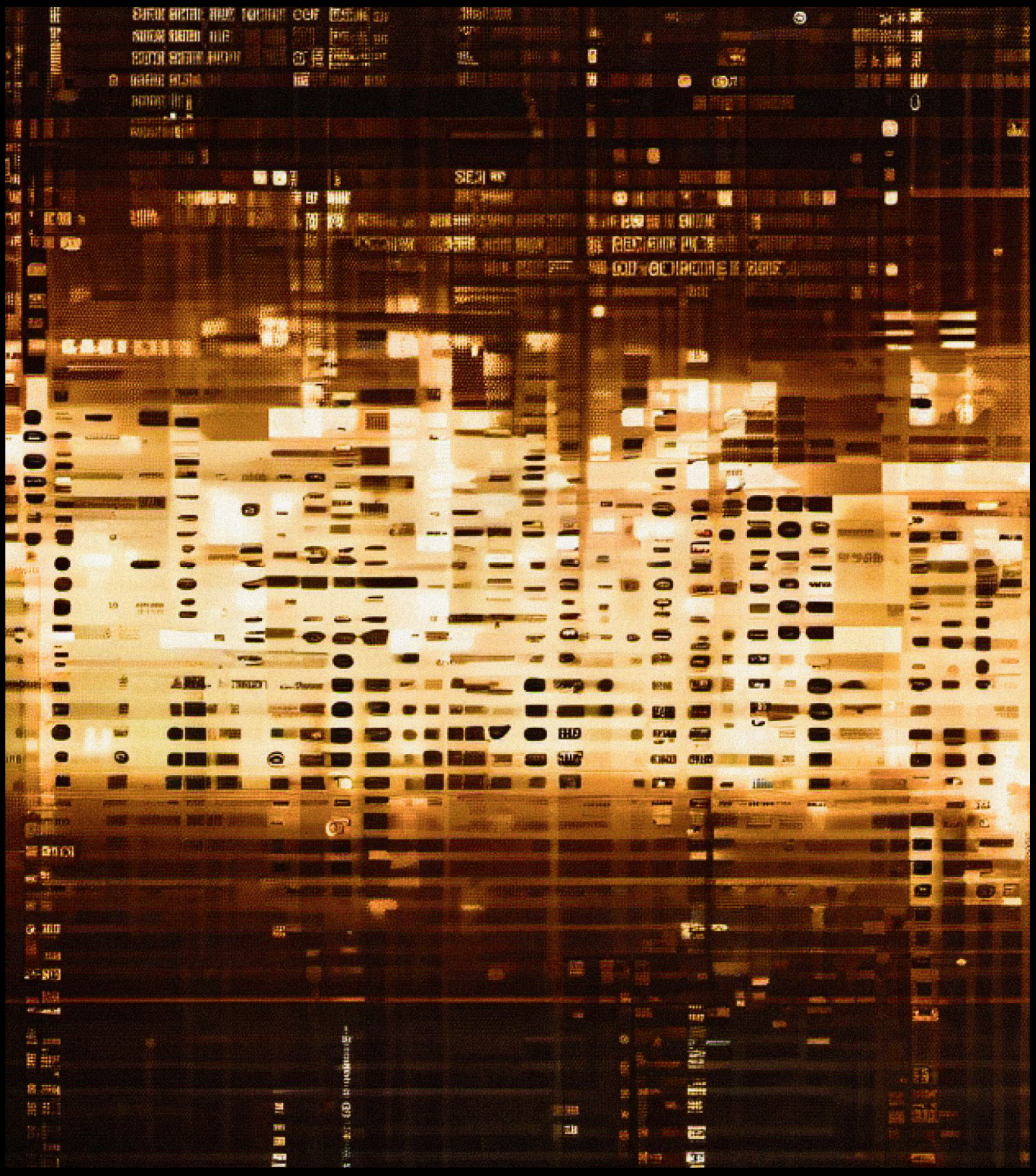
Intelligenza artificiale generativa e audiovisivo



un'indagine teorico-critica sulle immagini algoritmiche

Politecnico di Torino

Corso di laurea triennale
In Design e Comunicazione
A.a. 2025/2026
Sessione di Laurea Settembre 2026

Intelligenza artificiale generativa e audiovisivo:

un'indagine teorico-critica sulle immagini algoritmiche

Candidato: Denti Matteo
Relatore: Monetti Massimo

ABSTRACT

Lo sviluppo dell'intelligenza artificiale generativa sta trasformando la produzione, la circolazione e la percezione delle immagini audiovisive. La diffusione di modelli capaci di generare immagini e video sempre più realistici rende necessario interrogarsi sul futuro dell'industria cinematografica e dell'audiovisivo. In questo contesto, la tesi propone di leggere l'emergere di un possibile cinema algoritmico, nel quale le immagini non derivano esclusivamente dalla registrazione del reale, ma dalla generazione statistica di modelli addestrati su grandi dataset visivi. Il primo capitolo ricostruisce brevemente la storia dell'intelligenza artificiale e ne introduce i concetti fondamentali, distinguendo tra IA, machine learning, deep learning e reti neurali. Il secondo capitolo analizza l'IA generativa: foundation models, architetture VAE, GAN, Transformer e diffusion models, modalità di generazione audiovisiva e principali modelli video. Il terzo capitolo esamina gli impatti etici e sociali di queste tecnologie: la riproduzione di bias culturali e di genere nei dataset, le questioni di privacy e copyright, i deepfake e la tutela del lavoro creativo. L'analisi considera inoltre il consumo energetico e idrico dei datacenter, il lavoro invisibile del data labeling, le filiere delle risorse minerarie e la concentrazione del potere nelle grandi aziende tecnologiche. Il quarto capitolo affronta le trasformazioni dell'autorialità e dell'estetica nel cinema algoritmico, tra aspirazione al fotorealismo, rischio di omologazione stilistica e uso espressivo di glitch, deformazioni e allucinazioni. Attraverso casi studio e una sperimentazione pratica con ComfyUI, la tesi mostra come l'errore procedurale e l'imprevedibilità del modello possano diventare parte del processo creativo.

INDICE

Introduzione

CAPITOLO 1 – fondamenti dell'intelligenza artificiale

1.1 Breve storia dell'intelligenza artificiale
1.2 Definizione di intelligenza artificiale e le sue categorie
1.3 Machine learning, deep learning, reti neurali

CAPITOLO 2 – Intelligenza artificiale generativa

2.1 Defnizione di IA generativa
2.2 Architettura di modelli di IA generativa
2.3 Meccanismo funzionamento diffusion models
2.4 Le tipologie di generazione video
2.5 Principali modelli di IA generativa video presenti sul mercato

CAPITOLO 3 – Etica ed impatti delle immagini algoritmiche

3.1 Bias, pregiudizi e blackbox
3.2 Privacy, diritti e copywright nei dataset training: il caso LAION-5B
3.3 Il fenomeno dei deepfake
3.4 Costi ed impatti ambientali

CAPITOLO 4 – Autorialità ed estetica delle immagini algoritmiche

4.1 L'autore nell'era algoritmica
4.2 Le nuove forme del visibile
4.3 Open source vs closed source
4.4 Comfy-UI per creare contenuti generativi

CAPITOLO 5 – Conclusioni

Sitografi, bibliografia e Iconografia

Ringraziamenti



Introduzione

Lo sviluppo dell'intelligenza artificiale generativa sta trasformando la produzione, la circolazione e la percezione delle immagini audiovisive. La diffusione di modelli capaci di generare immagini e video **sempre più realistici** rende necessario interrogarsi sul futuro dell'industria cinematografica. In questo contesto, la tesi propone di leggere l'emergere di un possibile **cinema algoritmico**, nel quale le immagini non derivano esclusivamente dalla registrazione del reale, ma dalla generazione statistica di modelli addestrati su grandi dataset visivi.

Il lavoro si sviluppa **in quattro capitoli**, che ripercorrono la storia e **i concetti fondamentali dell'intelligenza artificiale**, ne analizzano le architetture generative e le modalità di produzione audiovisiva, ne esaminano le implicazioni etiche, sociali ed economiche, e ne indagano infine le conseguenze sull'autorialità e sull'estetica del cinema algoritmico, attraverso casi studio e una sperimentazione pratica con ComfyUI.

Con la seguente ricerca si cerca di analizzare la questione nel modo più aperto possibile, restituendo **l'argomento nella sua interezza e mostrandone le potenzialità quanto le criticità.** Non si tratta però di una trattazione neutra: il percorso che segue è inevitabilmente filtrato dalla mia sensibilità personale rispetto a un tema che, per sua natura, non ammette conclusioni univoche.

L'obiettivo di questa tesi è duplice: da un lato **fornire le conoscenze teoriche di base** necessarie a comprendere la generazione di contenuti mediante intelligenza artificiale, dall'altro sviluppare una **consapevolezza critica, di tipo etico ed estetico**, rispetto a questi strumenti, per poter guardare all'intelligenza artificiale con occhi più consapevoli.



Capitolo 1

Fondamenti dell'intelligenza artificiale

1.1 Breve storia dell'intelligenza artificiale

La storia e lo sviluppo dell'intelligenza artificiale non sono fenomeni recenti ma affondano le loro radici già nei **primi anni del Novecento**. Nel corso del tempo, l'intelligenza artificiale ha attraversato fasi alterne: periodi di forte entusiasmo e ingenti investimenti (come negli anni '50 e, in modo analogo, ai giorni nostri) si sono alternati a momenti di rallentamento, noti come gli **"inverni dell'intelligenza artificiale"**, caratterizzati da un calo di interesse e di finanziamenti.
Le principali tappe di sviluppo che hanno contribuito a definire l'intelligenza artificiale così come la conosciamo oggi sono le seguenti:

1943 -> La prima rete neurale artificiale

Lo psicologo cognitivo **Walter Pitts** e il cibernetico **Warren McCulloch** descrivono un sistema di neuroni artificiali capace di apprendere in modo analogo ai processi di apprendimento umani. La teoria che hanno proposto è il primo modello matematico di rete neurale artificiale, ed è tuttora **alla base degli algoritmi di deep learning**. In particolare gli autori dimostrano come il sistema possa apprendere autonomamente dai dati attraverso un meccanismo di tentativi ed errori che rafforza le connessioni corrette e indebolisce quelle errate. Tuttavia, a causa della mancanza di dati e dell'adeguata potenza computazionale dei computer, ne limitò l'applicazione pratica.

1950 -> Il test di Turing

Ideato dal matematico Alan Turing, il test consiste in un criterio volto a stabilire se una macchina sia in grado di pensare in modo analogo a un essere umano. Tale criterio è stato proposto da Turing nell'articolo **"Computing Machinery and Intelligence"**, pubblicato nel 1950 sulla rivista Mind. L'autore prende spunto da un esperimento noto come "gioco dell'imitazione", che coinvolge tre parteci-

panti, di cui uno, separato dagli altri due, deve identificare, tramite domande, chi sia l'uomo e chi la donna. Uno dei due interlocutori tenta di ingannarlo, mentre l'altro lo aiuta a giungere alla risposta corretta.

1956 -> La conferenza di Dartmouth

Il seminario avvenuto nel 1956 a Dartmouth che coinvolse un gruppo di ricercatori (John McCarthy, Marvin Minsky, Nathaniel Rochester e Claude Shannon) è considerato l'evento ufficiale che segna **l'inizio della ricerca nel campo dell'intelligenza artificiale.** Lo scopo della conferenza era lo studio di macchine capaci di compiere azioni considerate intelligenti; tale iniziativa fu presentata nel Dartmouth Proposal, documento in cui venne introdotto per la prima volta il termine intelligenza artificiale e furono **delineati i principali ambiti di ricerca**, tra cui le reti neurali, la teoria della computabilità e l'elaborazione del linguaggio naturale.

1966 -> ELIZA: il primo chatbot della storia

Ideato dall'informatico Joseph Weizenbaum, ELIZA è considerato il primo chatbot della storia. Il programma prende il nome da Eliza Doolittle e **simula uno psicoterapeuta** ed offre all'utente l'illusione di essere compreso attraverso la riformulazione degli input in frasi generiche o domande. Le sue capacità sono tuttavia molto limitate e si basano principalmente sulla ripetizione delle informazioni fornite dall'interlocutore.

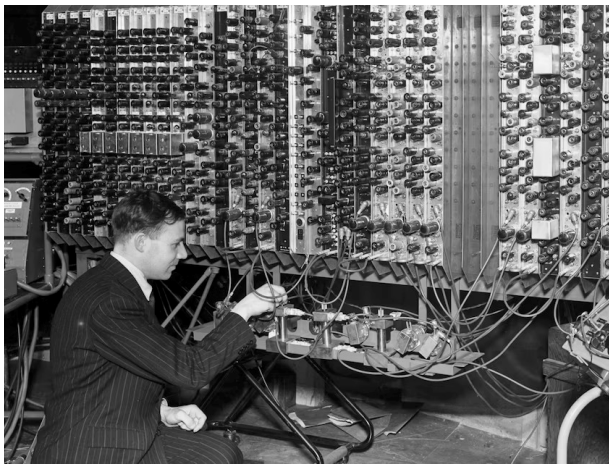


figura 1 *Alan Touring che ispeziona un prototipo della macchina computazionale automatica nel 1950, Londra.*



figura 2 *Alcuni dei ricercatori presenti alla conferenza di Dartmouth sull'intelligenza artificiale del 1956.*

Anni '70 –'80 -> Gli inverni delle Intelligenze artificiali

Gli "inverni dell'intelligenza artificiale" sono stati **periodi caratterizzati da sfiducia verso le IA**, con conseguente riduzione degli investimenti e rallentamento della ricerca nel campo dello sviluppo tecnologico. Il primo inverno dell'IA avvenne negli anni '70, quando i finanziamenti per la ricerca sull'IA diminuirono drasticamente e molti scienziati abbandonarono il campo. Ciò era dovuto al fatto che i primi sistemi di intelligenza artificiale non erano in grado di svolgere compiti complessi, limitandosi a operazioni semplici come il riconoscimento di oggetti o la comprensione di comandi basilari, deludendo le aspettative e facendo perdere fiducia nella tecnologia. Il secondo inverno si verificò tra la fine degli anni '80 e l'inizio degli anni '90, a causa di fattori economici come la crisi petrolifera, della crescente consapevolezza che molte promesse dell'IA non erano state mantenute e della tecnologia non ancora sufficientemente sviluppata. Questo portò a un calo dell'interesse per il settore e a ulteriori riduzioni dei finanziamenti.

1997 -> DeepBlue sconfigge scacchista Gary Kasparov

La svolta che segna l'uscita dal cosiddetto "inverno dell'intelligenza artificiale" avviene con la vittoria di Deep Blue, il **sistema sviluppato da IBM**, sul campione mondiale di scacchi Garry Kasparov. In seguito a questo evento, si iniziò a sostenere che gli scacchi potessero non rappresentare più un compito intrinsecamente "intelligente". L'episodio riportò l'attenzione sull'intelligenza artificiale.

2012 -> AlexNet vince la gara di riconoscimento immagini

Sviluppato dai dottorandi Alex Krizhevsky e Ilya Sutskever sotto la supervisione di Geoffrey Hinton, AlexNet è un **modello convoluzionale** che rappresentò un importante passo avanti nel **riconoscimento automatico delle immagini.** Nella ImageNet Challenge (competizione internazionale per testare sistemi di intelligenza artificiale nel riconoscimento di immagini) vinse con un margine netto rispetto agli altri partecipanti. La profondità del modello ha

dimostrato come il deep learning potesse superare drasticamente i metodi tradizionali nel riconoscimento delle immagini, grazie all'ampia disponibilità di dati e al crescente potere computazionale.

2016 -> Alphago sconfigge campione mondiale di Go

Il computer AlphaGo, **sviluppato da Google**, ha sconfitto il campione mondiale del gioco del Go (un antico gioco da tavolo cinese noto per la sua straordinaria complessità strategica), un risultato che fino ad allora sembrava impensabile, data l'estrema complessità e astrattezza del gioco. Questo evento ha **messo in discussione molti preconcetti**, suggerendo che alcune facoltà considerate tipicamente umane potrebbero non essere così uniche come si è a lungo creduto.

2020 -> OpenAI presenta GPT-3

OpenAI presenta GPT-3, un **sistema di elaborazione del linguaggio naturale** addestrato su 800 gigabyte di dati e dotato di una rete neurale digitale paragonabile a quella neurale biologica. GPT-3 mostra capacità di generazione di testi mai viste, segnando l'inizio dell'era delle intelligenze artificiali generative. Due anni più tardi viene lanciato **ChatGPT**, il chatbot basato su GPT-3, rendendo l'IA accessibile a un vasto pubblico.



figura 3 *Garry Kasparov durante una partita contro Deep Blue, il sistema di gioco degli scacchi sviluppato da IBM, 1997.*



figura 4 *AlphaGo sconfigge Ke Jie, campione mondiale di Go durante la seconda partita del Future of Go Summit a Wuzhen, 2017.*

1.2 Definizione di intelligenza artificiale e le sue categorie

Per comprendere appieno il funzionamento di questa tecnologia è fondamentale **partire dalle definizioni di base** e approfondire alcuni concetti chiave che ne costituiscono le fondamenta. Intelligenza artificiale, Machine Learning, Deep Learning e reti neurali non sono sinonimi: indicano concetti diversi tra loro.

Secondo la definizione della Treccani l'intelligenza artificiale (spesso abbreviata in IA) è ***"la disciplina che studia se e in che modo si possano riprodurre i processi mentali più complessi mediante l'uso di un computer".*** In termini pratici con questo termine si intendono macchine o algoritmi progettati per imitare comportamenti e facoltà umane. L'IA ha iniziato a influenzare la nostra vita e il nostro modo di pensare ben prima della diffusione delle intelligenze artificiali generative, come i modelli linguistici o di generazione di immagini. Essa è infatti presente in numerosi ambiti della nostra quotidianità, tra cui:
- **Sistemi di automazione industriale**, come le braccia meccaniche;
- **Software di guida assistita e navigazione** nelle automobili moderne;
- **Sistemi di riconoscimento vocale e facciale**, per lo sblocco dello smartphone e comandi vocali;
- **Algoritmi di selezione e gestione dei contenuti** delle piattaforme social;
- **Gli assistenti virtuali** come Siri e Alexa;
- **Sistemi di raccomandazione** per le piattaforme di servizi streaming e e-commerce;
- I personaggi dei videogiochi controllati dall'intelligenza artificiale **(NPC)**.

Inoltre, l'intelligenza artificiale può essere classificata in diverse tipologie in base al livello di capacità cognitive e di autonomia dei sistemi.

Si distinguono tre tipologie di intelligenza artificiale:

L'IA debole (o ristretta) comprende sistemi progettati per svolgere compiti e **risolvere problemi specifici** ma che **non possiedono alcun grado di coscienza**. Si tratta della tipologia di IA più diffusa oggi: addestrata su compiti circoscritti, spesso raggiunge una rapidità e una precisione superiori a quelle umane.Le altre due forme di intelligenza artificiale sono attualmente di natura teorica.

L'AGI (Artificial General Intelligence) indica un sistema dotato di **capacità cognitive paragonabili a quelle umane**, in grado di apprendere, ragionare, comprendere e adattarsi a qualsiasi contesto o dominio.

La Super AI o ASI (Artificial Superintelligence) indica un **livello di intelligenza superiore a quello umano** in quasi tutti gli ambiti, dalla creatività alla ricerca scientifica fino alle abilità sociali. Una simile forma di intelligenza sarebbe dotata di una coscienza propria, **capace di pensare ed esprimere giudizi** in modo autonomo.

1.3 Machine learning, deep learning, reti neurali

Il Machine Learning (abbreviato ML, in italiano "apprendimento automatico) è una **branca dell'intelligenza artificiale** e si fonda su algoritmi in grado di **apprendere dai dati e migliorare le proprie prestazioni nel tempo**, senza necessitare di una programmazione esplicita per ciascun compito. Grazie a questa **capacità di riconoscere schemi e pattern**, i modelli di machine learning possono prendere decisioni o formulare previsioni in modo autonomo, basandosi esclusivamente sui dati

disponibili.
Alla base dell'**apprendimento automatico** vi sono **due principi fondamentali** che ne guidano il funzionamento:

- **Apprendimento dai dati:** il sistema è in grado di acquisire competenze, abilità e conoscenze attraverso un processo di allenamento sui dati.

- **Processo di allenamento:** durante questa fase, lo sviluppatore non definisce manualmente ogni regola per risolvere un compito specifico ma programma l'algoritmo affinché possa **apprendere autonomamente dai dati forniti.** L'efficacia di tale processo è stata notevolmente incrementata dalla disponibilità di grandi quantità di dati e dall'aumento della potenza computazionale dei computer.

L'apprendimento automatico si distingue **in tre tipologie principali**: supervisionato, non supervisionato e per rinforzo. In tutti e tre i casi il principio comune consiste nel fornire dati all'algoritmo senza indicargli esplicitamente come risolvere il problema: il modello deve imparare autonomamente a estrarre informazioni e schemi di dati.

L'Apprendimento supervisionato prevede l'uso di set di **dati etichettati** per addestrare l'algoritmo a classificare informazioni o a prevedere risultati. Per ogni input è noto il risultato atteso, consentendo al modello di apprendere tramite esempi, come nel caso del riconoscimento facciale.

Nell'**Apprendimento non supervisionato** l'algoritmo lavora su dati non etichettati, senza esempi predefiniti, con l'obiettivo di **identificare schemi o cluster nascosti** all'interno dei dati, ad esempio nell'analisi delle abitudini degli utenti, per suggerire contenuti personalizzati.

L'**Apprendimento per rinforzo** è il tipo di apprendimento **ispirato al modo con cui gli esseri umani apprendono** tramite il sistema di **penalità e ricompensa.** L'algoritmo interagisce con un ambiente e riceve feedback numerici, positivi o negativi, in funzione delle azioni compiute, al fine di massimizzare la ricompensa cumulativa nel lungo periodo. Ad esempio, un'applicazione può essere nei videogiochi, dove si può lasciare giocare un'IA a mi-

gliaia di partite, partendo da mosse casuali; ogni volta che fa una mossa giusta ottiene un premio, mentre ogni volta che perde riceve una penalità. Con il tempo impara delle strategie che massimizzano i punti. Lo stesso principio si applica anche all'addestramento di robot.

Per il machine learning vengono impiegati principalmente **dati strutturati e facilmente gestibili**, come matrici numeriche o informazioni organizzate in tabelle. L'algoritmo utilizza questi dati per identificare pattern e comportamenti degli utenti, consentendo di personalizzare le pubblicità in modo mirato.
Il machine learning trova applicazione concreta in diversi ambiti: nei **sistemi di raccomandazione per contenuti su misura** (come su Netflix), nell'**analisi delle preferenze di mercato** per individuare trend e comportamenti dei consumatori e nel filtraggio dello spam, dove classifica automaticamente i messaggi legittimi da quelli indesiderati.

Il Deep Learning (abbreviato in DL, in italiano "apprendimento profondo") rappresenta un **sottoinsieme del Machine Learning** ed è basato su **reti neurali profonde**, caratterizzate da numerosi strati nascosti, consentono di elaborare dati molto più articolati rispetto agli algoritmi tradizionali di Machine Learning. Offre un **grado di astrazione superiore** e la capacità di affrontare compiti più complessi.

Alla base del deep learning vi sono i seguenti principali concetti:
– Gestione di dati indeterministici: il sistema è in grado di gestire **informazioni non strutturate**, come immagini o segnali audio, che non si prestano ad una rappresentazione tabellare.
– Astrazione elevata: le reti profonde **individuano schemi e relazioni nei dati** che sfuggirebbero a modelli più semplici, permettendo di affrontare compiti che richiedono un'elaborazione più sofisticata.

I modelli di Deep Learning costituiscono oggi la base dell'intelligenza artificiale più avanzata. Alcuni esempi di applicazione:
- **GAN (Generative Adversarial Networks)**: utilizzate

per la creazione di deepfake e per l'upscaling di immagini e video.
- **Generazione di immagini e video**: piattaforme come Midjourney, Stable Diffusion e Gen3.
- **Generazione di testo**: chatbot basati su grandi modelli linguistici (LLM), come ChatGPT.
- **Riconoscimento facciale**: i modelli elaborano dati ad alto livello di astrazione, consentendo l'identificazione accurata dei volti in contesti complessi.

