

A Month of Quality Content in a Day — by hand, with Claude Code

The exact build-first method — no gatekeeping, the actual steps.

By Enrique Marquez · INFINITX / Nave

nave.infinixai.com · @enriquemarq-0

A month of content is not a month of filming. It's one focused day where YOU drive the raw material and a trained system does the heavy lifting and holds the standard. This is the whole method, the real one, by hand, with your own keys. Read it, do it, keep it. I held nothing back.

The deal, up front

I don't gatekeep, so let me be straight about what this is before you spend a minute on it.

This is the full method. Not a teaser. Not the first three steps with the good part locked behind a checkout page. By the time you finish reading, you'll be able to sit down for one focused day, drive your own raw material, and walk out with a month of content. Your hands on the wheel, your own keys to the tools. If you never buy a thing from me, you still leave with all of it.

Here's the part I need you to hear before anything else, because it's the part everyone selling you software gets backwards. **You are in the driver's seat.** You bring the raw material and the direction. The system does the heavy lifting and holds the standard. It is not a magic button that makes content while you sleep, and anyone telling you "10 viral posts in 5 minutes, zero effort" is selling you the same no-code fantasy that already burned you. The point isn't that AI makes your content. The point is that you OWN the machine that makes it. Your hands, your keys, nobody skimming the middle.

I build first, then I teach. Everything in here is something I actually do, on real work, every month. Real receipts. The tools I name you bring yourself, no middleman, no markup. You pay them directly and you own the whole thing.

There's one honest catch and I'll say it now instead of springing it on you at the end: doing all of this by hand is a full day. A real one. It's doable. I did it for months. But it's a day, every month. Near the end I'll show you the system I built to squeeze that day down to about an hour. That's the only thing I sell, and the method works fine without it. You'll see exactly where it fits and you can take it or leave it. **Two promises, and I will never blur them: this free playbook gets you a month of content in a focused DAY, by hand. Nave compresses that day to about an HOUR. I'm not going to headline the by-hand method as "an hour." That's not honest, and honesty is the whole brand.**

One more thing that matters more than people realize. **Claude Code runs on your own existing Claude subscription. There is no Anthropic API key to buy.** The only keys you bring are for AssemblyAI, KIE.ai, and Zernio, each on your own account, paid directly, no me sitting between you and the tool. If you've been scared off by "bring your own API key," let that go right now. The brain runs on the subscription you may already have.

Here's the shape of the day. Drive one rich input. Transcribe it word-for-word. Mine it for the moments that actually land. Cut and reframe them for the phone. Edit each one to a standard people stop for. Spin that single input into every format. Then schedule the whole month and let a give-first distribution move handle the follow-up while you sleep.

And the raw material is broader than you think. It's not only recorded footage. You can drive the month with your real footage, with AI-generated images trained on your brand profile, with format and carousel examples you studied off the internet, with a guided concept, or with an avatar script. Footage is one example input, the proven one, not the only path. I'll name all five before we're done.

Let's build.

Why a content day beats a content grind

Here's the lie I believed for way too long: that staying visible meant showing up every day. Wake up, think of something to post, film it, caption it, fight the algorithm, do it again tomorrow. Forever. The work was scattered and there was no end to it. The second I got busy or sick or just tired, the whole thing fell apart, because it depended on me being "on" every single day.

That's a grind. And a grind is fragile by design.

The thing that changed it for me was realizing you don't need to post every day. You need one focused input day that produces a month. One day where you sit down, drive your real material once, then turn that single block into thirty-ish pieces. The posting spreads across the whole month. The work (the thinking, the driving, the heavy lifting the system carries for you) happens in one sitting.

Stop asking "**what do I post today.**" Start asking "**what do I capture this month.**" That one swap changes everything downstream of it.

Let me be straight: this is real work. A content day is a full day. You'll talk, you'll record or gather, you'll make decisions, you'll correct drafts. It's not a magic button, and the boring middle (the part everyone edits out of their highlight reel) IS the proof that it's real. The work doesn't disappear. What changes is *when* you do it and how many times.

- **Grind:** small effort, every day, forever. It never compounds. The day you stop, the content stops.
- **Content day:** one concentrated batch, once a month. You front-load the work, then it pays out for thirty days while you go run your actual business.

Same total output. Completely different cost to your life.

