

Creare immagini da testi con Copilot e Canva

L'intelligenza artificiale generativa rappresenta una delle aree più innovative e in rapida evoluzione nel campo dell'IA. Tra i modelli più diffusi troviamo quelli *text-to-image*, che sono in grado di generare immagini dettagliate e visivamente coerenti partendo da descrizioni testuali.

I modelli generativi *text-to-image* si basano su reti neurali profonde e utilizzano i trasformatori, la stessa tecnologia dei modelli linguistici, per tradurre testo in immagini. I **trasformatori**, una particolare architettura di reti neurali profonde proposta da Google nel 2017, hanno rivoluzionato l'analisi e la generazione dei testi in linguaggio naturale e sono stati adattati per gestire anche dati visivi. Questi modelli, allenati su vasti dataset di immagini e testi a queste associati, apprendono a **codificare relazioni complesse tra le descrizioni testuali e i contenuti visivi**, permettendo la generazione di immagini dettagliate e pertinenti rispetto al testo che le descrive.

I modelli generativi *text-to-image* stanno aprendo nuove frontiere nella creazione di contenuti, offrendo strumenti di considerevole potenza per artisti, designer, educatori e insegnanti. Tuttavia, l'entusiasmo per queste possibilità deve essere bilanciato da un impegno all'uso responsabile, critico ed etico di questi modelli.

Questa scheda tratta la generazione di immagini mediante modelli generativi *text-to-image* ad accesso libero, in particolare *Microsoft Copilot* e *Canva*.

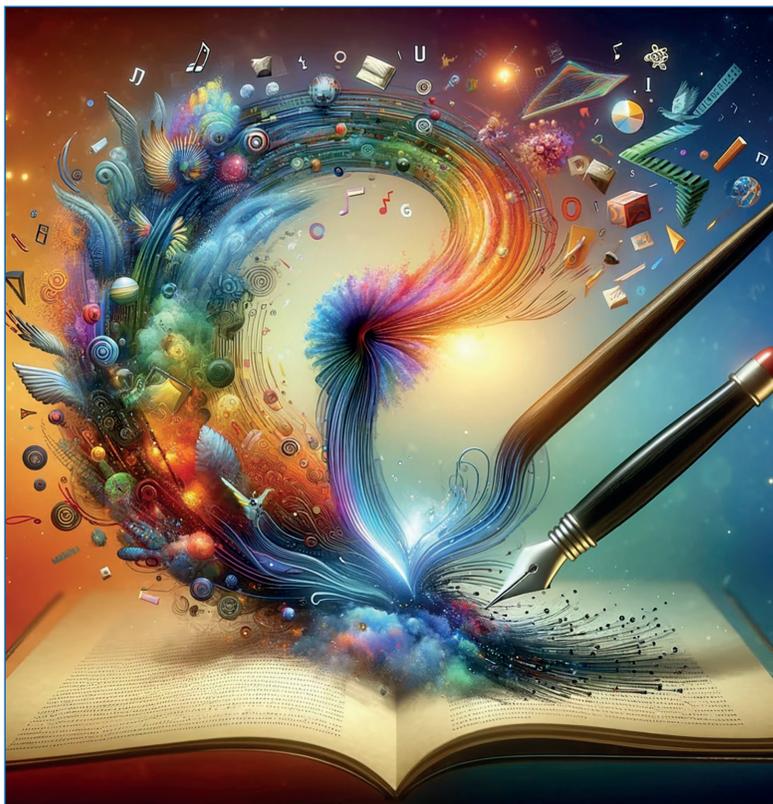


Immagine realizzata con *ChatGPT4*: “Genera un’immagine rappresentativa dei modelli di intelligenza artificiale *text-to-image*.”

Microsoft Copilot

Copilot è l'intelligenza artificiale di *Microsoft*: richiede di registrarsi per accedere al sito, ma offre un piano di utilizzo completamente gratuito che consente di:

- accedere sul web da computer o da dispositivi mobili;
- conversare con l'IA usando testo, voce e immagini;
- creare immagini con la funzione *Designer* (in precedenza denominata *Bing Image Creator*).

Dopo l'accesso il sito si presenta come nell'immagine sottostante.