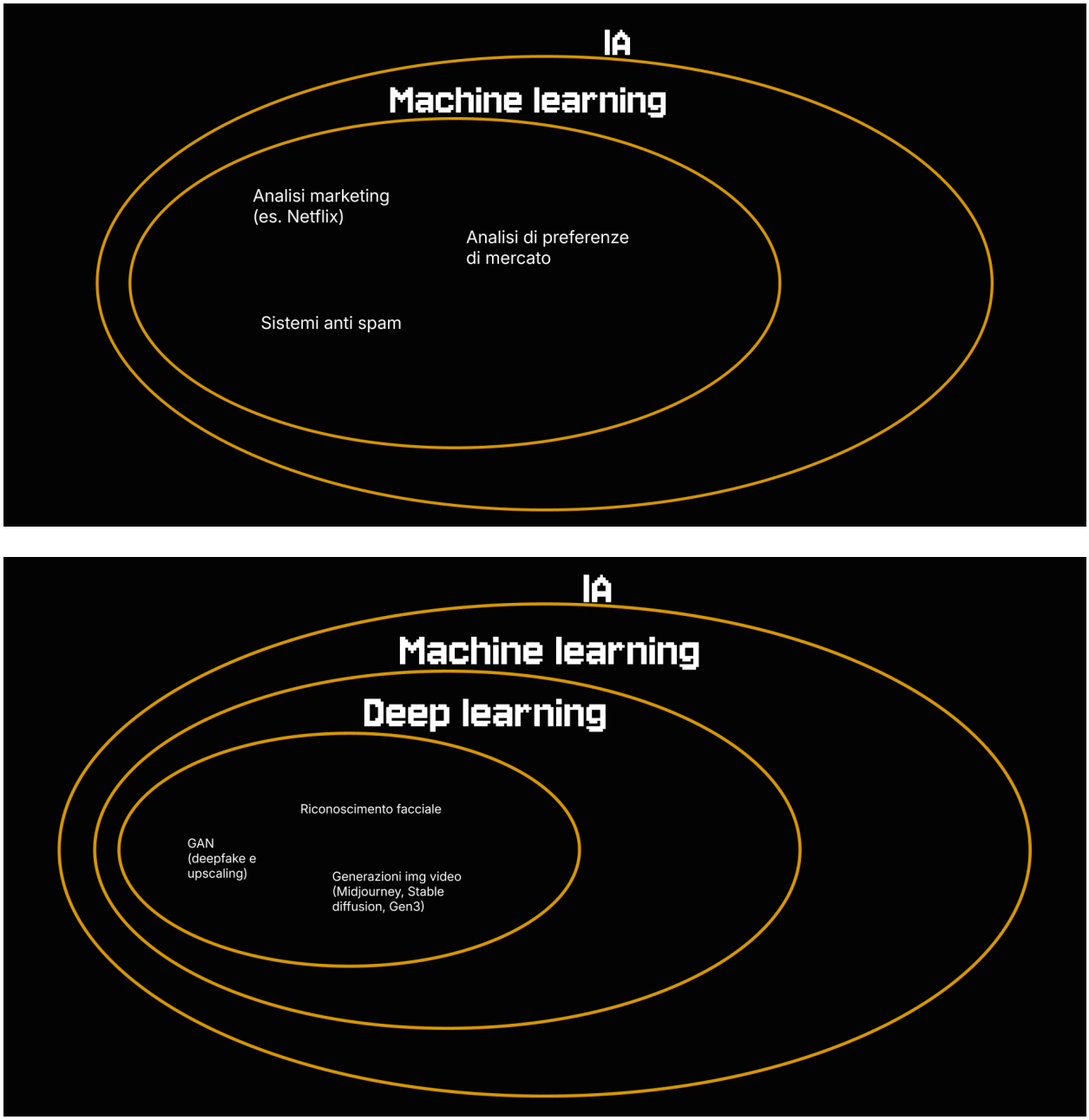


figura 5 e 6 *Rappresentazione schematica della relazione gerarchica tra intelligenza artificiale, machine learning e deep learning, con alcuni esempi di applicazione.*

Le reti neurali sono **modelli computazionali** ispirati alle reti di neuroni biologici, progettati per emulare in modo semplificato il funzionamento del cervello. Esse costituiscono la **base del Deep Learning.**

Un **neurone artificiale** si ispira alla struttura dei neuroni biologici e rappresenta l'unità fondamentale di una rete neurale. Una rete neurale artificiale è costituita da neuroni artificiali, chiamati **nodi**, organizzati in strati.
Lo strato di input riceve i dati iniziali in ingresso (come numeri, pixel di un'immagine, parole codificate).
Gli strati nascosti elaborano le informazioni, estraendo caratteristiche e relazioni dai dati.
Lo strato di output produce il risultato finale (come una classificazione o previsione).

Il funzionamento complessivo della rete avviene tramite la **propagazione in avanti (feedforward)** in cui l'input fornito alla rete viene elaborato attraverso gli strati nascosti fino a produrre il risultato finale.
I neuroni di uno strato sono connessi ai neuroni dello strato successivo tramite **connessioni pesate (pesi sinaptici)**, che determinano quanto una rete neurale consideri importante un'informazione che passa da un neurone all'altro.
Il peso sinaptico corrisponde ad un **valore numerico**: quando un segnale arriva al neurone successivo, viene moltiplicato per quel numero: se il peso è alto il segnale viene amplificato, se invece il peso è basso o negativo, il segnale viene attenuato o ridotto. I pesi consentono alla rete di attribuire maggiore o minore importanza alle informazioni, filtrando i dati rilevanti.

Questi segnali vengono quindi elaborati da una funzione di attivazione, dotata di una **soglia di attivazione**, che decide se il neurone deve trasmettere il segnale al livello successivo. Questo meccanismo consente alla rete di trasformare i dati grezzi di input nel risultato finale.

Una rete neurale si definisce profonda **(Deep Neural Network)** quando è costituita da **numerosi strati nascosti.** In questo caso, ogni strato non vede i dati come un'immagine percepita dall'occhio umano, ma come numeri.

Ogni strato costruisce una rappresentazione sempre più astratta dei dati, aumentando progressivamente il livello di astrazione e la capacità della rete di affrontare compiti complessi.

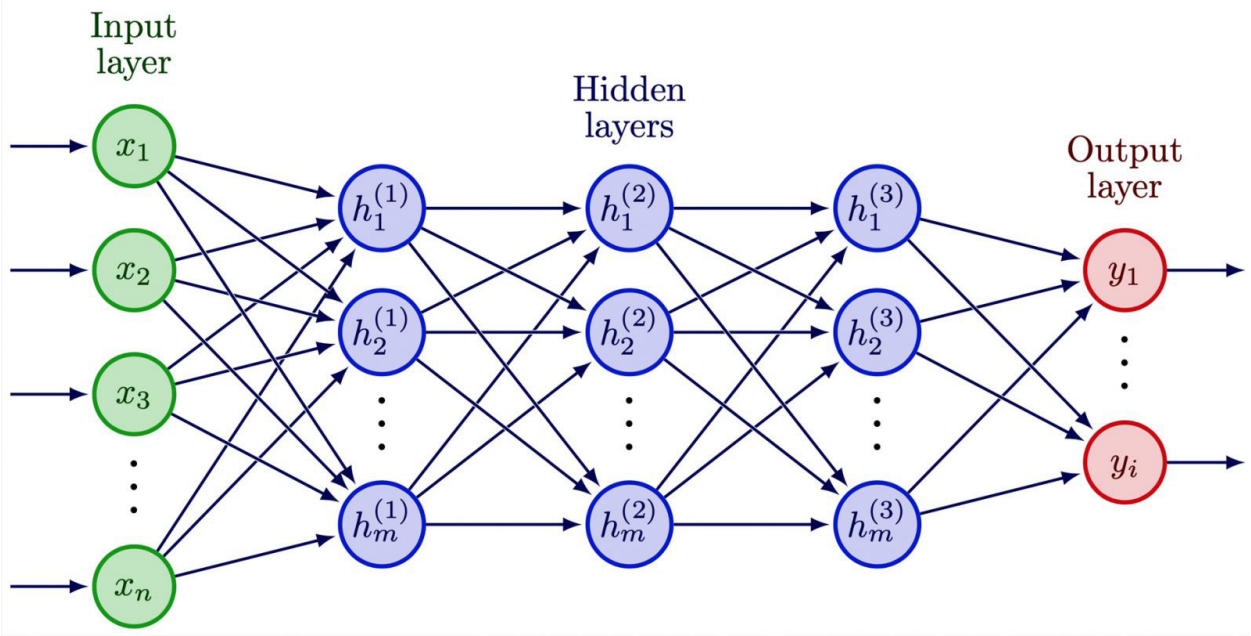


figura 7 *Rappresentazione schematica di una possibile architettura di rete neurale artificiale feed-forward. La rete è composta da uno strato di input con n variabili, tre livelli nascosti (hidden layers) formati da n neuroni artificiali e uno strato di output con i valori in uscita. Il pedice identifica la posizione del neurone all'interno del rispettivo livello, mentre l'apice tra parentesi ne indica il livello di appartenenza.*

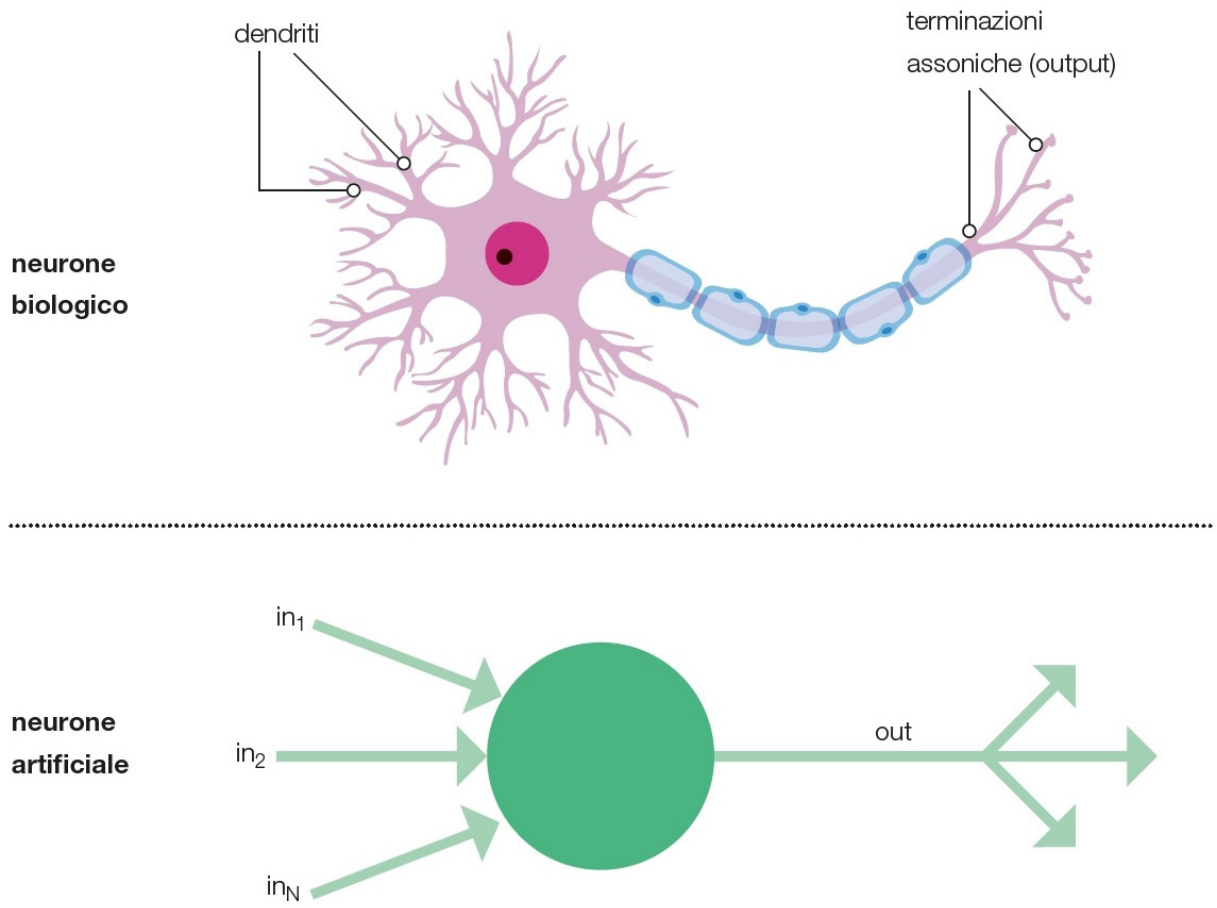


figura 8 *Confronto tra la struttura di un neurone biologico e quella di un neurone artificiale.*

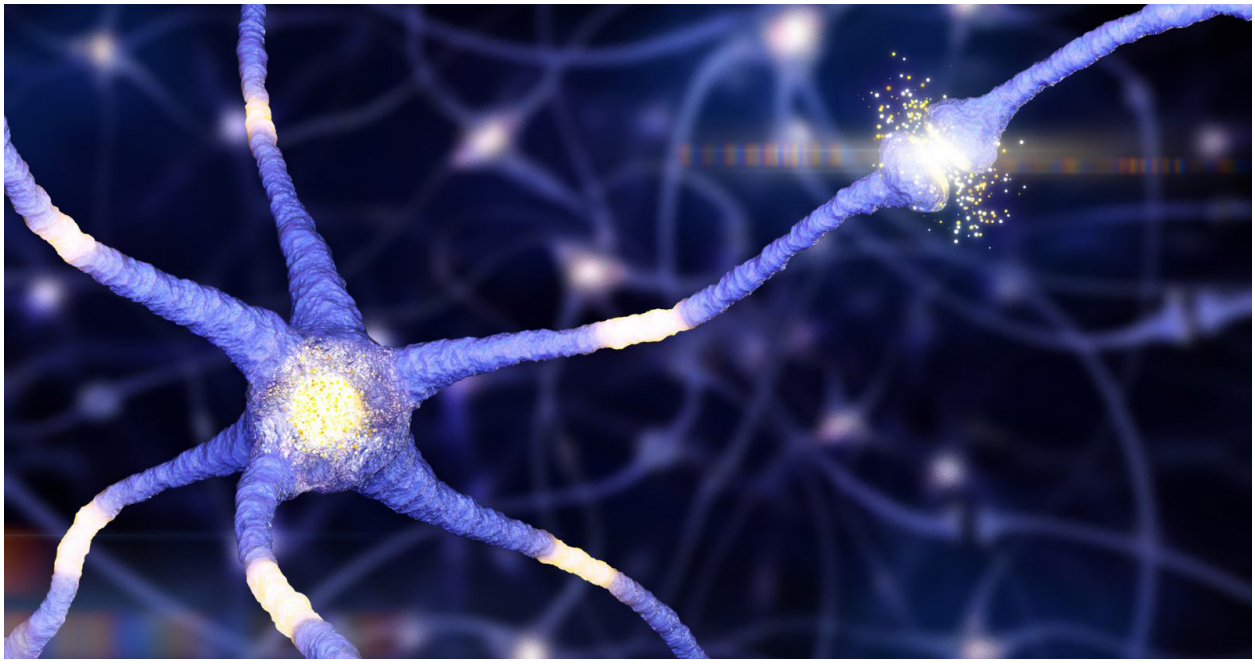


figura 9 *Ilustrazione dei neuroni del cervello umano.*



Capitolo 2

Intelligenza artificiale generativa

2.1 Definizione intelligenza artificiale generativa

L'intelligenza artificiale generativa (chiamata anche genIA) è una **branca dell'intelligenza artificiale** in grado di creare contenuti originali, come testi, immagini, video, audio o codice software, a partire da un prompt fornito dall'utente. Il suo funzionamento si basa su **modelli di deep learning**, addestrati a riconoscere pattern in grandi quantità di dati e a riutilizzarli per generare output coerenti con la richiesta ricevuta.

Al centro di questo funzionamento si trova il concetto di **foundation model**: modelli di deep learning di grandi dimensioni, addestrati su **dataset eterogenei e su scala molto ampia**, che imparano a rappresentare in modo generale il linguaggio, le immagini, l'audio o più modalità insieme.
A differenza dei modelli tradizionali, costruiti per un singolo compito, i foundation models non sono limitati a un'unica funzione: possono essere **adattati a molteplici compiti** tramite un ulteriore **addestramento leggero (fine-tuning)** o semplicemente con l'uso di prompt. Dopo l'addestramento, costituiscono una base comune ("foundation") su cui sviluppare diverse applicazioni.

Oggi i foundation model più comuni sono i modelli linguistici di grandi dimensioni **(LLM, Large Language Models)**, creati per applicazioni di generazione di testo, ma esistono anche foundation model per la generazione di immagini, video, audio e musica, nonché foundation model multimodali in grado di supportare diversi tipi di contenuti.

Alcuni esempi di foundation modello sono:
- **GPT**: generazione e comprensione del linguaggio;
- **BERT**: comprensione del linguaggio naturale;
- **CLIP:** elaborazione combinata di immagini e testo;
- **Stable Diffusion:** generazione di immagini;

- **Whisper:** riconoscimento vocale.

Il ciclo di vita di un modello generativo si può ricondurre a tre momenti: addestramento iniziale, adattamento a un uso specifico e allenamento continuo in fase di utilizzo (IBM, 2024).

Infatti, un foundation model, per quanto potente, non è immediatamente pronto per un'applicazione specifica: deve prima essere calibrato. Questo avviene principalmente in due modi complementari:

– Fine-tuning: il modello viene ulteriormente **addestrato su dati etichettati specifici** per il compito desiderato, così da specializzare le risposte su prompt e output attesi

– RLHF (Reinforcement Learning from Human Feedback): il modello viene affinato tramite **feedback diretto degli utenti,** che ne migliora progressivamente la precisione e la pertinenza.

Va notato che, mentre l'addestramento iniziale di un foundation model richiede **risorse computazionali enormi** e viene ripetuto raramente, l'allenamento tramite fine-tuning e RLHF è un processo molto più leggero e frequente, spesso condotto su base settimanale per correggere errori o adattarsi a nuovi casi d'uso.

Una volta distribuito un modello generativo non smette di essere osservato: sviluppatori e utenti ne valutano costantemente gli output, individuando errori o cali di qualità che richiedono correzioni mirate. Questo **meccanismo di verifica e correzione continua** consente a un'applicazione di restare affidabile nel tempo, anche quando il foundation model sottostante non viene aggiornato.

Una tecnica che si inserisce in questo processo di miglioramento è la **retrieval-augmented generation (RAG)**: anziché limitarsi alle informazioni apprese durante l'addestramento, il modello viene messo in condizione di **accedere a fonti esterne aggiornate**, integrando queste informazioni con le proprie rappresentazioni interne. In questo modo è possibile ridurre il rischio di risposte obsolete o imprecise, senza dover riaddestrare l'intero foundation model.

2.2 Architettura di modelli di IA generativa

Quando si parla di architettura di un modello di intelligenza artificiale, ci si riferisce alla **struttura e all'organizzazione interna** del modello: il tipo di unità computazionali che lo compongono, come sono disposte in strati, come sono collegate tra loro e in che modo elaborano i dati in ingresso. L'architettura definisce quindi le capacità, le prestazioni e i compiti per cui un modello può essere utilizzato.

Nel campo dell'IA generativa **quattro famiglie di architetture** hanno avuto un ruolo determinante, ciascuna con un principio di funzionamento diverso: gli **autoencoder** variazionali, che comprimono e ricostruiscono i dati passando per uno spazio numerico compresso; le **reti generative avversarie**, basate su una competizione interna tra due reti; i **Transformer**, che hanno rivoluzionato l'elaborazione di sequenze complesse grazie al meccanismo di attenzione; e i **modelli di diffusione**, oggi lo standard per la generazione di immagini fotorealistiche. Vale la pena analizzarle singolarmente, partendo dalla più semplice concettualmente.

Gli **Autoencoder variazionali (VAE, Variational Autoencoder)** funzionano attraverso **due reti neurali connesse** che operano in sequenza: un **encoder**, che riceve i dati in ingresso e li comprime in una rappresentazione numerica compatta, e un **decoder**, che compie il percorso inverso, ricostruendo il contenuto originale a partire da tali numeri.
Tra encoder e decoder si colloca lo **spazio latente**, costituito da variabili latenti che catturano le caratteristiche principali dei dati senza riprodurre direttamente l'immagine originale. Questo meccanismo consente alla rete di apprendere rappresentazioni compatte e significative dei dati.
I VAE trovano applicazione in diversi ambiti, come il rico-

noscimento facciale, il rilevamento di anomalie, la compressione dei dati e la rimozione del rumore dalle immagini. Un aspetto rilevante di questa architettura è che l'IA non "vede" o "legge" il contenuto originale: tutte le informazioni vengono trasformate in numeri.

Le **Reti generative avversarie (GAN, Generative Adversarial Network)** adottano una strategia radicalmente diversa: invece di un'unica rete che impara a ricostruire dati, ne mettono in campo due che competono tra loro. Una GAN infatti è composta da due reti neurali profonde in competizione: il **generatore**, che crea dati sintetici (immagini, video, audio) a partire da input casuali con l'obiettivo di ingannare il discriminatore, e il **discriminatore**, che valuta se i dati ricevuti sono reali o generati. Il processo di addestramento è iterativo e avversario: il generatore produce campioni sintetici, il discriminatore li confronta con dati reali e entrambi aggiornano i propri pesi tramite retropropagazione. Nel tempo, il generatore impara a produrre dati sempre più realistici, mentre il discriminatore migliora nella distinzione tra dati reali e sintetici. Questo meccanismo è alla base dei deepfake, in cui la continua competizione tra generatore e discriminatore consente di creare immagini facciali sintetiche quasi indistinguibili da quelle reali.

I Transformer, introdotti per la prima volta nel 2017 da Ashish Vaswani e colleghi, hanno segnato una rottura con il paradigma encoder-decoder tradizionale, diventando l'architettura di riferimento per gran parte degli strumenti di AI generativa più avanzati ad oggi, tra cui ChatGPT, GPT-4, Copilot, BERT, Gemini e Midjourney.
Il principio chiave dei Transformer è il **meccanismo di attenzione**, che consente al modello di individuare, all'interno di una sequenza di dati, quali parti siano più rilevanti per il compito da svolgere. Questo approccio permette di elaborare intere sequenze contemporaneamente, acquisire il contesto dei dati nella sequenza e codificarli in rappresentazioni numeriche (embedding o iperparametri) che catturano sia il contenuto sia le relazioni contestuali. Il risultato è una capacità di generazione testuale e di formulazione di risposte nettamente superiore, in termini di coerenza e precisione, rispetto alle architetture generative precedenti.

I modelli di diffusione (diffusion models) rappresentano l'architettura più recente tra quelle qui presentate, e oggi **predominante nella generazione di immagini**, è quella dei modelli di diffusione. Il loro funzionamento si basa su un principio controintuitivo: il modello viene addestrato a diffondere progressivamente campioni introducendo **rumore casuale**, per poi imparare a invertire questo processo e ricostruire immagini di alta qualità a partire dai dati degradati. Lo spazio latente viene così trasformato, passo dopo passo, consentendo al modello di catturare dettagli complessi e caratteristiche realistiche dei contenuti generati.

I modelli di diffusione sono alla base di molti strumenti text-to-image popolari, tra cui **Stable Diffusion** di Stability AI, DALL-E di **OpenAI**, **Midjourney** e **Imagen** di Google. Rispetto ad altre architetture, come gli autoencoder variazionali (VAE) o le GAN, i modelli di diffusione offrono prestazioni migliori e maggiore stabilità nella sintesi di immagini, consentendo di generare contenuti visivi più realistici e coerenti con i prompt forniti.

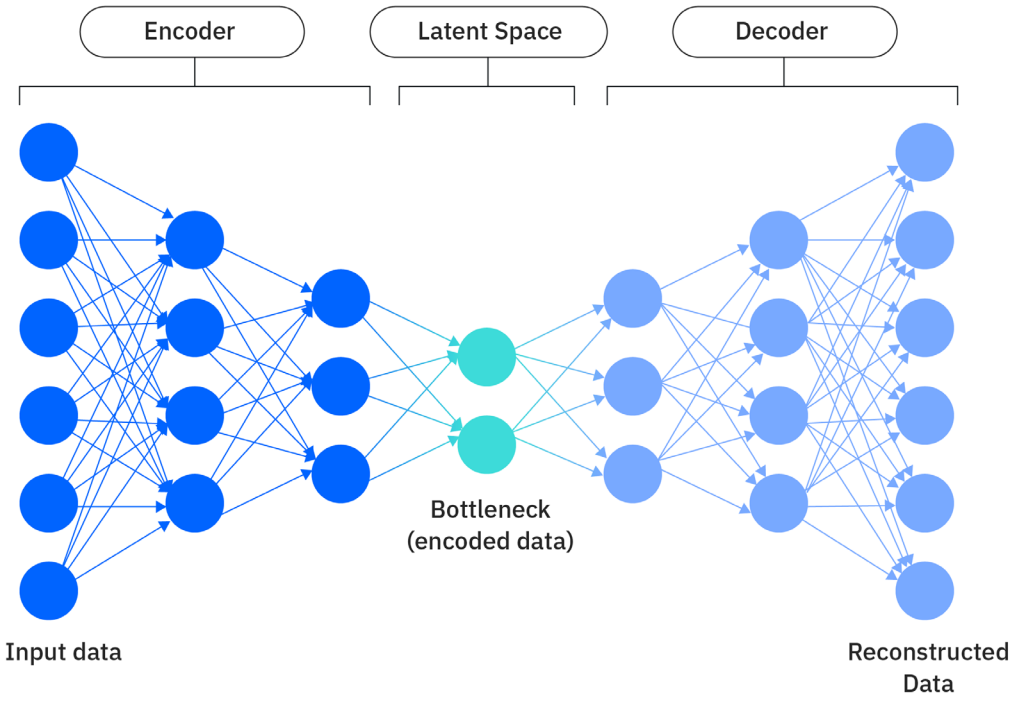


figura 10 *Rappresentazione schematica dell'architettura di un autoencoder: l'encoder comprime i dati di input in una rappresentazione nello spazio latente, mentre il decoder ricostruisce i dati a partire da tale rappresentazione.*

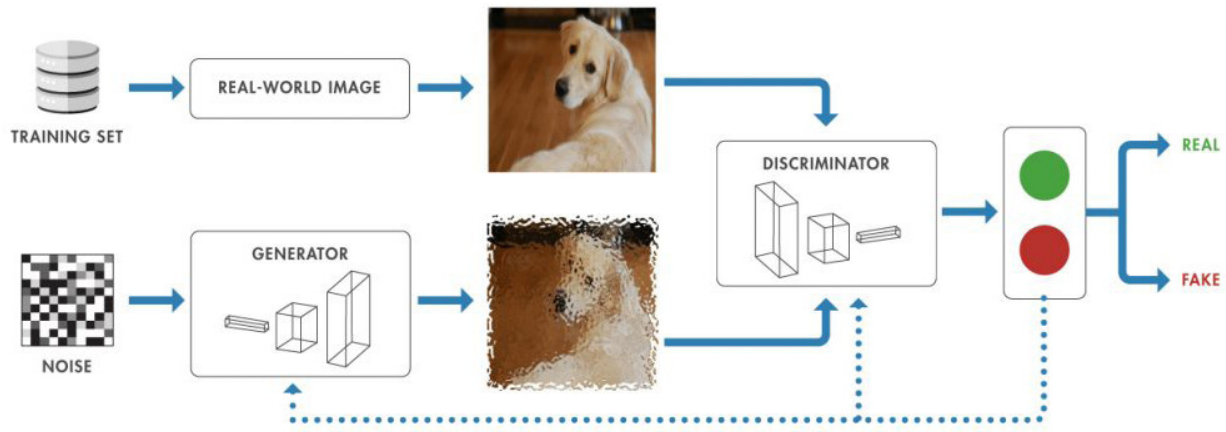


figura 11 *Rappresentazione schematica di una Generative Adversarial Network (GAN): il generatore produce immagini sintetiche a partire da rumore casuale, mentre il discriminatore le confronta con immagini reali provenienti dal dataset e apprende a classificarle come reali o artificiali.*

2.3 Meccanismo funziona-mento diffusion models

Il funzionamento dei modelli di diffusione si comprende meglio partendo da un'analogia fisica: l'idea di base è quella di insegnare al modello a generare dati realistici partendo da un input completamente casuale e progressivamente "pulito", invertendo il **processo di diffusione**. Durante l'addestramento, il modello apprende a degradare i dati reali aggiungendo **rumore casuale** in più passaggi e, contemporaneamente, a invertire questo processo, ossia a rimuovere il rumore e ricostruire i dati originali. È in questo **doppio movimento di degradazione e successiva ricostruzione** che si determina l'intero processo generativo. Questo doppio movimento si articola concretamente in due fasi distinte e complementari.

