



Tassonomia dei tool di *gen AI*

**Text 2 Text** (sintesi, traduzione, domande e risposte, creazione contenuti, rephrasing, creazione slide-email).

Tool Text 2 Text: [chatGPT-4o](#), [Claude 3.5 di Anthropic](#), [Gemini di Google](#) (ex Bard e Lamda); [training umano!](#); [DeepL Pro](#); [email assistants](#): [Google MED-Gemini](#) (con image analysis); [esempio elezioni americane](#).

**Text 2 Office** (XLS, PPT): [creazione slide PPT da prompt](#);

**Text 2 Image**: Dall-e 3 di OpenAI (integrato in chatGPT 4o), [Midjourney](#), [Stable Diffusion](#) (open-source); [face generation](#); [Grok image generator](#); [risk of model collapse](#); [virtual dressing](#); [Imagen3 di Google](#); [dreamlab di canva](#); [state-of-the-art video and image generation with Veo2 and Imagen 3](#); sfottò cinese;

Creazione diagrammi (direttamente o tramite Mermaid o plantUML): [chatGPT4o](#) di OpenAI; [accesso a Stable Diffusion tramite Python](#); nuovo o1 Image Generator di OpenAI (marzo 2025) → vedi sezione apposita. [ReveAI \(startup californiana, buona qualità, 20 immagini gratuite al giorno - Gaito\)](#)

**Text 2 Diagram**: [napkin.ai](#)

**Image 2 Image**: [Adobe Photoshop](#) e [Canva](#) per inpainting/outpainting; esempio di [image composition](#) con Adobe; [visualize cloting](#); [change positions of your eyes](#); [provare abiti con Google doppi](#);

**Image 2 Text** (*Image Captioning*): chatGPT4o ([esempi](#): orari parcheggi, sblocco sedile bicicletta, analisi specifiche Nvidia, quadro, device scheda audio, comando DOS, assistenza su dispositivo medico); [comprendere immagini nei PDF con Claude 3.5](#); [comprendere immagini nei PDF con chatGPT](#); [comprendere immagini con Google Ai Studio](#); tool di lettura immagini nei file PDF

**Image 2 Code**: [creazione mockup sito da lavagna](#);

**Text 2 Code**: [Gamma for website building](#); [Imagica per app](#); [AI-assisted coding con cursor](#); [Text 2 SQL \(con vanna\)](#); [Text 2 Site \(con Wordpress\)](#); [strumenti specifici per generare codice](#); assistenti AI integrati negli ambienti di sviluppo (Jupyter Notebook, Google Colab, ecc); GitHub in VS Code; [assistenti CLI Per T2C e C2T con chatGPT](#) e riflessioni generali sul vibe coding vedi la sezione «Vibe Coding» del PPT «1. ChatGPT – operativo»

**Code 2 Text** Vuol dire 2 cose: a) fornisco allo strumento un pezzo di codice e lui mi dice cosa esso fa? b) fornisco allo strumento un pezzo di codice con errori e lui li trova. C2T spesso non funziona bene. Esempio del primo tipo: [questa chat di Claude](#) genera un codice e lo anonimizza, la [seconda chat](#) (di chatGPT) cerca di capire il funzionamento.

**Text 2 Video** (generazione completa, cioè solo il prompt, senza editor e senza avatar): [Sora di OpenAI](#); altri due buoni tool di AI gen **T2V** sono [Runway](#) e [Pica](#) (Image2Image per face swap?); [veo2 di Google](#); [veo2 di DeepMind](#); state-of-the-art video and image generation with Veo2 and [Imagen 3](#); [sora vs veo3](#); [test Corriere della sera di riconoscimento video falsi di veo3](#); [demo di Kira](#) (perfetto *lip sync*); [fake di JFK](#); [GPT per prompt veo3](#); [veo3 con audio](#)

**Text 2 Video** (generazione parziale tramite editor con avatar): [Synthesia](#), [Fliki](#) (video Focus); [Heygen](#) (l'altro leader di mercato insieme a Synthesia: le funzioni base gratuite sono migliori in Heygen, le funzioni avanzate a pagamento sono migliori in Synthesia); esempio [pubblicità televisiva](#); [esempio InVideo AI v3](#); [generazione video \(con avatar\) multi-tool](#); [video di confronto avatar](#); [GPT sulla creazione T2V \(con veo3\)](#)

Novità!

**Text 2 Music:** vedi [slide ad hoc](#)

Novità!

**Text 2 Meme:** [supermeme.ai](#) (usato da DataPizza)

**Speech 2 Text (analogia umana)**: estensioni Chrome per YouTube, ad esempio: «YouTube Summary with ChatGPT & Claude»; OpenAI Whisper; [Otter.ai](#); [riverside.fm/transcription](#); [fathom.video](#)); [AssemblyAI](#); [Microsoft 365](#); [sentiment analysis vocale con AssemblyAI](#); [lo stato dell'arte S2T è ElevenLabs](#)

Novità!

**Text 2 Speech:** [il modello di OpenAI](#); [MeloTTS](#); [Scribe di ElevenLabs](#); la traduzione simultanea di [Google Meet](#); [Amazon Polly](#)

**Speech 2 Speech:** traduzione simultanea, anche multilingua (ad es. con chatGPT, [DeepL Voice](#), oppure con la app di Google Play [Voice Translator AI Translate](#)); [clonazione voce?](#); [ElevenLabs](#); [YouTube Aloud](#); [esempio WPC23](#); [esempio traduzione simultanea su TikTok](#); [YouTube auto-dubbing](#); [voice cloning di deepgram](#); [elevenlabs S2S](#) (ora Voice Changer).

**Video 2 Video:** esempio *motion transfer* («[Everybody dance now](#)»), [esempio Linkedin Generative AI](#); [esempio AnimateDiff](#); [Kendall Janner](#); [conferenza Adobe Max 2023](#); [gender/age transformation](#); [video transformation](#)

**Video 2 Text:** Whisper di OpenAI fa anche questo (vedi il notebook su Whisper); [CloudIA](#): da video nella «lingua dei segni» a testo; [esempio con GPT-4 Turbo with Vision \(Microsoft\)](#)

**Video 2 Image:** [mappe mentali con MyMap.ai](#),

**Image 2 Video:** [avatar](#), esempio [alta espressività](#); esempio [aging-deaging](#); esempio [ballo della vittoria](#); [esempio aging da Linkedin](#). In generale I2V (a partire da un'immagine propria o generata con T2I) funziona meglio che T2V; [confronto strumenti](#).

**Image & Audio 2 Video:** esempio [Berkeley](#); esempio [Adobe](#); [avatar avanzati di VASA](#)

**Text & Image 2 Video:** [Dream Machine di Luma AI](#); [Flux AI Vision Generator](#);

**Text 2 Audio:** [audio overviews con NotebookLM](#)

**Text 2 3D:** [meta 3D gen](#); Matrix di Alibaba; [Llama](#); Blender; Nvidia, Adobe (WPC24); [Google Genie](#); **Image 2 3D:** [Lorenzo Manzoni \(AI360\)](#);

**Catene di tool** (analisi dimensionale: entrambi i termini dell'equazione devono avere la stessa unità di misura – R. Stutzke):

- T2V con OpenAI chatGPT (T2T) + Synthesia (T2V);
- S2S con AssemblyAI (S2T) + OpenAI chatGPT (T2T) + ElevenLabs (T2S) – vedi DDoDS 2.12.2024 e slide;
- demo Deltatre su spatial computing (WPC24).
- [AI Speech classifier di ElevenLabs](#)
- [Demo AI week](#) (5 tool di gen AI in cascata)

**Image 2 3D:** [SAM3D di Google](#)

**Spatial Computing – Text to 3D - Realtà aumentata:** [WPC24/ComoNext](#), [Matrix-Game 2.0](#)

**Text 2 Robot Motion:** [esempio brainlink](#)

Novità!

Agenti Visuali: [anam](#), [tavus](#)

Novità!

Suite AI di [Amazon](#) ([Bedrock](#), [Sagemaker](#), [Q](#) e [Polly](#))

Novità!

[Study Mode](#)

Novità!

[Google AI Studio](#); [piani Google](#)

Novità!

[NotebookLM](#)

[audio overviews con NotebookLM](#)

Novità!

analisi immagini nei  
file PDF

## Tassonomia di modelli di gen AI

Models

Tasks

al 8.01.2025

Tasks Libraries Datasets Languages Licenses Other

Filter Tasks by name

Multimodal

- Audio-Text-to-Text
- Image-Text-to-Text
- Visual Question Answering
- Document Question Answering
- Video-Text-to-Text
- Any-to-Any

Computer Vision

- Depth Estimation
- Image Classification
- Object Detection
- Image Segmentation
- Text-to-Image
- Image-to-Text
- Image-to-Image
- Image-to-Video
- Unconditional Image Generation
- Video Classification
- Text-to-Video
- Zero-Shot Image Classification
- Mask Generation
- Zero-Shot Object Detection
- Text-to-3D
- Image-to-3D
- Image Feature Extraction
- Keypoint Detection

Natural Language Processing

- Text Classification
- Token Classification
- Table Question Answering
- Question Answering
- Zero-Shot Classification
- Translation
- Summarization
- Feature Extraction
- Text Generation
- Text2Text Generation
- Fill-Mask
- Sentence Similarity

Audio

- Text-to-Speech
- Text-to-Audio
- Automatic Speech Recognition
- Audio-to-Audio
- Audio Classification
- Voice Activity Detection

Tabular

- Tabular Classification
- Tabular Regression
- Time Series Forecasting

Reinforcement Learning

- Reinforcement Learning
- Robotics

Other

- Graph Machine Learning

# Aggiornamenti sulla genAI da Tikok (novembre 2024)



Inpainting (*I2I*)



Inpainting (*I2I*)

## Voice Translator: AI Translate

FeaturedApp

Contiene annunci

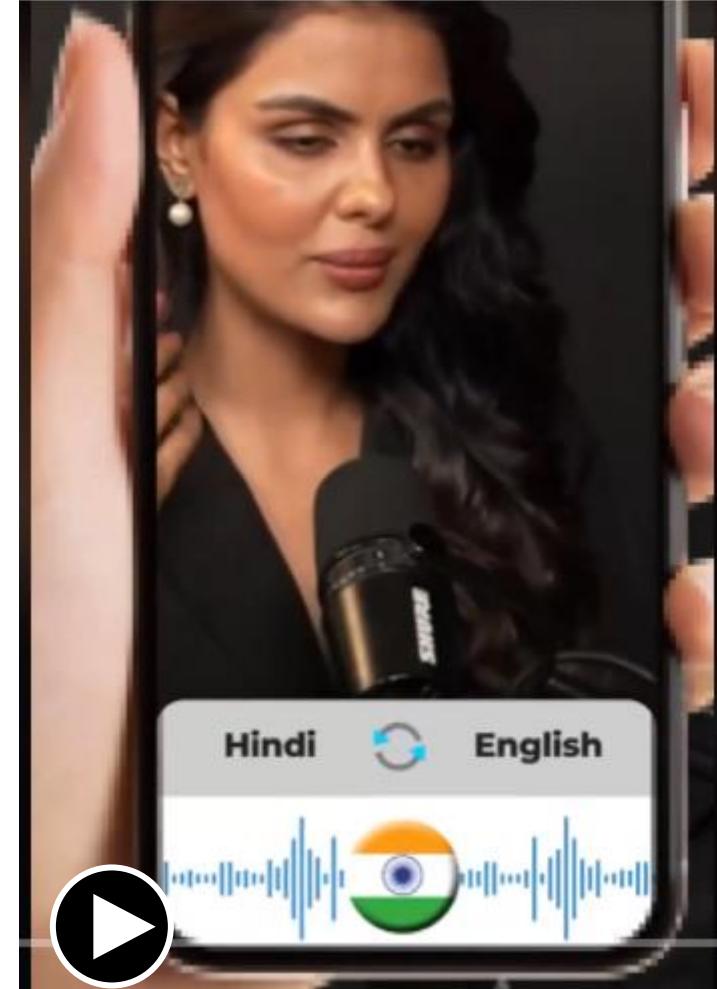
10 Min+



Download



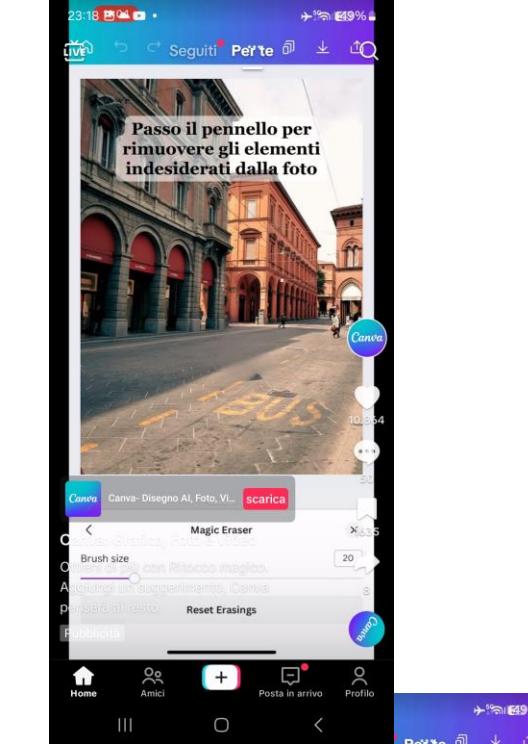
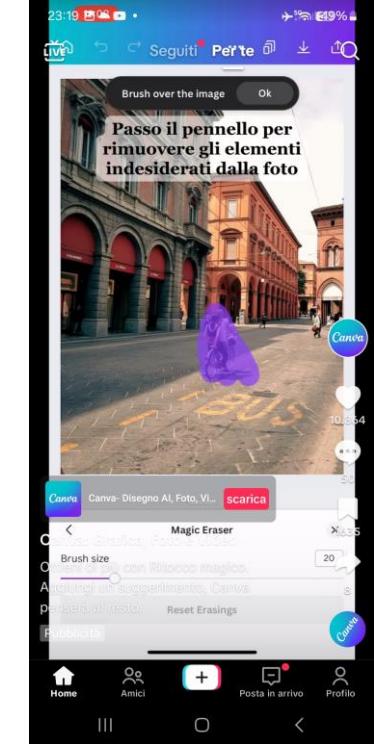
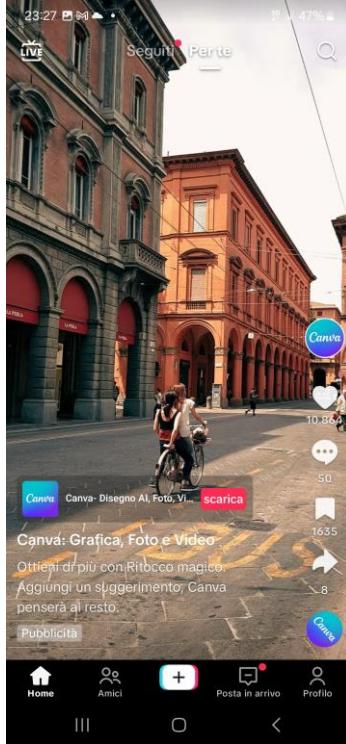
PEGI 3



traduzione multi-lingua  
*Speech 2 Speech*

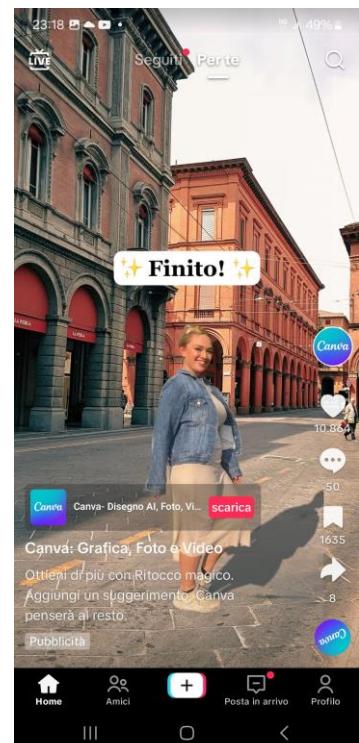
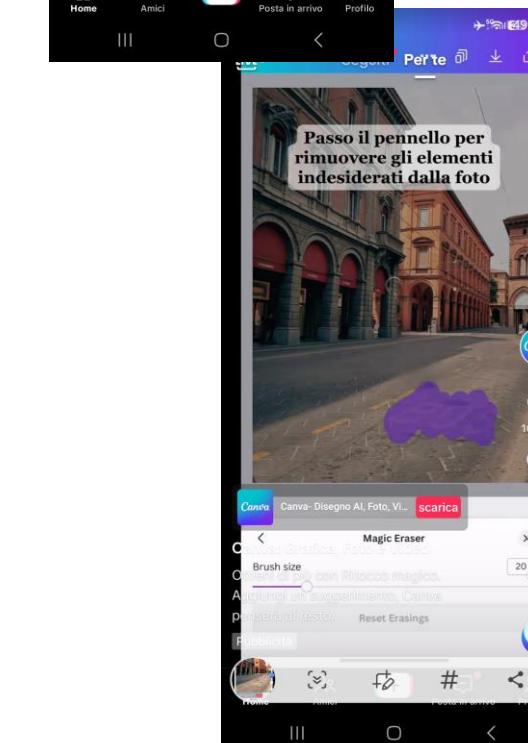


back to tassonomia



novembre 2024

## *Image 2 Image*: inpainting



[back to tassonomia](#)

*text 2 image*



# 1. Face Generation



back to tassonomia

<https://thispersondoesnotexist.com/>

- DEMO: <https://www.whichfaceisreal.com>
- Nvidia: StyleGANv2
- Nvidia: Alias-Free GAN —>StyleGANv3

One is real, the other is fake (sometimes not making sense a bit, especially about eyes and teeth; the background makes sense? ... → weird convolutions!)

} two great algorithms for face generation from Nvidia, based on GAN

Alcuni spot realizzati con l'AI sono stati trasmessi durante il big match dell'ultimo turno di Serie A. Tra cartelloni pubblicitari ancora rudimentali e propaganda politica, l'uso dell'intelligenza artificiale nella pubblicità è ancora un settore senza regole ma già in una fase esplosiva



Scatta la fine del primo tempo. Seduti sul divano, i tifosi che aspettano il fischio d'inizio di Inter-Napoli si sorbiscono minuti di pubblicità. Poi eccolo, lo **spot di un'azienda di scommesse**. Le immagini scorrono veloci, forse troppo perché il tifoso italiano possa rendersene conto. Davanti agli occhi degli spettatori ignari c'è **una pubblicità realizzata con l'intelligenza artificiale**.

Il volto di una donna su un corpo maschile, una pallina da tennis che non si allinea alla racchetta, mani distorte, espressioni innaturali. È difficile trovare certezze, ma un dubbio (fondato) si insinua: anche nelle pubblicità nostrane sta arrivando l'intelligenza artificiale. Ed è destinata a rimanere.

# Durante Inter-Napoli gli italiani hanno scoperto le pubblicità fatte (del tutto) con l'intelligenza artificiale

di Velia Alvich

## Le pubblicità realizzate con l'IA

Casi come quello di una nota catena di supermercati fanno persino sorridere. Sul cartellone si vede in primo piano una giovane donna mentre scatta un selfie. Fin qui **nulla di strano, specie se si guarda di fretta**. Ma non appena ci si sofferma su quello che succede alle sue spalle, ecco i problemi. **Volti che sbucano fuori dal nulla e mani che non hanno alcun senso**. Ecco ancora vecchi errori delle intelligenze artificiali tornati all'azione.

Forse il cartellone incriminato è stato realizzato quando le intelligenze artificiali generative erano ancora agli albori e avevano bisogno di essere raffinate. Tutt'altro risultato, invece, per le pubblicità di Mango. La catena di moda femminile ha annunciato a fine ottobre di voler

**rimpiattare le modelle delle pubblicità con avatar generati con l'IA**. «Si tratta di una creazione di contenuti più rapida», ha detto il Ceo Toni Ruiz a Bloomberg. Sull'immagine dell'ultima campagna spicca una giovane ragazza (la pubblicità è rivolta alle adolescenti) dai lineamenti mediterranei generata da un algoritmo. Almeno i vestiti, però, sono reali.

[link all'articolo](#)

*Text 2 Video*



La pubblicità della catena di fast fashion spagnola Mango. La modella è stata realizzata con l'IA, i vestiti sono reali



[back to tassonomia](#)

# Text2Music – 1/6

aprile 2024  
(update dicembre 2024)

## 1. riffusion

Si digita il testo, si sceglie il genere musicale -->  
si ottiene musica, voce, immagine sfondo.

## 2. udio

il parere di un esperto del settore



## 3. Suno v4

[HOME](#) > [Intelligenza artificiale](#) > [Articoli](#)

### L'IA MUSICALE È INCREDIBILE: SENTITE LE CANZONI CHE CREA, ANCHE IN ITALIANO

Abbiamo provato uno dei servizi di intelligenza artificiale del momento: si tratta di Suno, uno strumento IA che può generare canzoni in pochi secondi.

link  
all'articolo



### Udio: un'AI ASSURDA per creare musica (anche cantata) [Demo]



Raffaele Gaito

Admin

3d

Founder Growth Program

Il settore della musica generata dall'AI ha appena fatto un salto in avanti enorme con l'arrivo di Udio.

Si tratta di un tool di intelligenza artificiale capace di creare una canzone a partire da un prompt testuale.

La qualità è decisamente migliore rispetto ad altri tool in circolazione (incluso Suno AI) e la grande novità è che riesce anche a creare la parte cantata.

Eh sì, Udio crea una canzone completa, scrivendo anche il testo e cantandolo.

Oppure, volendo, possiamo caricargli noi il testo della canzone.

In questo video vi mostro una veloce demo con le funzionalità principali e provo a creare un paio di canzoni.

Buona visione 😊

This w  
Read c



back to tassonomia

## 🎵 Suno's Latest AI Music Generator



Suno has unveiled Version 4 (v4) of its AI music generation platform, introducing significant enhancements that elevate the quality and versatility of AI-composed music. This latest iteration enables users to create complete songs, including vocals, from simple text prompts, with improvements in audio clarity, lyrical sharpness, and dynamic song structures.

### Key Features of Suno v4:

- **Covers:** Users can upload any audio track and reimagine it through custom prompts, allowing the system to generate a fresh interpretation while preserving the original essence.

- **Personas:** This feature allows the creation and saving of distinctive musical identities, capturing unique characteristics such as vocal style, musical approach, and overall atmosphere for consistent application in future compositions.

- **Remaster:** Users can transform existing tracks created with previous versions to align with v4's superior audio quality while maintaining their core elements.

- **Lyrics by ReMi:** An AI-powered lyrics assistant designed to enhance the songwriting process, accessible by selecting "Write with Suno" in Custom Mode's lyrics mode and choosing ReMi from the model selection.

- **Cover Art:** Enhanced visual generation capabilities produce more sophisticated and stylistically matched artwork for tracks.

These advancements mark a significant leap forward in AI music creation, blurring the line between AI-generated and human-made music, and opening new possibilities for musicians and creators at every level.

fonte





Valerio Treviso



23h

Posted in Novità e Rilasci



...

## **Udio annuncia "Audio Uploads", Suno rilascia "Audio Inputs"**

A distanza di pochi giorni l'uno dall'altro, i due principali tools text-to-music Udio e Suno rilasciano la stessa funzionalità: la possibilità di caricare audio proprietari come integrazione a quelli generati dall'intelligenza artificiale.

Con **Audio Uploads** ora agli abbonati ai piani Standard e Pro di **Udio** possono caricare audio proprietari per un controllo maggiore sulle musiche generate: il modello udio-32 consente di utilizzare tutte le funzionalità di estensione, fino ad un lunghezza massima dell'audio di 15 minuti.



back to tassonomia

# Le grandi case discografiche fanno causa a Suno e Udio, due IA che generano musica

di Pier Luigi Pisa



*Il procedimento legale è stato avviato dalla RIAA (Recording Industry Association of America), che ha agito per conto di Sony Music, Warner Music e Universal Music. Creando nuova musica a partire da brani esistenti, Suno e Udio violerebbero il diritto d'autore*



Avv. Giuseppe Croari – Dott.ssa Francesca Gattarello

# Suno e Udio: quando l'AI rivoluziona la musica e sfida il copyright

*Suno e Udio: piattaforme AI che creano musica di alta qualità, suscitano polemiche per presunta violazione del copyright usando canzoni protette per addestrare algoritmi.*

## **Perché l'AI per creare musica violerebbe il copyright?**

La possibilità di creare musica attraverso l'AI ha creato numerose perplessità in capo alle grandi case discografiche, alcune delle quali hanno intentato delle cause nei confronti di Suno e Udio per **violazione del diritto d'autore**.

In particolare, la Recording Industry Association of America - che comprende case discografiche come Universal Music Group (UMG), Sony Music Entertainment e Warner Records - ha ritenuto che molti dei brani prodotti dalle applicazioni **non siano altro che un plagio** di alcune basi già esistenti in quanto le applicazioni utilizzerebbero illecitamente, ai fini dell'addestramento dell'algoritmo, canzoni protette dal diritto d'autore.

Le etichette discografiche sostengono che le applicazioni creerebbero dei modelli pressoché identici a canzoni del calibro di: American Idiot dei Green Day, Johnny B. Goode di Chuck Berry o Rock Around the Clock dei Bill Haley & His Comets.



## Lyria 3: Google Brings Music Generation Inside Gemini App



Google officially integrated its **Lyria 3 model into Gemini** this week, allowing users to spin up 30-second tracks from a simple text prompt or a photo. While niche players like Suno and Udio have dominated AI music for months, Google is the first to put high-fidelity audio generation directly inside a mainstream assistant used by millions. The update handles everything from lyrics to cover art and even allows users to upload a video to serve as the creative mood board for a custom soundtrack.

### How to create music in Gemini:

- **Multimodal Inputs:** Users can feed Gemini text, photos, or videos to generate tracks with automatic lyrics, vocals, and genre-matched cover art.
- **Audio Watermarking:** Every track is embedded with an imperceptible **SynthID** watermark, and users can upload any file to check if it was generated by Google AI.
- **YouTube Integration:** Creators are gaining access to the model through **Dream Track** for Shorts, making it easier to build custom, copyright-safe backing tracks.
- **Safety Guardrails:** The model is trained to avoid mimicking specific artists, instead using them as broad style inspiration to comply with partner agreements.

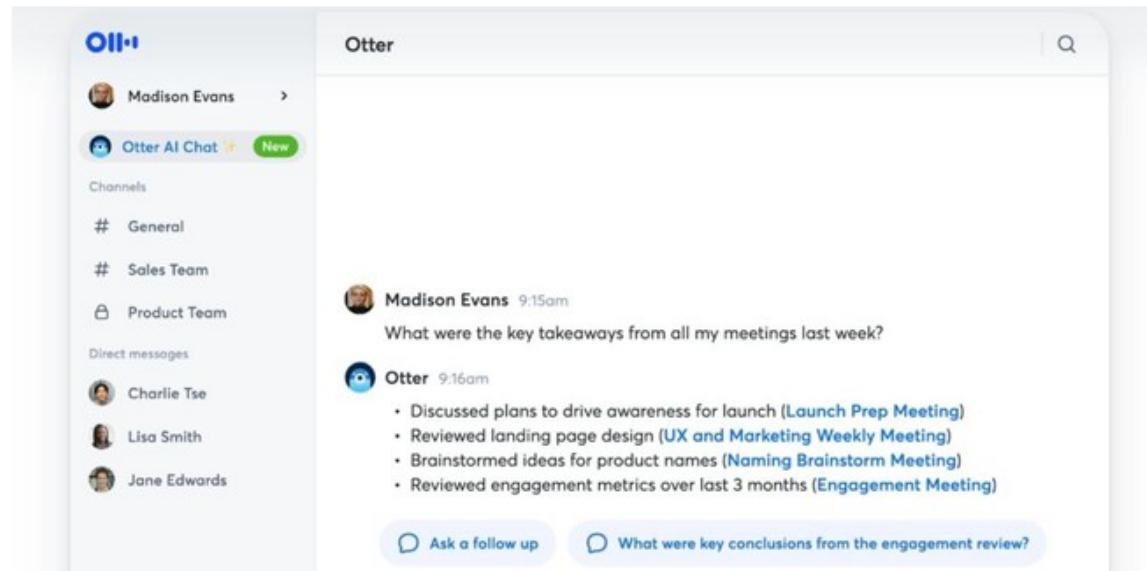
Platforms like Suno and Udio have existed for a year but embedding Lyria 3 into Gemini makes high-fidelity music generation a **default feature** for millions. Google is positioning this as a tool for "fun, unique expression" rather than a replacement for professional musicians. While the quality is high enough to fool casual listeners, the current 30-second limit suggests this is built for social sharing rather than replacing professional production.



**Otter.ai** è una piattaforma che si distingue per la sua accuratezza ed efficienza nella **trascrizione delle registrazioni vocali**. Importando un file audio, è possibile ottenere una trascrizione completa della conversazione in pochi minuti. L'assistente AI di Otter.ai è in grado di identificare gli interlocutori, inserire i marcatori temporali e produrre una trascrizione di elevata qualità.

settembre 2024

La piattaforma offre un'opzione gratuita, ma con alcune limitazioni: un massimo di 300 minuti di trascrizione al mese, una tariffa di 30 minuti per conversazione per tutte le conversazioni registrate dalla piattaforma stessa e solo tre importazioni a vita. Per coloro che necessitano di importazioni illimitate e di ricerche avanzate, Otter.ai propone un abbonamento annuale. Considerando il tempo risparmiato grazie all'utilizzo di questa piattaforma, l'investimento può rivelarsi vantaggioso per molti professionisti.



L'integrazione dell'**intelligenza artificiale** nei processi lavorativi è ormai una necessità per chi vuole rimanere competitivo e produttivo. Strumenti come ChatGPT, Canva Pro e Otter.ai offrono soluzioni pratiche e versatili che semplificano molte delle attività quotidiane, lasciando più tempo per dedicarsi alle attività centrali e strategiche. Adottare queste tecnologie oggi significa prepararsi al futuro del lavoro, dove automazione ed efficienza saranno sempre più centrali.



# DeepL presenta Voice, l'IA europea per la traduzione simultanea

*speech 2 speech*

ITALIAN TECH

di Andrea Nepori

*L'azienda di Colonia ha lanciato Voice, una piattaforma per la traduzione simultanea delle riunioni virtuali e delle conversazioni. Voice è già disponibile per i clienti commerciali e funziona anche in italiano*

14 NOVEMBRE 2024 ALLE 05:00

🕒 4 MINUTI DI LETTURA

"Vi chiederete come sia possibile che una multinazionale tecnologica non abbia sede nella Silicon Valley. In realtà, l'Europa è il luogo ideale per esplorare le sfide dell'IA applicata alle lingue: la traduzione fa parte delle nostre vite ogni giorno". A Berlino, il CEO di DeepL Jarek Kutylowski apre così, con un guizzo di orgoglio europeo, il keynote inaugurale della conferenza DeepL Dialogues.

Quando è nata nel 2017 a Colonia, DeepL era una piccola startup nel settore dell'intelligenza artificiale con l'ambiziosa missione di superare giganti come Google e Microsoft nel campo della traduzione automatica. Sette anni dopo, possiamo dire che la missione sia riuscita. DeepL è cresciuta fino a diventare una multinazionale tecnologica con una valutazione di circa due miliardi di euro, con uffici in Germania, Regno Unito, Stati Uniti e Giappone. Oggi, più di 100.000 utenti pagano ogni mese per utilizzare i servizi professionali della piattaforma, mentre decine di milioni di utenti accedono alla versione gratuita del traduttore.

## DeepL Voice

DeepL Dialogues, di cui si è tenuta la prima edizione a Berlino, è prima di tutto una customer conference, in cui i clienti hanno modo di familiarizzare con i prodotti di DeepL.

L'azienda ha però usato il palco dell'evento anche per presentare una novità molto importante: l'estensione della piattaforma al dominio della voce con l'introduzione di due nuovi servizi: DeepL Voice per le riunioni e DeepL Voice per le conversazioni.

"Il nostro obiettivo era garantire nella traduzione vocale simultanea la stessa qualità e sicurezza che contraddistinguono la piattaforma IA di DeepL", ha spiegato Kutylowski. "DeepL è già un punto di riferimento nella traduzione scritta, ma la traduzione vocale simultanea pone ben altre sfide. Quando si traduce il parlato in tempo reale, bisogna gestire input incompleti, problemi di pronuncia, latenza e altri fattori che possono causare traduzioni imprecise e un'esperienza utente mediocre. Si tratta degli stessi problemi che generano malintesi nelle conversazioni quotidiane, ed è per questo motivo che ne abbiamo tenuto conto sin dalle prime fasi di sviluppo. Con questa soluzione le aziende potranno comunicare in più lingue a seconda delle proprie esigenze, abbattendo le barriere linguistiche".



back to tassonomia

# DeepL presenta Voice, l'IA europea per la traduzione simultanea

ITALIAN TECH

di Andrea Nepori

*L'azienda di Colonia ha lanciato Voice, una piattaforma per la traduzione simultanea delle riunioni virtuali e delle conversazioni. Voice è già disponibile per i clienti commerciali e funziona anche in italiano*

14 NOVEMBRE 2024 ALLE 05:00

🕒 4 MINUTI DI LETTURA

## DeepL Voice Meetings

Il prodotto più enfatizzato da DeepL è il **traduttore simultaneo per le riunioni virtuali**. Al momento è disponibile all'interno di Microsoft Teams, ma l'azienda assicura che presto sarà compatibile anche con molte altre soluzioni per le videochiamate di lavoro. Il suo funzionamento è simile a quello di molti altri servizi analoghi già presenti sul mercato: sottotitola in tempo quasi reale, con una latenza di circa mezzo secondo, le conversazioni di tutti i partecipanti.