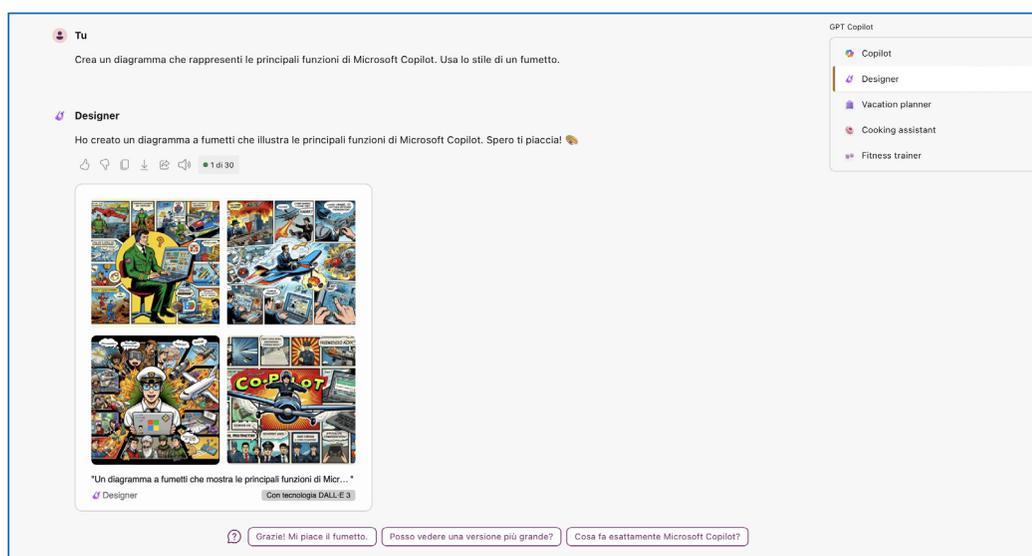


In questa finestra, all'interno della casella *Chiedimi qualsiasi cosa* possiamo scrivere all'intelligenza artificiale le nostre richieste, i cosiddetti **prompt**, esattamente come faremmo con *ChatGPT* o altri chatbot basati su modelli linguistici di grandi dimensioni.

Possiamo anche scrivere prompt che richiedano la rappresentazione di immagini, ma per avere il massimo dalla generazione *text-to-image* è preferibile usare la funzione *Designer* accessibile dal menù in alto a destra. Proviamo questa funzione inserendo il prompt:

*Crea un diagramma che rappresenti le principali funzioni di Microsoft Copilot.
Usa lo stile di un fumetto.*

L'immagine sotto riporta la schermata con la risposta.



Notiamo che *Copilot* risponde al nostro prompt con un breve **testo riassuntivo** della nostra richiesta seguito da **quattro diverse immagini**.

La creazione di immagini avviene con *DALL-E 3*, la tecnologia proprietaria di *OpenAI*, utilizzata anche in *ChatGPT4*, attualmente uno tra i prodotti di punta per quanto riguarda le intelligenze artificiali generative.

Obiettivo di questa scheda è trattare la generazione di immagini a partire da un prompt di testo: ci concentreremo dunque su questa funzionalità tralasciando le altre.

I PROMPT

Un prompt per immagini è l'input di testo che deve essere fornito per generare un'immagine. Può essere costituito da una breve frase oppure da un testo più complesso o un intero paragrafo.

Il formato del prompt e le informazioni in esso contenute hanno un impatto significativo su **qualità** e **varietà** delle immagini che si otterranno.

Per generare immagini, occorre che il prompt cominci con le parole appropriate, altrimenti *Copilot* potrebbe rispondere in forma testuale senza generare alcuna immagine.

Per indirizzare *Copilot* a generare un'immagine anziché una conversazione, possiamo usare prompt che iniziano nei modi seguenti:

- Crea un'immagine di...
- Genera un paesaggio di...
- Crea un ritratto realistico di...
- Genera una scena di fumetto con...
- Genera un'opera d'arte di...
- Illustra un (soggetto o scena) con...
- Immagina un design futuristico di...
- Crea un panorama dove...
- Crea un dipinto di...
- Crea una scena immaginaria di...
- Mostrami un'immagine ravvicinata di...
- Crea un diagramma di...
- Rappresenta (concetto o idea) in una immagine...

Questi sono soltanto alcuni esempi, la lista completa sarebbe molto più lunga.

Una caratteristica importante di *Designer* in *Copilot* è che ignora le **stop words**, esattamente come i motori di ricerca tradizionali. Quindi parole come **articoli**, **congiunzioni**, **preposizioni** sono **poco rilevanti** per la buona riuscita del prompt che stiamo scrivendo.

Non omettiamo le **stop words** dai nostri prompt perché ci sono utili per formularli correttamente dal punto di vista sintattico, ma concentriamoci sull'uso di termini descrittivi che siano significativi per l'immagine che desideriamo ottenere.

Per avere migliori risultati, *Microsoft* consiglia di utilizzare **almeno 6 parole descrittive direttamente collegate all'immagine desiderata**.

