# Craft better video, one lesson at a time

## Full course PDF

Learn framing, camera, light, prompts and delivery through short lessons and practice.

# 1. What generation actually is

## What does forward process mean in Diffusion models: generation by denoising?

The fixed schedule of adding noise to a real image, used to manufacture training examples.

https://shotcraft.video/learn/concepts/diffusion-basics/

## What does reverse process mean in Diffusion models: generation by denoising?

The learned sequence of denoising steps that turns pure noise into a sample.

https://shotcraft.video/learn/concepts/diffusion-basics/

## What does noise prediction mean in Diffusion models: generation by denoising?

The model's actual output at each step: an estimate of the noise in the current input, not the finished image.

https://shotcraft.video/learn/concepts/diffusion-basics/

## What does score mean in Diffusion models: generation by denoising?

The direction in which a noisy image should move to become more like real data; the quantity score-based formulations estimate.

https://shotcraft.video/learn/concepts/diffusion-basics/

## What does sample mean in Diffusion models: generation by denoising?

One generation drawn from the model, which differs run to run because the starting noise differs.

https://shotcraft.video/learn/concepts/diffusion-basics/

## What does latent mean in Latent space: working at proxy scale?

A compressed, learned representation of an image or clip that the diffusion model operates on instead of pixels.

https://shotcraft.video/learn/concepts/latent-space/

## What does latent diffusion mean in Latent space: working at proxy scale?

Running the denoising process inside a pretrained autoencoder's compressed space rather than in pixel space.

https://shotcraft.video/learn/concepts/latent-space/

## What does compression ratio mean in Latent space: working at proxy scale?

How many pixels of input each unit of latent stands for, trading detail for speed and length.

https://shotcraft.video/learn/concepts/latent-space/

## What does token mean in Latent space: working at proxy scale?

One position the model attends over; in image and video models, one patch of latent.

https://shotcraft.video/learn/concepts/latent-space/

## What does channels mean in Latent space: working at proxy scale?

The many values stored per latent cell, describing content rather than colour.

https://shotcraft.video/learn/concepts/latent-space/

## What does Seed mean in Seeds and reproducibility: re-running a generation?

The integer that initialises the pseudo-random number generator producing the starting noise for a generation.

https://shotcraft.video/learn/concepts/seeds-and-reproducibility/

## What does Generator mean in Seeds and reproducibility: re-running a generation?

The stateful random-number object a pipeline draws from, which must be re-seeded per call to repeat a result.

https://shotcraft.video/learn/concepts/seeds-and-reproducibility/

## What does Determinism mean in Seeds and reproducibility: re-running a generation?

The property that identical inputs produce identical outputs, which in practice holds only within a fixed software, hardware and model configuration.

https://shotcraft.video/learn/concepts/seeds-and-reproducibility/

## What does Version pinning mean in Seeds and reproducibility: re-running a generation?

Fixing the exact model checkpoint and library versions used for a job so it can be rerun later.

https://shotcraft.video/learn/concepts/seeds-and-reproducibility/

## What does Generation log mean in Seeds and reproducibility: re-running a generation?

The record of prompt, seed, model version, parameters and conditioning asset hashes needed to reconstruct a delivered frame.

https://shotcraft.video/learn/concepts/seeds-and-reproducibility/

## What does attention mean in Transformers and DiT: why the backbone changed?

An operation where each position pulls a weighted mixture of information from all other positions.

https://shotcraft.video/learn/concepts/transformers-and-dit/

## What does patch mean in Transformers and DiT: why the backbone changed?

A fixed-size square of latent treated as one token by the transformer.

https://shotcraft.video/learn/concepts/transformers-and-dit/

## What does DiT mean in Transformers and DiT: why the backbone changed?

Diffusion Transformer: a latent diffusion model whose denoising backbone is a transformer rather than a U-Net.

https://shotcraft.video/learn/concepts/transformers-and-dit/

## What does U-Net mean in Transformers and DiT: why the backbone changed?

The convolutional encoder-decoder backbone used by earlier diffusion models, which processes the frame through a hierarchy of resolutions.

https://shotcraft.video/learn/concepts/transformers-and-dit/

## What does scaling behaviour mean in Transformers and DiT: why the backbone changed?

The observed relationship between compute or model size and output quality, which lets teams predict the return on more spending.

https://shotcraft.video/learn/concepts/transformers-and-dit/

## What does encoder mean in The VAE: a learned codec between pixels and latents?

The half of the autoencoder that turns pixels into a latent, used during training and for image-to-video conditioning.

https://shotcraft.video/learn/concepts/vae/

## What does decoder mean in The VAE: a learned codec between pixels and latents?

The half that turns a finished latent back into pixels, run once at the end of generation.

https://shotcraft.video/learn/concepts/vae/

## What does 3D VAE mean in The VAE: a learned codec between pixels and latents?

A video autoencoder that compresses along time as well as height and width.

https://shotcraft.video/learn/concepts/vae/

## What does reconstruction artifact mean in The VAE: a learned codec between pixels and latents?

A defect introduced by decoding rather than by the diffusion process, so it repeats across seeds.

https://shotcraft.video/learn/concepts/vae/

## What does checkerboard artifact mean in The VAE: a learned codec between pixels and latents?

Regular grid patterning produced by certain upsampling layers in a decoder.

https://shotcraft.video/learn/concepts/vae/

## What does spacetime patch mean in The time axis: spacetime patches and the cost of duration?

A token covering a small area of frame across a short span of frames, the basic unit a video model operates on.

https://shotcraft.video/learn/concepts/video-time-axis/

## What does sequence length mean in The time axis: spacetime patches and the cost of duration?

How many tokens one generation contains, growing with resolution and duration.

https://shotcraft.video/learn/concepts/video-time-axis/

## What does context length mean in The time axis: spacetime patches and the cost of duration?

The maximum number of tokens a model was trained to process at once, which caps clip length.

https://shotcraft.video/learn/concepts/video-time-axis/

## What does joint spatial-temporal attention mean in The time axis: spacetime patches and the cost of duration?

Attention computed across all positions in space and time at once, rather than separately per frame.

https://shotcraft.video/learn/concepts/video-time-axis/

## What does temporal layers mean in The time axis: spacetime patches and the cost of duration?

Blocks added to a pretrained image model so it can relate frames to each other, an alternative to dicing along time.

https://shotcraft.video/learn/concepts/video-time-axis/

## Explain why this claim is wrong: the model draws the image progressively, like a painter adding strokes to a blank canvas.

In fact every step operates on the whole frame at once. What changes between steps is how much noise is left in the frame, not how much of the picture has been filled in.

https://shotcraft.video/learn/concepts/diffusion-basics/

## Explain why this claim is wrong: the noise is a random ingredient the model sprinkles in for variety.

The noise is the starting material and the training signal. The model's only output is an estimate of the noise present, and the picture is the residue after those estimates are subtracted.

https://shotcraft.video/learn/concepts/diffusion-basics/

## Explain why this claim is wrong: the forward noising process is something the model learns.

Only the reverse direction is learned. Adding noise is fixed arithmetic, which is what gives training a gradeable right answer at every noise level.

https://shotcraft.video/learn/concepts/diffusion-basics/

## Explain why this claim is wrong: diffusion models retrieve and blend training images.

No image library is consulted at generation time. What is stored is a function mapping a noisy frame to a noise estimate, and the resemblance to training data comes from that function.

https://shotcraft.video/learn/concepts/diffusion-basics/

## Explain why this claim is wrong: more steps means a more detailed picture.

Sampling converges. Past the point where each step moves the frame a negligible distance, extra steps buy render time and nothing else.

https://shotcraft.video/learn/concepts/diffusion-basics/

## Explain why this claim is wrong: the latent is a small version of the picture, so a good upscaler could recover the missing detail.

It is not an image at all, but a grid of learned descriptions. Detail that was never encoded can only be invented, not recovered.

https://shotcraft.video/learn/concepts/latent-space/

## Explain why this claim is wrong: a larger output resolution means the model worked at that resolution.

Denoising happens on the latent grid; pixel dimensions come from the decoder. Asking for more pixels does not add latent capacity.

https://shotcraft.video/learn/concepts/latent-space/

## Explain why this claim is wrong: latent space is a map you can navigate by moving in a direction toward a concept.

Some directions are meaningful, but the tidy pictures people show are low-dimensional projections, and projections of high-dimensional data mislead about distance and grouping.

https://shotcraft.video/learn/concepts/latent-space/

## Explain why this claim is wrong: small text fails because the model cannot spell.

Spelling is part of it, but compression is the harder limit. When a single token covers a 32 by 32 pixel patch, there is no room to carry the shape of a small glyph.

https://shotcraft.video/learn/concepts/latent-space/

## Explain why this claim is wrong: all models share one latent space, so latents are portable.

Each autoencoder defines its own; matching an existing latent space is a deliberate engineering goal, not a default.

https://shotcraft.video/learn/concepts/latent-space/

## Explain why this claim is wrong: the same seed gives the same picture.

The same seed gives the same picture only with the same model version, parameters and, in practice, the same hardware and library stack.

https://shotcraft.video/learn/concepts/seeds-and-reproducibility/

## Explain why this claim is wrong: a seed is a style or a look.

It is an index into the noise the process starts from; it correlates with composition and framing for a fixed prompt, but it carries no meaning that transfers to a different prompt or model.

https://shotcraft.video/learn/concepts/seeds-and-reproducibility/

## Explain why this claim is wrong: reusing the same generator object in a batch loop pins the result.

Generator state is consumed as it is used, so each call in the loop draws different noise unless you re-seed before every call.

https://shotcraft.video/learn/concepts/seeds-and-reproducibility/

## Explain why this claim is wrong: enabling deterministic algorithms makes runs bit-identical everywhere.

Deterministic modes remove specific sources of variation and cost throughput, while the underlying framework still refuses to guarantee identity across releases, commits or platforms.

https://shotcraft.video/learn/concepts/seeds-and-reproducibility/

## Explain why this claim is wrong: keeping the prompt is keeping the revision trail.

Without the model version and every parameter, a saved prompt reproduces a family of images rather than the shot that was approved.

https://shotcraft.video/learn/concepts/seeds-and-reproducibility/

## Explain why this claim is wrong: attention means the model looks where a human would look.

Attention is weighted retrieval between positions inside the network. It sometimes correlates with salience, but it is not a gaze and not an explanation.

https://shotcraft.video/learn/concepts/transformers-and-dit/

## Explain why this claim is wrong: transformers replaced diffusion.

DiT keeps the diffusion training objective and replaces only the backbone that performs the denoising.

