

HeyGen Setup Guide: API and MCP

Connect HeyGen to the Claude you already have, learn what it really costs, and get the watermark answer straight

HeyGen makes AI avatar videos. There are two ways in. One is a connector that plugs HeyGen straight into Claude, so you just ask for a video in plain words. The other is an API key, which is a password for software, and it suits big batches. Start with the connector. One warning first: the old v1 and v2 web addresses that most tutorials still show you get switched off on 31 October 2026. We run these calls in production at ATP AI, Sales and Marketing, so the notes here come from a real funded account.

YOU WILL NOT NEED A TERMINAL

You will not type a single command in this guide. You already have Claude, so Claude does the technical parts for you. Each step gives you words to copy and paste to Claude. If you are in Claude Code, it just runs the thing on your computer. If you are in Claude Desktop or Cowork, it may hand you one line to paste and tell you exactly where to paste it. That is the only difference, and I will not repeat it again.

Before you start

- A HeyGen account at app.heygen.com. The free one is enough to start.
- Claude on a paid plan. Free Claude cannot add custom connectors, and the connector is the easy route.
- A credit card. Since February 2026 HeyGen gives out no free API credits, so nothing renders until you put money in.
- A web browser and about 20 minutes.
- For the API route only: a safe place to keep a password. Claude will set this up for you. Not a Google Doc.
- If you want to clone a real person's face: their permission. HeyGen checks, and it will hold your avatar until it is happy.

The steps

1 Pick which door you walk through

This step saves you a wasted afternoon. There are two ways to drive HeyGen and they bill from two different pockets. Door one is the connector. It signs in as your normal HeyGen web account and spends your monthly web credits. Door two is an API key. It spends a separate prepaid wallet of real dollars and allows bigger jobs. HeyGen's own advice says the sign in route is for trying things out, and the key is for batches, pipelines and real production traffic. So: if you want Claude to make you a video while you chat, use the connector. If you want 200 personal videos made overnight, use the key. Plenty of businesses end up with both. You do not click anything yet. Just decide.

ASK CLAUDE

I want to use HeyGen for video. Here is what I need it for: [describe what you want to make and how many videos a month]. Tell me plainly whether I should start with the HeyGen connector in Claude or a HeyGen API key, and roughly what each one will cost me a month.

You know it worked when: You can say out loud which route you are starting with, and why.

WATCH OUT

The two pockets are not linked. Paying \$29 a month for a Creator web plan puts zero dollars in your API wallet. Topping up the API wallet gives you zero web credits. People lose hours to this.

2 Connect HeyGen to Claude

This is the no-code route and it is the one I recommend you start with. HeyGen runs an official connector. It is theirs, not a third party bolt on. In Claude, click the plus sign. Then click Connector. Then Manage Connector. Then + Add custom connector. Name it HeyGen. Paste this address in: <https://mcp.heygen.com/mcp/v1/>. Click save, then click Connect. HeyGen's own approval page opens in your browser. Approve it. That is the whole sign in. No API key anywhere. Last thing: set the connector permission to Always Allow, so Claude stops asking you every single time.

ASK CLAUDE

Walk me through adding a custom connector in Claude, step by step. The address is <https://mcp.heygen.com/mcp/v1/> and I want to call it HeyGen. Tell me exactly what to click. When I say I am done, check the connection by listing my HeyGen avatars back to me.

You know it worked when: HeyGen shows up in your connector list. Ask Claude to list your HeyGen avatars and real avatar names come back.

WATCH OUT

Claude Free cannot add custom connectors at all. You need a paid Claude plan for this step. Also remember the pocket split: the connector spends your web subscription credits, and those are the ones that run out fastest.

3 Make your first video by asking for it

Now you get to see the point of all this. You do nothing in a browser. You type a sentence to Claude and a video comes out. The connector gives Claude a big set of controls: make a video, list your videos, check a video, list avatar groups, list avatar looks, make an avatar, list voices, clone a voice, design a voice, make speech audio, build from a template, translate a video, cut AI clips, run a batch of videos, and list your brand kits. Claude picks the avatar, writes the script and renders it. The same connector also works in Claude web, Claude Desktop, Claude Code, Cursor and Gemini CLI.