Per esempio nel prompt:

Mostrami un'immagine ravvicinata di Marte. Nelle vicinanze di Marte deve essere visibile la sua luna Fobos,

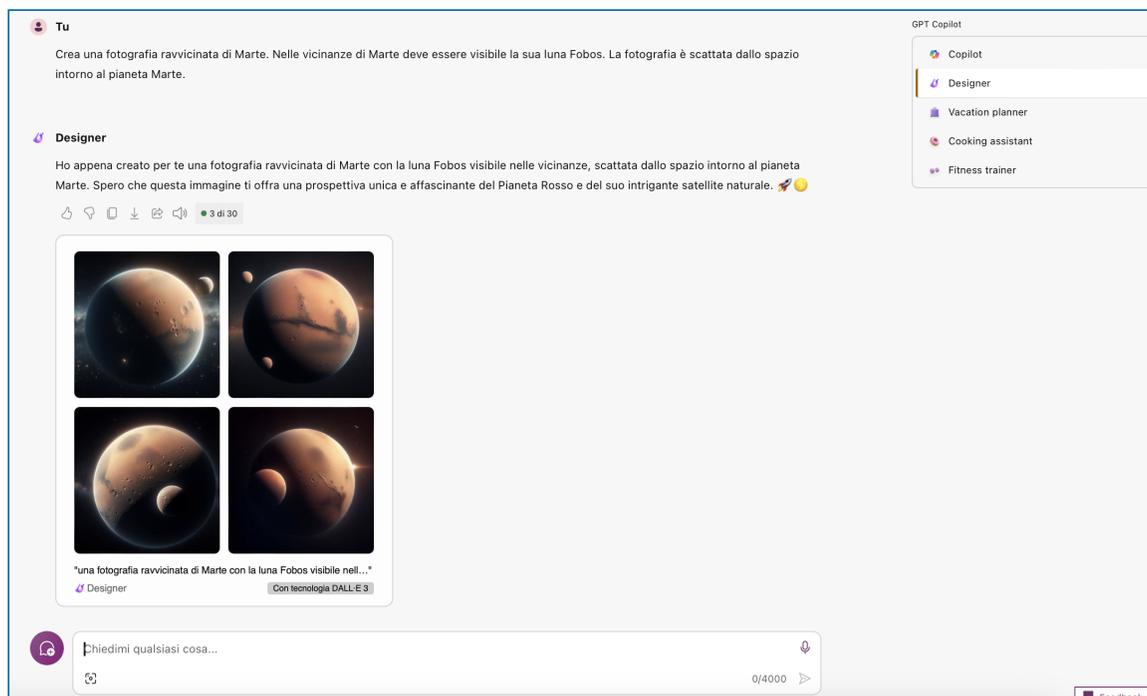
illustrato nella figura sotto, le parole descrittive sono:

Mostrami, immagine, ravvicinata, vicinanze, Marte, luna, Fobos.



Il risultato non è molto soddisfacente, ma possiamo migliorarlo specificando che la fotografia è scattata dallo spazio:

Crea una fotografia ravvicinata di Marte. Nelle vicinanze di Marte deve essere visibile la sua luna Fobos. La fotografia è scattata dallo spazio intorno al pianeta Marte.



Abbiamo ottenuto risultati migliori, ma sia il Pianeta Rosso, sia la sua luna sono molto diversi dalla realtà. Questo è dovuto al fatto che i modelli *text-to-image* difficilmente riescono a riprodurre immagini di oggetti reali, ma riescono a realizzarne soltanto imitazioni.

In genere, un prompt può contenere diversi elementi utili a caratterizzare l'immagine che desideriamo. Tra questi abbiamo:

- il tipo di immagine
- il soggetto
- il contesto
- eventuali dettagli
- l'illuminazione
- l'inquadratura
- il tipo di lente o fotocamera
- lo stile.

Per esempio la schermata a pagina 5 illustra il prompt

Crea un dipinto impressionista di un computer portatile appoggiato su un tavolo di fronte a una finestra oltre la quale si vede un bosco. La luce è naturale e lo stile quello di Claude Monet

nel quale definiamo:

- il **tipo di immagine**: un dipinto impressionista,
- il **soggetto**: un computer portatile appoggiato su un tavolo in legno di fronte a una finestra oltre la quale si vede un bosco,
- l'**illuminazione**: luce naturale,
- lo **stile**: quello di Claude Monet.

Tu

Crea un dipinto impressionista di un computer portatile appoggiato su un tavolo in legno di fronte a una finestra oltre la quale si vede un bosco. La luce è naturale e lo stile quello di Claude Monet.

Designer

Ho creato un dipinto impressionista che raffigura un computer portatile su un tavolo in legno, di fronte a una finestra che si affaccia su un bosco. La luce naturale filtra attraverso la finestra, illuminando la scena nello stile luminoso e colorato di Claude Monet. Spero che questa opera d'arte ti ispiri e porti un tocco di bellezza impressionista nel tuo ambiente. 🎨

🔗 📄 📥 🗑️ 🗣️ 4 di 30



"un dipinto impressionista di un computer portatile appoggiato su..."

GPT Copilot

- Copilot
- Designer**
- Vacation planner
- Cooking assistant
- Fitness trainer

Elenchiamo alcune buone pratiche per scrivere prompt efficaci.

