

Meta launches Muse Image model with agentic capabilities

Source copy kept by wheresthe.ai for the story "Meta launches Muse Image model with agentic capabilities" (<https://wheresthe.ai/d/us-meta-launches-muse-image-model-with-agentic-capabilities-2026-07/>).

Source: <https://ai.meta.com/blog/introducing-muse-image-muse-video-msl>

Text saved 24 Oct 2026 (the page text through a reader, because the site refused a direct fetch).

SHA-256 of the saved text: f5b038c3a4a228dc7efdfdb3bd1a92f92f98210e0013488c423ac72e6a456fe3

* Products * AI Research * Resources * About * AI Developers

* Try Muse *

BACK

* Muse * Meta AI * Media Generation * Vibes

* Overview * Projects * Resources & Tools * Publications * GitHub

* Blog * Learning Hub * Demos

* Overview * Open Source * Careers

Clear

* Clear * Products > * AI Research > * Resources > * About > * AI Developers

Try Muse

FEATURED

Research

Introducing Muse Image and Muse Video

July 7, 2026•

15 minute read

Video 27

Video 28

Video 29

Video 30

Video 31

Video 32

Video 33

Video 34

Video 35

Video 36

Video 37

Video 38

We're excited to launch Muse Image and preview Muse Video, the first media generation models developed by Meta Superintelligence Labs.

Muse Image is our most advanced image generation model yet: it follows instructions faithfully, edits with precision, and composes from multiple references. It also brings agentic tool use capabilities and integrates with

Muse Spark. Muse Video, built on the same pretraining base, delivers exceptional visual fidelity with native audio support.

Muse Image is available today across the Meta AI app and on meta.ai, Instagram Stories in the US, and WhatsApp in limited countries, and is coming soon to Facebook. Muse Video is coming soon to creators and Meta AI.

Muse Image: Agentic Image Generation

Instead of directly mapping prompts to images, Muse Image operates as an agent: it invokes search and coding tools to improve accuracy, self-refines its own generations, and improves through scaling test-time compute. Muse Image also integrates with Muse Spark, allowing the two models to share tools and plan jointly for powerful agentic media generation.

Tool Use

We provide Muse Image with access to tools to enhance its agentic capabilities.

****Coding.**** During reinforcement learning, Muse Image learns to write and execute code that produces accurate plots and QR codes, and condition on rendered figures to improve the accuracy of generated images. Muse Spark and Muse Image also integrate to use the combination of code and media generation to create animated GIFs, websites with embedded images, and interactive visual games.

Conference QR Code

Preparing the artwork

I'm setting up a landscape, Korean manhwa-style scene with an accurate QR code, checking composition and linework so the final image matches the user's request.

Conference QR Code

Generating QR code

I'm creating a scannable QR code for <https://meta.ai>, checking the encoding and error correction so it reads reliably before I place it in the scene.

Conference QR Code

Checking QR code image

I'm opening the generated QR code file to confirm it scans correctly, then I'm preparing to place it on the poster and set up the Korean manhwa scene with the user's subject and conference details.

Conference QR Code

Composing manhwa scene

I'm assembling a 3:2 landscape illustration in Korean manhwa style, placing the young woman scanning the QR code at the ICML 2025 poster, adding the conference backdrop, lighting, and details to match the brief.

Conference QR Code

Adjusting gaze to phone

I'm refining the illustration so the woman looks down at her phone screen instead of the QR code, updating her pose and expression while keeping the manhwa style, conference backdrop, and accurate QR details intact.

✓

Finished

Open Thread

Fractals Plot

Generating fractal base

I'm coding two panels, a Julia set and an IFS Sierpinski triangle, then merging them into one clean PNG so I can

apply the mid-century Swiss style next.

Fractals Plot

Rendering fractal base

I'm running the Python routine to compute the Julia set and Sierpinski points, then assembling them into a clean two-panel image and saving it as the base file for the next step.

Fractals Plot

Planning poster layout

I'm mapping a strict asymmetrical grid with generous white space, placing the two fractals and left-aligned labels, and choosing a flat red, blue, and black palette on warm off-white for a clean, functional modernist feel.

