



RESEARCH • 2026

The AI Avatar Market Report 2026

How AI avatars are actually used in production. Drawn from a census of **101,675 real renders** of our Aurora avatar model across one of its largest deployments (nine-week snapshot, mid-2026): use-case classification of ~1,000 prompts, a visual sample of 240 outputs, language spread, and the audience segments building with AI avatars today.

CENSUS

101,675 renders

WINDOW

9 weeks • mid-2026

BY

Creatify Labs Research

THE SHORT VERSION

This is a map of who's actually building with AI avatars, drawn from real production traffic rather than a survey. It resolves into six recurring creator segments, and what they make looks very different from the polished "corporate spokesperson" demo everyone expects.

- **The biggest, most commercial segment is the indie / SMB direct-response marketer** (~25% of volume), spinning up UGC-style video ads for e-commerce and DTC brands. They write full ad scripts, hooks, CTAs, sometimes a phone number baked right in.
- **Enterprise is already here, quietly.** Regional retail chains, telecoms, transit operators, and manufacturers are replacing filmed spokesperson and compliance-training videos with avatars, a sign this has crossed from novelty into production.
- **Brand safety is a demoable, pitch-ready requirement, not an edge case.** 132 prompts explicitly instruct "keep the logo intact," down to "hold up a business card so the logo is clearly visible."

101,675

renders analyzed (9 weeks)

98.7%

render success rate

~700h

of video generated

15s

median video length

20+

prompt languages

30K+

creators in the deployment

01

SECTION 01

The state of AI avatars



AI avatars have crossed out of the uncanny valley and into everyday production. Investors tracking the space, a16z among them, frame avatars as one of the defining consumer and enterprise categories of generative AI: tens of millions of people have already created one, and video models are increasingly specializing by use case rather than converging on a single generalist.

The market direction is steep. Synthetic-media software is growing at roughly 15% a year, the broader AI-video market far faster, and analysts expect synthetic media to make up the majority of enterprise video (training, support, internal comms) within the next year. In parallel, more than half of social content is projected to be AI-generated or AI-assisted, and brands report several times the engagement from synthetic creators over traditional ones.

Most of that story is forecast. This report is the opposite: a look at what people **actually** do when avatar video gets cheap and easy, grounded in real production usage rather than projections.

02

SECTION 02

How AI avatars actually get used



Nearly every render follows one shape: an **image** plus a **voiceover**, with optional **text direction**. Aurora's job is to animate the image and lip-sync it to the voice.

- **Voice comes first, animation second.** In this deployment essentially every request brought its own audio; the typical flow is that a creator generates a voiceover, uploads an avatar image, and gets a talking video in one step.
- **Motion control is a hard requirement.** The single most common thing creators write in a prompt isn't a creative brief, it's animation direction: "look at camera," "sync the lips," "don't drift." They iterate to get the motion exactly right.

