and background, rule of thirds composition`

Why: Prevents accidental close-ups and anchors how much environment the model must show.

Step 4. Style & finish (how it should look)

Choose the single style you want (photorealistic *or* painterly). If you try to demand both strongly, the model wavers.

Two clear options:

- **Photorealistic**: `photorealistic, cinematic lighting, ultra detailed, 8k, ultra sharp`
- **Painterly/fantasy**: `digital painting, whimsical fantasy illustration, highly detailed, painterly strokes.`

Pick one and tailor color notes:

Color palette: `lush greens, teal and blue tones with pink and purple accents, warm golden highlights, high saturation, high contrast`

Why: A coherent style block prevents the generator from merging incompatible aesthetics.

Step 5. Negative constraints (what *not* to do)

Short negatives prevent common drift: `humans, modern buildings, text, watermark, lowres`

Why: Reduces surprises like accidental tourists or signage.

Step 6. Assemble (modular prompt)

Put the blocks in priority order: Subject → Atmosphere → Camera → Style → Color → Quality → Negatives.

Final photorealistic prompt (clean, modular):

> `Japanese temple pagoda with a wooden bridge over a mossy stream in a dense ethereal forest, sunbeams filtering through the canopy, glowing butterflies and floating motes of light, warm golden lantern glow, soft ethereal light, water reflections, particles of light, wide cinematic shot, full-frame, strong depth of field, detailed foreground and background, rule of thirds composition, photorealistic, cinematic lighting, color palette: lush greens, teal and blue tones with pink and purple accents, warm golden highlights, high saturation, high contrast, ultra detailed, ultra sharp, 8k.`
>
> Negative: `humans, modern buildings, text, watermark, lowres.`

Why this works:

- The subject is first, so it dominates.
- The atmosphere is explicit (so lighting and small particle effects appear naturally).
- Camera and composition stop accidental closeups.
- Style is coherent (photorealistic), so “digital painting vs photo” conflict is avoided.
- Negatives block common garbage.

Variants: same idea, different finish

Painterly fantasy variant:

> `Japanese temple pagoda with wooden bridge over a mossy stream in a dense ethereal forest, magical sunbeams through the canopy, glowing butterflies and motes, lanterns emitting warm gold, flowing water reflections, wide cinematic shot, full-frame, layered depth, detailed foreground and background, rule of thirds, whimsical fantasy digital painting, painterly brushstrokes, lush greens and teal palette with pink and purple accents, soft ethereal glow, highly detailed, masterpiece.`
>
> Negative: `humans, modern structures, watermark.`

Mobile/vertical crop (for phone wallpapers)

> `Japanese pagoda and wooden bridge over mossy stream in a tall ethereal forest, vertical composition, strong depth of field, glowing butterflies, sunbeams filtered through canopy, warm lantern glow, photorealistic, cinematic lighting, ultra detailed, 9:16 aspect ratio, rich greens and teal with pink accents.`
>
> Negative: `people, text, lowres.`

Small surgical edits that preserve the idea

These show how removing words actually *sharpens* the image:

- Remove generic buzzwords like *"masterpiece"* unless you want artistic drift. They often add noise.
- Replace vague adjectives with concrete ones: *"soft ethereal light"* > *"sunbeams filtering through canopy + soft rim light from lanterns."*
- Drop redundant repeats: you don't need both *"magical sunbeams"* and *"sunbeams filtering"* — pick one precise phrase.

Troubleshooting quick fixes (if results still miss the mark)

Symptom: Model gives a close-up instead of wide shot.
Fix: Reinforce distance-only anchors: `full-body / full-structure / wide shot / visible landscape elements (mountains/large tree)` and move them earlier.

Symptom: Everything glows or everything is over-saturated.
Fix: Remove global words like *"high saturation"* or reduce emphasis; put `glow` on a single element: `(butterflies glow:1.2)` rather than "glowing effects" for everything.