Il software riconosce tredici lingue diverse (inglese, italiano, tedesco, francese, spagnolo, coreano, giapponese, olandese, polacco, portoghese, russo, svedese e turco) e può mostrare sottotitoli in 33 idiomi differenti, ovvero tutti quelli disponibili nel traduttore testuale.

A giudicare dalle demo viste alla conferenza, il sistema funziona molto bene. È affidabile, e sembra pronto per essere utilizzato in ambienti di lavoro fin da subito.



back to tassonomia

# Un nuovo sistema di traduzione automatica

È DeepL Pro, dell'omonima startup, ed è disponibile in 165 paesi. Obiettivo: aiutare le aziende ad accelerare la loro espansione nel mercato globale

È in uscita il DeepL Pro, la soluzione di traduzione automatica per le aziende, che adesso sarà disponibile in 165 nuovi mercati, garantendo un **netto abbattimento dei costi di traduzione e localizzazione**. Un'espansione internazionale che rappresenta un importante traguardo per DeepL, azienda leader dell'AI linguistica, il cui obiettivo è da sempre quello di soddisfare la crescente domanda di soluzioni per il mondo delle imprese. Infatti, a oggi, sono moltissime le società e gli enti che si affidano alla piattaforma per entrare in nuovi mercati. E se la lingua certe volte può costituire una barriera, DeepL semplicemente l'abbatte, creando di fatto nuove opportunità.



## Video 2 video (motion transfer)

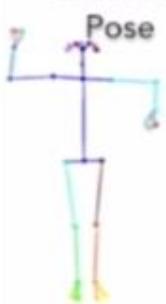
The model in use

A video from a professional dancer

Source Video



Detected Pose



It's just a detection step, and not a generative step

Source to Target Result



A video from a non-professional dancer dancing in a similar but not identical way (it's a bit behind of time, due to the translation)

## 2. do-as-I-do Motion Transfer<sup>1</sup>

- Video-to-video translation
- Pose estimation
- Discussion

1. [https://carolineec.github.io/everybody\\_dance\\_now/](https://carolineec.github.io/everybody_dance_now/)

<https://generativeai.net>



[Link al video](#)

[Link al paper](#)

# Motion transfer arricchito – luglio 2025



Alessio Pomaro [in](#) • Già segui

Head of Artificial Intelligence (AI) @ Search On Media Group, Docente, Speak...

[Visita il mio sito web](#)

19 minuti •

💡 LTX-Video è il primo modello di generazione video real-time basato su DiT, capace di produrre video HD a 30 FPS ( $1216 \times 704$ ) più velocemente della loro riproduzione.

👉 Perfetto per image-to-video, text-to-video, estensione video e animazioni keyframe, integra controlli avanzati su pose, profondità e bordi (canny).

⚙️ Con le versioni distillate e quantizzate, anche GPU consumer (es. RTX 4060) possono generare video in pochi secondi.

Supporta ComfyUI, HuggingFace Diffusers e può essere utilizzato come libreria Python o via CLI.

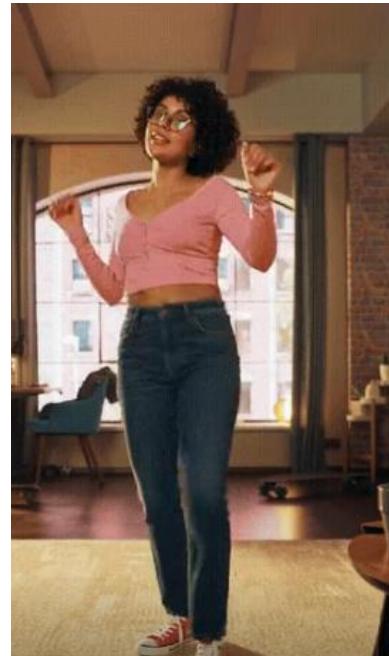
📦 Include modelli 2B e 13B, pipeline YAML, esempi pronti all'uso e strumenti per il fine-tuning (anche con LoRA).

Con TeaCache, ottieni inferenza fino a 2x più veloce senza sacrificare la qualità.

🔍 Ideale per creativi, sviluppatori e ricercatori che vogliono spingere i confini della generazione video AI.

Link al post  
[Linkedin](#)

Link al sito  
[github](#)





Generative AI  
4.945.706 follower  
[Contattaci](#)  
1 ora •

ottobre 2024

Video 2 Video

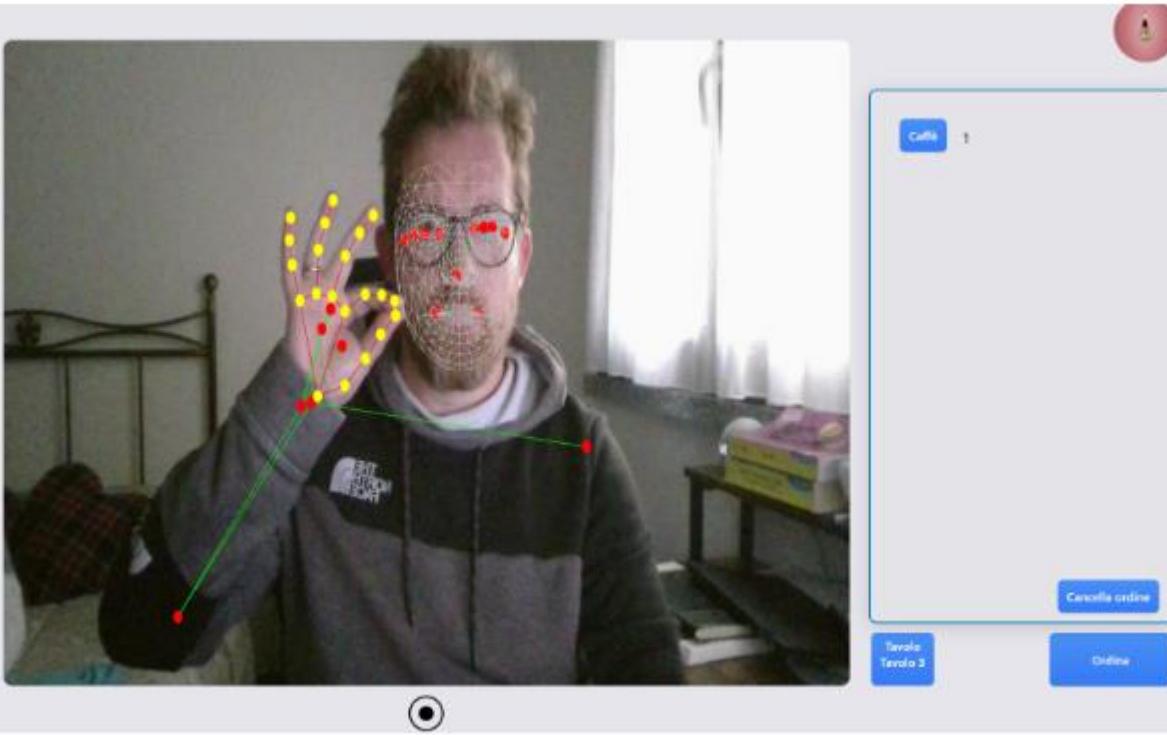
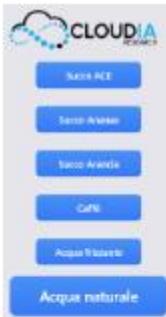
The Future of Reality is Here!

Imagine transforming your surroundings into a 1950s vibe or a fantasy world with just a click. Thanks to **#AR** and **#AI**, we're closer than ever to living in fully immersive environments!

This video shows what's possible with video-to-video transformation. And with the rapid speed of AI advancements, the line between virtual and real worlds is getting thinner every day.



[back to tassonomia](#)



Link al sito  
del PoliMI

*Video 2 Text*



back to tassonomia





Generative AI  
4.925.626 follower  
[Contattaci](#)  
8 ore •

ottobre 2024

AI Can Now Understand Human Emotions Better Than Ever 😊

Thanks to groundbreaking image-to-video technology, AI's ability to truly understand and interpret human emotions has made a massive leap forward!

This is a game-changer for how we interact with tech, making it more intuitive and connected to our feelings.

## Image 2 Video (*alta espressività*)



[back to tassonomia](#)

link a video  
TikTok di  
commento

video di Synthesia  
sugli avatar  
espressivi



JOY



Massimo Canducci • 1°

Innovation @ Engineering Group | Expert on "Innovation and Future" @ Sin...

2m • Modificato •

## The Future in a Photo, Thanks to AI

Thanks to generative AI algorithms, it's possible to predict a person's future physical appearance starting from a current image.

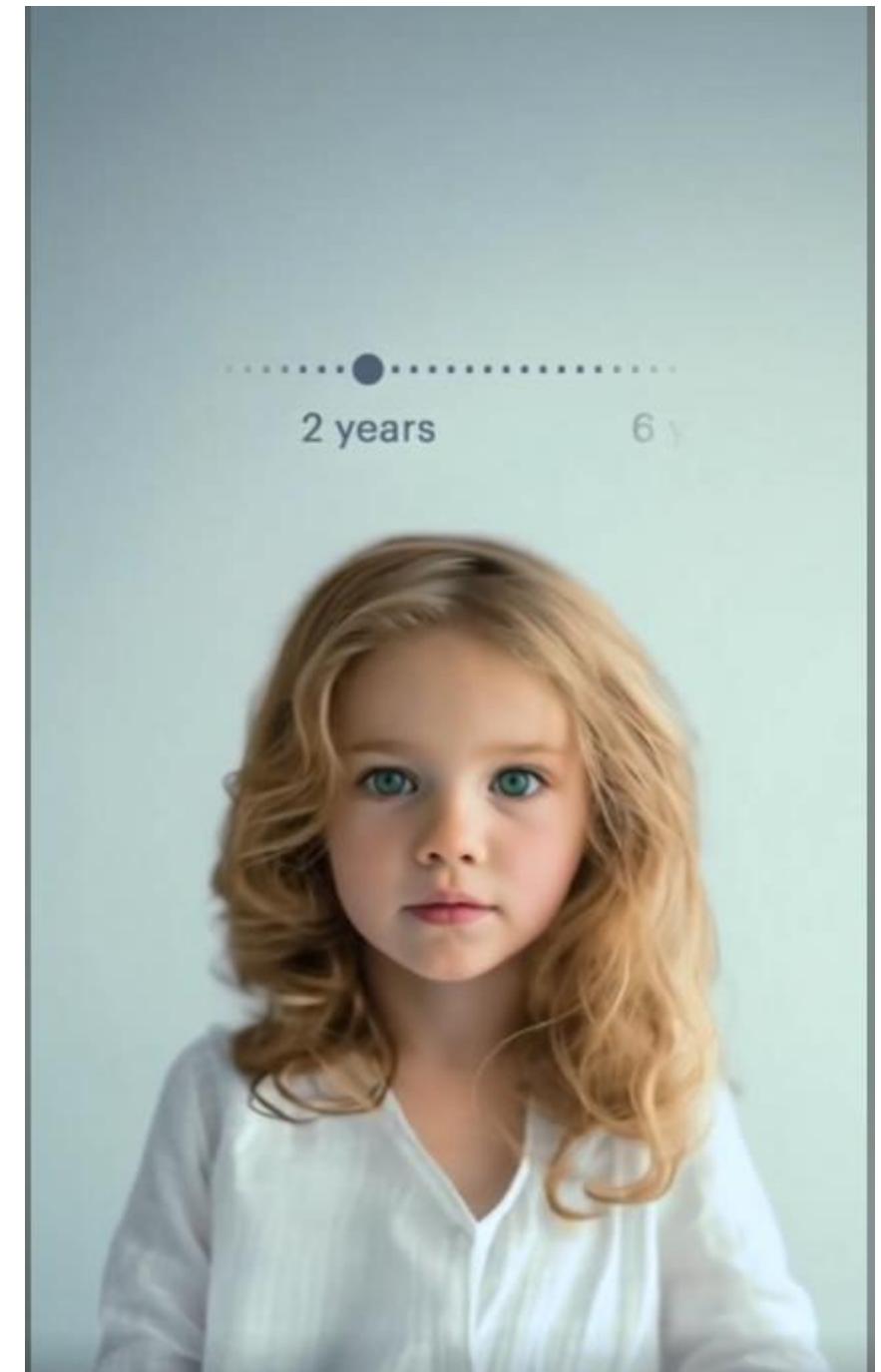
Of course, it would be necessary to test these algorithms on images of children from past years to understand if there is any reliability or if, as is likely, it's just a pure graphical speculation.

In any case, as technology and, more importantly, the available data progress, it will become increasingly possible to know our (and others') future physical appearance in advance.

Looking even further ahead, it might be possible to detect potential health issues from images. This is something already done with medical images, but it would be very interesting to identify disease markers simply in holiday photos.

settembre 2024

## Image 2 Video (*aging-deaging*)



[back to tassonomia](#)

*Image 2 Video*

Il ballo della vittoria  
alle elezioni

# village people

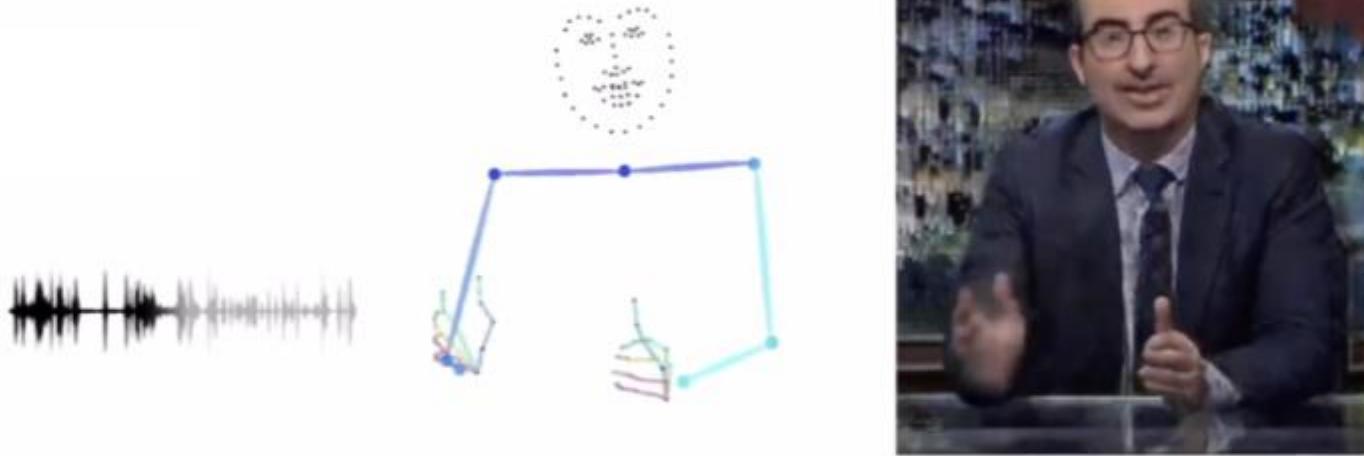


back to tassonomia

We **predict** arm and hand **gestures from speech**.  
Given only the **audio** of a speaker as **input**,  
we **predict** their **gestures**.

Use case: **predict gestures**<sup>1</sup>

### Speech/Image 2 Video



Audio input

Predicted gestures  
Face is ground truth

Synthesized video  
using Chan et al., 2018

1. <https://people.eecs.berkeley.edu/~shiry/projects/speech2gesture/index.html>

<https://generativeai.net>

[Link al video](#)



Big Data, Analytics, Business Intelligence & Visualization 18 ottobre 2024

Ultan O. • Già segui

[Visita il mio sito web](#)

8 minuti •

Adobe just announced Project Perfect Blend (AI Photo Harmonization)

This feature uses AI to seamlessly merge images, creating smooth, natural transitions between them. It's designed to eliminate that awkward 'photoshopped' look, to get stunning, realistic results effortlessly.



**Image 2 Image  
(image composition)**

BreakingNews

**Adobe** just made some big announcements at Adobe MAX 2024!

They've launched their first AI video tool in Premiere Pro, beating competitors like [OpenAI](#) and [Google](#)

Photoshop, Illustrator, and InDesign are getting new features, and [Frame.io](#) has received its biggest update since 2015.

One cool new feature, 'Project Know How', can track and prove ownership of an artist's work, whether it's online or printed on something like a tote bag or poster, making sure creators get the credit they deserve.



**Image & Audio 2 Video**



back to tassonomia



back to tassonomia

Tg1

Tg1 ore 20:00 del 27/10/2023



Video originale  
fornitomi da RAI



back to tassonomia

# PDF a nudo per Claude 3.5 Sonnet: ora può analizzare anche le immagini al loro interno

di Sergio Donato - 04/11/2024 12:41



A screenshot of a conversational interface. On the left, there is a circular flight plan diagram with various points labeled. In the center, a message bubble contains the text: "Hey Claude, can you please run me through the flight plan?". Below this, another message bubble says: "Scanning document...".

Di fronte ai PDF, di solito i modelli IA si fermano all'analisi del testo. Claude ora può anche "leggere" le immagini in essi contenute. Funzione molto importante specie per i documenti scientifici o con grafici

[link  
all'articolo](#)

Anthropic ha dato un'abilità molto importante al suo modello linguistico di grandi dimensioni Claude 3.5 Sonnet: quella di **leggere tutti i tipi di immagini all'interno di un PDF di meno di 100 pagine**.

Si tratta comunque di una funzione in anteprima attivabile dal pannello Feature Previews (l'icona del matraccio usato nei laboratori di chimica) chiamata Visual PDF.

I modelli multimodali più capaci hanno già la possibilità di analizzare immagini e descriverle, ma solitamente si tratta o di file immagini o di PDF di sole immagini. In generale, infatti, da un PDF viene estratto solo il testo anche se al suo interno sono presenti figure esplicative, come nei documenti scientifici dove la semplice didascalia può non essere un riassunto sufficientemente preciso della grafica a cui si riferisce.

ChatGPT e Gemini, per esempio, non riescono ancora a esaminare le immagini dentro i PDF. Abbiamo provato con [questo articolo scientifico](#) chiedendo di descrivere il grafico della figura 4 senza leggere la didascalia. ChatGPT ha fallito ma è stato onesto ipotizzando genericamente cosa poteva contenere sulla base del testo che la riguarda; mentre Gemini 1.5 Pro ha restituito dei valori come se li leggesse nella figura, ma poi ci siamo accorti che ha fatto un calcolo sfruttando il testo del documento a sua disposizione e facendo credere di avere interpretato l'immagine.

Questa capacità di intrepretare le immagini nei PDF ora pare ce l'abbia solo Claude 3.5 Sonnet che inoltre adesso può **gestire documenti con dimensioni complessive di 30 MB, rispetto al limite precedente di 10 MB**.



back to tassonomia



guida-bias-cognitivi-raffaele-gaito.pdf

PDF

qual è il rapporto percentuale tra istinto e razionalità, citato nel PDF allegato?

Nel PDF "Guida ai bias cognitivi" di Raffaele Gaito, viene citato un dato interessante sul rapporto tra **istinto e razionalità** nel nostro cervello:

- ◆ "Il sistema 2 non è, in ogni caso, quello predominante ma corrisponde solo al 5% dell'attività cerebrale."

Quindi, secondo quanto riportato:

- Sistema 1 (istintivo, veloce, intuitivo) → 95%
- Sistema 2 (razionale, analitico, lento) → 5%

📌 Questo concetto si basa sulle teorie di **Daniel Kahneman**, che distingue due modalità di pensiero:

- **Sistema 1:** reazioni automatiche, basate su intuizioni, abitudini e scorciatoie mentali.
- **Sistema 2:** pensiero logico, calcolo, riflessione cosciente.

👉 Quindi, nella vita quotidiana, la stragrande maggioranza delle decisioni che prendiamo è guidata da **automatismi e intuizioni**, non da ragionamenti logici e ponderati.

Se vuoi posso aiutarti a creare un'infografica o una slide che visualizzi bene questo rapporto, utile anche per corsi o presentazioni. Fatti sapere!



Statistical\_flaws\_in\_Excel.pdf  
PDF

dimmi cosa riporta la Figure 2

La **Figure 2** nel documento "Statistical flaws in Excel" mostra un grafico con una trendline applicata ai dati del cosiddetto **J. Simonoff example**, un caso in cui Excel fallisce clamorosamente nell'adattare una retta di regressione.

### In breve, cosa mostra la Figure 2:

- Un tentativo di Excel di inserire una retta di regressione su un grafico XY scatter.
- Il risultato è totalmente errato: la retta non rappresenta affatto la distribuzione dei dati.
- Excel calcola persino un valore di **R<sup>2</sup> negativo**, che è assurdo, visto che il coefficiente di determinazione dovrebbe essere  $\geq 0$  in regressione lineare standard.

👉 Il problema segnalato è che Excel usa algoritmi numericamente instabili per la regressione, in particolare quando lavora con numeri molto grandi e con variazioni minime nei dati (che è proprio il caso di questo esempio). Il difetto è amplificato quando si usa la funzione "Add Trendline".

👉 L'autore sottolinea che nessuna delle funzioni di trendline di Excel (eccetto la media mobile) riesce a restituire un risultato accettabile per questo dataset.

Se ti serve, posso anche ricreare quel grafico in un ambiente Python (tipo matplotlib), per farti vedere la differenza tra la regressione corretta e quella sbagliata di Excel. Vuoi?



back to tassonomia



The Blondoner

@TheBlondoner · 2890 iscritti · 93 video

Un punto di vista diverso sul mondo ...altro

instagram.com/iamtheblondoner

Iscriviti

Abbonati



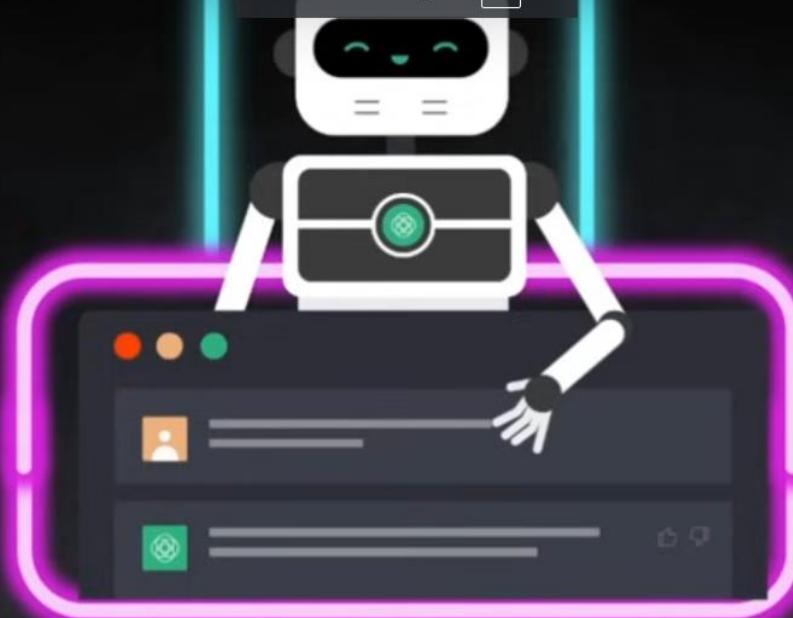
Intelligenza Artificiale: IMBUCATA TRA GLI AI TRAINERS

To exit full screen, press Esc



RISPOSTA  
1

“  
Roma è più grande  
di Firenze in termini  
di popolazione  
”



PROMPT

“  
Quale città è più grande?  
Roma o Firenze?  
”

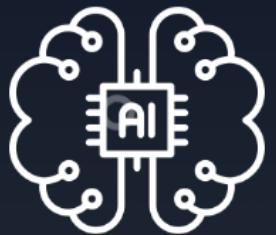


RISPOSTA  
2

“  
Secondo Wikipedia,  
Firenze è più grande  
di Roma  
”



back to tassonomia



# Grok Image Generator

Welcome to Grok Image Generator, where creativity meets visualization! Turn your text into an art using Grok AI.

Grok per immagini è basato sul modello T2I [Flux 1](#), realizzato da un gruppo di progettisti uscito da Stable Diffusion.



back to tassonomia



Generative AI  
4.303.880 follower  
1 ora •

luglio 2024

## 🚀 Introducing Meta 3D Gen – a breakthrough in text-to-3D generation by [Meta's AI researchers](#)

This new research delivers high-quality 3D assets with advanced geometry and textures.

📄 Research paper: <https://lnkd.in/dp9mBHJJ>

Meta 3D Gen enables text-to-mesh generation with top-tier geometry, texture, and PBR materials. It produces high-resolution textures and material maps end-to-end, offering superior results at 3-10x the speed of previous solutions.

📄 In addition, Meta is publishing research on two components of the Meta 3D Gen system:

Meta 3D AssetGen: Generates 3D models from text.

Meta 3D TextureGen: Provides high-quality texture generation and AI-assisted retexturing.

📄 Meta 3D AssetGen paper: [https://lnkd.in/dchB\\_CKD](https://lnkd.in/dchB_CKD)  
📄 Meta 3D TextureGen paper: <https://lnkd.in/dAHzH6-E>



back to tassonomia



**Steve Nouri** • Già segui

Generative AI Founder | Advisor @ Fortune 500 | 1.6 Millio...

[Richiedi servizi](#)

2 giorni •

This is truly wild. 🔥

Luma AI has just released Dream Machine, a groundbreaking tool that generates AI videos from text and images, now open to the public!

[link al post](#)

Unlike Sora, Dream Machine is available to everyone starting today. This marks a significant leap forward for image-to-video models. Personally, I find image-to-video models preferable over text-to-video as they offer greater control over aesthetics.

For those looking for alternatives while waiting for Sora, Runway Gen-2 is another text-to-video generative AI currently available via web and mobile platforms. This tool also allows users to create videos from text input, offering a viable option for content creators in the interim ([dragganaitool.com](http://dragganaitool.com)).

Da chatGPT



back to tassonomia

# Così cambia la posta elettronica per aiutarci a risparmiare tempo

La promessa, almeno in parte mantenuta, dei nuovi strumenti integrati con le mail è aiutarci a restare aggiornati sulla nostra agenda e suggerirci come rispondere

di Alessandro Longo

26 maggio 2024

29 maggio 2024

24 ORE

Tecnologia

1/4

Uno degli strumenti più noti è [Copilot Pro di Microsoft](#), che integra l'AI (basata su Gpt di OpenAI) nelle [app di Office](#), compresa Outlook, dove però funziona solo su account sotto il [dominio Microsoft](#). Possiamo riassumere i contenuti delle mail, generare una bozza (dando alcune indicazioni su quello che vogliamo dire) oppure ricevere suggerimenti (coaching) per migliorare il nostro testo. Forse è questa la funzione che, nella nostra prova, si dimostra ora più interessante.

Copilot Pro costa 22 euro a utente, per mese, e mette Gpt-4 in molte app per generare testi, immagini.

Se abbiamo [Gmail](#), invece, la scelta più ovvia è [Shortwave](#), che ha una vasta gamma di funzioni AI e un prezzo contenuto (da 7 dollari al mese); peccato che non funzioni con altre mail.

È uno dei pochi ad avere la funzione ricerca AI delle e-mail, che si fa in linguaggio naturale come su ChatGpt, «fammi un elenco di tutti i compiti a me assegnati questo mese dal mio supervisore».

Possiamo fare domande sui contenuti delle mail, inoltre. E, sempre in linguaggio naturale, programmare riunioni o nuovi eventi. Anche Gemini di Google funziona (solo) con Gmail, ma è ancora agli inizi, quindi diamogli qualche altro mese per maturare.

 shortwave

clicca sull'immagine



back to tassonomia



### Done

Mark emails as done  
when there are no more  
action items



### Pin

Pin emails that need  
follow-up or contain  
to-dos



### Snooze

Snooze emails you'll  
need later



### Delete

Delete the emails you  
don't need anymore

You're probably familiar with snooze and delete from other email clients, but done and pin are unique to Shortwave.

Pin is your means of separating emails with to-dos from the rest of your inbox so that they never get lost in the noise. Even as new emails come in, your pinned section will stay tidy.



## Shortwave:

- categorizza le email per tipologia
- permette di fare triage delle email: *pin, snooze, all done*
- assistente AI per: sintesi thread, supporto risposta, traduzione, ecc

link al video  
di tutorial di  
shortwave

Link alla pagina introduttiva  
sull'assistente AI

## How the AI Assistant can help

The AI Assistant is more than just a tool; it's your partner in managing your email workflow efficiently. It's like having a personal assistant who knows your preferences inside out.

### Here's what it can do for you:

- Search your email history for answers
- Write drafts in your unique voice with Ghostwriter
- Translate threads & analyze data
- Instantly summarize every thread in your inbox
- Schedule and create calendar events
- Research general knowledge questions
- Suggest completions to your drafts while you type using AI Autocomplete

## Access the AI Assistant



oppure *ctrl+j*

**Personalized writing style**  
Upgrade to personalize the AI's writing style  
to match your own  
[Upgrade](#)



back to tassonomia

UPGRADE SUCCESSFUL

## Enjoy our premium features

### Full email history

Import all of your emails so you can search and view every message

### AI search

Unlock full AI Assistant capabilities with a customized search experience

### AI writing style

Let Shortwave AI craft replies in your style, drawing from your past sent emails

[Activate premium features](#)

back to tassonomia



Steve Nouri • Già segui

Generative AI Founder | Advisor @ Fortune 500 | 1.6 ...

[Visita il mio sito web](#)

20 ore •

**Google** has launched Med-Gemini, an advanced AI fine-tuned for medical Tasks.

It significantly outperforms earlier models, including GPT-4, on most medical benchmarks.

Achieves top scores, particularly on the MedQA-USMLE benchmark with a groundbreaking 91.1% accuracy.

Demonstrates superior performance over GPT-4 by 44.5% on average across seven multimodal benchmarks.

Excels in tasks such as medical summarization, generating doctor referrals, and simplifying medical documents.

It is a preferred method over human expert analyses for complex text-based medical tasks.

This marks a significant advancement in AI for healthcare, suggesting potential improvements in medical diagnostics and patient care.

Positions healthcare as a leading field in AI development, promising substantial impacts on clinical practices and education.

2 maggio 2024



[back to tassonomia](#)

## Conversation with Med-Gemin

For the past few weeks, I have noticed **multiple itchy lumps** on my legs and arms. The lumps are solid and don't contain fluid. They are very itchy and I have been waking up at night scratching them. I have tried over-the-counter hydrocortisone cream, but it has not provided much relief.



I understand your concern. Can you **send me a picture** of the lumps?



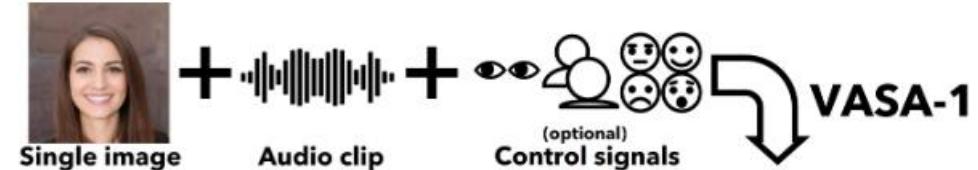
Have you recently **been exposed** to any new plants, animals, or medications?



# VASA-1: Lifelike Audio-Driven Talking Faces Generated in Real Time

link alla url  
(con demo video)

TL;DR: single portrait photo + speech audio = hyper-realistic talking face video with *precise lip-audio sync, lifelike facial behavior, and naturalistic head movements*, generated in *real time*.



## Abstract

We introduce VASA, a framework for generating lifelike talking faces of virtual characters with appealing visual affective skills (VAS), given a single static image and a speech audio clip. Our premiere model, VASA-1, is capable of not only producing lip movements that are exquisitely synchronized with the audio, but also capturing a large spectrum of facial nuances and natural head motions that contribute to the perception of authenticity and liveliness. The core innovations include a holistic facial dynamics and head movement generation model that works in a face latent space, and the development of such an expressive and disentangled face latent space using videos. Through extensive experiments including evaluation on a set of new metrics, we show that our method significantly outperforms previous methods along various dimensions comprehensively. Our method not only delivers high video quality with realistic facial and head dynamics but also supports the online generation of 512x512 videos at up to 40 FPS with negligible starting latency. It paves the way for real-time engagements with lifelike avatars that emulate human conversational behaviors.



back to tassonomia



A.I. images generated by Sina Alemohammad and others.