- Possiamo usare un **processo iterativo** per arrivare a ottenere prompt ottimali, sperimentando: una strategia è iniziare con prompt semplici e migliorarli progressivamente, aggiungendo tutti gli elementi tra quelli visti poc'anzi, necessari a caratterizzare l'immagine desiderata.
- Dobbiamo impartire sempre **istruzioni chiare** come quelle degli incipit proposti in questo paragrafo.
- Dobbiamo essere molto specifici riguardo all'immagine che vogliamo ottenere, quindi cerchiamo di ripulire i nostri prompt da eventuali imprecisioni e di non tralasciare **dettagli importanti**, come colori, dimensioni degli oggetti rappresentati, forme, posizioni, luci, inquadratura.
- Evitiamo di scrivere nei prompt cosa non fare, scriviamo invece **cosa fare**.
- Evitiamo le **ambiguità**.

UN ESEMPIO: L'IMMAGINE DI UN PENDOLO SEMPLICE PER UNA LEZIONE DI FISICA

Vediamo all'opera il processo iterativo per la creazione dell'immagine di un **pendolo semplice**, che potremmo inserire in una lezione di fisica.

Iniziamo con il prompt dell'immagine sotto:

Crea una fotografia di un pendolo semplice da usare per una lezione di fisica.

Tu

Crea una fotografica di un pendolo semplice, da usarsi per una lezione di fisica.

Designer

Ho creato un'immagine di un pendolo semplice che potrebbe essere utilizzata per una lezione di fisica. Spero che sia di tuo gradimento e che aiuti a rendere la lezione più interessante e visivamente stimolante! ✨

🔗 📄 📥 🗑️ 🗣️ 1 di 30



"un pendolo semplice"

Designer

Con tecnologia DALL·E 3

GPT Copilot

- Copilot
- Designer**
- Vacation planner
- Cooking assistant
- Fitness trainer

Applichiamo il **metodo iterativo** e miglioriamo questo primo prompt arricchendo di **dettagli** il soggetto:

Crea un'illustrazione editoriale di un pendolo semplice composto da una sfera in acciaio, un supporto in legno e un filo. Il pendolo sta su un tavolo all'interno di un laboratorio di fisica.

Tu

Crea un'illustrazione editoriale di un pendolo semplice composto da una sfera in acciaio, un supporto in legno e un filo. Il pendolo sta su un tavolo all'interno di un laboratorio di fisica.

Designer

Ho creato un'illustrazione editoriale di un pendolo semplice come hai descritto: con una sfera in acciaio, un supporto in legno e un filo, posizionato su un tavolo all'interno di un laboratorio di fisica. Spero che questa immagine possa essere un ottimo supporto visivo per la tua lezione di fisica! 🌟

👍 🗨️ 📄 ⬇️ 📧 🔊 2 di 30



"un pendolo semplice composto da una sfera in acciaio, un supp..."
👍 Designer Con tecnologia DALL·E 3

GPT Copilot

- Copilot
- Designer
- Vacation planner
- Cooking assistant
- Fitness trainer

Possiamo anche definire lo **stile** dell'immagine, per esempio nel caso sottostante abbiamo scelto lo stile *stop motion*.

Tu

Crea un'illustrazione in stile stop motion di un pendolo semplice composto da una sfera in acciaio, un supporto in legno e un filo. Il pendolo sta su un tavolo all'interno di un laboratorio di fisica.

Designer

Ho creato un'illustrazione in stile stop motion di un pendolo semplice, con una sfera in acciaio, un supporto in legno e un filo, posizionato su un tavolo all'interno di un laboratorio di fisica. Spero che questa immagine possa essere utile per la tua lezione e che catturi l'attenzione dei tuoi studenti! 🌟 📸

👍 🗨️ 📄 ⬇️ 📧 🔊 3 di 30



"un'illustrazione in stile stop motion di un pendolo semplice com..."
👍 Designer Con tecnologia DALL·E 3

GPT Copilot

- Copilot
- Designer
- Vacation planner
- Cooking assistant
- Fitness trainer

È anche possibile creare **immagini 3D** simili a quelle create con un software di grafica tridimensionale, dobbiamo semplicemente specificare:

Crea un'immagine 3D di un pendolo semplice. Il pendolo sta su un tavolo. Luce fredda. Lente 50 mm.

Tu

Crea un'immagine 3D di un pendolo semplice. Il pendolo sta su un tavolo. Luce fredda. Lente 50mm.

Designer

Ho creato un'immagine 3D di un pendolo semplice, posizionato su un tavolo, con una luce fredda e una prospettiva equivalente a quella di una lente da 50mm. Spero che questa immagine possa essere esattamente ciò che cercavi per illustrare il concetto di pendolo nelle tue lezioni di fisica! 🤖👉

👍 🗨️ 📄 ⬇️ 🗑️ 🗣️ 5 di 30



"un'immagine 3D di un pendolo semplice, su un tavolo, con luce ..."