https://shotcraft.video/learn/concepts/transformers-and-dit/

## Explain why this claim is wrong: a token in an image or video model is an object, a word or a meaningful unit.

It is a fixed-size patch of latent, drawn on a grid with no regard for what happens to be inside it.

https://shotcraft.video/learn/concepts/transformers-and-dit/

## Explain why this claim is wrong: the DiT result proves bigger is always better.

It reports a scaling trend between compute per forward pass and FID on image benchmarks. That is a prediction about spending, not a guarantee about any particular shot.

https://shotcraft.video/learn/concepts/transformers-and-dit/

## Explain why this claim is wrong: U-Nets are obsolete.

They remain in wide use for images and in some video pipelines, and Lumiere's space-time U-Net is a recent design. The shift to transformers is about scaling behaviour, not about one architecture being wrong.

https://shotcraft.video/learn/concepts/transformers-and-dit/

## Explain why this claim is wrong: the VAE is part of the diffusion model.

It is a separate network, trained separately and sometimes swapped independently, so two pipelines can share a diffusion model and differ in their decoder.

https://shotcraft.video/learn/concepts/vae/

## Explain why this claim is wrong: every artifact comes from the diffusion model, so another seed will clear it.

Decoder artifacts are deterministic given the latent, which is why grid patterning, texture buzz and waxy skin reappear across every seed.

https://shotcraft.video/learn/concepts/vae/

## Explain why this claim is wrong: a video VAE is just an image VAE applied frame by frame.

Early pipelines did that and it leaves motion between consecutive frames unsmooth. Current video models compress along time too, which is what a 3D VAE means.

https://shotcraft.video/learn/concepts/vae/

## Explain why this claim is wrong: more latent compression is free efficiency.

It buys speed and duration and costs fine detail, which is why LTX-Video states the limitation directly and gives extra work to the decoder to compensate.

https://shotcraft.video/learn/concepts/vae/

## Explain why this claim is wrong: the decoder simply enlarges the latent.

It reconstructs pixels from a learned description, and in some designs it also performs the last denoising step, so it is making decisions about texture rather than scaling anything.

https://shotcraft.video/learn/concepts/vae/

## Explain why this claim is wrong: a video model generates frames one after another the way a camera exposes them.

Most current models denoise the entire clip at once, with all frames present in the same latent at every step.

https://shotcraft.video/learn/concepts/video-time-axis/

## Explain why this claim is wrong: doubling the clip length doubles the cost.

Token count grows with duration and attention cost grows faster than linearly in token count, so the bill grows worse than double.

https://shotcraft.video/learn/concepts/video-time-axis/

## Explain why this claim is wrong: frame rate is effectively free because interpolation is cheap.

In a spacetime-patch model, temporal resolution consumes the same token budget as spatial resolution. Some pipelines do add temporal super-resolution afterwards, but that is a separate model with its own failure modes.

https://shotcraft.video/learn/concepts/video-time-axis/

## Explain why this claim is wrong: the model maintains an internal 3D scene and renders frames from it.

It holds spacetime patches and attention between them. Any apparent geometry is a byproduct of appearance statistics, not a maintained representation.

https://shotcraft.video/learn/concepts/video-time-axis/

## Explain why this claim is wrong: a longer clip is simply a longer version of the same generation.

Length changes what fits in context, which changes how the model can hold the start and end of a move together.

https://shotcraft.video/learn/concepts/video-time-axis/

# 2. Reading a frame

## The subject’s size within the frame. Name the term: ____.

Shot scale

The subject’s size within the frame.

Working definitions with sources; labels describe practices rather than guaranteed model behaviour.

https://filmanalysis.yale.edu/cinematography/ · 2026-09-16

## The subject occupies little of an expansive view. Name the term: ____.

Extreme long shot

The subject occupies little of an expansive view.

Working definitions with sources; labels describe practices rather than guaranteed model behaviour.

https://filmanalysis.yale.edu/cinematography/ · 2026-09-16

## A standing figure fits nearly the full frame height. Name the term: ____.

Long shot

A standing figure fits nearly the full frame height.

Working definitions with sources; labels describe practices rather than guaranteed model behaviour.

https://filmanalysis.yale.edu/cinematography/ · 2026-09-16

## A person is framed approximately from the chest upward. Name the term: _____.

Medium close-up

A person is framed approximately from the chest upward.

Working definitions with sources; labels describe practices rather than guaranteed model behaviour.

https://filmanalysis.yale.edu/cinematography/ · 2026-09-16

## A head or comparably sized object dominates the frame. Name the term: _____.

Close-up

A head or comparably sized object dominates the frame.

Working definitions with sources; labels describe practices rather than guaranteed model behaviour.

https://filmanalysis.yale.edu/cinematography/ · 2026-09-16

## A small detail dominates the image. Name the term: _____.

Extreme close-up

A small detail dominates the image.

Working definitions with sources; labels describe practices rather than guaranteed model behaviour.

https://filmanalysis.yale.edu/cinematography/ · 2026-09-16

## The camera looks from above the subject. Name the term: _____.

High angle

The camera looks from above the subject.

Working definitions with sources; labels describe practices rather than guaranteed model behaviour.

https://filmanalysis.yale.edu/cinematography/ · 2026-09-16

## The camera looks from below the subject. Name the term: _____.

Low angle

The camera looks from below the subject.

Working definitions with sources; labels describe practices rather than guaranteed model behaviour.

https://filmanalysis.yale.edu/cinematography/ · 2026-09-16

## The frame is rolled away from level. Name the term: _____.

Canted framing

The frame is rolled away from level.

Working definitions with sources; labels describe practices rather than guaranteed model behaviour.

https://filmanalysis.yale.edu/cinematography/ · 2026-09-16

## The image approximates what a character sees. Name the term: _____.

Point of view

The image approximates what a character sees.

Working definitions with sources; labels describe practices rather than guaranteed model behaviour.

https://filmanalysis.yale.edu/cinematography/ · 2026-09-16

## Shot-size labels depend on context; they are not universal cuts. Name the term: _____.

Scale boundaries

Shot-size labels depend on context; they are not universal cuts.

Working definitions with sources; labels describe practices rather than guaranteed model behaviour.

https://filmanalysis.yale.edu/cinematography/ · 2026-09-16

## How would you use this principle: Operate the intention, not the actor?

a move that anticipates a turn reads as authorship, a move that chases it reads as an accident. Learn the action well enough to be very slightly early, every time.

https://shotcraft.video/learn/crafts/camera-operator/

## How would you use this principle: Composition is the management of empty space?

where you leave room tells an audience where to expect something. A walking figure with lead room feels ordinary; the same figure pressed against the leading edge feels pursued, and nothing else in the frame has to change.

https://shotcraft.video/learn/crafts/camera-operator/

## How would you use this principle: Match the look, not just the side of the line?

coverage cuts together when the height and angle of an eyeline agree, not merely when both singles stay within the same 180 degrees. Two performances photographed hours apart become one conversation only if the looks line up.

https://shotcraft.video/learn/crafts/camera-operator/

## How would you use this principle: The settle is part of the move?

audiences forgive a rough start far more readily than a frame that never comes to rest. Decide the final frame first and work backwards to the start position.

https://shotcraft.video/learn/crafts/camera-operator/

## How would you use this principle: The operator is the last human filter before recording?

nobody in post can put back what the frame excluded, so a problem seen in the eyepiece has to be said out loud before the take, not after it.

https://shotcraft.video/learn/crafts/camera-operator/

# 3. The camera

## The camera turns sideways. Name the term: _____.

Pan

The camera turns sideways.

Working definitions with sources; labels describe practices rather than guaranteed model behaviour.

https://filmanalysis.yale.edu/cinematography/ · 2026-09-16

## The camera turns upward or downward. Name the term: _____.

Tilt

The camera turns upward or downward.

Working definitions with sources; labels describe practices rather than guaranteed model behaviour.

https://filmanalysis.yale.edu/cinematography/ · 2026-09-16

## The camera travels through space. Name the term: _____.

Tracking

The camera travels through space.

Working definitions with sources; labels describe practices rather than guaranteed model behaviour.

https://filmanalysis.yale.edu/cinematography/ · 2026-09-16

## Rapid sideways rotation creates streaked blur. Name the term: _____.

Whip pan

Rapid sideways rotation creates streaked blur.

Working definitions with sources; labels describe practices rather than guaranteed model behaviour.

https://filmanalysis.yale.edu/cinematography/ · 2026-09-16

## Sharpness transfers between depths. Name the term: _____.

Rack focus

Sharpness transfers between depths.

Working definitions with sources; labels describe practices rather than guaranteed model behaviour.

https://filmanalysis.yale.edu/cinematography/ · 2026-09-16

## Focal length changes. Name the term: _____.

Zoom

Focal length changes.

Working definitions with sources; labels describe practices rather than guaranteed model behaviour.

https://filmanalysis.yale.edu/cinematography/ · 2026-09-16

## An arm carries the camera through space. Name the term: _____.

Crane

An arm carries the camera through space.

Working definitions with sources; labels describe practices rather than guaranteed model behaviour.

https://filmanalysis.yale.edu/cinematography/ · 2026-09-16

## Small adjustments keep the subject in view. Name the term: _____.

Reframing

Small adjustments keep the subject in view.

Working definitions with sources; labels describe practices rather than guaranteed model behaviour.

https://filmanalysis.yale.edu/cinematography/ · 2026-09-16

## A wheeled platform moves the camera through space. Name the term: _____.

Dolly

A wheeled platform moves the camera through space.

Physical camera equipment, as described by the American Society of Cinematographers. These names do not guarantee a model will reproduce the move.

https://theasc.com/article/shot-craft-camera-movement/ · 2026-09-16

## Camera travel and an opposing zoom hold the subject’s size while the background changes scale. Name the term: _____.

Dolly zoom

Camera travel and an opposing zoom hold the subject’s size while the background changes scale.

Physical camera equipment, as described by the American Society of Cinematographers. These names do not guarantee a model will reproduce the move.

https://theasc.com/article/shot-craft-camera-movement/ · 2026-09-16

## A small crane carries the camera on a pivoting arm. Name the term: _____.

Jib

A small crane carries the camera on a pivoting arm.

Physical camera equipment, as described by the American Society of Cinematographers. These names do not guarantee a model will reproduce the move.

https://theasc.com/article/shot-craft-camera-movement/ · 2026-09-16

## A branded body-mounted stabilizer uses a vest and spring arms to isolate movement. Name the term: ____.

Steadicam

A branded body-mounted stabilizer uses a vest and spring arms to isolate movement.

Physical camera equipment, as described by the American Society of Cinematographers. These names do not guarantee a model will reproduce the move.

https://theasc.com/article/shot-craft-camera-movement/ · 2026-09-16

## A compact camera platform travels along a short track. Name the term: _____.

Slider

A compact camera platform travels along a short track.