**After four generations, the faces all appeared to converge.**

## Risk of Model Collapse



Shea Brown [in](#) • 3° e oltre

AI & Algorithm Auditing | Founder & CEO, BABL AI Inc. | ForHumanity...

1 giorno •

agosto 2024

Fantastic [The New York Times](#) article covering the problem of Generative AI feeding on its own outputs. Short version; it leads to the collapse of the model into bad or uniform (and biased) outputs.

The data used to illustrate this is quite nice (but on the extreme end) and easy to understand for a non-technical person.

Slightly longer version:

AI is creating a lot of content on the internet, like articles, reviews, and social media posts.

It's hard to tell what content is made by AI, which can lead to confusion and misinformation.

If AI systems use their own content to learn, their quality can decrease over time, making their outputs repetitive and less accurate.

This problem, called "model collapse," means AI outputs can become less diverse and more distorted.

It can also lead to a loss of diversity, amplify existing biases in the data, and make it more likely for AI to overlook or erase information about minorities.

To avoid these issues, companies need to use more real, high-quality data and find better ways to spot AI-generated content.



back to tassonomia



link all'articolo

## Further Stable Diffusion Pipeline with Diffusers

by Abid Ali Awan on June 11, 2024 in Stable Diffusion

2

[Share](#)  [Tweet](#)  [Share](#)

There are many ways you can access Stable Diffusion models and generate high-quality images. One popular method is using the Diffusers Python library. It provides a simple interface to Stable Diffusion, making it easy to leverage these powerful AI image generation models. The diffusers lowers the barrier to using cutting-edge generative AI, enabling rapid experimentation and development. This library is very powerful. Not only you can use it to generate pictures from text prompts, but also to leverage LoRA and ControlNet to create a better picture.

In this post, you will learn about Hugging Face's Diffusers, how to generate images, and how to apply various image generation techniques similar to Stable Diffusion WebUI. Specifically, you will learn how to:

- Build a Diffusers Pipeline and generate a simple image with a prompt.
- Loading LoRA weights of fine-tuned models and generating IKEA-style images.
- Build ControlNet OpenPose pipeline to generate an image using a reference image.

Let's get started.





Business Intelligence, Analytics & Data Visualization (...)

Ultan O' • Già segui

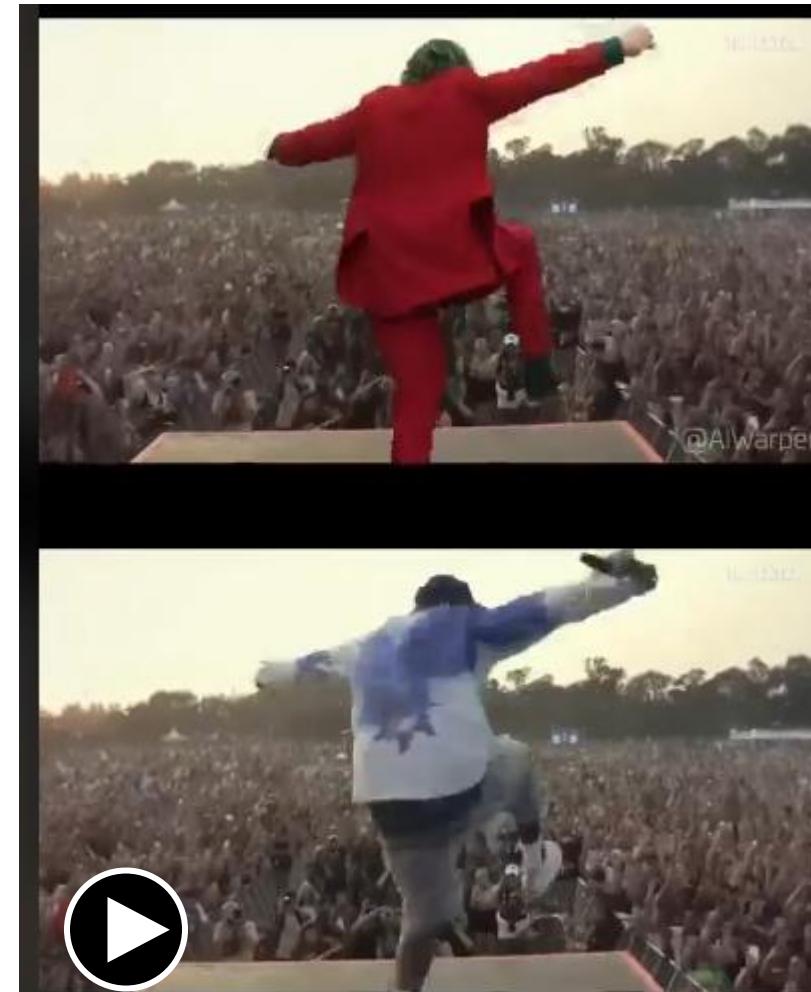
[Visita il mio sito web](#)

7 ore •

AnimateDiff, a diffusion model to replace anyone from a video💡

💬 AnimateDiff Link: <https://lnkd.in/g-He6g5q>

link al sito





Generative AI  
5.136.823 follower  
[Contattaci](#)  
5 ore •

novembre 2024

## The Matrix: Alibaba's New AI Model Creates Endless Virtual Worlds at 16 FPS

Researchers from Alibaba, University of Hong Kong, and University of Waterloo introduced a game-changing AI model called "The Matrix".

Until now, creating high-quality simulated environments has been incredibly expensive and time-consuming. Game developers spend countless hours manually crafting environments, and simulation creators face similar challenges.

Key highlights:

- >Generates infinite-length 720p videos in real-time
- >Offers frame-by-frame control
- >Works in both gaming and real-world environments
- >Runs at up to 16 frames per second

The numbers are impressive:

- >Move-PSNR: ~28.98 (that's tech-speak for "really good quality")
- >Speed: 8-16 FPS in real-time
- >Can handle tasks like simulating a BMW X3 driving through an office (something it never saw in training!)

The applications are massive:

- From video game development to training autonomous vehicles, and even creating virtual reality experiences.

link al  
paper



back to tassonomia

[link  
all'articolo](#)**Speech 2 Speech**

1/2

# Bastano 15 secondi per clonare la tua voce: come funziona il nuovo programma di OpenAI

*Non solo per scopi educativi o per parlare in altre lingue, la nuova intelligenza artificiale sorella di ChatGPT permetterà anche a chi soffre di malattie oncologiche o degenerative di recuperare l'uso della voce. Ma non mancano i rischi.*

A cura di **Velia Alvich**

In un futuro non troppo lontano potremo usare la nostra voce per parlare lingue che non abbiamo mai imparato. Non utopia, è una realtà firmata da **OpenAI** e si chiama **Voice Engine**. L'unico problema è che il progetto per adesso non è aperto al pubblico ma solo a un ristretto gruppo di associazioni e centri di ricerca.

Il funzionamento è semplice: bastano appena quindici secondi di voce registrata e l'intelligenza artificiale fa tutto il resto. Una voce che può anche essere stata registrata per altri scopi, come è capitato a una giovane paziente che non parla più come una volta e che ha fornito al software un campione audio preso da un vecchio video.



Per adesso si tratta solo di una presentazione. Ancora non si sa, infatti, quando Voice Engine verrà rilasciato a un pubblico più ampio. A tenere lontane le mani (e le voci) degli utenti di tutto il mondo ci sono forse le **preoccupazioni per gli usi impropri** che possono essere fatti di questa tecnologia. Intanto, però, la stessa OpenAI ha lanciato alcune idee su come potrebbe essere utilizzato (bene) in futuro.

## Quando verrà rilasciato

OpenAI non è nuova all'idea di combinare l'intelligenza artificiale con la voce umana. Dal lancio di **ChatGPT**, il prodotto di punta dell'azienda, sono state presentate diverse possibilità per trasformare il testo in un pezzo parlato (compresa la funzionalità per leggere quello che l'IA generativa scrive). Lo stesso progetto di Voice Engine è nato alla fine del 2022 e in meno di due anni ha raccolto adesioni di associazioni e centri di ricerca per mettere alla prova lo strumento.



Banias Baabe • 3° e oltre  
I make learning Data Science effortless for you | Data ...  
7 ore •

11 marzo 2024

Do you want high-quality text-to-speech in Python?

Use **MeloTTS**.

**MeloTTS** supports various languages for speech generation without needing a GPU.

You can use it via CLI, Python API, or Web UI.

--  
 Get the Free Data Science Online Book (200+ tricks) by subscribing to my weekly newsletter today: <https://lnkd.in/dS22kf9Q>

[Vedi traduzione](#)

```
melotts.py
#!/git clone https://github.com/myshell-ai/MeloTTS.git
#!cd MeloTTS
#!pip install -e .
#!python -m unidic download
from melo.api import TTS

text = "Le chat noir se promène silencieusement dans le jardin nocturne."
model = TTS(language='FR', device='cpu')
speaker_ids = model.hps.data.spk2id

output_path = 'fr.wav'
model.tts_to_file(text, speaker_ids['FR'], output_path, speed=1.0)
```

https://www.linkedin.com/in/banias/



back to tassonomia



Generative AI  
3.065.507 follower  
19 ore •

29 febbraio 2024

## 🌟 Exciting News: **Gamma**, a leading innovator in AI-powered presentation creation, has now launched a groundbreaking no-code AI website builder

Gamma's pioneering presentation tool has rapidly become one of the most swiftly expanding AI web applications, attracting **10 million users** within a mere 9 months. Building on this success, they have now extended their platform to include **AI website creation** capabilities.

Current website builders on the market are bulky, over-templated, and fragile.

So Gamma built something better.

Introducing: Gamma sites. Now, anyone can spin up a website and maintain it completely on their own.

**No design or coding skills required.**

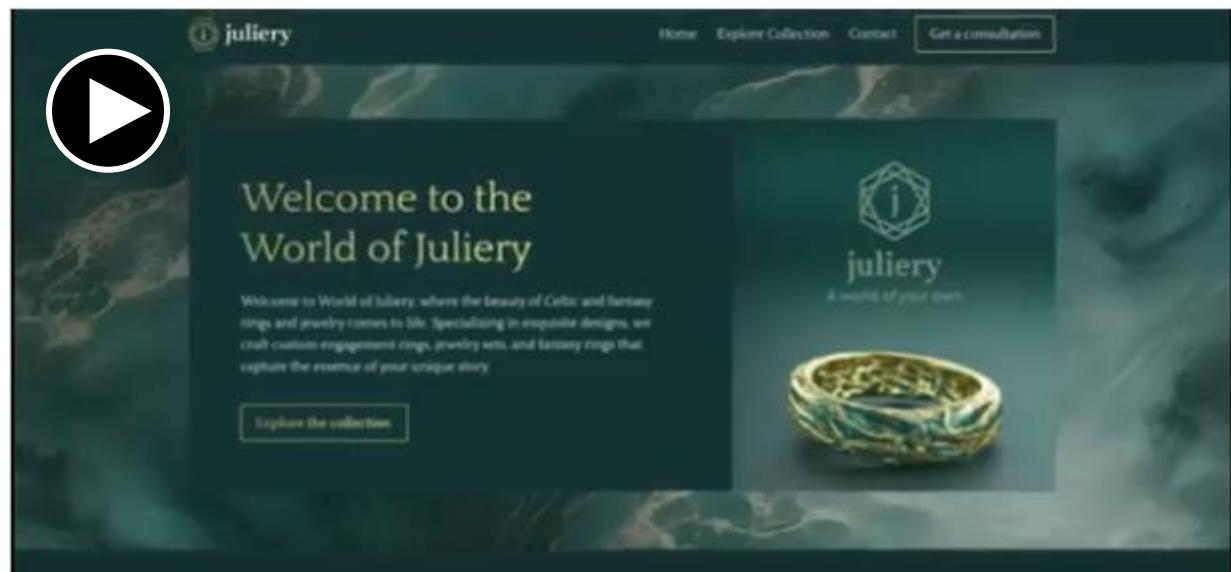
Features:

- Mobile friendly
- SEO optimized
- Easy to edit

**For Everyone:** Ideal for launching portfolios, side hustles, projects, courses, and more.

**Simple and Powerful:** Building and publishing a website has never been easier.

Try it today at <https://gamma.app/>



back to tassonomia

# Sora, il nuovo modello di OpenAI per generare video realistici da prompt di testo

16 Febbraio 2024 

Link all'articolo di HDBlog.it, che contiene alcuni esempi di video generati con Sora



Business Intelligence, Analytics &amp; Data Visualiza

Ultan O. • Già segui  
1 ora • 

15 febbraio 2024

⚠️ OpenAI has just launched Sora, their video generation model

Here's one example of the results 

Runway, Pika and other tools currently generate a few seconds at a time. Sora can generate up to 60 seconds.



Prompt: A stylish woman walks down a Tokyo street filled with warm glowing neon and animated city signage. She wears a black leather jacket, a long red dress, and black boots, and carries a black purse. She wears sunglasses and red lipstick. She walks confidently and casually. The street is damp and reflective, creating a mirror effect of the colorful lights. Many pedestrians walk about.



back to tassonomia

Molto più avanzato rispetto a Lumiere di Google?!

Non ancora disponibile perché OpenAI sta implementando misure di sicurezza, ad esempio contro i deepfake. Data solo a red teamers ed artisti



Ultan O. (Sharing Education Tools) • Già segui

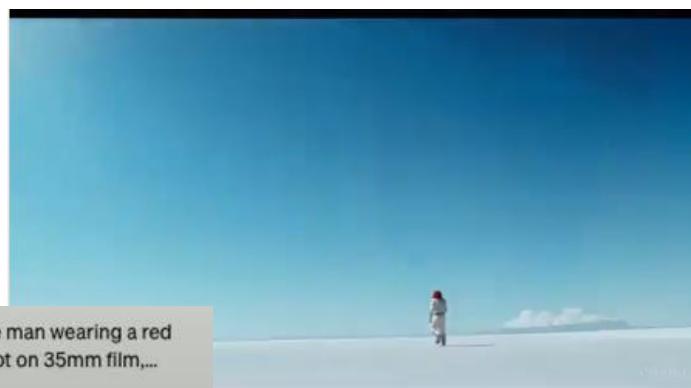
 Follow for Daily Innovations, Resources & Tools → ...  
1 ora • 

16 febbraio 2024

⚠️ OpenAI's New Text to Video tool, Sora

This is one example of the results from Sora, OpenAI's video generation model 

Prompt: "A movie trailer featuring the adventures of the 30 year old space man wearing a red wool knitted motorcycle helmet, blue sky, salt desert, cinematic style, shot on 35mm film, vivid colors."



Prompt: A movie trailer featuring the adventures of the 30 year old space man wearing a red wool knitted motorcycle helmet, blue sky, salt desert, cinematic style, shot on 35mm film,...  
more

why doesn't my sora-created video produce audio track (just video)?

My media · Turin's Transformation Story

480p Apr 17, 10:26AM



Image prompt genera un video nel quale questa persona parla e dice la seguente frase: "vivo a Torino, mi piace molto la mia città ..."

Tutti

Video

Video brevi

Immagini

Notizie

Libri

Web

Altro



Search Labs | AI Overview

Sora, OpenAI's text-to-video tool, currently doesn't include audio. However, adding sound to Sora videos is a planned future feature. [🔗](#)

### Explanation [🔗](#)

- Sora is currently only able to generate videos from text input.
- Adding sound and syncing it to the video would give creators more control over the final product.
- Some say that Sora could incorporate models that target video-to-audio and audio-to-video.

### Workarounds

- You can upload your Sora video to YouTube and add closed captions and a soundtrack. [🔗](#)
- You can use other AI tools, such as lipsync technology, to automate and accelerate various stages of the video production workflow. [🔗](#)

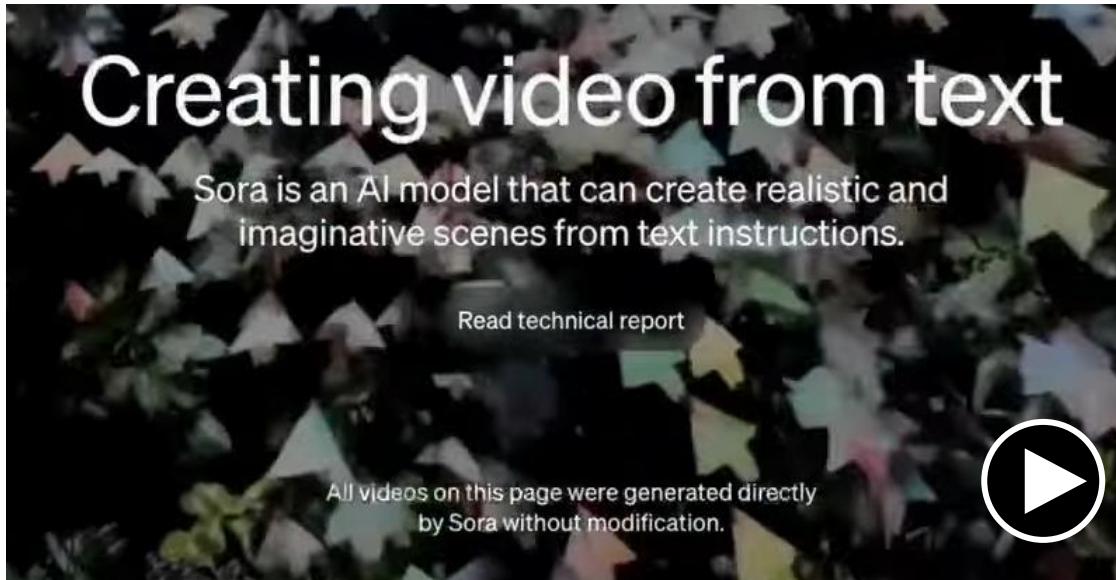
### About Sora

Sora is a tool that uses advanced AI to create realistic and smooth videos. It can capture complex scenes and actions, and ensure consistency in videos. Sora is available to ChatGPT Plus, Team, and Pro users. [🔗](#)



back to tassonomia

L'annuncio di Sora di OpenAI (con gli esempi)



L'annuncio di Sora di OpenAI alla CNN



Sora di OpenAI è un T2V.

Genera video di lunghezza sino ad 1' (Pica e Runway solo di alcuni secondi).

Alta qualità e definizione del primo piano.

Aderenza al prompt.



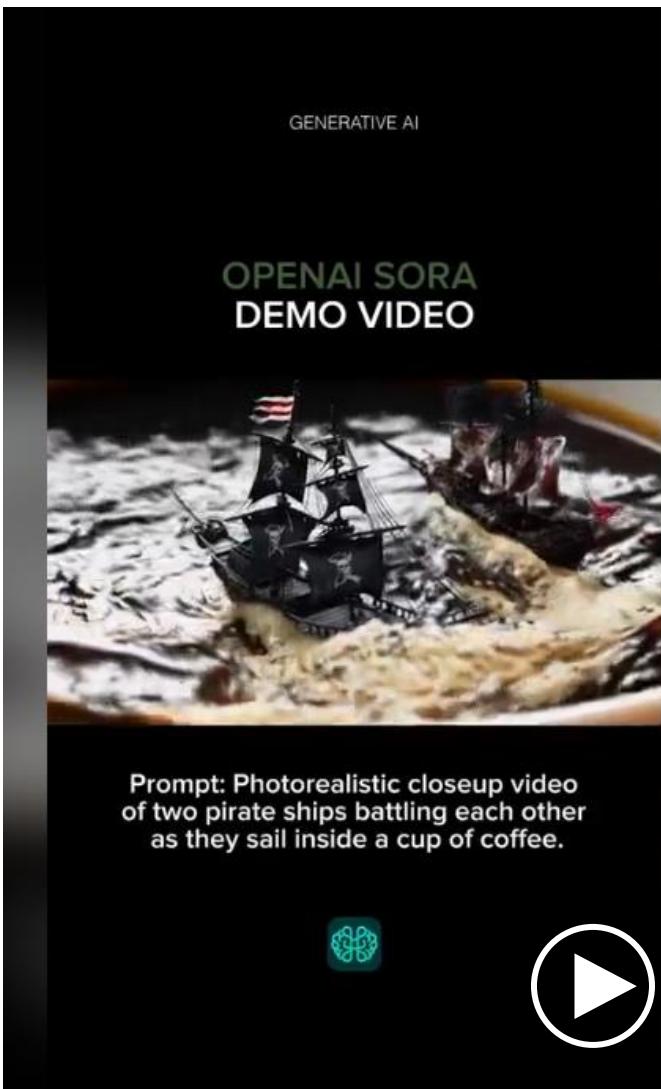


Generative AI

2.920.946 follower

6 ore •

Sora detailed video of 2 battling ships in the cup of coffee ☕



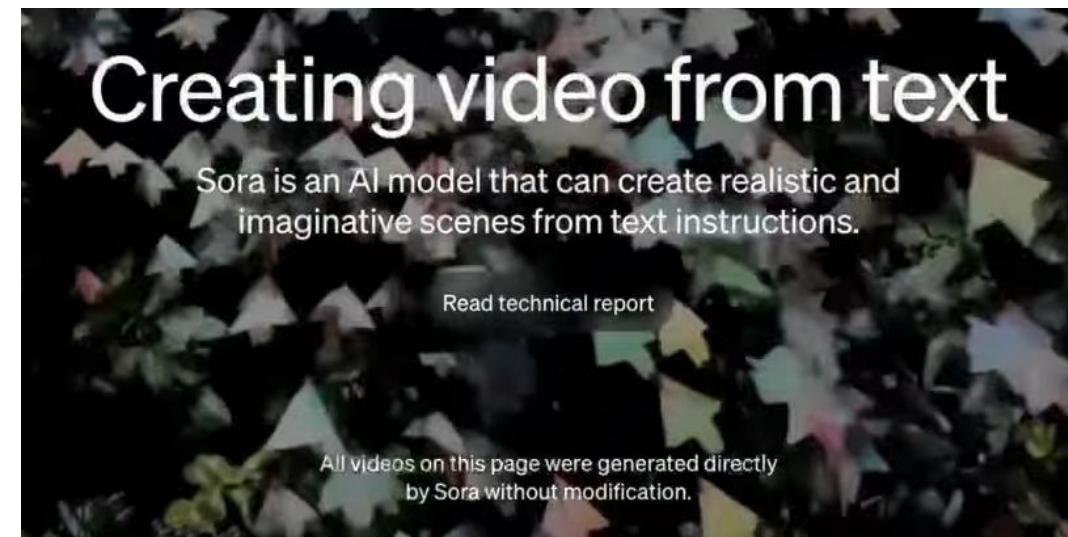
# OpenAI annuncia SORA per video PAZZESCHI [Analisi completa]



Raffaele Gaito

ia360.academy ➜ Studia l'IA

19 febbraio 2024

[link al post  
di Gaito](#)

back to tassonomia

Steve Nouri • 2°  
Generative AI Founder | Advisor @ Fortune 500 | 1.5 ...  
11 ore • Modificato • 🔥

I am speechless with OpenAI's new release!🔥

# 16 febbraio 2024

These impressive clips are all generated by OpenAI's Sora model without any modifications, highlighting its ability to create realistic and imaginative scenes.

The goal behind Sora is to develop models that understand and simulate real-world dynamics, assisting in solving practical problems.

Here is what you need to know:

⌚ Ability to generate detailed videos from complex prompts, like a stylish woman walking in Tokyo or a movie trailer.

✓ Maintains visual quality and consistency throughout the video.

👉 Understands the physical world, ensuring characters and scenes behave in a believable manner.

🧠 Users can guide what Sora remembers and forgets, enhancing personalization in future interactions.

💻 Utilizes a diffusion model and transformer architecture for superior scaling, capable of extending videos and animating still images accurately.



back to tassonomia

Business Intelligence, Analytics & Data Visualization 16 febbraio 2024  
Jaeden Schafer • 3° e oltre  
3 ore • 🔁

🚨 BREAKING: OpenAI has just announced their new Text-To-Video model called Sora.

The quality is absolutely insane.

My year long prediction was correct. We will be able to direct our own movies VERY soon.

I think I speak for everyone when I say this is incredibly exciting and also concerning.

This is the prompt used for this video:

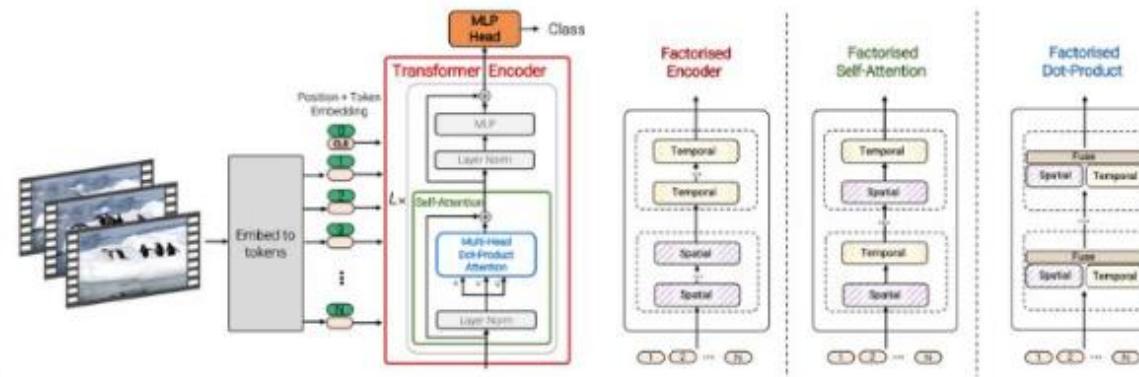
Several giant wooly mammoths approach treading through a snowy meadow, their long wooly fur lightly blows in the wind as they walk, snow covered trees and dramatic snow capped mountains in the distance, mid afternoon light with wispy clouds and a sun high in the distance creates a warm glow, the low camera view is stunning capturing the large furry mammal with beautiful photography, depth of field.



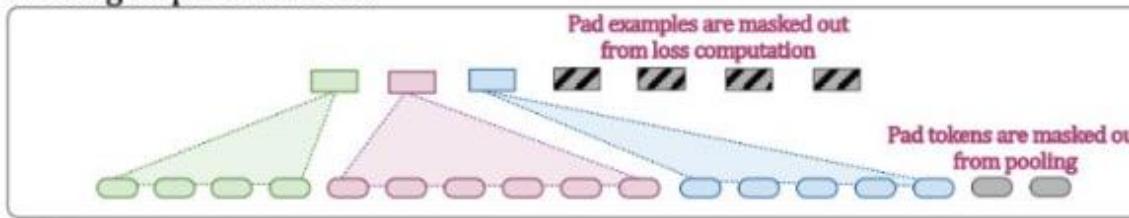
Prompt: Several giant wooly mammoths approach treading through a snowy meadow, their long wooly fur lightly blows in the wind as they walk, snow covered trees and dramatic sno...  
more

3/16

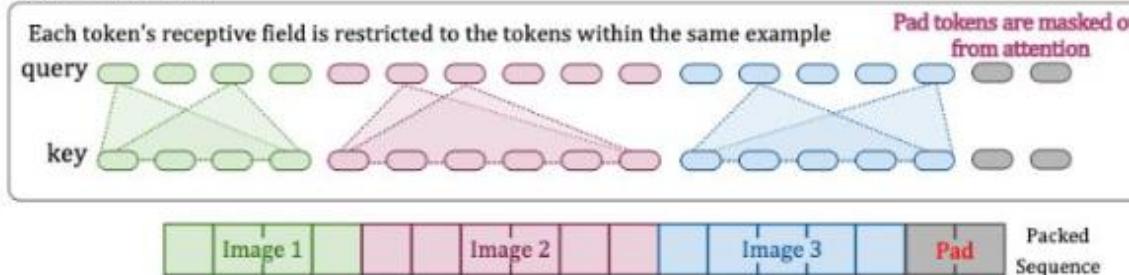
# Big Ideas in OpenAI Sora: Video Compression, Space-Time Patches, Transformer model



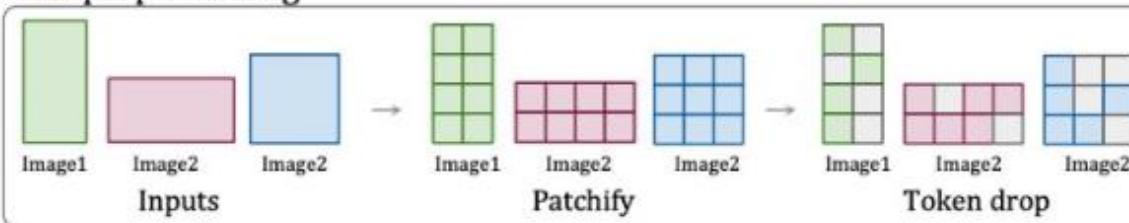
## Pooling Representations



## Self-Attention



## Data preprocessing



Machine Learning Community (Moderated)

Aishwarya Naresh Reganti • 3° e oltre

3 ore •

What makes OpenAI's **#Sora** so good at creating realistic videos? Here are the big research ideas behind the model architecture:

### 1. Turning Visual Data into Patches

Inspired by the success of LLMs, Sora employs a similar approach by transforming visual data into patches. This method involves compressing videos into a lower-dimensional latent space (using a compression network) and then decomposing them into spacetime patches, which serve as the basic building blocks for the model.

This patch-based approach allows Sora to efficiently handle diverse types of videos and images in terms of duration, aspect ratio, and resolution. Sora leverages the transformer architecture, known for its effectiveness in various domains like language and image processing, adapted here for video generation.

It uses the popular diffusion approach where the model is trained to de-noise input patches based on conditioning information, such as text prompts. This enables the generation of high-quality videos from noisy inputs.

### 2. Training on Variable Durations, Resolutions, Aspect Ratios

Unlike traditional methods that may require standardizing video dimensions, Sora can handle native-sized data, offering benefits like improved composition, framing, and the ability to generate content for different devices at their native resolutions.

Beyond generating new content, Sora can edit existing videos or images, animate static images, and extend videos in time, creating seamless loops or transformations based on textual descriptions.

Sora exhibits emergent capabilities such as 3D consistency, long-range coherence, object permanence, and interaction simulations. These aspects indicate its potential as a simulator for various scenarios in the physical and digital worlds.

The below image shows an example of how space-time visual patches can be used to encode videos/images.

#sora technical report by OpenAI: <https://lnkd.in/e5ixUAsD>

Papers that use similar architectural ideas as #sora :

1. <https://lnkd.in/ekBf-UG6>
2. <https://lnkd.in/eZF6Hyev>

articolo 1

articolo 2



back to tassonomia



Generative AI  
2.886.905 follower  
7 ore • 🌎

Wow. This is Sora, a new text-to-video model by [OpenAI](#).

Here are 10 minutes of creations from Sora 🔥



Generative AI  
2.896.522 follower  
1 ora • 🌎

19 febbraio 2024

**BREAKING: AI Sound Effects are coming 🎵**

Sora by OpenAI makes great videos, but there's no sound.

What if you could describe a sound and generate it with AI?

ElevenLabs has just announced that AI-generated sound effects are coming soon! Simply describe a sound, and the AI will handle the rest.





Steve Nouri • 2°

Generative AI Founder | Advisor @ Fortune 500 | 1.6 ...

13 ore •

👉 Sora also changed video editing forever! 🎉

AI-generated videos by Sora are just mindblowing but this new feature is also as impressive.

Sora chops up visuals into bits called patches, kinda like breaking down a sentence into words, and uses them to whip up new videos.