03

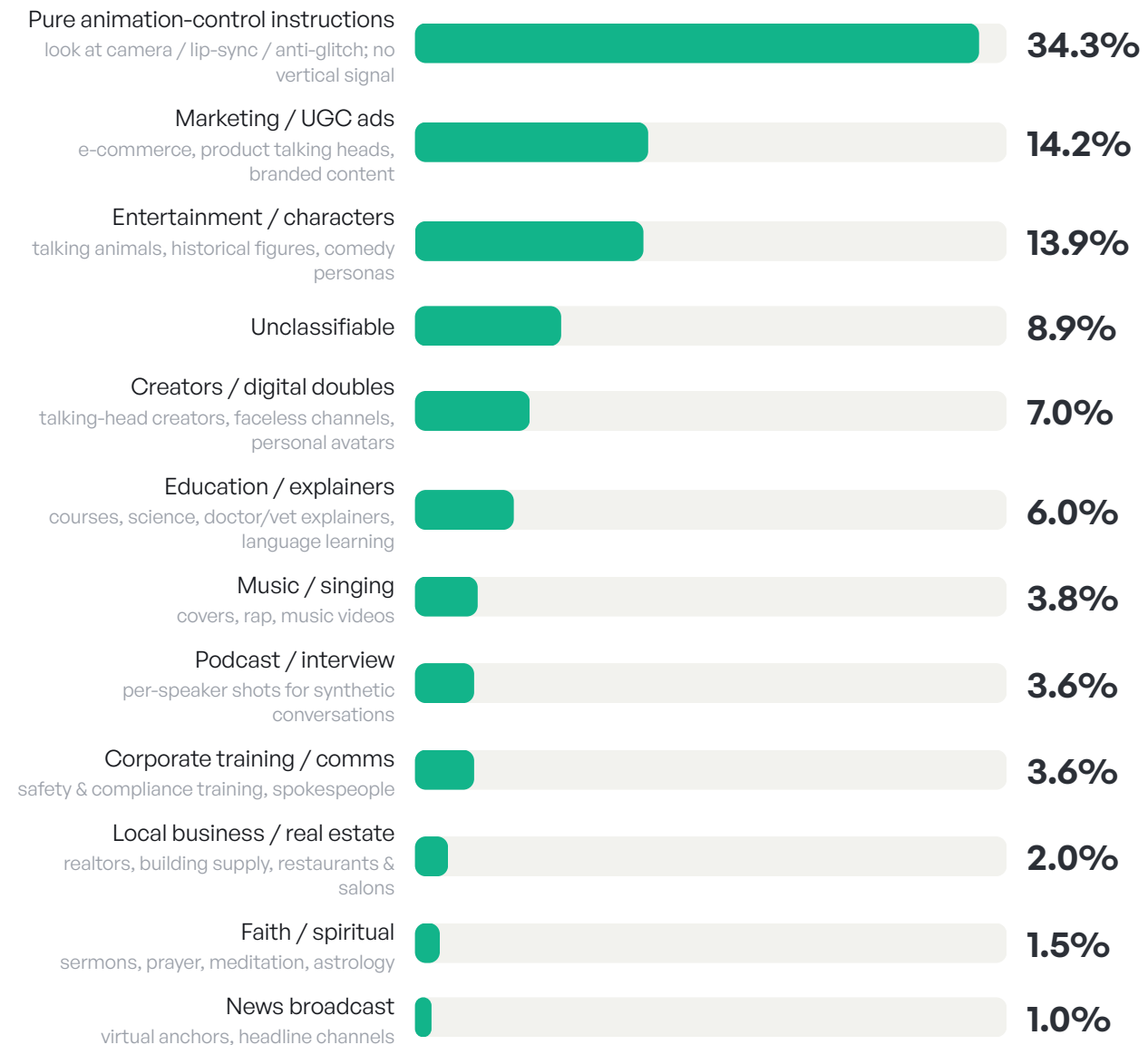
SECTION 03

What people make



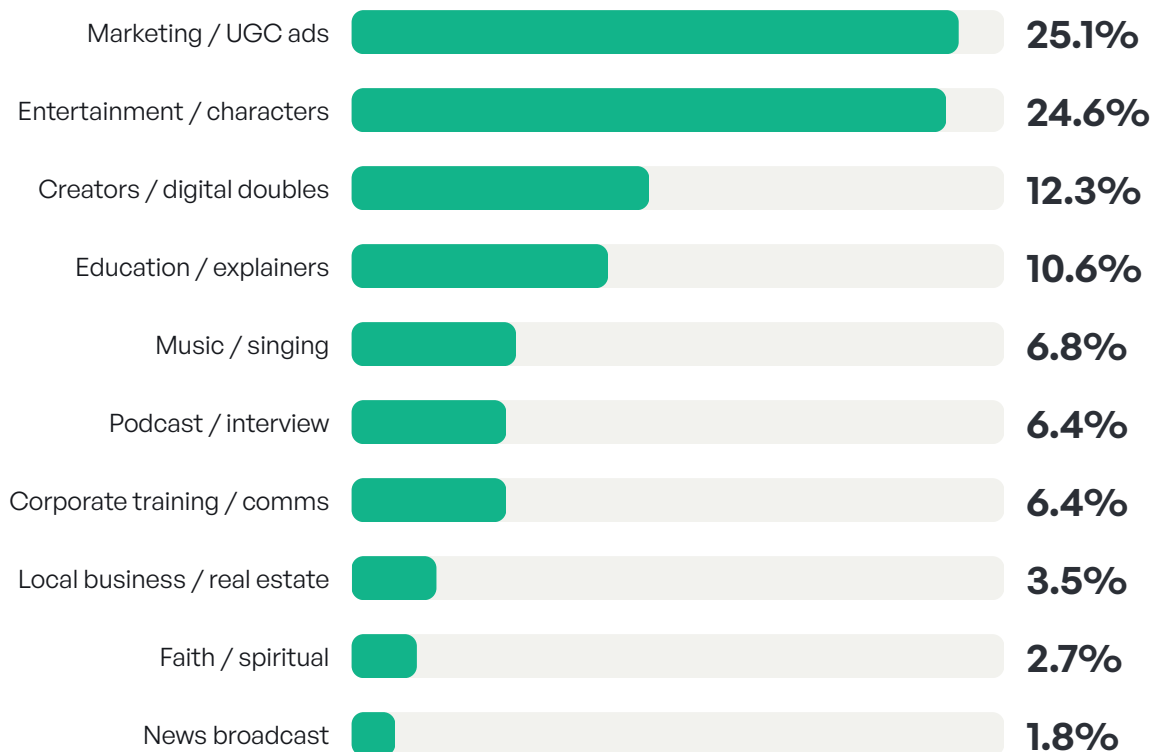
About a third of requests carry a text prompt. We sampled ~1,000 at random and labeled them across 12 use-case categories, weighted by volume:

