

So entstand unser Werbefilm „Der Satz vom Anfang“

Dieser Film soll Agenturen zu einem Termin bringen. Dass er das kann, liegt weniger an den Bildern als am Aufbau: Jeder Satz hat eine Aufgabe, und die Frage vom Anfang wird erst kurz vor Schluss beantwortet. Hier sehen Sie, wie wir ihn mit OMVIN gebaut haben, vom Vorbild bis zur fertigen Datei.



Länge	68 Sekunden, 9:16 und 1:1
Sprechertext	15 Sätze
Animation	30 Einstellungen, 24 Bilder pro Sekunde
Stimme	ElevenLabs Eleven v4
Beispiel-Spot	GPT Image 2.5 und Kling V3 Pro
Musik	Lyria 3 Pro
Geräusche	ElevenLabs Sound Effects v2

Ein Vorbild zerlegen, nicht kopieren

Wir haben nicht bei null angefangen, sondern bei einer Story-Anzeige aus einer ganz anderen Branche, deren Aufbau uns überzeugt hat. Der Agent hat sie in drei Durchgängen zerlegt: Bild (jede Einstellung mit Zeit, Schnitt und Bewegung), Psychologie (welche Aufgabe jeder Satz hat und wie viel Zeit er bekommt) und Ton (Musik, Geräusche, Sprechtempo).

Heraus kam eine Schablone mit 17 Takten und Platzhaltern. Vom Vorbild übernommen haben wir nur den Aufbau: keinen Satz, kein Bild, keinen Ton.

Vorbild analysieren VORSCHLAG

Das Video vorher im Reiter „Desktop“ hochladen.

Analysiere den Werbefilm in Desktop/vorbild.mp4 in drei Durchgängen. 1. Bild: jede Einstellung mit Start, Dauer, Inhalt und Bewegung. 2. Psychologie: Welche Aufgabe hat jeder Satz (Hook, Versprechen, Einwand, Beweis, Handlungsaufforderung) und wie lange dauert er? 3. Ton: Musik, Geräusche, Sprechtempo in Silben pro Sekunde. Mach daraus eine Schablone mit Platzhaltern in eckigen Klammern, die ich auf [mein Angebot] übertragen kann. Übernimm keinen Satz aus dem Film.

SCHRITT 2

Das Paradox finden

Die Schablone trägt nur, wenn das eigene Angebot ein echtes Paradox hat: Der Kunde tut etwas, das nach Verlust aussieht, und man gewinnt trotzdem. Ohne diese offene Frage hört niemand eine Minute lang zu.

Unser Paradox: Ein Kunde schreibt seiner Agentur „Machen wir jetzt selbst. Mit KI.“ Das klingt nach einem verlorenen Auftrag. Der zweite Satz verspricht das Gegenteil: „Ich zeig dir, warum er dann trotzdem dein Kunde bleibt.“ Aufgelöst wird das erst in Sekunde 47.

Das Paradox muss ohne Einschränkung stimmen. Bei uns steht es so in den Bedingungen für Agenturen: Der Kunde bleibt Kunde der Agentur, und die Agentur bekommt Provision auf seinen Zugang.

Die Formel VORSCHLAG

Dein Kunde [tut etwas, das nach Verlust aussieht], und trotzdem [gewinnst du etwas]. Ich zeig dir, warum.

Fakten und Verbote vor dem ersten Satz

Bevor ein Satz geschrieben wurde, hat der Agent alle Aussagen gesammelt, die wir belegen können, jede mit ihrer Fundstelle auf unserer Website: Preis, Bedingungen, Serverstandort, Vertrag. Dazu kam eine Liste mit Aussagen, die nie vorkommen dürfen. Was nicht auf der ersten Liste steht, kommt nicht in den Film.

Faktenbasis VORSCHLAG

Sammele alle Aussagen über [mein Angebot], die auf [meiner Website] belegt sind, jeweils mit Fundstelle. Schreib dazu eine Liste mit Aussagen, die wir nicht machen dürfen. Verwende im Konzept nur Aussagen aus der ersten Liste.