Symptom: Too painterly when you asked for photorealism.
Fix: Move `photorealistic` to the *very front* of style block and add `photo-real photograph, sharp detail, realistic textures.`

Symptom: Model adds people / temple looks modern.
Fix: Add `humans, modern structures` to negatives and add `traditional Japanese architecture, Edo period details` to subject.

Mini exercise (apply this once)

1. Take your messy prompt. Extract skeleton (3 lines).
2. Write a one-line atmosphere block (3–6 phrases).
3. Add camera/composition (1–2 phrases).

4. Pick one style and write 2 style phrases.
5. Build the final prompt by concatenating blocks in priority order.
6. Run it, inspect the output, and adjust only one block (lighting or style) for iteration.

Final note: You are not “losing” parts of your idea when you cut; you are moving them to the right place. A well-structured prompt preserves every intention by giving it hierarchy and a role. The temple, the stream, the butterflies, the glow, all survive, but now each has a stage direction and a priority.

Beyond Generation: Working With Images You Already Have

Sometimes the hard part isn't making a picture — it's keeping it consistent, telling its story, or even remembering how you made it in the first place.

Keeping Characters Consistent Across a Comic Series

You want to create a comic series. But every time you hit Generate, your "wise old mage" looks like a different person.

Without a fixed **seed**, this is nearly impossible to avoid completely. But you can make it easier by asking AI to **analyze your existing image** and turn it into a reusable prompt.

Use instructions like:

```
Describe the character in this image with maximum detail.
Follow these rules:
1. Analyze every facial feature in depth.
2. Include anatomical and stylistic details:
face shape, eye color/shape, brows, nose, lips, jawline,
chin, skin texture, age markers, hair color/style/length,
distinctive marks (moles, scars).
3. Be objective — no opinions like "beautiful face."
4. Write the result as a single prompt line suitable for
image generators.
5. Use words that best help replicate this face.
6. If multiple images are provided, merge information from
all angles.
```

With this, you'll get a prompt like:

```
Old male mage, oval face, deep wrinkles around eyes,
hooked nose, thick grey eyebrows, long silver beard,
narrow pale-blue eyes, high forehead, thin lips,
shoulder-length white hair, weathered skin with freckles
and age spots
```

Discovering the Story Behind an Image

Sometimes you create an image that hits you right in the heart, but you don't know why. Maybe you'd love to know the story behind it — or write it yourself.

Ask a chatbot to play **creative director and storyteller**:

```
You are a creative director and storyteller.
Your task: deeply analyze the provided image,
describe the world it belongs to, and invent a short story
that could take place there or with these characters.
Principles:
- Analyze every detail: architecture, clothing style,
environment, lighting.
- Build a coherent world:
 Genre/Era, Environment, Rules of physics or magic.
- Define society/creatures: who lives here, their norms or
abilities.
- Highlight unique features: what makes this world
distinct.
- Write a 3–5 sentence story fragment that fits naturally
into this world.
```

This will give you something like:

```
Analysis:
A rusted tower rises from a swamp, glowing symbols etched on its stone. Two cloaked figures approach in silence under pale mist.
World:
Fantasy realm, drowned cities reclaimed by nature. Ancient runes suggest forgotten magic tied to water and decay. Small nomadic clans survive among marshlands, scavenging artifacts of fallen empires.
Story:
At dawn, the mage whispered to his apprentice: “If we open these gates, the old gods will finally speak again.” The marsh water rippled, glowing softly, as if the tower itself had been waiting for centuries to hear those words.
```

Lost Prompts? How to Recreate Them

You’ve generated hundreds of images, thinking you’d never need the prompts again. Now you wish you had them.