ASK CLAUDE

Using HeyGen, make me a 45 second video explaining [your topic]. Pick a suitable avatar and voice. Show me the script first and wait for my yes. Then render it and give me the link when it is done.

You know it worked when: The finished video shows up on the Projects page inside your HeyGen dashboard, and Claude hands you a link.

WATCH OUT

Every render spends credits, so read the script before you say yes. Ask for the script first, every time. It is much cheaper than re-rendering.

4 Get the watermark answer before you promise a client anything

Buyers always ask this and most vendors dodge it, so here is the straight answer. There are two answers because there are two products. On the web app: the Free plan puts a watermark on everything and you cannot remove it. Creator at \$29 a month, Pro at \$49, Business at \$149 and Enterprise all include watermark removal. You turn it off with the watermark toggle at the bottom left of the Submit Video menu when you click Generate. Once it is off, it stays off. On the API: HeyGen's v3 documentation has no HeyGen branding watermark at all. The only watermark setting there is one you control, for putting your own PNG or JPEG on top of your video. There is also no test flag any more. The old v2 test:true preview, which gave you a free render with a stamp on it, is gone. I checked this by hand. I downloaded a finished video paid for from a funded API wallet and pulled a still frame out of it. No HeyGen mark. No corner logo. Clean 1080p.

ASK CLAUDE

Take this finished HeyGen video file and pull three still frames out of it: one near the start, one in the middle, one near the end. Show them to me so I can check for any watermark or logo.

You know it worked when: You look at real frames from a real render and see no HeyGen logo.

WATCH OUT

The connector route spends web subscription credits, not the API wallet. So if your web plan is Free, you get Free plan rules on that video, watermark included. If clean video is written into a client contract, spend a dollar on one test render through the exact route you will use for real, and look at the actual frames.

5 Put money in the API wallet

Only do this if you want the API route for volume. Go to app.heygen.com and open the API dashboard at app.heygen.com/home?from=&nav=API. Click to add balance. Pay as you go starts at \$5. No subscription, no monthly commitment. The balance is held in real dollars, not points. Credits you buy this way expire 12 months after you buy them.

ASK CLAUDE

Remind me in 11 months that my HeyGen prepaid API credits expire 12 months after purchase, and tell me how to check the balance left.

You know it worked when: The API dashboard shows a dollar balance above zero.

WATCH OUT

Free API credits ended in February 2026. A brand new key with a zero balance signs in perfectly fine and then fails every single render with a 402 insufficient_credit error. It looks exactly like a broken key. It is not one. It is an empty wallet.

6 Create the API key

The key is a long password that lets software use your HeyGen account. Stay on the same screen: `app.heygen.com/home?from=&nav=API` . Click the button to generate your API key. There is no plan gate on the key itself, so even a free web account can make one. Copy it the moment it appears and put it somewhere permanent. HeyGen will never show it to you again. Your only option later is to make a new one.

ASK CLAUDE

I am about to generate a HeyGen API key. Before I copy it anywhere, set up a safe place to keep it and tell me exactly where to paste it. Do not print the key back to me in the chat.

You know it worked when: You have the key saved somewhere safe, and you have not pasted it into a chat, a document or an email.

WATCH OUT

Making a new key kills the old one instantly. If the old key is wired into something that is running live, that thing stops the second you click the button.

7 Store the key where nobody can see it

A key is money. Anyone who reads it can spend your wallet. The safe way is to keep it in a settings file that never leaves your machine or your server. HeyGen names it `HEYGEN_API_KEY` in every one of their examples, so use that name. In a settings file it is one line. In a hosting control panel it is a setting with that name. HeyGen's own security notes are blunt: never save the key into shared code, never put it in anything a web page loads, and change it now and then.

ASK CLAUDE

Store my HeyGen API key as `HEYGEN_API_KEY` in the right settings file for this project, and make sure that file is never shared or uploaded anywhere. Then tell me in plain words where you put it.