Konzept in vier Kritikrunden

Das Konzept haben wir nicht einmal geschrieben, sondern viermal. Ein zweiter Agent hat jede Fassung bis zur letzten als strenger Kritiker bewertet, mit Punkten und Begründung. Vier Dinge haben wir dabei gelernt:

Ein Werbespot, den jeder erwartet, ist kein Paradox. Eine frühe Idee hatte keine offene Frage und damit keinen Grund, dranzubleiben.

Rhetorik vom Vorbild ohne echtes Paradox wirkt leer. „Klingt zu gut, oder?“ trägt nur, wenn es wirklich zu gut klingt.

Die Reihenfolge entscheidet. Wer im Hook einen ganzen Auftrag verliert, hält 30 % Provision für einen Trostpreis. Deshalb kommt zuerst „Er bleibt dein Kunde“ und erst danach die Provision.

Der Satz des Kunden steht in einer Nachrichtenkarte, nie im Eingabefeld von OMVIN. Sonst wirkt das eigene Produkt ohne Ton wie der Gegner.

Kritiker VORSCHLAG

Bewerte dieses Konzept als strenger Kritiker aus Sicht von [Zielgruppe]: Wo steigt jemand aus, welche Aussage glaubt er nicht, was versteht man ohne Ton nicht? Gib Punkte je Takt, begründe sie und schlag konkrete Änderungen vor. Lob nichts, was nicht trägt.

Der Aufbau in 15 Sätzen

So sieht die Schablone fertig aus. Jede Zeile ist ein Takt mit genau einer Aufgabe. Die Zeiten sind die Startsekunden im fertigen Film.



0,4 s · Hook: das Problem des Zuschauers ab dem ersten Bild

„Wenn dein Kunde das sagt, klingt das nach einem verlorenen Auftrag.“



4,5 s · Paradox-Versprechen, die offene Frage

„Ich zeig dir, warum er dann trotzdem dein Kunde bleibt.“

8,0 s · Skepsis aussprechen, bevor der Zuschauer es tut

„Das klingt nach dem nächsten KI-Versprechen, oder?“



11,0 s · Marke zeigen statt behaupten

„Dann schau dir an, was in OMVIN gerade entsteht.“

14,1 s · Mechanismus in einem Satz

„Dort bauen KI-Agenten Websites, Werbespots und Posts für deine Kunden.“

20,2 s · Drehpunkt: die Lage des Zuschauers neu deuten

„Deine Kunden wollen KI so oder so. Die Frage ist nur, ob sie dafür zu dir kommen.“



26,1 s · „Stell dir vor“ mit Bildbeweis

„Stell dir vor, die Goldschmiede um die Ecke fragt nach einem Spot. Bisher hättest du vielleicht abgesagt, jetzt schickst du ihr einen Entwurf.“



37,8 s · Einwand in der Sprache des Zuschauers, sofort beantwortet

„Was kostet das, und wo liegen deine Projekte? Ein Zugang kostet neunundsechzig Euro netto im Monat, und deine Projekte liegen auf Servern in Deutschland.“



47,5 s · Auflösung: der Satz vom Anfang kommt zurück

„Bleibt noch der Satz vom Anfang. Will dein Kunde selbst ran, richtest du ihm sein Projekt ein. Er bleibt dein Kunde, und für seinen Zugang bekommst du jeden Monat dreißig Prozent Provision.“



60,1 s · Eine Handlung mit niedriger Hürde

„Buch dir jetzt einen unverbindlichen Termin, danach testest du dreißig Tage kostenfrei.“

Die offene Frage aus Sekunde 4 wird in Sekunde 54 beantwortet, fast mit denselben Worten: „Trotzdem dein Kunde.“ wird zu „Er bleibt dein Kunde.“

Zwei Bildebenen: helle Tafeln für Sätze, die hängen bleiben sollen, dunkle Szenen für den Inhalt.

Pausen nur an zwei Stellen: nach dem Paradox und nach dem Markennamen. Danach läuft der Film durch.

Eine einzige Handlung: Termin buchen. Dasselbe Wort steht im Sprechertext, auf dem Knopf und auf der Endkarte.