Designer Con tecnologia DALLE 3

GPT Copilot

- Copilot
- Designer
- Vacation planner
- Cooking assistant
- Fitness trainer

Per concludere vediamo un ultimo esempio

Crea una Claymation di un pendolo semplice posto su un tavolo all'interno di un laboratorio di fisica. Il tavolo è circondato da alunni e alunne che hanno sguardi curiosi. Lente grandangolare, fuoco soft, ombre leggere, inquadratura dalla stessa altezza degli alunni.

Il prompt contiene

- lo **stile**: Claymation
- il **soggetto**: un pendolo semplice posto su un tavolo all'interno di un laboratorio di fisica
- il **contesto**: il tavolo è circondato da alunni e alunne che hanno sguardi curiosi
- il **tipo di lente**: grandangolare
- la **messa a fuoco**: soft
- l'**illuminazione**: ombre leggere
- il tipo di **inquadratura**: alla stessa altezza degli alunni.

Tu

Crea una Claymation di un pendolo semplice posto su un tavolo all'interno di un laboratorio di fisica. Il tavolo è circondato da alunni e alunne che hanno sguardi curiosi. Lente grandangolare, fuoco soft, ombre leggere, inquadratura dalla stessa altezza degli alunni.

Designer

Ho creato una Claymation di un pendolo semplice, come hai richiesto. Il pendolo è posizionato su un tavolo all'interno di un laboratorio di fisica, circondato da alunni e alunne con sguardi curiosi. Ho utilizzato una lente grandangolare per catturare l'intera scena, un fuoco soft per dare un aspetto più delicato all'immagine, e ombre leggere per aggiungere profondità senza sovrastare i dettagli. L'inquadratura è stata impostata alla stessa altezza degli alunni per far sentire lo spettatore parte della scena. Spero che questa immagine possa essere un'aggiunta coinvolgente e educativa per la tua lezione di fisica! 🤖👉

👍 🗨️ 📄 ⬇️ 🗑️ 🗣️ 7 di 30

La generazione dell'immagine richiede tempo. Controlla lo stato di creazione dell'immagine in [Image Creator](#).

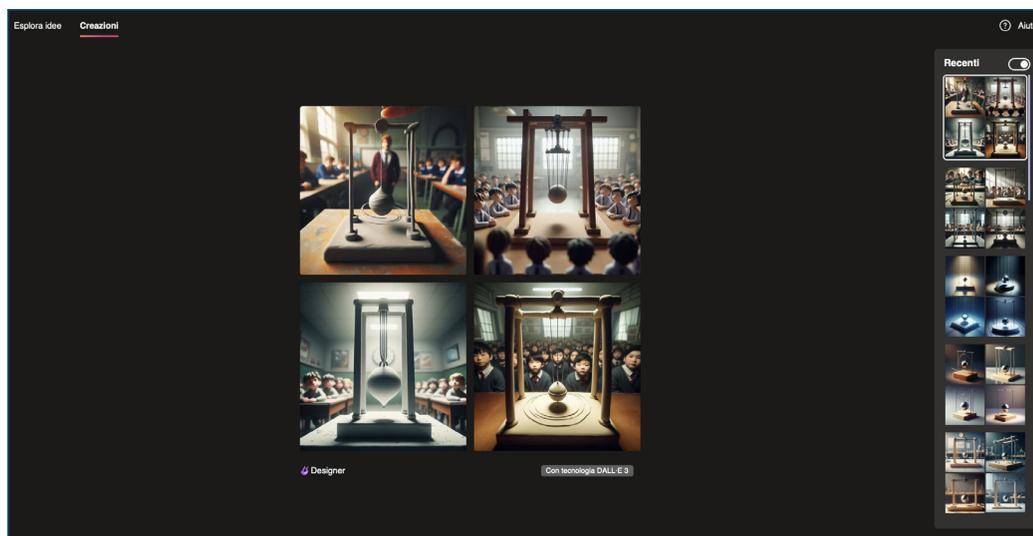


GPT Copilot

- Copilot
- Designer
- Vacation planner
- Cooking assistant
- Fitness trainer

In questo caso *Copilot* impiega più tempo del dovuto per generare le immagini e non le mostra all'interno della conversazione, come in tutti gli esempi precedenti.

Per ottenere le immagini dobbiamo cliccare sul link *Image Creator* e si aprirà una pagina come quella sottostante. All'interno di questa pagina vedremo le immagini al termine della generazione e nella barra destra avremo a disposizione tutta la cronologia delle immagini recenti.