You know it worked when: Claude confirms where the key lives, and the key does not appear anywhere on screen.

WATCH OUT

If the key ends up in the code a web page downloads, anyone can view source and read it. Every HeyGen call has to run from a server, from your own machine, or from an automation tool. Never from a public web page.

8 Prove the key works before you build on it

One quick check, and it costs nothing. There is an address that just tells you who you are: <https://api.heygen.com/v3/users/me> . The key goes in a label called X-API-Key. A good key gives back your name, email, billing_type and your wallet balance. Mine comes back with billing_type of wallet and remaining_balance in real dollars. That one answer tells you three things: the key is good, which billing pocket you are on, and how much money is left. A wrong key gives back 401, which just means not allowed.

ASK CLAUDE

Using my stored HEYGEN_API_KEY, call GET <https://api.heygen.com/v3/users/me> with the header X-API-Key set to that key. Tell me my account name, my billing type and my remaining balance in plain English. Do not show me the key itself.

You know it worked when: You get your own name back, plus a dollar balance.

WATCH OUT

The label is X-API-Key. It is not Authorization and it is not a Bearer token, unless you deliberately went the sign in route. Capital letters do not matter. The name does. Also watch for a stray quote mark or a blank line copied in with the key.

9 Find your avatar and its look ID

An avatar is the person in your video. HeyGen splits this in two, and this trips everyone. A group is the character. A look is one outfit or pose of that character. The list lives at <https://api.heygen.com/v3/avatars> . Add `ownership=public` for HeyGen's stock library, `ownership=private` for avatars you made yourself, or leave it off to see both. You can also set `limit`, from 1 to 50, which starts at 20, and page through with a cursor. Each group shows a `looks_count` and a `default_voice_id`. The single look sits at https://api.heygen.com/v3/avatars/looks/{look_id} . The look ID is the one you use when you make a video. Not the group ID.

ASK CLAUDE

List my HeyGen avatars using GET <https://api.heygen.com/v3/avatars> with the X-API-Key header. Show me a simple table: character name, how many looks it has, its default voice, and the look ID for each look. Also tell me which engines each look supports.

You know it worked when: You have a short list of look IDs written down, with a name beside each one.

WATCH OUT

Group IDs and look IDs both look like random letters and numbers, so they are easy to mix up. Passing the group ID gets you a 400 error or `avatar_not_found`. Also check `supported_api_engines` on the look. Asking for an engine the look does not support fails outright.

10 Pick a voice and listen to it

The voice list lives at <https://api.heygen.com/v3/voices> . You can narrow it by type, which is public or private, and by language, gender and engine. Each voice gives you a `voice_id`, a name, a `preview_audio_url` you can actually listen to, and two flags: `support_pause` and `support_locale`. Pick two or three. Play the previews. Write down the IDs. You will use the same handful forever.

ASK CLAUDE

List my HeyGen voices using GET <https://api.heygen.com/v3/voices> with the X-API-Key header. Filter to English. Give me a table of name, `voice_id`, gender and the preview link so I can listen to each one.

You know it worked when: You have listened to at least two previews and saved the `voice_id` of the one you like.

WATCH OUT

Previews always sound better than a full 90 second script. Before you lock a voice in for a client, render one real test with your actual words. Voices that sound great for eight seconds can grate for a minute and a half.

11 Have Claude make the video

This is the volume route. Videos are created at <https://api.heygen.com/v3/videos>. A simple avatar video needs a type of avatar, the `avatar_id` set to your look ID, your script, your `voice_id`, an `aspect_ratio` and a resolution. HeyGen suggests `aspect_ratio` of `auto` and resolution of `1080p`. Instead of a script and a voice you can pass `audio_url` or `audio_asset_id` and it will lip sync to your own recording. You can also set the engine: `avatar_iv` is the default, `avatar_v` is the sharpest but is digital twins only and needs an eligibility check, and `avatar_iii` is the photo pipeline. You get back a `video_id` and a status of `waiting`. Then you wait. Status goes `pending`, then `processing`, then `completed` or `failed`. If it fails you get a `failure_code` and a `failure_message` saying why.