Jede Bedingung steht im Bild, während sie gesprochen wird, etwa „Nach den 30 Tagen 69 € netto (82,11 € brutto) je Zugang und Monat, monatlich kündbar.“

SCHRITT 6

Ein Sprechertext, der sich gut spricht

Ein Text kann gut zu lesen sein und trotzdem schlecht klingen. Diese Regeln haben wir beim Schreiben eingehalten:

Volle Sätze mit mindestens sieben Silben, das wichtigste Wort am Satzende.

Etwa 4,2 bis 4,4 Silben pro Sekunde. Das ist ruhiger als viele Creator-Videos und bleibt beim ersten Hören verständlich.

Zahlen so ausschreiben, wie sie gesprochen werden: „neunundsechzig Euro netto“.

Den Markennamen so schreiben, wie er klingen soll. Im Sprechertext stand „Omwinn“, damit die Stimme OMVIN nicht buchstabiert.

Eine Anrede durchhalten. Im Film durchgehend du.

SCHRITT 7

Die Stimme

Gesprochen hat die Stimme „Helmut“ aus der Stimmbibliothek von ElevenLabs, mit dem Modell Eleven v4, Sprache Deutsch, Stability 0,45. Die Regieanweisung steht in eckigen Klammern vor dem Text.

Bevor ein einziges Bild entstand, kam die Lesung: Erst wenn die Stimme passt, lohnt sich jedes weitere Bild. Wir haben den ganzen Text dreimal sprechen lassen und je Satz den besten Take genommen. Die Pausen setzt die Montage, nicht die Stimme. Nach dem Goldschmiede-Satz bleiben vier Sekunden frei für den Spot.

Lesung (Anfang) ORIGINAL

Stimme Helmut, ID g1jpii0iyvtRs8fqXsd1 · Modell Eleven v4 · Sprache Deutsch · Stability 0,45

[direct, warm, conversational, energetic, talking to one person] Wenn dein Kunde das sagt, klingt das nach einem verlorenen Auftrag. Ich zeig dir, warum er dann trotzdem dein Kunde bleibt.

Das klingt nach dem nächsten KI-Versprechen, oder? Dann schau dir an, was in Omwin gerade entsteht.

SCHRITT 8

Ein Film aus Code

Bis auf den Goldschmiede-Spot ist der ganze Film eine Animation, die der Agent als Webseite gebaut hat: Texte, Karten, das OMVIN-Fenster, Licht. Eine Zeitachse legt fest, wann was erscheint, und jede Einblendung hängt an einem gesprochenen Wort. Gerechnet wird der Film Bild für Bild, 24 Bilder pro Sekunde.

Der Vorteil: Jede Schrift ist echte Schrift, ohne KI-Tippfehler, und jede Änderung ist ein Satz an den Agenten statt eines neuen Drehs. Zuerst entstand ein Animatic mit Platzhaltern, das wir im Player mit Kommentaren an der Zeitstelle abgenommen haben. Erst danach wurde der Spot erzeugt.

Für den ganzen Film gelten dieselben Bewegungskurven, so wirkt er wie aus einem Guss:

Eintritt: `cubic-bezier(0.23, 1, 0.32, 1)`

Bewegung im Bild: `cubic-bezier(0.77, 0, 0.175, 1)`

Austritt: `cubic-bezier(0.32, 0, 0.67, 0)`, etwa drei Viertel so lang wie der Eintritt

Federn nur für Objekte wie Karten und Knöpfe, nie für Text und Zahlen

Untertitel mit höchstens fünf Wörtern, das gesprochene Wort hervorgehoben

Animation bauen VORSCHLAG

Bau den Film als HTML-Animation in 1080 × 1920 mit 24 Bildern pro Sekunde. Die Zeitachse kommt aus den Wortzeiten der Sprecherspur: Jede Einblendung hängt an einem Wort. Helle Tafeln für Signalsätze, dunkle Szenen für Inhalt, Untertitel mit höchstens fünf Wörtern. Halte das obere und untere Fünftel frei für die Instagram-Oberfläche. Rechne das Video danach Bild für Bild.

