



DICEMBRE 2023

# Login to Web3

## Dai Lumière all'AI: guida alla scelta dei tool text to video

Wavemaker

## Trasformare un testo in un video con l'IA

Nell'era in cui le intelligenze artificiali generative stanno rivoluzionando il modo in cui creiamo contenuti, un nuovo capitolo entusiasmante sta prendendo forma: **la generazione di video a partire da testo**. Dopo aver esplorato i benefici degli strumenti di Text to Text e di Text to Image, questa volta proviamo a capire a che punto siamo nella creazione di immagini in movimento.

### Dai Lumière all'IA

Diciamo subito che a differenza degli altri due campi precedentemente esplorati, già in fase avanzata, quello del Text to Video è da poco uscito dall'ambito della ricerca. Siamo in una fase che ricorda molto quella di fine ottocento, inizi novecento con i primi esperimenti di cinema dei fratelli Lumière. In pratica oggi esistono software commerciali che permettono di **manipolare video e di scrivere una descrizione testuale sulla cui base viene generato un video**. Attualmente i risultati sono brevi spezzoni di immagini in movimento che ricordano molto i film degli inventori del "cinematografo".

Proviamo a capire quali sono i software che possono immediatamente essere utilizzati nella realizzazione di contenuti per il marketing e la comunicazione aziendale.



## Runway, Pika Labs e Kaiber

**Runway** è il prodotto di punta dell'azienda Runway AI che ha guidato gli sforzi di ricerca dietro Stable Diffusion, uno dei più importanti modelli di generazione di immagini da testo. Runway nasce come uno **strumento di editing per la manipolazione di video con tecniche di intelligenza artificiale**. Permette di eliminare o sostituire elementi da una scena, definire il percorso di oggetti (motion tracking), aggiungere effetti speciali e "color grading" (la colorazione di scene sulla base di una certa palette).

**Nell'ultimo anno sono stati aggiunti due moduli basati su tecniche di intelligenza artificiale generativa: GEN-1 e GEN-2.**

**GEN-1** è un tool di video to video che da un video caricato, permette di generare un video simile al quale possono essere applicati degli effetti. In pratica, ci dà la possibilità di caricare un video e di trasformarlo sulla base di tre opzioni: un'immagine che diamo come riferimento-guida, un preset (uno effetto predefinito dal sistema) o un prompt/descrizione testuale. Per esempio posso caricare un video di una persona con un tubo di ferro in mano e chiedere di trasformarlo in un guerriero Jedi con una spada laser.

**GEN-2** è uno strumento di text to video che consente di generare un video di 4 secondi (estendibile a 16) a partire da una descrizione testuale, da un'immagine fissa (che verrà animata) o da un mix di entrambe.

Runway può essere testato gratuitamente fino ad un certo numero di output (125 crediti corrispondenti a 8 secondi di video creati con GEN-1 e 25 secondi con GEN-2) per poi passare a dei piani di abbonamento che partono da 12 dollari al mese.