ASK CLAUDE

Make a HeyGen video for me. Use `avatar_id` [your look ID], `voice_id` [your voice ID], `aspect_ratio` `auto` and resolution `1080p`. The script is: [paste your script]. Send an `Idempotency-Key` so a network hiccup cannot bill me twice. Then check the status every 15 seconds and tell me when it is done or if it fails.

You know it worked when: Claude reports a status of `completed` and hands you a link.

WATCH OUT

Only one visual source is allowed. Send both an `avatar_id` and an image and you get a 400 saying Exactly one visual source required. The script cap is 5,000 characters. The output cap is 30 minutes and 50 scenes. Also, do not let anything check status in a tight loop. Every check counts against your rate limit, and a 429 makes a perfectly healthy render look broken. For a long batch, ask Claude to set `callback_url` instead, so HeyGen tells you when it is done.

12 Grab the finished file and keep your own copy

The finished video comes with a lot in the box: `video_url` for the MP4, `thumbnail_url`, `gif_url`, `captioned_video_url` with captions burned in, `subtitle_url` for the SRT caption file, plus duration, `completed_at` and `video_page_url`. Your whole library sits at <https://api.heygen.com/v3/videos>, with the same fields, paged in blocks of up to 100, and you can filter by `folder_id` or title. Download the MP4 to your own storage the moment it says completed. Do not leave it sitting on HeyGen's link.

ASK CLAUDE

```
When that HeyGen video finishes, download the MP4 and the SRT caption
file to [your folder], name them after the script topic, and tell me the
file paths. Then list my last 10 HeyGen videos with their titles and how
long each one is.
```

You know it worked when: The MP4 is sitting in your own folder and it plays.

WATCH OUT

`video_url` is a link with a timer on it. It expires. If you paste that link into a client email, it will be dead when they finally click it next Tuesday. Save the file first, put it on your own storage, then send that link.

13 Know your three avatar types, what they cost, and the consent rule

There are three kinds and they are not interchangeable. Studio Avatars are HeyGen's stock library, roughly 500 ready made characters, available right away with ownership=public. Photo Avatars are built from one still photo. They are cheap and fast and fine for volume. Digital Twins are trained from real video of a real person. That is the one that actually looks like you, and it is what founder led content needs. Making either a Digital Twin or a Photo Avatar costs \$1.00 each time. Rendering is billed by the second, and the rate depends on both the avatar type and the engine. All the numbers are in the costs section below.

ASK CLAUDE

```
I want a HeyGen digital twin of myself. Explain the consent step in plain
English, tell me what footage I need to record, and start the process
using POST https://api.heygen.com/v3/avatars/{group_id}/consent once I
give you the group ID.
```

You know it worked when: Consent clears and the twin moves out of pending into training.

WATCH OUT

Digital twins need consent verified first. You record through a webcam, upload a consent video, or hold a skip-consent permission. Budget a full day for this on a client job. It is not instant, and training does not even start until consent passes.

14 Move off the old v1 and v2 addresses now, not in October

This is the bit that bites anyone following an older tutorial. Right now the old addresses still work. In my own account `https://api.heygen.com/v1/video.list` , `https://api.heygen.com/v1/video_status.get?video_id=...` , `https://api.heygen.com/v2/avatars` , `https://api.heygen.com/v2/voices` , `https://api.heygen.com/v2/video/generate` and `https://api.heygen.com/v2/user/remaining_quota` all still answer normally. But every one of them also sends back a warning in the reply saying this v1 endpoint is Legacy and will be removed on 2026-10-31, along with a `sunset_date` and the name of the new address. The swaps are one for one. `v1/video.list` becomes `GET /v3/videos`. `v1/video_status.get` becomes `GET /v3/videos/{video_id}`. `v2/avatars` becomes `GET /v3/avatars`. `v2/voices` becomes `GET /v3/voices`. `v2/video/generate` becomes `POST /v3/videos`. The X-API-Key label is the same on both, so that part does not change.