There’s no perfect way to recover them, but you can ask AI to **reverse-engineer** the prompt:

```
You are an expert in recreating prompts for images.
Instructions:
- Identify the main subject: who/what is depicted.
- Determine art style or medium: photo, painting, sketch, anime, 3D render.
- Analyze composition and framing: focus, arrangement, camera angle.
- Describe lighting: source, quality, atmosphere, time of
```

```
day.
- Note color palette and mood.
- Extract unique details: textures, clothing, objects,
background elements.
- Infer negative prompts: what should be avoided.
- Use this structure:

Stylistics/Art style: …
(Optionally: by [Artist])
Subject: …
Pose/Clothing: …
Actions/Details: …
Setting/Environment: …
Lighting/Background: …
Color palette/Mood: …
Camera angle/Composition: …
Additional keywords: …
Quality enhancers: …
```

You'll get a reusable approximation like:

```
Digital painting, epic fantasy landscape,heroic knight
with flowing crimson cape and shining silver armor,
standing on a mossy cliff edge overlooking stormy ocean
waves, dramatic golden hour backlighting, cinematic
wide-angle shot, mist swirling around rocks,
hyper-detailed, 8K, masterpiece
```

AI isn't just for generating images. It can help you **document, recreate, and even narrate** your existing art, keeping your worlds consistent and alive.

Bringing Your Still Image to Life

You've generated an image that feels perfect... and now you wish it could move. Maybe the wizard could brush dust off the map, or a candle could flicker in the corner. But how do you even start turning it into video?

Here's a simple helper prompt for any chatbot or AI assistant:

```
You will receive a prompt describing an image or its main elements.
Convert it into a structured prompt for video generation.
Follow this exact format:
(Overall scene & visuals — expanded environment and style),
(Main subject — actions, gestures, movements),
(Camera — shot type, camera motion, angles),
(Duration & pacing — how long and how fast),
(Lighting & atmosphere — dynamic changes and effects),
(Optional: sound design & music),
(Optional: transitions & editing style).
Add creative details if necessary to make the video feel cinematic and alive.
```

Example:
Image prompt:

```
Old wizard sitting at a wooden table in a candlelit room, studying an ancient map covered in mystical symbols, fantasy illustration, warm glowing atmosphere
```

Video prompt output:

```
(A dimly lit stone chamber with shelves of dusty tomes and flickering candles, warm fantasy illustration style,
```

```
medieval interior, glowing runes on the walls), (A wise
old wizard with a long silver beard slowly tracing paths
on an ancient map, adjusting his round spectacles, his
robe's sleeves swaying slightly), (Camera starts with a
close-up of the map, then slowly pans upward to reveal the
wizard's focused face, final shot in a soft
over-the-shoulder angle), (8-second clip with gentle
pacing, subtle movements of the wizard's hands and
shifting candlelight), (Warm golden lighting with soft
shadows, atmospheric dust particles drifting in the air, a
faint magical aura glowing from the map),
(Background sound of crackling fire and turning parchment,
soft mystical music playing), (Smooth cinematic
transitions, slight vignette effect).
```

Note: Video generators evolve quickly — new tools appear almost every month. Each one has its own rules: prompt syntax, supported effects, maximum duration, even how it handles camera movement.

Before using the video prompt from this guide, **check the parameters of your chosen generator** and adjust the structure if needed. A few words changed can make the difference between a static slideshow and a cinematic, flowing scene.

Beyond the Last Page

Remember that feeling it all began with?
The sense of tossing words into a mysterious black box and hoping for a miracle?

Now, it's different.

You're no longer standing at a locked door. The keys are in your hand. You know how to shape the light, frame the shot, tell the story, even breathe life into your worlds. You've learned to speak to the machine in the language of creativity.

This book is only the start of your conversation with it. The real magic begins when you close this page and open that empty prompt field, this time with the confidence of a director, not the tremor of a gambler.

The world is waiting for your stories.
Now you know how to tell them.

And if you ever forget a term — flip to the Quick Reference Cheat Sheet at the back. Keep it close, keep experimenting.