And here's the part I care about more than the content itself. When you batch like this you're not feeding a treadmill anymore. You're building a system you own: a repeatable process that doesn't live or die on whether you woke up motivated. Right now I'm teaching you to run it by hand, your own keys, your own tools. Later I'll show you where it gets compressed. But even if you never touch anything I sell, you walk away owning the method. You get your time back, and the thing works for you instead of you working for it.

Do this now: Open your calendar and block your Content Day in the next seven days. Up front, block a focused four-hour window for the capture and setup — recording or gathering the one rich input and getting it ready. The rest of the day is the mining, editing, and scheduling by hand; by-hand, the whole thing fills a real day. Don't plan what you'll make yet, that's the next few sections. Just claim the time. The grind stops the moment that block exists on your calendar.

The system: a trained environment (the four primitives)

Before the steps, I want you to see the thing you're actually building. Because here's the reframe that changed everything for me: **you don't make content. You build the thing that makes it.**

A prompt asks AI to guess. You type "make me a clip," it guesses what you meant, and you get generic slop you have to fix by hand. A *trained environment* removes the guessing. It already knows your rules, your face, your colors, your bar, so the tenth piece comes out as clean as the first. One is a tool you rent by the prompt; the other is a machine you own.

The machine has four primitives. You don't need to build them to use the method by hand (every step below works with just an editor and your own two hands) but I want you to understand the shape, because it's what Nave is and it's what you can build yourself.

SKILLS — the recipe, where every rule is a scar. A skill is a written-down recipe for one job: how to caption, how to reframe, how to build a compilation. The reason it's reliable is that every rule in it is a fix for a real mistake I already made. The captions used to glue a comma onto the front of every line, so the skill now says *the events format must include the Name field*. The video used to die on certain frames, so the skill now says *encode on the GPU*. The voice used to get buried under the music, so the skill now says *lift the voice to about -15 LUFS, music whispers at a tenth of the volume*. The cover image used to get forgotten, so the skill now says *a reel is not shippable without an authored thumbnail*. Every scar becomes a line. That's why a trained skill makes the 10th piece as clean as the 1st: you're not relearning the lesson, the recipe already holds it.

AGENTS — one narrow specialist per format. Not one do-everything AI. A spotlight specialist that only knows how to cut conversational interview clips. A compilation specialist that only knows the name-chip-then-CTA build. A carousel specialist that only knows to pull real lines straight from the transcript. On the Tiempo build there are 16 of these. Narrow is what makes it reliable: a specialist that does one thing can't drift into ten. And because they're separate, they run in parallel. Five formats getting built at once while you make tea.

TOOLS — the hands that touch the files. The agents don't conjure things out of air. They reach for real tools: AssemblyAI for the transcript, a face-tracking reframe for the vertical crop, ffmpeg for the cuts, headless Chrome to render the title and CTA cards, ElevenLabs for the music bed. Your keys, your accounts. The agents do the reaching. You never open a command line.

THE DYNAMIC WORKFLOW — the line that routes it all. This is the part that makes it feel alive. The workflow reads what's actually *in* your input (who's talking, what they said, where the good moments are) and routes the right clips to the right agent. The spotlight moments go to the spotlight agent, the question-and-answer beats go to the compilation agent, the quotable lines go to the carousel agent. Then a QA gate checks the output before anything ships: voice leveled, no black frames, captions inside the frame, cover image present. You say "make my month," the

line runs. **And it stages, never auto-posts. You approve the slate before a single thing goes live.** That's the human staying in the driver's seat, by design.

Two things are what make it *trained* instead of generic, and they're worth naming on their own:

- **The footage catalog — its eyes.** A document that identifies every clip by who's in it, what's happening, and what's said. Without it, the system is blind and grabs the wrong person's b-roll. With it, every clip uses the right material. On the Tiempo build this got promoted to a global skill, because nothing about it is Tiempo-specific. You point it at any footage drop and it reads it.
- **The brand bible — its taste.** Your colors, your voice, your sacred phrases, your CTA, and what's *not* allowed. Every step reads it. That's how the output comes out on-brand without you art-directing each clip.

Take those two away and you're back to a generic AI guessing. Keep them, and you have a machine that knows you.

The template: The Footage-Drop Method

Here's the method, named so you can hold onto it: **The Footage-Drop Method — one drop, a month of content.** Said the same way every time.