Physical camera equipment, as described by the American Society of Cinematographers. These names do not guarantee a model will reproduce the move.

https://theasc.com/article/shot-craft-camera-movement/ · 2026-09-16

## The operator’s body supports the camera. Name the term: ____.

Handheld

The operator’s body supports the camera.

Physical camera equipment, as described by the American Society of Cinematographers. These names do not guarantee a model will reproduce the move.

https://theasc.com/article/shot-craft-camera-movement/ · 2026-09-16

## What does Camera trajectory conditioning mean in Camera control: a move without a camera?

Supplying a sequence of camera poses as an explicit input so the generated clip approximates a specified move.

https://shotcraft.video/learn/concepts/camera-control/

## What does Motion prior mean in Camera control: a move without a camera?

The statistical tendency a model has learned about how images move, which fills in everything the prompt and conditioning do not specify.

https://shotcraft.video/learn/concepts/camera-control/

## What does Parallax mean in Camera control: a move without a camera?

The differential apparent motion of near and far objects under a moving viewpoint, which a generative model imitates rather than computes.

https://shotcraft.video/learn/concepts/camera-control/

## What does Appearance-free condition mean in Camera control: a move without a camera?

A control signal such as a camera pose or trajectory chosen so it alters motion without changing the shape or look of objects.

https://shotcraft.video/learn/concepts/camera-control/

## How do Motion prior and Camera trajectory conditioning differ?

Motion prior: The statistical tendency a model has learned about how images move, which fills in everything the prompt and conditioning do not specify. Camera trajectory conditioning: Supplying a sequence of camera poses as an explicit input so the generated clip approximates a specified move.

https://shotcraft.video/learn/concepts/camera-control/

## Explain why this claim is wrong: the model builds a 3D scene and renders a camera through it.

It predicts pixels across space and time; the Sora report presents 3D consistency as an emergent behaviour of scale, not as an explicit geometric stage in the pipeline.

https://shotcraft.video/learn/concepts/camera-control/

## Explain why this claim is wrong: naming a focal length in the prompt sets a focal length.

It biases the image toward the compression, distortion and depth cues typical of images captioned that way, which is why 24mm and 85mm read differently even though no lens exists.

https://shotcraft.video/learn/concepts/camera-control/

## Explain why this claim is wrong: trajectory conditioning guarantees the move you specified.

Camera pose modules such as CameraCtrl markedly improve fidelity to a trajectory but sit on top of a frozen generator, so the move is followed approximately and degrades on trajectories unlike the training data.

https://shotcraft.video/learn/concepts/camera-control/

## Explain why this claim is wrong: camera motion and subject motion are one control.

MotionCtrl exists precisely because the two were conflated, and separating them was the contribution; in prompt-only workflows they still bleed into each other.

https://shotcraft.video/learn/concepts/camera-control/

## Explain why this claim is wrong: a longer generation gives a longer move.

Clip length is set by the model's token budget, and stitching two generations to extend a move gives you two similar moves, not one continuous one.

https://shotcraft.video/learn/concepts/camera-control/

# 4. Light

## The main source establishes the subject’s lighting. Name the term: _____.

Key light

The main source establishes the subject’s lighting.

SouthWorks’ practical lighting guidance, published by Adobe. A setup is a starting point, not a universal requirement.

https://helpx.adobe.com/ie/stock/contributor/help/artist-hub-migration/lighting-for-video.html · 2026-09-16

## Additional illumination lifts the shadows cast by the key. Name the term: _____.

Fill light

Additional illumination lifts the shadows cast by the key.

SouthWorks’ practical lighting guidance, published by Adobe. A setup is a starting point, not a universal requirement.

https://helpx.adobe.com/ie/stock/contributor/help/artist-hub-migration/lighting-for-video.html · 2026-09-16

## Light behind the subject helps separate it from the background. Name the term: _____.

Backlight

Light behind the subject helps separate it from the background.

SouthWorks’ practical lighting guidance, published by Adobe. A setup is a starting point, not a universal requirement.

https://helpx.adobe.com/ie/stock/contributor/help/artist-hub-migration/lighting-for-video.html · 2026-09-16

## Dark flags or fabric absorb stray light to deepen shadows. Name the term: _____.

Negative fill

Dark flags or fabric absorb stray light to deepen shadows.

SouthWorks’ practical lighting guidance, published by Adobe. A setup is a starting point, not a universal requirement.

https://helpx.adobe.com/ie/stock/contributor/help/artist-hub-migration/lighting-for-video.html · 2026-09-16

## Increasing the apparent size of the source softens shadow edges. Name the term: _____.

Soft light

Increasing the apparent size of the source softens shadow edges.

SouthWorks’ practical lighting guidance, published by Adobe. A setup is a starting point, not a universal requirement.

https://helpx.adobe.com/ie/stock/contributor/help/artist-hub-migration/lighting-for-video.html · 2026-09-16

## A pale surface redirects light into the scene. Name the term: _____.

Bounce

A pale surface redirects light into the scene.

SouthWorks’ practical lighting guidance, published by Adobe. A setup is a starting point, not a universal requirement.

https://helpx.adobe.com/ie/stock/contributor/help/artist-hub-migration/lighting-for-video.html · 2026-09-16

## A source remains illuminated during recording. Name the term: _____.

Continuous light

A source remains illuminated during recording.

SouthWorks’ practical lighting guidance, published by Adobe. A setup is a starting point, not a universal requirement.

https://helpx.adobe.com/ie/stock/contributor/help/artist-hub-migration/lighting-for-video.html · 2026-09-16

## The source’s warm or cool appearance is described in kelvins. Name the term: _____.

Colour temperature

The source’s warm or cool appearance is described in kelvins.

SouthWorks’ practical lighting guidance, published by Adobe. A setup is a starting point, not a universal requirement.

https://helpx.adobe.com/ie/stock/contributor/help/artist-hub-migration/lighting-for-video.html · 2026-09-16

## Weather and the changing sun can alter exposure and colour during a shoot. Name the term: _____.

Natural-light continuity

Weather and the changing sun can alter exposure and colour during a shoot.

SouthWorks’ practical lighting guidance, published by Adobe. A setup is a starting point, not a universal requirement.

https://helpx.adobe.com/ie/stock/contributor/help/artist-hub-migration/lighting-for-video.html · 2026-09-16

## How would you use this principle: Concept belongs to the cinematographer, implementation belongs to the gaffer?

on the ASC's Camerimage panel on exactly this question, Vilmos Zsigmond described the cinematographer as in charge of the concept of the light and the gaffer as responsible for making it happen, and James Plannette described roughing the light in and then finishing it with the cinematographer, who is, in his phrase, the decider. That does not mean there is no argument, only that the argument ends.

https://shotcraft.video/learn/crafts/gaffer/

## How would you use this principle: Know which system you are working in?

the same panel spelled out that in the American division of labour the grips set the flags, diffusion and bounce while the electricians handle the fixture, whereas in the European system the gaffer and electricians set everything involving the fixture, flags and diffusion included. Turning up with the wrong assumption wastes an hour and irritates two departments.

https://shotcraft.video/learn/crafts/gaffer/

## How would you use this principle: Plan the power before the picture?

a lighting design a location's switchboard cannot carry is not a design. Read the electrical installation condition report, look at the board for signs of overload and age, and size distribution, protection and emergency lighting for the whole rig.

https://shotcraft.video/learn/crafts/gaffer/

## How would you use this principle: Repeatability beats brilliance?

a look nobody can rebuild after a unit move is a look you shot once. Numbered or colour-coded channels, desk notes filed against slate numbers, and metered colour temperature are how a night interior matches three weeks later on another stage.

https://shotcraft.video/learn/crafts/gaffer/

## How would you use this principle: Speed is a lighting value?

the fastest gaffer buys the cinematographer the ten minutes that make the last shot of the day look like the first. Order the setups so the fewest fixtures have to move.

https://shotcraft.video/learn/crafts/gaffer/

## How would you use this principle: Movement is dramatic punctuation or it is noise?

a push that begins on a line and settles on a reaction means something, while the same push drifting through a scene only says the camera is restless. Ask what the move is for before deciding how to build it.

https://shotcraft.video/learn/crafts/key-grip/

## How would you use this principle: The department's real product is repeatability?

a dolly move that can be hit identically on take nine, same start mark, same speed, same settle, is what lets a director choose a take on performance rather than on mechanics.

https://shotcraft.video/learn/crafts/key-grip/

## How would you use this principle: The division of labour is regional, not universal?

the ASC's cinematographers-and-gaffers panel put it plainly, that American grips set the flags, diffusion and bounce while in the European system the gaffer's electricians handle everything attached to the fixture. Read the room before you touch a stand.

https://shotcraft.video/learn/crafts/key-grip/

## How would you use this principle: Design from the location, not the catalogue?

uneven ground, a protected floor, a doorway three inches too narrow. Recce notes with measurements and photographs are what turn a shot list into a build that works on the day.

https://shotcraft.video/learn/crafts/key-grip/

## How would you use this principle: Quote the honest number early?

install and strike times belong in the schedule before the day, not on it. A key grip who says an hour and means an hour keeps their safety veto credible for the moment it is needed.

https://shotcraft.video/learn/crafts/key-grip/

## How would you use this principle: Motivate the key?

every light the audience can perceive should have a plausible cause inside the world of the shot, a window, a fire, a practical, the sky. Motivation is what lets the audience stop noticing lighting and start believing the space. When you need a light with no source, hide it in a direction the eye is not auditing.

https://shotcraft.video/learn/crafts/lighting-td/

## How would you use this principle: Light the story point, not the set?

a CG set is fully modelled and fully lit-able, which is the trap. Decide what the audience must see in this shot and let the rest fall away. Wide-open even illumination is a failure of decision, not a neutral choice.

https://shotcraft.video/learn/crafts/lighting-td/

## How would you use this principle: Continuity across a sequence is the actual deliverable?

a shot that is beautiful alone and half a stop off its neighbour is a broken shot. Build the sequence rig first, review shots side by side and in cut order rather than one at a time, and treat any drift in key direction or level as a bug.

https://shotcraft.video/learn/crafts/lighting-td/

## How would you use this principle: Guide the eye with contrast and colour temperature?

the brightest, warmest, most saturated area in frame is where the audience looks. Use that deliberately, and if the thing they should look at is not that area, you have a note to give yourself before the director gives it to you.

https://shotcraft.video/learn/crafts/lighting-td/

## How would you use this principle: Render cost is a design constraint you own?

volumes, area lights, transmissive materials and deep ray depths all buy realism with time. Knowing what a lighting idea costs, and finding the version that reads for a fraction of it, is as much of the craft as taste.

https://shotcraft.video/learn/crafts/lighting-td/

# 5. The lens and the look

## The focal length is fixed. Name the term: _____.

Prime lens

The focal length is fixed.