Appendix A: Quick Reference Cheat Sheet

Core Art Styles & Movements

Term	Meaning	When to Use
Photorealistic	A style that aims to reproduce an image as realistically as possible, indistinguishable from a photograph.	For creating lifelike portraits, scenes, and objects where realism is the primary goal.
Impressionism	Characterized by small, thin brushstrokes, an emphasis on the accurate depiction of light, and open composition.	For dreamy, soft-focus scenes, especially landscapes and candid moments, with a focus on light and color.
Surrealism	Features bizarre, dreamlike, and illogical scenes. Think of artists like Salvador Dalí and René Magritte.	To create strange, fantastical, and thought-provoking images that defy logic.
Abstract	Art that does not attempt to represent external reality, but seeks to achieve its effect using shapes, forms, colors, and textures.	For creating non-representational images, backgrounds, textures, or conceptual pieces.
Minimalism	A style characterized by extreme simplicity of form and a deliberate lack of expressive content.	For clean, simple compositions, focusing on basic shapes, negative space, and a limited color palette.
Pop Art	Known for its use of bold colors, imagery from popular culture (e.g., comics, ads), and a sense of irony.	For creating vibrant, graphic, and modern images with a retro or commercial feel.
Cyberpunk	A subgenre of science fiction featuring a high-tech, futuristic urban setting, often with a dark, dystopian mood, neon lights, and cybernetics.	For creating futuristic cityscapes, characters with cybernetic enhancements, and moody, neon-lit scenes.
Steampunk	A style inspired by 19th-century industrial steam-powered machinery. Features gears, clockwork, and Victorian aesthetics.	For creating retro-futuristic technology, fashion, and environments with a Victorian-era mechanical twist.

Digital & Modern Aesthetics

Term	Meaning	When to Use
Concept Art	Art used to convey an idea for use in films, video games, animation, or other media before it is put into the final product.	To generate character designs, environments, and props that look like they belong in the pre-production of a game or movie.
Digital Art / Digital Painting	Art made using digital tools, which can mimic traditional painting or have a distinctly clean, digital look.	A broad term for creating images with a modern, digitally-created feel, from painterly styles to crisp vector art.
Vaporwave	An aesthetic characterized by 1980s/90s internet imagery, retro computer graphics, Greco-Roman statues, and a pastel color palette.	For nostalgic, surreal, and retro-tech images with a distinct pink, purple, and cyan color scheme.
Low Poly	A 3D modeling style that uses a small number of polygons, resulting in a blocky, faceted appearance.	To create a stylized, retro video game look or a minimalist, geometric representation of a subject.
Pixel Art	Digital art created at the level of individual pixels, characteristic of early video games.	For creating retro 8-bit or 16-bit video game-style characters and scenes.

Professional Photography Styles

Term	Meaning	When to Use
Portrait Photography	Photography of a person or group of people that captures the personality of the subject by using effective lighting, backdrops, and poses.	For generating classic, professional-looking photos of people, focusing on the individual.
Landscape Photography	Captures the presence of nature and is often grand in scale. The art is in capturing the perfect moment of light and weather.	For creating vast, beautiful, and dramatic images of natural scenery like mountains, deserts, and coastlines.
Macro Photography	Extreme close-up photography, usually of very	To capture intricate details of small objects like insects,

	small subjects, in which the size of the subject in the photograph is greater than life-size.	flowers, water droplets, or textures.
Product Photography	A branch of commercial photography that is about accurately and attractively representing a product. Often has a clean background.	For creating clean, commercial-style images of objects, often isolated on a white or simple background.
Fashion Photography	A genre devoted to displaying clothing and other fashion items. Often enhanced by exotic locations or accessories.	For generating stylized images of models wearing specific clothing, focusing on glamour, style, and mood.
Architectural Photography	The photography of buildings and similar structures that is both aesthetically pleasing and accurate in its representations.	For capturing the exterior or interior of buildings, emphasizing lines, geometry, and form.
Drone Photography / Aerial View	Photography taken from an aircraft or other flying object. Provides a top-down or high-angle perspective.	For creating sweeping, bird's-eye views of landscapes, cityscapes, or events.
Wildlife Photography	Concerned with documenting various forms of wildlife in their natural habitat.	To generate realistic images of animals in nature, often with a focus on capturing a specific action or behavior.