It's seven steps. Each one ends in a concrete **"Do this now"** you can run today. I'm calling it the Footage-Drop Method because footage is the proven, load-bearing input. But read "drop" broadly: a drop is *one rich seed*, and a seed can be footage, brand-trained images, a studied reference, a guided concept, or an avatar script. I'll widen it fully in the inputs section. For now, walk the seven steps with me, and run them on one real input as you go.

Step 1 — One rich input (the drop)

A month of content is not a month of filming. It's one good input, mined properly. So the first move isn't to make thirty things. It's to drive *one* rich thing the whole month gets built from.

Most flexible and proven version: one deep session of you talking or working. A 20-to-40-minute talking-head, a walkthrough, a podcast or interview you were on, a screen-record of real work. Loose and unscripted is fine. You record it once, then spend the rest of the day turning it into ten, fifteen, twenty pieces. You film once. The mining is where the month comes from.

Not every recording gives you a month, though. A 4-minute "here's a tip" is one piece, not a drop. Run this lens before you commit:

- **Multiple distinct points?** If you can list five or six separate things you cover, you've got five or six spotlight clips, a compilation, and a carousel hiding in there. One point gives one clip. Many points give a month.
- **At least one real money line?** A moment where you say the thing crisp and quotable. That's your montage anchor and probably your best short. If you can't picture the standout sentence, the input is thin, so go deeper.
- **Can you talk 20+ minutes without running dry?** That's the honest test of whether you actually know it. The topics you can ramble on forever mine the richest, because the depth is real and it shows.
- **Would you actually watch it?** If the raw thing bores *you*, no edit saves it. Energy is the one thing you can't add in post.

And don't over-produce it. No lighting rig, no script, no five takes. Substance over polish; the polish happens later, format by format. The drop just needs the substance in it.

Do this now: On your Content Day, spend the first four-hour capture window recording or gathering one deep input: the topic you could talk about for 30 minutes without notes, with five distinct points in it. Phone on a tripod is fine. One deep thing, not ten small things. That single input is the seed for everything that follows — and the mining, editing, and scheduling that fill out the rest of the day all hang off it.

Step 2 — Transcribe with word-level timing

Before I cut a single frame, I get the words. Not the gist, the *words*, each one stamped with the exact millisecond it's spoken. That timed transcript is the spine the whole day hangs off. Captions snap to it. Cuts land on it. The "find me the best 45 seconds" pass reads it. Skip this and you're back to scrubbing a timeline by ear, which is the slow way I used to lose entire afternoons.

I run every drop through **AssemblyAI** on its **Universal-3 Pro** model. It's the most accurate speech engine I've used on real creator audio (phone mics, room echo, two people talking over each other, Spanglish, all of it) and it gives you word-level timestamps for free. You bring your own AssemblyAI key, pay them directly, and it's cents: about **\$0.27 an hour of audio**. A 30-minute interview runs you about fourteen cents. That's the whole bill for this step.

Why word-level is the part that matters:

- **Captions snap to the audio.** Each word lights up the instant it's said. No drift, no nudging subtitles a frame at a time.
- **Cuts land on real word boundaries.** You trim between words, not mid-syllable. Clean edit vs. chopped one.
- **Clip-picking gets smart.** Because every line is timed, you (or Claude Code reading the file) can say "give me 4:12 to 4:51" and it's a real, usable range, not a vague "somewhere in the middle."

You don't need to write code. Make a free account at assemblyai.com, copy your API key, keep it private (treat it like a password). Then hand the audio file and your key to Claude Code (running on your own subscription) and say it plainly: *"Transcribe this with AssemblyAI Universal-3 Pro, word-level timestamps, save to a file."* It uploads the file, runs the job, waits, saves the result. Be picky about exactly one thing: the model is `universal-3-pro` and you want the word-level output, not just the plain paragraph. Let it default to a lesser model and you lose accuracy; grab only the plain text and you lose the timestamps, which are the entire point.

What you end up with is a file (I name mine after the speaker) holding every word and its timing. One detail so nothing surprises you: AssemblyAI gives those times in milliseconds (120 and 320, not 0.12 and 0.32). Anything downstream that wants seconds divides by 1000, and Claude Code handles that. I only mention it so the numbers make sense when you peek.

Do this now: Make your free AssemblyAI account, copy your key into a private note, take your longest input from Step 1, hand it and your key to Claude Code, and get a words file on disk with real start and end times. When that file exists, Step 2 is done, and you never opened a terminal.