Nella **fase di forward diffusion** il modello riceve i dati originali e **li corrompe gradualmente**, introducendo rumore in una sequenza di passaggi controllati, fino a ottenere una rappresentazione praticamente indistinguibile da un rumore puro. Questa fase **serve al modello a imparare**, passo dopo passo, quali trasformazioni avvengono tra un'immagine nitida e la sua versione degradata.

Nella **fase di reverse diffusion** avviene la parte più rilevante dal punto di vista generativo: il modello applica in senso inverso ciò che ha appreso nella fase precedente, partendo da rumore casuale e **"pulendolo" progressivamente** fino a ottenere un'immagine coerente e definita. Questo meccanismo consente di produrre immagini ad alta qualità, coerenti con i prompt forniti, e permette di controllare caratteristiche specifiche del contenuto generato intervenendo sullo spazio latente.

Il vantaggio principale di questa architettura, rispetto a GAN e VAE trattati in precedenza, risiede proprio nella **natura iterativa e graduale** del processo di ricostruzione: an-

ziché generare un'immagine in un unico passaggio, il modello la costruisce attraverso **decine o centinaia di piccoli aggiustamenti successivi**, ottenendo generalmente una maggiore stabilità e fedeltà nella sintesi di dettagli complessi. È questa caratteristica a spiegare perché i diffusion models si siano affermati come architettura di riferimento non solo per la generazione fotorealistica di immagini, ma anche per compiti più ampi di computer vision e per applicazioni multimodali che combinano testo e immagine, come si vedrà nella sezione dedicata ai principali modelli text-to-image e text-to-video attualmente disponibili sul mercato.

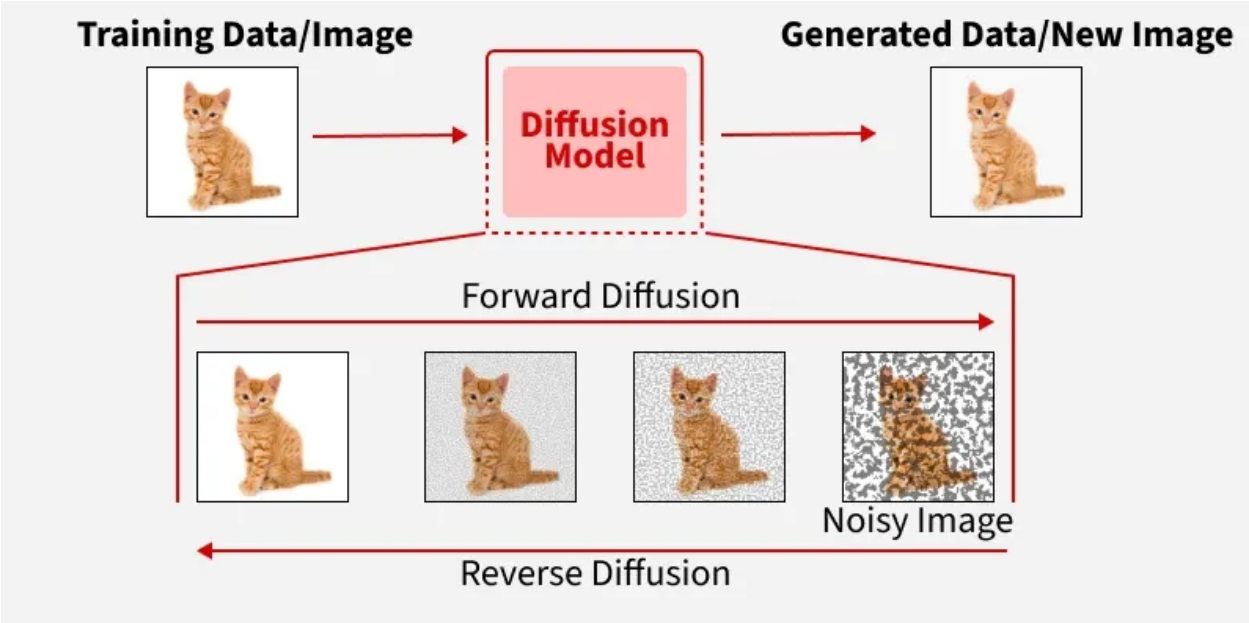


figura 12 *Rappresentazione schematica di un modello di diffusione: durante la forward diffusion (diffusione diretta), un'immagine di addestramento viene progressivamente contaminata da rumore; durante la reverse diffusion (diffusione inversa), il modello apprende a rimuovere il rumore per generare una nuova immagine.*

2.4 Le tipologie di generazione video

L'intelligenza artificiale generativa ha aperto nuove possibilità nella creazione di contenuti video e le principali categorie sono:

Video-to-Video: modelli che elaborano un input visivo preesistente trasformandolo stilisticamente, applicando effetti o generando versioni alternative del contenuto originale. Questo approccio consente di sperimentare nuovi linguaggi visivi e stilizzazioni, mantenendo però struttura e caratteristiche fondamentali del materiale di partenza.

Text-to-Video: modelli in grado di creare sequenze video complete a partire da input testuali. Hanno ampliato le possibilità del cinema amatoriale e della produzione indipendente, riducendo tempi e costi. Tuttavia, per ottenere risultati coerenti ed efficaci, restano fondamentali la conoscenza della grammatica cinematografica, della regia e della narrazione audiovisiva.

Image-to-video: modelli che partono da un'immagine statica e la animano, generando movimento coerente per soggetti, sfondo e camera pur mantenendo la fedeltà visiva al frame di partenza. È l'approccio più diffuso per chi lavora già con generatori di immagini e vuole portare quel materiale in movimento, garantendo un controllo maggiore sulla composizione e sull'estetica finale rispetto al text-to-video puro, dove l'aspetto del risultato resta più imprevedibile.

Clonazione vocale (Voice Cloning): Tecnologie che replicano o trasformano la voce umana generando contenuti audio realistici. Si distinguono principalmente in Text-to-Speech (TTS) che converte un testo scritto in voce sintetica o Speech-to-Speech (STS) che modifica una voce esistente per farla sembrare quella di un'altra persona.

2.5 I principali modelli di IA generativa video presenti sul mercato



figura 13 *Logo Kling.*

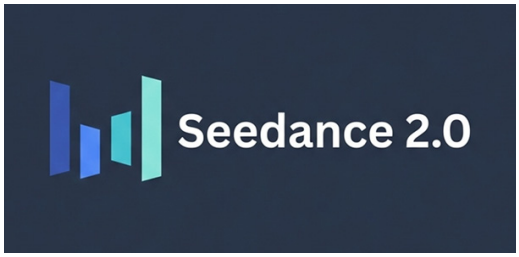


figura 14 *Logo Seedance (ByteDance).*

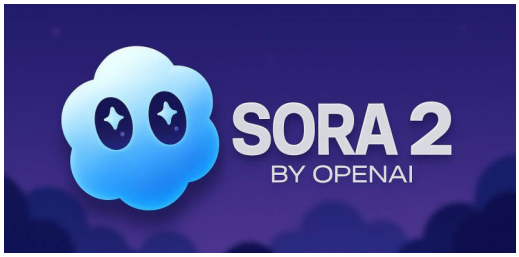


figura 15 *Logo Sora2 (OpenAI).*



figura 16 *Logo WAN.*

A data inizio 2026 i principali strumenti di generazione video sono i seguenti:

Kling: modello cinese basato su architettura diffusion-transformer con VAE 3D proprietario; con Kling 3.0 (feb. 2026) genera nativamente audio sincronizzato e video 4K a 60fps, con forte controllo del movimento (Motion Control, Elements per la coerenza dei personaggi);

Seedance (ByteDance): modello a doppio ramo diffusivo (video+audio) integrato nel chatbot Doubao, particolarmente effiace nell'image-to-video: mantiene dettagli di packaging, texture e tipografia dei prodotti fino a risoluzione 2K;

Sora2 (OpenAI): genera clip fino a 20-25 secondi con dialoghi e audio sincronizzati e maggiore accuratezza fisica; l'app standalone Sora è stata chiusa il 26 aprile 2026, ma l'accesso via API resta attivo fino al 24 settembre 2026;

WAN: modello open-source (licenza Apache 2.0) orientato alla narrazione cinematografica multi-shot, eseguibile anche in locale; generazione rapida (circa 20 secondi per clip nella versione 2.6);



figura 17 *Logo Google VEO.*

Google VEO (veo 3.1): modello multimodale nativo che genera audio e video come un unico output, con lip-sync accurato e foley coerente con l'azione fisica. È considerato tra i più solidi per l'aderenza al prompt e la qualità complessiva, ed è integrato anche in Adobe Firefly;



figura 18 *Logo Hailuo AI (MiniMax).*

Hailuo (MiniMax): Specializzato in stile anime, 3D CG e movimenti espressivi (danza, arti marziali), con buon supporto audio nativo e prezzo competitivo rispetto ai modelli occidentali equivalenti;



figura 19 *Logo Runaway.*

Runaway (Gen-4/Gen-4.5): piattaforma AI-video "full-stack" (non solo generazione ma anche editing), con Motion Brush per il controllo direzionale del movimento e clip fino a 4K/60 secondi con buona coerenza temporale;



figura 20 *Logo Luma Ray3.*

Luma Ray3: è il primo modello video a generare HDR nativo a 16 bit in formato ACES2065-1 EXR, con output 1080p e una gamma dinamica pensata per il color grading professionale;



figura 21 *Logo Hunyuan.*

Hunyuan: modello video open-source (HunyuanVideo), eseguibile localmente su GPU consumer; la licenza Tencent limita però l'uso commerciale in ambito UE;



figura 22 *Logo Midjourney.*

Midjourney: nato come generatore di immagini, ha lanciato nel 2025 il modello video V1 (image-to-video via Discord), tuttora più limitato rispetto ai concorrenti nativamente video-first;

Tuttavia non conviene iscriversi singolarmente ai singoli Ai video generators ma è più conveniente utilizzare i tool che raggruppano insieme i diversi modelli generativi, come ad
esempio:



figura 23 *Logo Magnific.*

Magnific: ex Freepik, è una piattaforma di stock media evoluta in hub di generazione AI, dà accesso a più modelli esterni (immagine/video) integrati nel proprio catalogo.



figura 24 *Logo Artlist.*

Artlist: combina libreria di stock footage/ musica con strumenti di generazione AI, pensata per workflow di content creation più ampi del solo video generativo.



figura 25 *Logo Adobe Firefly.*

Firefly (Adobe): integra oltre 30 modelli partner (Veo 3.1, Kling, Runway, Luma) dentro l'ecosistema Creative Cloud/Premiere Pro, con posizionamento "commercially safe" e modelli custom addestrabili sul brand.



figura 26 *Logo Higgsfield.*

Higgsfield: aggrega più di 15 modelli (Sora 2, Veo 3.1, Kling) attorno a un "Cinema Studio" che simula controlli di camera/lente professionali per una resa cinematografica ripetibile.



figura 27 *Logo ImagineArt.*

ImagineArt: pipeline unica per generazione, editing, animazione e automazione dei workflow, con crediti giornalieri gratuiti generosi rispetto ai concorrenti.



Capitolo 3

Etica e impatti delle immagini algoritmiche

3.1 Bias, pregiudizi e black box

Nei capitoli precedenti abbiamo analizzato come funzionano le IA (le architetture, i modelli e le tipologie di generazione). In questo capitolo ci interrogheremo sulle **conseguenze e sugli impatti** delle IA generative nel mondo: quali dinamiche di potere innescano, quali diritti mettono in discussione, quali effetti lasciano sul territorio e sulle persone. Per affrontare l'argomento è importante definire un concetto fondamentale: le AI vengono addestrate su **enormi dataset**, ovvero ampie raccolte di dati, nel caso delle immagini, milioni di fotografie accompagnate da didascalie e metadati, che costituiscono la base da cui il modello apprende a riconoscere pattern e associazioni. In base a come sono costruiti questi dataset, si generano diverse visioni dell'IA. Dunque, le IA non sono strumenti neutrali, ma ogni sistema di intelligenza artificiale porta con sé una **visione del mondo**, un insieme di scelte, spesso invisibili all'utente, che determinano ciò che il modello impara a vedere, a rappresentare e a considerare "normale" all'interno di una data categoria di immagini.

A conferma di quanto osservato finora, è utile richiamare le parole di **Flavia Barca**, che nel suo saggio I pregiudizi delle IA sintetizza efficacemente il nodo concettuale qui affrontato:
"La tecnologia non è mai neutra, e non lo sono nemmeno i suoi effetti sociali."

In quest'ottica appare particolarmente significativa questa affermazione: i sistemi di intelligenza artificiale non si limitano a osservare la realtà, ma la modellano attivamente, decidendo cosa sia rilevante e cosa no.

Si introduce perciò il concetto di **Bias IA** (in italiano "distorsione" o "pregiudizio"), che indica una **deviazione sistematica** nei risultati prodotti da un modello, causata da

squilibri o distorsioni presenti nei dati di addestramento. Il bias è un **errore sistematico** che si ripete in modo coerente e prevedibile ogni volta che il modello si trova di fronte a determinate categorie di input.

Le origini del bias di un modello IA possono essere ricondotte a due livelli:

1) La raccolta dei dati: se il dataset è costruito a partire da **fonti che sovra-rappresentano** certi gruppi sociali, culture o punti di vista, il modello apprenderà quella distribuzione come se fosse la realtà.
I modelli di apprendimento "assorbono" come spugne gli **squilibri culturali e di genere**, che codificano il mondo secondo pregiudizi di razza e classe sociale.
Secondo uno studio condotto da ricercatori di **Hugging Face e dell'Università di Lipsia (2023)**, basato sull'analisi di oltre 96.000 immagini generate, è stato rilevato che DALL-E 2 rappresentava figure come "CEO" o "director" come uomini bianchi nel 97% dei casi. Stable Diffusion associava le parole chiave "Black" e "woman" prevalentemente a professioni come assistente sociale, domestica o tassista.
Un ulteriore esempio riguarda il confronto tra le immagini reperite tramite due motori di ricerca web con il prompt **"Office in Ethiopia"** e quelle generate dai modelli DALL-E 2 e Stable Diffusion con lo stesso prompt **(figura 28)**. Il confronto evidenzia che, in entrambi i casi, i modelli generativi tendono a rappresentare l'Etiopia come un paese arretrato e in condizioni di povertà economica, riproducendo uno stereotipo diffuso anziché una rappresentazione accurata della realtà contemporanea del paese.



figura 28 *Una serie di immagini generate da motori di ricerca per immagini, DALL·E 2 e Stable Diffusion (SD) a partire dal prompt "Office in Ethiopia". Rispetto ai risultati della ricerca per immagini, entrambi i modelli generativi rappresentano l'Etiopia attraverso condizioni economiche precarie (Naik e Nushi, 2023).*

2) La classificazione dei dati: la classificazione dei dati all'interno di un training set non è mai un'operazione neutrale: implica decidere quali categorie utilizzare, quali soggetti includere e quali escludere, cosa considerare "normale" o rilevante. Le **etichette usate** per organizzare i dati non sono categorie assolute e oggettive, ma riflettono le scelte, la **cultura e i valori di chi le ha create**. Risulta quindi difficile ottenere modelli realmente "rappresentativi" di una visione del mondo plurale: il modello finisce per apprendere una **versione parziale e distorta** della realtà. Da qui il rischio di ottenere un dataset impreciso e sbilanciato, cioè un modello che funziona bene per alcuni gruppi sociali, ma male per altri o addirittura in modo discriminatorio.

Infatti, la logica di classificazione è, per sua natura, intrinsecamente discriminatoria: classificare significa "separare" e separare significa "creare categorie rigide". In questo senso, la discriminazione viene automatizzata. Si parla, a questo proposito, dei **principi bias-in e bias-out**: se i dati storici da cui il modello apprende sono già viziati da pregiudizi, l'algoritmo **risulterà matematicamente discriminatorio.**
Ad esempio, un dataset che sovrarappresenta corpi bianchi, maschili e occidentali genera inevitabilmente un modello visivo che normalizza quella particolare visione del mondo, riproducendola su larga scala ogni volta che un utente genera un'immagine.

A questo si aggiunge il concetto di **Ghost Work**: una vasta schiera di **lavoratori sottopagati**, spesso invisibili, incaricati di **etichettare i dati** che alimentano l'intelligenza artificiale. Il loro compito è catalogare l'oceano di contenuti senza i quali l'IA non potrebbe funzionare inclusi, spesso, contenuti sensibili e disturbanti.

Le implicazioni diventano ancora più problematiche quando si affronta il **tema della sicurezza**: le criticità si aggravano quando il bias algoritmico viene portato in ambito giudiziario e l'AI viene impiegata per la sorveglianza e il controllo, come nel **caso COMPAS** (Correctional Offender Management Profiling for Alternative Sanctions), uno strumento algoritmico usato nei tribunali statunitensi per

stimare il rischio di recidiva degli imputati e orientare le decisioni sulla pena.
Nel 2016 un'inchiesta di ProPublica ha rivelato che il software produceva **errori sistematicamente sbilanciati** in base al **colore della pelle**: gli imputati afroamericani venivano classificati come ad alto rischio di recidiva il doppio delle volte rispetto agli imputati bianchi con profili analoghi. I falsi positivi, ovvero persone classificate ad alto rischio che poi non avrebbero commesso nuovi reati, erano pari al 45% tra gli imputati neri, contro il 23% tra quelli bianchi.

In questo contesto, il meccanismo è chiaro: l'algoritmo era stato addestrato su **dati storici che riflettevano le disuguaglianze** già presenti nel sistema penale americano (tassi di arresto più elevati per la popolazione afroamericana, condizioni socioeconomiche più precarie, maggiore esposizione al controllo della polizia). Il modello non ha fatto altro che imparare e riprodurre queste disparità, amplificandole.

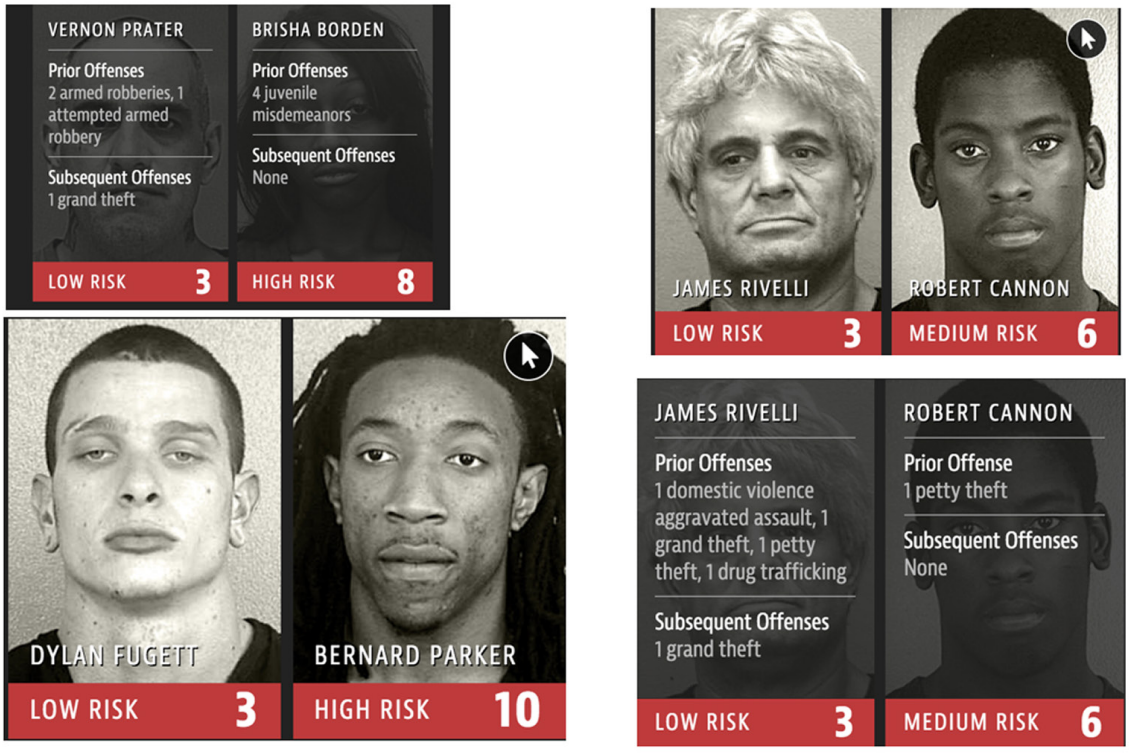


figura 29 *Il caso COMPAS: vengono mostrate una serie di schede che riportano precedenti penali, valutazione algoritmica del rischio di recidiva e reati successivi. Il confronto evidenzia alcuni casi di classificazione errata, come quello di Brisha Borden, classificata ad alto rischio ma non nuovamente arrestata, e di Vernon Prater, classificato a basso rischio nonostante la successiva commissione di un ulteriore reato.*

Il caso COMPAS rivela qualcosa di più profondo di un semplice errore tecnico. Come osservano i giuristi, che hanno analizzato la sentenza della Corte Suprema del Wisconsin del 2016, che ha validato l'uso del software pur ponendo alcune limitazioni, il problema non riguarda solo l'accuratezza del sistema, ma anche la **trasformazione del paradigma decisionale** che esso introduce: un algoritmo che produce un punteggio numerico tende ad acquisire un'autorità epistemica che raramente spetta all'argomentazione umana. Il rischio, insomma, non è solo che la macchina sbagli, ma che il suo errore venga percepito come oggettivo e quindi accettato senza essere messo in discussione.

Ed è per questo che è importante interrogarsi su cosa comportino queste scelte: accettare un'AI senza porsi queste domande significa, come scriveva **Norbert Wiener** già nel 1950, cedere il controllo senza nemmeno accorgersene. Wiener, matematico, statistico statunitense e **padre della cibernetica** (la disciplina che studia i sistemi di comunicazione e controllo ed autoregolazione negli esseri viventi e nelle macchine) metteva in guardia in maniera "profetica" nel suo testo "L'uso umano degli esseri umani. Introduzione alla cibernetica":
"Ho già detto che per quanto l'uomo moderno, e in specie l'americano, conosca preziosi accorgimenti tecnici, ha ben poca coscienza di ciò che vuole. Egli accetterà perciò la superiore abilità delle decisioni prese dalla macchina, senza preoccuparsi eccessivamente dei motivi e dei principi che sono dietro di esse" (Pag 227 - cap XI "Alcune macchine di comunicazione e il loro futuro")

Queste parole hanno anticipato i rischi con cui oggi l'intelligenza artificiale generativa si manifesta nella società e, più in generale, l'intelligenza artificiale nel suo complesso. Il rischio che i sistemi algoritmici vengano adottati come se fossero realtà trasparenti e affidabili, senza chiedersi chi li abbia costruiti, con quali dati, per quali fini e secondo quale visione del mondo è elevato. Bisognerebbe, dunque, adottare un'**intelligenza artificiale pluriversale**, in cui ci si approcci in maniera critica alla tecnologia e si diversifichino le figure che sviluppano, progettano e governano i modelli di IA.