Fundamental Lighting Concepts

Term	Meaning	When to Use
Natural Light	Light that comes from the sun or moon.	For outdoor scenes or interiors lit by windows, creating a realistic and organic feel.
Soft Light	Diffused light that creates soft, subtle shadows with gradual transitions.	For portraits, beauty shots, and scenes that require a gentle, flattering, or dreamy mood.
Hard Light	Direct, focused light that creates sharp, well-defined shadows and high contrast.	To create drama, tension, or a sense of mystery; emphasizes texture and form.
High Key Lighting	Bright, even lighting with minimal shadows, resulting	For an optimistic, cheerful, clean, or ethereal atmosphere.

	in a low-contrast image.	Common in comedies and product shots.
Low Key Lighting	Dark, shadowy lighting with high contrast, where much of the scene is in shadow.	For dramatic, mysterious, suspenseful, or moody scenes. Classic for film noir and thrillers.
Ambient Light	The general, non-direct light already present in a scene before any specific lights are added.	To establish the base level of illumination and the overall mood of an environment.

Directional & Stylistic Lighting

Term	Meaning	When to Use
Backlight / Rim Light	Light placed behind the subject, creating a bright outline or halo effect.	To separate the subject from the background, adding depth and a dramatic or angelic feel.
Front Light	Light source is directly in front of the subject, minimizing shadows.	For a flat, clear, and revealing look. Often used in passport photos or basic vlogging.
Side Light	Light that hits the subject from the side, creating strong highlights and shadows.	To reveal texture, create drama, and add a three-dimensional quality to the subject.
Volumetric Lighting	Makes light beams visible as they travel through the air (e.g., sunbeams through dust).	To add atmosphere, depth, and a sense of magic or divinity. Think of light rays in a forest.
Golden Hour	The period shortly after sunrise or before sunset, known for its warm, soft, reddish light.	For romantic, nostalgic, and beautiful outdoor scenes. Highly flattering for portraits.
Blue Hour	The period just before sunrise or after sunset when the light is a cool, deep blue.	For serene, tranquil, or moody cityscapes and landscapes with a calm, quiet feeling.
Cinematic Lighting	A general term for dramatic, stylized lighting that looks like it's from a movie.	When you want a scene with high production value, mood, and a narrative quality.

Advanced & Niche Lighting

Term	Meaning	When to Use
Rembrandt Lighting	A specific portrait technique creating a small, inverted triangle of light on the cheek opposite the light source.	For classic, dramatic, and moody portraits that feel timeless and artistic.
Split Lighting	Lights exactly half of the subject's face while keeping the other half in shadow.	To create a dramatic, conflicted, or mysterious look for a character.
Caustic Lighting	The patterns of light created when light reflects or refracts off a surface (e.g., light through water).	For underwater scenes, or to add visual complexity and realism to scenes with water or glass.
Chiaroscuro	An Italian term for the use of extreme contrasts between light and dark in art.	For highly dramatic, intense, and emotional scenes, mimicking Renaissance or Baroque painting styles.
Gobo (Go-Between Object)	A stencil or pattern placed in front of a light source to project a specific shape or texture.	To simulate light coming through a window, leaves, or to add interesting background patterns.
Neon Lighting	Light produced by neon signs, often in vibrant, saturated colors like pink, blue, and green.	For cyberpunk, futuristic, urban nightlife, or retro 80s aesthetics.
Bioluminescent	Light produced by living organisms, like fireflies or deep-sea creatures.	For fantasy, sci-fi, or nature scenes that need a magical, glowing, or otherworldly quality.

Common Camera Shots

Term	Meaning	When to Use
Close-up	Focuses on a single subject or object, typically the face.	To emphasize facial expressions, emotions, or intricate details.
Medium Shot	Shows the subject from the waist up.	To display body language and facial expressions together, good for conversations.