Step 3 — Mine the moments

Now read your transcript like a stranger and mark the few seconds worth attention. Here's what nobody tells beginners: a "month of content" is not a month of ideas. It's a handful of strong moments you already said once, pulled out and cleaned up. Your job isn't to write anything new. It's to find what's already there.

This is the taste layer, and it sounds like the hard part, like something that needs ten years of editing. It doesn't. Three questions, in this order. Run every chunk that makes you pause through them.

- **Would I stop scrolling for this?** Picture this line in your own feed between a hundred other videos. Does it make you stop, or keep going? Most of your transcript is keep-going. That's fine. You're hunting for the few seconds that are *stop*.
- **Does it teach something or prove something?** A moment earns its place one of two ways. It teaches (a real step, a how, a thing the viewer can go do) or it proves (a result, a number, a story, a receipt). If it does neither, it's filler. Cut it loose, no matter how good it sounded coming out of your mouth.
- **Is there a clean hook in the first 3 seconds?** People skip this one and it matters most. Does the moment *open* on the point, or is it a warm-up, "so, um, what I usually do is..."? You can't keep a windup. If the good line is buried 15 seconds in, fine, just mark where the real start is. You'll cut to it in Step 4.

Clears all three, it's a piece. Clears two, it's a maybe (keep it lower on the list). Clears one, it's a no. That's the whole filter.

As you read, the same shapes show up — tag each keeper by which one it is, with its timestamp:

- **Hook** — a line sharp enough to be the first 3 seconds by itself.
- **Money line** — the one-sentence version of something true and a little contrarian. The line you'd quote to a friend.
- **Teachable step** — anywhere you explain how to do a thing in order. These become carousels and "here's exactly how" clips.
- **Proof beat** — a result, a before/after, a number, a short story where something actually happened. Trust gets built here, not in the teaching.

Don't force the weak parts in to hit a number. A 40-minute recording might give you 12 real moments. Twelve is a great number. A feed full of "meh" trains people to scroll past you, so fewer, stronger moments win.

Do this now: Read your transcript clean once. Read it again and pull every moment that clears all three questions into a notes file, timestamp and shape on the line above each. Shortlist 8 to 15 moments, star your top 5. That list *is* your month.

Step 4 — Cut and reframe to vertical

Turn a wide 16:9 frame into a 9:16 vertical that looks like it was *shot* that way, not a lazy center crop with the head jammed against the top. Most people give up here and slap a center crop on it. Don't. The gap between a cheap reframe and an intentional one is five rules. Here are all five.

- **Cut on a completed sentence, with a short tail.** Set your in and out points off the *transcript*, not by ear. Start right as the first real word lands (trim the "um," the breath). End on a finished thought, never mid-sentence; a clip that cuts off "...and that's why you should—" feels broken and people bounce. Leave a ~1.5-to-2-second tail after the last word so it feels edited, not yanked. You're cutting on *meaning*. The transcript is your map.
- **Face-track the reframe, never static-crop.** A static crop assumes the person never moves. They always move. Use a face-tracking reframe, software that finds the face every few frames and slides the crop to keep it centered. Two things make tracked motion look pro instead of seasick: **smoothing** (the crop drifts, doesn't snap, so it lags the face slightly and small head bobs don't whip the frame) and a **deadzone** (tiny shifts in the seat shouldn't move the crop at all, only real movement follows).
- **Keep a caption-safe center band.** This is the one nobody tells you, and it's the single biggest "why does mine look cheap" fixer there is. Your captions live in the lower-middle. If your face is *also* dead-center, the captions land on your mouth and chin: cramped, amateur. So drop the face a little higher: eyes in the upper third, lower third clear for text. Now the captions have a clean lane and the face has room to breathe.
- **Never leave black gaps.** When the crop slides to follow a face near the edge, a naive tool runs *off* the source and fills the gap with black bars. That reads as broken instantly. The rule: clamp the crop to the edge of the source frame. When the face nears the edge, the crop stops at the boundary and lets the face drift within frame instead of exposing black. Full frame, edge to edge, every second.
- **Music ducks under the voice.** If you lay music under the clip (you usually should, it lifts the energy), keep it low the whole time: a bed, not a track, whisper level. Duck it further whenever the person speaks so the voice always sits on top. The voice is the product; music is the floor. You should catch every word without reaching for the volume.