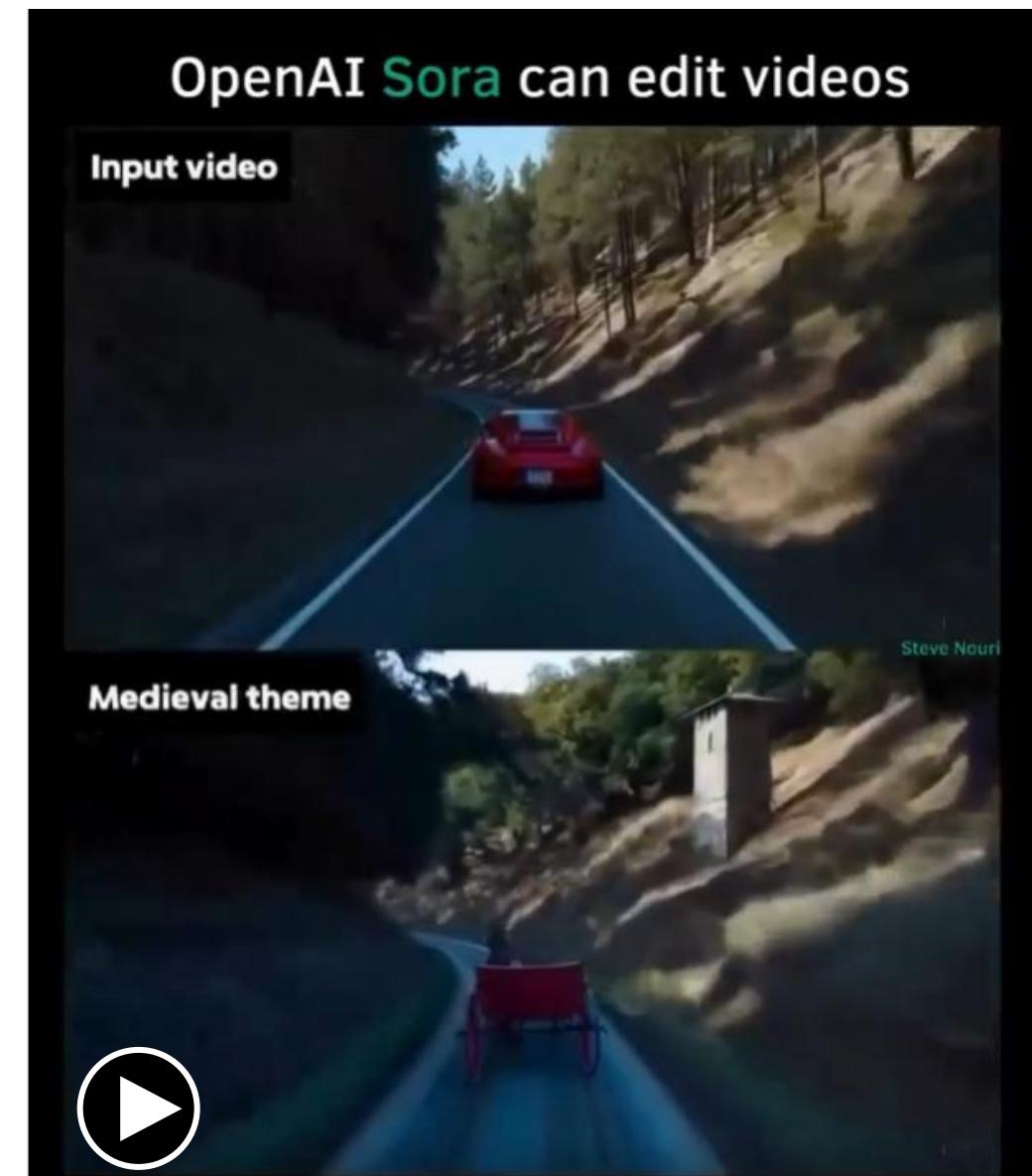
It's like it's practicing to become a mini-universe simulator, showing us bits of the real and digital worlds, though it's still learning the ropes on getting all the physics and interactions just right.

You can tell it to change a video's setting, like turning a city scene into a lush jungle, just with a simple prompt.

It can connect two totally different videos smoothly

Or maybe you want to make a video loop perfectly or stretch a short clip into a longer story.

Sora can tweak videos to do all that, making it feel like you're crafting a whole new world or extending the storyline of your favorite moments.



19 febbraio 2024



New Will Smith eating spaghetti / pasta video | OpenAI Sora | AI video



Steve Nouri • 2°

Generative AI Founder | Advisor @ Fortune 500 | 1.6 ...

6 ore • Modificato • 🔍

OpenAI dropped the bomb on video generation a couple of days ago.🔥

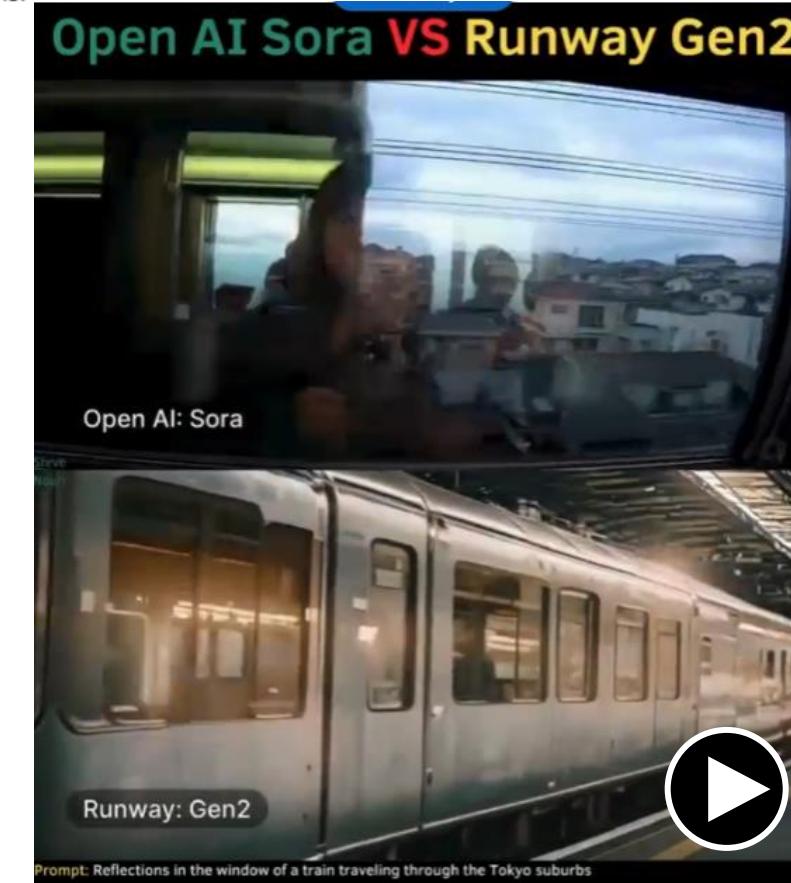
But how do they compare to their closest rival?

OpenAI: Sora vs Runway: Gen2

The space is so brutal, that openAI is going for all generative models (text, Image, Video), and maybe soon sounds.

At least as a result of such a fierce competition users are getting access to dreamy tools.

### Open AI Sora VS Runway Gen2



video SORA più veloci ed animati? video RunWay con migliore definizione?



back to tassonomia

7/16

22 febbraio 2024

## SORA DI OPENAI, L'IA CHE GENERA VIDEO ULTRA REALISTICI, SECONDO META È DESTINATA A FAILIRE

Sono passati pochi giorni dalla [presentazione di Sora da parte di OpenAI](#), l'IA che ha stupito a lasciato tutti a bocca aperta per la sua incredibile capacità di generare video ultra realistici in pochi minuti, ma non tutti sono d'accordo sulle sue potenzialità.

Yann LeCun, il capo scienziato della divisione IA di Meta, infatti, **non si è detto molto colpito dal modello text-to-video**. LeCun in particolare contesta alcune affermazioni di OpenAI secondo cui Sora porterà alla fine di "*simulatori generici del mondo fisico*", e se così fosse a detta del ricercatore di Meta l'approccio di OpenAI è completamente sbagliato.

*"Modellare il mondo per l'azione generando pixel è altrettanto dispendioso e destinato al fallimento quanto l'idea in gran parte abbandonata dell'analisi per sintesi"* scrive LeCun su un post pubblicato sul proprio account X. Dichiarazioni importanti, che arrivano da quello che secondo molti è uno dei padroni dell'intelligenza artificiale, che mai prima d'ora aveva criticato il lavoro portato avanti dai concorrenti.

Secondo LeCun, **l'approccio di Sora che genera pixel da variabili latenti esplicative** sarebbe inefficiente e non permette di affrontare l'incertezza che deriva dalle previsioni di uno spazio 3D, che è caratterizzato da una serie di variabili molto complesse. Sostanzialmente, quindi, ritiene che i modelli come Sora stanno cercando di dedurre troppi dettagli non rilevanti non porteranno a nulla: l'esempio è quello di una persona che vuole calcolare la traiettoria di un pallone da calcio prendendo in esame come entra in gioco ogni materiale da cui è composto, piuttosto che concentrarsi esclusivamente sulla massa e velocità della palla.

LeCun in un secondo post sottolinea che **"non c'è niente di sbagliato in questo se lo scopo è effettivamente generare video"**, ma **"se lo scopo è capire come funziona il mondo, è una proposta destinata a fallire"**.

Tale approccio si è rivelato vincente per modelli linguistici come ChatGPT, dal momento che il "testo è discreto con un numero finito di simboli", ma la simulazione del mondo è molto più complessa.



## Analisi

# Perché Sora di OpenAi fa così paura all'industria culturale?

Il nuovo modello di Ai che trasforma in video brevi descrizioni di testo ha messo a dura prova la pazienza dei detrattori dell'intelligenza artificiale generativa.

 [Servizio di Luca Tremolada](#) 25 febbraio 2024

Sora è il nuovo modello di intelligenza artificiale generata text-to-video presentata da OpenAi la settimana scorsa. Come funziona?

Scrivi la trama di un video, decidi dove metti le telecamere, cosa si vede, cosa accade, in quale stile e l'Ai di Sam Altman realizza un video della durata un minuto. Il servizio non è ancora aperto al pubblico anche perché devono ancora capire come proteggerla da eventuali problemi dovuti alla produzione di video falsi, ma è già nelle mani di una selezionata schiera di registi, videomaker e addetti ai lavori che hanno incominciato a produrre sui social le loro opere.





Generative AI

3.144.629 follower

12 ore • 🔍

**First ever 1-minute long video generated by OpenAI revolutionary text-to-video model - Sora 🔥**

Prompt: fly through tour of a museum with many paintings and sculptures and beautiful works of art in all styles



back to tassonomia



Steve Nouri • 2°

Generative AI Founder | Advisor @ Fortune 500 | 1.6 ...  
5 ore • Modificato •

26 marzo 2024

The Balloon Man!

A short movie made by OpenAI Sora

Soon, anyone will generate clips, shorts, movies, TV shows, and ads at low, in minutes.

This is not about helping Hollywood, it's about enabling everyone to express themselves ( hopefully responsibly!)



back to tassonomia

# Stable Cascade è il generatore di immagini IA tutto nuovo. Ecco cos'è e come provarlo

di Sergio Donato - 14/02/2024 10:49

15



Lo ha realizzato Stability AI e il suo addestramento ha un costo 16 volte inferiore rispetto a Stable Diffusion. Lo si può provare con una demo.

Stability AI, la società di IA generativa che ha realizzato Stable Diffusion,

X Posta ha rilasciato un modello tutto nuovo text-to-image per creare immagini dalla loro descrizione testuale. **Si chiama Stable Cascade e si propone di ridurre di 16 volte i costi necessari al suo addestramento.**

La riduzione dei costi di addestramento è determinata soprattutto dall'architettura Würstchen su cui si basa Stable Cascade.

Würstchen è un sistema di codifica autoregressivo per la compressione di immagini. Per sistema autoregressivo si intende modello statistico che rappresenta una sequenza di dati come una serie di variabili casuali dipendenti l'una dall'altra. In parole povere, **il sistema predice il valore di ogni elemento della sequenza in base ai valori degli elementi precedenti.**



# Primo spot pubblicitario realizzato con Sora di OpenAI, le reazioni

link all'articolo  
con il video



ADVERTISEMENT

Il celebre rivenditore americano di giocattoli Toys 'R' Us ha recentemente annunciato di essere il primo marchio ad aver realizzato un **video promozionale** utilizzando **Sora**, il modello AI text-to-video sviluppato da OpenAI.



Tiziana  
Foglio  
Pubblicato il



back to tassonomia

[OpenAI has officially launched Sora](#), its first text-to-video AI model, now available to **ChatGPT Plus** and **Pro** subscribers. Sora allows users to create highly realistic videos by transforming text prompts into visually compelling content. Here's a detailed breakdown of what Sora offers:

## Key Features of Sora

### 1. Text-to-Video Generation:

- Users can generate videos up to **20 seconds long** with prompts.
- Supported resolutions include up to **1080p** with **16:9, 1:1, and 9:16** aspect ratios for versatility.

### 2. Image Animation:

- Static images can be animated based on descriptive prompts, adding dynamic effects and realistic motion.

### 3. Video Remixing:

- Existing video clips can be modified or extended using prompts, enhancing creativity in content creation.

### 4. AI Safeguards:

- **Watermarking:** All AI-generated videos include visible watermarks to ensure transparency.
- **C2PA Metadata:** Videos carry secure metadata indicating they were generated by AI, supporting authenticity checks.

## Availability & Rollout

- **Global Access:** Sora is accessible in several countries, excluding the UK, Switzerland, and the European Economic Area, with expansion plans set for 2024.

## Use Cases for Sora

- **Content Creation:** Marketers, creators, and businesses can generate promotional content and dynamic visuals.
- **Film & Media Production:** Sora can assist filmmakers in pre-visualization, concept design, and animation.
- **Educational & Training Videos:** Educators can produce engaging learning materials quickly.

Sora represents OpenAI's next step in generative AI, aiming to revolutionize video content creation with professional-quality outputs powered entirely by AI. This move marks a significant leap forward in text-to-video technology, opening new creative possibilities for businesses and creators alike. 

annuncio  
ufficiale di  
OpenAI



link all'articolo  
con video  
ufficiale

# OpenAI ha lanciato Sora, l'AI che genera video realistici da un semplice testo

10 Dicembre 2024  30

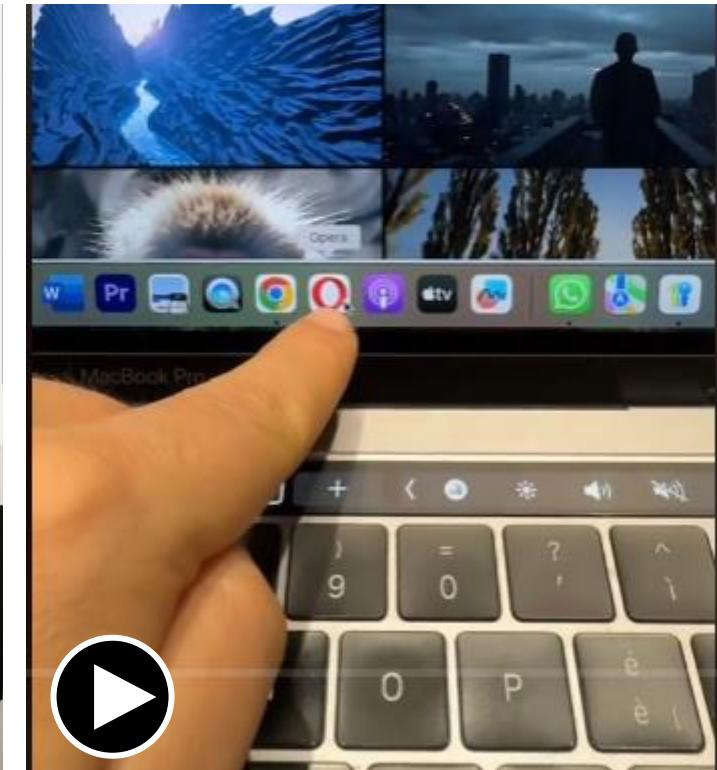
**SORA Turbo**  
**dicembre 2024**

15/16



ia360.academy

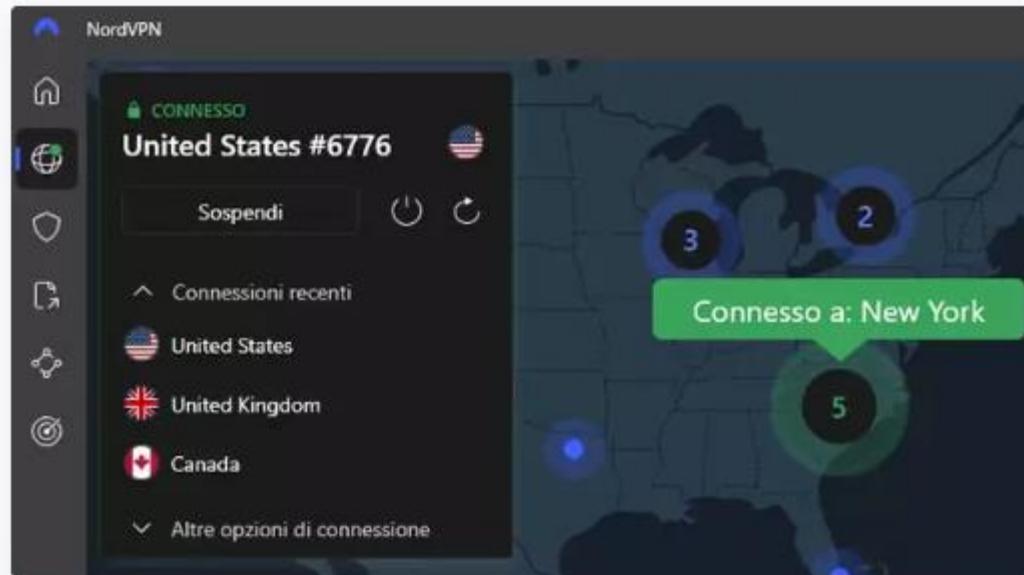
FINALMENTE Sora è arrivato: tutorial completo in italiano



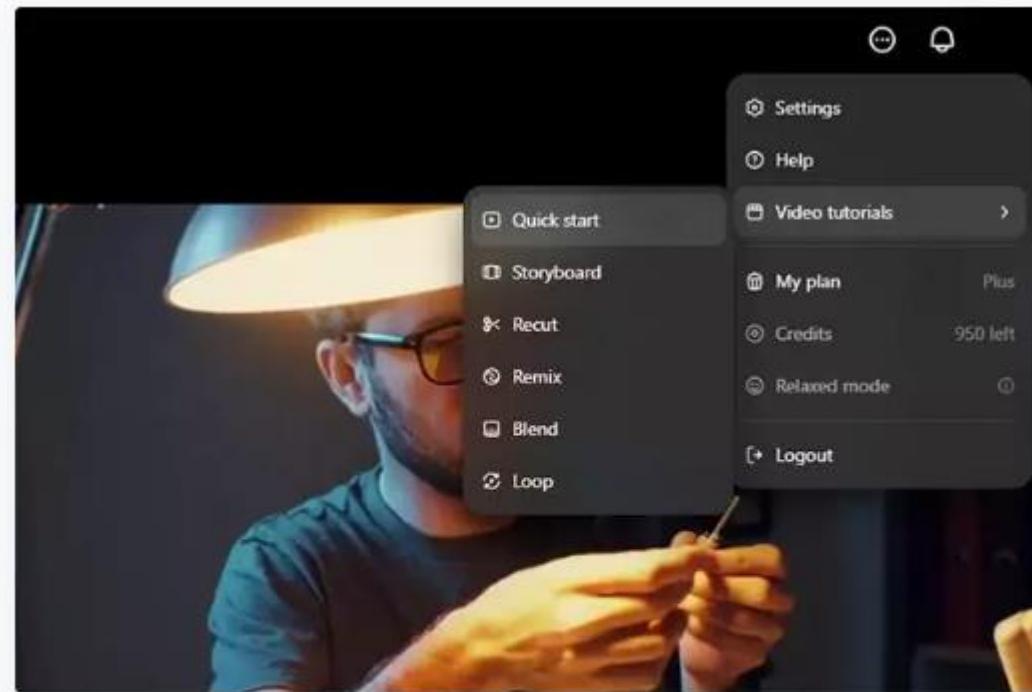
# Sora è disponibile: come accedere subito dall'Italia

## OpenAI regala Sora per Natale: accedi dall'Italia

Vuoi sapere **come fare?** È molto semplice: non devi far altro che aprire l'app del servizio e connetterti a un server collocato in uno dei tanti paesi supportati, ad esempio gli Stati Uniti, per ottenere un indirizzo IP geolocalizzato in quel territorio. Poi, ti basta collegarti al sito [sora.com](https://sora.com) per iniziare a **creare video con l'intelligenza artificiale**.



Una volta effettuato l'accesso (l'account è lo stesso utilizzato per ChatGPT Plus) ci si trova di fronte all'**editor** in cui inserire il **prompt testuale** da trasformare in un filmato, impostando parametri come l'aspect ratio, la risoluzione e la durata. Al termine del rendering è possibile intervenire con modifiche mirate: ci sono diversi tool per farlo, da Recut a Remix, da Blend a Loop.



Ci sarà tempo per approfondire tutte le **funzionalità** integrate e per valutarne l'efficacia, anche considerando le alternative proposte dalla concorrenza. In questo momento, l'accesso dall'Italia attraverso una soluzione come la rete globale di [NordVPN](#) (guarda l'offerta) permette di prendere confidenza con una tecnologia che presenta al tempo stesso enormi potenzialità e ampi margini di miglioramento.

Vale la pena sottolineare che, secondo le pagine del supporto ufficiale di **OpenAI**, l'utilizzo dei propri strumenti come Sora da paesi per i quali non è ancora previsto l'accesso (Italia compresa) può comportare il blocco o la sospensione dell'account.





Donald Trump and Kamala Harris's Campaign Speeches analysed with a word cloud and summarized with artificial intelligence.

- *Full speeches coming from New York Times*
- *Word cloud made with freewordcloudgenerator.com*
- *Summaries made by Anthropic Claude 3.5*

Massimo Canducci • 1°

Innovation @ Engineering Group | Expert on "Innovation a...

1s •

### From a Photo to a Complete Digital Avatar

- ◆ Using AI, it is possible to transform a painting or an old picture into a photorealistic digital avatar.
- ◆ When experimenting with portraits from centuries ago, it's an interesting endeavor—WOW, we can see what Mozart or Leonardo looked like.
- ◆ At some point, we'll start doing this with pictures of our grandparents, parents, children, friends, and loved ones who are no longer with us, and this will have a completely different effect on us.
- ◆ The next step will be equipping them with appropriate knowledge, voice, and consistent movements. At that point, we will have digital clones of our deceased loved ones, capable of interacting with us in a manner very similar to how they did in the past, and this will take things in a very, very, very different direction.
- ◆ Would you like to have digital clones of deceased people available to you?
- ◆ And would you accept, if in the future, to be cloned in this way?
- ◆ Do you think these aspects should be regulated in some way?

agosto 2024

Image2Video



back to tassonomia

# GPT4 Turbo with Vision

Enhancements



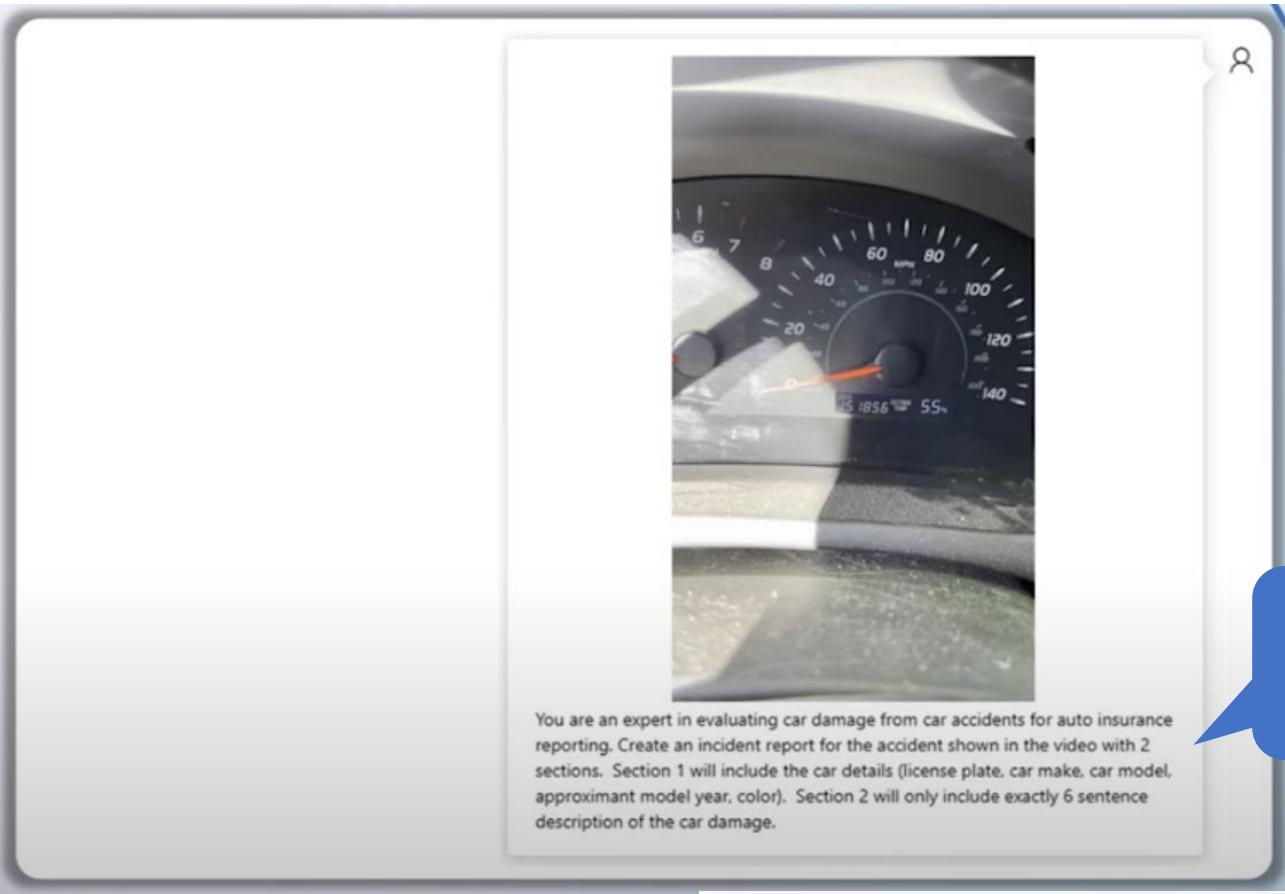
Roberta Bruno – Microsoft –  
25'39'' → 28'54''

## Video 2 Text.

GPT-4 Turbo with Vision accetta in entrata anche immagini e video (oltre al prompt testuale), a differenza di Dall-e 3 (T2I). Fruibile da Azure o tramite API di OpenAI



back to tassonomia



You are an expert in evaluating car damage from car accidents for auto insurance reporting. Create an incident report for the accident shown in the video with 2 sections. Section 1 will include the car details (license plate, car make, car model, approximant model year, color). Section 2 will only include exactly 6 sentence description of the car damage.

Il system message e la richiesta

L'output

### Section 1: Car Details

- License Plate: MAT 251, Minnesota
- Car Make: Toyota
- Car Model: Camry
- Approximate Model Year: Based on the design and style, it appears to be a model from the late 2000s to early 2010s.
- Color: Blue

### Section 2: Description of Car Damage

The rear right side of the blue Toyota Camry has sustained significant damage, characterized by deep scratches and scuff marks. These marks are most prominent around the wheel arch and extend to the rear door and the area around the fuel cap. The rear bumper displays noticeable scuff marks, with the paint chipped off in various spots, revealing the underlying surface. Indentations on the bodywork are visible, especially around the rear right corner, indicating a forceful impact. The left side of the car is not spared, showcasing long, curved scratches on both rear and front doors. The extent and nature of the damage suggest the vehicle was involved in a sideswipe accident or scraped against a rough surface.



Generative AI  
1.514.379 follower  
1s •

Link alla pagina  
wikipedia IT della  
modella Kendall Jenner

We're treading into the realms of Black Mirror with Meta's new venture that is as intriguing as it is eerie! 00 ↗

Say hello to Billie, a.k.a Kendall AI, a virtual persona crafted with the visage of Kendall Jenner, brought to life by Meta AI.

This isn't a whimsy name change but a full-fledged AI avatar, a part of Meta's audacious new AI experiences.

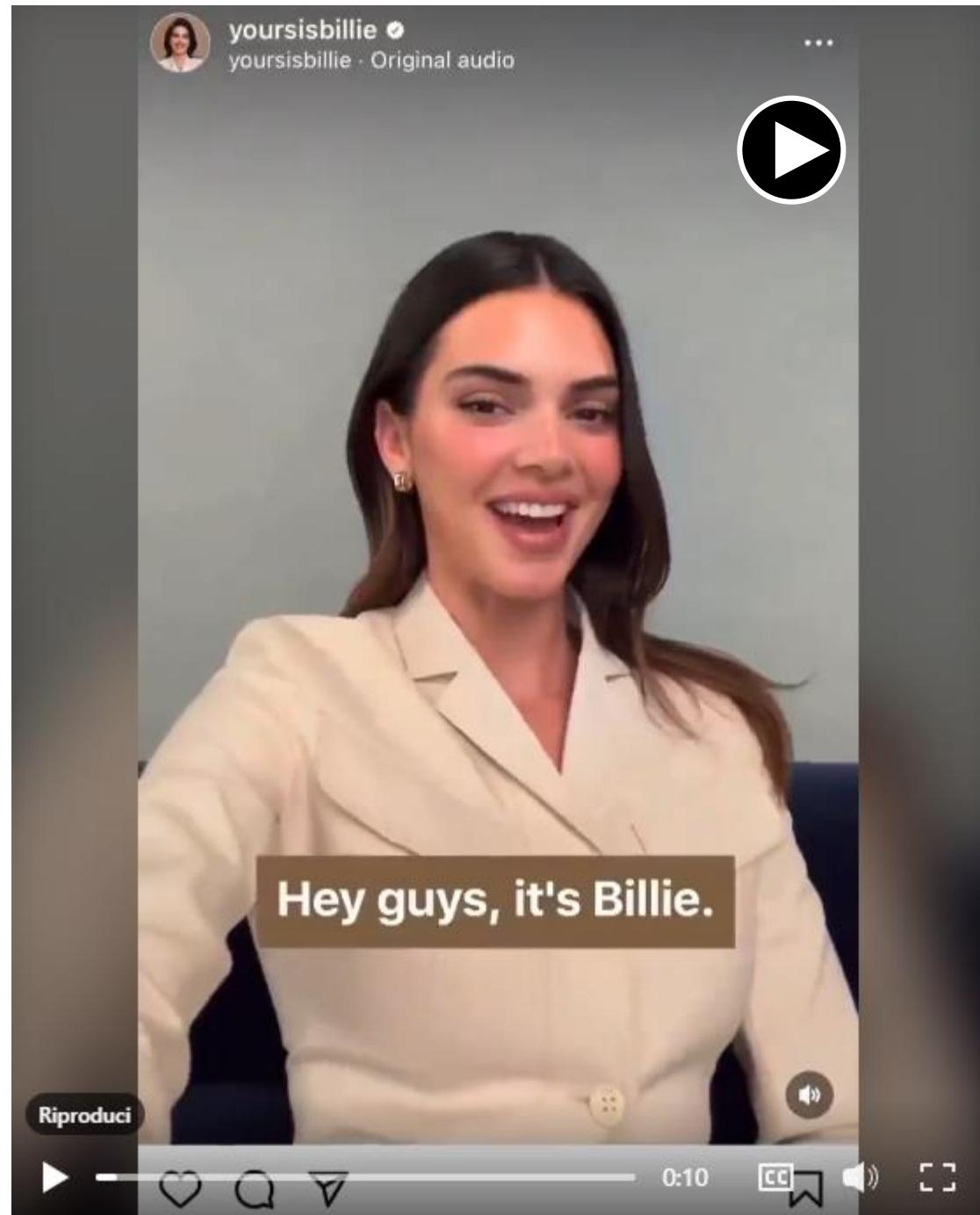
The outcome? A blend of amusement and bemusement as followers find 'Kendall' asserting her name as Billie. This venture is driving us straight into the heart of the Uncanny Valley, with followers finding it hard to distinguish the real Kendall from her AI counterpart.

Swipe through to see the uncanny resemblance and the futuristic spectacle that Meta AI is unfolding:

- ⌚ 1 - Billie, not Kendall, greeting her 106,000 followers.
- ⌚ 2 - An eerily lifelike video introduction by Billie, leaving followers befuddled and intrigued.
- ⌚ 3 - Comments flooding in, expressing awe and dread in equal measure.
- ⌚ 4 - A glimpse into the future? Other celebs like Tom Brady and Paris Hilton are also on board!

The lines between reality and virtuality are blurring, and as Mark Zuckerberg envisions, these AI entities reflect a 'huge need' for celeb AI chatbots, bridging the gap between stars and their global fanbase.

But as we slide into this new era, the question looms - Are we ready to embrace our virtual counterparts?



25 ottobre 2023



*Video2Video ed altro*



Raffaele Gaito • 1°

growthprogram.academy ➜ Studia con me

4 giorni • ⓘ

Le novità mostrate da Adobe sono FANTASCIENZA pura.

Se non sapete di cosa sto parlando, vedetevi il mio ultimo video dove commento la conferenza Adobe Max.

