

دليل خطوة بخطوة + كل البرمببات

سوّ فيديو بشخصيتك الكرتونية

من صورة وحدة وصوتك... بدون ستوديو وبدون كاميرا.



شنو تحتاج؟

هذي الطريقة اللي سويت فيها فيديو «هالفيديو كله ما تصوّر بكاميرا». ما أوعدك بخمس دقائق، أول مرة تحتاج تجربة وتعديل وصبر.

١. صورة وحدة لوجهك

واضحة، إضاءة جيدة، بدون نظارة شمس، وجهك قدام الكاميرا.

٢. ChatGPT (أو أي أداة صور)

تحوّل الصورة لشخصية كرتونية وتسوي لها وجوه وخلفية. تحقق من خطتك لأن توليد الصور له حدود.

٣. ElevenLabs

تستنسخ صوتك من تسجيل قصير (Instant Voice Clone). هذي الميزة تحتاج خطة مدفوعة، تأكد من خطتك قبل لا تسجل.

٤. Claude

يحوّل الصور والصوت لفيلم: يفصل الشخصيات من الخلفية الخضراء، يحرك الفم على الصوت، يخلي الشخصيات تتحرك وتمثّل، ويسوي حركة الكاميرا والمؤثرات. أنا استخدمت Claude Code. وإذا ما تبّي كود تقدر تركّب نفس الصور والصوت في CapCut أو Canva.

٥. دقيقتين تسجيل لصوتك

في غرفة هادية، والجوال قريب من فمك.

الفكرة بجملة وحدة: صورة وحدة + صوتك + AI = أنت شخصية في قصة، بدون ستوديو وبدون كاميرا.



الصورة ← الشخصية

ارفق صورتك في ChatGPT، وإذا تقدر ارفق معها صورة من فيلم أنميشن كمرجع للاستايل (هذا اللي يفرق بين كرتون وصورة معدلة).

البرومبت (انسخه كما هو، وغيّر الأقواس)

Turn the attached photo into a stylized 3D animated character in the style of the attached style reference: soft Pixar-like 3D, NOT a photo and NOT photorealistic.
Stylized proportions: head ~30% larger relative to the body, large glossy expressive eyes, small simple nose, soft rounded cheeks, smooth matte clay-like skin with no pores, hair and beard simplified into clean sculpted shapes.
Keep my [glasses / beard / hairstyle / skin tone] so it is clearly me.
Outfit: [your outfit].
Chest-up portrait, slightly 3/4 angle, neutral relaxed face, mouth closed.
Soft neutral lighting, no colored light.
Background: perfectly flat solid bright green (#00FF00), no shadows on it, no green anywhere on the character. Portrait 1024x1536.

إذا طلعت النتيجة شبيهة بالصورة أكثر من اللازم

اكتب له: «More stylized. Push it toward an animated film character: bigger eyes, smoother simplified shapes. Reduce all photo detail. والبرومبت هذا هو اللي صلّحها.

ليش خلفية خضراء؟ عشان نفصل الشخصية عن الخلفية بسهولة وبدون حواف. لا تخلي أي لون أخضر في ملابس الشخصية.



الوجوه: الفم والعين

الفيديو يتحرك بتبديل صور الوجه: فم مفتوح، فم عريض، فم مدور، رمشة. كل صورة في رسالة منفصلة، في نفس المحادثة عشان تبقى الشخصية نفسها.

ابدأ كل رسالة بهذي الجملة

Same character, same outfit, same framing, same head angle and position, same lighting, same flat green background. Change ONLY the face as described.

وبعدها سطر واحد في كل رسالة:

- 1 Mouth open wide as if saying "ah".
- 2 Mouth wide and flat as if saying "ee".
- 3 Mouth small and round as if saying "oh".
- 4 Eyes closed (a blink), mouth closed.
- 5 Surprised: eyebrows raised, eyes wide, mouth open in an "o".
- 6 Thinking: eyes looking up and to the left, mouth closed.