Do this now: Take one clip pick from Step 3. Set the in-point on the first real word and the out-point on a finished sentence plus a ~2-second tail. Run a face-track reframe to 9:16. Then check the thing everyone misses: is the face sitting above center so the lower third is clear for captions? If the head's dead-center, nudge the crop up. That's a cheap reframe turned intentional in two minutes.

Step 5 — Edit to a real standard

Reframing makes it the right *shape*. Editing makes it the thing someone actually *watches*. A clip that's only reframed still gets scrolled past. This step is the gap, and the gap is everything. None of it needs code; all of it can be done by hand in any caption-capable editor (CapCut, Descript, Premiere, whatever you've got). The deliberate pass on every clip:

- **Cut the windup so it starts ON the point.** Almost everyone buries the good part: "so basically what I want to talk about today is..." Find the first sentence that actually lands (the claim, the number, the hot take) and make it your first frame. Gut check: play the first 3 seconds on mute. If a stranger wouldn't know what this is about, you started too early. Move the cut forward.
- **Burn in bold, word-level, lower-third captions.** About 80% watch with the sound off. "Burn-in" means the captions are baked into the pixels, not a separate file the platform might not show. Bold and big, lower third (high enough to read, low enough to clear the face), each word appearing as it's spoken (this is why Step 2's word-level transcript matters: the caption snaps in sync), high-contrast (white with a dark stroke or subtle box).
- **Pop the 1-2 keyword words per sentence.** Pick the words that carry the meaning, the number, the verb, the punchline word, and emphasize them a little bigger, in your accent color. Don't pop every word; if everything's emphasized, nothing is. This pulls the eye exactly where the meaning is.
- **One brand accent, everywhere, never drifting.** Pick one accent color and use it on the popped keywords, the title card, the handle in the corner. Mine is orange, **#FF7614**. When every clip carries the same color, your stuff gets recognizable in the feed before anyone reads a word. Pick *yours* and lock it.
- **A clean one-line title card.** A simple text title over your opening frame, telling the viewer in one line what they're about to get. Short, in your accent color, gone in a couple seconds. Not a fancy animated intro (those make people leave), just a one-line "here's what this is" so the clip feels finished instead of raw.

Do this now: Run the full pass on one clip: trim to the point, burn word-level lower-third captions, pop the single most important word in your accent hex, drop a one-line title card. Then play the first 3 seconds on mute. If a muted stranger gets it in 3 seconds, that's the bar, and that's the bar for every clip after this one.

Step 6 — Spin one drop into many formats

You already did the hard part. One input, a clean word-level transcript, the moments mined, a clip edited to a real standard. That raw material is worth far more than one post. The mistake most people make is treating a shoot as one video. It isn't. It's a month waiting to be cut a bunch of different ways.

Here's the shift: **a format is just a recipe**. Same steps, same order, every time. You build the recipe once (the cuts, the captions, the music level, the end card) and from then on you never think about it again. You just run it on the new material. **You're not creating five things from scratch. You're running five recipes you already own**. That's how a month of content fits in a day.

The five I run every drop:

- 1. Spotlight clips.** The bread and butter. One strong moment, 20-45 seconds, vertical 9:16, word-by-word captions, the face kept centered as they move. 3 to 6 per drop. (This is Steps 4 and 5, run again and again.) Funnel job: **trust** → **DM**.
- 2. The montage.** A fast, rhythmic stitch of the best 8-15-second hits across several moments: the energy reel. No single full thought, just punch after punch cut to the beat. Grab the shortest, hardest hits, trim each to its peak, hard-cut them back to back, one music bed driving the pace, a title card on the front. **This is the one people share**. Funnel job: **cold reach**.
- 3. The compilation.** Several speakers (or several takes) answering the same question, stacked into one piece. Title card up top ("We asked them...") then segment, segment, segment, then your CTA card. Cut each answer to its value beat, drop a name chip on each so viewers know who's talking, concat in order, one music bed across the whole thing, voice leveled and music whispered under. This is the exact recipe behind the Tiempo "advice" compilation: same build, new segments fed in. Funnel job: **reshare / social proof**.
- 4. The carousel.** The material becomes words on slides. Pull the 6 to 10 best lines straight from your transcript. They're already the speaker's real phrasing, so it doesn't sound written. One idea per slide, hook slide first, soft CTA slide last, on-brand colors and fonts. This format costs almost nothing because the writing is already done for you in the transcript. Funnel job: **authority** → **DM**.
- 5. The long-form package.** The full piece, lightly cleaned: dead air and "ums" cut, a title card, captions, your intro and outro. This is the anchor every short points back to, and it's what lives on YouTube and keeps getting found for months. Funnel job: **depth / the searchable asset**.