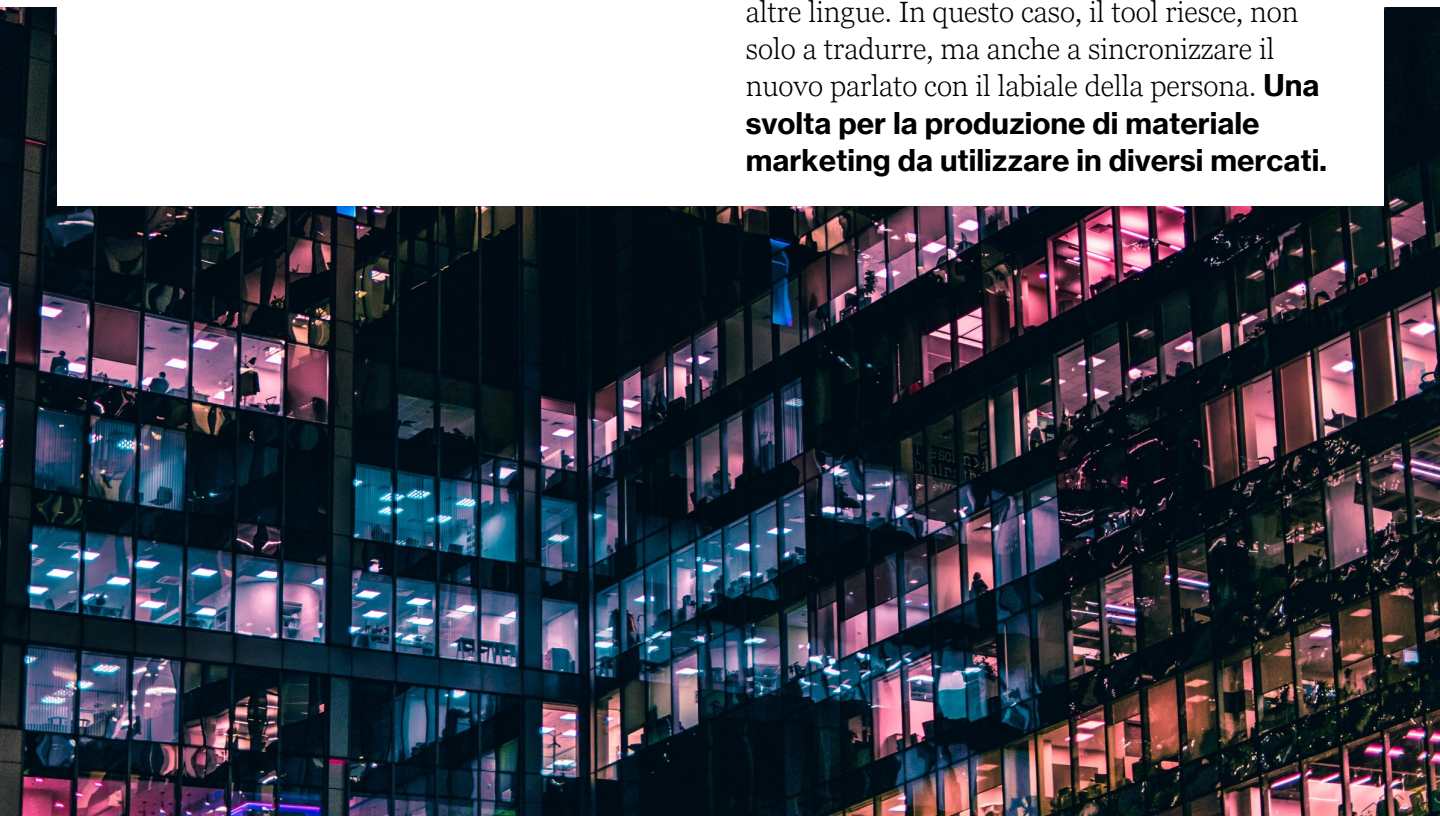


Tra i software alternativi a Runway, vale la pena di tenere d'occhio **Pika Labs e Kaiber**. Il primo è gratuito, ma ancora in beta e si può utilizzare solo attraverso l'interfaccia di Discord. Qui dentro basta scrivere un prompt con la sintassi /create seguita da una descrizione e da alcuni parametri per ottenere dei video di soli 3 secondi. Il risultato è ancora acerbo, ma usabile, anche se richiede diversi tentativi.

Il secondo tool, **Kaiber**, è già stato utilizzato per **la generazione di videoclip musicali** come quello della nota band Linkin Park. Tende a non produrre scene fotorealistiche, ma che ricordano animazioni fantasiose e oniriche. Anche qui si può partire da un'immagine fissa, magari generata con un software come Midjourney o Dall-E, da un video oppure da un prompt testuale. Nativamente, è anche possibile aggiungere un tappeto sonoro e far muovere le immagini in sincrono con la musica. Si può usare gratuitamente, anche da applicazione mobile, per poi passare ad un piano a pagamento quando si hanno esigenze più sofisticate.

**Nell'ambito della categoria Text To Video**, una menzione speciale meritano software come **HeyGen o Synthesia** che sono pensati principalmente per la creazione di presentazioni video. Per esempio, quelle che prevedono la presenza di una persona che spiega l'utilizzo di un prodotto o che espone un certo tema (usate per i training aziendali). In particolare HeyGen spicca per la **rapidità delle innovazioni che ha prodotto nei mesi scorsi**. Il software permette di creare video con oltre cento avatar digitali ai quali mettere in bocca qualunque testo. Alla presentazione, poi, si possono aggiungere testi, immagini e video. Inoltre, con la funzione "Instant Avatar" è possibile caricare il video di una persona reale e poi "clonarlo digitalmente" in modo da utilizzare il suo avatar/clone all'infinito facendogli declamare qualunque nuovo testo.

Strabiliante anche la funzione "**Video Translator**" che ci dà la possibilità di caricare un video di una persona che parla in italiano e di ottenere un nuovo video nel quale la stessa persona, con lo stesso tono di voce, parlerà in altre lingue. In questo caso, il tool riesce, non solo a tradurre, ma anche a sincronizzare il nuovo parlato con il labiale della persona. **Una svolta per la produzione di materiale marketing da utilizzare in diversi mercati.**





**Questi strumenti di text to video, ancora poco maturi ma in rapida evoluzione, potrebbero essere rivoluzionari per i creativi e i marketer. Assisteremo a un abbattimento delle barriere creative, che consentirà a un pubblico più ampio di esprimersi visivamente con più facilità e senza grandi mezzi materiali? Ci troveremo di fronte ad una resistenza da parte di artisti e direttori creativi o emergeranno nuove figure in grado di sfruttare queste tecnologie per esplorare nuovi orizzonti della creatività?**



Any questions? We're here to help.  
Please contact  
[positive.provocations@wmglobal.com](mailto:positive.provocations@wmglobal.com)

**Vincenzo Cosenza &  
Wavemaker Innovation Team**

**Wavemaker**

**Grow fearless**