Full Shot / Long Shot	Captures the entire subject from head to toe.	To show the subject's entire body and their immediate surroundings.
Wide Shot / Establishing Shot	Shows the entire scene or a large part of it.	To establish the setting, location, and atmosphere of the scene.

Variations and Specialized Shots

Term	Meaning	When to Use
Extreme Close-up	Frames a specific detail of a subject, like the eyes or mouth.	To create a dramatic, intimate, or intense feeling.
Choker Shot	A shot framed from just above the eyebrows to the mouth.	For a tight, intense focus on a character's emotions.
Medium Close-up	Frames a subject from the chest or shoulders up.	A classic shot for interviews and dialogue-heavy scenes.
Cowboy Shot / American Shot	A variation of a medium shot, framing from mid-thighs up.	To show a character's confidence and readiness, often showing hands near holsters.
Knee Shot	Frames the subject from the knees up.	A good balance between a full shot and a medium shot, showing action and some environment.
Two-Shot / Group Shot	A shot with two or more subjects in the frame.	To show the relationship and interaction between characters.

Camera Angles and Perspectives

Term	Meaning	When to Use
Eye-Level	The camera is at the same height as the subject's eyes.	To create a neutral, direct, and relatable perspective.
High-Angle Shot	The camera looks down on the subject.	To make the subject appear smaller, vulnerable, or insignificant.
Low-Angle Shot	The camera looks up at the subject.	To make the subject look powerful, heroic, or intimidating.

Bird's-Eye View / Top-Down Shot	The camera is directly overhead, looking straight down.	To show a pattern, an overview of a location, or a sense of surveillance.
Worm's-Eye View	A very low-angle shot, as if from the perspective of a worm.	To create an extreme sense of scale, making the subject look monumental.
Dutch Angle / Canted Angle	The camera is tilted, causing the horizon to be slanted.	To create a sense of unease, disorientation, or tension.
Over-the-Shoulder Shot	Filmed from behind a person, looking over their shoulder at the subject.	To show a conversation from one character's point of view.
Point of View (POV)	The shot shows what a character is looking at.	To immerse the viewer directly into the character's experience.

Core Quality & Detail Enhancers

Term	Meaning	When to Use
High Quality / 4K / 8K	A general instruction to produce a high-resolution, clear, and sharp image, free of artifacts. 4K/8K specifically reference screen resolutions.	As a standard, all-purpose booster for almost any prompt to signal a desire for a polished, high-resolution final image.
Highly Detailed	Asks the AI to focus on rendering intricate details, textures, and small elements within the scene.	For complex subjects like ornate architecture, detailed character designs, intricate machinery, or textured landscapes.
Sharp Focus / In Focus	Instructs the AI to ensure the main subject is crisp and clear, without blurriness. It mimics a perfect camera focus.	Essential for portraits, macro shots, and any image where the clarity of the subject is the top priority.
Intricate Details	A more intense version of "highly detailed," pushing the AI to generate extremely fine and complex patterns and textures.	For close-ups of jewelry, fantasy armor, complex fabric patterns (like paisley or damask), or hyper-realistic insect wings.
Masterpiece / Best Quality	Subjective terms that encourage the AI to generate an image in its highest possible quality, with better composition and lighting.	Use as a general "try your best" command to improve the overall artistic merit and technical execution of an image.