Every one of those is the same source material. One drop, five formats, and none of them is "make a video" anymore. Each is a checklist you've already written. When the material has a genuine gap, a shot you never filmed, an image you're actually missing, that's where **KIE.ai** comes in (AI image and video on your own key, to fill the hole instead of re-shooting). Use it only where you're truly missing something, not to manufacture content from nothing.

The honest Nave note, right here. Building each recipe once is the easy part: a couple hours and you have it forever. But running all five *by hand* on a fresh drop is a full day. You cut clips one at a time, level audio, place chips, harvest carousel lines, clean the long-form. The recipe doesn't run itself. You run it, five times. The way I escape that day is I don't run the recipes, I own the agents that know them. A spotlight specialist that already knows cut-on-a-sentence. A compilation specialist that already knows chip-then-CTA. A carousel specialist that already knows to pull real lines from the transcript. Each runs its own recipe in parallel while I make tea. That's Nave: the same recipes you just learned, owned as workers instead of run by hand. **You don't need Nave to do any of this.** Everything above works with your own hands and your own keys today.

Do this now: Take the one input you've been working through. Pick two of the five recipes (start with Spotlight and Carousel; they share the most material). Write each as a literal numbered checklist in a note, the way I wrote them here. Run both, by hand, on that material. Time yourself. That number is exactly what you're weighing when you decide whether to keep doing it by hand or hand it to a worker.

Step 7 — Schedule the month + comment-to-DM

One day of work becomes a month that runs itself, staged across platforms, with a distribution move that quietly out-performs how most people post. I'll give you the whole thing, the why and the how, so you can do it by hand tonight if you want.

One drop feeds three channels. I don't make different content for different platforms. I make one strong idea, then let each platform do the one job it's good at:

- **LinkedIn — conviction.** Where the belief lands: the take, the teardown, the "here's what I think and here's why." People read LinkedIn to think. Give them something to argue with.
- **Instagram — reach.** Same idea, short and punchy: a clip, a hook, a carousel. Your widest net and your lead engine. The comment-to-DM move lives here.

- **YouTube — depth.** The long version, the full build shown on screen. One good YouTube video gets found for months, the only piece that works while you sleep a year from now.

Stage the whole month at once. Don't post by hand every day; that's the trap that kills consistency. Lay the month out in one sitting while your head's already in it:

1. **Pick a rhythm you'll actually sustain.** Whatever you can hold beats whatever sounds impressive. Don't copy my numbers; pick numbers you won't quietly abandon by week two.
2. **Drop every piece onto a spreadsheet.** One row per piece, five columns: **Date · Platform · Hook · File · Keyword.** Seeing the whole month on one screen is the point. Gaps become obvious and you stop staring at a blank "what do I post today."
3. **Schedule it, don't babysit it.** Load the posts into a scheduler so they go out on their own. I use **Zernio** because it publishes to all three platforms and runs the comment-to-DM mechanic from the same place, on your keys and your accounts. Starting out you don't even need that: the native scheduler inside each app queues posts for free. Schedule by hand your first month, automate once it's working.

The give-first distribution move: keyword in the comments, not links in the post. This is the part most people get backwards, and it's free. **Do not drop your external link in the post body.** Every platform wants to keep people on the platform, so the second you paste a link that sends them away, the algorithm quietly throttles your reach. You pay for that link with views you'll never see. Do this instead: tell people to comment a keyword, and you DM them the thing.

*"Want the free template? Comment **GUIDE** and I'll send it to you."*

Two real reasons it wins. **It protects your reach:** no outbound link means the platform isn't punishing you, so the post gets shown to more people. And **comments are the strongest signal there is:** a comment counts for far more than a like; it tells the algorithm people are engaging and pushes the post wider. So the ask doesn't just collect leads, it grows the post *while* it collects them. And it's more honest, not less: the person asked, they typed the keyword, you hand them the real thing in the DM exactly like you said.

Set it up by hand first (zero tools): pick a dead-simple keyword per post (GUIDE, PLAYBOOK, MAP), watch the comments, DM the resource personally. Manual, but for your first month it teaches you what people actually want and a real human DM converts beautifully. Automate with Zernio when the volume gets real: you tell it *this post, this keyword, this DM*, and anyone who comments gets the DM automatically with a tappable button. Same keyword logic, just at scale, on your own account.