Der Beispiel-Spot: vom Bild zum Video

Der einzige gefilmt wirkende Teil ist ein kurzer Spot für eine erfundene Goldschmiede. Damit die drei Einstellungen zusammenpassen, gibt es eine Produktionsbibel: einen Hausstil, einen Schauplatz mit Lichtsatz und eine Kamera-Regel. Diese Sätze stehen wortgleich in jedem Prompt.

Keine Hände und keine Gesichter, das ist die häufigste Fehlerquelle; Werkzeuge kommen vom Bildrand ins Bild. Das obere Drittel bleibt dunkel und leer, dort liegt später der Text. Keine Schrift und kein Logo im Bild: Die erfundene Marke steht nur als Fußzeile „Beispiel · erfundene Marke“.



1 · Casting: drei Varianten in mittlerer Qualität (GPT Image 2.5)



2 · Schlüsselbild: die beste Variante per Bearbeitung in hoher Qualität, Lampe aus der Textzone entfernt



3 · Video: Kling V3 Pro macht daraus fünf Sekunden Bewegung, ohne Ton

Aus jedem Clip laufen im Film nur die sauberen Sekunden. Was sich verformt oder springt, fliegt raus.

Ein Bild, das fast passt, wird bearbeitet statt neu erzeugt. So bleiben Licht und Komposition erhalten.

Hausstil Bild ORIGINAL

Photorealistic feature-film still, anamorphic lens character, motivated low-key light from practical sources, atmospheric haze, natural surface texture.

Hausstil Video ORIGINAL

Real weight and inertia, cause before effect, objects keep their shape. Light direction, props and materials stay exactly as in the start frame. 180-degree shutter motion blur.

Casting: Ring im Lichtkegel ORIGINAL

GPT Image 2.5 · 1024 × 1536 · Qualität medium · 3 Bilder

Photorealistic feature-film still, anamorphic lens character, motivated low-key light from practical sources, atmospheric haze, natural surface texture, vertical portrait frame.

SCENE: the workbench of a small independent goldsmith's workshop at night, nobody in frame. A worn leather bench skin under the bench pin, dusted with fine gold filings, lies far out of focus behind the velvet pad, its gold filings turning into soft points of light.

LIGHT: one small brass task lamp camera-left, about 3200 K, hard pool of light on the bench pin, everything else falls off into deep shadow; the torch flame is the only other light when it burns; no fill light. In this shot the hard pool of light falls on a small velvet pad; the haze makes the lamp's beam faintly visible only right around the ring.

SUBJECT: a plain, wide gold ring with a hand-hammered surface and no stone stands upright on a small round pad of dark petrol-blue velvet, its lower edge pressed into the velvet pile, slightly below the centre of the frame. Every hammered facet catches its own small highlight from the lamp, the inside of the band glows warm, and the velvet around it falls off into darkness.

PALETTE: deep petrol-blue shadow, warm gold, brass, dark walnut, soft haze grey.

CAMERA: Macro and short-tele only: 100 mm macro and 50 mm, full-frame equivalent. Camera at bench height, slightly above the work surface. Locked-off tripod or one slow motorised slide per shot, never handheld. Main direction left to right.

Shallow depth of field in macro (f/2.8), f/5.6 at 50 mm. This shot: 100 mm macro at f/2.8, close-up at the height of the ring, the ring sharp, everything behind it soft. All key content inside the central 84 % of the width.

EMPTY AREAS: the upper 35 % of the frame stays dark, empty shadow with nothing in it; no signs, no labels, no packaging.

TEXTURE: fine dust and a few loose fibres on the velvet, tiny tool marks on the inside of the band, one gold filing caught in the pile.

CONSTRAINTS: no text, no letters, no numbers, no logos, no signs, no watermark; no hands, no fingers, no people, no faces; no glamorization, no retouching, no studio lighting.

Schlüsselbild: Bearbeitung ORIGINAL

GPT Image 2.5 Edit · Qualität high · Bild 1 ist die gewählte Casting-Variante

Image 1 is the image to edit.