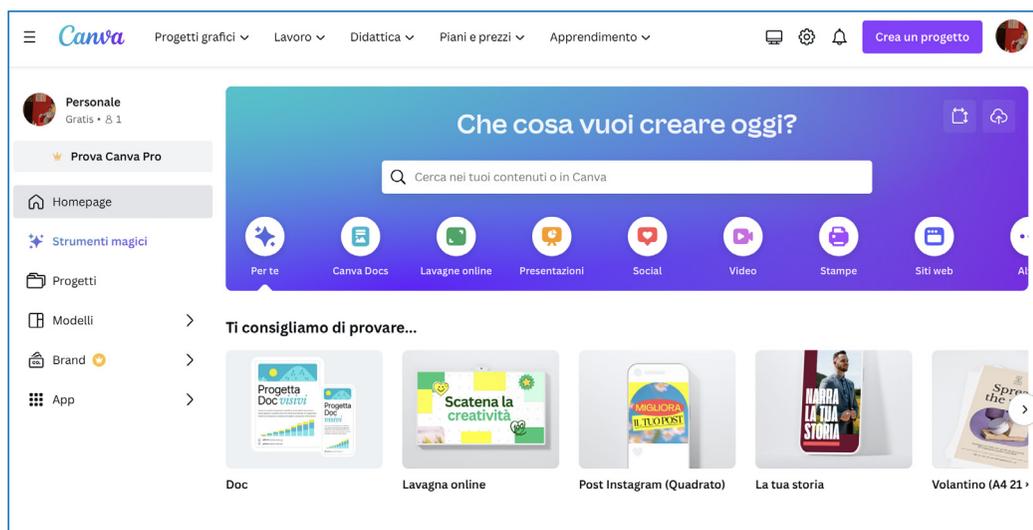
Canva

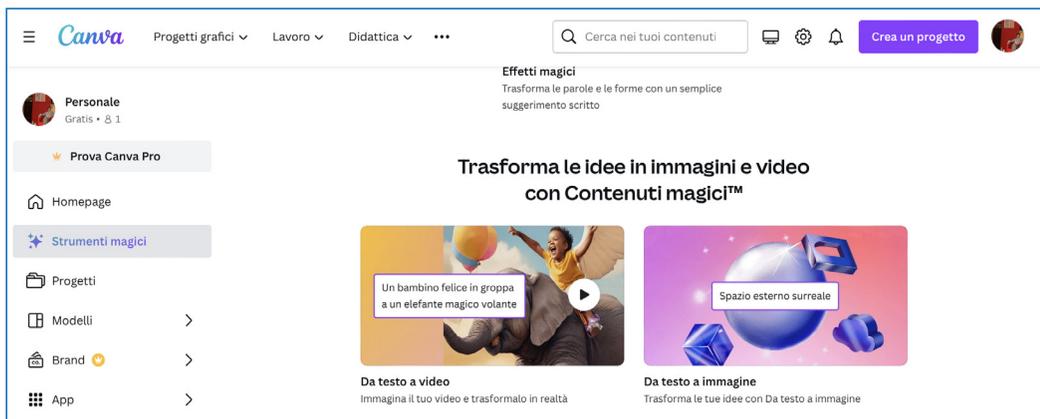
Canva è una piattaforma famosa per i suoi potenti strumenti di progettazione grafica.

Tra questi recentemente è stata inserita una IA *text-to-image* sviluppata a partire dal modello di deep learning *Stable Diffusion*, che è di pubblico dominio.

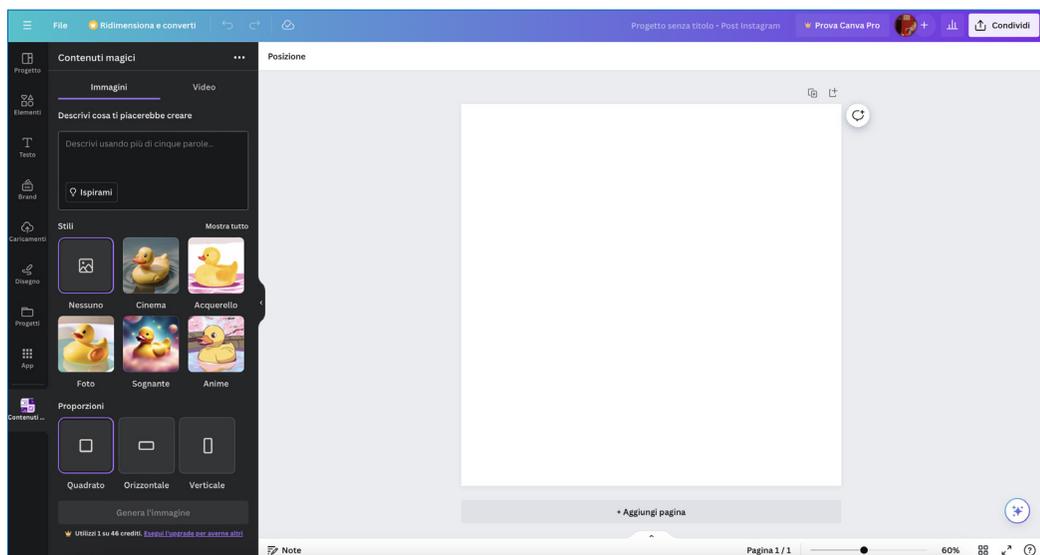
Per accedere a *Canva* occorre registrarsi. Il sito prevede un piano gratuito che offre molti strumenti, accessibili con alcune limitazioni.