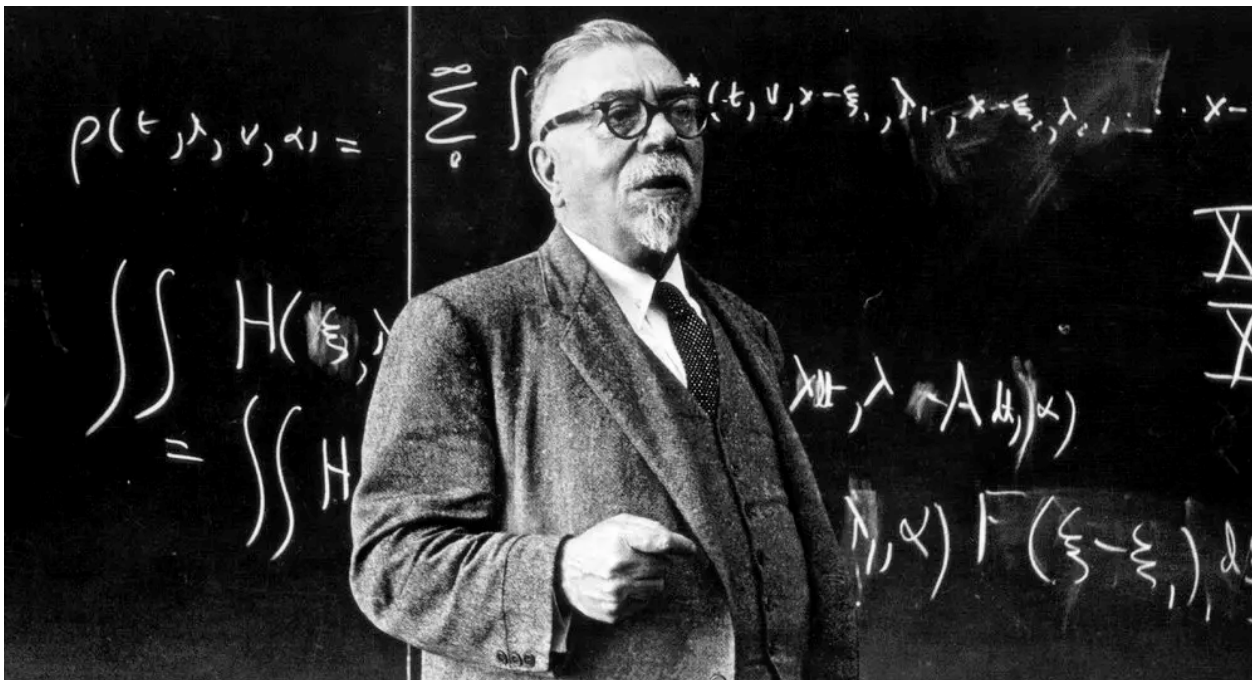


figura 30 *Norbert Wiener, matematico, fondatore della cibernetica e autore de "L'uso degli esseri umani. Introduzione alla cibernetica".*

Tuttavia, contrastare questi rischi non è affatto semplice, poiché il problema risiede nella natura stessa dei modelli di deep learning: le IA sono vere e proprie **black box**, scatole nere. Il termine indica qualcosa di cui è possibile osservare l'input (ad esempio un prompt testuale per la generazione di un'immagine) e l'output (la risposta prodotta dal sistema), ma non i processi interni che li collegano.

All'interno di una rete neurale operano miliardi di neuroni artificiali distribuiti su strati multipli, che si attivano in modo complesso, interagiscono e cambiano continuamente. Ogni neurone possiede un valore numerico, che si modifica a ogni iterazione durante l'addestramento, ma **risulta impossibile mappare con precisione questi cambiamenti**: i neuroni possono interagire tra loro in modi non lineari e imprevedibili, arrivando persino ad aggregarsi spontaneamente in cluster, che gestiscono il dato in maniera collettiva, senza che chi ha progettato il sistema abbia esplicitato alcuna strategia di calcolo.

Questa è un'aggregazione che avviene in totale autonomia da parte dell'AI.
Tutto questo rende impossibile ricostruire con precisione ciò che accade all'interno di una rete neurale in risposta a un input specifico. In questo senso il paragone con il cervello umano è calzante: sappiamo che il cervello elabora le informazioni, ma comprendere come ogni singolo neurone contribuisca al pensiero resta un'impresa quasi impossibile.

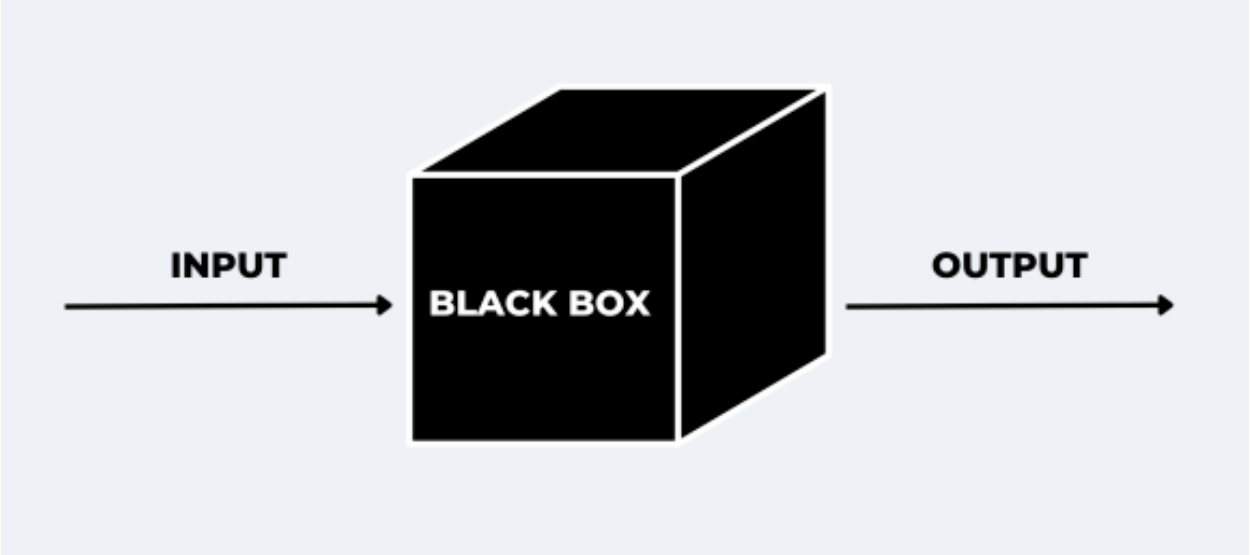


figura 31 *Rappresentazione visiva del concetto di black box: input e output sono visibili, mentre il processo decisionale interno resta opaco.*

3.2 Privacy, diritti e copyright nei dataset di training

Come già descritto in precedenza, i sistemi di IA basati sul deep learning necessitano di raccogliere una **quantità estremamente elevata di dati**, non solo numerici, ma anche di alta qualità appartenenti a categorie specifiche. Esistono aziende che si occupano di creare i dataset da fornire in pasto alle IA, che possono essere di vario genere, pubblici o privati. Due esempi principali sono Common Crawl e LAION-5B.



figura 32 *Logo Common Crawl.*

Common Crawl è un'organizzazione non profit americana, attiva dal 2011, che raccoglie automaticamente grandi quantità di contenuti web tramite **scraping**, ovvero scansione automatica delle pagine web. Questi dati vengono poi messi gratuitamente a disposizione per la ricerca e lo sviluppo di algoritmi di IA non commerciali. Il dataset è composto prevalentemente da testi e include oltre 3,1 miliardi di pagine web, per un totale di **420 terabyte di dati**. Modelli come ChatGPT e molti altri grandi modelli linguistici sono stati addestrati su Common Crawl.



figura 33 *Logo LAION-5B.*

LAION-5B, invece, è un **dataset pubblico e open-source** che raccoglie circa 5 miliardi di immagini accompagnate dalle relative descrizioni testuali, costruito tramite **crawler** (programmi automatici che scandagliano il web passando da una pagina all'altra per raccogliere ed estrarre contenuti, in questo caso coppie immagine-testo).

La sua funzione principale è addestrare modelli di generazione visiva ed è il dataset su cui è stato **addestrato Stable Diffusion.**
LAION-5B è stato al centro di un acceso dibattito etico, noto come lo **scandalo LAION-5B**. Nel dicembre 2023 i ricercatori dello Stanford Internet Observatory hanno pubblicato uno studio sul dataset, secondo cui al suo interno erano presenti 1.008 immagini verificate di abuso sessuale su minori **(CSAM**, Child Sexual Abuse Material), sebbene il numero effettivo potesse essere molto più alto.

Le immagini erano state indicizzate automaticamente dai crawler, senza alcun filtro manuale.
Ma il problema non si limitava solo al CSAM. Il dataset conteneva anche **immagini protette da copyright**, referti medici e cartelle cliniche private, fotografie personali e sensibili raccolte da profili social e siti web senza il consenso degli interessati.
Google aveva già rifiutato in precedenza di rendere pubblico il proprio modello Imagen proprio perché il dataset LAION-400M, il predecessore di LAION-5B, conteneva, come dichiarato dagli stessi ricercatori di Google, contenuti pornografici, insulti razzisti e stereotipi sociali dannosi.
Questo caso ha messo in luce una contraddizione strutturale dei big data: più un dataset è grande, più diventa difficile da analizzare, filtrare e controllare. LAION ha successivamente pubblicato una versione ripulita del dataset, **Re-LAION-5B**, rimuovendo oltre 2.000 link problematici.
Tuttavia, il danno era ormai avvenuto: Stable Diffusion era già stato addestrato, distribuito e scaricato da milioni di utenti in tutto il mondo.

Il caso LAION-5B ha portato in superficie una questione legale su cui l'industria dell'AI è chiamata a riflettere: è lecito utilizzare opere protette da copyright per addestrare un modello commerciale senza il consenso dell'autore?

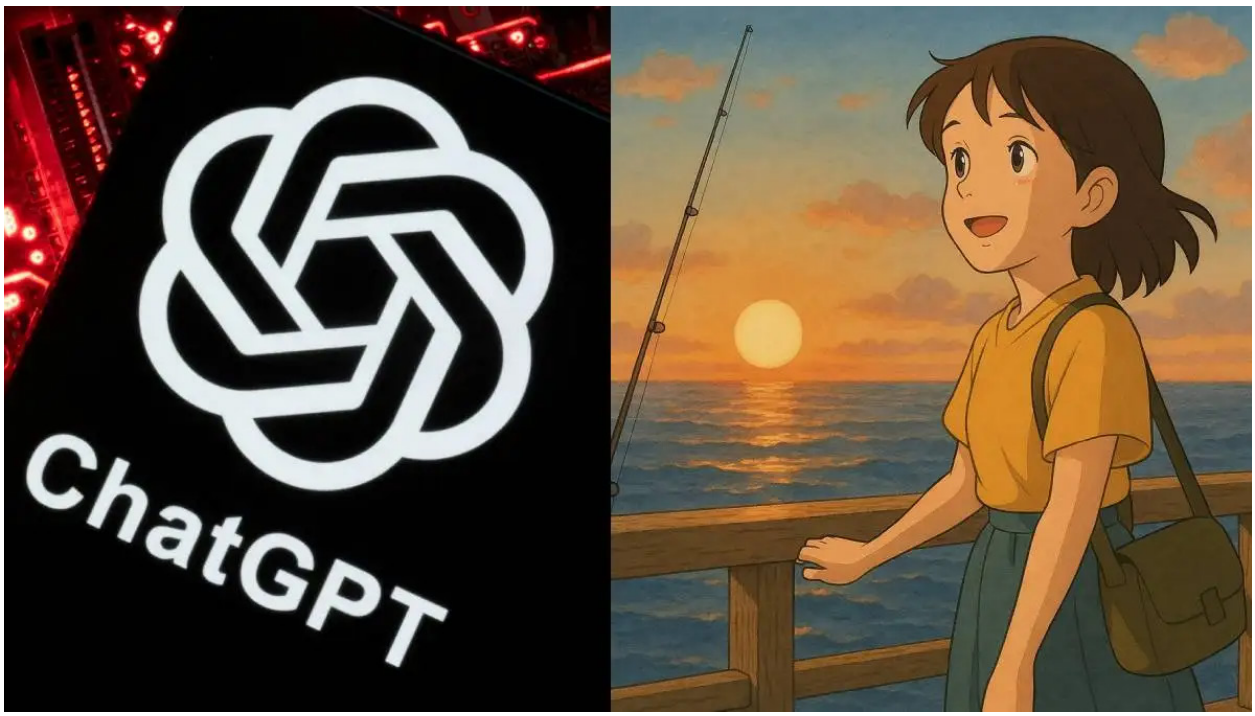
La risposta non è univoca. Negli Stati Uniti sono in corso diverse cause legali, tra cui un'azione di **Getty Images contro Stability AI** e azioni collettive di illustratori e artisti contro Midjourney e Stability AI, volte a stabilire se il training su opere protette costituisca una violazione del copyright.

In Europa, il Tribunale di Amburgo si è pronunciato il 27 settembre 2024 nel caso **Kneschke contro LAION**, stabilendo che LAION poteva avvalersi dell'eccezione per il text and data mining a fini di ricerca scientifica non commerciale (art. 3 della direttiva UE 2019/790). Il confine con gli usi commerciali resta però ambiguo: nel 2025 la Corte d'Appello ha confermato la liceità dell'addestramento, ma ha spostato la base giuridica sull'eccezione TDM commerciale (art. 4), poiché l'opt-out del sito di origine non era in formato leggibile da macchina.

Un caso emblematico, che ha reso visibile al grande pubblico la questione del copyright nell'AI generativa, è quello legato allo **Studio Ghibli**.
Nel marzo 2025, a poche ore dal rilascio del nuovo modello generativo **GPT-4o di OpenAI**, si sono rapidamente diffuse sui social media immagini generate nello stile del regista **Hayao Miyazaki**: fotografie trasformate in atmosfere soffuse di film come Il mio vicino Totoro o Il castello errante di Howl. Il caso ha sollevato riflessioni in ambito legale sui diritti di copyright: per replicare con tale precisione lo stile visivo dello studio giapponese, è plausibile che OpenAI abbia addestrato il modello su milioni di fotogrammi tratti dai film di Miyazaki.

Come spiega però **Evan Brown**, avvocato specializzato in proprietà intellettuale intervistato da TechCrunch, lo stile non è esplicitamente protetto dal diritto d'autore, il che colloca prodotti come GPT-4o in una zona grigia legale ancora da definire in modo specifico.

Il caso Ghibli, tuttavia, non è isolato: gli utenti hanno replicato lo stesso esperimento con lo stile Pixar e con quello di Dr. Seuss, confermando che la questione non riguarda un singolo autore, ma l'intera cultura visiva su cui questi modelli vengono addestrati. Come sintetizza Brown, il nodo centrale riguarda le implicazioni legali della prassi di raccogliere dati da tutto il web per costruire questi database, e se ciò configuri di fatto una violazione sistematica del diritto d'autore. Una domanda a cui decine di cause legali in corso contro OpenAI, Meta e Midjourney stanno cercando di dare risposta, in assenza di una normativa chiara che regoli l'addestramento dei modelli di IA.



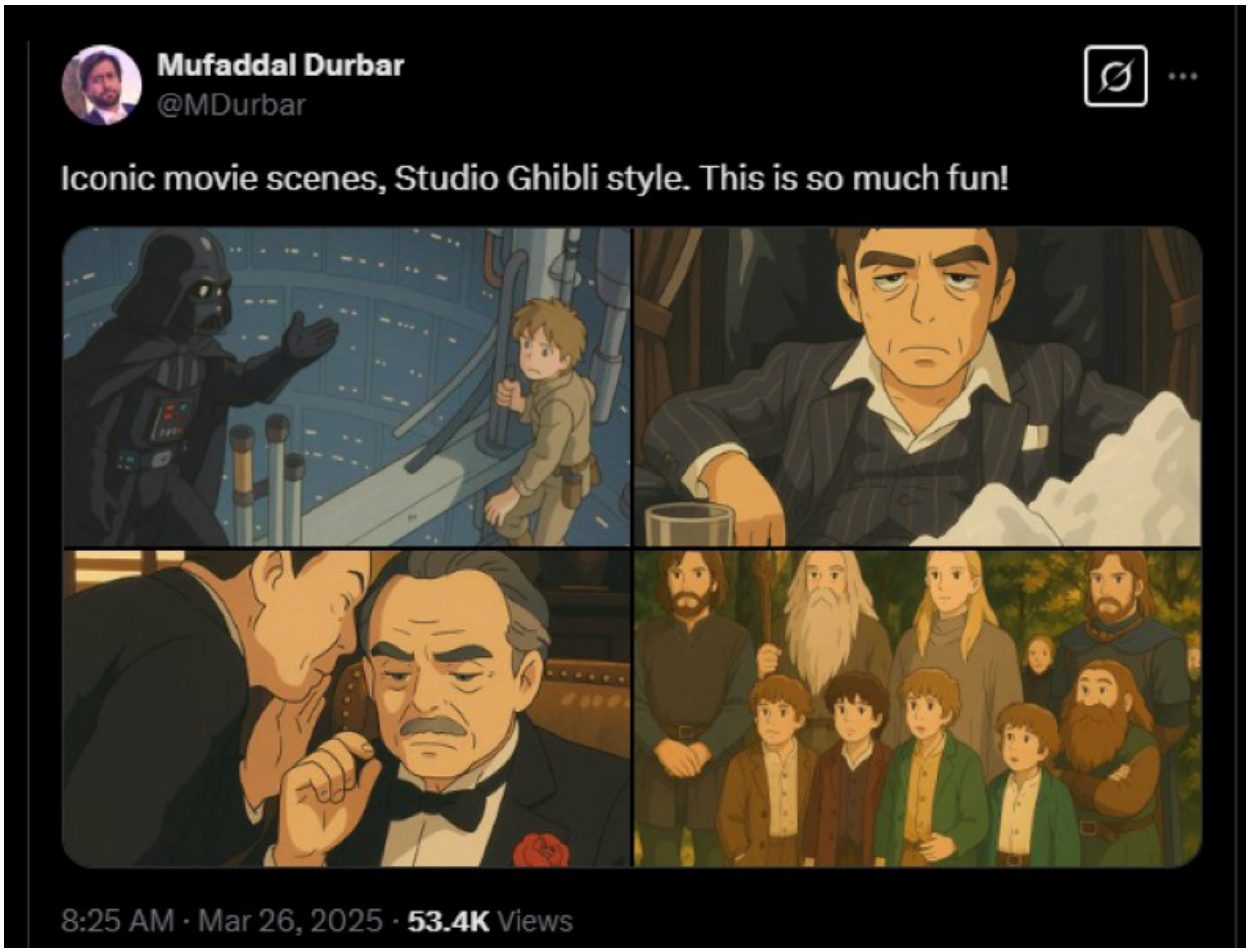


figura 34 e 35 *Immagini generate con ChatGPT/GPT-4o nello stile visivo associato allo Studio Ghibli, marzo 2025.*

Proprio per colmare questo vuoto normativo, l'**AI Act dell'Unione Europea**, approvato nel 2024, ha introdotto l'obbligo per i fornitori di modelli IA di adottare una **policy per il rispetto del copyright** e di rispettare la riserva dei diritti espressa dagli autori, inclusa la possibilità di **opt-out**, ovvero la facoltà di escludere le proprie opere dai dataset di training. Si tratta di un primo passo normativo significativo, sebbene la sua applicazione pratica rimanga complessa: come si è visto nel caso Kneschke contro LAION, la validità stessa di un opt-out può dipendere da dettagli tecnici, come il formato in cui viene espresso, capaci di ribaltare l'esito di una causa legale.

Le questioni di diritto e copyright non sono però rimaste confinate nelle aule di tribunale. Nel 2023 hanno attraversato i set cinematografici di Hollywood, dando vita alla più grande interruzione della storia dell'industria dell'intrattenimento americana, con due scioperi sovrapposti che hanno paralizzato le produzioni in tutto il mondo.

Il 2 maggio 2023 la **Writers Guild of America (WGA)**, il **sindacato degli sceneggiatori americani**, ha indetto uno sciopero durato 148 giorni, conclusosi il 27 settembre. L'uso crescente dell'AI generativa da parte degli studios era uno dei nodi centrali: le major volevano poter usare modelli IA per generare o riscrivere copioni, riducendo il numero di sceneggiatori necessari. I principali punti di conflitto erano il timore di sostituzione lavorativa, la preoccupazione che le aziende potessero usare IA addestrata sui lavori degli



figura 36 *Logo dell'AI Act, il regolamento europeo sull'intelligenza artificiale.*

sceneggiatori senza compensarli adeguatamente, il rischio di standardizzazione dei contenuti e l'incertezza legale su chi possedesse i diritti d'autore di una sceneggiatura generata da un algoritmo. L'**accordo finale**, ratificato dal 99% dei membri della WGA, ha stabilito che l'IA non può scrivere né riscrivere copioni o qualsiasi materiale letterario; che il materiale generato dall'IA non può essere considerato fonte su cui basare il lavoro di uno sceneggiatore; che le aziende non possono obbligare uno sceneggiatore a usare strumenti di IA; e che devono dichiarare esplicitamente se i materiali consegnati includono contenuti generati dall'IA.

Il 14 luglio 2023 si è aggiunto lo sciopero della **SAG-AFTRA (Screen Actors Guild American Federation of Television and Radio Artists)**, il **sindacato degli attori**, durato 118 giorni che si è concluso il 9 novembre dello stesso anno. Al centro del conflitto c'era una proposta degli studios che volevano acquisire la **scansione digitale del corpo** di un attore, pagandolo per una sola giornata, per poi riutilizzarne il volto e la voce all'infinito, anche dopo la sua morte, senza ulteriore compenso né consenso. L'accordo conclusivo ha introdotto **tutele significative**: gli studios devono ottenere il consenso esplicito di ogni attore prima di creare una sua replica digitale, l'attore deve essere compensato per ogni utilizzo, le repliche di attori deceduti richiedono il consenso degli eredi e i consensi devono riferirsi a progetti specificamente definiti.

Entrambi gli scioperi hanno avuto un impatto economico stimato in oltre **6 miliardi di dollari di perdite** sull'economia della California. A causa degli scioperi un gran numero di produzioni cinematografiche e televisive è stato sospeso, rinviato o cancellato. Tra i titoli colpiti dalle interruzioni vi erano Deadpool 3 di Marvel, Mission Impossible - Dead Reckoning Part Two con Tom Cruise, il sequel di Avatar e intere stagioni di serie tv come Stranger Things, The Last of Us, Euforia, Severance, rinviate al 2025 e oltre.
Ma il loro significato va oltre i numeri: per la prima volta, i lavoratori creativi di un'industria globale hanno **negoziato collettivamente i propri diritti** di fronte all'IA, ottenendo risultati concreti che potrebbero diventare un modello per altri settori esposti alle stesse pressioni.





figura 37 e 38 *Manifestazioni WGA e SAG-AFTRA durante gli scioperi del 2023, contro l'uso non regolamentato dell'intelligenza artificiale nell'industria audiovisiva.*

3.3 Il fenomeno dei deepfake

Definizione e origine del termine

Il termine deepfake è un neologismo che nasce dalla fusione delle parole **"deep learning"** e **"fake"**, coniato nel 2017 da un utente anonimo della piattaforma di reddit. Si tratta di una tecnica per la **sintesi dell'immagine umana** basata sull'intelligenza artificiale, utilizzata per **combinare e sovrapporre immagini e video esistenti** a video o immagini originali tramite le GAN (Generative Adversarial Networks).

Tuttavia non esiste una definizione univoca del termine: in letteratura se ne moltiplicano diverse, senza mai convergere del tutto:
- ***"Video sintetici che assomigliano molto a video reali"*** (Vaccari e Chadwick, 2020)
- ***"Video iper-realistici nei quali il volto di una persona è stato analizzato da un algoritmo di deep learning e sovrapposto sul volto di un attore"*** (Öhman, 2020)
- ***"Un tipo di media sintetici (immagini, audio e video) manipolati o completamentegenerati attraverso l'intelligenza artificiale"*** (Schick, 2020)



figura 39 *Esempio di manipolazione facciale mediante tecniche deepfake (face reenactment). La sequenza illustra la sintesi di differenti espressioni e movimenti del volto a partire da immagini di soggetti reali e da riproduzioni artistiche.*

Sebbene nell'uso comune il termine venga spesso associato ai soli video, in realtà copre qualsiasi tipologia di media manipolato o generato tramite l'intelligenza artificiale, il cui obiettivo è la mimesi totale di un medium originale.

La tassonomia comprende quattro grandi categorie:

1) Immagini deepfake: non sono semplici immagini ritoccate, ma nuove immagini generate a partire da immagini esistenti. Il caso più emblematico è **www.thispersondoesnotexist.com**, sito creato dall'ingegnere Philip Wang nel 2019, che sfrutta le GAN per produrre ritratti di persone inesistenti: un volto inedito a ogni refresh della pagina, pressoché indistinguibile da fotografie reali.

Un caso più inquietante è **DeepNude (2019)**, un'applicazione che usava reti neurali generative per "rimuovere virtualmente" i vestiti dalle foto di donne, generando immagini finte di nudo. Funzionava solo su foto di donne (l'algoritmo era stato addestrato su un dataset sbilanciato in tal senso) e produceva risultati piuttosto approssimativi, spesso con artefatti evidenti. Permetteva di creare immagini di nudo non consensuali (spesso definite "revenge porn sintetico" o rientranti nella categoria più ampia dei deepfake pornografici) a partire da qualsiasi foto, anche di persone reali tratte dai social. Nonostante la chiusura ufficiale, il codice fu clonato e ricreato da altri (i cosiddetti "cloni DeepNude") e, negli anni successivi, sono nate app e bot simili, spesso più sofisticati che si sono diffusi.