Nikon’s explanation of physical lenses. A lens name in a generation prompt does not install that optical system.

https://www.nikonusa.com/learn-and-explore/c/tips-and-techniques/understanding-focal-length · 2026-09-16

## The focal length is adjustable. Name the term: _____.

Zoom lens

The focal length is adjustable.

Nikon’s explanation of physical lenses. A lens name in a generation prompt does not install that optical system.

https://www.nikonusa.com/learn-and-explore/c/tips-and-techniques/understanding-focal-length · 2026-09-16

## For the same camera format and viewpoint, it narrows the view and increases magnification. Name the term: _____.

Longer focal length

For the same camera format and viewpoint, it narrows the view and increases magnification.

Nikon’s explanation of physical lenses. A lens name in a generation prompt does not install that optical system.

https://www.nikonusa.com/learn-and-explore/c/tips-and-techniques/understanding-focal-length · 2026-09-16

## For the same camera format and viewpoint, it widens the view and reduces magnification. Name the term: _____.

Shorter focal length

For the same camera format and viewpoint, it widens the view and reduces magnification.

Nikon’s explanation of physical lenses. A lens name in a generation prompt does not install that optical system.

https://www.nikonusa.com/learn-and-explore/c/tips-and-techniques/understanding-focal-length · 2026-09-16

## Designed to focus close to small subjects. Name the term: _____.

Macro lens

Designed to focus close to small subjects.

Nikon’s explanation of physical lenses. A lens name in a generation prompt does not install that optical system.

https://www.nikonusa.com/learn-and-explore/c/tips-and-techniques/understanding-focal-length · 2026-09-16

## The opening through which light enters the lens. Name the term: _____.

Aperture

The opening through which light enters the lens.

Nikon’s explanation of aperture. Compare changes with the other camera and scene conditions held constant.

https://www.nikonusa.com/learn-and-explore/c/tips-and-techniques/what-is-aperture · 2026-09-16

## A larger lens opening admits more light. Name the term: _____.

Low f-number

A larger lens opening admits more light.

Nikon’s explanation of aperture. Compare changes with the other camera and scene conditions held constant.

https://www.nikonusa.com/learn-and-explore/c/tips-and-techniques/what-is-aperture · 2026-09-16

## A smaller lens opening admits less light. Name the term: _____.

High f-number

A smaller lens opening admits less light.

Nikon’s explanation of aperture. Compare changes with the other camera and scene conditions held constant.

https://www.nikonusa.com/learn-and-explore/c/tips-and-techniques/what-is-aperture · 2026-09-16

## The span of distances that appears acceptably sharp. Name the term: _____.

Depth of field

The span of distances that appears acceptably sharp.

Nikon’s explanation of aperture. Compare changes with the other camera and scene conditions held constant.

https://www.nikonusa.com/learn-and-explore/c/tips-and-techniques/what-is-aperture · 2026-09-16

## It reduces depth of field when other conditions stay fixed. Name the term: _____.

Wide aperture

It reduces depth of field when other conditions stay fixed.

Nikon’s explanation of aperture. Compare changes with the other camera and scene conditions held constant.

https://www.nikonusa.com/learn-and-explore/c/tips-and-techniques/what-is-aperture · 2026-09-16

## How would you use this principle: Grade from a reference and the story's arc, not shot by shot?

choose the anchor frame for a sequence and match to it. A reel graded one pretty shot at a time has no shape, and an audience should feel the film brighten or cool because the story does.

https://shotcraft.video/learn/crafts/colorist/

## How would you use this principle: Respect the display transform?

in a colour-managed pipeline the path from scene-referred data to the display is fixed, and the grade belongs in front of it. Fighting the transform, clipping highlights to force contrast and then rescuing them, is why a grade stops travelling between HDR, SDR and a phone.

https://shotcraft.video/learn/crafts/colorist/

## How would you use this principle: Contrast before colour?

putting black, white and the mid-tone where they belong does most of what people call the look. Hue moves made before the tonal range is set have to be made again.

https://shotcraft.video/learn/crafts/colorist/

## How would you use this principle: Shape the eye's path?

a frame is a hierarchy. Windows, tracked mattes and small desaturations decide what a viewer looks at first, second and never. This is closer to lighting than to painting.

https://shotcraft.video/learn/crafts/colorist/

## How would you use this principle: Latitude is finite and belongs to the story?

log camera negative holds a great deal but not everything. Spend the recoverable highlight or shadow range where the scene needs it and stop reaching for more.

https://shotcraft.video/learn/crafts/colorist/

## How would you use this principle: Shots match when the faces match?

scopes keep you consistent and legal, but the judgement is made on skin, sky and any object the audience already knows the colour of.

https://shotcraft.video/learn/crafts/colorist/

## How would you use this principle: A focus pull is a rehearsed performance, not a reaction?

it is a set of physical marks and a rhythm learned in rehearsal. When the actor changes the timing, you follow their intention rather than their distance, which is why the good ones watch the performer and not the barrel.

https://shotcraft.video/learn/crafts/first-ac/

## How would you use this principle: The subject is a decision, not a fact?

in a two-shot at a wide stop, somebody is soft. Knowing which eye of which actor the scene needs sharp on which line is the difference between a focus puller and a rangefinder.

https://shotcraft.video/learn/crafts/first-ac/

## How would you use this principle: Prep is where the day is won?

a lens with drifting back focus, a loose mount or an unbalanced head produces failures that look like operator error and cost hours on the floor. Collimation and tracking get checked before the camera leaves the rental house, not on set.

https://shotcraft.video/learn/crafts/first-ac/

## How would you use this principle: Report the soft take?

the only expensive mistake is the one found in the edit. Saying 'that one was soft' within five seconds is the job, and the role rewards people who are comfortable being briefly unpopular.

https://shotcraft.video/learn/crafts/first-ac/

## How would you use this principle: Own the physical order of the truck and the cart?

kit packed in shooting order buys minutes at every setup, and those minutes are what let the cinematographer light the last shot of the day properly.

https://shotcraft.video/learn/crafts/first-ac/

# 6. Writing a prompt

## What does step mean in Sampling and guidance: the two dials that shape a generation?

One pass of the model plus one scheduler update, moving the latent partway from noise toward a picture.

https://shotcraft.video/learn/concepts/sampling-and-guidance/

## What does scheduler mean in Sampling and guidance: the two dials that shape a generation?

The rule that decides step sizes and how each noise prediction is applied; also called the sampler.

https://shotcraft.video/learn/concepts/sampling-and-guidance/

## What does classifier-free guidance mean in Sampling and guidance: the two dials that shape a generation?

Running the model with and without conditioning and pushing away from the unconditional prediction to strengthen prompt adherence.

https://shotcraft.video/learn/concepts/sampling-and-guidance/

## What does guidance scale mean in Sampling and guidance: the two dials that shape a generation?

How strongly that push is applied, trading prompt fidelity against variety.

https://shotcraft.video/learn/concepts/sampling-and-guidance/

## What does distillation mean in Sampling and guidance: the two dials that shape a generation?

Training a student model to reproduce a many-step sampler's output in far fewer steps.

https://shotcraft.video/learn/concepts/sampling-and-guidance/

## What does text encoder mean in Text encoders and conditioning: how a prompt steers the picture?

A separate network that converts prompt tokens into vectors the diffusion model can attend to.

https://shotcraft.video/learn/concepts/text-encoders-conditioning/

## What does contrastive training mean in Text encoders and conditioning: how a prompt steers the picture?

Learning by matching images to their captions among many wrong options, as CLIP does.

https://shotcraft.video/learn/concepts/text-encoders-conditioning/

## What does cross-attention mean in Text encoders and conditioning: how a prompt steers the picture?

The mechanism by which each image patch reads a weighted mixture of the prompt token vectors at every denoising step.

https://shotcraft.video/learn/concepts/text-encoders-conditioning/

## What does conditioning mean in Text encoders and conditioning: how a prompt steers the picture?

Any extra input that biases generation, including text, a start image or structural maps.

https://shotcraft.video/learn/concepts/text-encoders-conditioning/

## What does frozen encoder mean in Text encoders and conditioning: how a prompt steers the picture?

A text encoder whose weights stay fixed while the diffusion model trains around it.

https://shotcraft.video/learn/concepts/text-encoders-conditioning/

## In the prompt reference’s suggested order, which comes first, second and third? Shot size / Aspect ratio / Duration

Duration → Aspect ratio → Shot size

Work down the aspects in order. The order is the recommended order in the prompt too.

https://shotcraft.video/prompts/

## In the prompt reference’s suggested order, which comes first, second and third? Lens and focal length / Camera movement / Camera angle

Camera angle → Camera movement → Lens and focal length

Work down the aspects in order. The order is the recommended order in the prompt too.

https://shotcraft.video/prompts/

## In the prompt reference’s suggested order, which comes first, second and third? Colour and grade / Lighting / Depth of field

Depth of field → Lighting → Colour and grade

Work down the aspects in order. The order is the recommended order in the prompt too.

https://shotcraft.video/prompts/

## In the prompt reference’s suggested order, which comes first, second and third? Mood / Subject motion / Medium and style

Medium and style → Subject motion → Mood

Work down the aspects in order. The order is the recommended order in the prompt too.

https://shotcraft.video/prompts/

## In the prompt reference’s suggested order, which comes first, second and third? Resolution / quality tier / Pacing and cuts / Audio bed

Audio bed → Pacing and cuts → Resolution / quality tier

Work down the aspects in order. The order is the recommended order in the prompt too.

https://shotcraft.video/prompts/

## In the prompt reference’s suggested order, which comes first, second and third? Edit instruction (video editing workflows) / Ad format / mode (branded video) / First/last frame anchoring

First/last frame anchoring → Ad format / mode (branded video) → Edit instruction (video editing workflows)

Work down the aspects in order. The order is the recommended order in the prompt too.

https://shotcraft.video/prompts/

## Study this original prompt example: A luxury fragrance film. A plain rectangular amber glass perfume bottle with a brushed brass cap stands on dark stone. Begin in a close product shot and slowly move the camera toward the bottle, keeping its whole silhouette visible. A broad soft light makes a clean reflection along the glass; a narrow warm backlight defines the edges. Deep charcoal surroundings, restrained amber colour, crisp cap detail. The bottle remains still as the camera settles on the final composition. Which choice describes its Framing?