Non crederete ai vostri occhi...



back to tassonomia



Generative AI  
1.114.461 follower  
5 ore •

💡 Google Shopping now utilizes cutting-edge generative AI to display clothes on diverse body types! This innovative step gets us closer to the concept of a 'magic mirror' that reveals how different clothes would look on you. 🎀👚

As we all know, online shopping has seen a tremendous growth but the challenge of high return rates due to mismatched expectations continues to persist. Google's AI solution is aimed at overcoming this hurdle by increasing buyer's confidence, and also paving the way for easily customizable clothing. 💡

This technology not only lets us visualize clothing on different body types but also improves the overall shopping experience with new filters that help find exactly what you're looking for. 💪

23 giugno 2023

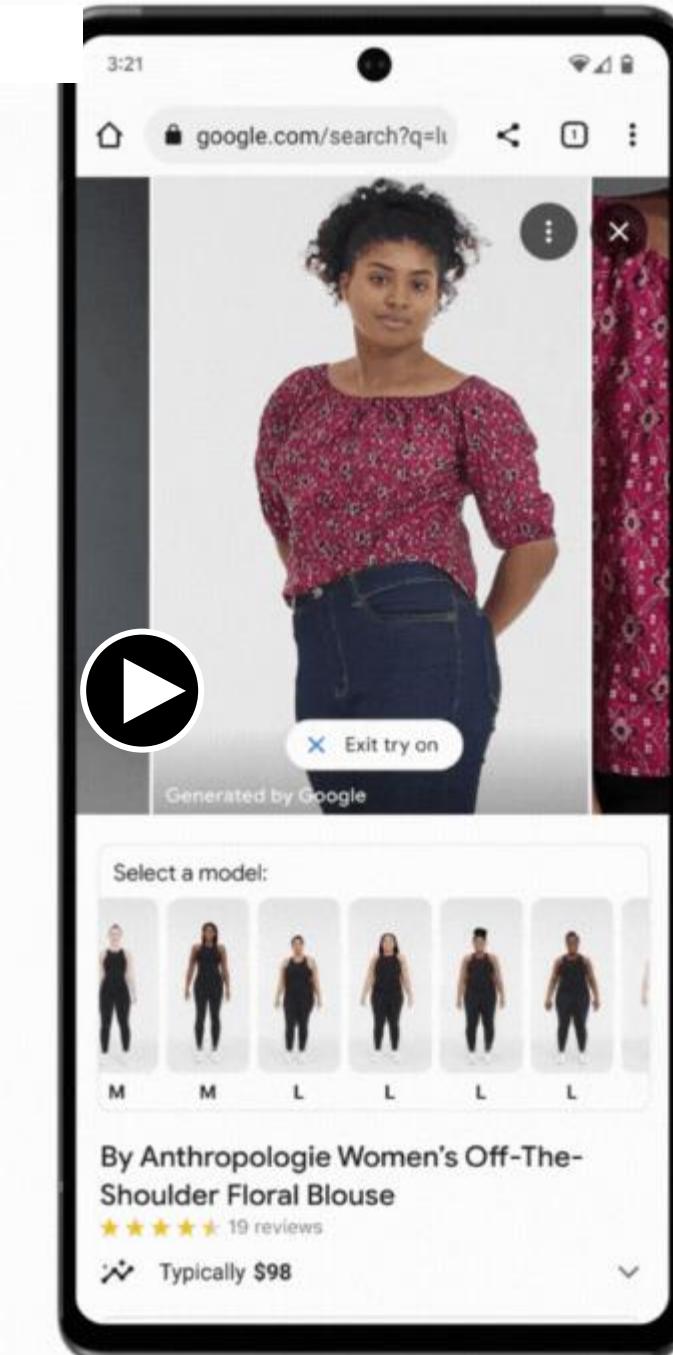


Image2Image



back to tassonomia

27 marzo 2023



Business Intelligence, Analytics & Data Visualization (Moder... • • • X

Aditi Takale • 3° e oltre

15 ore • 15

Amazing new AI technology from NVIDIA called Eye Contact...

It uses AI to change the position of your eyes to make them appear to be focused on the camera!

Read more here: <https://lnkd.in/dqyZhtiv>

credits: @1030 (t/w)

#deeplearning #artificialintelligence #innovation #ml #machinelearning  
#datascience #technology #ai

[Vedi traduzione](#)



Image2Image



back to tassonomia

Massimo Canducci • 1°

Chief Innovation Officer | Executive MBA Professor | Auth...  
22 ore • 🔍

marzo 2023

### AI-generated virtual pharmacist

Using three different generative AI tools, a virtual pharmacist has been created, capable of generating responses to health-related questions and advising on medicines and health treatments.

- ChatGPT for generating information to be transmitted
- ElevenLabs for generating the voice
- Photoleap for generating the final video

The concatenated and cross-used of different generative tools is very interesting. It is a bit concerning that this prototype was proposed in the medical field, as we must always keep in mind that these tools were designed to APPEAR reliable, not to BE reliable, therefore the "advice" of this machine could be completely wrong. We might soon move from "I read it on the internet" to "the virtual pharmacist told me."

Don't try this at home!

Farmacista virtuale generato dall'intelligenza artificiale

Utilizzando tre diversi strumenti di intelligenza artificiale generativa, è stato creato un farmacista virtuale, in grado di generare risposte a domande relative alla salute e consigliare farmaci e trattamenti sanitari.

- ChatGPT per generare informazioni da trasmettere
- ElevenLabs per generare la voce
- Photoleap per generare il video finale

La concatenazione e l'utilizzo incrociato di diversi strumenti generativi è molto interessante.

È un po' preoccupante che questo prototipo sia stato proposto in campo medico, in quanto dobbiamo sempre tenere presente che questi strumenti sono stati progettati per APPARIRE affidabili, non per ESSERE affidabili, quindi il "consiglio" di questa macchina potrebbe essere completamente sbagliato.

Potremmo presto passare da "l'ho letto su Internet" a "mi ha detto il farmacista virtuale".

Non provare questo a casa!



1 ottobre 2023

# Spotify'a AI Voice Translation, YouTube Aloud (5'17")

*Speech2Speech*

The screenshot shows the YouTube channel page for 'For the Record' on Spotify. The channel has 1.1M subscribers. The main video thumbnail features two hosts, a woman in a yellow hoodie and a man in a white shirt, sitting in a studio. Below the thumbnail is a portrait of a man with a beard. The channel banner is purple with green soundwave graphics and the text 'Introducing Voice Translation for podcasters'. The video title 'Spotify's AI Voice Translation Pilot Means Your Favorite' is visible at the bottom.

For the Record

Stream On 2023 News Podcasts Company Info Communities Get Spotify

Introducing  
Voice Translation  
for podcasters

STORIES FOR THE WORLD

Spotify's AI Voice Translation  
Pilot Means Your Favorite



back to tassonomia

Massimo Canducci • 1°

Chief Innovation Officer | Executive MBA Professor | Author and International ...

...

Generative AI for Video Production

We are now accustomed to the automatic generation of texts and high-quality images using simple textual prompts.

The next step is video production. I've written about it in the past, but this tool (link in the first comment) really seems like a step forward.

It is (or will be) possible to generate and edit videos starting from simple prompts or to animate static images provided to the machine.

Three important things to consider:

1) As with images, the creation of simple content will be accessible to everyone, but the creation of professional videos will require huge skills, and professionals in this field should immediately study how to get the best out of these platforms.

2) The generation of images is for today, with all its limits, although it will still improve. Video generation is here now.

3) The next generative wave will involve immersive environments for virtual reality and metaverse.

## Text 2 Video con Pika

30.11.23 – 1.12.23



## PIKA: video con l'intelligenza artificiale PAZZESCHI



Raffaele Gaito

ia360.academy Studia l'IA

176 articoli

Già segui

1 dicembre 2023

Da giorni nell'ambiente sta circolando questo video incredibile di Pika.

Cosa è? Si tratta di un sistema *text-to-video*, quindi che permette di creare dei video da zero, semplicemente scrivendo del testo.

Dalla demo presente sul sito la qualità sembra molto superiore a Runway e altre IA viste fino ad oggi.



back to tassonomia



UNA PRODUZIONE  
**Foeus**  
MEDIASET

ChatGPT



Sito Focus Mediaset  
Infinity → Dirette TV →  
Guida TV → Giorno → Ora >>  
(23.49) → canale (focus)  
→ 20'20" – 23'40"

scrivi un testo sul pericolo delle fake news con l'avvento dell'intelligenza artificiale



Free Research Preview. ChatGPT may produce inaccurate information about people, places, or facts. [ChatGPT May 24 Version](#)

Questo video illustra un caso d'uso della gen AI del tipo T2V (con elaborazione manuale) realizzato con lo strumento [Fliki](#) – un'alternativa a Synthesia (inferiore) ma con l'interessante funzionalità delle scene scelte automaticamente da Fliki in base al testo della scena («pick random»)

Attenzione: il video non risulta visibile da Jmeet. Mettere la URL in chat di Jmeet.

Nuovo link al video  
(settembre 2024)



# Generative AI

## Audio-To-Audio

«doppiaggio  
vocale»

non più  
disponibile su  
YouTube

IIElevenLabs – Voice Dubbing demo

[https://www.youtube.com/watch?v=17\\_xLsqny9E](https://www.youtube.com/watch?v=17_xLsqny9E)



Oggi vi mostro un'AI semplice, potente e soprattutto utilissima.

**Text2Diagram**

Si chiama Napkin AI e vi crea grafici in un click.

dicembre 2024

Voi gli mettete dentro un testo di qualsiasi tipo e in pochi secondi vi tira fuori dei grafici pazzeschi.

Poi potete personalizzarli ed esportarli per usarli altrove: nelle vostre slide, sul vostro sito, sui vostri social, e così via.

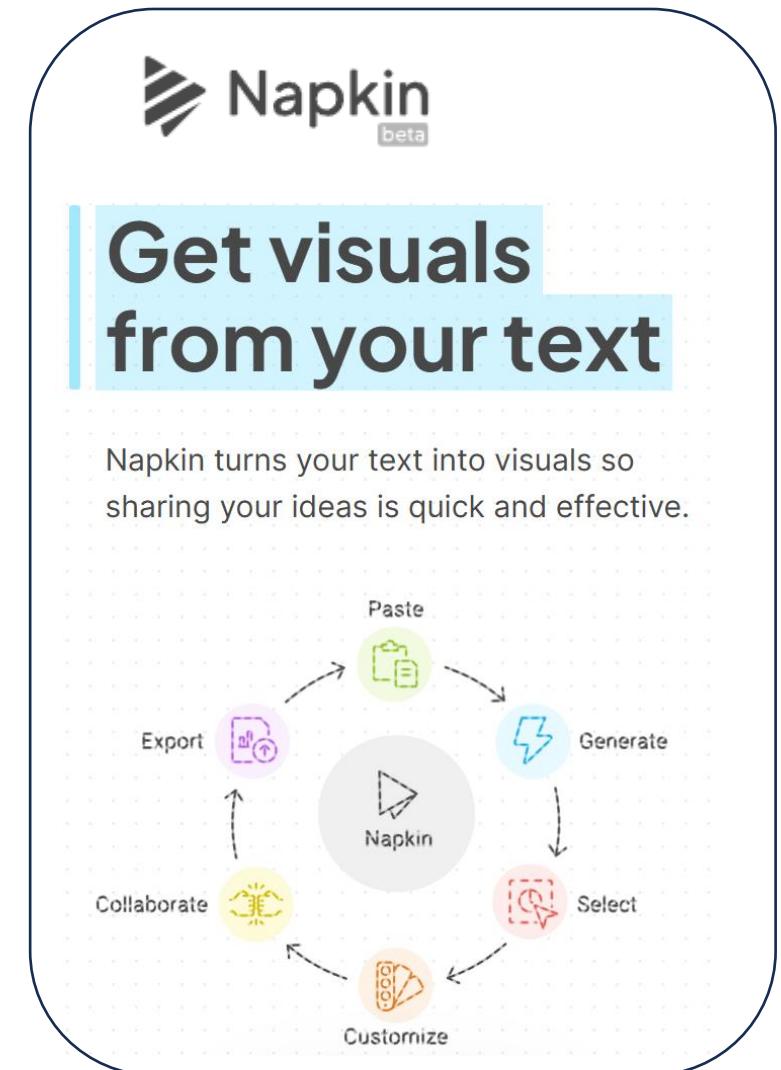
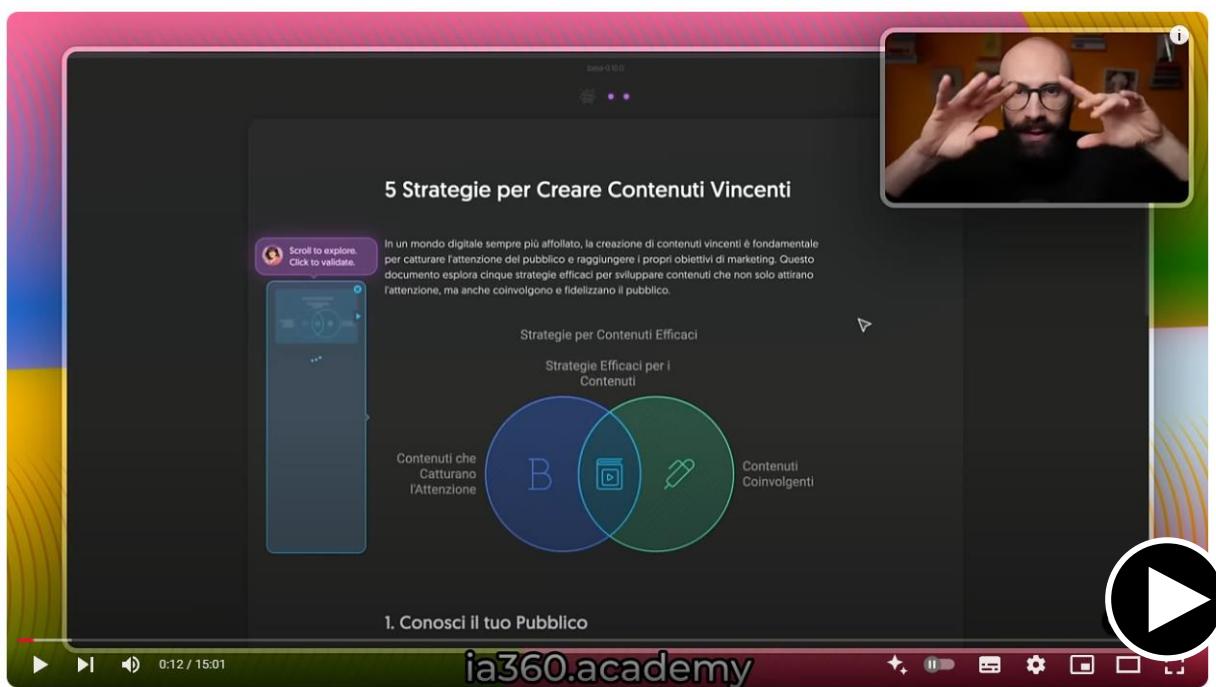
Ed è completamente gratis.

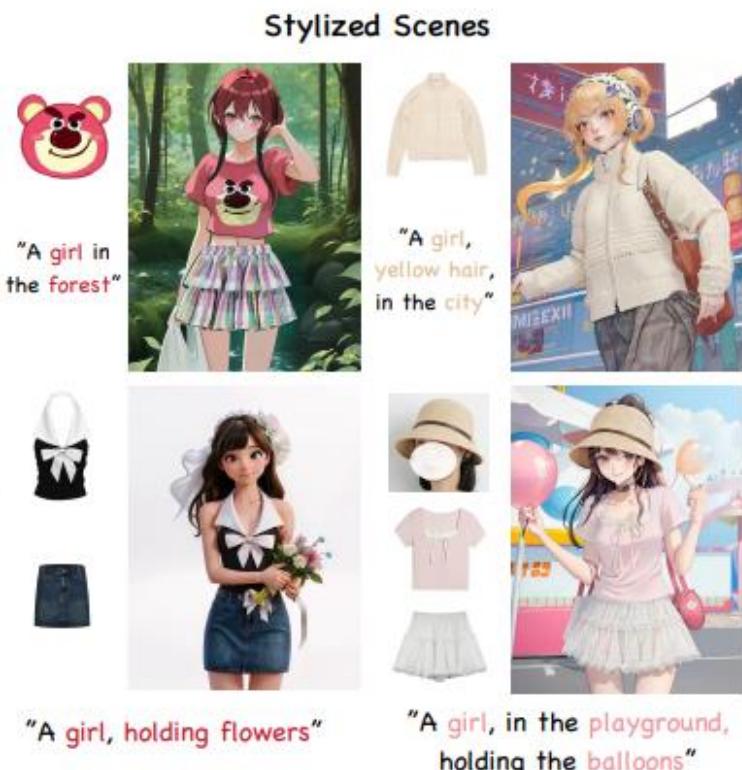
Guardate subito il tutorial e vedetela in azione.

Buona visione 😎

napkin.ai  
(letter. «tovagliolo»)

Il sito ufficiale è: <https://www.napkin.ai/>



**Reliability**

**Virtual dressing  
(Text 2 Image)**

**Compatibility**

Figure 1. Customizable virtual dressing results of our AnyDressing. **Reliability:** AnyDressing is well-suited for a variety of scenes and complex garments. **Compatibility:** AnyDressing is compatible with LoRA [15] and plugins such as ControlNet [55] and FaceID [54].





# ByteDance's AnyDressing is Making

## Virtual Clothes Shopping Smarter

### [Open Source]

ByteDance researchers have introduced AnyDressing, a new AI system that lets users visualize how multiple clothing items will look together on a digital model.

Unlike previous approaches that could only handle one garment at a time, AnyDressing can seamlessly combine multiple clothing items while maintaining high fidelity to both garment details and customized scene settings.

AnyDressing can:

- Show multiple items together (like shirts with pants and hats)
- Keep all the small details and patterns clear
- Create realistic-looking combinations
- Work alongside other popular AI tools

Online retailers can now show clothing combinations more effectively, create product images more efficiently, and provide a better shopping experience overall.

Since ByteDance has made AnyDressing open source, any company can use and improve the technology.

[link  
all'articolo](#)



# YouTube's new auto-dubbing feature is now available for knowledge-focused content

Lauren Forristal — 1:28 PM PST · December 10, 2024

YouTube [announced](#) on Tuesday that its auto-dubbing feature, which allows creators to generate translated audio tracks for their videos, is now rolling out to hundreds of thousands more channels.

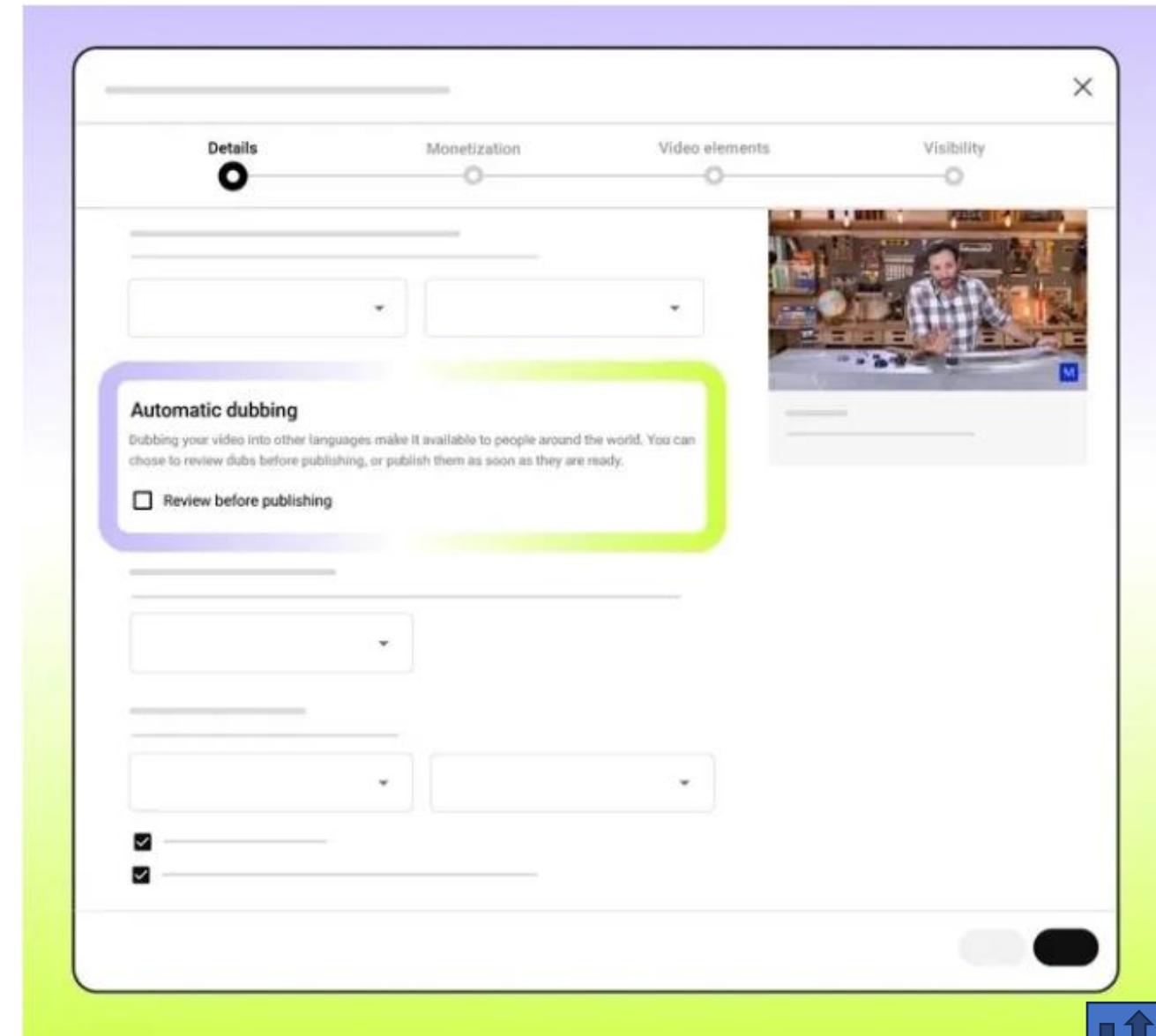
YouTube first [introduced](#) its AI-powered auto-dubbing tool at Vidcon last year, which was only being tested with a limited group of creators. This tool could help make content on the platform more accessible and easier to understand for people all over the world.

The auto-dubbing feature is now available to channels that are focused on informational content, such as videos that teach viewers how to cook or sew. It'll expand availability to other types of content soon.

To use the feature, simply upload a video as you normally would. YouTube will automatically detect the language and create dubbed versions in other languages. The tool supports English, French, German, Hindi, Indonesian, Italian, Japanese, Portuguese, and Spanish.

link  
all'articolo

TC TechCrunch



back to tassonomia

TOGETHER WITH ASSEMBLYAI

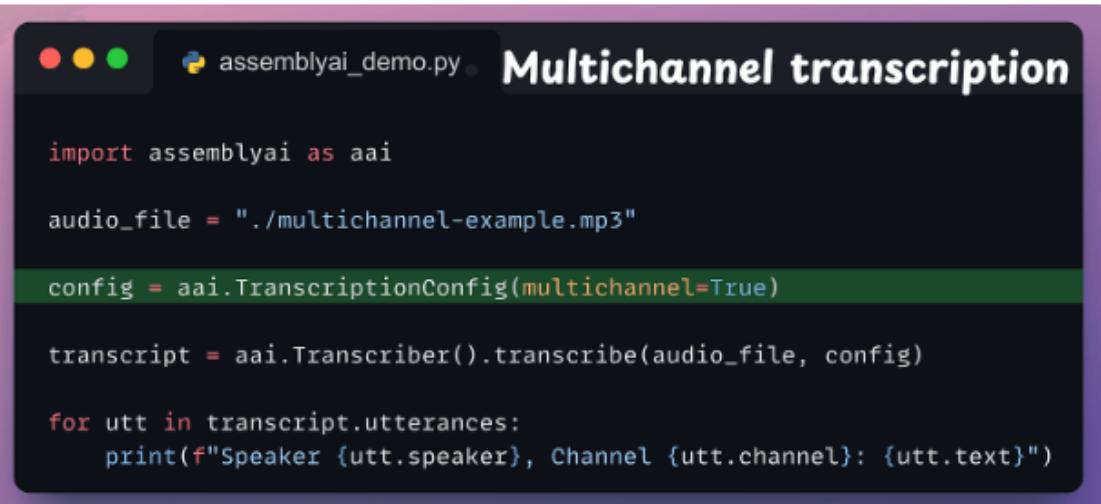
## Speech-to-text at unmatched accuracy with AssemblyAI

There are two techniques to identify and generate speaker-labeled transcripts from speech files:

- **Multichannel transcription:** Each speaker has a dedicated channel (conference calls, zoom meetings, etc.), leading to better accuracy.
- **Speaker diarization:** All speakers belong to the same channel, which requires sophisticated techniques.

[AssemblyAI's SOTA transcription models](#) can tackle both of them seamlessly in just six lines of code:

- For multichannel transcription, set `multichannel=True`:



```
assemblyai_demo.py Multichannel transcription

import assemblyai as aai

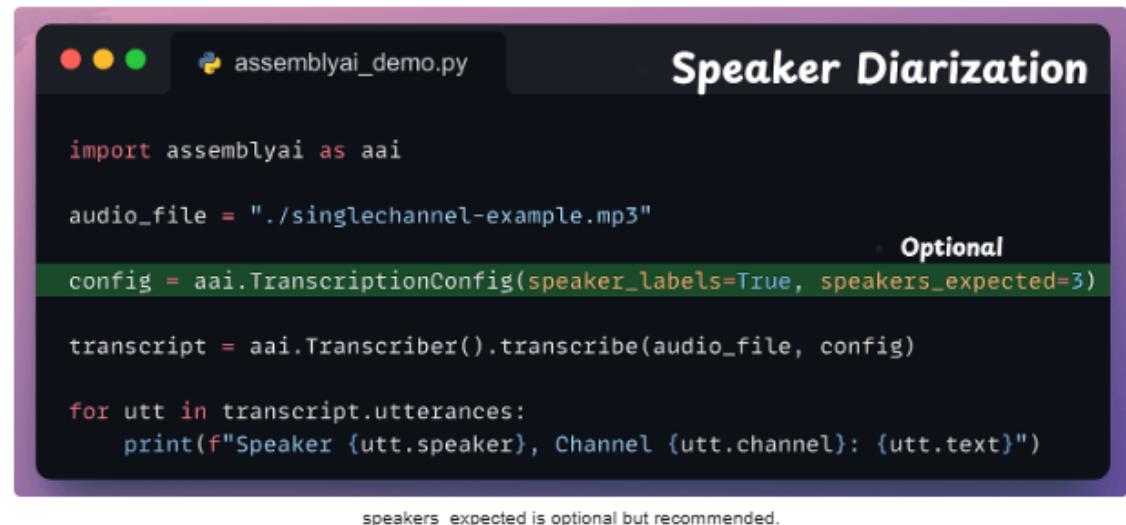
audio_file = "./multichannel-example.mp3"

config = aai.TranscriptionConfig(multichannel=True)

transcript = aai.Transcriber().transcribe(audio_file, config)

for utt in transcript.utterances:
    print(f"Speaker {utt.speaker}, Channel {utt.channel}: {utt.text}")
```

- For speaker diarization, set `speaker_labels=True` and `speakers_expected`:



```
assemblyai_demo.py Speaker Diarization

import assemblyai as aai

audio_file = "./singlechannel-example.mp3"

config = aai.TranscriptionConfig(speaker_labels=True, speakers_expected=3) Optional

transcript = aai.Transcriber().transcribe(audio_file, config)

for utt in transcript.utterances:
    print(f"Speaker {utt.speaker}, Channel {utt.channel}: {utt.text}")
```

speakers\_expected is optional but recommended.

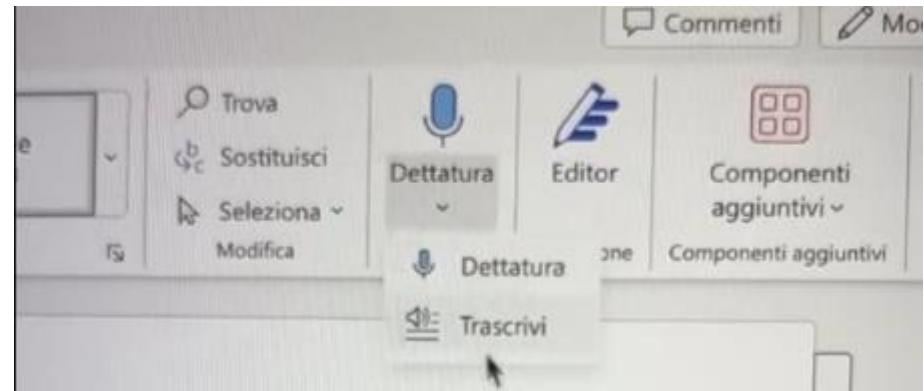
[AssemblyAI's](#) speech-to-text models rank top across all major industry benchmarks. You can transcribe 1 hour of audio in ~35 seconds at an industry-leading accuracy of >93% (for English).

Start building with  
AssemblyAI

[Sign up today](#) and get \$50 in free credits!



1. Vai su Edge (od anche su Chrome)
2. Digita *word.new* (per aprire un nuovo documento con Microsoft 365)
3. Da «dettatura»: detta oppure trascrivi un audio



**Microsoft 365**  
dicembre 2024  
(gratuito, per ora)  
Microsoft 365 gratuito?  
[Ecco perché!](#)



1. Vai a «Google AI Studio»

**Google AI Studio**

2. Scegli a sx «Stream Realtime»

**Stream Realtime**

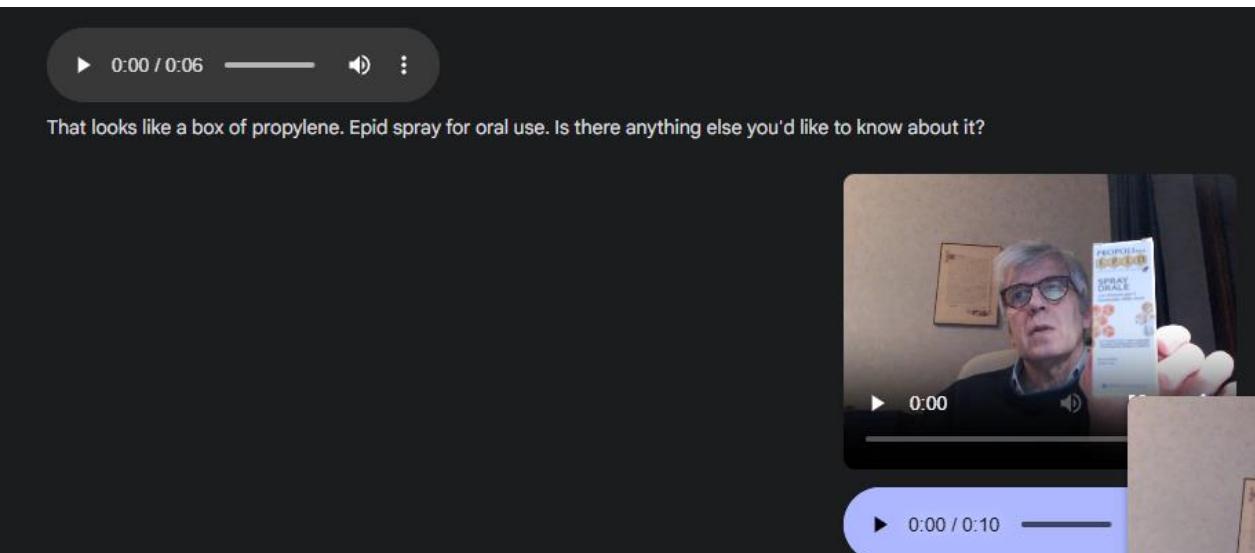
3. Premi il bottone al centro «Show Gemini»

**Show Gemini**

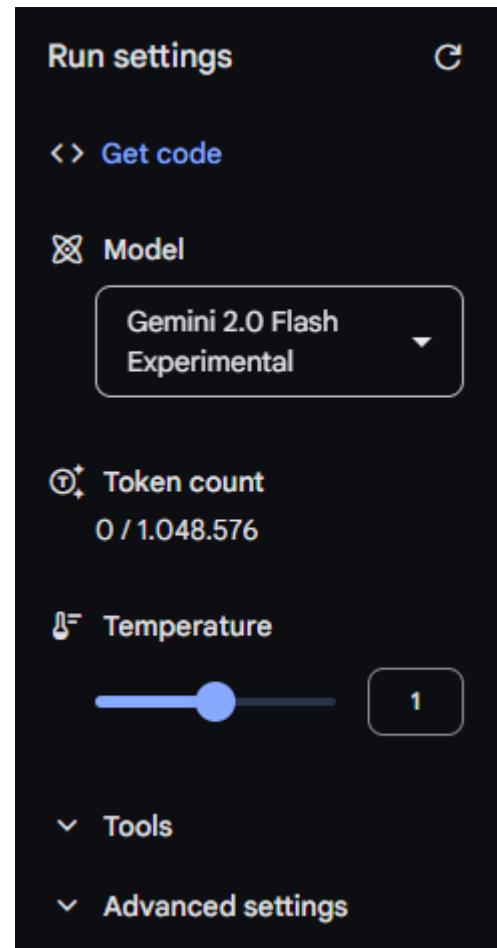
4. Abilita microfono e webcam *per questa sessione*

5. Fai vedere qualche oggetto; lo riconosce! **I2T**

6. Premi il bottone a destra «share your screen» **Share your screen**, scegli i contenuti, abilita e fatti spiegare cosa vedi **I2T**



link ad un  
video TikTok di  
demo





## Generative AI

5.507.760 follower

[Contattaci](#)

1 ora •

This isn't a car commercial shot in Italy—it's 100% AI-generated.

Using [DeepMind](#)'s Veo2 model, creator [László Gaál](#) crafted every frame from a simple text prompt. No cameras, no film crew, no location scouting.

This milestone marks a pivotal moment in creative production. The line between AI-generated and human-created content is becoming harder to distinguish.

Tools like Veo2 are reshaping advertising, storytelling, and visual communication—pushing the boundaries of what's possible.

What excites you most about the future of AI in creative industries?