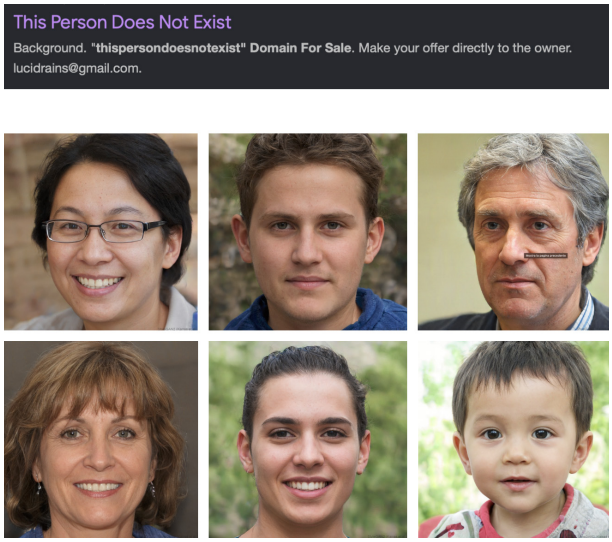


figura 40 *Volti umani generati artificialmente e randomicamente tramite il sito thispersondoesnotexist.com.*

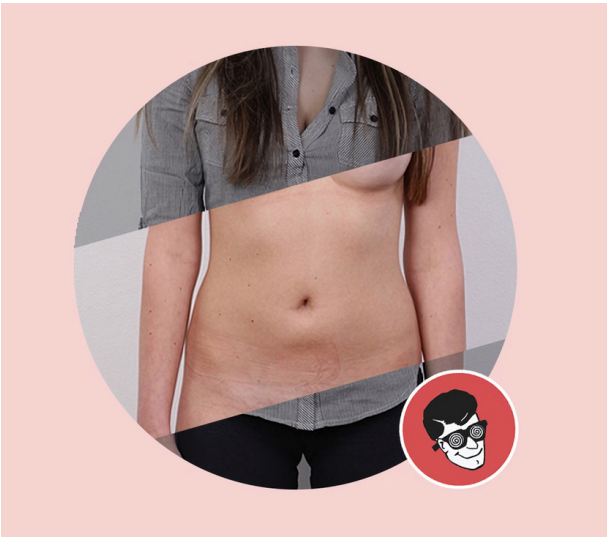


figura 41 *Deepnude era un'applicazione che generava immagini deepfake intime non consensuali tramite IA.*

2) Audio deepfake: la possibilità di **clonare la voce** di una persona reale per farle pronunciare parole mai dette. Audio o canzoni generati dall'IA a partire dalle voci di persone realmente esistenti. Un caso emblematico è quello di **"Heart on My Sleeve"** (2023): un utente anonimo di TikTok, noto come Ghostwriter977, ha scritto e prodotto un brano originale, sostituendo la propria voce con repliche generate dall'IA di Drake e The Weeknd, con il tag sonoro caratteristico del produttore Metro Boomin. Il brano ha raggiunto oltre 20 milioni di visualizzazioni su Twitter e 11 milioni su TikTok, prima che Universal Music Group ne ottenesse la rimozione da tutte le piattaforme di streaming.

3) Testi deepfake: modelli come GPT-2 e GPT-3 di OpenAI rientrano a pieno titolo in questa categoria.

4) Video deepfake: la forma più diffusa e più discussa, su cui ci concentreremo nei paragrafi successivi.

Ma come vengono generati i video deepfake? Come già anticipato nel Capitolo 2, i deepfake vengono prodotti principalmente tramite le **GAN (Generative Adversarial Networks)**, reti antagoniste generative.
Si tratta di un sistema composto da **due reti neurali** che **lavorano in competizione** tra loro: il generator, che tenta di produrre immagini false indistinguibili dagli originali, e il discriminator, che prova a smascherare il falso.
Come descrisse il loro ideatore Ian Goodfellow, il processo è "simile a un team di contraffattori che tenta di produrre valuta finta mentre la polizia prova a smascherarli: la competizione porta entrambi i team a migliorarsi fino a quando i dati contraffatti risultano indistinguibili da quelli genuini".

Nel caso dei video deepfake, la rete viene addestrata su grandi quantità di materiale visivo (foto, video, registrazioni) della persona "vittima", da cui apprende quali caratteristiche del volto replicare e sovrapporre al soggetto del video di destinazione.
Il software più utilizzato oggi tra i deepfake designer è **DeepFaceLab**, pubblicato su GitHub dall'utente russo Iperov verso la fine del 2018; esistono anche alternative open source come FaceSwap e MyFakeApp, oltre alla pionieristica FakeApp, lanciata nel gennaio 2018, prima

applicazione a rendere la produzione di deepfake accessibile anche a utenti privi di competenze tecniche avanzate. A queste si affiancano numerose altre app disponibili su Google Play e App Store, a pagamento o gratuite, con risultati di qualità variabile.

La funzione di faceswap è stata inoltre integrata anche in applicazioni e software di uso più comune, come Canva.

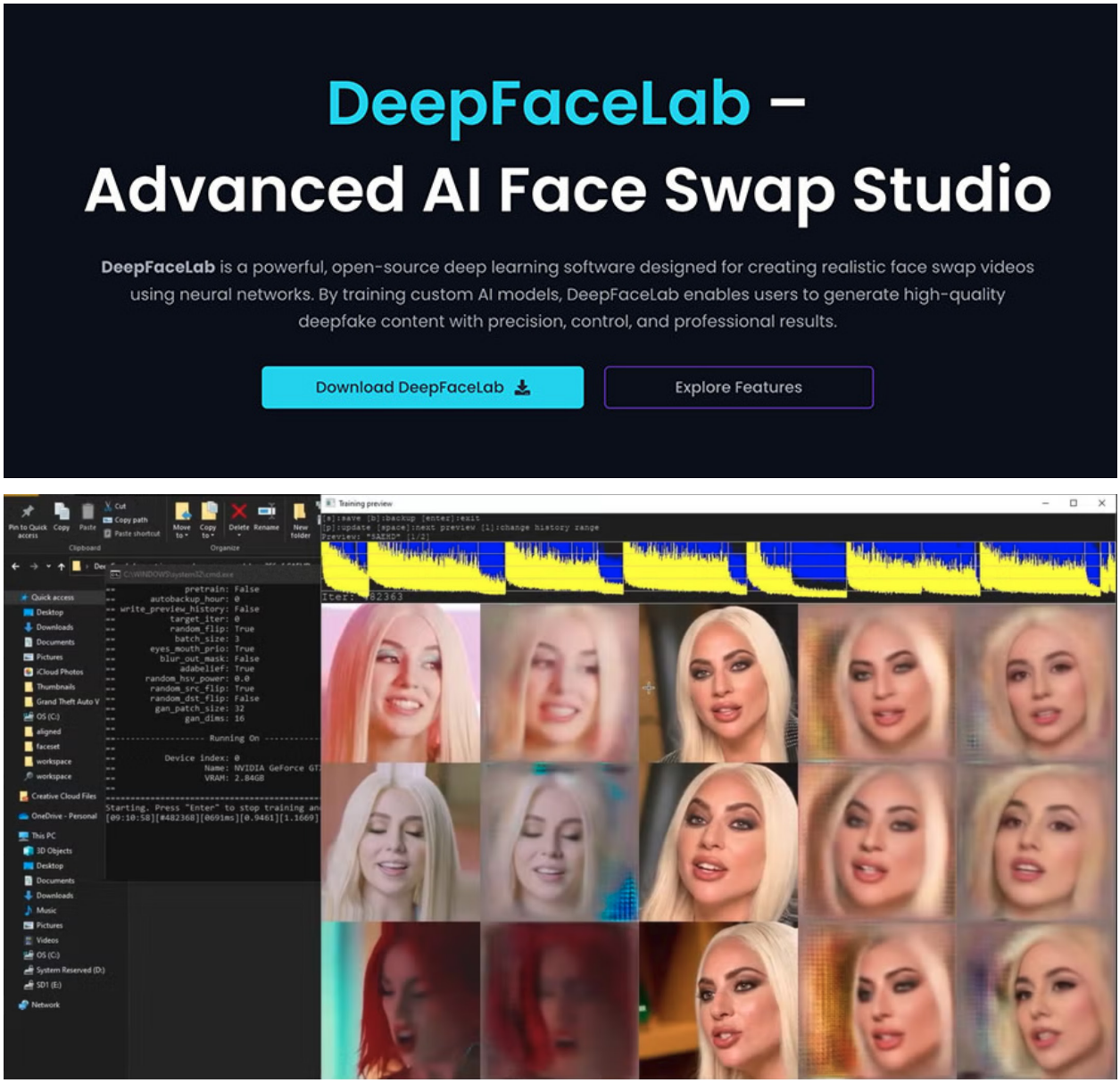


figura 42 e 43 *DeepFaceLab è un software open source per la creazione di video deepfake mediante face swap; nella seconda immagine è mostrato un esempio di face swap realizzato con DeepFaceLab: il volto della persona di origine viene appreso e sovrapposto al volto del soggetto nel video di destinazione.*

Inoltre è fondamentale distinguere i deepfake dai cosiddetti **cheapfake (o shallowfake)** cioè falsi prodotti non attraverso l'AI ma con **strumenti di editing tradizionali**, attraverso tecniche come il rallentamento o l'accelerazione del video, il ritaglio o la ricontestualizzazione delle immagini.

Alcuni **esempi celebri di cheapfake** hanno avuto conseguenze politiche reali. Il caso più noto è il video di **Nancy Pelosi** di maggio 2019, allora Speaker della Camera dei Rappresentanti USA, tra le figure democratiche più autorevoli, in aperto scontro con Trump. Il video è stato semplicemente **rallentato a tre quarti della velocità originale** per farla apparire ubriaca e confusa. Nonostante la semplicità della manipolazione, il video ottenne milioni di visualizzazioni su Facebook e fu condiviso da diversi esponenti politici prima di essere smascherato come falso.
I deepfake e cheapfake dimostrano come spesso i falsi più pericolosi non richiedano una grande sofisticazione tecnica, ma sfruttino i **bias cognitivi dello spettatore** e la velocità di diffusione dei social media.



figura 44 *Confronto tra il video originale di Nancy Pelosi e la sua versione rallentata e manipolata (cheapfake), diffusa nel 2019 per alterare la percezione della sua immagine pubblica.*

I potenziali rischi dei deepfake

L'immagine di una persona, intesa come aspetto fisico, è un diritto inviolabile: nessuno può diffonderla senza il suo consenso. L'uso improprio dei deepfake viola sistematicamente questo diritto, esponendo le vittime a rischi concreti e molteplici:

Un primo rischio riguarda la **clonazione dell'identità:** il volto e la voce di una persona vengono usati per farle compiere azioni o pronunciare dichiarazioni mai avvenute, con ricadute che vanno dal danno reputazionale a veri e propri usi criminali, come truffe o l'accesso non autorizzato a sistemi di riconoscimento facciale.

Un secondo ambito, già affrontato in precedenza, è quello dei **DeepNude e del revenge porn sintetico**: contenuti a sfondo sessuale generati a partire dal volto di una persona reale senza il suo consenso, un fenomeno che si configura a tutti gli effetti come violenza di genere, colpendo in maniera prevalente le donne.

Un terzo rischio riguarda la diffusione di **fake news e disinformazione**: dichiarazioni false attribuite a politici o figure pubbliche allo scopo di orientare l'opinione pubblica. Il caso più emblematico è il **video deepfake del presidente ucraino Zelensky**, diffuso nel marzo 2022, in cui il leader sembrava invitare le truppe ucraine alla resa: Zelensky fu costretto a smentire pubblicamente il video per arginarne gli effetti.

Il rischio forse più insidioso, tuttavia, riguarda il **problema della verità percepita**: la diffusione dei deepfake produce un effetto collaterale su ciò che è autentico. Quando chiunque può creare un video falso convincente, diventa impossibile fidarsi ciecamente di qualsiasi contenuto audiovisivo, e si diffonde una tendenza opposta e altrettanto pericolosa: negare l'autenticità di prove reali, spacciandole per deepfake. È il cosiddetto liar's dividend (il "dividendo del bugiardo"): il beneficio che chi mente ottiene dal clima di sfiducia generalizzata generato dalla tecnologia stessa.

Dal **punto di vista giuridico,** l'uso improprio dei deepfake è già perseguibile in Italia attraverso diverse norme:
- Il deepnude può integrare il reato di revenge porn (art. 612-ter c.p.), che prevede la reclusione da 1 a 6 anni e una multa fino a 15.000 euro, sia per chi produce sia per chi diffonde i contenuti.
- Le dichiarazioni false o le immagini che danneggiano la reputazione di una persona configurano il reato di diffamazione, punito con la reclusione da 6 mesi a 3 anni, aggravato dall'uso dell'intelligenza artificiale quale strumento del reato.
- La diffusione di notizie false tramite deepfake può integrare il reato di procurato allarme (art. 656 c.p.), punito con l'arresto fino a 3 mesi o un'ammenda fino a 309 euro.

A livello europeo, due normative costituiscono il quadro di riferimento: il **GDPR (Regolamento UE n. 679/2016)**, che definisce le regole per il trattamento dei dati personali, inclusi i dati biometrici alla base dei deepfake, e l'**AI Act (2024)**, che regolamenta l'utilizzo dell'intelligenza artificiale e introduce obblighi specifici di trasparenza per i sistemi generativi, tra cui l'obbligo di etichettare i contenuti sintetici come tali.

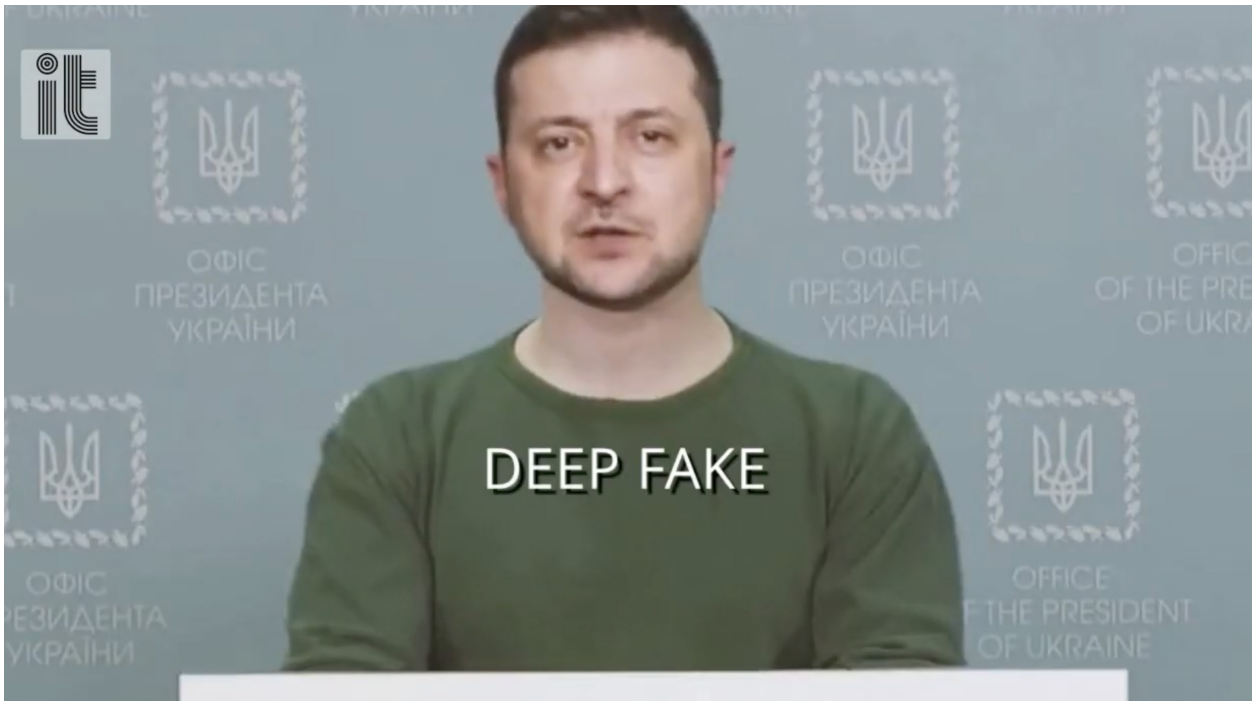


figura 45 *Deepfake di Volodymyr Zelensky diffuso nel marzo 2022, nel quale il presidente ucraino veniva falsamente mostrato mentre invitava le forze armate alla resa.*

Come riconoscere un deepfake?

Nonostante la qualità dei deepfake sia in costante miglioramento, esistono ancora segnali che possono aiutare a identificarli

Nei video, i segnali più frequenti riguardano **artefatti o glitch nei movimenti della bocca e degli occhi**, un labiale spesso fuori sincrono con l'audio, bordi del viso innaturali e irregolarità nella texture della pelle, soprattutto nelle zone di transizione tra il volto e il resto dell'inquadratura.

Nelle immagini, gli indizi più comuni sono una **pelle eccessivamente liscia** e priva di imperfezioni naturali, insieme a dettagli anatomici spesso incoerenti o deformati, come orecchie, denti e l'attaccatura dei capelli sulla fronte.

Negli audio deepfake, infine, il tallone d'Achille resta la difficoltà nel riprodurre fedelmente **inflessioni dialettali, cadenze locali** e le particolarità di pronuncia tipiche della voce reale da clonare.

Per ridurre il rischio di diventare vittime di deepfake è fondamentale limitare la condivisione di foto e video su app o siti poco affidabili, leggere con attenzione le condizioni sulla privacy, in particolare le clausole relative alla gestione e al riutilizzo dei dati e non concedere consensi generici e illimitati al trattamento del proprio materiale biometrico.

A questo si aggiunge una dimensione più culturale del problema. Nell'era dell'**IA slop**, termine con cui ci si riferisce a contenuti generati dall'IA a **basso valore educativo,** prodotti in serie e senza particolare cura, è necessario che l'utente sviluppi uno sguardo critico su ciò che sta guardando: non accettare passivamente il video, ma analizzarlo, chiedendosi se quel personaggio o quel prodotto abbia effettivamente senso in quel contesto. In questo senso si parla anche di **brainrot**, ovvero **contenuti a consumo rapidissimo** e privi di reale valore informativo o narrativo, il cui unico scopo è generare engagement attraverso l'assurdo.

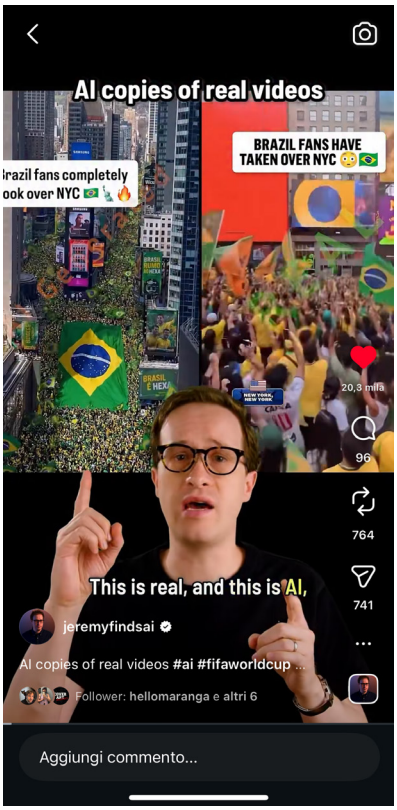
Un esempio calzante è il fenomeno dell'**Italian Brainrot**, il filone di personaggi generati dall'IA diventato virale su TikTok nel 2025, da **Tralalero Tralala** (lo squalo con tre gambe e scarpe Nike) a **Bombardiro Crocodilo** (l'ibrido tra un bombardiere e un coccodrillo) fino a **Tung Tung Tung Sahur**, che pur nella loro apparente innocuità, dimostrano quanto sia ormai semplice produrre contenuti sintetici capaci di ottenere milioni di visualizzazioni.

Proprio per rispondere a questa **esigenza di alfabetizzazione visiva**, stanno emergendo divulgatori specializzati nel riconoscimento di contenuti generati dall'AI. Un caso rappresentativo è quello di **Jeremy Carrasco**, produttore video statunitense noto sui social con il nome **Jeremy-FindsAI.** Nei suoi contenuti analizza video virali, spiegando nel dettaglio i segnali che ne rivelano la natura sintetica: posizioni di oggetti sullo sfondo poco plausibili, riflessi alterati su superfici di vetro o metalliche, movimenti innaturali di animali e persone, maniglie delle porte posizionate in modo anomalo e principi di fisica dei materiali che non corrispondono a quelli della realtà.



figura 46 *Alcuni personaggi dell'Italian Brainrot generati con l'intelligenza artificiale, tra cui Tralalero Tralala, Cappuccina Ballerina e Tung Tung Tung Sahur.*









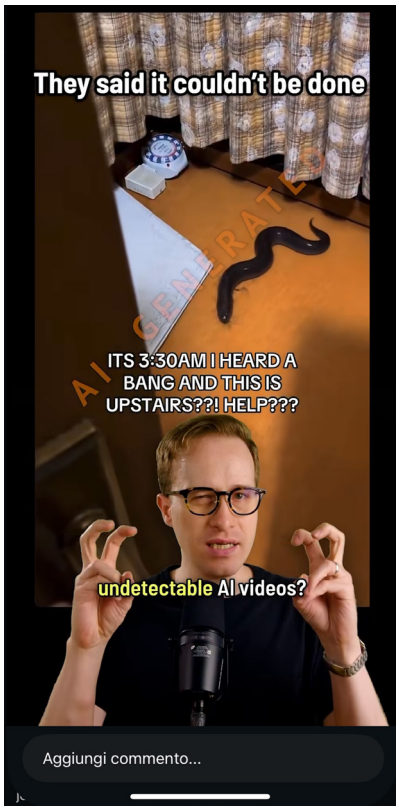


figura 47 *A sinistra, Jeremy Carrasco, noto sui social come JeremyFindsAI; a destra, una serie di screenshot dei suoi video divulgativi dedicati al riconoscimento di contenuti generati dall'intelligenza artificiale.*

Ma **come reagisce l'opinione pubblica** di fronte a contenuti generati dall'intelligenza artificiale, in particolare a quelli che rientrano nella categoria del cosiddetto AI slop?

Negli ultimi anni, man mano che l'IA è diventata sempre più pervasiva nella vita quotidiana e al **centro di discussioni sui social**, nei media e nei forum online, una parte del pubblico più attento ha iniziato a guardare con sospetto al suo impiego nella produzione creative ed artistiche, temendo che ne indebolisca la percezione di autenticità.

Un caso emblematico è quello del rapper Tyga che, con il suo **album Starface (2026)**, un concept ispirato agli anni ottanta, è stato realizzato dichiaratamente con l'ausilio dell'IA ed ha ricevuto **forti critiche negative**, tra cui uno zero su dieci da parte della testata musicale Pitchfork.
Lo stesso artista ha difeso la propria scelta paragonando l'uso dell'IA a quello dell'Auto-Tune: uno strumento tecnologico che molti inizialmente rifiutano, ma che gli artisti possono integrare nel proprio linguaggio invece di sostituirlo. Ma la stroncatura non è arrivata solo dalla critica professionale, ma anche da una parte del pubblico più consapevole e attento, come testimoniano diversi commenti sulla piattaforma di recensioni Album of the Year (AOTY).
Questo caso mediatico mostra come il web reagisca spesso con durezza quando artisti, musicisti o game developer dichiarano apertamente di utilizzare l'intelligenza artificiale nei loro processi creativi. Allo stesso tempo, è plausibile che molte persone ascolteranno brani prodotti con l'ausilio dell'AI perché proposti dagli algoritmi delle piattaforme di streaming o trasmessi in radio, senza saperlo. In questo scenario, come si è osservato in precedenza, l'utente non ha solo il diritto, ma anche il dovere, di non consumare passivamente ciò che gli viene proposto, interrogandosi su chi sia l'artista, con quali collaboratori abbia lavorato e in che modo gli strumenti di intelligenza artificiale siano stati integrati nel processo creativo.