ASK CLAUDE

Go through everything I have connected to HeyGen and find any use of the old v1 or v2 addresses. Swap each one to the v3 replacement, and fix the way the answers are read too, because the shape of the reply is different. Then test each one and show me the results side by side.

You know it worked when: Nothing you own still calls a v1 or v2 address, and your videos and avatars still list correctly afterwards.

WATCH OUT

The replies are shaped differently. The old ones wrap everything in a data object with a code field and hide the list under a name like `videos` or `avatars`. The new ones give a flat data list with `has_more` and `next_token`. If you only change the address and not the reading part, everything breaks quietly. Quietly is the dangerous word there.

15 Optional: the command line tool

Skip this unless you are on a Mac or Linux and you like shortcuts. HeyGen ships an official command line tool that wraps the whole v3 API and prints plain results. You install it with `curl -fsSL https://static.heygen.ai/cli/install.sh | bash`, then check it with `heygen --version`. You sign in with `heygen auth login --api-key`, or just set `HEYGEN_API_KEY` in your shell. It has a `--wait` flag that watches the render for you and hands back the finished video, and it retries rate limits on its own. Claude Code can install and run this for you.

ASK CLAUDE

```
Install the HeyGen CLI on this machine using the install script at
https://static.heygen.ai/cli/install.sh , confirm the version, sign it in
with my stored HEYGEN_API_KEY, and tell me if my computer is not
supported.
```

You know it worked when: `heygen --version` prints a version number.

WATCH OUT

It is Mac and Linux only right now. HeyGen says Windows is coming soon and suggests WSL in the meantime. On a Windows machine, forget the tool and use the connector or the API instead. Neither one loses you anything.

What it costs

There are two separate wallets and this is where everyone gets confused. WEB PLANS, from heygen.com/pricing as listed in August 2026: Free is \$0, gives you 3 videos a month, and puts a watermark on them that you cannot remove. Creator is \$29 a month with 600 credits and watermark removal. Pro is \$49 a month with 1,000 credits. Business is \$149 a month plus \$20 per seat, with 1,500 credits. Enterprise is priced on request. API BILLING is a completely separate prepaid wallet held in real US dollars. Pay as you go starts at \$5, has no subscription, and the credits expire after 12 months. Buying a web plan gives you zero API access. Topping up the API wallet gives you zero web credits. HeyGen stopped handing out free API credits in February 2026. Now the render rates, charged per second, from HeyGen's own self serve pricing page. Avatar IV, which is the default engine, costs \$0.05 a second for a Photo Avatar, about \$3.00 a minute, and \$0.0667 a second for a Digital Twin or Studio Avatar, about \$4.00 a minute. Avatar III, the photo to video pipeline, costs \$0.0167 a second for a Digital Twin or Studio Avatar, about \$1.00 a minute, and \$0.0433 a second for a Photo Avatar, about \$2.60 a minute. Avatar V is the sharpest engine, digital twins only, at \$0.0667 a second. Cinematic Avatar is a flat \$7.00 per video, for videos of 4 to 15 seconds. Video Agent, which turns a prompt into a video, is \$0.0333 a second. Video translation is \$0.0333 a second on Speed mode and \$0.0667 a second on Precision lip sync. Lip sync matches those same rates. AI clipping is \$0.15 per clip. Filler word removal is \$0.01 for every second of your source video, with a one minute minimum, and you get refunded if nothing changes. Text to speech on the Starfish model is \$0.000667 a second. Creating a Digital Twin or a Photo Avatar is \$1.00 per call. HyperFrames cloud rendering runs \$0.10 to \$0.30 a minute, depending on resolution and frame rate. Here is a sanity check

from my own account. The old credit view showed 2,972 credits against a \$49.53 wallet balance. So one old credit is one sixtieth of a dollar, which is about one second of standard avatar video. That is just the old rule of \$1 equals one minute, said another way. All of these numbers change. Read developers.heygen.com/docs/pricing before you quote a client a price.