## ElevenLabs Enters Speech-to-Text Race with Scribe

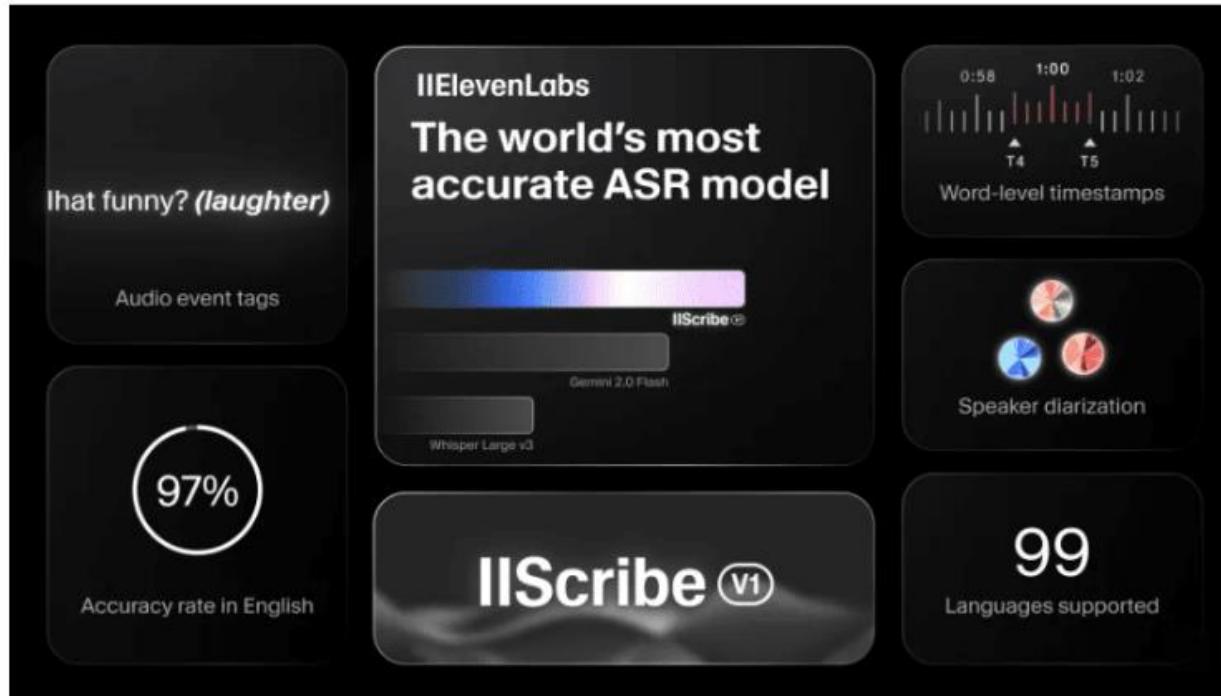


Image Source: ElevenLabs

Fresh off a **\$180M funding** round, ElevenLabs just launched **Scribe**, a record-breaking speech-to-text model that outperforms Google Gemini 2.0 Flash, OpenAI Whisper v3, and Deepgram Nova-3.

### Why this matters:

- Scribe boasts **96.7% accuracy** for English, setting a new standard in automated transcription.
- Speaker diarization, word-level timestamps, and auto-tagged sound events for enhanced content accessibility.
- **Covers 99+ languages**, with 25+ hitting sub-5% word error rates, including French, German, Hindi, Japanese, and Portuguese.
- A real-time version is in the works, which could make Scribe a serious contender for live meeting transcriptions.
- The launch coincides with **Hume AI's Octave**, a text-to-speech model that allows customizable AI-generated voices with emotional tuning.

As ElevenLabs and Hume AI go head-to-head, the AI voice & transcription space is heating up—who will dominate?

altre  
info



back to tassonomia

Avi Chawla • Già segui  
Co-founder DailyDoseofDS | IIT Varanasi | ex-AI Engineer MastercardAI |...  
4 ore •

marzo 2025

Everyone believed the hype around Cursor.

But it's not even close to what I found:

Cursor, Copilot, and Augment Code all promise AI-powered coding—but how they actually perform is a different story.

I ran a head-to-head test between Cursor and Augment Code. See the video below.

Being capped at 10-12k tokens, Cursor struggles to answer questions requiring full codebase context.

This context window is not enough for real-world projects.

And it's terribly verbose as shown in the video below!!

**Augment Code** is a powerful AI Assistant built for developers working with large, evolving codebases—needing an assistant that understands the full context of their projects.

Check it here: <https://fnf.dev/4i13YnN>.

It indexes your entire repo upfront, so it answers questions instantly, no matter how large your project is.

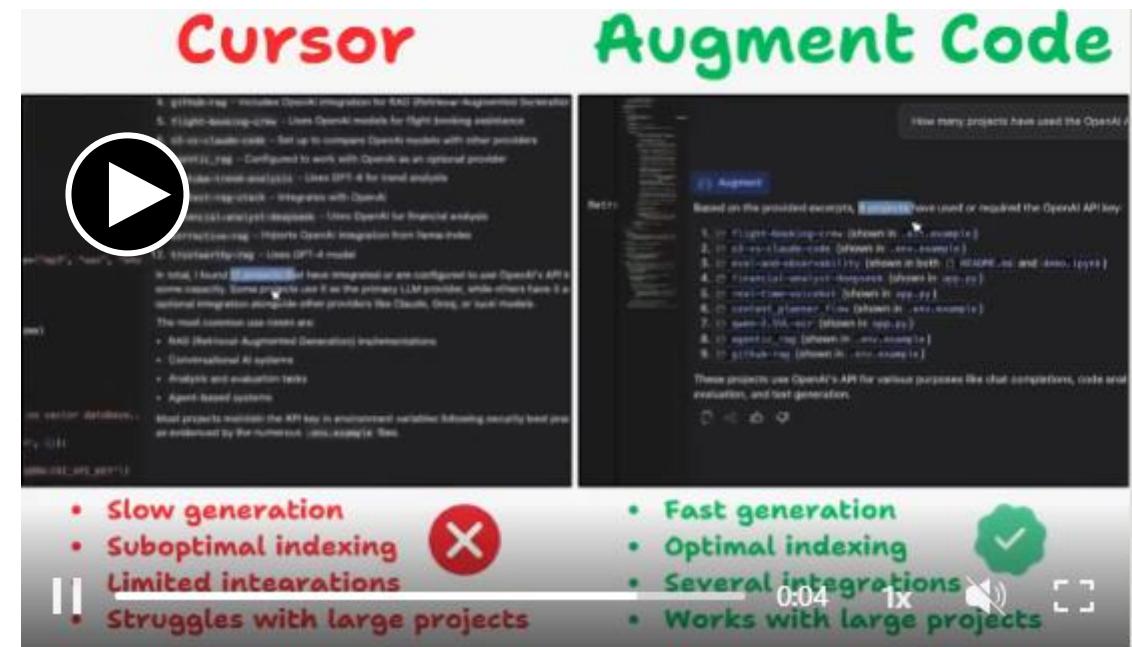
This is evident from the generation speed in the video.

If you look closely, it also takes 6-8x fewer tokens to generate the right response (more cost-saving!).

Cursor, even after generating many more tokens, produces a wrong response.

Lastly, Augment Code is fully compatible with VSCode, JetBrains, Vim, Slack, and more.

Thanks to Augment Code for showing me their powerful AI coding assistant and working with me on this post.



back to tassonomia

# State-of-the-art video and image generation with Veo 2 and Imagen 3

Dec 16, 2024  
7 min read

We're announcing new versions of Veo and Imagen, and introducing our latest experiment in image generation: Whisk.

Google updated its **image-to-video generation tools**, improving output precision, quality, and creative flexibility.

- The tools now handle complex scenes with **greater realism**, catering to content creators and marketing teams.
- Part of Google's broader effort to lead in AI-driven creativity and visual applications.

## Why It Matters:

AI-generated video tools have vast implications for industries like media, marketing, and education. As these tools become more refined, businesses can reduce production costs while boosting creative capabilities for video-based content.

link alla pagina





**Generative AI**  
5.414.221 follower  
[Contattaci](#)  
6 giorni •

gennaio 2025

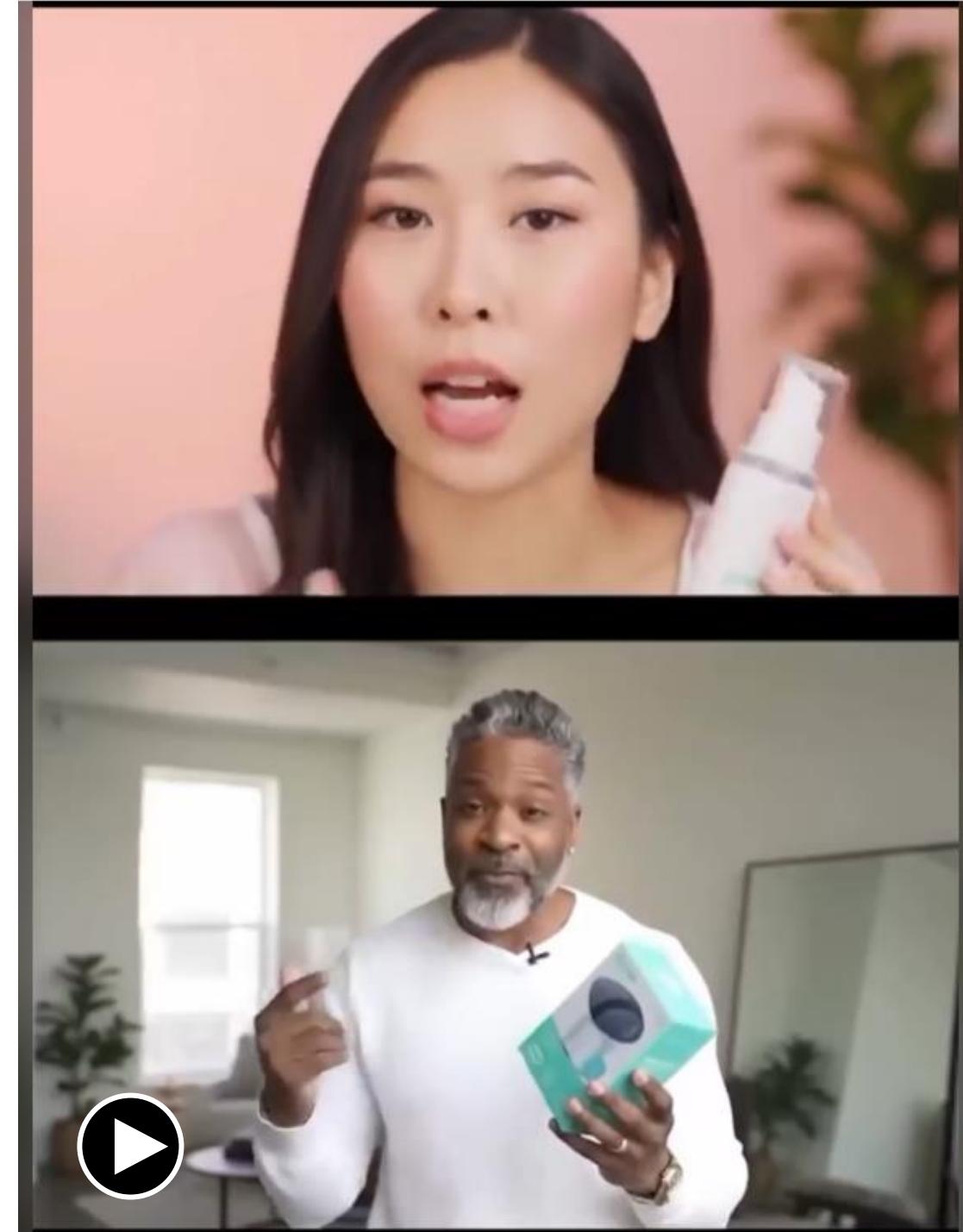
### Google's Veo2: The Future of AI-Generated Content?

The more we see videos from Google's Veo2, the more it blows our minds! 😱

This latest clip? It's an almost flawless AI-generated influencer-style video—complete with natural flow, visuals, and style.

It's a game-changer for content creation. Imagine the possibilities:

- Brands creating tailored videos at scale.
- Influencers generating unique, AI-enhanced content in minutes.
- Redefining how we think about creativity and AI.





Generative AI  
4.816.090 follower  
[Contattaci](#)  
28 minuti •

ottobre 2024

Anyone concerned about revealing their face should explore AI technology.

With AI, even an elderly man can transform into a stunning woman! What was once science fiction is now a reality!



back to tassonomia

28 gennaio 2024



Generative AI  
2.549.002 follower  
16 ore • 🔍

Quanto è folle? 🤯

Il video mostra cosa può fare l'IA generativa.



back to tassonomia



marzo 2025

Just In: Google Makes A Striking Connection Between LLMs And Human Brain Activity

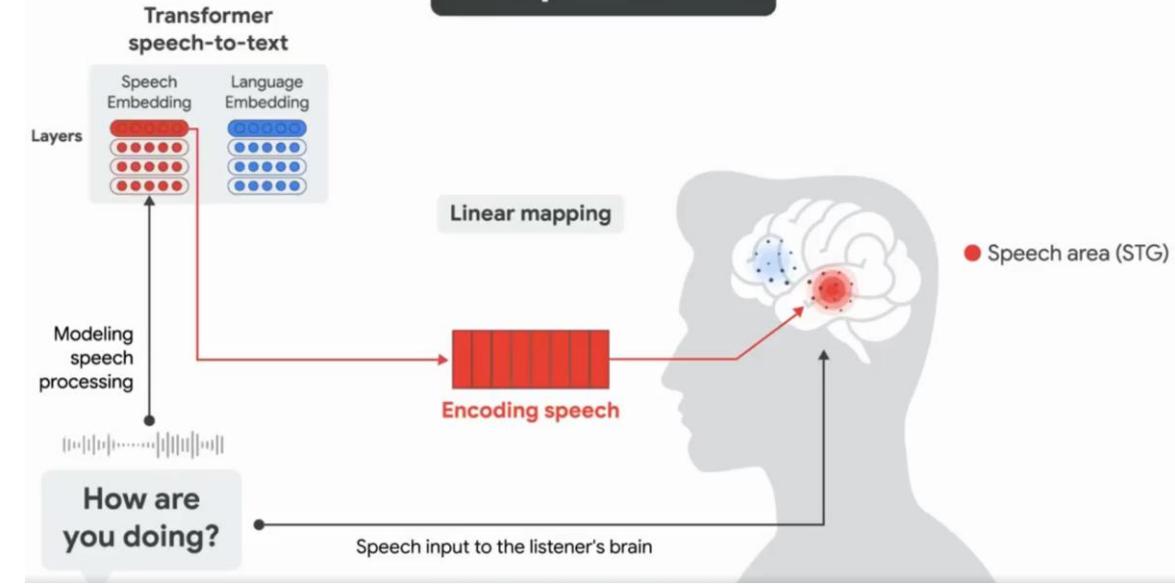
Here are the key findings:

- 1 Neural Activity Alignment: LLMs' internal representations align with neural activity in the brain's speech and language areas.
- 2 Shared Computational Principles: Both LLMs and the human brain use prediction and embedding-based representations to process language.
- 3 Differences in Architecture: Despite shared principles, the neural architectures of LLMs and the brain differ significantly.

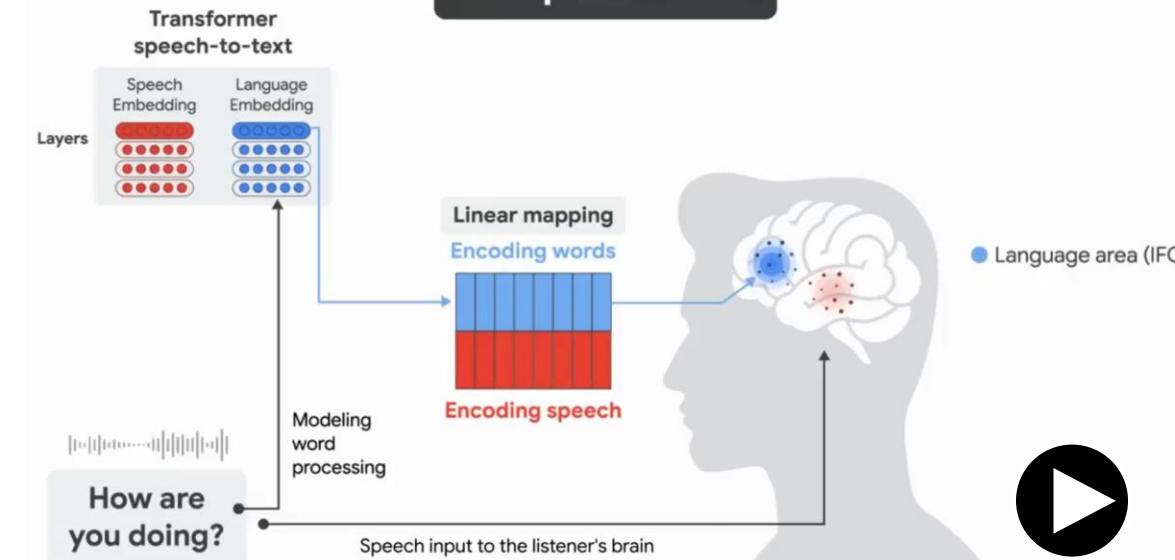
This research suggests that AI can be designed more effectively by mimicking human cognitive processes, potentially leading to more intuitive and powerful AI systems.

As we move forward, the question remains: How can we harness these insights to create AI that truly complements human intelligence?

## Comprehension



## Comprehension





## Image 2 3D

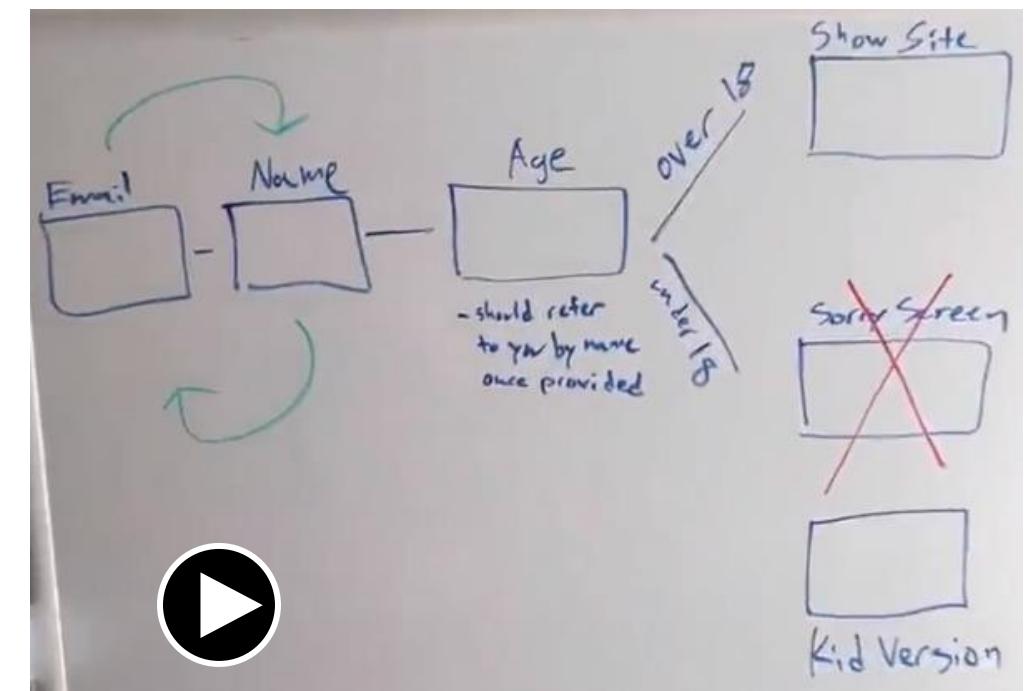
- Show di un prodotto in 3D (o anche video) che non esiste
- Inimmaginabile 2 anni fa
- Si fa spesso l'esempio di una macchina o di un paio di scarpe.



## Image 2 Code

Massimo Canducci • 1°  
Chief Innovation Officer | Executive MBA Professor | Aut...  
2 giorni • 0

ChatGPT developer  
With ChatGPT-4 it is possible to generate complete working code directly from a diagram on the board.  
This is a diagram that has corrections in it, but ChatGPT interprets them correctly and writes the necessary code as required.  
We are at the beginning of a big revolution in software writing as well, but this does not mean that programmers will lose their jobs, it means that programmers have one more tool to work better.



«whiteboard example»



# 4 ottobre 2023



Generative AI  
1.423.264 follower  
2 giorni • 🌎

Less than 3 days since ChatGPT gained vision, the discovery is blowing! 🚀🚀

Prepare to be amazed by these 10 awesome examples:

👉 From a SaaS dashboard to code:

[https://lnkd.in/d4C\\_nHhE](https://lnkd.in/d4C_nHhE)

👉 Diagram master reader:

[https://lnkd.in/d4xJ9r\\_C](https://lnkd.in/d4xJ9r_C)

👉 The future of education. Right there:

<https://lnkd.in/degbhkYv>

👉 Can you get park here?

<https://lnkd.in/dtTckF6X>

👉 Understand complex images:

<https://lnkd.in/d2gHdrZh>

👉 Yet another insane diagram to understand:

<https://lnkd.in/dUwJZMDD>

👉 You like a style? Name it:

<https://lnkd.in/dy49A4A6>

👉 From whiteboard to editable notes:

<https://lnkd.in/dVmMfdT>

👉 Why is it funny?

<https://lnkd.in/dRg4CzF9>

👉 ChatGPT knows any film. Any sentence:

<https://lnkd.in/dvJ27dCe>



It's Wednesday at 4 pm. Can I park at this spot right now? Tell me in 1 line.



Yes, you can park for up to 1 hour starting at 4 pm.





Massimo Canducci • 1°

Chief Innovation Officer | Executive MBA Professor | Aut...  
12 minuti • 

... X

## ChatGPT and the senses

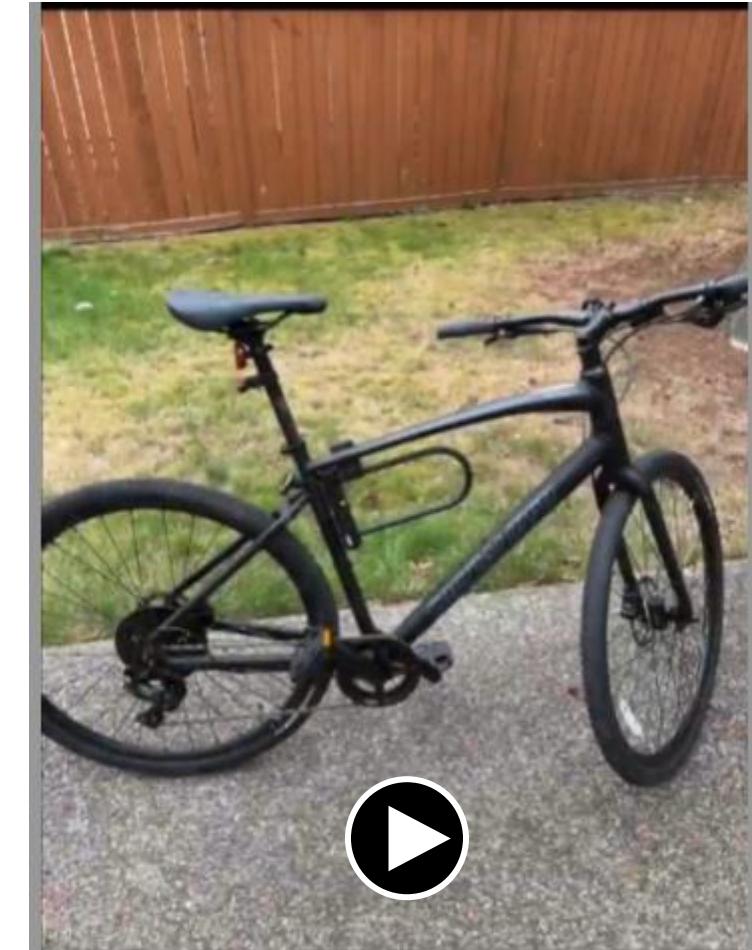
Providing senses to machines has very interesting implications.

In this case ChatGPT has the ability to interact directly with the images as input information, as if it had vision, in order to better assist us in the task at hand.

Obviously the bar of possibilities is being raised higher and higher and, in a very short time, we will arrive at the general personal assistant that many of us would like to have in our lives.

Now try to think about having this potential directly in your smart glasses, without the need for a smartphone. That will be another huge step forward.

In this field, too, we are only at the beginning.



```
1 !nvidia-smi
executed in 94ms, finished 14:52:13 2024-04-18
```

Thu Apr 18 14:52:13 2024

NVIDIA-SMI 536.23			Driver Version: 536.23		CUDA Version: 12.2		
GPU	Name	TCC/WDDM	Bus-Id	Disp.A	Volatile	Uncorr.	ECC
Fan	Temp	Perf	Pwr:Usage/Cap	Memory-Usage	GPU-Util	Compute M.	MIG M.
0	NVIDIA GeForce GT 1030	WDDM	00000000:01:00.0	On		N/A	
31%	25C	P8	N/A / 30W	930MiB / 2048MiB	0%	Default	N/A

il modello

l'informazione cercata  
(poco chiara!)

## primo prompt

You

How much video RAM has the following NVIDIA card: "NVIDIA-SMI 536.23  
536.23 CUDA Version: 12.2"

Driver Version:

domanda incompleta, non c'è il modello! (ho inserito nel prompt le specifiche tecniche incomplete: la prima riga)

Processes:							
GPU	GI	CI	PID	Type	Process name	GPU Memory Usage	
ID	ID						
0	N/A	N/A	6172	C+G	...ekyb3d8bbwe\PhoneExperienceHost.exe	N/A	
0	N/A	N/A	6600	C+G	....Search_cw5n1h2txyewy\SearchApp.exe	N/A	
0	N/A	N/A	8488	C+G	C:\Windows\explorer.exe	N/A	
0	N/A	N/A	9492	C+G	...046_x64_8wekyb3d8bbwe\ms-teams.exe	N/A	
0	N/A	N/A	9948	C+G	...ams\Cisco Spark\CiscoCollabHost.exe	N/A	
0	N/A	N/A	10032	C+G	...CBS_cw5n1h2txyewy\TextInputHost.exe	N/A	
0	N/A	N/A	10696	C+G	...oogle\Chrome\Application\chrome.exe	N/A	
0	N/A	N/A	10932	C+G	...2txyewy\StartMenuExperienceHost.exe	N/A	
0	N/A	N/A	11908	C+G	...crosoft\Edge\Application\msedge.exe	N/A	
0	N/A	N/A	12064	C+G	...al\Discord\app-1.0.9042\Discord.exe	N/A	
0	N/A	N/A	12836	C+G	...64_8wekyb3d8bbwe\CalculatorApp.exe	N/A	
0	N/A	N/A	13300	C+G	...tionsPlus\logioptionsplus_agent.exe	N/A	
0	N/A	N/A	13932	C+G	...nap Inc\Snap Camera\Snap Camera.exe	N/A	
0	N/A	N/A	14748	C+G	...aam7r\AcrobatNotificationClient.exe	N/A	
0	N/A	N/A	15016	C+G	...5n1h2txyewy\ShellExperienceHost.exe	N/A	
0	N/A	N/A	15208	C+G	...siveControlPanel\SystemSettings.exe	N/A	
0	N/A	N/A	15680	C+G	...on\123.0.2420.97\msedgewebview2.exe	N/A	
0	N/A	N/A	20092	C+G	...t Office\root\Office16\POWERPNT.EXE	N/A	
0	N/A	N/A	21436	C+G	...Desktop\app-2.1.6\GoSignDesktop.exe	N/A	

master su slide corso chatGPT





You

How much video RAM has the following NVIDIA card: "NVIDIA-SMI 536.23  
536.23 CUDA Version: 12.2"



ChatGPT

The NVIDIA System Management Interface (NVIDIA-SMI) tool information you provided indicates the software and driver version, but it does not specify which NVIDIA graphics card you are using. The version of NVIDIA-SMI (536.23) and the CUDA version (12.2) tell us about the software environment, but not about the hardware specifics like video RAM.

To find out how much video RAM (VRAM) your NVIDIA graphics card has, you would need to identify the model of the graphics card first. Once you know the model, you can easily find the VRAM details either from NVIDIA's official specifications or other reputable tech websites.

If you have access to the system where the card is installed, you can use the NVIDIA-SMI tool to check the VRAM. Here's how you can do it on a machine running Linux or Windows with command line access:

- 1. Open the Command Prompt or Terminal.**
- 2. Run the NVIDIA-SMI command:**

```
bash
Copy code
nvidia-smi
```

This command will display information about all NVIDIA GPUs in the system including the GPU name and the total amount of memory.

If you follow these steps, you should see the amount of total available GPU memory listed under the "Memory-Usage" or similar heading. If you need help executing this command or interpreting the output, feel free to ask!



You  
secondo prompt

```
1 !nvidia-smi
executed in 94ms, finished 14:52:13 2024-04-18
Thu Apr 18 14:52:13 2024
+-----+
| NVIDIA-SMI 536.23 | Driver Version: 536.23 | CUDA Version: 12.2 |
+-----+
| GPU Name TCC/WDDM Bus-Id Disp.A Volatile Uncorr. ECC |
| Fan Temp Perf Pwr:Usage/Cap | Memory-Usage GPU-Util Compute M. |
| | | | | | MIG M. |
+-----+
| 0 NVIDIA GeForce GT 1030 WDDM 00000000:01:00.0 On N/A |
| 31% 25C P8 N/A / 30W 930MiB / 2048MiB 0% Default N/A |
+-----+
+-----+
| Processes: GPU GI CI PID Type Process name GPU Memory |
| ID ID ID | Usage |
+-----+
| 0 N/A N/A 6172 C+G ...ekybd8bbwe\PhoneExperienceHost.exe N/A |
| 0 N/A N/A 6600 C+G ...Search_cw5n1h2txyewy\SearchApp.exe N/A |
| 0 N/A N/A 8488 C+G C:\Windows\explorer.exe N/A |
| 0 N/A N/A 9492 C+G ...046_x64_8wekybd8bbwe\ms-teams.exe N/A |
| 0 N/A N/A 9948 C+G ...ams\Cisco Spark\CiscoCollabHost.exe N/A |
| 0 N/A N/A 10032 C+G ...CBS_cw5n1h2txyewy\TextInputHost.exe N/A |
| 0 N/A N/A 10696 C+G ...oogle\Chrome\Application\chrome.exe N/A |
| 0 N/A N/A 10932 C+G ...2txyewy\StartMenuExperienceHost.exe N/A |
| 0 N/A N/A 11908 C+G ...crosoft\Edge\Application\msedge.exe N/A |
| 0 N/A N/A 12064 C+G ...al\Discord\app-1.0.9042\Discord.exe N/A |
| 0 N/A N/A 12836 C+G ...64_8wekybd8bbwe\CalculatorApp.exe N/A |
| 0 N/A N/A 13300 C+G ...tionsPlus\logioptionsplus_agent.exe N/A |
| 0 N/A N/A 13932 C+G ...nap Inc\Snap Camera\Snap Camera.exe N/A |
| 0 N/A N/A 14748 C+G ...aam7r\AcrobatNotificationClient.exe N/A |
| 0 N/A N/A 15016 C+G ...5n1h2txyewy\ShellExperienceHost.exe N/A |
| 0 N/A N/A 15208 C+G ...siveControlPanel\SystemSettings.exe N/A |
| 0 N/A N/A 15680 C+G ...on\123.0.2420.97\msedgewebview2.exe N/A |
| 0 N/A N/A 20092 C+G ...t Office\root\Office16\POWERPNT.EXE N/A |
| 0 N/A N/A 21436 C+G ...Desktop\app-2.1.6\GoSignDesktop.exe N/A |
+-----+
```

Here's the output to the command you said (nvidia-smi). Can you tell me the memory?



ChatGPT

From the `nvidia-smi` output you've provided, the NVIDIA GeForce GT 1030 graphics card has 2048 MB of video RAM (VRAM). This information is shown under the "Memory-Usage" section, where it indicates that the total available memory is 2048 MiB.