Use-case distribution (n≈999, request-weighted)



The largest single bucket is pure animation-control instructions ("look at camera, sync the lips, don't move"), evidence of heavy iterative tuning and how much creators care about motion control.

Share of identifiable verticals (no-signal buckets removed)



Marketing/ads and character entertainment tie for first at around a quarter each; education and corporate training together make up the most standardized, highest-value segment.

What the prompts look like (scrubbed excerpts)

UGC / E-COMMERCE ADS

"Realistic UGC-style vertical video... 'If your mascara is flaking onto your cheeks by noon...' — a hook-and-CTA ad script.

CORPORATE TRAINING

"An experienced corporate trainer... the first frame and the last frame must be visually identical, ready for seamless editing between clips." (reused as a production-line template)

EDUCATION / EXPLAINER

"An educator introducing a new tool... make the mouth movement follow the uploaded voiceover naturally, like a lecturer." (non-English voiceover)

ENTERTAINMENT / CHARACTERS

"A highly realistic orange tabby cat hosting a professional podcast in a modern recording studio, studio headphones around its neck."

LOCAL BUSINESS

"Here's how the pros install architectural shingles... get all your materials at your local building-supply store." (with an address/phone CTA)

MUSIC / SINGING

"The avatar should sing actively with precise lip synchronization... gently sway to the rhythm, the energy of a live singer."

04

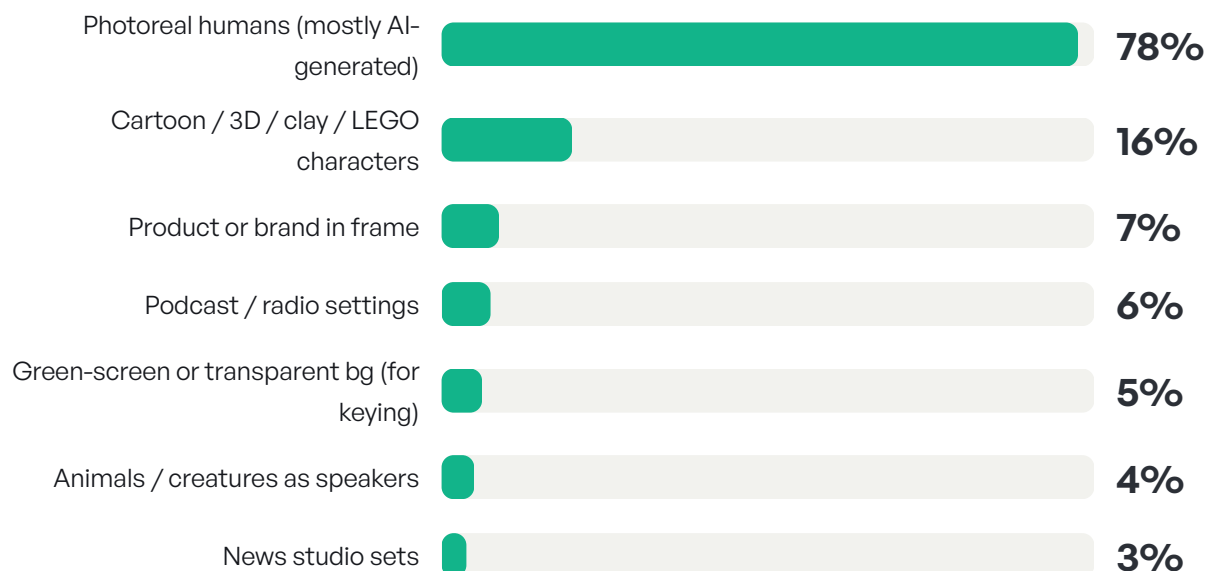
SECTION 04

What the inputs look like



Because outputs are short-lived, we pulled first frames ($\approx$ the input image) from 240 finished videos in a 24-hour window and classified them by hand:

Input image types (n=240, manual review; categories overlap)



Most "humans" are AI-generated portraits (Midjourney/Flux-style); real selfies and footage are roughly a fifth. Green-screen / transparent backgrounds show up because creators want to key the avatar into editors like CapCut, a mature workflow.