إذا تغيّر الإطار بين صورة وصورة: أعد الطلب وقول «keep the exact same framing as the master». فرق صغير (بكسل أو اثنين) أقدر أعذّله بعددين، بس فرق كبير لا.

شنو تتوقع

الصور الست تكفي لفيديو كامل. الفم هو الأهم، والرمشة تخلي الشخصية «حية».



الجلسة والمكان

الشخصية التي تطفو بالهوا تبين كأنها ستيكر. خلها جالسة في المكان. الوجه القريب يكفي للقطات الحميمة، وصورة الجسم كاملة للقطات الواسعة.

الشخصية جالسة (جسم كامل، مربع 1024×1024)

Same character as the master image, same Pixar-like 3D style, same outfit. FULL BODY, sitting cross-legged on the ground in a relaxed storyteller pose, body in a 3/4 view turned slightly toward the right, hands resting on the knees, friendly closed-mouth smile. Show the whole body including feet. No cushion, no ground, no furniture, no shadow. Soft neutral lighting. Background: perfectly flat solid bright green (#00FF00).

المشهد (غير المكان لمكانك: بحر، كافيه، مكتب...)

Stylized 3D animated film environment, same Pixar/Disney-like soft 3D style as the attached character. Starry desert night, vertical 9:16, wide establishing shot: a warm-lit tent on the left, a campfire in the center, two empty cushions facing each other by the fire, big dunes, the Milky Way above. Warm orange firelight on the sand, tiny sparks rising. No characters, no text.

خلفية القطة القريبة (ضبابية، الوسط فاضي للشخصية)

Same world and style, vertical 1024x1536. Low camera angle near the ground, looking toward the warm-lit tent on the left, softly out of focus. The campfire glows as orange bokeh in the lower-left corner with floating sparks. Shallow depth of field, cinematic. The center and right of the frame are calm and uncluttered because a character sits there. No characters, no text.



أشياء ثلاثية الأبعاد بدل الكلام

بدل ما تكتب كلمة على الشاشة، خل الفكرة تطلع كشيء صغير يطير بجانب الشخصية: جوال للصورة، مايك للصوت، لابتوب لـ Claude.

A set of 9 separate small objects in the same Pixar-like 3D toy style, soft glossy materials, warm and friendly, each one on its own with wide empty space around it, arranged in a 3x3 grid on a perfectly flat solid bright green background (#00FF00), no shadows: 1) a smartphone with a selfie photo on screen, 2) a camera, 3) a face-emoji card showing several expressions, 4) a painted canvas with a starry desert scene, 5) a microphone with small sound waves, 6) a laptop with a glowing sparkle on screen, 7) a film clapperboard, 8) a gift box with a ribbon and a PDF document sticking out, 9) a glowing lightbulb. No text on any object.



نصيحة: إذا ChatGPT قرّب الأشياء من بعض، اطلب 5 في الصورة بدل 9. وكل شيء يطلع بدون نص أحسن، عشان الكتابة تضيفها أنت بعدين.



سجل دقيقة إلى ثلاث دقائق، وصوتك الحقيقي هو الذي يوصل. الاستنساخ ينسخ الطاقة التي سجلت فيها، فلا تقرأ بصوت ثابت.

خطوات ElevenLabs

Instant Voice Clone ← Add a new voice ← Voices ← ارفع التسجيل، سم الصوت، أكد إنه صوتك ← انسخ الـ Voice ID.

وصفة التسجيل

- غرفة هادئة، بدون موسيقى ولا مروحة. الجوال قريب (٢٠ سم) وبزاوية قليلة عشان ما تطلع نفخة.
- نوع: جملة هادئة، جملة متحمسة، سؤال، تفاجؤ، جملة جدية، وجملة تحكي فيها قصة.
- اكتب الكلمات مثل ما تنطقها بلهجتك، واخل أسماء الأدوات بالإنجليزي (Claude, ChatGPT).
- إذا تعثرت، كرر الجملة وكمل، لا توقف التسجيل.