Dopo aver effettuato l'accesso, *Canva* si presenta come nella schermata **a**. Per accedere al generatore di immagini clicchiamo su *Strumenti magici* e poi scorriamo verso il basso finché non troviamo il box *Da testo a immagine* su cui clicchiamo (schermata **b**).





Una volta entrati nello strumento di progettazione delle immagini, questo si presenta come possiamo vedere nella schermata **c**. Rispetto a *Copilot*, con *Canva* siamo di fronte a un vero strumento di progettazione grafica.



Possiamo assegnare un **titolo** al progetto compilando la casella *Progetto senza titolo*.

A sinistra, nella stessa barra blu, troviamo il menu *File*, tramite il quale possiamo accedere a funzioni come il salvataggio, l'importazione di immagini e le impostazioni.

La barra strumenti a sinistra contiene:

- lo spazio per il **prompt** alla voce *Descrivi cosa ti piacerebbe creare*;
- numerosi **stili** tra i quali scegliere (clicchiamo su *Mostra tutto* per visualizzarli tutti);
- la possibilità di scegliere le **proporzioni** dell'immagine generata;
- il pulsante che avvia la **generazione** dell'immagine.

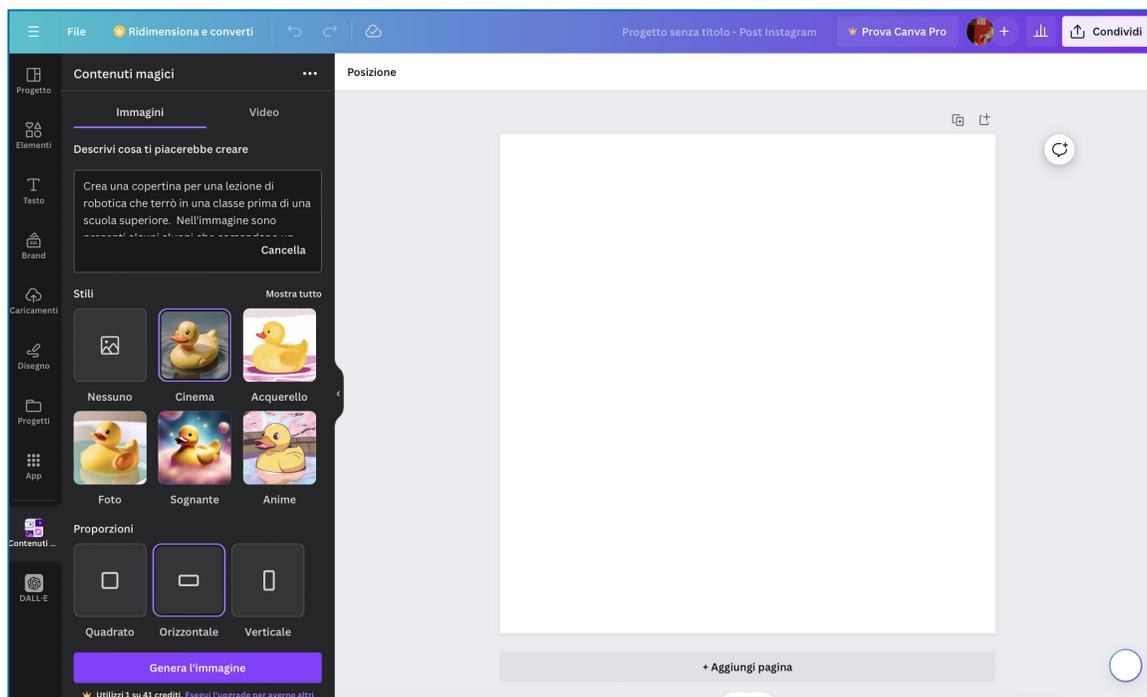
Per quanto riguarda i prompt, vigono regole e buone pratiche simili a quelle indicate per *Copilot*: la differenza saliente sta nel fatto che possiamo **omettere lo stile e le proporzioni nel prompt** e specificarli nelle sezioni *Stile* e *Proporzioni*.