Fractals Plot

Crafting poster composition

I'm arranging the fractals on a warm off-white field with a strict grid, adding red and blue geometric accents, and placing left-aligned labels and captions in clean sans-serif type for a balanced, modernist layout.

✓

Finished

Open Thread

Flipbook Animation

Refining character details

I'm refining the two fighters' looks and poses, matching the blue-haired male's gi and the red-ponytailed female's outfit while keeping the dojo lighting and camera angle consistent for the next frame. I'm polishing the fighters' looks and stances, tightening the dojo's lighting and textures, and double-checking the composition so everything feels cohesive before I finalize the scene.

Flipbook Animation

Adjusting fighter positions

I'm repositioning both fighters closer together in the dojo, keeping the background, lighting, and outfits identical while shifting their stances forward so they're nearly face to face in a tighter, more dynamic confrontation.

Flipbook Animation

Generating punch frame

I'm creating the next image where the blue-haired male throws a quick right jab, arm extended forward, while the red-ponytail female holds her defensive stance. I'm keeping the dojo, lighting, and all background details exactly the same.

Flipbook Animation

Generating dodge frame

I'm creating the next frame where the female dodges left as the male's punch whiffs past, keeping the dojo background, lighting, and character details consistent while adjusting poses and motion for the action.

✓

Finished

Open Thread

Open Artifact

Pet Game

Crafting teen cat pose

I'm turning the cream Persian into a lanky, gangly teenager, leaning forward with one paw raised and a mischievous smirk. I'm keeping the fluffy beige-cream coat and warm coral gradient while shaping the playful, energetic Pixar-style scene.

Pet Game

Refining the portrait

I'm adding subtle silver-gray highlights around the muzzle and cheeks, adjusting the pose to a regal upright sit, and shifting the background to a muted burgundy gradient with warm golden lighting for a distinguished, calm look.

Pet Game

Creating newborn kitten

I'm turning the cream Persian into a tiny newborn kitten, curled up with barely open golden-amber eyes and soft baby-fine fur against a warm buttercream gradient, in a cozy Pixar-style illustration.

Pet Game

Encoding the cat images

I'm converting the six cat photos into base64 data URIs so they can be embedded directly in the HTML, ensuring the game loads instantly without external files.

Pet Game

Verifying the checklist

I'm reviewing the self-contained file to confirm the six embedded images, controls, animations, win and lose screens, scoring, and responsive layout all work as specified before I prepare the delivery note.

✓

Done

Open Thread

****Search.**** Muse Image learns to search the web to ground generated images in factual and real-time information and visual references. Enabling search improves factual accuracy on knowledge-intensive prompts, particularly those involving current events and real-world facts.

Muse Image improves with search tool use. Win rate from internal ablation.

Summer 2026 trends

Researching summer trends

I'm looking up the hottest women's summer 2026 looks, scanning runway reports and style guides to spot three distinct trends I can translate into outfits for the user.

Summer 2026 trends

Searching product catalogues

I'm scanning product catalogues for linen dresses, cinched-waist styles and buttery-cream shades, matching the summer 2026 trends I've noted so I can pull shoppable links for the user's first look.

Summer 2026 trends

Finding product matches

Summer 2026 trends

Running code

✓

Finished

Open Thread

Moon Formation

Researching formation diagrams

I'm looking up existing scientific diagrams and reputable sources on the giant impact hypothesis to gather accurate facts and visual inspiration for the infographic's stages and details.

Moon Formation

Searching for moon formation facts

Moon Formation

Reviewing impact hypothesis diagrams

Moon Formation

Drafting infographic panels

I'm mapping the vertical sequence of proto-Earth, Theia, and the Moon, weaving in collision timing, orbital dynamics, and the present-day evidence and counterpoints so the layout stays accurate and visually consistent.

Moon Formation

Composing infographic prompt

I'm drafting the exact prompt for the image generator, specifying six vertically stacked panels, photorealistic cinematic lighting, and precise text labels that describe each stage of the giant impact hypothesis.