- **Human presenters ~78%:** spokespeople, UGC talking heads, news anchors, and doctor/coach personas. A large share is fully synthetic, AI face + AI voice + Aurora animation.
- **Cartoon, 3D and mascot characters ~16%:** animal podcast hosts, Pixar-style teachers, claymation doctors, LEGO stop-motion. Brand-mascot content is a clear sub-market.
- **Real company branding appears in a meaningful share of frames** (pharmacies, showrooms, retail counters, safety-training sites), a sign that enterprise content teams are already using avatar video in production.
- **Language and geography span 20+:** Spanish, Portuguese, German, French, Polish, Swahili, Hebrew, Arabic, Japanese, Korean, Vietnamese, Thai and more. English is only ~81%.
- **Length:** median 16s, mean 27s; ~31% are quick ≤ 8 s probes, ~9% run over a minute (up to multi-minute training pieces).

05

SECTION 05

Brand and logo signal



We scanned a large random prompt sample (plus the visual review above) for explicit brand, company, product and website mentions. The takeaway: brand-safe generation is a named, demoable requirement, not a rare request.

11.3%

of prompts (random baseline)
name a specific real brand or
company

492

distinct companies, products
and sites referenced

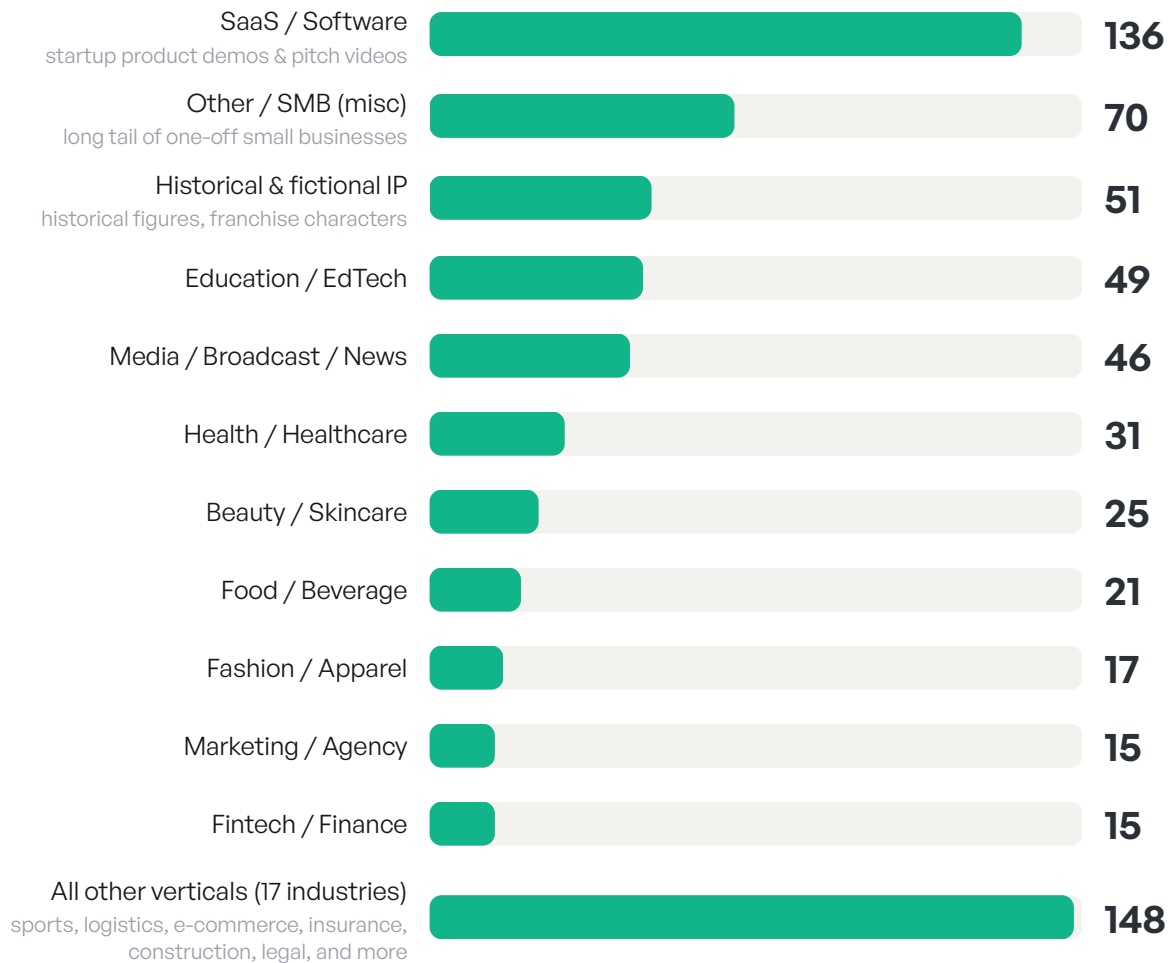
132

prompts explicitly say
"preserve / don't distort the
logo"

27

prompts request a green-screen / keyable background

Brand mentions by industry (n=624, top 12 buckets)



Software/SaaS is the single largest bucket, a long tail of founders using avatars for product demos and pitch videos. "Historical and fictional IP" (historical figures, franchise characters) is comparable to any single industry vertical.

06

SECTION 06

Who's building with AI avatars



Six recurring segments explain almost all of the traffic. Together they're a map of who reaches for AI avatar video when it's easy to make:

~25% OF VOLUME

The indie / SMB direct-response marketer. Runs UGC-style ads for e-commerce and DTC brands, skincare, supplements, gadgets. Writes ad scripts with hooks and CTAs, sometimes with a phone number baked in.

LARGEST BRAND-MENTION GROUP

The SaaS founder / solopreneur. Builds product-demo and pitch videos instead of filming themselves, the single largest brand-mention category. Technical and API-comfortable.

ENTERPRISE SEGMENT

The corporate L&D / compliance lead. Replaces filmed safety and onboarding videos with avatar training at scale, reusing an identical first/last-frame loop template hundreds of times.