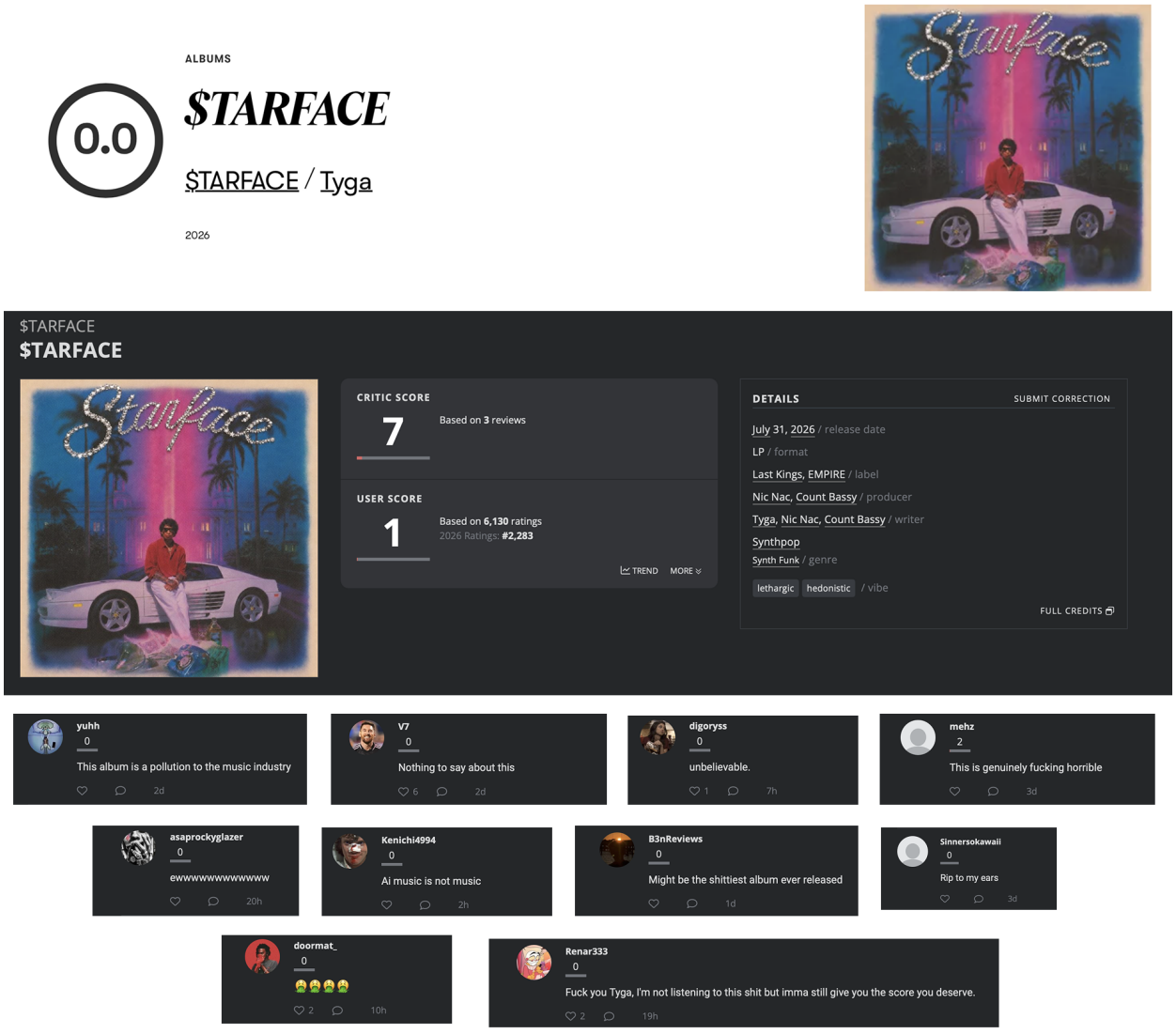


figura 48 *In alto, la recensione di Pitchfork che assegna all'album Starface dell'artista Tyga un punteggio di 0,0; in basso, la pagina di Album of the Year, con la valutazione della critica, il voto del pubblico e una selezione di commenti rappresentativi della ricezione negativa dell'album.*

3.4 Costi ed impatti ambientali

Quando si parla di intelligenza artificiale, è fondamentale tenere conto che l'allenamento dei modelli richiede ingenti risorse computazionali. Il luogo dove le IA vengono effettivamente addestrate e "vivono" è il **data center**: strutture fisiche, spesso enormi capannoni industriali, che **ospitano migliaia di server** e che forniscono lo spazio di calcolo e di archiviazione necessario al funzionamento di questi sistemi, operando ininterrottamente **24 ore su 24.**
All'interno di un datacenter, l'unità organizzativa fondamentale è il **rack**: uno **scaffale metallico standardizzato**, verticale, pensato per ospitare in modo ordinato e compatto l'hardware informatico: server, schede video, sistemi di rete o storage, impilati uno sopra l'altro per massimizzare lo spazio disponibile. Un singolo rack destinato al training di modelli AI può contenere fino a 8 GPU ad alte prestazioni, per un costo complessivo compreso tra 50.000 e 100.000 dollari, pari a circa 10.000 dollari per singola GPU.



figura 49 *Rack di server: infrastruttura hardware impiegata per l'archiviazione e l'elaborazione di grandi quantità di dati.*

Nel 2026 i prezzi delle GPU top di gamma Nvidia (H100, H200, B200) restano su questi ordini di grandezza o superiori, a seconda della configurazione richiesta per il training. Un datacenter di dimensioni industriali può ospitare centinaia o migliaia di questi rack, ciascuno dei quali genera calore che deve essere **costantemente dissipato.**
Da qui l'importanza del raffreddamento, che analizzeremo nel dettaglio più avanti. L'impatto dei datacenter si manifesta principalmente attraverso il **consumo energetico e idrico**, le emissioni dirette dei generatori di emergenza e quelle indirette legate alla filiera tecnologica.





figura 50 e 51 *Interno ed esterno di un data center: infrastrutture fisiche che ospitano server, sistemi di archiviazione e impianti di raffreddamento necessari all'elaborazione di grandi quantità di dati.*

I consumi energetici e di acqua

l'addestramento dei modelli di deep learning richiede **ingenti quantità di energia**. Uno studio dell'Università del Massachusetts Amherst ha stimato che l'addestramento di un singolo grande modello linguistico, come Transformer, ha consumato 656.347 kWh di elettricità, emettendo 626.155 kg di CO2 equivalente, quasi cinque volte le emissioni nell'intero ciclo di vita di un'automobile americana media.

A livello globale, il consumo energetico dei data center legati all'AI è passato da circa 85 TWh nel 2024 a una stima di **200 TWh nel 2026.** Sapere con certezza i consumi complessivi dell'AI è difficile, poiché i colossi tech sono poco trasparenti sulle proprie stime interne.
Più agevole è invece calcolare i consumi dei datacenter nel loro complesso: secondo l'**IEA (International Energy Agency)**, a livello globale nel 2024 i datacenter hanno consumato circa l'1,5% dell'energia elettrica mondiale, con una crescita del 12% annuo negli ultimi cinque anni.

Un problema aggiuntivo riguarda la **concentrazione geografica** di queste infrastrutture in cluster densi, che genera disagi concreti per la produzione e distribuzione locali di energia elettrica nelle aree circostanti. A ciò si aggiunge la pressione esercitata sulle **risorse idriche locali:** l'elevata domanda d'acqua necessaria al raffreddamento può entrare in competizione con gli usi domestici, agricoli e industriali del territorio, soprattutto nelle aree già soggette a stress idrico. In alcuni casi, le comunità residenti nelle vicinanze hanno segnalato **criticità relative alla qualità dell'acqua**, ai pozzi privati e alla capacità delle infrastrutture locali di sostenere una domanda così elevata. segnalato problemi relativi ai pozzi privati e alla capacità della rete idrica locale.

Un ulteriore effetto riguarda **l'inquinamento acustico**. Ventole, chiller, sistemi di ventilazione e generatori di emergenza producono infatti un rumore di fondo continuo che, se il datacenter è collocato in prossimità di abitazioni, può essere percepito anche all'interno delle case. Pur non superando necessariamente le soglie normative previste, questo **ronzio costante** può incidere negativamente sulla





figura 52 e 53 *Data center in prossimità di abitazioni e Beverly Morris, residente in Georgia, che denuncia problemi al proprio pozzo privato dopo la costruzione di un vicino data center di Meta, costringendola a usare acqua trasportata in secchi per lo scarico del bagno.*

qualità della vita, disturbando il sonno, la concentrazione e la possibilità di usufruire degli spazi esterni.

Inoltre, le proiezioni IEA indicano che, entro il 2030, l'elettricità consumata dai data center supererà i 945 TWh, più del triplo del fabbisogno energetico annuo dell'Italia, che si attesta intorno ai 300 TWh. Secondo lo MIT Technology Review, entro il 2028 più della metà dell'elettricità destinata ai data center sarà utilizzata specificamente per carichi di lavoro legati all'intelligenza artificiale.

Ma da cosa dipendono i consumi dell'AI? Per comprendere questi numeri è utile distinguere due fasi:

- **Addestramento dei modelli**: la quantità di dati e calcoli necessari è cresciuta esponenzialmente negli anni. Secondo alcune stime, GPT-4 sarebbe stato addestrato su circa 4,9 trilioni di datapoint, richiedendo un numero enorme di GPU operanti in parallelo per settimane o mesi.

- **Utilizzo dei modelli (inferenza):** dipende da fattori come la lunghezza del prompt e della risposta generata, le dimensioni del modello utilizzato e la modalità di input/output richiesta. Emerge qui una differenza cruciale: rispondere a domande testuali comporta emissioni significativamente inferiori rispetto alla generazione di immagini o video, che richiedono una potenza di calcolo molto più elevata per ogni singola unità di output prodotta. Inoltre data la grande quantità di mole di calcolo il raffreddamento delle GPU richiede grandi quantità d'acqua poiché il calore generato dai processori deve essere dissipato per evitarne il danneggiamento.

Nel 2023 i data center statunitensi hanno consumato complessivamente oltre 283,9 miliardi di litri d'acqua, con una media giornaliera per impianto di 1,7 milioni di litri.
Questo significa che un singolo data center di medie dimensioni consuma ogni giorno una quantità d'acqua pari a quella consumata quotidianamente da circa 1.500 famiglie americane.
A questo consumo diretto si aggiunge un **consumo idrico indiretto**, legato alla produzione dell'elettricità necessaria al funzionamento delle server farm: secondo una

ricerca pubblicata su Nature, questa componente "invisibile" rappresenta in realtà la fetta più ampia dell'impronta idrica complessiva di queste strutture.
Se per la generazione di energia i data center possono ricorrere anche a fonti rinnovabili, l'acqua utilizzata per il raffreddamento viene invece prelevata direttamente da fiumi e bacini d'acqua dolce limitrofi e, una volta impiegata, evapora in gran parte a causa delle alte temperature dei server, rendendola non riutilizzabile.

Un altro impatto riguarda il **consumo di suolo**, dato che queste strutture occupano superfici enormi. Vi sono inoltre impatti meno noti: i datacenter dispongono di generatori diesel di backup che, pur essendo destinati all'uso in emergenza, vengono accesi regolarmente per test e manutenzione, emettendo comunque gas serra. Infine, la costruzione stessa richiede **grandi quantità di materie prime** critiche come litio e cobalto, mentre la frequente sostituzione dell'hardware genera enormi quantità di rifiuti elettronici.

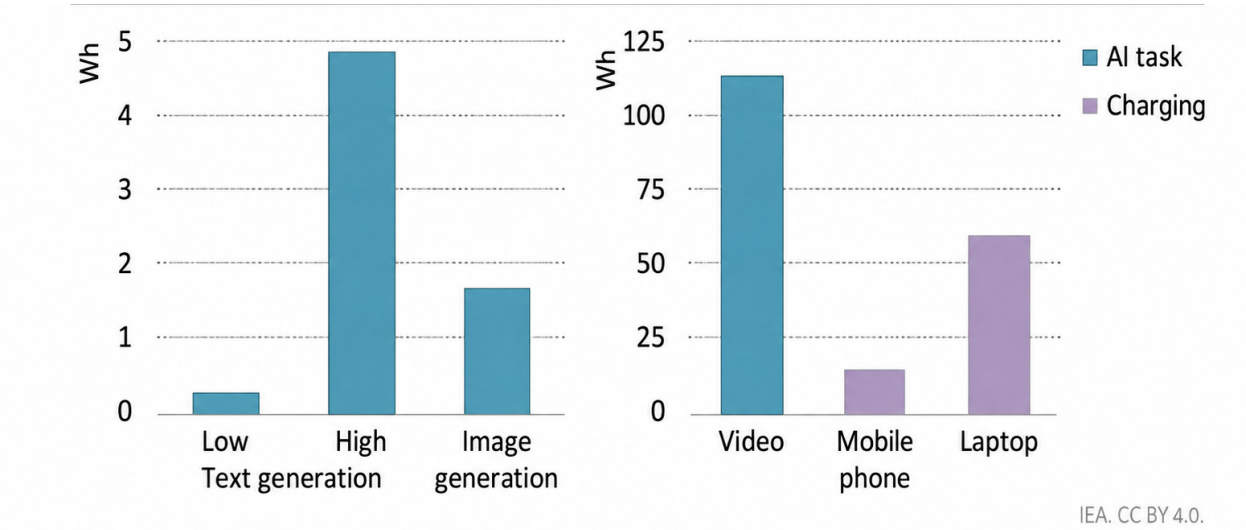


figura 54 *Grafico che mostra le stime dei consumi dell'intelligenza artificiale in base all'output generato (testo o immagini), confrontato con i consumi per ricaricare un cellulare e un computer. La generazione di immagini e soprattutto di video, richiede consumi sensibilmente più elevati rispetto alla generazione di testo (fonte: IEA).*

La trasparenza sui datacenter

Un aspetto critico riguarda la trasparenza dei dati raccolti sulla **distribuzione geografica dei datacenter**: secondo diverse infografiche disponibili pubblicamente, gli Stati Uniti ospitano di gran lunga il maggior numero di data center al mondo, oltre 3.700, circa 8 volte più di qualsiasi altro paese seguiti da Germania, Regno Unito e Cina.

Questo dato solleva un interrogativo metodologico rilevante: è plausibile che la Cina, seconda potenza tecnologica mondiale, disponga di un numero di data center così nettamente inferiore rispetto a paesi come Germania o Regno Unito? La discrepanza suggerisce che i dati raccolti da queste classifiche, spesso basate su registri commerciali pubblici occidentali, non colgano l'effettiva distribuzione delle infrastrutture cinesi, gestite in modo meno trasparente e spesso direttamente dallo Stato. Se i dati alla base della nostra stessa analisi non sono realmente rappresentativi, diventa difficile condurre una riflessione etica solida sull'impatto ambientale e geopolitico dell'IA: il problema della trasparenza non riguarda solo i dataset di addestramento, ma anche i dati che dovrebbero permetterci di valutare criticamente l'industria stessa.

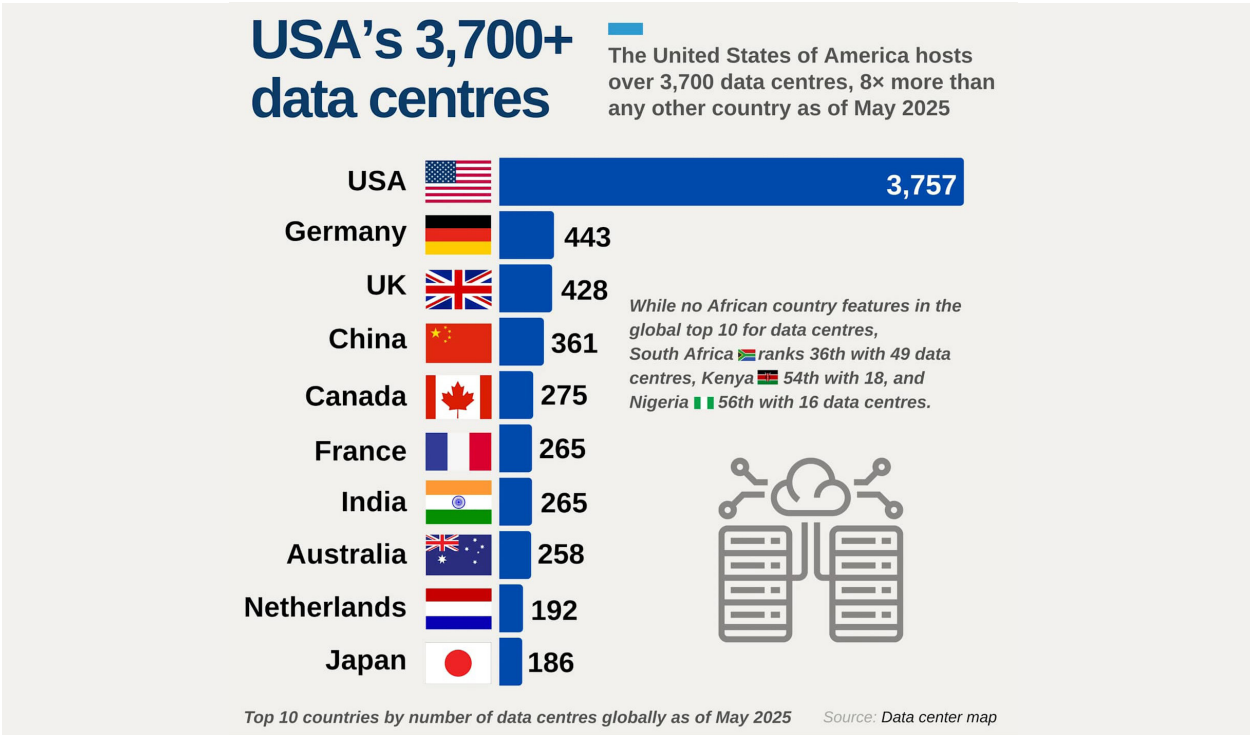


figura 55 *Distribuzione mondiale dei data center secondo i dati pubblici di Data Center Map (maggio 2025): gli Stati Uniti risultano al primo posto con 3.757 strutture, seguiti da Germania, Regno Unito e Cina. La rappresentazione evidenzia tuttavia i limiti di trasparenza e completezza delle classifiche basate su registri pubblici e commerciali (fonte: Data center map).*

Temi centrali di riflessione

Da questa riflessione emergono diversi nodi strutturali del settore.

Il primo riguarda **l'oligopolio privato**: nel panorama occidentale, lo sviluppo dell'AI è concentrato nelle mani di un **ristretto numero di aziende** estremamente potenti: OpenAI, Google, Microsoft, Meta, Amazon che rischiano di plasmare unilateralmente il mondo a venire, limitando la competizione e creando barriere d'ingresso pressoché insormontabili per le realtà più piccole.

Strettamente connesso a questo è lo **sfruttamento delle risorse minerarie**: l'hardware avanzato necessario all'AI, GPU e TPU in primis, dipende da risorse come silicio, litio e altri **metalli rari**, la cui **estrazione intensiva** solleva preoccupazioni ambientali e sociali significative, in un contesto di risorse naturali sempre più limitate.

A questi due nodi, di natura più propriamente economica e industriale, si affianca una terza criticità di ordine culturale: la **non-neutralità tecnologica**. Come già discusso a proposito dei bias nei training set, le tecnologie AI non sono mai neutrali, ma riflettono i pregiudizi e gli stereotipi impliciti nei dati su cui vengono addestrate.

Il quarto nodo, infine, riguarda lo **sfruttamento del lavoro** nella fase di addestramento dei modelli: come abbiamo già scritto in precedenza parlando del **Ghost Work**, la preparazione e l'etichettatura dei dati (data labeling) sono spesso affidate a lavoratori sottopagati nei paesi in via di sviluppo, un aspetto della filiera dell'AI tanto essenziale quanto raramente discusso, se confrontato con l'attenzione mediatica riservata all'innovazione tecnologica che ne deriva.

Ciò che tiene insieme questi nodi, in ultima analisi, è un problema comune e trasversale: la mancanza di trasparenza da parte delle aziende, che rende estremamente difficile formulare previsioni accurate tanto sui consumi energetici quanto sull'impatto reale di queste tecnologie sulla società.



Capitolo 4

Autorialità ed estetica delle immagini algoritmiche

4.1 L'autore nell'era algoritmica

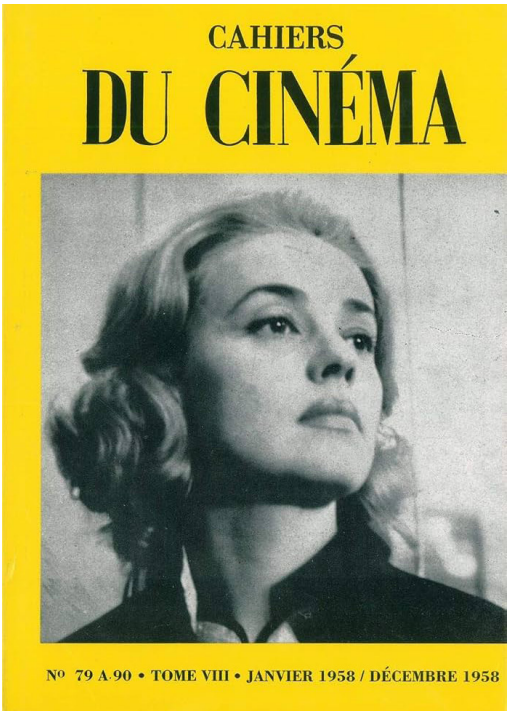


figura 56 *Copertina del numero 79 dei Cahiers du cinéma (1958), storica rivista francese di critica cinematografica.*

In questo ultimo capitolo analizzeremo la questione delle immagini algoritmiche dal punto di vista di chi crea contenuti, come si evolve la figura del creativo/filmmaker in quest ambito e quali sono le caratteristiche estetiche del linguaggio AI che possono essere sfruttate in maniera creativa.
Il cinema, e più in generale il linguaggio audiovisivo, ha sempre presupposto un **imponente lavoro collettivo**. Basta guardare, ad esempio, alla fine di un film, lo scorrere dei titoli di coda: registi, sceneggiatori, produttori, tecnici della fotografia e del suono, montatori, esperti di effetti speciali, costumisti, truccatori, compositori e via dicendo. Il cinema, così come altre forme di linguaggio audiovisivo affni, dagli spot pubblicitari ai videoclip, dai documentari alla videoarte, è un'**arte intrinsecamente collettiva**, in cui il dialogo e la collaborazione tra le persone sono tanto fondamentali quanto la componente tecnica.
Con l'avvento dell'IA generativa, questa collettività si concentra progressivamente nelle mani di una sola persona, o di un gruppo estremamente ristretto. Il filmmaker non è più colui che coordina e dialoga con i diversi reparti produttivi, ma diventa, potenzialmente, lui stesso l'intera troupe. Si delinea così un **cinema autarchico e solitario**, in cui la figura del regista si avvicina sempre più a quella di un **"programmatore"** che **dialoga e interroga la macchina** alla ricerca dello shot perfetto, piuttosto che a quella del regista nel senso tradizionale del ter-

mine. Questa dinamica trova un precedente teorico nella **politique des auteurs**, la corrente critica sviluppata in Francia negli anni Cinquanta da François Truffaut sulle pagine dei **Cahiers du Cinéma**, secondo cui "non ci sono opere, ci sono solo autori". Se la politique rivendicava già negli anni Cinquanta la centralità assoluta del regista come unico vero autore dell'opera cinematografica, l'IA generativa porta questa visione alle sue estreme conseguenze pratiche: oggi un singolo individuo può letteralmente scrivere, dirigere, "fotografare" e comporre la colonna sonora di un intero film, senza il bisogno di alcuna troupe.

Un caso emblematico è quello del regista americano Paul Trillo che nel 2023 ha realizzato **Thank You For Not Answering**, uno dei primi cortometraggi interamente generati tramite l'AI Runway Gen-2.
Nei titoli di coda del film si scrive "**Written and 'Directed' by Paul Trillo"**, con la parola "Directed" tra virgolette, un dettaglio che segnala esplicitamente l'ambiguità del ruolo autoriale in un processo in cui, come lo stesso Trillo ha dichiarato, l'IA "sta prendendo molte decisioni che io non ho preso".

Nel 2024, sempre Trillo ha diretto **The Hardest Part**, il primo video musicale ufficialmente commissionato e realizzato con Sora di OpenAI: il processo ha previsto la generazione di circa 700 clip, di cui solo 55 sono state selezionate e assemblate nel montaggio finale.

Nonostante questa radicale trasformazione dei processi produttivi, permangono importanti continuità con la tradizione cinematografica: sebbene l'IA generativa consenta di snellire notevolmente i costi di produzione e renda accessibili a filmmaker indipendenti o a basso budget realizzazioni altrimenti impossibili, il **linguaggio visivo** resta sostanzialmente **lo stesso del cinema tradizionale**: movimenti di macchina, composizione dell'inquadratura, gestione della luce, scelta cromatica. Il **regista-programmatore** deve dunque possedere una solida padronanza degli elementi tecnici e grammaticali del linguaggio cinematografico, per tradurli efficacemente in istruzioni testuali.
In questa prospettiva, **il prompt diventa il nuovo spa-**



figura 57 *Sequenza di fotogrammi di Thank You for Not Answering, cortometraggio di Paul Trillo realizzato con Sora di OpenAI. L'opera alterna immagini evocative accompagnate da una voce narrante e si conclude con il cartello «Written and "Directed" by Paul Trillo», nel quale il termine "Directed" è posto significativamente tra virgolette. Disponibile in streaming su YouTube.*

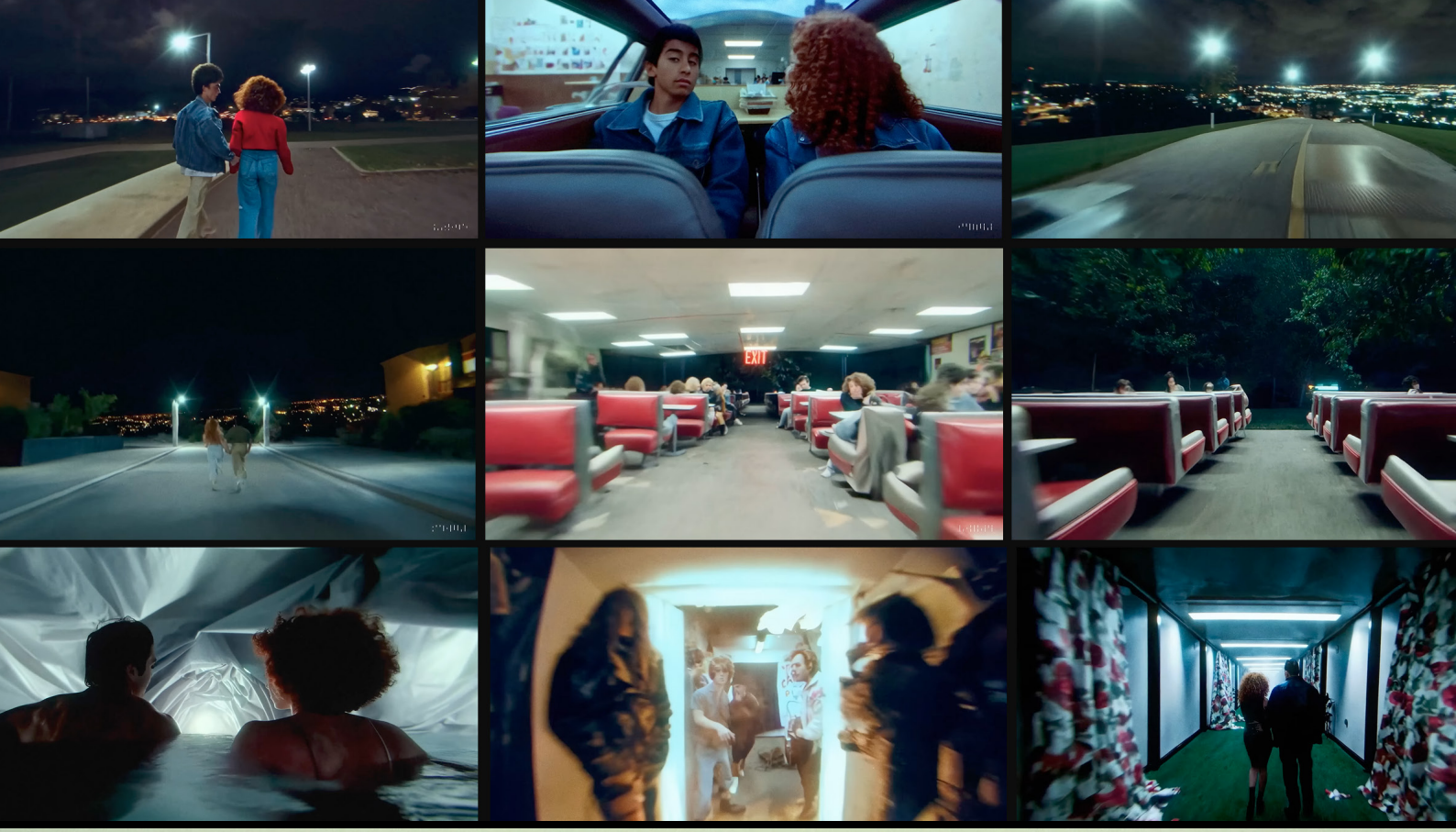




figura 58 *Sequenza di fotogrammi di The Hardest Part, videoclip dell'artista Washed Out diretto da Paul Trillo. Il video sviluppa un immaginario ispirato agli anni Ottanta attraverso una successione continua di immagini che si trasformano e confluiscono l'una nell'altra, creando un effetto di morphing assimilabile a un tunnel visivo. Disponibile in streaming su YouTube.*

zio di regia: il filmmaker può dare la propria visione del racconto, decidendo non solo cosa raccontare, ma anche come raccontarlo, attraverso la scelta precisa di termini tecnici e stilistici capaci di orientare la macchina verso l'estetica desiderata.