Moon Formation

Finalizing infographic prompt

I'm polishing the full infographic prompt, stitching together the six panels and evidence boxes into a cohesive layout with captions, visual cues, and styling notes so the image generator can produce the complete scene.

✓

Finished

Open Thread

Redecorate with FB Marketplace items

Planning marketplace search

I'm mapping out a search for unbranded vintage pieces on Facebook Marketplace around San Francisco, focusing on warm, simple styles that fit the user's room and location.

Redecorate with FB Marketplace items

Loading shopping workflow

I'm loading the shopping skill to follow its room restyle workflow, using the user's uploaded bedroom photo and location to plan unbranded vintage finds from Facebook Marketplace that fit a warm, simple aesthetic.

Redecorate with FB Marketplace items

Preparing marketplace queries

I'm drafting marketplace queries for unbranded nightstands, dressers, lamps, rugs, and décor in San Francisco, matching the user's warm, simple vintage style with wood, rattan, brass, and linen accents.

Redecorate with FB Marketplace items

Searching marketplace listings

Redecorate with FB Marketplace items

Picking vintage pieces

I'm selecting six warm, simple vintage items from the marketplace results, balancing tone and budget across nightstand, dresser, rug, lamp, rattan chair, and autumn landscape art to fit the room's style.

✓

Finished

Open Thread

Self-Refinement

Muse Image reflects on and improves upon its own work within its chain of thought. This self-refining behavior can take different forms: a local edit to the current image draft when a small detail is off, a new image generation from scratch when larger parts are wrong, or a different tactic like tool use for more factually accurate generation. We didn't design this behavior. Instead, it emerged during RL training simply because self-refinement produced better images and therefore higher reward.

Muse Image improves with emergent self-refinement. Win rate from internal ablation.

Self-refinement

Finding reference images

Self-refinement

Assembling the spread

I'm putting together the glossy magazine page, weaving the proof steps, headline, and fashion portrait into a cohesive layout while checking typography, spacing, and the gold-accented details for a polished editorial feel.

Self-refinement

Refining the formula

I'm correcting the formula to include the missing division slash, updating the layout so the expression reads $S = n(n + 1) / 2$, and double-checking the surrounding text for clarity and consistency.

Self-refinement

Checking the image

I'm reviewing the generated magazine image to confirm the macro detail, glossy layout, and accurate proof before sharing it, and I'm ready to tweak the typography or swap the portrait if you'd like. What would you like to adjust next?

✓

Finished

Open Thread

Test-Time Compute Scaling

Like language models, Muse Image improves the more it thinks at inference time. With more test-time compute, the model reasons more, uses more tool calls, and uses more self-refinement steps to improve its generations. Increasing reasoning strength (and thus test-time compute) improves human-preference Elo scores and shows an approximately log-linear scaling relationship. Notably, this compute spans two very different kinds of work — text tokens for reasoning, visual tokens for generation — yet quality is a function of the combined total compute.

We find that using the token budget judiciously matters just as much for effective test-time scaling. Best-of-N (BoN), where the model generates several images and keeps the best, improves quality early but saturates quickly. Spending that same compute on deliberate reasoning scales considerably better. Reasoning and tool use

compound when combined. Tools let the model reach beyond what it already knows, whether by searching for references it lacks or writing code to get precise details right, filling gaps that reasoning alone can't.

Muse Image improves with scaling test-time compute. Elo from internal ablation.

Image Editing

Muse Image edits images with precision, changing exactly what the user asks for. It can follow a variety of instructions as our examples show.

Restore this image

[View thread](#)

Edit this to clear up the fog and reveal the beautiful valley below

[View thread](#)

Change the flower so the petals form a rainbow gradient

[View thread](#)

Can you change this image to be '\$3.00 ALL DAY', change the no free parking text to instead say 'FREE PARKING ON WEEKENDS' and change the phone number to be 555-5555

[View thread](#)

Zoom out slightly to share the chaos the dog has seen, giving context to its sheepish look

[View thread](#)

Muse Image maintains coherence across editing turns, supporting iterative refinement and open-ended brainstorming toward a target result.