When it goes wrong

Every call comes back 401 unauthorized and you are sure the key is right.	Check the label name first. It is X-API-Key, not Authorization, and not Bearer. Then check for a stray quote mark or a blank line that came along when you copied the key out of a settings file. Ask Claude to call GET https://api.heygen.com/v3/users/me for you. A good key gives back 200 and your account details.
402 insufficient_credit or subscription_required on a brand new key.	The wallet is empty. A perfectly good key with no money behind it signs in fine and then refuses to render anything. Top up at app.heygen.com/home?from=&nav=API . There have been no free API credits since February 2026.
429 rate_limit_exceeded or quota_exceeded.	Read the Retry-After label and wait that many seconds. Pay as you go allows 10 video jobs at once, and that pool is shared across avatar renders, video agent sessions and translations. Queue the work instead of firing 50 requests at the same time, and stop anything that checks status in a tight loop.
400 saying Exactly one visual source required.	You sent more than one of avatar_id, image_url or image_asset_id. Pick one. You get the same kind of error if you send a script plus voice_id and an audio_url together. One audio source only.
400 as soon as you name an engine.	That look does not support that engine. Ask Claude to call GET https://api.heygen.com/v3/avatars/looks/{look_id} and read supported_api_engines, which lists which of avatar_iii, avatar_iv and avatar_v that look accepts. Avatar V is digital twins only and needs an eligibility check.
avatar_not_found on an avatar you can clearly see in the dashboard.	You are passing the group ID instead of the look ID. The group is the character. The look is the outfit or pose. Only the look ID works as avatar_id.
The video link you sent a client stopped working.	video_url is a link with a timer on it and it expires. Download the MP4 to your own storage as soon as the status hits completed, then share your own copy.

A watermark showed up on a video you expected to be clean.	Work out which pocket actually paid for it. The connector and the sign in route spend your web subscription credits, so a Free web plan means Free web plan rules. An API key spends the API wallet, and the v3 docs show no HeyGen branding on that output. Also check you did not set the watermark field yourself. In v3 that field puts your own image on top.
Your setup worked last month and now fields are missing or odd.	You are probably still on v1 or v2, and the shape of the reply changed under you, or you are close to the 31 October 2026 switch off. Ask Claude to print the warning object in the reply. HeyGen names the exact v3 replacement address right there in it.
A digital twin sits at pending forever.	Consent has not cleared. Finish POST https://api.heygen.com/v3/avatars/{group_id}/consent by recording through a webcam or uploading a consent video. Training does not start until consent passes.
The heygen CLI install script fails on your machine.	You are on Windows. The tool is Mac and Linux only at the moment. Use WSL, or just skip it and use the API directly or the connector in Claude. You lose nothing either way.

Where this came from

Last verified: 22 August 2026. Vendor screens change constantly. If a step does not match what you see, the tool moved.

<https://developers.heygen.com/docs/api-key>
<https://developers.heygen.com/docs/quick-start>
<https://developers.heygen.com/docs/pricing>
<https://developers.heygen.com/docs/usage-limits>
<https://developers.heygen.com/docs/error-codes>
<https://developers.heygen.com/docs/avatars>
<https://developers.heygen.com/docs/for-ai-agents>
<https://developers.heygen.com/reference/create-video>
<https://developers.heygen.com/reference/get-video>
<https://developers.heygen.com/reference/list-videos>
<https://developers.heygen.com/reference/list-avatar-groups>
<https://developers.heygen.com/reference/list-voices>
<https://developers.heygen.com/models>
<https://developers.heygen.com/mcp/overview>
<https://developers.heygen.com/mcp/claude-web>
<https://developers.heygen.com/cli>
<https://developers.heygen.com/features>
<https://developers.heygen.com/lms.txt>
<https://www.heygen.com/pricing>

<https://www.heygen.com/api-pricing>

<https://www.heygen.com/faq>

<https://help.heygen.com/en/articles/10060327-heygen-api-pricing-explained>

<https://help.heygen.com/en/articles/11057301-how-to-remove-the-heygen-watermark>

WHO MADE THIS

Built by ATP AI, Sales and Marketing. We set this up inside businesses for a living, so if you would rather not do it yourself, that is what we are for. atpsalesandmarketing.com/contact-us/