Close product shot

Original starting points, not tested generation results. Adapt them to your model and inspect the output.

https://shotcraft.video/prompts/luxury-video-prompts/

## Study this original prompt example: A luxury fragrance film. A plain rectangular amber glass perfume bottle with a brushed brass cap stands on dark stone. Begin in a close product shot and slowly move the camera toward the bottle, keeping its whole silhouette visible. A broad soft light makes a clean reflection along the glass; a narrow warm backlight defines the edges. Deep charcoal surroundings, restrained amber colour, crisp cap detail. The bottle remains still as the camera settles on the final composition. Which choice describes its Movement?

Slow camera push-in

Original starting points, not tested generation results. Adapt them to your model and inspect the output.

https://shotcraft.video/prompts/luxury-video-prompts/

## Study this original prompt example: A luxury fragrance film. A plain rectangular amber glass perfume bottle with a brushed brass cap stands on dark stone. Begin in a close product shot and slowly move the camera toward the bottle, keeping its whole silhouette visible. A broad soft light makes a clean reflection along the glass; a narrow warm backlight defines the edges. Deep charcoal surroundings, restrained amber colour, crisp cap detail. The bottle remains still as the camera settles on the final composition. Which choice describes its Light?

Soft reflection and warm edge light

Original starting points, not tested generation results. Adapt them to your model and inspect the output.

https://shotcraft.video/prompts/luxury-video-prompts/

## Study this original prompt example: A luxury fragrance film. A plain rectangular amber glass perfume bottle with a brushed brass cap stands on dark stone. Begin in a close product shot and slowly move the camera toward the bottle, keeping its whole silhouette visible. A broad soft light makes a clean reflection along the glass; a narrow warm backlight defines the edges. Deep charcoal surroundings, restrained amber colour, crisp cap detail. The bottle remains still as the camera settles on the final composition. Which choice describes its Palette?

Amber and charcoal

Original starting points, not tested generation results. Adapt them to your model and inspect the output.

https://shotcraft.video/prompts/luxury-video-prompts/

## Explain why this claim is wrong: more steps always improves the image.

Sampling converges, and past the plateau the frame stops changing while the render time keeps climbing.

https://shotcraft.video/learn/concepts/sampling-and-guidance/

## Explain why this claim is wrong: the scheduler is a quality setting where one option is simply better.

It is a different route through the same trained weights. Different schedulers give genuinely different pictures from the same seed, and none is a strict upgrade.

https://shotcraft.video/learn/concepts/sampling-and-guidance/

## Explain why this claim is wrong: higher guidance means better prompt adherence.

Only up to a point. Beyond it, contrast and saturation blow out, composition collapses to one reading, and variety across seeds disappears, which is exactly the fidelity-versus-diversity trade the guidance paper describes.

https://shotcraft.video/learn/concepts/sampling-and-guidance/

## Explain why this claim is wrong: few-step distilled models are just fast versions of the same model.

They are separately distilled models with their own look, and they often fix guidance internally, so a control you relied on may not be exposed.

https://shotcraft.video/learn/concepts/sampling-and-guidance/

## Explain why this claim is wrong: the seed is the only thing that decides the output.

Scheduler, step count, guidance value and model version all change the result, so all of them belong in the record if a shot has to be reproduced.

https://shotcraft.video/learn/concepts/sampling-and-guidance/

## Explain why this claim is wrong: the model reads the prompt as a sentence and follows it as an instruction.

It receives a set of token vectors, and each image patch weights those vectors for itself. There is no stage where the instruction is parsed and obeyed as a whole.

https://shotcraft.video/learn/concepts/text-encoders-conditioning/

## Explain why this claim is wrong: longer, more detailed prompts always help.

Adding tokens spreads attention across more of them and can weaken the ones already working. Specific concrete nouns usually beat length.

https://shotcraft.video/learn/concepts/text-encoders-conditioning/

## Explain why this claim is wrong: negative words remove things from the image.

Attention has no natural mechanism for not. Where systems do support negatives, they implement them as a separate conditioning pass, not as a word inside the prompt.

https://shotcraft.video/learn/concepts/text-encoders-conditioning/

## Explain why this claim is wrong: any competent text encoder is interchangeable with any other.

Encoders capture different things. Imagen's result that scaling the text encoder beat scaling the diffusion model shows how much rides on this, and models built on different encoders answer the same phrasing differently.

https://shotcraft.video/learn/concepts/text-encoders-conditioning/

## Explain why this claim is wrong: the text encoder is trained alongside the image model.

It is often frozen, trained separately on its own objective, and simply queried by the diffusion model.

https://shotcraft.video/learn/concepts/text-encoders-conditioning/

# 7. What models can and cannot do

## What does Identity adapter mean in Character and set consistency: the weakest link?

A small trained module that lets a frozen base model read a reference image of a subject as a prompt, IP-Adapter being the common example.

https://shotcraft.video/learn/concepts/character-consistency/

## What does Subject fine-tuning mean in Character and set consistency: the weakest link?

Updating model weights, or an adapter on them, from a small image set so a specific person, object or set can be called up by name.

https://shotcraft.video/learn/concepts/character-consistency/

## What does Textual Inversion mean in Character and set consistency: the weakest link?

Learning a new embedding, effectively a new word, for a concept while leaving the model itself unchanged.

https://shotcraft.video/learn/concepts/character-consistency/

## What does Prior preservation mean in Character and set consistency: the weakest link?

A training term that keeps a fine-tune from overwriting the model's general knowledge of the class your subject belongs to.

https://shotcraft.video/learn/concepts/character-consistency/

## What does Character sheet mean in Character and set consistency: the weakest link?

A prepared set of reference frames covering angles, expressions, wardrobe and lighting, used as conditioning across a sequence.

https://shotcraft.video/learn/concepts/character-consistency/

## What does Conditioning mean in Conditioning and control: everything you hand the model besides noise?

Any input other than the initial noise that biases what the model generates, including text, images, masks, structure maps and camera trajectories.

https://shotcraft.video/learn/concepts/conditioning-and-control/

## What does Cross-attention mean in Conditioning and control: everything you hand the model besides noise?

The layer type that lets a diffusion model read a conditioning signal such as an encoded text prompt while denoising.

https://shotcraft.video/learn/concepts/conditioning-and-control/

## What does Structure conditioning mean in Conditioning and control: everything you hand the model besides noise?

Conditioning on a spatial map derived from a reference, such as depth, edges or human pose, to fix geometry while leaving appearance to the model.

https://shotcraft.video/learn/concepts/conditioning-and-control/

## What does Conditioning scale mean in Conditioning and control: everything you hand the model besides noise?

A weight that sets how strongly a structure control overrides the model's own tendencies, traded against image quality.

https://shotcraft.video/learn/concepts/conditioning-and-control/

## What does Outpainting mean in Conditioning and control: everything you hand the model besides noise?

Generating new content beyond the existing frame edges, used to extend a plate or change aspect ratio.

https://shotcraft.video/learn/concepts/conditioning-and-control/

## What does LoRA mean in LoRAs and fine-tuning: teaching a model your subject?

Low-rank adaptation: a training method that freezes the base weights and learns small rank-decomposition matrices inserted into each layer.

https://shotcraft.video/learn/concepts/loras-and-finetuning/

## What does Rank mean in LoRAs and fine-tuning: teaching a model your subject?

The size of the low-rank update, which sets how much the adapter can learn and how large the resulting file is.

https://shotcraft.video/learn/concepts/loras-and-finetuning/

## What does Base checkpoint mean in LoRAs and fine-tuning: teaching a model your subject?

The specific pretrained model an adapter was trained against and must be loaded with.

https://shotcraft.video/learn/concepts/loras-and-finetuning/

## What does Trigger word mean in LoRAs and fine-tuning: teaching a model your subject?

The token a fine-tune binds a subject or style to, used to call that concept up in a prompt.

https://shotcraft.video/learn/concepts/loras-and-finetuning/

## What does Overfitting mean in LoRAs and fine-tuning: teaching a model your subject?

Training so far into the dataset that the model reproduces its incidental details and loses responsiveness to prompts.

https://shotcraft.video/learn/concepts/loras-and-finetuning/

## What does temporal flicker mean in Temporal consistency: flicker, drift and morphing?

Frame-to-frame instability in texture, brightness or fine detail that was never committed to in the latent.

https://shotcraft.video/learn/concepts/temporal-consistency/

## What does identity drift mean in Temporal consistency: flicker, drift and morphing?

A face or character gradually changing across a shot because the model stores no persistent identity.

https://shotcraft.video/learn/concepts/temporal-consistency/

## What does texture crawl mean in Temporal consistency: flicker, drift and morphing?

Patterns and fine surfaces reshuffling each frame instead of holding still on the object.

https://shotcraft.video/learn/concepts/temporal-consistency/

## What does keyframe cascade mean in Temporal consistency: flicker, drift and morphing?

Generating distant frames first and then filling the gaps with temporal super-resolution models.

https://shotcraft.video/learn/concepts/temporal-consistency/

## What does temporal super-resolution mean in Temporal consistency: flicker, drift and morphing?

A separately trained model that increases frame rate by generating intermediate frames.

https://shotcraft.video/learn/concepts/temporal-consistency/

## What does Spacetime patch mean in The economics of generated video: why iteration is a budget line?

The unit a video transformer processes, a small block of latent covering an area of frame across a span of frames.

https://shotcraft.video/learn/concepts/video-economics/

## What does Context length mean in The economics of generated video: why iteration is a budget line?

The maximum number of tokens a model can attend over at once, which sets the longest clip it can generate in one pass.

https://shotcraft.video/learn/concepts/video-economics/

## What does Compression ratio mean in The economics of generated video: why iteration is a budget line?

How much smaller the latent representation is than the pixels, traded directly against retained fine detail.

https://shotcraft.video/learn/concepts/video-economics/

## What does Denoising steps mean in The economics of generated video: why iteration is a budget line?

The number of iterative passes over the whole latent clip, each one a full-cost forward pass.

https://shotcraft.video/learn/concepts/video-economics/

## What does Accept rate mean in The economics of generated video: why iteration is a budget line?

The proportion of generations a human keeps, which is the number that actually determines cost per finished shot.

https://shotcraft.video/learn/concepts/video-economics/

## What does Native audio mean in Video model capability matrix?

The model returns a soundtrack with the picture, generated in the same pass, rather than leaving you to add sound afterwards.

https://shotcraft.video/reference/model-capabilities/ · 2026-08-17

## What does Negative prompt field mean in Video model capability matrix?

A separate field for what the model must leave out. Where there is none, exclusions have to sit inside the prompt itself, where they compete for length and are read as description.