NVIDIA-SMI 536.23			Driver Version: 536.23			CUDA Version: 12.2		
GPU	Name	TCC/WDDM	Bus-Id	Disp.A	Volatile	Uncorr.	ECC	
	Fan	Temp	Perf	Pwr:Usage/Cap	Memory-Usage	GPU-Util	Compute M.	MIG M.
0	NVIDIA GeForce GT 1030	WDDM	00000000:01:00.0	On				
31%	25C	P8	N/A / 30W	930MiB / 2048MiB	0%	Default	N/A	

Processes:								
GPU	GI	CI	PID	Type	Process name	GPU	Memory	Usage
ID	ID							
0	N/A	N/A	6172	C+G	...ekyb3d8bbwe\PhoneExperienceHost.exe	N/A		
0	N/A	N/A	6600	C+G	....Search_cw5n1h2txyewy\SearchApp.exe	N/A		
0	N/A	N/A	8488	C+G	C:\Windows\explorer.exe	N/A		
0	N/A	N/A	9492	C+G	...046_x64_8wekyb3d8bbwe\ms-teams.exe	N/A		
0	N/A	N/A	9948	C+G	...ams\Cisco Spark\CiscoCollabHost.exe	N/A		
0	N/A	N/A	10032	C+G	...CBS_cw5n1h2txyewy\TextInputHost.exe	N/A		
0	N/A	N/A	10696	C+G	...oogle\Chrome\Application\chrome.exe	N/A		
0	N/A	N/A	10932	C+G	...2txyewy\StartMenuExperienceHost.exe	N/A		
0	N/A	N/A	11908	C+G	...crosoft\Edge\Application\msedge.exe	N/A		
0	N/A	N/A	12064	C+G	...al\Discord\app-1.0.9042\Discord.exe	N/A		
0	N/A	N/A	12836	C+G	...64_8wekyb3d8bbwe\CalculatorApp.exe	N/A		
0	N/A	N/A	13300	C+G	...tionsPlus\logioptionsplus_agent.exe	N/A		
0	N/A	N/A	13932	C+G	...nap Inc\Snap Camera\Snap Camera.exe	N/A		
0	N/A	N/A	14748	C+G	...aam7r\AcrobatNotificationClient.exe	N/A		
0	N/A	N/A	15016	C+G	...5n1h2txyewy\ShellExperienceHost.exe	N/A		
0	N/A	N/A	15208	C+G	...siveControlPanel\SystemSettings.exe	N/A		
0	N/A	N/A	15680	C+G	...on\123.0.2420.97\msedgewebview2.exe	N/A		
0	N/A	N/A	20092	C+G	...t Office\root\Office16\POWERPNT.EXE	N/A		
0	N/A	N/A	21436	C+G	...Desktop\app-2.1.6\GoSignDesktop.exe	N/A		

chatGPT ha saputo estrarre l'informazione da qui

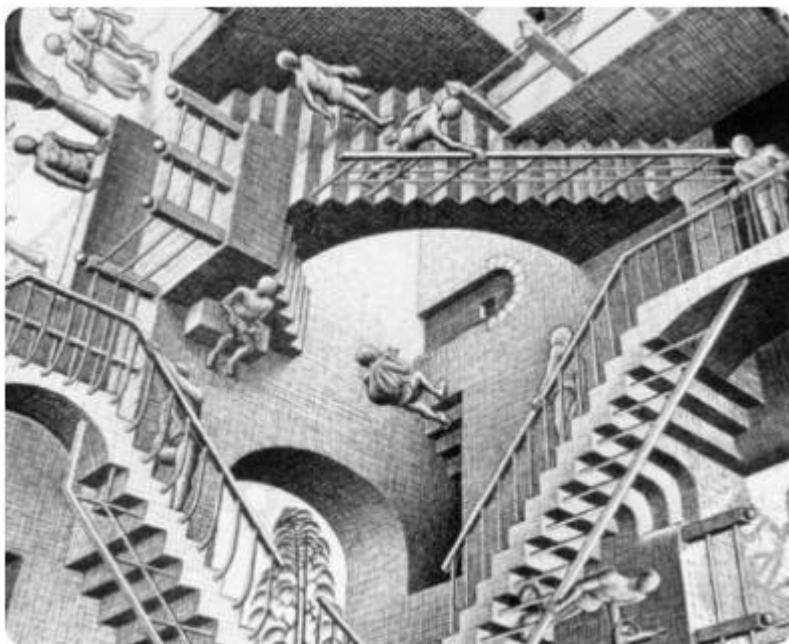
La verifica della correttezza della risposta su un sito specializzato

Data di uscita: 2017  
NVIDIA GeForce GT 1030



La scheda GeForce GT 1030 di NVIDIA è stata introdotta sul mercato nel 2017, usa una GPU GP108 con 1,8 miliardi di transistor ed è ovviamente basata su architettura Pascal. Mette a disposizione un totale di 384 CUDA core, abbinate a 24 TMU e 16 ROPs. La scheda è dotata di 2GB di memoria GDDR5, su bus a 64 bit. La memoria opera a 6000MHz.





di chi è questo quadro e cosa rappresenta?

Il quadro che hai allegato è "Relativity" di Maurits Cornelis Escher, artista olandese noto per le sue incisioni, litografie e mezzetinte che giocano con la prospettiva, la simmetria e le geometrie impossibili.

### Descrizione dell'opera:

- **Titolo:** *Relativity* (1953)
- **Tecnica:** Litografia
- **Stile:** Arte concettuale / geometria impossibile

### Cosa rappresenta?

L'opera raffigura un mondo paradossale dove le leggi della **gravità** non seguono le regole a noi familiari:

- Ci sono tre diversi punti di gravità all'interno della stessa scena.
- Le scale vanno in tutte le direzioni, alcune "in su" per un personaggio sono "in giù" per un altro.
- Gli ambienti si intersecano e i personaggi camminano o siedono su piani inclinati in modo diverso rispetto allo spettatore.

### Significato:

L'opera può essere letta come una riflessione:

- sull'**ambiguità della percezione visiva**,
- sulla **relatività del punto di vista** (da cui il titolo),
- oppure come una rappresentazione surreale dell'**assurdità o complessità della realtà**.





dimmi cosa è

Quello che hai fotografato è un'interfaccia audio Focusrite Scarlett, un dispositivo molto usato per registrazioni audio professionali o semi-professionali. Più precisamente, sembra uno dei modelli della serie **Scarlett Solo o 2i2** (non si vede il modello esatto da questa angolazione).

## A cosa serve

Serve per collegare strumenti musicali, microfoni e cuffie al computer, offrendo una qualità audio molto superiore rispetto alle schede audio integrate nei PC.

## Componenti visibili nella foto:

- **Porta USB:** per il collegamento al computer, serve anche per l'alimentazione.
- **Slot per il lucchetto Kensington:** per la sicurezza fisica del dispositivo.
- **Output L/R:** uscite audio per collegare monitor da studio o casse.
- **Input 2 (XLR/TRS combo):** per collegare un microfono o uno strumento (come una chitarra elettrica o tastiera).

## Come si usa

1. Collega via USB al computer.
2. Installa i driver dal sito Focusrite (se richiesto).
3. Collega microfoni o strumenti agli ingressi frontali.
4. Registra o ascolta tramite il tuo software DAW preferito (Ableton, Logic, Cubase, ecc.).
5. Collega le casse monitor alle uscite R e L posteriori.



back to tassonomia

## C:\ Prompt dei comandi - python

```
Microsoft Windows [Versione 10.0.19045.5965]
(c) Microsoft Corporation. Tutti i diritti sono riservati.

C:\Users\Utente>python
Python 3.12.3 | packaged by conda-forge | (main, Apr 15 2024, 18:20:11) [MSC v.1938 64 bit (AMD64)] on win32
Type "help", "copyright", "credits" or "license" for more information.
>>>
```

ho eseguito questo comando da un terminale dos; spiegami l'output

### Cosa mostra l'output

pgsql

 Copy  Edit

Python 3.12.3 | packaged by conda-forge | (main, Apr 15 2024, 18:20:11) [MSC v.1938 64 bit (AMD64)]

Significato riga per riga:

Python 3.12.3

Hai installato Python versione 3.12.3, una delle versioni più recenti attualmente disponibili.

packaged by conda-forge

Python è stato installato tramite **conda-forge**, cioè un canale della community per Anaconda/Miniconda. Quindi molto probabilmente stai usando **Anaconda** o **Miniconda** per la gestione degli ambienti Python.

(main, Apr 15 2024, 18:20:11)

Data e ora di compilazione dell'interprete Python.

[MSC v.1938 64 bit (AMD64)]

- **MSC v.1938**: il compilatore Microsoft (Microsoft C) usato per compilare Python.
- **64 bit (AMD64)**: architettura della build. Anche se c'è scritto **win32**, il Python è a 64 bit, come confermato da "64 bit (AMD64)".

on win32

Indica il sistema operativo in uso: **Windows**. Il termine **win32** qui è "storico" e viene usato anche per build a 64 bit su Windows.



back to tassonomia



devo fare una fototerapia di mezzora tutte le mattine con questa macchina che mi ha fornito l'ospedale. purtroppo non c'è il manuale delle istruzioni. mi puoi spiegare come funziona?

Ho capito 

Dalla foto vedo che si tratta di una **lampada per fototerapia (light therapy lamp)**, usata spesso per trattare disturbi come la depressione stagionale (SAD) o problemi di ritmo circadiano.

Sulla parte bassa ci sono tre pulsanti principali:

1. **A sinistra** → simbolo di **accensione/spegnimento** 
  - Serve per accendere e spegnere la lampada.
2. **Al centro** → simbolo di una **lampadina** 
  - Regola l'intensità/luminosità della luce (di solito ci sono 3 livelli: basso, medio, alto).
3. **A destra** → simbolo di un **orologio** 
  - È un timer: puoi impostare la durata della sessione (tipicamente 10, 20 o 30 minuti).
  - Quando il tempo è finito, la lampada si spegne da sola.

**Come usarla (indicazioni generali):**

- **Distanza:** siediti a circa **30–50 cm** dalla lampada.
- **Angolazione:** non devi guardarla direttamente negli occhi; basta che la luce arrivi al tuo campo visivo laterale mentre leggi, lavori al PC, o fai colazione.
- **Durata:** nel tuo caso l'ospedale ti ha prescritto **30 minuti ogni mattina**.
- **Orario migliore:** entro 1 ora dal risveglio, per aiutare a regolare il ritmo circadiano.
- **Costanza:** funziona meglio se usata **tutti i giorni alla stessa ora**.

 Nota:

- Se senti fastidio agli occhi, mal di testa o insonnia, va riferito al medico.
- Non usarla la sera, perché potrebbe disturbare il sonno.
- Segui sempre le indicazioni mediche: la potenza luminosa è in genere 10.000 lux, ma la distanza e la durata possono cambiare a seconda del modello  delle tue condizioni cliniche.



29 maggio 2023

Machine Learning Community (Moderated)  
Billy Gareth • 3° e oltre  
4 ore • ⓘ

👉 Write a PowerPoint Presentation with ChatGPT 🚀  
👉 <https://lnkd.in/gM65CNAr> - ChatGPT Mastery - Prompt 👉 👉

...vedi altro

Vedi traduzione

The Dawn of AI  
The concept of 'Artificial Intelligence' was first introduced by the British mathematician Alan Turing. The birth of AI is often dated to a conference at Dartmouth.

0:50

U back to tassonomia

# Dazi, l'America torna a lavorare in fabbrica: la satira cinese a colpi di intelligenza artificiale



TikTok  
@axiang67  
CORRIERETV

axiang67 ·  
Make america great again

Nel video, viene mostrata un'America decadente e triste: uomini e donne americani, rappresentati in modo caricaturale come obesi e trasandati, lavorano in fabbriche e officine in condizioni faticose, con una musica tradizionale cinese in sottofondo che accentua il tono grottesco e beffardo.

Tutti i personaggi appaiono esausti e oppressi dal lavoro, un chiaro tentativo di rafforzare stereotipi negativi sull'etica del lavoro americana. Il video si conclude con un'inquadratura sullo slogan della campagna di Trump: *Make America Great Again*, usato in tono ironico e provocatorio.

Un video generato con l'intelligenza artificiale, apparentemente creato da un utente cinese su TikTok, sta facendo discutere per il modo in cui ridicolizza gli Usa a seguito dei dazi imposti da Trump





Akshay Pachaar • Già segui

Co-Founder DailyDoseOfDS | BITS Pilani | 3 Patents | X (187K+)  
1 giorno • 5

Finally! A Text-to-SQL tool that actually works!

(100% open-source, 14k+ stars)

The screenshot shows a Jupyter Notebook interface. On the left is a sidebar with three items: 'Install Vanna' (indicated by a right-pointing arrow), 'Import Vanna' (indicated by a downward-pointing arrow), and 'Log In to Vanna'. The main area displays a code cell with the following content:

```
import vanna
```

Below the code cell, there is descriptive text: 'Vanna provides a function to get an API key. You'll get a code sent to your e-mail. You can get answers from your database faster than ever before'.

aprile 2025

Vanna is an open-source RAG framework for complex Text-to-SQL generation, designed for handling dynamic datasets.

Works in 2 easy steps:

- 1 Train a RAG "model" on your data.
- 2 Ask questions in natural language which will return SQL queries that can be set up to automatically run on your database.

Key features:

- ⌚ High accuracy on complex datasets
- 🤖 Self-learning: improves with each query
- 🔒 Secure: data never leaves your environment
- 🌐 Connect to any SQL DB (Snowflake, Redshift, etc.)
- ✳️ Multiple front-end integrations (Jupyter, Slack, etc.)



# WordPress: con l'intelligenza artificiale ora si può creare un sito web da zero in pochi minuti e chiunque può riuscirci



Il generatore elabora layout, immagini e testi a partire da semplici istruzioni, con l'obiettivo di semplificare l'avvio di nuovi progetti digitali. Attualmente disponibile in accesso anticipato, il tool supporta solo nuove installazioni WordPress.

di Rosario Grasso pubblicata il 10 Aprile 2025, alle 16:42 nel canale [WEB](#)

### Digitare una descrizione e ottenere un sito completo, con immagini, testi e sezioni già pronte.

È ora realtà con il nuovo [AI Site Builder](#) di [WordPress](#), disponibile da oggi in accesso anticipato. Il sistema si basa su un'**interazione diretta con un chatbot**, al quale l'utente può fornire informazioni dettagliate per avviare la creazione di un sito personale o aziendale.

Il builder consente di definire ogni elemento fondamentale semplicemente tramite input testuali al chatbot: **header, palette cromatica, tono dei testi, tipo di contenuti e organizzazione delle pagine**. Più dettagliata è la descrizione iniziale, più precisa sarà la proposta del sistema, che genera anche immagini e contenuti testuali completamente inediti. **Il chatbot resta disponibile per ogni modifica successiva**, incluse le variazioni di stile, struttura e materiali visivi.

link al  
sito



Un sito web è unico,  
proprio come lo  
sei tu.

Per accedere al servizio **è necessario registrare un account WordPress**. La fase di prova iniziale è gratuita e non richiede una carta di credito, ma per pubblicare il sito serve un piano di hosting WordPress, con prezzi a partire da **8 euro al mese** (con sconti disponibili su abbonamenti annuali o pluriennali).

**Lo strumento è attualmente compatibile solo con nuove istanze WordPress**. Non è previsto, per ora, l'uso su siti già esistenti, né la possibilità di gestire configurazioni più complesse come piattaforme e-commerce o con integrazioni avanzate. WordPress ha comunque confermato l'intenzione di **ampliare in futuro le capacità del tool**.

Una volta completata la prima versione del sito, l'utente può decidere se trasferirla direttamente su WordPress per procedere con ulteriori modifiche manuali oppure continuare a interagire con il chatbot per rifinire l'aspetto e i contenuti. Anche l'associazione a un dominio immediata, in modo tale da democratizzare al massimo la creazione dei siti web, **ora realmente alla portata di chiunque e con pochissimo lavoro necessario**.



TOGETHER WITH ASSEMBLYAI

## Analyze audio sentiment in just 5 lines of code!

[AssemblyAI's audio intelligence](#) transcribes audio and analyzes sentiment, turning speech into actionable insights.

Here's how it works:

The screenshot shows a terminal window with the following Python code:

```
import assemblyai as aai

aai.settings.api_key = "<YOUR_API_KEY>"

audio_file = "https://assembly.ai/wildfires.mp3"

config = aai.TranscriptionConfig(sentiment_analysis=True)

transcript = aai.Transcriber().transcribe(audio_file, config)

for sentiment_result in transcript.sentiment_analysis:
    print(sentiment_result.text)
    print(sentiment_result.sentiment)
    print(sentiment_result.confidence)
    print(f"Timestamp: {sentiment_result.start} - {sentiment_result.end}")
```

A callout bubble from the text `sentiment_analysis=True` points to the text "Enable Sentiment Analysis".

Below the code, there is a flow diagram with four stages: "Audio file" (represented by a sound波 icon), "Transcript" (represented by a document icon), "Sentiment Analysis using AssemblyAI" (represented by a gauge with colored segments), and "Result" (represented by a dashboard icon).

- Provide an audio file URL.
- Transcribe the audio using [AssemblyAI's speech-to-text API](#) while enabling sentiment analysis.
- Display sentiment, confidence, and timestamps for each text segment.

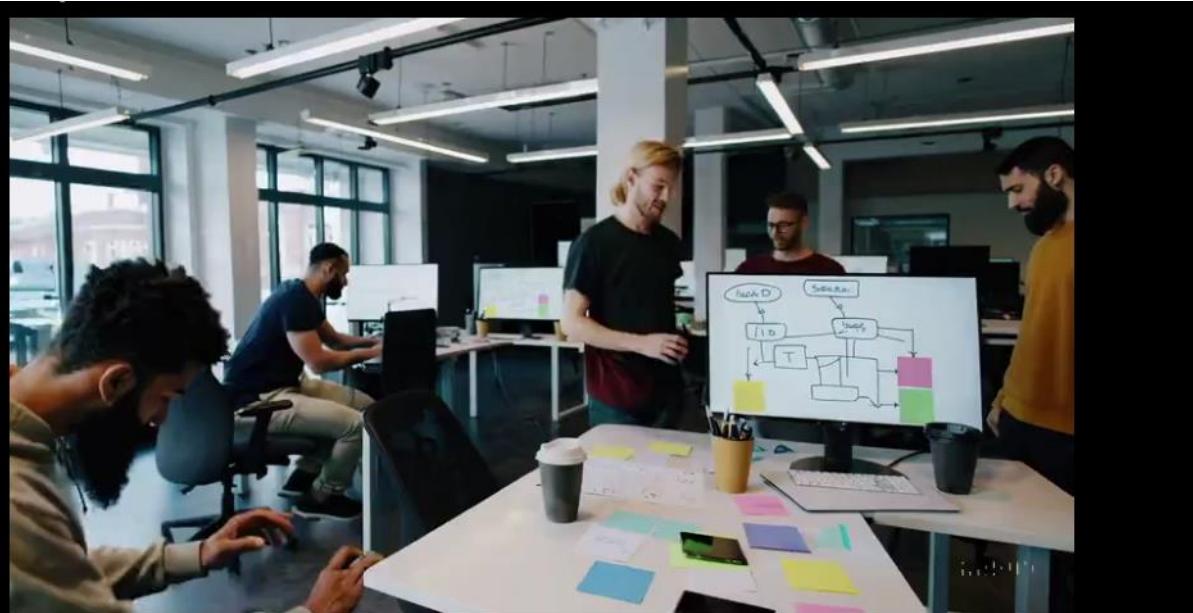
prova



giugno 2025

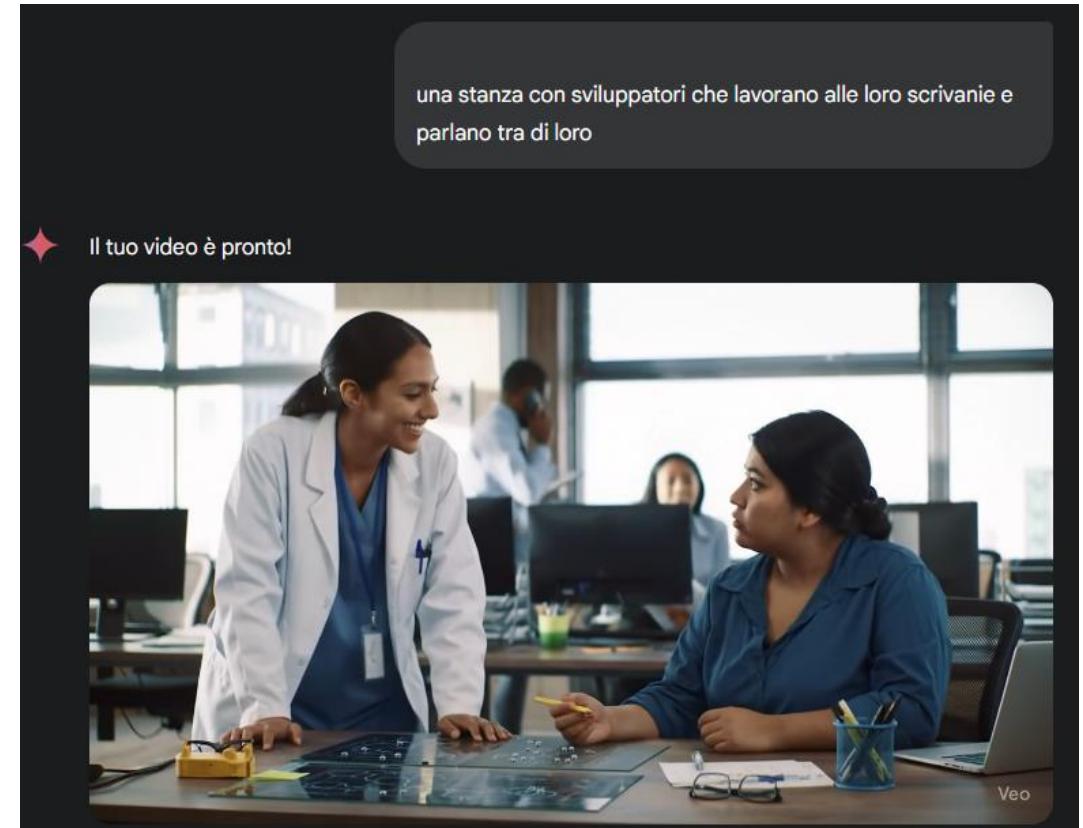
Poche settimane fa era stato il turno di **Google** di raccontare le sue novità alla sua community di sviluppatori. L'evento si chiama **I/O** e una delle funzionalità in arrivo più interessanti è la **traduzione simultanea su Google Meet**. Significa, in pratica, poter parlare con qualunque interlocutore che comunichi in qualunque lingua del mondo. Perché la piattaforma traduce, grazie all'intelligenza artificiale, ciò che stiamo dicendo noi nella sua lingua e - viceversa, ciò che sta dicendo lui nella nostra. Alice Scaglioni ha **provato in anteprima** la nuova funzionalità e potete vedere com'è andata [in questo video](#) o leggere le sue considerazioni [in questo articolo](#). Un passaggio: "Innanzitutto possiamo dire che la traduzione da inglese a italiano e viceversa funziona bene: nelle conversazioni che abbiamo intrattenuto non sono stati segnalati errori nella traduzione, **la conversazione è stata tutto sommato fluida** e il tempo di latenza nelle risposte dato dall'attesa dalla traduzione completa è davvero minimo. Un dettaglio rivoluzionario - e che lascia senza fiato appena udiamo le prime parole generate dalla traduzione - è che **la voce che sentiamo tradotta** nella lingua da noi scelta (nel nostro caso inglese, e poi italiano) **è praticamente la stessa della persona che parla**, solo tradotta. Quello che cambia è unicamente la lingua: tono, intensità, emozione nella voce sono le medesime".





Dynamic Developer Collaboration Prompt una stanza con sviluppatori che lavorano alle loro scrivanie e parlano tra di loro 480p 10s 10:25am

**sora**



**veo3**



genera un video di un signore di mezza età che percorre una strada in montagna

Il tuo video è pronto!



veo3

sora



Serene Mountain Walk Prompt un signore di mezza età che percorre una strada in montagna 720p 5s 3:56pm



back to tassonomia



1960s Chess Victory Prompt genera un breve video di una ragazza che gioca a scacchi in un albergo degli anni 60 480p 5s 9:57am

sora

genera un video con una ragazza che gioca un torneo di scacchi in un albergo negli anni '60

Il tuo video è pronto!

A screenshot of a video player showing a woman in a floral dress playing chess. The video player interface includes a progress bar and a play button.

veo3





Epic Tennis Showdown Prompt: genera un video di due tennisti che giocano una partita di tennis nel campo centrale del roland garros 480p 10s 10:23am

*sora*



*veo3*





una coppia di persone anziane che cammina mano nella mano  
su una spiaggia, in lontananza si vedono delle colline



<https://www.synthesia.io/>

## AI video assistant

Turn any source material into a video in minutes

File

URL

Prompt

Script

Describe il fenomeno dei cambiamenti climatici, le sue cause e le sue conseguenze.

### Template



Elegant Personal Avatar Training

Landscape

### Length

5 minutes

### Objective

Describe il fenomeno dei cambiamenti climatici, le sue cause e le sue conseguenze. In modo oggettivo, basato su studi ed evidenze scientifiche maggioritarie, senza pregiudizi ideologici.

### Audience

Studenti della scuola media superiore

### Language

Italian

### Speaker

Un ricercatore universitario che si occupa di climatologia da

### Tone

Informativo



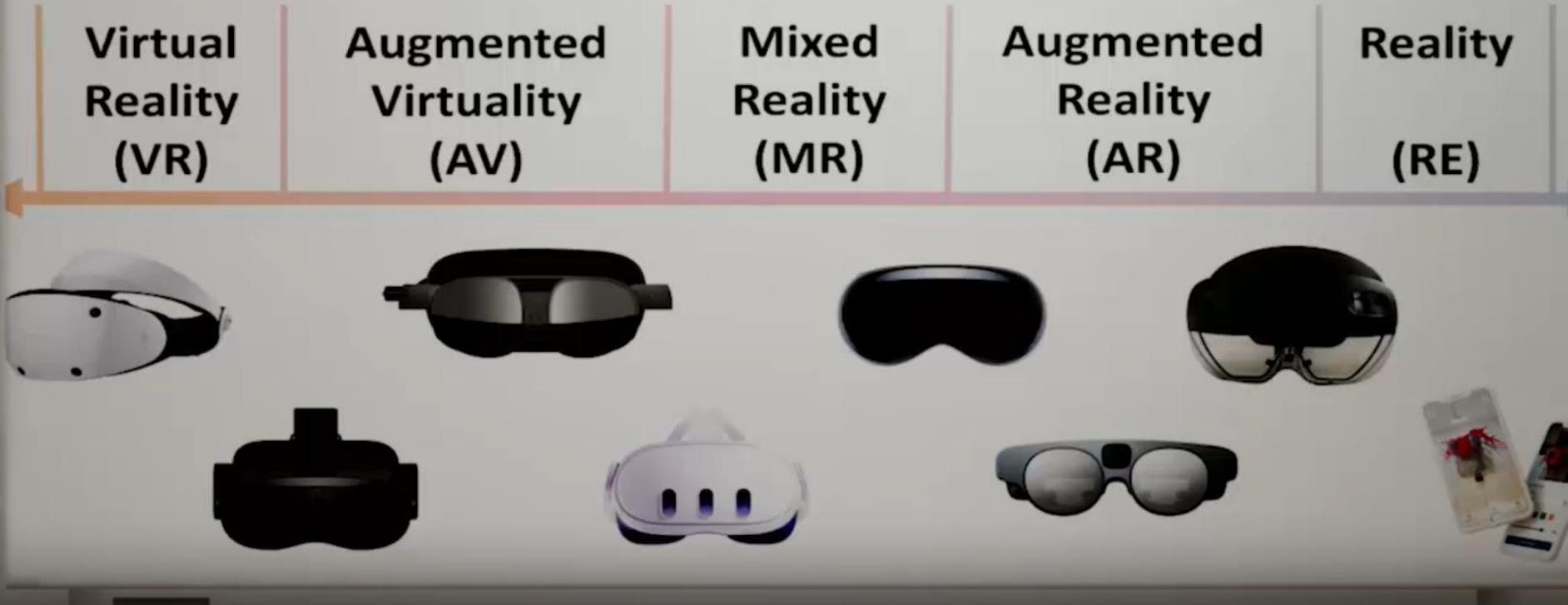
Create outline



back to tassonomia

# Spatial Computing

Spatial Computing | Extended Reality (XR) Immersive Technologies (IT)



## Tool AI specifici per il codice



back to tassonomia

## 👗 Google's New App Lets You Try On Any Outfit With AI

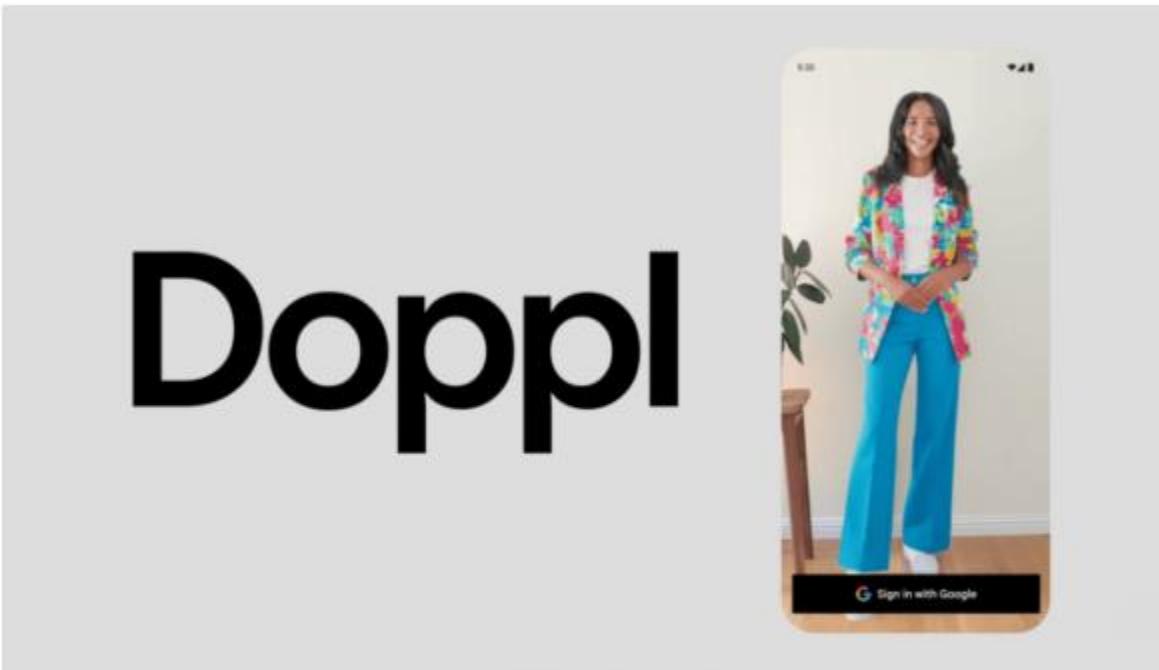


Image Credits: Google

Google just launched **Doppl**, an experimental app that turns a photo of you into a virtual mannequin so you can preview how any outfit might look on your own body. Available now on **iOS and Android in the US**, Doppl turns screenshots or street photos of clothing into realistic try-on experiences, complete with AI-generated video.

### What's new:

- **Upload yourself:** Users snap a full-body photo and Doppl builds a 3D version of them to try on outfits
- **Try anything, from anywhere:** Use a thrift store photo, a social media screenshot, or a friend's look — no purchase needed
- **Watch it in motion:** Doppl generates short clips so you can see how the clothes move on your body
- **Save and share:** Users can bookmark outfits or send their virtual looks to friends
- **Built on Google Shopping tech:** Doppl moves the try-on feature out of search and into a fun standalone app

Unlike Google's previous try-ons that used preset models, Doppl puts **you** at the center while giving Google a fresh stream of body data in return. It's an early Labs experiment, so expect some clunky fits. But the direction is clear: Google wants to be the mirror you scroll through before you dress.

link  
all'articolo  
con video



I **Code Assistant CLI** (Command Line Interface) sono strumenti da riga di comando che sfruttano l'intelligenza artificiale per assisterti nello sviluppo software **direttamente dal terminale**. In pratica, invece di usare un'interfaccia grafica (come Copilot in VS Code), li interroghi scrivendo comandi nel terminale, e loro rispondono con codice, suggerimenti, o modifiche.

## OpenAI Codex CLI (non ufficiale)

### Cosa fanno

A seconda dello strumento, possono:

- Generare codice a partire da una descrizione in linguaggio naturale
- Spiegare un pezzo di codice o una funzione
- Tradurre codice da un linguaggio a un altro
- Scrivere test unitari per una funzione
- Eseguire refactoring o debugging automatico
- Interagire con il contenuto del tuo progetto locale

## Claude Code – installato sul mio PC

Install [Node.js 18+](#), then run:

```
npm install -g @anthropic-ai/claude-code
```

## Gemini CLI



the task I gave it and it is  
using the particular Codex Cle which

# OpenAI Codex CLI Coding Agent



@cristiano\_rizzo98

Iscriviti

Code

OpenAI è un coding assistant incredibile! #codexcli  
#openaiagent

**Claude Code** è considerato da molti come il miglior strumento per la programmazione oggi disponibile. Non solo è capace di creare web app da zero, ma si dimostra estremamente efficiente anche in compiti complessi legati all'infrastruttura. In ambito tech viene spesso citato come il cuore del cosiddetto "Vibe coding", una filosofia di sviluppo che unisce potenza tecnica e accessibilità creativa." (Vedi [qui](#)).