Here is my living room Can you imagine it in Japandi style?

I like this image but please use the lamp and the cabinet under it from the first image.

Great. Can you make a before-after picture of my first image and the last one?

[View thread](#)

Make an image of this cat and this dog as best friends having a picnic on a sunny day. Vintage 35mm style

Now show this exact picnic image as a framed photo hanging on the wall of a cozy cafe. We see a table and two empty chairs right in front of this wall.

Show the front of the cafe with its name and match the vibe of the cafe from the earlier photo. We should be able to see the framed photo from the window too.

Design this cafe's paper menu using its exact name, and add a 'The Picnic Special' item with a tiny illustration of the same cat and dog from the photo.

Place this menu on the table from the close up empty table image we made earlier

[View thread](#)

Multi-Reference Image Composition

Muse Image can compose elements from many input reference images in the prompt, including people, objects, clothing, styles, and environments. It supports interleaving text and images inline in prompts for complex image compositions.

Make an image of riding this bike and wearing while passing by sitting on a park bench. Make it look like a drawing in the style of this image.

[View thread](#)

Change the color of my living room to this pattern on the wall. Then add on the coffee table, 2 on the credenza

and a large super flat wall mounted 65 inch tv, with the following image on the screen . Then add sitting on the couch wearing , and with no socks.

[View thread](#)

Image Benchmarks

Muse Image holds the No. 2 spot on Arena for text-to-image, single-image editing, and multi-image editing as measured by human preference Elo rankings at the time of writing.

Arena Elo rankings as of July 5, 2026.

Arena Elo rankings as of July 5, 2026.

Arena Elo rankings as of July 5, 2026.

Previewing Muse Video

Alongside the release of Muse Image, we're sharing an early preview of Muse Video. It offers competitive performance in prompt adherence, visual fidelity, and temporal consistency. We're investing in areas with current performance gaps, such as audio-video synchronization and physically accurate fast motion. Muse Video is coming soon to creators and in Meta AI.

Video 39

A baby panda tumbling head over heels down a small grassy slope

Video 40

A handmade paper-cutout / construction-paper collage stop-motion animation explaining Bernoulli's principle, calm warm male narrator with a clear natural voice and relaxed, well-enunciated pacing, 16:9. Layered textured cut paper, subtle paper grain, soft drop shadows, gentle flat lighting, warm pastel palette, slight stop-motion jitter, light paper-rustle foley, soft acoustic background music kept low under the voice. Visuals: a cut-paper title 'Bernoulli's Principle', then a proper cambered AIRFOIL cross-section ...

Video 41

first person point of view strolling along the edge of a small half-frozen pond in a snowy suburban park at night, snow-dusted reeds and cattails along the bank, fresh snow falling gently. Warm orange lamplight reflects softly on the dark open water. Serene and peaceful, with quiet nature sounds: the faint trickle of unfrozen water, a soft breeze through the reeds, distant snow-muffled stillness, and gentle footsteps in the snow. Natural handheld walking camera, cinematic, meditative.

Video 42

A man juggles three oranges, adds a fourth, drops them all, and takes a bow anyway. Warm natural morning light, gentle slow-motion. A single continuous ~10-second moment with a clear beginning, turn, and payoff. AUDIO: quiet room tone with crisp foley. Photorealistic, natural lighting and physics, believable real-world footage. Not a cartoon, not stylized.

Video 43

Handheld vertical 9:16 UGC-style ad: REC icon visible, a woman in her cluttered kitchen films herself with natural autofocus, slow-motion shot of steam curling off her Brivo kettle as she pours; she grins, 'Okay this Brivo kettle is actually insane, boils in like a minute!'; AUDIO: diegetic pour and steam-hiss plus her voice in sync under soft lo-fi music; bright neutral palette, neutral accurate white balance, clean true-to-white highlights, natural skin tones, no yellow or amber cast; premium high-end commercial look

Video 44

A mother duck leads a line of exactly FIVE ducklings toward a curb; the fifth and smallest duckling struggles at the step, then hops up to rejoin the others. There are exactly five ducklings in total — the same five throughout the entire video — never add, remove, duplicate or change the number of ducklings; keep all five visible and consistent from start to finish. Cozy indoor lamplight, authentic home-video look, slightly imperfect handheld framing with natural shake. A single continuous ~10-second moment with a clear beginning ...