~25% OF VOLUME

The faceless / character entertainment creator. Builds recurring fictional personas, talking pets, historical figures, horror characters, for YouTube and TikTok. High posting frequency, lower budget per video.

~11% OF VOLUME

The global educator / explainer. Teachers, doctors and coaches making explainer content, with heavy multi-language signal across 20+ languages.

~3.5% OF VOLUME

The local service-business owner. Realtors, roofers and suppliers whose prompts bake in a phone number or address CTA directly.

07

SECTION 07

Where this is heading



Read together, the usage points to a few clear directions for avatar video, the gaps creators are already engineering around by hand:

- **Brand-safe generation becomes table stakes.** With 132 prompts explicitly asking to preserve a logo, "keep the brand intact" moves from a hopeful instruction to an expected capability.
- **Native green-screen / alpha output.** Creators manually request green backgrounds today so they can key the avatar into their editor. Producing a clean alpha channel removes the workaround entirely.
- **Loop and first/last-frame presets.** The template enterprise trainers rebuild by hand, identical opening and closing frames for seamless stitching, is an obvious feature to productize.

08

SECTION 08

Method and caveats



- **Scope:** a full census of 101,675 renders over a nine-week production window, cross-validated against an independent event stream (daily counts match).
- **Use-case classification:** ~999 random prompts out of ~25,000 distinct, labeled by multiple parallel LLM classifiers into 12 mutually exclusive categories, weighted by request count. Only ~35% of requests carry a prompt, so vertical shares are lower-bound estimates over prompt-bearing traffic; the rest are covered by the visual sample.
- **Visual sample:** 240 finished videos from a single 24-hour window, first-frame manual review, proportions reflect that day's traffic mix.
- **Brand extraction:** a ~1,764-prompt sample processed by parallel LLM extractors. The 11.3% brand-mention rate is computed on the purely random subset to avoid keyword-filter inflation.
- **Privacy:** all figures are aggregate. Individual company names, end-user identifiers and verbatim customer content have been removed.

Build with the models behind this report.

[Explore the API](#)

The AI Avatar Market Report 2026 · Creatify Labs · Aggregate production analysis, mid-2026. Figures are rounded and anonymized. © 2026 Creatify.