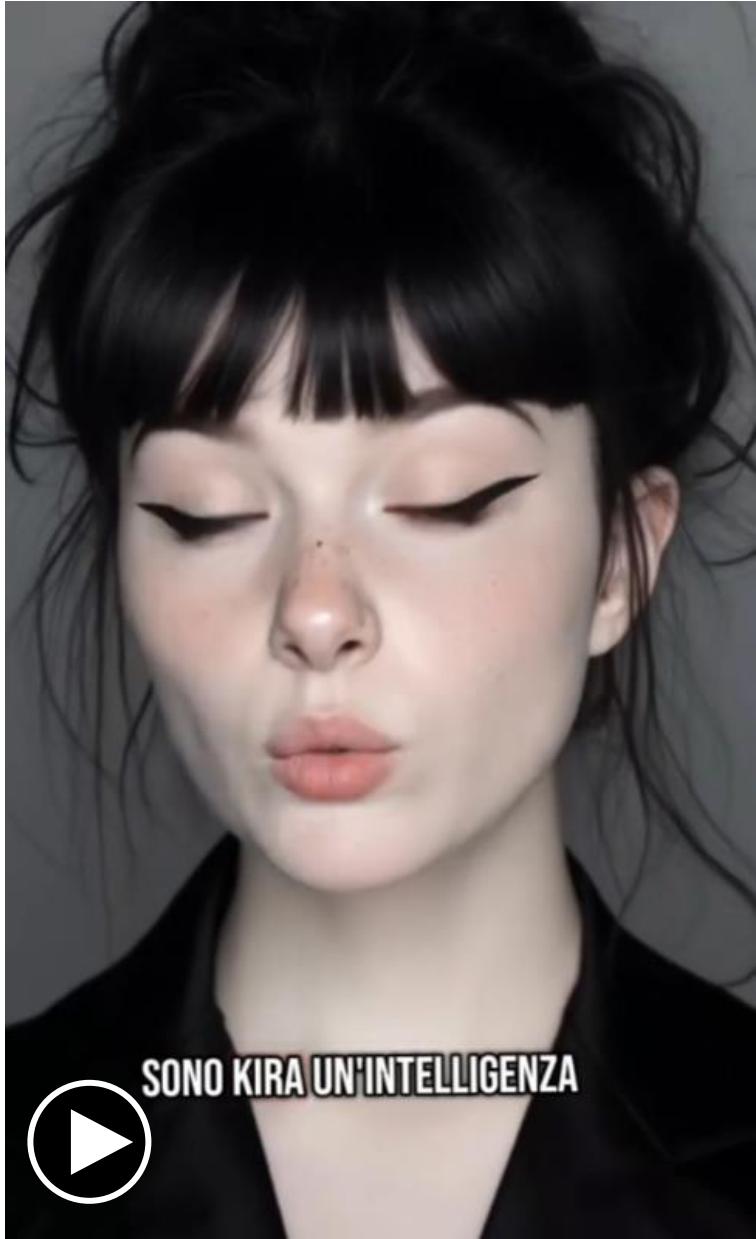




k\_i\_r\_a\_art

Kira

luglio 2025



Lip sync perfetto

Link al Kira  
Universe di  
TikTok



back to tassonomia

**Sky riporta in vita JFK grazie all'AI: il deepfake dell'ex presidente Usa parla di Putin e Ucraina**



Purtroppo



back to tassonomia



## Veo 3 Prompt Director

By Claudio Zagarini &

Writes visually rich, cinematic Veo prompts

GPT sulla creazione T2V:

- «[veo 3 prompt director](#)» di Claudio Zagarini



# State-of-the-art video and image generation with Veo 2 and Imagen 3

Dec 16, 2024

7 min read

We're announcing new versions of Veo and Imagen, and introducing our latest experiment in image generation:  
Whisk.



Aäron van den Oord

Research Scientist, Google DeepMind



Elias Roman

Senior Director, Product Management, Google Labs



Read AI-generated summary



Share



## Matrix-Game 2.0: The New Open-Source Alternative to Genie

3



Image Credit: Skywork AI

Skywork AI has launched **Matrix-Game 2.0**, the first open-source interactive world model capable of **real-time, minutes-long, high-frame-rate generation**. The release lands just a week after DeepMind's closed Genie 3 update, offering the AI community a transparent alternative for building interactive virtual worlds.

Here's what's inside the upgrade:

- **Fully open-source:** Code, models, and documentation available on GitHub, Hugging Face, and the project's homepage for unrestricted use.
- **High-speed rendering:** Maintains stable **25 FPS** for minute-long sequences, with smooth, lifelike motion and responsive camera control.
- **Cross-scenario adaptability:** Handles environments from cities to wilderness, and visual styles from photorealistic to painterly.
- **Physics-aware movement:** Characters navigate stairs, obstacles, and complex terrains with physical plausibility.
- **Interactive freedom:** Users can move, rotate views, and manipulate environments in real time through simple commands.

By delivering low-latency, long-sequence generation with robust scene semantics, Matrix-Game 2.0 blurs the line between content creation and interaction. It also positions Skywork AI as a champion for open research in a space increasingly dominated by closed, corporate models. If Genie 3 showed what's possible behind closed doors, Matrix-Game 2.0 just handed the keys to the entire AI developer community.

link

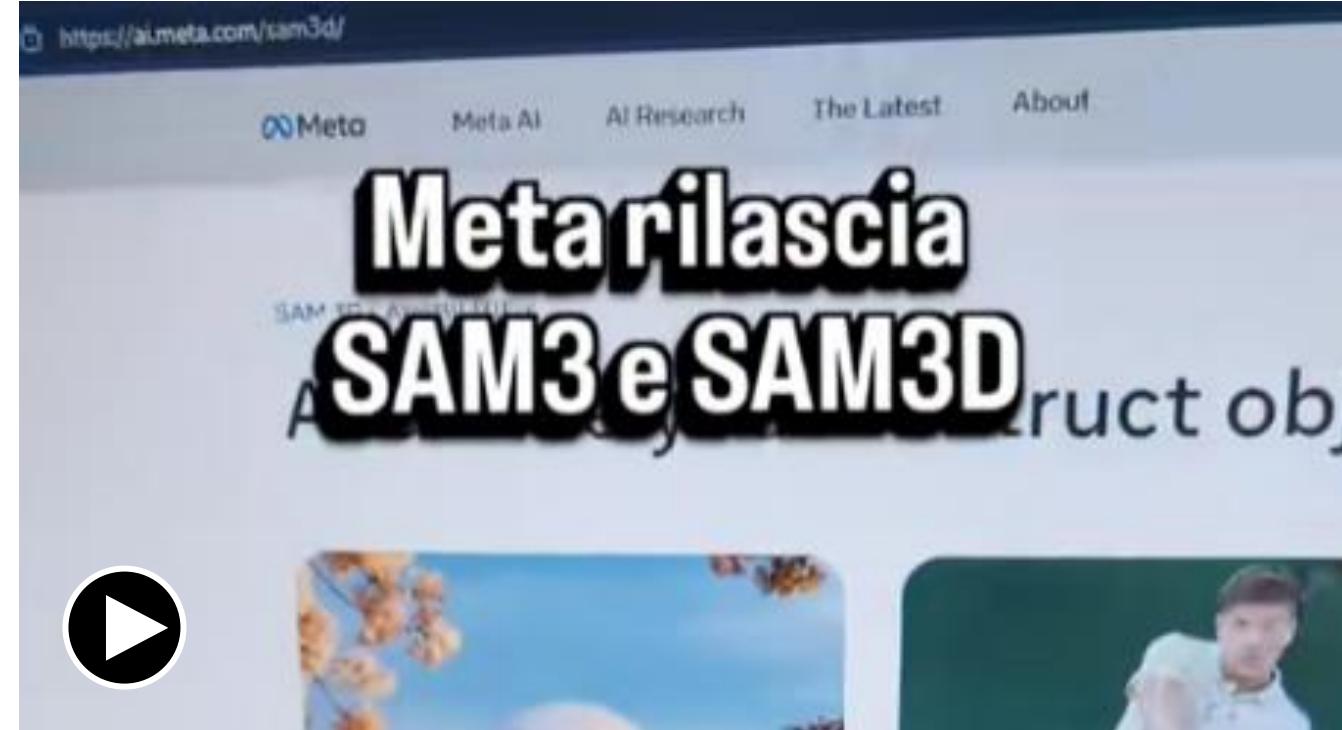


back to tassonomia



novembre 2025

un modello che  
trasforma le immagini in  
modelli 3D

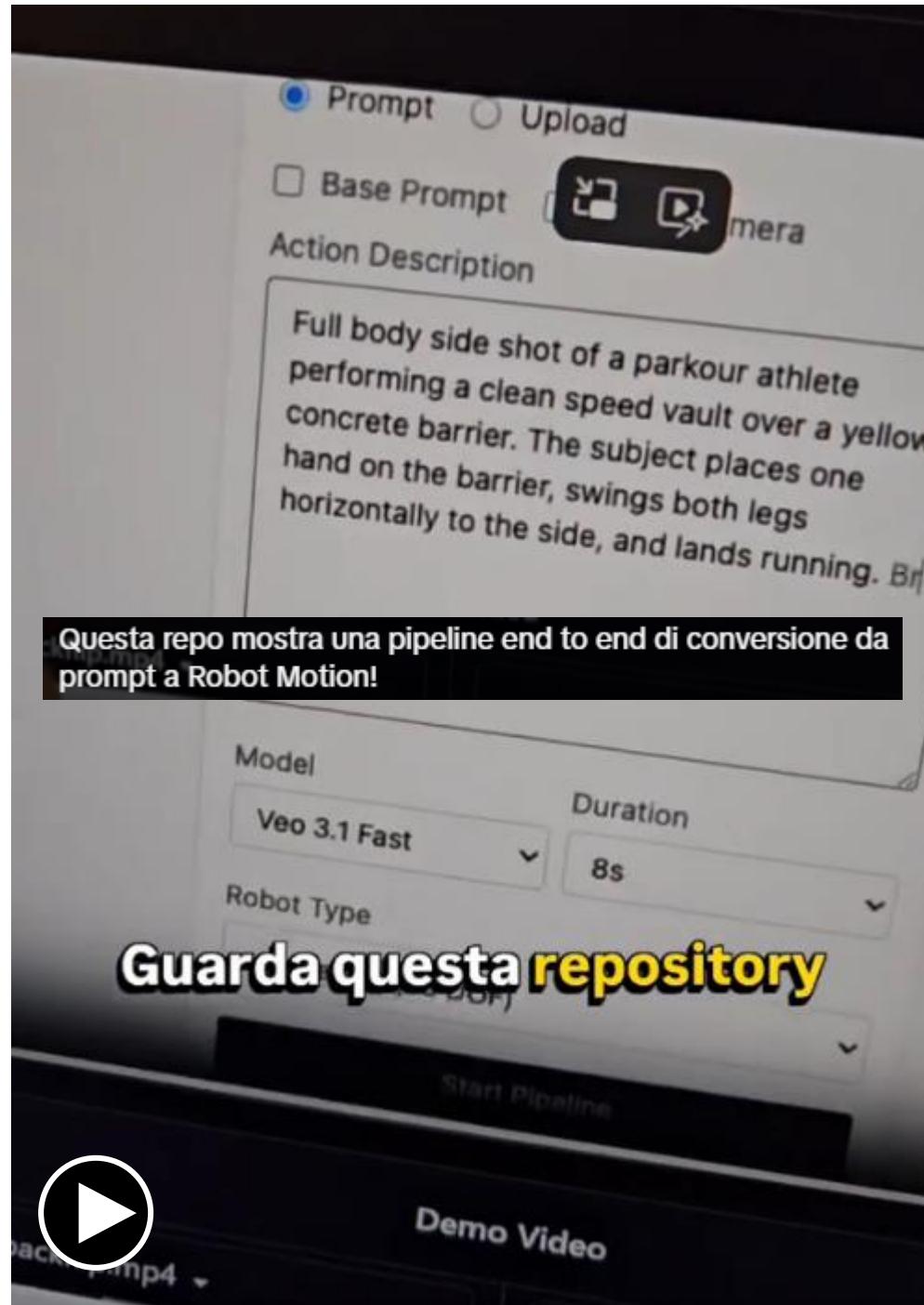


back to tassonomia

## Robot motion conversion

brainlink\_project

dicembre 2025



Questa repo mostra una pipeline end to end di conversione da prompt a Robot Motion!



back to tassonomia

# Google Genie: Text 2 3D

## Project Genie: Real-Time Worlds From a Text Prompt



**Project Genie** is a Google DeepMind demo for Genie 3, a 'world model' that creates an explorable environment from a prompt. When you type a scenario, it produces a 720p world and lets you navigate around at 20–24 fps. It allows quick iteration when you need a plausible scene, camera path, or mood reference without a 3D tool.

### Core functions (and how to use them):

- **Scene blocking:** describe a location and lighting, then explore it to grab reference for a shot list or storyboard.
- **UI-in-context checks:** generate a "storefront" or "app kiosk" setting and use it to sanity-check signage, placement, and visual hierarchy.
- **Perspective testing:** switch between first-person and third-person to see which viewpoint communicates your idea better.
- **Level mood drafts:** Generate three variations of the same place by changing one detail like weather or time of day, then pick the version that supports your tone.
- **Iteration control:** Prompt a weather or object change to explore alternate moods or constraints.

### Try this yourself:

Create one prompt with three lines: an environment line describing a narrow interior space with lighting and weather, a camera line describing a slow first-person stroll with one turn, and an event line that changes anything minor after 10 seconds, like a light flicker or a door. Repeat the prompt three times, changing one variable (lighting, signage or object placement). Compare how well the world stays consistent as the mood changes and which features drift first.



## 1) Piattaforma core di Machine Learning

### Amazon SageMaker

È il centro della strategia ML AWS.

Componenti principali:

- SageMaker Studio (ambiente di sviluppo)
- Training (anche distribuito su GPU)
- JumpStart (modelli e template pronti)
- Autopilot (AutoML)
- Feature Store
- Pipelines (MLOps)
- Model Registry
- Deployment endpoint e monitoring

Uso tipico:

→ costruzione e gestione di modelli proprietari.

## 2) Generative AI (LLM e foundation models)

### Amazon Bedrock

Accesso a foundation models senza gestire infrastruttura.

Include:

- modelli Anthropic, Meta, Amazon Titan, ecc.
- fine-tuning
- RAG
- agent orchestration

### Amazon Titan

Famiglia di modelli foundation proprietari AWS:

- Titan Text
- Titan Embeddings
- Titan Image

### Agents for Bedrock

Per costruire agenti autonomi collegati a API e dati.

Uso tipico:

→ chatbot, copiloti, generazione contenuti, AI conversazionale.

## 3) Assistenti AI e copilot

### Amazon Q

Assistente AI per lavoro e sviluppo.

Varianti:

- Q Developer (coding, cloud)
- Q Business (knowledge aziendale)
- Q in QuickSight (BI conversazionale)

### CodeWhisperer

Generazione codice e suggerimenti per sviluppatori.

#### **Schema mentale semplice**

Tre grandi blocchi:

##### A. Costruire modelli

- SageMaker
- Trainium / GPU
- MLOps

##### B. Usare modelli già pronti

- Bedrock
- Titan
- Rekognition, Comprehend, Textract...

##### C. AI per lavorare meglio

- Amazon Q
- CodeWhisperer
- QuickSight AI



## **Amazon Polly**

**Implementa voci umane di alta qualità e dal suono naturale in decine di lingue**

Amazon Polly (il software da testo a voce degli Amazon Web Services)

<https://aws.amazon.com/it/polly/>



GPT sulla creazione T2V:

- «veo3 prompt director» di Claudio Zagarini
- «veo 3 director» di Claudio Zagarini

30.7.2025

Anche questo:

<https://chatgpt.com/g/g-68755521d2348191a5ea8f6457412d51-mj-fh-buddy>

Slide con ai: canva, altri.

Browser ai:

- [comet di perplexity](#): chat sulla pagina immediata, senza dover fare copia & incolla.
- Edge
- aspettiamo google - è difficile integrare i browser con ai

Consistenza personaggi: tra 2 giorni video RG di guida finale.

Agenti: DR legge solo le fonti, gli agenti possono interagire con pagine e file, tramite connettori, chatgpt e claude ne hanno tanti, sempre nuovi.

*Creative score*: tool fantastico per ottimizzare campagne, landing page, copertine.

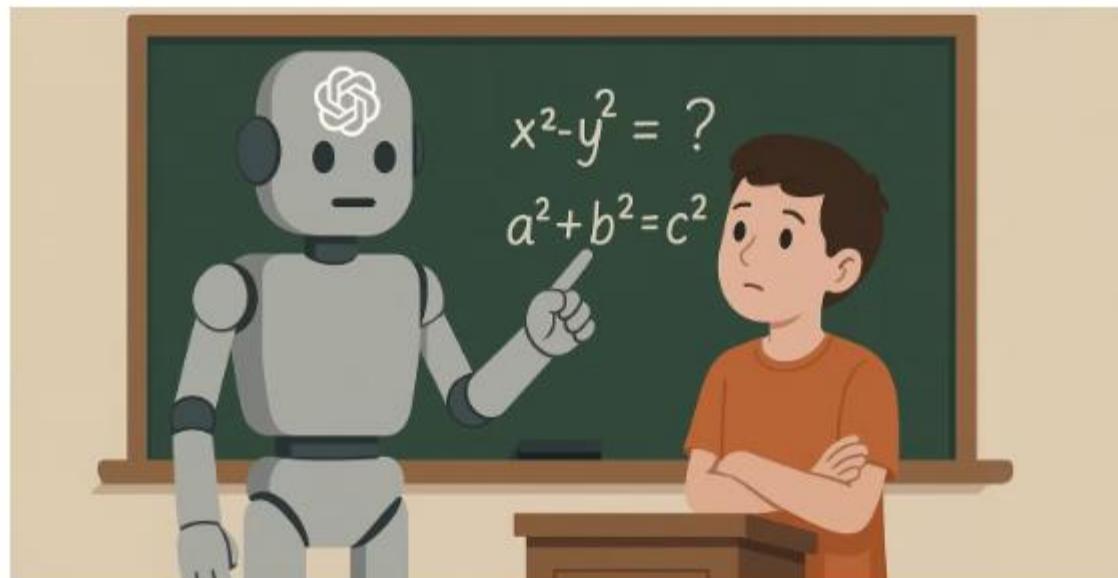


back to tassonomia

# L'IA non ti fa più i compiti, ora ti insegna a farli da solo: le novità di ChatGPT e NotebookLM per l'apprendimento

di Sergio Donato - 30/07/2025 12:24

2



Con "Studia e Impara" e le Video Overview, l'IA non fa più i compiti al posto degli studenti, ma li guida passo passo, stimola il ragionamento e trasforma ChatGPT e NotebookLM in tutor interattivi. Vorremmo crederci

Inutile negarlo, gli studenti hanno iniziato a usare da subito l'intelligenza artificiale per trovare un'alleata nei compiti a casa. Ieri, due delle piattaforme IA più avanzate, ChatGPT di OpenAI e NotebookLM di Google, hanno però introdotto **funzionalità pensate espressamente per supportare lo studio**, la comprensione profonda e la preparazione agli esami, in modo che l'IA non sia solo "uno strumento che fa i compiti al posto dei ragazzi".

Da una parte, "**Studia e Impara**", ora disponibile in **ChatGPT**, affianca gli studenti nel ragionamento passo dopo passo; dall'altra, **Video Overviews** di **NotebookLM** rende visivi e narrabili anche i concetti più complessi. Due approcci diversi ma complementari che segnano un passo deciso verso un'educazione mediata dall'IA, almeno per certi scopi.

## ChatGPT diventa tutor con la modalità "Studia e Impara"

La modalità "Studia e Impara" è ora accessibile agli utenti ChatGPT (Free, Plus, Pro, Team), anche in UE, con supporto imminente per gli ambienti educativi di ChatGPT Edu: una scelta questa quanto meno bizzarra, visto il tipo di prodotto.



Allora, Sergio. Ti butti?

Fai una domanda

+ Strumenti

Modalità agente Nuovo

Deep Research

Crea immagine

Studia e impara

Ricerca sul web

Canvas

ma ti posso caricare anche delle dispense?

Sì, assolutamente! Puoi caricare dispense, appunti, schede PDF, immagini, presentazioni — tutto quello che stai studiando.

Una volta caricato un file, posso:

- aiutarti a capire i concetti spiegati lì dentro,
- farti domande a partire dal testo (come un quiz),
- riassumere o chiarire i punti difficili,
- guidarti a rispondere a eventuali domande di fine capitolo, ma sempre senza fare il compito al posto tuo 😊

Non si tratta di una semplice risposta in stile “compito svolto”, ma di **un’interazione guidata da domande, suggerimenti strategici e sollecitazioni alla riflessione**. Il cuore del sistema è un set di istruzioni sviluppate con il contributo di pedagogisti, docenti e scienziati cognitivi. Queste istruzioni personalizzano l’interazione tenendo conto del livello dello studente, del suo storico nei precedenti dialoghi e della sua capacità di comprensione.

Tecnicamente, “Studia e Impara” si articola in prompt interattivi che stimolano il pensiero critico; risposte a sezioni che semplificano l’assimilazione di concetti complessi e quiz e verifiche integrati per la ritenzione delle conoscenze. **Si possono anche caricare materiali di studio come dispense, appunti, PDF, immagini e presentazioni**, in modo che poi ChatGPT possa “interrogare” lo studente o aiutarlo a capire la lezione, e soprattutto senza risolvere i compiti al posto suo.

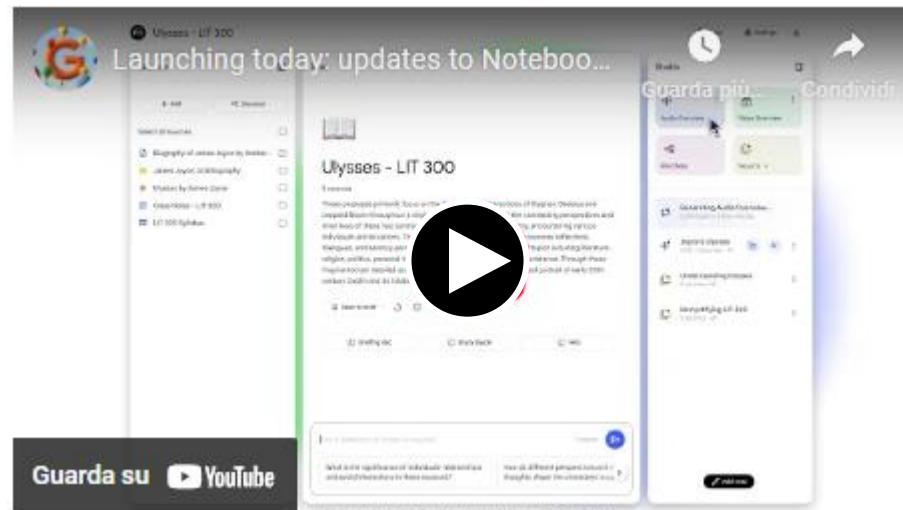
## NotebookLM, i concetti diventano video

Dall’altro lato, Google espande le capacità di NotebookLM con le Video Overview, pensate per aiutare chi impara attraverso elementi visivi. A differenza dei riassunti testuali o audio, questa nuova funzione **genera slideshow narrati che illustrano concetti estratti direttamente dai documenti dell’utente, arricchendoli con grafici, immagini, citazioni e numeri**. Una funzione che Google aveva già anticipato qualche mese fa in un tweet e che aveva suscitato aspettative comprensibilmente alte.



Le Video Overviews si adattano a diversi profili. **Dallo studente che affronta per la prima volta un argomento tecnico, fino al ricercatore che chiede un focus su una parte specifica.** Il sistema è flessibile, multiformato, e si integra in un pannello Studio rinnovato, che ora permette di generare più output dello stesso tipo (video, audio, mappe concettuali), navigare tra contenuti in parallelo (per esempio, ascoltare un audio mentre si esplora una mappa mentale), costruire contenuti multilingue o personalizzati per singoli team o corsi.

In termini di disponibilità, Video Overview è però (e purtroppo) **in fase di rollout per tutti gli utenti in lingua inglese, con supporto per altre lingue previsto a breve.**



Tra l'altro, bisogna aggiungere che, sempre ieri, Google **ha annunciato** un importante potenziamento di Gemini for Education, rendendo gratuitamente disponibili i limiti avanzati del modello Gemini 2.5 Pro per studenti ed educatori di ogni livello.

Basato su LearnLM (che si può provare anche in Google Ai Studio), **Gemini 2.5 Pro può essere sfruttato dagli insegnanti per creare piani didattici, adattare testi e aumentare il coinvolgimento degli studenti**, mentre gli studenti possono usarlo per chiarire concetti complessi, prepararsi agli esami e ricevere feedback personalizzato su testi e ricerche.

## Due visioni convergenti, ma poi gli studenti...

OpenAI e Google, sebbene con stili differenti, sembrano convergere su un'idea condivisa: **l'IA non dovrebbe sostituire l'apprendimento, ma amplificarlo.**

"*Studia e Impara*" fa leva sull'interazione dialogica e riflessiva, mentre Video Overview punta sulla chiarezza visuale e sulla sintesi multimediale.

In entrambi i casi, lo scopo non è fornire una scorciatoia, ma costruire strumenti che rafforzino la comprensione, stimolino la curiosità e rendano l'autonomia dello studente il centro dell'esperienza.

Ovviamente, nessuno potrà davvero vedere se gli studenti preferiranno farsi risolvere i compiti dall'IA o usare l'IA per allenare il proprio apprendimento. Probabilmente, però, conosciamo già tutti la risposta.



**claudio.hypergrowth**

Claudio | Hypergrowth

## Meet Study Mode



**lanciato la Study**

ChatGPT lancia la 'Study Mode': un tutor virtuale per studenti e professionisti

Parliamo della nuova modalità 'Study Mode' di ChatGPT, lanciata da OpenAI. Questa funzione può aiutare sia gli studenti che i professionisti a studiare e apprendere in modo più efficace, guidandoli con domande, indizi e feedback personalizzati. È come avere un tutor virtuale che utilizza il metodo socratico, potenziato dall'intelligenza artificiale. La modalità si adatta al tuo livello e obiettivi, e può essere attivata e disattivata quando vuoi. È disponibile per gli utenti free, plus e premium, e presto arriverà anche una versione per le scuole con quiz e prompt interattivi. Questo strumento non è un aiuto per fare i compiti, ma piuttosto un supporto per studiare, fare pratiche e costruire maggiore fiducia nelle tue capacità.



back to tassonomia

link a Google AI Studio



marcelloscacchettiai

Marcello Scacchetti | AI



## Google AI Studio



Stream realtime

Studio



Stream



History



Dashboard



Documentation



settembre 2025



Talk to Gemini live



back to tassonomia

come posso creare un audio con notebook lm?

NotebookLM ti permette di creare un "Audio Overview" a partire dalle tue fonti, che è un riassunto dei tuoi documenti sotto forma di conversazione tra due voci AI. È un po' come un podcast generato dalle tue note.

Ecco come puoi farlo:

#### Passaggi per creare un Audio Overview in NotebookLM

1. **Accedi a NotebookLM:** Vai su <https://notebooklm.google.com/> e accedi con il tuo account Google.

#### 2. Crea o apri un blocco note:

- Se è la prima volta, ti verrà chiesto di creare un "New Notebook" (Nuovo blocco note).
- Se hai già dei blocchi note, puoi aprirne uno esistente o crearne uno nuovo.

3. **Aggiungi le tue fonti:** Per generare un audio, NotebookLM ha bisogno di materiale su cui basarsi. Puoi aggiungere diverse fonti, tra cui:

- File PDF, TXT, Markdown dal tuo computer.
- Link a siti web.
- Testo copiato e incollato.
- File di Google Docs o Google Slides.
- Link a video di YouTube (NotebookLM ne estrae la trascrizione).
- File audio (MP3 o WAV, che verranno trascritti).

#### 4. Genera l'Audio Overview:

- Una volta aggiunte le fonti, vedrai un pannello chiamato "Studio" o "Notebook guide".
- Cerca la sezione "Audio Overview" e fai clic su "Generate" (Genera).
- Potrebbe volerci qualche minuto, soprattutto per blocchi note più grandi.

#### 5. Personalizza (Opzionale):

- Prima di generare, o dopo, puoi scegliere la **lunghezza** dell'audio (più breve, predefinito o più lungo).
- Puoi anche impostare un "**steering prompt**" (prompt di direzione) per focalizzare la discussione su argomenti specifici o regolare il livello di competenza (massimo 500 caratteri).
- NotebookLM supporta la generazione di Audio Overview in **oltre 50 lingue**. Puoi scegliere la lingua di output nelle impostazioni.

#### 6. Interagisci o scarica l'audio:

- Una volta generato, puoi **ascoltare** l'Audio Overview direttamente in NotebookLM.
- In modalità interattiva (Beta), puoi persino **porre domande** agli host AI mentre ascolti, e loro risponderanno basandosi sulle tue fonti.
- Puoi anche **scaricare** il file audio per ascoltarlo offline o condividerlo.

#### Vantaggi dell'Audio Overview:

- **Comprensione migliorata:** Ascoltare le informazioni può aiutare a elaborare concetti complessi.
- **Apprendimento accessibile:** È utile per chi preferisce imparare attraverso l'ascolto o per ripassare i contenuti in movimento.
- **Creazione rapida di contenuti:** Trasforma le tue fonti in una discussione coinvolgente, utile per riassunti, podcast educativi, ecc.

Tieni presente che la funzione "Audio Overview" è in continua evoluzione e Google aggiunge regolarmente nuove funzionalità e miglioramenti.



back to tassonomia

# Google AI Pro

## ◆ App Gemini

Ottieni un livello di accesso superiore a 3.1 Pro, il nostro modello più intelligente, a Deep Research e alla generazione di immagini con Nano Banana Pro, oltre a usufruire della generazione di video con Veo 3.1 Fast<sup>6</sup>

*Da Gemini, non da Google AI Studio*

## (○) 1000 crediti AI mensili<sup>4</sup>

Crediti utilizzati per la generazione di video su Flow e Whisk

## ■ Flow<sup>5</sup>

Accesso superiore al nostro strumento AI di creazione video per realizzare scene e storie cinematografiche, con l'accesso limitato a Veo 3.1 incluso<sup>6</sup>

## . Whisk<sup>7</sup>

Maggiore accesso alla creazione di video da immagini con Veo 3

## ☒ Gemini Code Assist e interfaccia a riga di comando di Gemini

Limiti giornalieri di richieste più elevati nelle estensioni IDE di Gemini Code Assist e nell'interfaccia a riga di comando di Gemini

## ▲ Google Antigravity

Limiti di frequenza più elevati per il modello di agente in Google Antigravity, la nostra piattaforma di sviluppo agentico

## ☒ NotebookLM

Assistente di ricerca e scrittura con un numero quintuplicato di overview audio, notebook e altro

## ☒ Gemini in Gmail, Documenti, Vids e altro

Accedi a Gemini direttamente nelle app Google

## ⌂ Google Home Premium (piano Standard)

30 giorni di cronologia degli eventi e funzionalità di Gemini

## ☁ Spazio di archiviazione

2 TB di spazio di archiviazione totale per Foto, Drive e Gmail

link al piano





is there in the AI market any tool able to read PDF files and tables inside it too?

Yes, absolutely — there are several tools in the AI market that can read PDF files including tables, and extract the content in a structured way. Here's a quick breakdown of some solid options, depending on your needs:

### 1. Adobe Acrobat with AI (Acrobat AI Assistant)

- **What it does:** Uses generative AI to summarize, answer questions, and extract info from PDFs — including tables.
- **Strength:** Great UI, built into Acrobat ecosystem, and very user-friendly.
- **Limitations:** Not ideal for large-scale automation or integration into pipelines.

### 2. ChatGPT (with code interpreter or file upload)

- **What it does:** You can upload a PDF, and it will read the content, extract tables, and even convert them into CSV/JSON if needed.
- **Strength:** Very flexible. Can write code (e.g., Python with `pdfplumber`, `PyMuPDF`, `tabula`) to extract data precisely.
- **Limitations:** Manual interaction needed unless you integrate via API or custom solution.

Lo fa Claude,  
chatGPT NO!!

### 3. Amazon Textract

- **What it does:** Extracts text, form data, and tables from scanned PDFs/images.
- **Strength:** Strong OCR, very scalable, ideal for enterprise use.
- **Limitations:** AWS setup required, and cost may increase for high volumes.

### 4. Microsoft Azure Form Recognizer

- **What it does:** Reads PDFs (including forms and tables) and converts them into structured JSON.
- **Strength:** Integrates well in enterprise Azure environments.
- **Limitations:** Requires some setup and tuning.

### 5. PDFplumber (Python Library)

- **What it does:** Programmatic tool that can read PDF content and detect tables with high accuracy.
- **Strength:** Free, great for custom workflows, supports table extraction well.
- **Limitations:** Requires coding knowledge.

### 6. LangChain + LLMs

- **What it does:** Combines LLMs (like OpenAI or Claude) with PDF parsers to do semantic search, Q&A, or summarization over PDF content.
- **Strength:** Very customizable, ideal for chat-over-PDF type applications.
- **Limitations:** Needs a bit of setup, more advanced.