Video 45

A baby otter floating on its back, holding a smooth pebble on its belly

Video 46

A slick, aspirational cinematic TV commercial for Blush Fizz, a sparkling lemonade brand. Opens with a top-down slow rotation of the glass bottle then a side tracking shot as it's poured in slow motion with citrus pulp and effervescent bubbles swirling upward. AUDIO: diegetic — the fizzy pour, faint citrus zest crackle, and light glass-clink foley, crisp and in sync, rising into a bright pastel-pop music bed; a confident voiceover says, 'Blush Fizz. Sip the light side.' Soft pastel pink and mint palette, neutral accurate white balance with clean true-to-white highlights and no ...

Video 47

a close-up with a subtle handheld sway of a single podcast host speaking into a large studio microphone, wearing headphones, in a cozy podcast studio with warm lighting and acoustic foam on the wall. The host says, 'hey chat, controversial take incoming, but the alphabet has been in the wrong order this entire time and we all just accepted it. Why is Q sitting right next to R? Who decided this? I demand answers.' On the wall behind them are two posters reading 'Reform The ABCs' and 'Why Is Q Next To R'. Deadpan-intense, slightly unhinged, comedic, realistic.

On Arena, Muse Video ranks No. 3 in human-preference Elo for text-to-video at the time of writing.

Arena Elo rankings as of July 5, 2026

Content Seal

To help people verify whether an image is AI-generated, Muse Image includes Content Seal, our invisible watermarking system. Images created by Muse Image in the Meta AI app and on meta.ai carry a hidden provenance signal that stays intact — even when cropped, compressed, resized, or screenshotted. We plan to extend Content Seal to video soon. We're previewing a detection tool that lets you check whether an image carries a Content Seal watermark, providing an initial way to help you better understand if an image was made with Meta AI.

Muse Image in Meta Products

Muse Image connects deeply with the Meta ecosystem. Our ongoing investments in image and video generation will further enable creators and businesses to generate dynamic content across Meta products.

Marketing assets for small businesses

Video 48

Personalized presets directly in Instagram

Gallery

Video 49

Video 50

Video 51

Video 52

Explore additional resources

Try Muse Image in Meta AI

Check Content Seal

Explore Meta AI Research

Read More Meta AI Updates

Learn About Meta AI

Work on AI at Meta

* * *

Written by:

Meta Superintelligence Labs

Share:

* * *

Our latest updates delivered to your inbox

Subscribe to our newsletter to keep up with Meta AI news, events, research breakthroughs, and more.

Join us in the pursuit of what's possible with AI.

See all open positions

Related Posts

Computer Vision

Introducing Segment Anything: Working toward the first foundation model for image segmentation

April 5, 2023

Read post

FEATURED

Research

MultiRay: Optimizing efficiency for large-scale AI models

November 18, 2022

Read post

FEATURED

ML Applications

MuAViC: The first audio-video speech translation benchmark

March 8, 2023

Read post

Meta AI

Assistant

Media Generation

Vibes

Muse

Agent

AI agents explained

What is agentic AI

Agentic AI examples

AI Research

Overview

Projects

Resources & tools
Publications
GitHub
Resources
Blog
Learning Hub
Demos
About
Overview
Open Source
Careers
Meta AI
Meta AIAssistantMedia GenerationVibes
Muse
MuseAgentAI agents explainedWhat is agentic AIAgentic AI examples
AI Research
AI ResearchOverviewProjectsResources & toolsPublicationsGitHub
Resources
BlogLearning HubDemos
About
AboutOverviewOpen SourceCareers
Privacy Policy
Terms
Cookies
Meta © 2026
Close
Close
Close
Close
Close
Close
Close
Close
Close
Close
Close
Close