È in questo senso che si può parlare anche di un **cinema espanso**, recuperando la nozione teorizzata nel 1970 da **Gene Youngblood** nel suo **Expanded Cinema**; in questo testo l'autore ha anticipato l'idea di un cinema che si espande oltre i propri confini tecnologici e mediali tradizionali, includendo progressivamente nuovi strumenti (il video, il computer, l'olografia) come altrettanti "medium" con cui praticare la settima arte.

In questo contesto, il **montaggio** assume una connotazione particolarmente rilevante: sebbene la creazione delle singole immagini sia delegata alla macchina, l'atto di unire i diversi frammenti **rimane un gesto umano**. C'è sempre un autore dietro che decide quale storia raccontare e come ricomporre i diversi frammenti generati in una struttura narrativa coerente, un principio che richiama direttamente la teoria del montaggio di **Sergej Ejzenštejn**, per cui il montaggio non è mai una semplice operazione tecnica di taglio e incollatura, ma il vero e proprio **"metodo compositivo"** attraverso cui si costruisce il significato stesso dell'opera cinematografica. Anche nel cinema generato dall'AI, dunque, è nel montaggio che si concentra ancora oggi il nucleo irriducibile del gesto autoriale e non solo nella generazione della singola immagine.





figura 59 e 60 *In alto una fotografia dell'installazione Movie-Drome di Stan VanDerBeek, ambiente audiovisivo immersivo a proiezione multipla; in basso, Sergej Ejzenštejn al lavoro alla moviola. Regista e teorico sovietico, Ejzenštejn concepì il montaggio come un metodo compositivo capace di produrre il significato dell'opera cinematografica.*

4.2 Le nuove forme del visibile

Analizzare l'estetica dei filmati prodotti con l'intelligenza artificiale è un'operazione complessa, poiché entrano in gioco numerosi fattori: la tipologia di modello utilizzato, infatti alcuni modelli funzionano meglio per certe categorie di immagini o determinati movimenti di camera, e la costante evoluzione tecnologica, con nuovi modelli rilasciati ogni pochi mesi, sempre più performanti e capaci di restituire sempre di più un look progressivamente più fotorealistico.
Per semplificare l'analisi, è utile distinguere **due macrocategorie** estetiche ricorrenti nella produzione contemporanea: da un lato le opere che cercano di **emulare il fotorealismo**, dall'altro quelle che sfruttano le **allucinazioni dell'IA** come strumento tecnico ed espressivo del linguaggio algoritmico. Analizzeremo pro e contro di ciascuna scelta, portando casi studio concreti.

Estetica fotorealistica

In questa categoria rientrano tutti quei filmati che cercano di replicare, nel modo più veritiero e credibile possibile, **l'esperienza della ripresa dal vero**: movimenti corporei naturali, riflessi di superfici e gestione accurata della luce. Da un punto di vista pratico, questa tipologia di generazioni di AI offre una serie di vantaggi concreti.

Anzitutto, **l'abbattimento dei costi**: non è necessario affittare location, acquistare materiali di scena o retribuire attori e professionisti, con un conseguente snellimento del workflow produttivo, poiché, se il processo diventa più leggero, si riduce anche il bisogno di personale. A questo si aggiunge il **dialogo in linguaggio naturale**: per lavorare con questi algoritmi si utilizza la lingua parlata comune, l'italiano o l'inglese, dialogando con la macchina come si farebbe con un collaboratore umano. Infine, questi strumenti abilitano l'esplorazione di immagini inedite, aprendo

a nuove idee di cinema e a nuove possibilità formali, con un potenziale di grande soddisfazione per gli sperimentatori. Tuttavia, questa tipologia di produzione presenta **criticità significative.**

La prima riguarda l'**accentramento del potere** nelle mani di pochi: si perde l'intero apporto delle maestranze, ovvero di tutti quei creativi e tecnici che permettono la messa in scena concreta delle idee di regista e sceneggiatore, così come si perde la dimensione collettiva e collaborativa tra i diversi reparti. Con l'IA, l'autore condensa su di sé facoltà che prima appartenevano a professionalità distinte: questo apre possibilità interessanti, come il regista come unico ideatore dell'intero universo narrativo, ma comporta anche problematiche significative, prima fra tutte la **perdita di quel rapporto collaborativo** con attori e direttori della fotografia che da sempre costituisce uno dei nodi centrali della produzione cinematografica.

A questo si aggiunge il **rischio di un appiattimento estetico**: adottando uno specifico algoritmo di generazione, si eredita automaticamente anche la grana e la specificità stilistica. Inquadrature realizzate con lo stesso algoritmo tendono a condividere punti di fuga simili, composizioni analoghe, una grana pressoché identica, un identico trattamento fotografico e ambientazioni ricorrenti, come ad esempio luci al neon che conferiscono sistematicamente la stessa sensazione atmosferica. Questo fenomeno rappresenta un **attacco al concetto di autorialità**: se tutti possono accedere alla stessa idea di cinema attraverso lo stesso strumento, il rischio è un'omologazione complessiva dell'estetica del cinema algoritmico.
È in questo contesto che risultano centrali tre concetti tecnici ricorrenti nella critica al cinema algoritmico: il flicker, il morphing e l'uncanny motion.

Il flicker (letteralmente "sfarfallio") indica l'incapacità del modello di mantenere una coerenza visiva stabile tra un fotogramma e l'altro, che tecnicamente si definisce **coerenza temporale (temporal consistency).** Poiché molti modelli di generazione video basati su diffusione elaborano ogni fotogramma in modo relativamente indipendente, tutto ciò che il prompt non specifica con precisione, ad

esempio la trama esatta di un tessuto o la posizione di un dettaglio sullo sfondo, viene stabilito casualmente a ogni fotogramma. Il risultato visibile è uno **sfarfallio percettibile** soprattutto nelle **texture ad alta frequenza di dettaglio**, come capelli, tessuti o fogliame, che sembrano vibrare o "respirare" pur restando teoricamente immobili nella scena.

Il morphing, invece, è un effetto speciale utilizzato nel cinema e nell'animazione per **trasformare un'immagine o una forma in un'altra** attraverso una **transizione fluida e continua.** Tradizionalmente, questo tipo di effetto veniva ottenuto tramite tecniche di dissolvenza incrociata direttamente sulla pellicola, ma dai primi anni Novanta tale approccio è stato progressivamente sostituito da software digitali dedicati. Nel contesto dell'IA generativa, tuttavia, il morphing smette di essere una scelta stilistica intenzionale e diventa un **artefatto involontario**: caratteristiche di un volto, di un oggetto o di uno sfondo mutano gradualmente e senza controllo lungo la sequenza video, generando quella che viene definita **deriva temporale (temporal drift)**, a differenza del flicker, che è un fenomeno casuale e localizzato, il drift è invece progressivo e direzionale: un volto che cambia lentamente forma, un colore che vira impercettibilmente verso un'altra tonalità, un elemento di sfondo che "migra" nel corso dell'inquadratura.



figura 61 *Esempio di morphing facciale a diversi gradi di intensità (originali agli estremi, sinistra e destra).*

L'uncanny motion ("movimento perturbante") richiama, infine, il celebre concetto di uncanny valley ("valle perturbante") teorizzato nel **1970** dallo studioso di robotica giapponese **Masahiro Mori**. Secondo la sua ipotesi, il grado di familiarità e gradimento che un osservatore prova verso un robot o un automa antropomorfo cresce con l'aumentare della somiglianza con l'essere umano ma solo fino a un certo punto. Superata questa soglia, un realismo quasi perfetto ma non completo produce l'effetto opposto: un crollo improvviso del gradimento, sostituito da repulsione e disagio.

Come illustra il grafico, le due curve, quella degli oggetti fermi e quella tratteggiata degli oggetti in movimento, salgono insieme fino quasi al realismo totale per poi crollare bruscamente nella "valle" (zombie, cadavere, protesi) prima di risalire verso l'essere umano vero, con la linea tratteggiata che scende più in profondità e più bruscamente di quella continua.

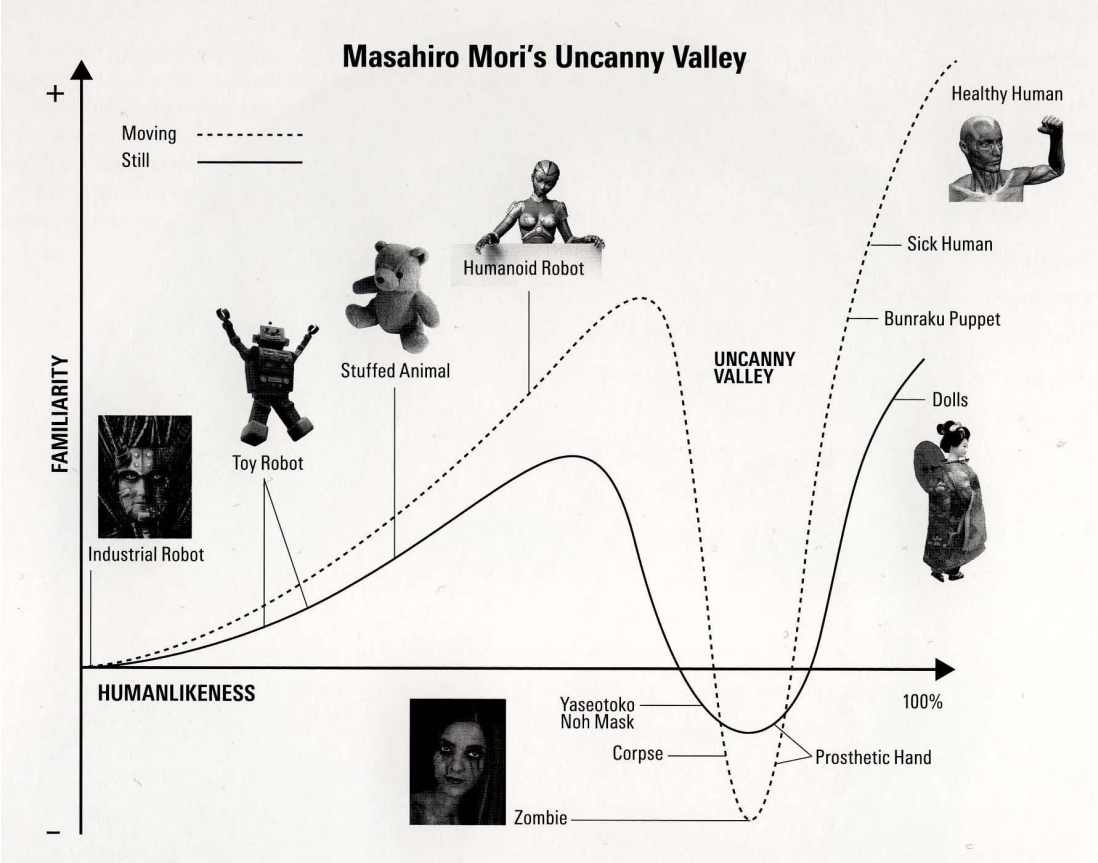


figura 62 *Rappresentazione grafica dell'uncanny valley teorizzata da Mori (1970): sull'asse orizzontale il grado di somiglianza con l'essere umano, su quello verticale il grado di familiarità, con le due curve (oggetti fermi e oggetti in movimento) che precipitano bruscamente prima di risalire verso l'essere umano reale.*

Mori sottolineava inoltre come il movimento amplifichi drasticamente questo effetto: un oggetto realistico ma immobile può risultare semplicemente inquietante, ma lo stesso oggetto animato con un movimento innaturale, rigido o leggermente sfasato rispetto alla fisica reale, precipita bruscamente nel fondo della "valle".
Applicato al video generato dall'IA, l'uncanny motion descrive proprio quei movimenti corporei, facciali o di camera che risultano tecnicamente fluidi ma percettivamente "sbagliati": micro-esitazioni innaturali, accelerazioni improvvise, gesti che non rispettano pienamente le leggi della fisica, capaci di tradire istantaneamente la natura sintetica dell'immagine anche quando il singolo fotogramma, osservato staticamente, appare perfettamente fotorealistico.

Adesso analizzeremo una serie di casi studio legati a prodotti che cercano di essere "fotorealistici"

Un primo esempio è **The Woman with Red Hair** (Staircase Studios AI, regia di Brett Stuart, 2025), **dramma sulla resistenza olandese** nella Seconda guerra mondiale basato su una sceneggiatura Black List del 2016, con character design curati dall'ex Pixar Teddy Newton (Gli Incredibili) e dall'animatore Alfred Gimeno.
Lo studio ha esplicitamente dichiarato l'obiettivo di **raggiungere un fotorealismo** "quasi da studio" attraverso la propria pipeline proprietaria ForwardMotion, capace, secondo l'azienda, di produrre film per meno di 500.000 dollari. Il primo assaggio pubblico ha però ricevuto **accoglienze quasi unanimi e negative**: le espressioni facciali e i movimenti dei personaggi sono apparsi particolarmente inquietanti, con un'evidente perdita di risonanza emotiva nonostante la drammaticità del tema trattato.

figura 63 *Poster e fotogramma estrapolato da "The Woman with red hair" (Staircase Studios AI, regia di Brett Stuart, 2025). Disponibile in streaming su YouTube.*

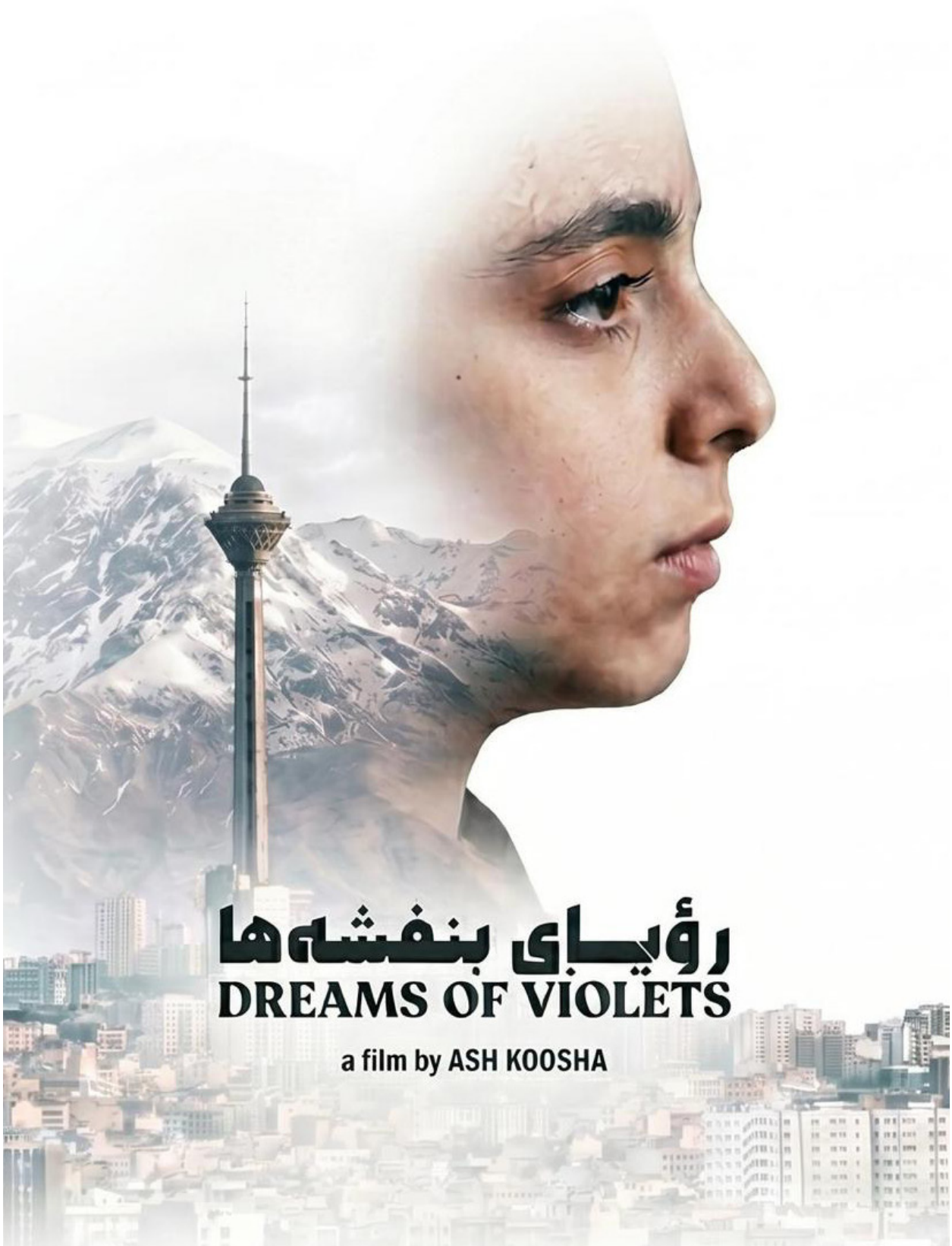

DIRECTED BY BRETT STUART
A STAIRCASE STUDIOS AI FILM
WRITTEN BY MICHAEL SCHATZ
THE WOMAN WITH Red HAIR
MAYA-NIKA BEWLEY
LEANDER VYVEY
GEOFFREY BRETON
JULIAN DEUSTER
MIRIAM LUCIA
EMILIA ROSE ARNÓ
MIRA PRANGER
ANGUS CASTLE-DOUGHTY
KOLA KRAUZE
TOR LIGHTEN
JOE REISIG
WILLIAM LUDWIG
MAX MARSTON
DANA SMIT
KILKE JOHN
PAUL MCEWAN
WILLIAM BRAND
LINN JOHANSSON

Un altro esempio significativo è **Dreams of Violets** di Ash Koosha, presentato in prima mondiale al **Tribeca Festival** nel **giugno 2026**: è il primo lungometraggio interamente generato tramite intelligenza artificiale a essere inserito nella selezione ufficiale di un grande festival cinematografico. Il film, realizzato con un **budget di circa 2.000 dollari** e senza attori, set o camere fisiche, ricostruisce gli eventi del massacro dei civili iraniani del gennaio 2026 attraverso lo sguardo di cinque personaggi, un'operazione che il regista ha definito necessaria proprio perché, altrimenti, impossibile da realizzare, data l'impraticabilità di girare fisicamente in Iran.

Dal punto di vista stilistico, il film mostra un uso di inquadrature, **profondità di campo, blocking** dei soggetti e **movimenti di camera estremamente fotorealistici**, tanto che in alcuni momenti risulta difficile riconoscere la natura generativa delle immagini. Koosha ha utilizzato principalmente **KlingAI**; anche nella produzione IA, del resto, si parla comunque di "costi di produzione", riferiti ai crediti necessari per l'utilizzo delle piattaforme di generazione.

Anche in questo caso la **ricezione critica** dell'opera è stata **molto divisa**. Da un lato vi è la lettura "democratizzante" proposta dal Guardian, che ha inquadrato il film attorno all'idea che l'equivalente in CGI tradizionale sarebbe costato milioni di dollari contro i soli 2.000 spesi da Koosha. Dall'altro, buona parte della critica ha bocciato duramente il risultato, descrivendo volti spesso spettrali, sospesi tra esistenza e sparizione, e strade di Teheran che appaiono insieme reali e immaginate.

figura 64 *A destra la locandina di Dreams of Violets (Ash Koosha, 2026), presentato in prima mondiale al Tribeca Festival.*


رؤیای بنفشه‌ها
DREAMS OF VIOLETS
a film by ASH KOOSHA

Un ulteriore esempio è **Hell Grind** un **action-fantasy di 95 minuti** realizzato da un team di 15 persone in circa due settimane dalla startup Higgsfield AI, con un **budget complessivo di 500.000 dollari**, di cui 40.000 destinati ai soli costi di calcolo. Il film è visionarie gratuitamente su YouTube. Nonostante l'ambizione fotorealistica, diversi giornalisti che hanno visto il film in anteprima a Cannes hanno descritto le scene d'azione come intense ma a tratti robotiche, pur risultando complessivamente credibili.



figura 65 *Locandina e fotogramma estrapolato da Hell Grind (Higgsfield AI, 2026). Disponibile in streaming gratuito su youtube.*

Un ultimo riferimento significativo è l'annuncio tecnico pubblicato da Runway nel maggio 2026, intitolato **Project Luxo**, in cui l'azienda sostiene esplicitamente di aver superato l'uncanny valley con i propri modelli di generazione video. Tra questi, il caso più emblematico è il cortometraggio **Last Night** di durata di 5 minuti e 26 secondi pubblicato su YouTube dal **canale ufficiale di Runway**, che racconta l'ultima notte di una relazione attraverso il filtro di ricordi frammentati, ambientata in una Tokyo notturna. Il dato più significativo, enfatizzato dalla stessa Runway nella propria comunicazione, riguarda le condizioni di produzione: il film è stato realizzato interamente da una sola persona, in appena 7 ore di lavoro. Si tratta di un caso emblematico del "cinema autarchico e solitario" descritto nel paragrafo precedente, portato qui alle sue estreme conseguenze produttive: un intero cortometraggio dotato di risoluzione narrativa, coerenza emotiva e ambientazione internazionale, realizzato in un'unica giornata lavorativa da un unico autore, comodamente seduto a casa propria.