Technical & Artistic Enhancers

Term	Meaning	When to Use
UHD (Ultra High Definition)	Similar to 4K/8K, this is a technical term asking for an image with very high resolution and pixel density.	A good alternative to 4K/8K. Useful for generating images that will be viewed on large screens or need to be cropped.
Ray Tracing	A rendering technique that produces realistic lighting effects by simulating the physical behavior of light, creating realistic reflections, shadows, and refractions.	For scenes with reflective surfaces like metal, glass, or water, or for achieving hyper-realistic and cinematic lighting.
Unreal Engine / Octane Render	Invokes the aesthetic of high-end 3D rendering software, known for photorealistic lighting, materials, and cinematic quality.	To create images that look like stills from a modern, high-budget video game or a CGI film.
Flickr / 500px	Referencing high-quality photography websites. The AI associates these with well-composed, high-resolution, and often professionally edited photographs.	A great way to subtly ask for a "professional photograph" aesthetic without being overly technical.
Film Grain	Adds a subtle texture to the image that mimics the look of traditional photographic film, which can add a sense of realism or nostalgia.	To give digital images a more organic, vintage, or cinematic feel and to avoid an overly clean, "digital" look.
Chromatic Aberration	A subtle visual effect that mimics a real camera lens imperfection, where colors are slightly misaligned, especially at the edges.	To add a layer of realism or a retro/lo-fi photographic effect to an image.

Appendix B: Sample Prompts for Quick Start

These prompts are structured for clarity. Each follows the same logic:
Style – Subject – Lighting – Mood – Camera/Angle – Quality
Use them as starting points and adapt them to your own creative goals.

I. Portrait

Beginner Prompt:

```
Photorealistic style, a young woman smiling softly, natural daylight from a window, warm and friendly mood, close-up shot at eye level, sharp focus and high detail
```

Advanced Prompt:

```
Cinematic photography style, a man with short hair wearing a casual jacket, dramatic side lighting, thoughtful and introspective mood, medium close-up shot, soft background blur, high-resolution and realistic textures
```

II. Landscape

Beginner Prompt:

```
Photorealistic style, a calm mountain lake at sunrise, soft golden light, peaceful and serene mood, wide establishing shot, high detail and clarity
```

Advanced Prompt:

```
Concept art style, a desert canyon with towering cliffs and glowing sky, strong directional sunlight casting long shadows, epic and dramatic mood, wide-angle shot from ground level, highly detailed with textured rocks and atmospheric haze
```

III. Product / Object

Beginner Prompt:

```
Product photography style, a ceramic coffee mug on a wooden table, bright natural light, clean and simple mood, centered composition at eye level, sharp focus with realistic textures
```

Advanced Prompt:

```
Studio photography style, a vintage wristwatch on a dark velvet surface, soft diffused spotlight, elegant and luxurious mood, close-up shot with shallow depth of field, intricate details and high-resolution rendering
```

IV. Fantasy / Concept Art

Beginner Prompt:

```
Fantasy illustration style, a glowing sword stuck in a stone, soft ambient light with mist, mysterious and magical mood, medium shot framed at waist height, crisp details with luminous effects
```

Advanced Prompt:

```
Concept art style, a massive ancient tree with glowing runes on its bark, beams of light breaking through fog, mystical and awe-inspiring mood, wide establishing shot from below, ultra-detailed textures and atmospheric lighting
```

V. Cinematic Scene

Beginner Prompt:

```
Cinematic style, a person walking through a rainy street at night, illuminated by neon lights, moody and urban atmosphere, medium full-body shot, high detail and natural reflections on wet pavement
```

Advanced Prompt:

```
Cinematic style, two characters talking in a dimly lit café, warm candlelight with soft shadows, intimate and emotional mood, over-the-shoulder shot, realistic textures and high-resolution detail
```

Bonus Prompt: The Book Cover

Even the cover of this book began as a prompt, here it is, in case you want to try your own version.

```
Book cover illustration for "From Sketch to Shot: Writing Prompts That Work", split-screen concept: a pencil in the foreground draws a gently curved line that splits the left and right sides.
Left side: a monochrome pencil sketch of an artist's workspace — desk, computer, chair, window.
Through the window, a simple line drawing of a tree and landscape.
Right side: a vivid, painterly colorful landscape with bright sunlight, green trees, and a small river.
Soft, textured paper background, clean composition, harmonious color transition, highly detailed.
```