A field is not the same as an effect. Some models accept the field and act on it weakly. See the prompt length guide for how much room the field gets, which differs from the prompt's own budget.

https://shotcraft.video/reference/model-capabilities/ · 2026-08-17

## What does Image as start frame mean in Video model capability matrix?

The generation can be anchored to a still you supply, which is how you control composition, cast and colour instead of describing them.

https://shotcraft.video/reference/model-capabilities/ · 2026-08-17

## What does Last frame or keyframes mean in Video model capability matrix?

The generation can be given an end frame, or frames at set points, so a shot lands somewhere specific. This is what makes a cut between two shots match.

https://shotcraft.video/reference/model-capabilities/ · 2026-08-17

## What does Longest single generation mean in Video model capability matrix?

Seconds in one call, before extension. Longer clips are usually assembled from several calls, and the joins are visible.

Models you run yourself are usually specified as a frame count and a frame rate rather than a duration, and the ceiling belongs to the inference pipeline rather than to the weights. Those rows are left not established unless the repository states a figure in seconds.

https://shotcraft.video/reference/model-capabilities/ · 2026-08-17

## What does Open weights mean in Video model capability matrix?

The weights are published, so the model can run on your own hardware. It does not mean you may use the output commercially. Read the licence.

https://shotcraft.video/reference/model-capabilities/ · 2026-08-17

## What does API mean in Video model capability matrix?

The vendor offers a first-party programmatic endpoint. A model marked no may still be reachable through a reseller.

https://shotcraft.video/reference/model-capabilities/ · 2026-08-17

## How do Conditioning and Structure conditioning differ?

Conditioning: Any input other than the initial noise that biases what the model generates, including text, images, masks, structure maps and camera trajectories. Structure conditioning: Conditioning on a spatial map derived from a reference, such as depth, edges or human pose, to fix geometry while leaving appearance to the model.

https://shotcraft.video/learn/concepts/conditioning-and-control/

## Explain why this claim is wrong: the model remembers the character from the last generation.

It has no persistent memory of your cast or set between generations; continuity comes from what you re-supply as conditioning each time.

https://shotcraft.video/learn/concepts/character-consistency/

## Explain why this claim is wrong: one good reference image is enough.

Single references pin a look from one angle and one light; identity fails first on profile, extreme angles, and lighting far from the reference, which is why fine-tuning methods use small sets rather than single frames.

https://shotcraft.video/learn/concepts/character-consistency/

## Explain why this claim is wrong: a DreamBooth-style fine-tune is strictly better than an adapter.

A fine-tune changes the weights and can degrade everything else the model knew, which is why DreamBooth carries a prior-preservation loss; an adapter such as IP-Adapter leaves the base model frozen and stays compatible with other controls.

https://shotcraft.video/learn/concepts/character-consistency/

## Explain why this claim is wrong: consistency means the face.

Wardrobe detail, set dressing, prop position, hair, light direction and eyeline all drift independently, and audiences read a change of light direction across a cut as fast as they read a change of face.

https://shotcraft.video/learn/concepts/character-consistency/

## Explain why this claim is wrong: this is a solved problem in current models.

Cross-shot continuity remains the weakest part of the pipeline and is currently bought with references, fine-tunes and human selection rather than with model memory.

https://shotcraft.video/learn/concepts/character-consistency/

## Explain why this claim is wrong: a good enough prompt can specify anything.

In practice text is one conditioning channel competing with image, structure and motion inputs, and spatial precision comes from the image-side overlays, not from more adjectives.

https://shotcraft.video/learn/concepts/conditioning-and-control/

## Explain why this claim is wrong: image-to-video means the model keeps your image intact.

It uses the frame as a starting condition and re-synthesises everything after it, so wardrobe, set dressing and background text drift as the clip runs.

https://shotcraft.video/learn/concepts/conditioning-and-control/

## Explain why this claim is wrong: ControlNet-style depth and pose conditioning give you camera and lighting control.

They constrain where surfaces and limbs are; the lens character, exposure and light direction still come out of the model's learned appearance.

https://shotcraft.video/learn/concepts/conditioning-and-control/

## Explain why this claim is wrong: stacking more controls always helps.

Multiple conditionings can contradict each other, and the standard advice with multi-ControlNet setups is to mask the inputs so they do not overlap and to tune each conditioning weight separately.

https://shotcraft.video/learn/concepts/conditioning-and-control/

## Explain why this claim is wrong: inpainting a face fixes it.

Inpainting resamples inside the mask with the same priors that produced the error, so the usual fix is a better reference or a different seed rather than a repeated local repair.

https://shotcraft.video/learn/concepts/conditioning-and-control/

## Explain why this claim is wrong: a LoRA is a smaller model.

It is a set of additional low-rank weights that only works on top of the specific base model it was trained against, so the base checkpoint is part of the deliverable.

https://shotcraft.video/learn/concepts/loras-and-finetuning/

## Explain why this claim is wrong: more training images always give a better LoRA.

A small, clean, varied set beats a large set with repeated backgrounds or inconsistent lighting, because the model will happily learn the wallpaper along with the face.

https://shotcraft.video/learn/concepts/loras-and-finetuning/

## Explain why this claim is wrong: full fine-tuning is the quality ceiling and LoRA is the compromise.

The original LoRA work reports performance on par with or better than full fine-tuning at a fraction of the trainable parameters, and adapters have the practical advantage of composing with other controls.

https://shotcraft.video/learn/concepts/loras-and-finetuning/

## Explain why this claim is wrong: if the LoRA works it will work at any weight.

Adapter strength is a dial with a usable window; too low and identity slips, too high and the adapter overrides prompt, composition and camera behaviour.

https://shotcraft.video/learn/concepts/loras-and-finetuning/

## Explain why this claim is wrong: overfitting shows up as bad images.

It usually shows up as loss of control: the trigger word stops being optional, the poses repeat, and the model can no longer place the subject anywhere it did not see during training.

https://shotcraft.video/learn/concepts/loras-and-finetuning/

## Explain why this claim is wrong: flicker is a bug awaiting a patch.

It follows from the model re-deriving detail it never committed to. Scale and architecture have reduced it considerably and have not removed it, and benchmarks still measure it as a standing dimension.

https://shotcraft.video/learn/concepts/temporal-consistency/

## Explain why this claim is wrong: the model tracks characters and objects from frame to frame.

It keeps frames mutually plausible. Identity is not stored, which is why a face can drift inside one shot and will not match across shots by default.

https://shotcraft.video/learn/concepts/temporal-consistency/

## Explain why this claim is wrong: higher resolution reduces morphing.

Resolution and temporal stability are separate axes, and a crisper frame often makes crawl and drift easier to see.

https://shotcraft.video/learn/concepts/temporal-consistency/

## Explain why this claim is wrong: interpolating between keyframes gives consistent motion.

It gives smooth motion. Lumiere's authors argue that generating distant keyframes and filling between them makes global consistency harder, which is why they generate the whole duration at once.

https://shotcraft.video/learn/concepts/temporal-consistency/

## Explain why this claim is wrong: continuity within a shot and continuity across shots are the same problem.

Within a shot, attention between latents does real work. Across shots nothing is shared, so matching has to be bought with references, fine-tunes and selection.

https://shotcraft.video/learn/concepts/temporal-consistency/

## Explain why this claim is wrong: video costs about as much as an image times the frame count.

Tokens span space and time and attention cost grows faster than linearly with sequence length, so doubling duration costs more than double.

https://shotcraft.video/learn/concepts/video-economics/

## Explain why this claim is wrong: the model generates frame by frame, so a longer clip is just more of the same work.

Sampling is iterative over the whole clip at once, and every denoising step processes the full spacetime latent.

https://shotcraft.video/learn/concepts/video-economics/

## Explain why this claim is wrong: compute is the dominant cost of a generative sequence.

Human selection is usually the larger line, because someone must watch and reject the majority of generations and that time does not scale down.

https://shotcraft.video/learn/concepts/video-economics/

## Explain why this claim is wrong: generating at final resolution from the start saves a step.

The standard economy is to prototype small and finish large, which the Sora report describes as an explicit benefit of training on native sizes.

https://shotcraft.video/learn/concepts/video-economics/

## Explain why this claim is wrong: more compression is free efficiency.

LTX-Video is explicit that its high compression ratio inherently limits the representation of fine detail, which is why the model puts effort back into the decoder.

https://shotcraft.video/learn/concepts/video-economics/

# 8. Sound

## What does the ElevenLabs v3 audio tags guide say about The voice you choose limits what a tag can do?

ElevenLabs states that the voice and its training samples affect tag effectiveness, and gives the example that a whispering voice will not suddenly shout because you added a shout tag. Tag choice follows voice choice, not the other way round.

documented by the vendor

https://shotcraft.video/prompts/audio-tags/ · 2026-08-17

## What does the ElevenLabs v3 audio tags guide say about Some tags are less consistent than others?

The documentation describes experimental tags as less consistent across voices and advises testing thoroughly before production use. Treat an unusual tag as something to audition, not something to schedule around.

documented by the vendor

https://shotcraft.video/prompts/audio-tags/ · 2026-08-17

## What does the ElevenLabs v3 audio tags guide say about Punctuation does work that tags cannot?

Ellipses add pauses and weight, and capitalisation increases emphasis. These operate alongside tags, so a line often needs fewer tags than it first appears to.

documented by the vendor

https://shotcraft.video/prompts/audio-tags/ · 2026-08-17

## What does the ElevenLabs v3 audio tags guide say about One or two per line, at the point of effect?

The documented examples place one or two tags per phrase rather than stacking them. A tag applies from where it sits, so its position in the line is part of the instruction.

documented by the vendor

https://shotcraft.video/prompts/audio-tags/ · 2026-08-17

## What does the Suno structure tags guide say about They are hints, and the model may ignore them?

Every account of these tags, including the ones that document them most thoroughly, describes them as probabilistic. Suno follows them much of the time and sometimes does not, so a structure tag is not a guarantee of an arrangement.

community convention, no official list found

https://shotcraft.video/prompts/audio-tags/ · 2026-08-17

## What does the Suno structure tags guide say about A section tag goes on its own line?

Placed above the lyrics for that section rather than inline, and repeated wherever the section recurs, so a chorus that comes back three times carries the tag three times.

community convention, no official list found

https://shotcraft.video/prompts/audio-tags/ · 2026-08-17

## What does the Suno structure tags guide say about This vocabulary is not verified against a vendor document?

Treat the exact spellings as convention. If a tag does nothing, the likeliest explanation is that it was never a documented tag, not that the model failed.

community convention, no official list found

https://shotcraft.video/prompts/audio-tags/ · 2026-08-17

## How would you use this principle: Design for what the scene means, not for what is on screen?

an effect's job is the audience's feeling about an event, not an inventory of it. A correctly recorded real punch sounds like almost nothing, which is why nobody uses one.