Reply to every comment, fast. Aim for inside about 90 minutes of the post going out. The first hour or two decides how far a post travels: early replies pull more comments, more comments tell the algorithm the post is alive, and it gets pushed wider right when it counts. It's also just decent. This is the cheapest reach you'll ever buy.

Do this now: Open a blank spreadsheet. Make five columns — **Date · Platform · Hook · File · Keyword** — and fill the next 7 days from what you already built. For every Instagram row, write one keyword and rewrite the caption to "comment [KEYWORD] and I'll DM it to you" — no link in the post. Schedule those 7 into the native app schedulers (or set them in Zernio with one comment-to-DM automation). Post the first today and reply to every comment within 90 minutes. That's your month started, and the loop that runs it.

The inputs: what you can drive with, and how each feeds the month

I keep saying "drive your raw material," and I mean broader than footage. Here are the five inputs you can seed a month from. The first is the proven spine; the rest widen what counts as a "drop." Every one of them converges on the same mechanic from Step 6: **one input, five recipes, fanned out into a month.**

1. Your real footage — the spine. One rich session of you talking or filming once: an interview, a walkthrough, a podcast you were on, a day of filming, or pure action b-roll. The transcript *and* the footage catalog become the supply: your top-5 starred moments become Spotlights, the rest become montage fuel, carousel slides, and supporting cuts. This is the load-bearing, proven path with a full shipped pipeline behind it (it's the Tiempo receipt below). It's *one example input*, and the only one with a complete end-to-end pipeline today. Not the only path, but the one I'd start you on.

2. AI-generated images trained on your brand profile. Your brand profile holds your colors, voice, sacred phrases, CTA, and what's *not* allowed, and every step reads it, so output comes out on-brand without you art-directing a thing. You generate on-brand images with **KIE.ai on your own key** and feed them in: carousel heroes, result shots, gap-fills for a shot you never filmed. Honest frame, because I won't oversell it: the brand profile tunes the *voice, colors, and CTA* for every step. It does **not** train an image model, and Nave has no shipped image-gen button. You generate with KIE and feed them in, and you use it only where you're actually missing something.

3. Format / carousel examples studied from the internet — references, not clones. Pull one outlier creator open, watch five of their best posts back to back, and name the *one* mechanical thing that's working: the hook, the format, the cadence. Then rebuild that mechanic with *your* own substance. **Copy the carpentry, keep the house.** References don't have to be competitors; pull craft from anyone who's an outlier at the mechanic. This shapes your format recipes and your hooks. To be clear, this is a human studying and directing, not auto-scraping and cloning someone's post.

4. Guided concepts / your direction — how YOU steer. You set the direction once: your topic, the angle, the brand colors and the alignment (what makes this entity who it is, why it exists, what success looks like) and the trained system holds that standard across every piece. This is the give-direction-not-grunt-work layer. You correct drafts rather than fill a 30-question form. Because every step is bound to the brand bible, the output stays on-brand without you steering each individual clip. This is the input that proves the premise: you drive, the system supports.

5. Optional AI avatars — a delivery target the script feeds. The system can write a spoken-word-only script you paste into an avatar tool like HeyGen; you take the script there and render the avatar. Honest frame: this writes the *script*. It does **not** render the avatar inside the system, and it never replaces the human in the driver's seat. One optional input among several, never the centerpiece. Mark this one *your call*.

Five inputs, one mechanic. Whatever you drive with, it converges the same way: mine it, cut it, edit it to your bar, run the five recipes, fan it across the month.

Design your own formats

Here's the thing I want you to take more than the list: **my formats are not your formats.** The five recipes above are a model, not a prescription. The real skill isn't running my five. It's *designing* the formats your business actually needs. And the rule that makes a format worth building is simple: **every format has a funnel job.**

Format	Funnel job
Spotlight	Trust → DM
Pulse / short	Reach + community
Compilation	Reshare / social proof

Format	Funnel job
Montage	Cold reach / discovery
Carousel	Authority → DM

Read that as a model, not a menu. Maybe your business needs a "client result" format and a "myth-buster" format more than it needs a montage. Build those. The method is the same for any format you invent: write the recipe once (the cuts, the captions, the cards, the music), then, if you're building the machine, give it a skill and a narrow agent, and it runs forever.