CHANGE ONLY: remove the brass lamp shade at the left edge of the frame; where it was, show the same dark, empty shadow as the rest of the upper part of the frame. The lamp is now outside the frame camera-left; its warm beam still enters from the left, faintly visible in the haze, and lands on the ring exactly as before.

PRESERVE: everything else in Image 1 exactly: the ring, its shape, width, hammered facets, highlights, inner glow and three-quarter orientation; the velvet pad, its shape, colour and pile; the bench and every background object with its position and blur; light direction, colour temperature, contrast, haze, depth of field; camera position, framing and scale of every object.

CONSTRAINTS: do not zoom, do not crop, do not move the camera; do not add any object; no text, no letters, no numbers, no logos, no watermark; no hands, no fingers, no people.

Photorealistic feature-film still, anamorphic lens character, motivated low-key light from practical sources, atmospheric haze, natural surface texture.

Video: der Ring dreht sich ORIGINAL

Kling V3 Pro, Bild zu Video · 5 Sekunden · ohne Ton · Negativ-Prompt: blur, distort, low quality, text, logo, ring outline, extra objects, deformed, flicker, camera shake, zoom, morphing

Locked-off camera on a tripod, absolutely no camera movement, no zoom. The velvet pad turns slowly on a hidden turntable and carries the upright ring with it, about a quarter turn from left to right over the whole shot; its hammered facets catch the lamp one after another, so small highlights travel across the band. The ring keeps its exact shape and hammered pattern; the beam in the haze and the background stay still. End state: the ring, turned a quarter, stands upright in the centre of the pad. The dark upper part of the frame stays empty. Real weight and inertia, cause before effect, objects keep their shape. Light direction, props and materials stay exactly as in the start frame. 180-degree shutter motion blur.

SCHRITT 10

Musik, Geräusche, Mischung

Für die Musik haben wir drei Fassungen im fertigen Schnitt gegeneinander gehört. Gewonnen hat nicht eine der beiden, die wir eigens für diesen Film bestellt hatten, sondern ein Stück mit Lyria 3 Pro, das schon im Animatic als Platzhalter lief. Darum lohnt der Vergleich im Schnitt: Erst dort hört man, welche Musik Spannung macht.

Die Musik ist auf ihre Takte geschnitten, damit Einsätze auf Bildwechseln liegen. Direkt nach dem Markennamen setzt sie kurz ganz aus, vor der Auflösung dünnt sie aus. Der Sprecher liegt vorn, die Musik deutlich darunter; der fertige Film hat -14 LUFS.

Die Geräusche des Spots kommen von ElevenLabs Sound Effects v2, je ein Klang pro Ereignis. Für Klicks, Wischer und Tippen haben wir Geräusche aus einem früheren Film wiederverwendet. Jeder Klang sitzt genau auf dem Bildwechsel, zu dem er gehört.

Musik ORIGINAL

Lyria 3 Pro

Instrumental film score, no vocals, 84 BPM, cold, sly and tense with a touch of irony: a low pulsing synth bass, sparse muted piano notes with long silences, a quiet ticking texture like a meter running, slowly tightening, ending on an unresolved held note.

Geräusch: Ringklang ORIGINAL

ElevenLabs Sound Effects v2 · 3 Sekunden · Prompt-Einfluss 0,6

A solid gold ring struck once lightly with a tiny steel rod, a single pure bright ringing tone with a long smooth shimmering decay, close perspective, dry close-mic foley, no reverb tail, one-shot

SCHRITT 11

Prüfen, kennzeichnen, ausliefern

Bevor der Film rausging, haben wir diese Punkte abgehakt:

KI-Kennzeichnung ab dem ersten Bild: „Stimme, Bild und Musik KI-generiert“.

Jede Zahl und jede Bedingung steht so auf unserer Website.

Stumm verständlich: Wer ohne Ton schaut, liest jeden wichtigen Satz als Tafel oder Untertitel.

Jeder Zwischenstand abgenommen im Player, der Agent setzt die Kommentare um.

Zum Schluss eine zweite Fassung in 1:1 für den Feed, mit demselben Ton und denselben Ereignissen.