https://shotcraft.video/learn/crafts/sound-designer/

## How would you use this principle: Sound is the cheapest way to build the world outside the frame?

what an audience hears off screen sets how big the world is. One distant train can do more for a location than a set extension.

https://shotcraft.video/learn/crafts/sound-designer/

## How would you use this principle: Contrast is the only loud there is?

nothing sounds big unless something before it was small. Design the quiet and the drop-out first, then place the impact into the space you made.

https://shotcraft.video/learn/crafts/sound-designer/

## How would you use this principle: Layer by role, not by count?

every element in a stack should be doing a specific job, whether that is attack, body, texture or room. Piling on more of the same makes mud, and a mixer can only turn mud down.

https://shotcraft.video/learn/crafts/sound-designer/

## How would you use this principle: Leave the mixer room?

ten overlapping atmosphere tracks with no plan is not detail, it is a problem passed downstream. Deliver organised pre-dubs with headroom and legible intent.

https://shotcraft.video/learn/crafts/sound-designer/

## How would you use this principle: Everything competes with the voice?

build effects and backgrounds with the dialogue's frequency range in mind, or they will simply be pulled down at the dub and the work will be inaudible.

https://shotcraft.video/learn/crafts/sound-designer/

## Which items make up ElevenLabs v3 audio tags / Voice and delivery / 1?

[laughs], [laughs harder], [starts laughing], [wheezing], [whispers]

documented by the vendor

https://shotcraft.video/prompts/audio-tags/ · 2026-08-17

## Which items make up ElevenLabs v3 audio tags / Voice and delivery / 2?

[sighs], [exhales], [sarcastic], [curious], [excited]

documented by the vendor

https://shotcraft.video/prompts/audio-tags/ · 2026-08-17

## Which items make up ElevenLabs v3 audio tags / Voice and delivery / 3?

[crying], [snorts], [mischievously]

documented by the vendor

https://shotcraft.video/prompts/audio-tags/ · 2026-08-17

## Which items make up ElevenLabs v3 audio tags / Sound effects / 1?

[gunshot], [applause], [clapping], [explosion], [swallows]

documented by the vendor

https://shotcraft.video/prompts/audio-tags/ · 2026-08-17

## Which items make up ElevenLabs v3 audio tags / Sound effects / 2?

[gulps]

documented by the vendor

https://shotcraft.video/prompts/audio-tags/ · 2026-08-17

## Which items make up ElevenLabs v3 audio tags / Unique and special / 1?

[strong X accent], [sings], [woo], [fart]

documented by the vendor

https://shotcraft.video/prompts/audio-tags/ · 2026-08-17

## Which items make up Suno structure tags / Song sections / 1?

[Intro], [Verse], [Verse 1], [Verse 2], [Pre-Chorus]

community convention, no official list found

https://shotcraft.video/prompts/audio-tags/ · 2026-08-17

## Which items make up Suno structure tags / Song sections / 2?

[Chorus], [Hook], [Bridge], [Outro], [End]

community convention, no official list found

https://shotcraft.video/prompts/audio-tags/ · 2026-08-17

## Which items make up Suno structure tags / Instrumental and dynamics / 1?

[Instrumental], [Instrumental Break], [Break], [Interlude], [Drop]

community convention, no official list found

https://shotcraft.video/prompts/audio-tags/ · 2026-08-17

## Which items make up Suno structure tags / Instrumental and dynamics / 2?

[Build-Up], [Breakdown]

community convention, no official list found

https://shotcraft.video/prompts/audio-tags/ · 2026-08-17

# 9. Delivery

## What does Super-resolution mean in Upscaling and interpolation: finishing at size?

Increasing pixel dimensions by predicting the detail a higher-resolution version of the image would plausibly contain.

https://shotcraft.video/learn/concepts/upscaling-and-interpolation/

## What does Blind super-resolution mean in Upscaling and interpolation: finishing at size?

Upscaling when the degradation that damaged the source is unknown, trained on synthetic degradation pipelines instead.

https://shotcraft.video/learn/concepts/upscaling-and-interpolation/

## What does Detail hallucination mean in Upscaling and interpolation: finishing at size?

Invented high-frequency content that looks correct in isolation but does not correspond to anything in the source.

https://shotcraft.video/learn/concepts/upscaling-and-interpolation/

## What does Frame interpolation mean in Upscaling and interpolation: finishing at size?

Synthesising frames between existing ones, usually by estimating intermediate optical flow, to raise frame rate or slow motion.

https://shotcraft.video/learn/concepts/upscaling-and-interpolation/

## What does Cadence mean in Upscaling and interpolation: finishing at size?

The rhythm of motion set by frame rate and shutter, which stops matching the picture when frames are synthesised after the fact.

https://shotcraft.video/learn/concepts/upscaling-and-interpolation/

## How would you use this principle: The offline is a description, not a source?

Every event in it is a claim about what should happen, and each claim has to be verified against the original. Trusting a relink because it looked right at proxy resolution is how a wrong take reaches a master.

https://shotcraft.video/learn/crafts/online-conform-editor/

## How would you use this principle: Compare, do not remember?

Put the offline and the conform on top of each other in difference mode and step through. The eye is unreliable across two viewings and completely reliable at spotting a difference matte that is not black.

https://shotcraft.video/learn/crafts/online-conform-editor/

## How would you use this principle: Preserve the data, deliver the display?

The version approved in the room is display-referred and standard-specific. An archival master rendered without the output transform is what allows the show to be remastered later without regrading it.

https://shotcraft.video/learn/crafts/online-conform-editor/

## How would you use this principle: Every deliverable is a contract with numbers in it?

Frame rate, active picture, audio channel order, loudness and wrapper are not preferences. Read the spec for the specific platform before rendering, not after QC fails.

https://shotcraft.video/learn/crafts/online-conform-editor/

## How would you use this principle: Fix at the right altitude?

A tracked patch in finishing is cheaper than a vendor round-trip, and a shot that needs a real rebuild is more expensive to patch badly twice than to send out once.

https://shotcraft.video/learn/crafts/online-conform-editor/

## In the dated TikTok feed (In-Feed / For You) guide, what bottom reserve is recorded for the 1080x1920 working canvas?

660 px

The safe area is not a plain rectangle. Above the action rail the right margin is only 120px (11.11%); the 300px (27.78%) right margin applies only across the vertical band the rail occupies, measured at y 840-1260 on a 1080x1920 frame. The conservative rectangle is x 120-780, y 240-1260 = 660x1020px. official-template-measured

TikTok's stated minimum for vertical is 540x960 (9:16); it also accepts 1:1 (>=640x640) and 16:9 (>=960x540), but non-vertical assets are letterboxed and lose a large amount of usable frame (see crops). TikTok's own safe zone templates are authored on 720x1280 (standard) and 540x960 (with-anchor) canvases; all values below are scaled to 1080x1920. TikTok publishes no resolution 'recommendation' beyond the minimum, so 1080x1920 is a production convention, not a TikTok spec.

https://ads.tiktok.com/help/article/tiktok-auction-in-feed-ads?lang=en · 2026-08-14

## In the dated Instagram Reels guide, what bottom reserve is recorded for the 1440x2560 (Meta's ads guide recommendation); 1080x1920 is the common production canvas working canvas?

672 px

Meta expresses this only as percentages, so it scales cleanly. On 1080x1920: top 269, bottom 672, sides 65. On Meta's own recommended 1440x2560: top 358, bottom 896, sides 86. Usable box on 1080x1920: 950x979px at x 65-1015, y 269-1248. official

Meta's ads guide for the Instagram Reels placement specifies 1440x2560 and 9:16, with 1% aspect ratio tolerance, min width 250px (500px for assets over 30s), max 4GB, MP4/MOV. Instagram's consumer help center states organic Reels accept '1.91:1 or 9:16' with a 720px minimum and 30fps minimum.

https://www.facebook.com/business/ads-guide/update/video/instagram-reels?locale=en_US · 2026-08-14

## In the dated YouTube Shorts guide, what bottom reserve is recorded for the 1080x1920 working canvas?

672 px

These are Google's own pixel figures, stated on a 1080x1920 reference frame - the only platform in this set that publishes actual pixels. Usable box: 840x960px at x 48-888, y 288-1248. Note the pronounced left-right asymmetry: 48px left against 192px right. official

1080x1920 is stated on Google's own safe zone diagram for vertical video ads. YouTube's Shorts classification is broader than 9:16: per YouTube Help, videos uploaded on or after 15 October 2024 with a square OR vertical aspect ratio and up to three minutes are categorised as Shorts. Google Ads Help adds that horizontal assets 'serve with blurred top and bottoms in the vertical Shorts experience', and that aspect ratios between 16:9 and 9:16 are allowed with the player adapting. So Shorts accepts 1:1 through 9:16, but only 9:16 uses the whole canvas.

https://storage.googleapis.com/support-kms-prod/jbb4hOyc2nGdunbKddqpFAEQpHOQvy582jIR · 2026-08-14

## In the dated Instagram Stories guide, what bottom reserve is recorded for the 1440x2560 (Meta's ads guide recommendation); 1080x1920 in common use working canvas?

672 px

Identical percentages to Instagram Reels. On 1080x1920: top 269, bottom 672, sides 65. On 1440x2560: top 358, bottom 896, sides 86. official

Meta's ads guide Instagram Stories video placement: Ratio 9:16, Resolution 1440 x 2560, aspect ratio tolerance 1%, min width 250px, duration 1s to 60 minutes, max 4GB, MP4/MOV/GIF.

https://www.facebook.com/business/ads-guide/update/video/instagram-story?locale=en_US · 2026-08-14

## In the dated Facebook Reels guide, what bottom reserve is recorded for the 1440x2560 working canvas?

672 px

Same percentages as Instagram Reels and Instagram Stories. On 1080x1920: top 269, bottom 672, sides 65. On 1440x2560: top 358, bottom 896, sides 86. official

Meta's ads guide, Facebook Reels video placement: Resolution 1440 x 2560 pixels, Ratio 9:16, MP4/MOV/GIF, H.264, square pixels, fixed frame rate, progressive scan, stereo AAC 128kbps+. Distinct from the separate 'Ads on Facebook Reels' OVERLAY placement, which is 1:1 at 1440x1440 - a different format that sits on top of someone else's reel.

https://www.facebook.com/business/ads-guide/update/video/facebook-facebook-reels?locale=en_US · 2026-08-14

## In the dated Facebook Stories guide, what bottom reserve is recorded for the 1440x2560 working canvas?

340 px

Meta states this one as a percentage AND a pixel pair - '14% (250 pixels)' and '20% (340 pixels)' - and the two do not agree with each other: 250/1920 = 13.0% and 340/1920 = 17.7%. Both readings are recorded rather than reconciled. Meta gives NO side figure at all for Facebook Stories, unlike the other three placements. disputed

Meta's ads guide, Facebook Stories video placement: Resolution 1440 x 2560 pixels, Ratio 9:16, MP4/MOV/GIF, duration 1 second to 3 minutes, max 4GB.

https://www.facebook.com/business/ads-guide/update/video/facebook-story?locale=en_US · 2026-08-14

## In the dated Snapchat (Spotlight, Stories, Commercials - one shared spec) guide, what bottom reserve is recorded for the 720x1280 stated by Snap; 1080x1920 in common use working canvas?