Steal the method, not my list. Pick the funnel jobs YOUR audience needs filled, name a format for each, write its recipe, and you own a content engine shaped to your business instead of mine.

Proof (honest receipts only)

I build first, then I teach, so here's the real receipt, and only what actually happened. No invented numbers.

One Tiempo footage drop, **5 player and coach interviews plus about 100 training clips**, became **12 finished, on-brand pieces** in one pass: 8 interview reels, 1 compilation, 1 montage, and 2 carousels. The client (Coach Fernando) reviewed the whole batch on a **2026-06-01** Google Meet and gave the green light to start posting. What he loved, in his words: the training **montage** and the **compilation** were "straight gasolina." Fire.

And here's the part that proves the premise better than any view count could: **it was human-in-the-loop, not push-button.** On that review, the client caught AI transcription errors (a player's name is **HYDEN, not "Hayden"**; a soccer term is "**lace shot, not "lay shot"**") and surfaced a crucial beat that had been dropped. That's the whole point. The human drives and holds the final taste. The system supports and holds the standard. A push-button would have shipped "Hayden." A driver doesn't.

The architecture under it is real and tenant-agnostic: **16 format agents**, each backed by a skill, all reading a shared footage-intelligence catalog and a brand bible. The footage-intelligence piece was promoted to a *global* substrate skill, because nothing about it is Tiempo-specific. You point it at any drop and it reads it. And the cost receipts are real: AssemblyAI Universal-3 Pro at about **\$0.27 an hour of audio**, all tools bring-your-own-key, paid directly, no markup.

One honesty guardrail I'll hold even though it'd be easy to fudge: **there are no published view, reach, or conversion numbers here.** That batch was being scheduled when this was written, not yet measured. I'm not going to invent engagement stats to sell you something. The receipt is "one drop, twelve finished pieces, client-approved." That's the true claim, so that's the only one I'll make.

You own the machine, you rent the fuel

Quick honest note on ownership, because I won't oversell this either. You own the machine: the repo, the recipes, the system. You rent the fuel: the tool costs, paid directly to AssemblyAI, KIE.ai, and Zernio, in cents, no markup. That's the real frame: **pay cost, not markup, and swap any single tool without re-buying the whole system.** If a better transcription engine shows up next year, you swap it in and the machine doesn't change.

I won't tell you "no vendor can ever throttle you," because that's not true. A heavy interactive Claude Code run can still hit your plan's limits, like anything else. What's true is the thing that actually matters: you're not renting your *output* from someone by the month, and no middleman is skimming your bill. You own the part that makes the content. That's the part you can't be priced out of.

This is also where I'll be generous about the tools that do the rented version of this: the Blotato/Sabrina crowd and the IG automation folks. They're genuinely good tools for people who want to *rent*. This playbook is for people who want to *own*. Same goal, different relationship to the machine, and which one's right depends on you, not on me dunking on anybody.

The honest Nave close

So here it is straight. By hand, this method is a full day, every month. I did it for months. Then I did what I always do: I built first, then I taught. I took the day I just handed you and wired it into a system that runs it for me. That system is **Nave**.

Nave doesn't know a secret you don't. It runs this *exact* method: same tools, same steps, your own keys, your own Claude Code subscription, no middleman skimming your bill. The only thing it changes is the clock. The day becomes a one-time setup, and after that it's about an hour a month. You drive the input, you say "make my month," you approve the slate, and it stages,

never auto-posts, because you stay in the driver's seat. Same output, a day's work compressed into an hour, on a machine you own and a repo you can read.

Nave isn't the gate in front of the method. The method is right here, free, finished. Close this PDF right now and never hear from me again, and you've still got everything you need to make a month of quality content with your own two hands, your own keys. I mean that. No "advanced version" held back, no good part paywalled.

If the hour sounds better than the day, the **\$89 template** that compresses it lives at **nave.infinixai.com**, and I'm the one who answers when you reply. But the thing I actually want you walking out with isn't the tool. It's the thing I built all of this to get back for myself in the first place: my time, and a system that's mine instead of rented by the month. That survives any tool rename. That's what I'm actually giving you.

Do this now: Decide which person you are this month. Want to own the day? Block a real one, title it "make my month," and run the seven steps end to end. You have all of them. Want to own the hour? Go to nave.infinixai.com and grab a setup seat. Pick one, put it on the calendar today, and stop renting your output.

Build first,

Enrique Marquez · @enriquemarq-0 · nave.infinixai.com