للنص: تافات المشاعر (استخدمها بخفة)

[excited] [laughs] [curious] [whispers] [proud] [shocked]

مثال: [excited] هالفيديو كله... ما تصوّر بكاميرا!

صوت الشخصية الثانية (مثل الروبوت)

استخدم Voice Design في ElevenLabs واكتب وصف صوت. انتبه: كلمة «child» و«kid» تنحظر. أنا
a small, cute, playful cartoon robot sidekick, light and bright high-pitched voice, warm, صفته:
«very expressive, Kuwaiti Gulf dialect».

اسمع قبل لا تعتمد. لهجتك ولهجة الشخصية يحكم عليها أذنك، مو أنا ولا الأداة.



Claude يحولها لفيلم

لا تعطيه خطوات، أعطه الهدف والإحساس والقيود، وخله هو يقرر شلون، أول نسخة سويتها بتعليمات خطوة بخطوة، وطلعت الشخصيات ثابتة مثل الستيك. لما عطيته الهدف بس، صارت تتحرك وتمثل.

برومبت ل Claude (Claude Code أو أي نسخة تقدر تشغل كود)

Goal: turn this into a fun animated short where my characters feel alive, like a real animated film, not stickers on a background.
In the folder [path]: my characters as PNG images on a flat green background (neutral, mouth "ah" / "ee" / "oh", blink, surprised, thinking; chest-up and full body sitting), the background images, and one audio file per spoken line.
The viewer should laugh at least twice and finish thinking "this was made without a studio?".
Choose whatever tools, techniques, shots, timing and camera moves give the strongest result.
Rules: keep the characters looking exactly like the images. Sound effects only, no music. Vertical 1080x1920, 30 fps. Ask me before anything that costs money.
Build the whole film, check it yourself, then show me.

شنو سوّي عندي من نفس الصور الثابتة

الراس يلتفت ويهز مع الكلام، الشخصية تتنفس وتنط وتنضغط، والهوائي يتمرجح. الكاميرا تتحرك مثل الأفلام: زوم، لف، وقطع سريع بين اللقطات. ونكت بصرية: العالم يصير رسمة رصاص، والمكان بيني نفسه كلمة بكلمة.

شنو لازم تتأكد منه قبل النشر

- الفم بيدّل حسب الصوت، والشخصية تبقى نفس الشكل.
- ما في حواف خضراء على الشخصية.
- اسمع كل سطر: اللهجة، الأرقام، أسماء الأدوات.
- إذا ما تبي كود: CapCut أو Canva، وبدّل صور الفم يدويًا. ياخذ وقت أكثر بس يشتغل.

أخطاء صارت معي، عشان ما تكررها

١. أول شخصية طلعت صورة مو كرتون

الحل: صورة ستايل مرجعية + جملة «NOT a photo».

٢. شارب مزدوج عند تبديل الفم

لما بدلت جزء صغير بس من الوجه طلع شبح. الحل: أخذت منطقة الفم كاملة (الشفاف والشارب والذقن) من صورة الفم، وخليت الحواف تتدرج.

٣. ElevenLabs رفض وصف الصوت

بسبب كلمة «child». غيّرتها لـ «small, cute cartoon robot».

٤. الفيديو طلع أطول من المخطط

خططت لـ ٥٥-٦٥ ثانية، والكلام وحده طلع ٦٨ ثانية (الفيلم كامل ٧٧). اقطع النكتة اللي ما تخدم الفكرة.

٥. عطيته تعليمات أكثر من اللازم

قلت له كل خطوة: بدّل الفم، اقطع هني، زوم هني. طلعت الشخصيات ثابتة. لما عطيته الهدف والإحساس والقيود بس، هو اختار الطريقة وصارت تمثّل. القاعدة: أنت تحدد **شنو وليش**، وهو يقرر **شلون**.



جرب، وإذا طلعت لك شخصيتك شاركني

أبي أشوف شنو سويت. ما أضمن لك نتيجة من أول مرة، بس الطريقة تشتغل.

@aliturnover