672 px

Snap publishes percentages on a diagram, not pixels; the pixel column above is those percentages applied to 1080x1920. The diagram also marks a DEEPER bottom bracket of 45% (864px on a 1920 frame) alongside the 35% one, without labelling which formats each applies to. The right-hand 15% applies only across the lower portion where the action rail sits - the diagram's red safe rectangle is a stepped shape, full width to 6% above the rail and stepping in to 15% across it. Usable box on 1080x1920, taking 35% bottom: 953x1056px at x 65-1018, y 192-1248; taking 15% right across the rail band, 853px wide there. official

Snap's own ad format page states Aspect ratio 9:16 and Resolution 720x1280 for Single Image or Video ads, Story Ads and Sponsored Snaps, with .mp4/.mov/.jpg/.png accepted. 720x1280 is Snap's stated figure, not a minimum-versus-recommended distinction - Snap does not publish a 1080x1920 recommendation.

https://businesshelp.snapchat.com/s/article/safe-zones · 2026-08-14

## Explain why this claim is wrong: upscaling recovers detail.

It reconstructs plausible detail using priors learned from training degradations, which is restoration where the information is present and invention where it is not.

https://shotcraft.video/learn/concepts/upscaling-and-interpolation/

## Explain why this claim is wrong: a 4K upscale is a 4K deliverable.

It meets a pixel count; it does not meet the detail, edge behaviour or grading latitude of footage acquired at that size.

https://shotcraft.video/learn/concepts/upscaling-and-interpolation/

## Explain why this claim is wrong: interpolation makes motion look better.

It makes motion smoother, which is a different thing. Frames synthesised between long-exposure frames keep the original blur, so cadence and blur stop agreeing.

https://shotcraft.video/learn/concepts/upscaling-and-interpolation/

## Explain why this claim is wrong: interpolation and upscaling are interchangeable finishing steps in any order.

Temporal work before spatial work generally holds up better, because upscalers amplify the flow artifacts interpolation leaves at occlusion edges.

https://shotcraft.video/learn/concepts/upscaling-and-interpolation/

## Explain why this claim is wrong: more finishing passes always improve the frame.

Stacked super-resolution and sharpening compound ringing and overshoot, the exact artifacts a blind super-resolution model such as Real-ESRGAN is trained to suppress rather than to create.

https://shotcraft.video/learn/concepts/upscaling-and-interpolation/

# 10. Judgement

## What does Open weights mean in Open weights versus closed API: renting or owning the model?

A model whose trained parameters can be downloaded and run locally, regardless of whether its licence meets any open-source standard.

https://shotcraft.video/learn/concepts/open-vs-closed-weights/

## What does Open Source AI Definition mean in Open weights versus closed API: renting or owning the model?

The Open Source Initiative's 1.0 standard, which asks for information about training data, the training and inference code, and the weights.

https://shotcraft.video/learn/concepts/open-vs-closed-weights/

## What does Version pinning mean in Open weights versus closed API: renting or owning the model?

Locking a production to one exact model build so results stay stable across a revision cycle.

https://shotcraft.video/learn/concepts/open-vs-closed-weights/

## What does On-premises inference mean in Open weights versus closed API: renting or owning the model?

Running generation on hardware you control so confidential material never leaves your network.

https://shotcraft.video/learn/concepts/open-vs-closed-weights/

## What does Community licence mean in Open weights versus closed API: renting or owning the model?

A vendor-written licence that permits broad use with carve-outs, such as territorial exclusions or separate commercial terms, and is not an open-source licence.

https://shotcraft.video/learn/concepts/open-vs-closed-weights/

## What does Fair use mean in Training data and copyright: what is not settled?

A US statutory defence that can excuse otherwise infringing copying based on a multi-factor analysis, applied case by case rather than as a blanket rule.

https://shotcraft.video/learn/concepts/training-data-and-copyright/

## What does Fair dealing mean in Training data and copyright: what is not settled?

The narrower UK and Commonwealth equivalent, available only for specified purposes rather than as an open-ended balancing test.

https://shotcraft.video/learn/concepts/training-data-and-copyright/

## What does Provenance record mean in Training data and copyright: what is not settled?

The log of model, model version, prompt, seed, parameters and input references behind a generated shot, kept so a delivery can be explained later.

https://shotcraft.video/learn/concepts/training-data-and-copyright/

## What does Indemnification mean in Training data and copyright: what is not settled?

A vendor's contractual promise to cover defined legal costs arising from use of its outputs, usually subject to caps, conditions and correct use of the product.

https://shotcraft.video/learn/concepts/training-data-and-copyright/

## What does Multidistrict litigation mean in Training data and copyright: what is not settled?

A US procedure that consolidates related federal cases before one judge for pretrial purposes, which is why several separate complaints can appear under a single docket.

https://shotcraft.video/learn/concepts/training-data-and-copyright/

## What does Object permanence mean in Why models fail: hands, text, physics and permanence?

Whether things persist correctly through occlusion and after leaving frame, which current models hold often but not reliably.

https://shotcraft.video/learn/concepts/why-models-fail/

## What does Object state change mean in Why models fail: hands, text, physics and permanence?

The correct transformation of a thing under an action, such as food being eaten, which the Sora report lists as unreliable.

https://shotcraft.video/learn/concepts/why-models-fail/

## What does Out-of-distribution generalisation mean in Why models fail: hands, text, physics and permanence?

Performance on cases unlike anything in training, where video models have been shown to fail rather than extrapolate a rule.

https://shotcraft.video/learn/concepts/why-models-fail/

## What does Case-based generalisation mean in Why models fail: hands, text, physics and permanence?

Producing an answer by imitating the closest training example instead of applying an abstracted rule.

https://shotcraft.video/learn/concepts/why-models-fail/

## What does Temporal flickering mean in Why models fail: hands, text, physics and permanence?

Frame-to-frame instability in texture, colour or detail, tracked as its own evaluation dimension in VBench.

https://shotcraft.video/learn/concepts/why-models-fail/

## Explain why this claim is wrong: open weights means open source.

The Open Source Initiative's definition requires information about the training data, the training and inference code, and the weights. Most downloadable video models ship under custom licences that do not satisfy it.

https://shotcraft.video/learn/concepts/open-vs-closed-weights/

## Explain why this claim is wrong: open means free to use commercially anywhere.

HunyuanVideo's community licence states it does not apply in the European Union, the United Kingdom or South Korea, and Stable Video Diffusion's community licence directs commercial use to separate Stability terms. Licences differ per model and sometimes per model size.

https://shotcraft.video/learn/concepts/open-vs-closed-weights/

## Explain why this claim is wrong: the closed model is always better.

Hosted flagships often lead on quality, but the gap moves with every release, and quality is one axis alongside control, confidentiality, reproducibility and cost.

https://shotcraft.video/learn/concepts/open-vs-closed-weights/

## Explain why this claim is wrong: pinning the model name in an API call gives identical results forever.

Hosted endpoints are updated underneath the name. Only weights you hold can be genuinely pinned.

https://shotcraft.video/learn/concepts/open-vs-closed-weights/

## Explain why this claim is wrong: self-hosting is cheaper.

You swap a per-generation charge for GPUs, storage, engineering time and idle capacity. It pays off at volume and on confidential work, not automatically.

https://shotcraft.video/learn/concepts/open-vs-closed-weights/

## Explain why this claim is wrong: a court has already ruled that training on copyrighted work is fair use.

No US appellate court has ruled on training fair use. Thomson Reuters v. Ross was argued in the Third Circuit in June 2026 and remains undecided, and the most-discussed district-court analysis, in Bartz v. Anthropic, ended in a settlement approved in July 2026, so it was never appealed and binds nobody.

https://shotcraft.video/learn/concepts/training-data-and-copyright/

## Explain why this claim is wrong: Getty lost in the UK, so training is lawful there.

Getty abandoned its training claim mid-trial for want of evidence that training took place in the UK, and the court made no finding on how many works were used. That is a narrow judgment about territoriality and evidence, not a ruling that training is lawful.

https://shotcraft.video/learn/concepts/training-data-and-copyright/

## Explain why this claim is wrong: if the output does not resemble any single training image, there is nothing to argue about.

Output similarity and the copying performed to build the training set are separate questions, and both are being litigated.

https://shotcraft.video/learn/concepts/training-data-and-copyright/

## Explain why this claim is wrong: open weights means the training data was licensed.

A licence on the weights says nothing about the provenance of the data used to produce them.

https://shotcraft.video/learn/concepts/training-data-and-copyright/

## Explain why this claim is wrong: a model trained on licensed data removes the risk.

It narrows one risk. Indemnities carry caps and conditions, and prompting for a protected character or a trade mark can still produce an output problem.

https://shotcraft.video/learn/concepts/training-data-and-copyright/

## Explain why this claim is wrong: these are bugs awaiting a patch.

The Sora technical report presents physical-interaction and object-state errors as current limitations of the approach, alongside long-sample incoherence and objects appearing spontaneously.

https://shotcraft.video/learn/concepts/why-models-fail/

## Explain why this claim is wrong: a more realistic-looking model has better physics.

Physics-IQ reports that physical understanding across current models is severely limited and unrelated to visual realism, so photorealism is no guide to whether an interaction will be correct.

https://shotcraft.video/learn/concepts/why-models-fail/

## Explain why this claim is wrong: scaling will fix physics.

A study on a controlled mechanics testbed found in-distribution generalisation but out-of-distribution failure, and concluded that scaling alone is not sufficient for these models to recover fundamental physical laws.

https://shotcraft.video/learn/concepts/why-models-fail/

## Explain why this claim is wrong: hands and text are the whole problem.

Hands and small text are the visible tip; continuity across shots, weight, contact and object permanence fail more often and are harder to spot in a single viewing.

https://shotcraft.video/learn/concepts/why-models-fail/

## Explain why this claim is wrong: one quality score tells you whether a model suits your shot.

VBench exists because a single number hides the trade-offs; models differ dimension by dimension, so the useful question is how a model scores on the dimension your shot depends on.

https://shotcraft.video/learn/concepts/why-models-fail/

Progress can be exported to a file and imported again. With no server, that file is the only real guarantee — browsers do clear storage.