Proviamo a usare *Canva* per creare l'immagine di copertina per il materiale di una lezione di robotica. Il prompt utilizzato è:

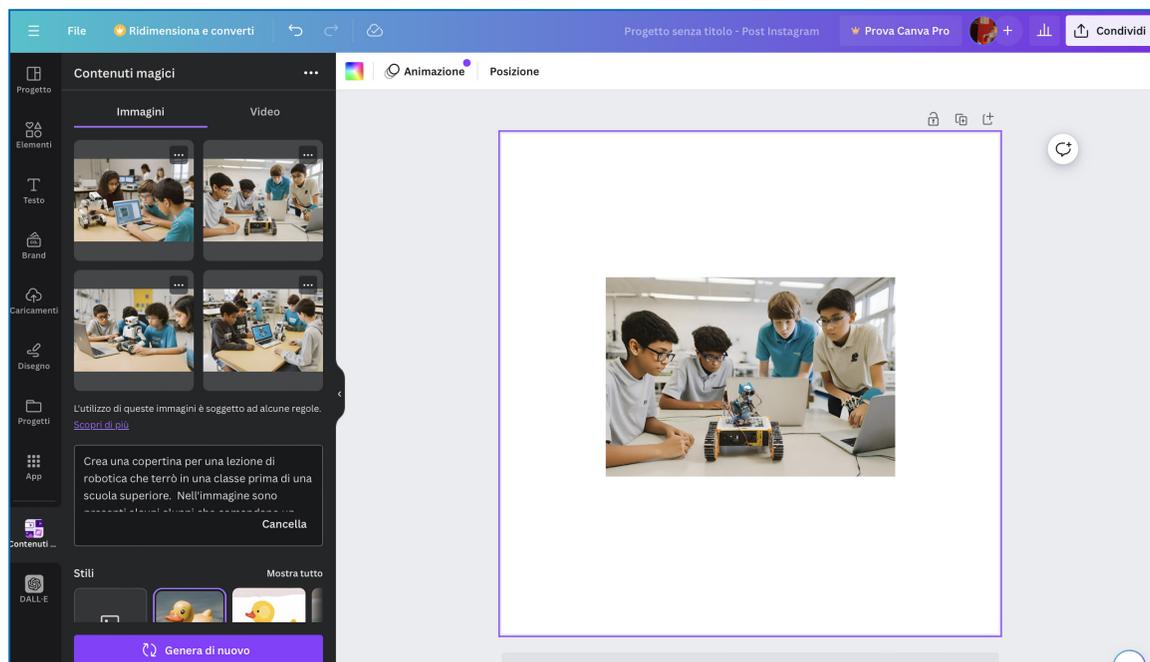
Crea una copertina per una lezione di robotica che terrò in una classe prima di una scuola superiore. Nell'immagine sono presenti alcuni alunni che comandano un piccolo robot mediante un laptop, all'interno di un luminoso laboratorio di robotica.

Lo stile selezionato è *Cinema* e tra le proporzioni abbiamo selezionato *Orizzontale*.

Al momento di premere il pulsante di generazione la schermata si presenta come in questa immagine:



Al termine della generazione, nella parte alta della barra degli strumenti, vediamo 4 immagini generate a partire dal prompt indicato. Cliccando su una o più di esse, queste vengono inserite nell'area di lavoro a destra.



A questo punto entrano in gioco tutti gli strumenti consueti di *Canva*, che ci permettono di modificare le immagini o trasformarle e inserirle all'interno dei nostri progetti.

Conclusioni

In conclusione, i modelli *text-to-image* rappresentano uno sviluppo tecnologico notevole nell'ambito dell'intelligenza artificiale, aprendo orizzonti inediti alla creatività digitale e alla generazione di contenuti. Nonostante il loro potenziale, questi modelli portano con sé una serie di problemi e limitazioni che non dobbiamo sottovalutare. Innanzitutto, la tendenza dei modelli a incorporare **bias** e **pregiudizi** esistenti nei dati di addestramento pone seri interrogativi sulla neutralità e l'equità delle immagini generate, riflettendo spesso **stereotipi** e visioni del mondo distorte.

Un'altro problema a cui prestare attenzione riguarda le potenziali violazioni del **copyright**. I modelli *text-to-image*, avendo appreso da vasti database di immagini esistenti, possono generare immagini che assomigliano molto a opere protette da diritto d'autore, sollevando questioni sulla proprietà intellettuale.

Ci sono anche problemi di natura etica che hanno rilevanza sociale, infatti la capacità di generare immagini realistiche alimenta il rischio di creare contenuti falsi, come **deepfake**, che possono avere conseguenze dannose nella disinformazione e nella manipolazione dell'opinione pubblica. Inoltre, la facilità con cui possono essere prodotti contenuti offensivi o inappropriati solleva questioni sulla responsabilità dei creatori di tali modelli e sulla necessità di regolamentazioni efficaci.

Quando usiamo queste tecnologie dobbiamo sempre assicurarci che il loro impiego da parte nostra e dei nostri alunni e alunne sia non solo innovativo, ma anche responsabile e rispettoso dei diritti e della dignità di tutte le persone.



Immagine realizzata con *ChatGPT4*: “Rappresenta una sintesi allegorica dei problemi associati ai modelli *text-to-image*, illustrando visivamente le sfide legate ai bias e ai pregiudizi, alle questioni di copyright, ai dilemmi etici e alla necessità di un dibattito aperto sull'uso responsabile di queste tecnologie “