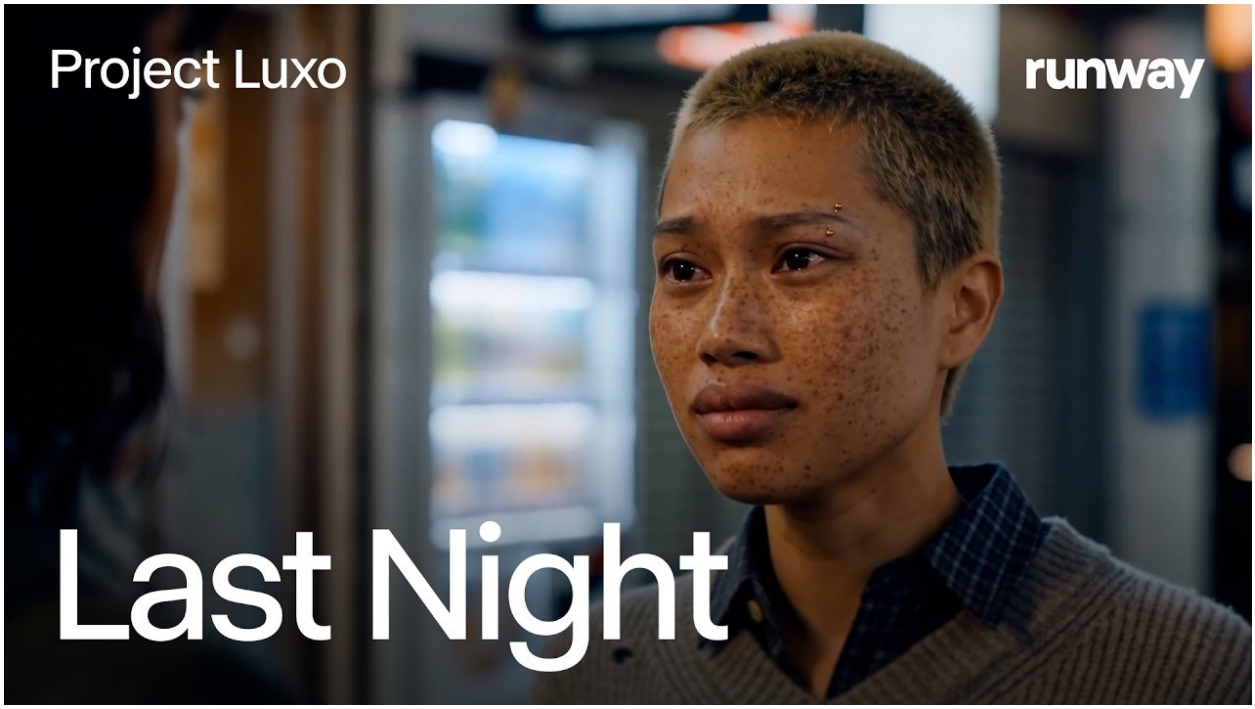


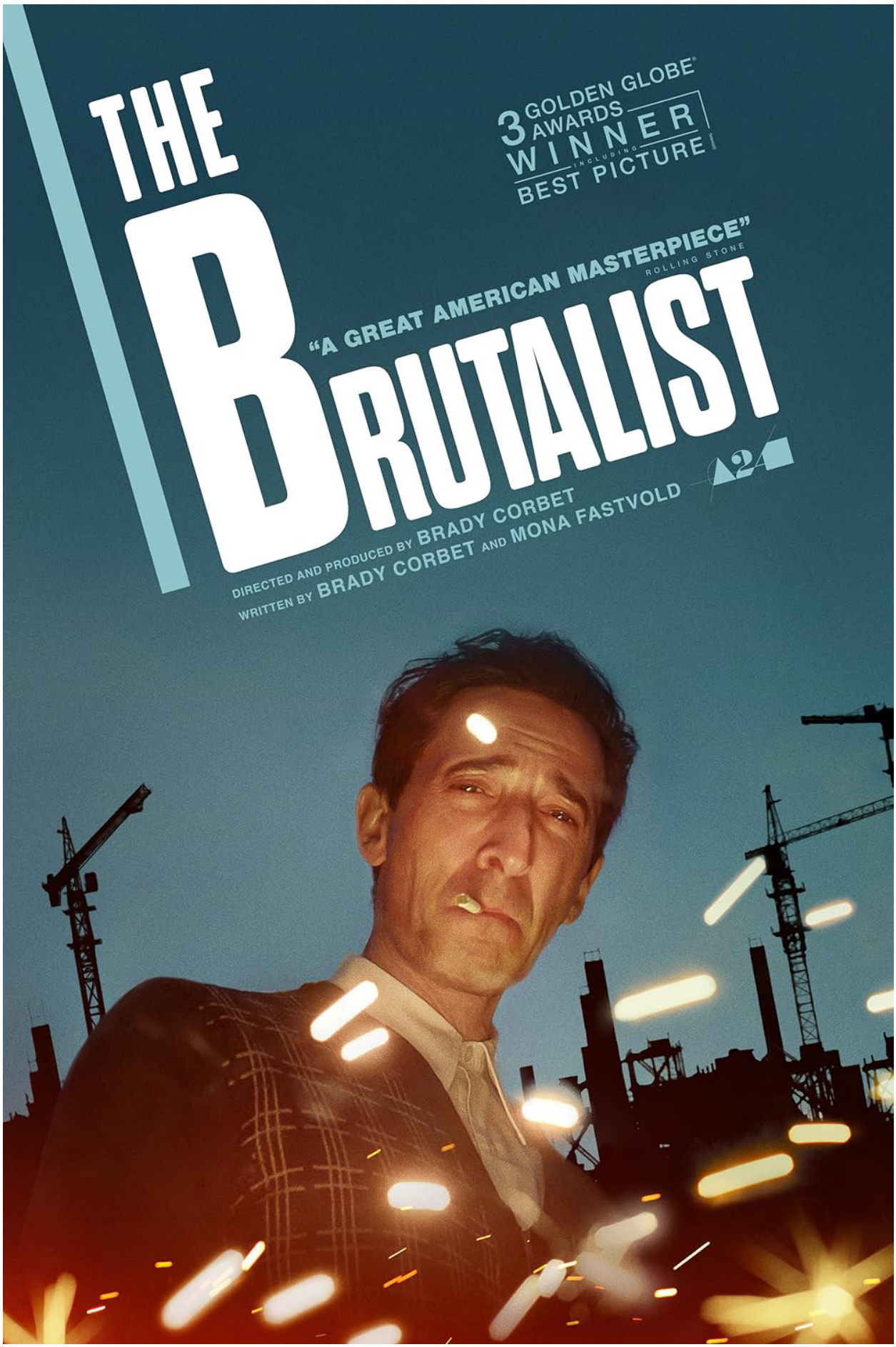
figura 66 *Fotogramma estrapolato da Last Night (Project Luxo, Runway, 2026), disponibile in streaming su YouTube.*

È comunque da tenere in mente che tutte queste produzioni, seppur siano molto più economiche rispetto a produzioni cinematografiche standard, prevedono dei **costi di generazione video**, dovuti ai **token** che si utilizzano sulle piattaforme di generazione, inoltre si deve anche aspettare il tempo che si genera il video

Infine, è da tenere in conto un ultimo utilizzo delle IA ovvero quello legato ad **effetti speciali e visivi**, che vengono impiegati con l'obiettivo di ridurre costi e tempistiche senza sostituire integralmente il lavoro umano.

Un esempio pragmatico è **The Brutalist** (2024) di Brady Corbet, film vincitore di tre Golden Globe e candidato a diecipremi Oscar, che nel gennaio 2025 è stato al centro di una **controversia mediatica** per l'uso dell'AI in post-produzione. Il montatore del film, l'ungherese Dávid Jancsó, ha rivelato in un'intervista a Red Shark News di aver utilizzato **Respeecher**, un software ucraino di voice conversion, per perfezionare la pronuncia ungherese degli attori protagonisti Adrien Brody e Felicity Jones.
È importante chiarire la natura tecnica dell'intervento: non si è trattato di sostituire le voci degli attori con un doppiaggio sintetico, ma di correggere, lettera per lettera, alcune vocali e consonanti particolarmente ostiche di una lingua notoriamente difficile da pronunciare per i non madrelingua, mantenendo intatte l'interpretazione, il ritmo e l'intensità emotiva della recitazione originale registrata sul se.

figura 67 *A destra la locandina di The Brutalist (Brady Corbet, 2024 al centro delle polemiche per l'uso di Respeecher in post-produzione.*


THE BRUTALIST
3 GOLDEN GLOBE® AWARDS
WINNER
INCLUDING
BEST PICTURE
"A GREAT AMERICAN MASTERPIECE"
ROLLING STONE
DIRECTED AND PRODUCED BY BRADY CORBET
WRITTEN BY BRADY CORBET AND MONA FASTVOLD
A24

Possiamo citare anche un altro esempio: il film di Robert Zemeckis **Here** (2024), con Tom Hanks e Robin Wright, che ha utilizzato la tecnologia **Metaphysic Live** per invecchiare e ringiovanire in tempo reale gli attori attraverso circa sei decenni di narrazione. A differenza delle tecniche tradizionali, che richiedono mesi di compositing in post-produzione, il sistema mostrava il **volto trasformato direttamente sul monitor durante le riprese**, con un ritardo di pochi fotogrammi, permettendo a Zemeckis di dirigere gli attori vedendo istantaneamente il risultato finale. Gli stessi interpreti utilizzavano uno "specchio della giovinezza" (youth mirror) per calibrare la propria performance osservando in diretta la versione ringiovanita del proprio volto.
Il caso è significativo perché dimostra come l'AI, in questo contesto, non sostituisca il lavoro degli attori ma lo integri nel momento stesso della recitazione, **eliminando i costosi processi di VFX tradizionali** che avevano fatto lievitare, ad esempio, il budget di The Irishman di Scorsese fino a oltre 200 milioni di dollari.



figura 68 e 69 *A sinistra il monitor sul set con l'applicazione in tempo reale del de-aging tramite Metaphysic Live; a destra la locandina ufficiale del film.*

Tom Hanks Robin Wright



FROM THE WRITER, DIRECTOR AND STARS OF
FORREST GUMP

Joy, Hope, Loss, Love, Life happens...

HERE

Exclusively In Movie Theatres
November 1

MIRAMAX PG-13 #HereMovie Here-Movie.com

Nel giro di pochi anni la tecnologia si è evoluta a una velocità impressionante, e il divario qualitativo rispetto alle produzioni tradizionali si è ridotto sensibilmente. Tuttavia, come dimostrano i casi analizzati, l'annuncio di un definitivo superamento della uncanny valley resta più una dichiarazione di posizionamento commerciale che un dato confermato dalla ricezione critica, che continua a rilevare gli stessi sintomi: volti spettrali, movimenti robotici, mancanza di una continuità effettiva tra le diverse inquadrature. A questo si somma una resistenza di natura diversa, non tecnica ma valoriale: una parte della critica e pubblico contesta a queste opere una legittimità artistica indipendentemente dalla loro qualità visiva, in quanto prive di quella componente umana e interpretativa considerata insostituibile nell'atto creativo.

Estetica "allucinatoria"

In questa categoria rientrano tutte quelle produzioni che rendono evidente ed esplicito l'uso dell'IA, trasformando le allucinazioni del modello in uno strumento creativo ed espressivo. Per **allucinazione** si intende la produzione, da parte di un sistema generativo, di contenuti plausibili ma non pienamente aderenti ai dati, al prompt o alle regole del mondo reale: nelle immagini e nei video ciò può manifestarsi attraverso deformazioni anatomiche, oggetti incongruenti, transizioni impossibili, mutazioni improvvise o scenari privi di una logica fisica coerente.
L'intento **non è emulare la realtà**, ma costruirne una nuova, sfruttando appieno le potenzialità specifiche del medium algoritmico.
Il vantaggio principale di questo approccio risiede nella sua **natura dichiaratamente artistica:** non si tratta di ottimizzare i costi emulando la produzione tradizionale, ma di esplorare le ragioni espressive proprie del mezzo.
Un riferimento centrale per questa concezione è il **Manifesto Umanista per un Cinema in Intelligenza Artificiale**, pubblicato dal **collettivo milanese Hariel (hariel.xyz).** Il manifesto rappresenta un tentativo di **formalizzare un approccio culturale all'immagine algoritmica**, individuando strategie per produrre immagini interessanti ed espressive all'interno di una disciplina che tende, strutturalmente, ad
appiattire originalità e autorialità.

Questi sono i diversi punti del manifesto:
0) Non ci sono più scuse. Chiunque, ora, può fare cinema.

Hai una responsabilità.
1) L'errore è tuo alleato. L'unica immagine interessante prodotta da un'IA è la sua allucinazione. Quell'istante in cui la macchina cessa di imitarci.
2) Non farti sedurre dall'alta definizione. Non temere il rumore e le imperfezioni. Saranno loro a rendere attraente la tua immagine algoritmica.
3) Il dialogo con la macchina è in linguaggio naturale. Devi conoscere le parole. E scegliere quelle giuste.
4) Ogni algoritmo comporta una diversa estetica. Lascia che esploda.
5) Mantieni un elemento umano in fase di produzione. Non c'è bisogno di altre specifiche.
6) Torna alla storia del cinema prima che la macchina la rimastichi per te.
7) La nostalgia per gli autori che hai amato è un veleno. L'IA open-source [il cui codice è accessibile e modificabile] è il suo antidoto. Desidera forme che siano solo tue.
8) Ogni immagine algoritmica ha un suo suono algoritmico. Ogni suono algoritmico ha una sua immagine algoritmica.
9) Studia la teoria, conosci il linguaggio cinematografico. Il resto sono processi automatici.
10) La tua identità d'autore risiede nella lotta contro la macchina. Non assecondarla. Mai. A meno che tu non la stia ingannando.
11) Persegui la deformità. Non c'è niente di più noioso, falso



figura 70 *Copertina del Manifesto Umanista per un Cinema in Intelligenza Artificiale, pubblicato dal collettivo milanese Hariel.*

e retorico del fotorealismo.
12) Un'immagine iniziale è tutto ciò di cui hai bisogno per la tua storia.
13) Un film con l'IA è sempre un film sull'IA. Ma fai attenzione: se la tua scrittura lo rivela, hai già fallito.

Tra i tredici punti che compongono il testo, alcuni risultano particolarmente rilevanti: **Il punto #4** afferma che ogni modello ha la propria estetica. **Ciascun modello generativo**, infatti, produce **immagini riconoscibili per inquadrature ricorrenti, palette cromatiche e movimenti di camera specifici**, dovuti alla propria programmazione e al proprio dataset di addestramento. Il compito dell'artista diventa allora quello di **ricercare modelli e tecniche diversi**, combinandone più di uno all'interno della stessa opera per aumentarne la ricchezza espressiva.
Il punto #7 riguarda la concezione di cinema come open source, dove l'artista deve poter accedere al codice del modello per **praticare il fine tuning**, ovvero il riaddestramento con dati specifici che permette di adattare un modello generico verso una visione personale e unica. Tecniche come la **LoRA (Lelow-Rank Adaptation)** permettono questo adattamento con dataset più piccoli e costi computazionali ridotti, rendendo possibile una vera rottura rispetto ai modelli closed source, che impongono invece una propria visione precostituita di cinema e dell'estetica.
Il punto #10 afferma l'identità autoriale del regista come lotta con la macchina. Secondo il manifesto, l'identità dell'autore risiede proprio nello **scontro con lo strumento algoritmico:** la macchina si aspetta un input interpretabile, mentre l'artista le fornisce qualcosa che sfugge parzialmente alla sua comprensione. Prima di poter forzare i limiti della macchina, tuttavia, è **necessario comprenderla e dialogare con essa** a fondo.

Sempre a partire da questo collettivo è nato **The Pornographer** (titolo stilizzato THE PØRNØGRAPHƏR), cortometraggio selezionato alla Settimana Internazionale della Critica della 82ª Mostra del Cinema di Venezia e vincitore del Premio Centro Nazionale del Cortometraggio. Il film **fonda la propria autorialità sul concetto di allucinazione dell'AI**: nessuno degli autori ha mai inserito nel prompt istruzioni per generare **"corpi deformi"**: le imma-



figura 71 *Locandina di The Pornographer (collettivo Hariel), selezionato alla Settimana Internazionale della Critica della 82ª Mostra del Cinema di Venezia.*

gini di corpi mutanti e irriconoscibili che caratterizzano il film sono **emerse spontaneamente dal comportamento imprevedibile del modello**, al punto che gli stessi autori non sono mai riusciti a riprodurre due volte la medesima immagine.
Le allucinazioni, in questo senso, si verificano quando un modello di AI genera output che **si discostano dal prompt fornito** o dalla propria struttura di addestramento prevista, ad esempio un modello nato per il fotorealismo che finisce per inventare forme nuove, imprecise e inattese. Questo accade perché i modelli generano statisticamente la sequenza più probabile di elementi, senza una reale comprensione semantica della realtà rappresentata; le cause principali risiedono in **dati di addestramento insufficienti o in bias** e interpretazioni errate intrinseche all'algoritmo. Se in ambiti come quello sanitario o dell'informazione le allucinazioni rappresentano un rischio serio, in ambito artistico esse offrono a registi e designer un approccio del tutto nuovo alla direzione artistica, aprendo la strada a scene surreali e oniriche.

Per la realizzazione del film, il collettivo Hariel ha sperimentato con **Stable Diffusion**, utilizzando modelli customizzati e riaddestrati da altri utenti della community, proprio per generare l'effetto di scarto rispetto al prompt originale: le immagini risultanti, infatti, non hanno quasi nulla a che vedere con le istruzioni testuali fornite, sfuggendo al controllo e ai limiti che gli autori intendevano imporre originariamente. Il risultato restituisce una **dimensione atemporale**, quasi onirica, che richiama tanto l'estetica del sogno quanto quella dell'incubo, in un registro apertamente espressionista.
L'errore, l'imprecisione e la deformazione tipiche delle immagini generate dall'intelligenza artificiale non vengono in alcun modo occultati, ma ostentati e trasformati in vera e propria materia estetica, un approccio che richiama da vicino la **filmografia di David Cronenberg**: (autore di La mosca, Videodrome, The Shrouds-Segreti sepolti) il regista canadese ha da sempre indagato il rapporto tra corpo, media e tecnologia, ed è proprio questa eredità che The Pornographer sembra raccogliere e trasporre nel dominio dell'AI, dove la **"nuova carne"** non nasce più da una mutazione biologica ma **da un processo generativo algoritmico.**



figura 72 *.Fotogramma tratto da The Pornographer: dettaglio di una fusione innaturale tra parti anatomiche (orecchie, piccolo braccio) generata algoritmicamente.*

Da un punto di vista teorico, il cortometraggio si oppone frontalmente alla logica di gerarchizzazione e indicizzazione dei corpi su cui si fonda il funzionamento stesso dell'AI: ogni immagine di training viene normalmente etichettata e classificata secondo un'ideologia implicita, cosa definisce un uomo, cosa definisce una donna, quali sono i connotati "normali" di un corpo. Tuttavia, le immagini prodotte da The Pornographer sfuggono deliberatamente a ogni tentativo di classificazione, complice anche la bassa risoluzione, che introduce una sfumatura indefinita nei contorni e nei generi rappresentati.



figura 73 *Fotogramma tratto da The Pornographer: più corpi si fondono in un'unica massa biologica indistinta e inquietante, dai tratti body horror che richiamano l'estetica del regista David Cronenberg.*

Un altro esempio che segue questo filone è **The Eggregores' Theory** (2024) del regista italo-greco Andrea Gatopoulos, primo cortometraggio realizzato quasi interamente con l'intelligenza artificiale ad aprire la Settimana Internazionale della Critica del Festival di Venezia.
Il film racconta, attraverso una **sequenza di fotografie statiche in bianco e nero** accompagnate da una voce narrante, la storia distopica della città di Lublich, dove alcune parole diventano letteralmente "velenose" e uccidono chi le pronuncia o le ascolta, fino a portare al rogo dei libri e alla creazione di un nuovo alfabeto artificiale. Tutte le immagini del film, comprese le lettere di un alfabeto inedito, generato combinando glitch tipografici prodotti dall'AI, sono state **create con Midjourney** e poi editate manualmente una a una dallo stesso regista per garantire coerenza visiva, in un processo che lui stesso descrive come il tentativo di "immaginare la storia come se fosse stata estratta da un archivio pre-censura, già masticato e distorto da una macchina" Il cortometraggio ha vinto il Gran Premio al festival ZINEBI di Bilbao ed è stato candidato agli European Film Awards e ai David di Donatello, con la giuria che ne ha elogiato "l'uso creativo e originale di tecniche come l'intelligenza artificiale, generalmente impiegate per sostituire gli artisti". Anche la colonna sonora del film è stata generata tramite AI, rendendo The Eggregores' Theory un caso emblematico opposto rispetto a produzioni hollywoodiane come The Brutalist o Here: mentre queste ultime usano l'AI come strumento discreto a supporto di una produzione tradizionale con budget elevati, il cortometraggio di Gatopoulos dimostra come la stessa tecnologia possa **abbattere drasticamente i costi di produzione** nel cinema indipendente e sperimentale.

figura 74 *Locandina di The Eggregores' Theory (Andrea Gatopoulos, 2024), presentato in apertura della Settimana Internazionale della Critica alla Mostra del Cinema di Venezia. Disponibile in streaming su MUBI.*

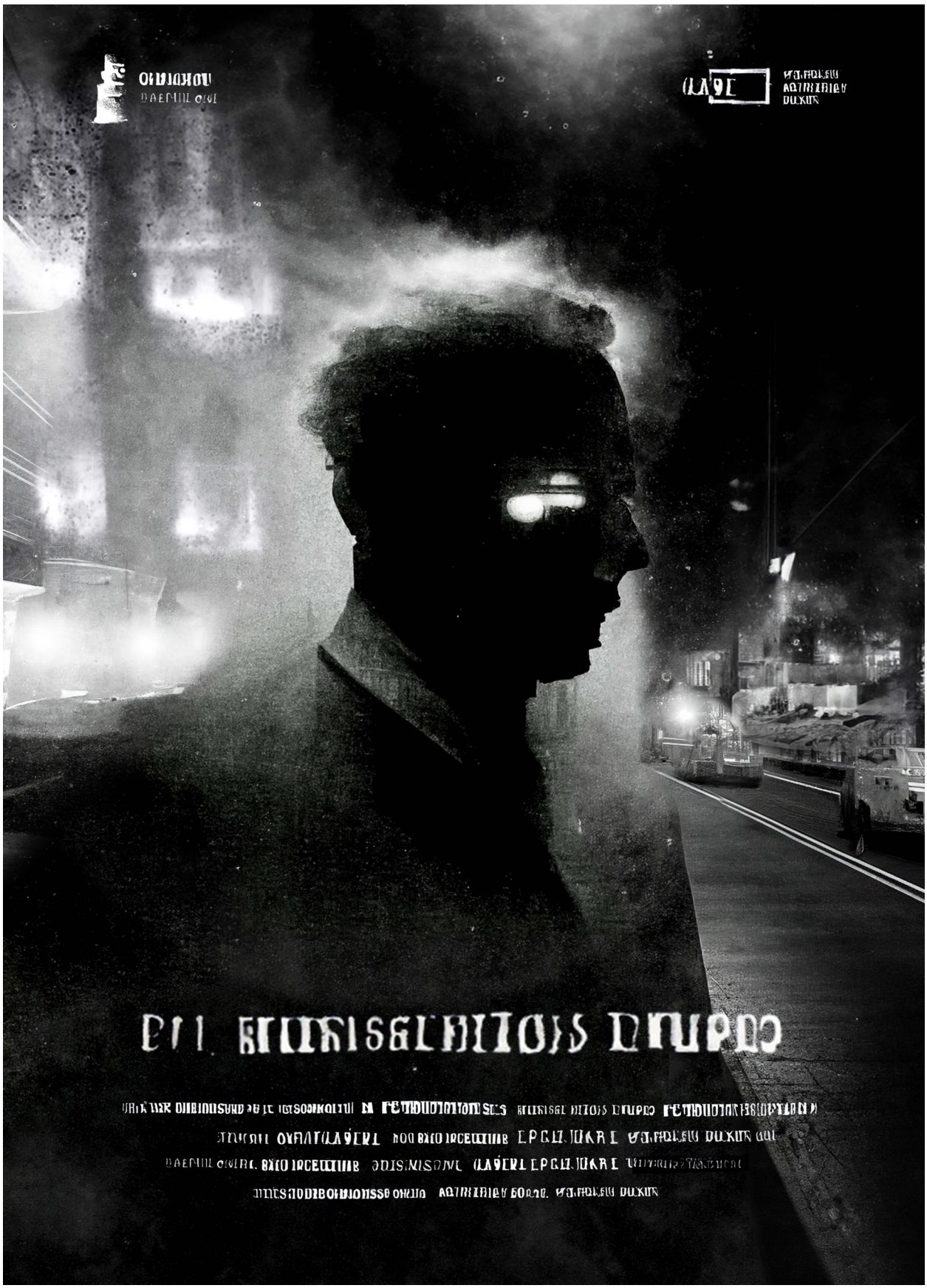




figura 75 *Sequenza di fotogrammi tratti da The Eggregores' Theory: una serie di immagini in bianco e nero che si susseguono nella storia grana sporca e imperfetta delle immagini generate dall'IA diventa un mezzo espressivo ed artistico, con un'estetica che si richiama a*



scandite da una voce narrante fuori campo che costruisce un immaginario distopico in cui la parola e l'arte sono state messe al bando. La "La Jetée" di Chris Marker (1962).

Un altro esempio è il documentario **Post Truth** (2025) diretto dall'artista turco Alkan Avcıoğlu, è il **primo documentario lungometraggio della storia interamente generato con intelligenza artificiale**, con immagini, suono, voce e musica compresi. il film esplora come tecnologia, social media, intelligenza artificiale e flussi informativi incessanti abbiano ridefinito la nostra comprensione di verità, realtà e fiducia e lo fa costruendo un ambiente narrativo completamente sintetico, costruito per rispecchiare l'ambiente digitale sintetico che critica, spingendo lo spettatore a interrogarsi non solo su cosa sta guardando ma su come interpreta il mondo che lo circonda. Usa lo stesso **linguaggio sintetico dell'epoca "post-verità"** per parlare di quell'epoca, in questo modo la forma diventa argomento. Il film finale nasce da oltre **60 ore di girato generato dall'IA** e più di 1.000 brani musicali composti dall'intelligenza artificiale ed è è diventato il primo film AI della storia a ottenere una **distribuzione cinematografica vera e propria**, uscendo in oltre 20 città turche a luglio 2025 tramite Başka Sinema. L'incasso resta comunque minuscolo (poco meno di 7.000 dollari complessivi), ma il precedente distributivo è quello che conta.

In questo caso infatti, il film non cerca di nascondere la propria natura sintetica per raggiungere il fotorealismo credibile; qui invece, **l'estetica sintetica è rivendicata apertamente** come contenuto stesso del discorso, che ti permette di superare la sola questione tecnica (flicker, morphing, uncanny motion) per affrontare la dimensione più propriamente filosofica del cinema algoritmico. Anche la **critica è stata sorprendentemente positiva** su alcune testate: viene descritto come un film denso e illuminante che richiede più visioni, capace di stratificare idee, immagini e argomentazioni continuando a rivelare significato. Si definisce come un **"a fake film about a real world".**

figura 76 *Locandina di Post Truth (Alkan Avcıoğlu, 2025), primo lungometraggio documentario della storia interamente generato dall'IA.*


"ONE OF THE BEST FILMS OF THE YEAR"
Warsaw Film Festival
OFFICIAL SELECTION
FDP FESTIVAL DEI POPOLI
2025
AUDIENCE AWARD
FRONTDOC
WINNER
BEST FEATURE
ART FILM
BEST AI FILM
FANTASPORTO
OFFICIAL SELECTION
BIFAN
POST
TRUTH
A / Fake / Film / About / The / Real / World
a SPONGEWORTHY production
IN ASSOCIATION WITH başka sinema IN HOMAGE TO godfrey reggio, adam curtis
WITH CINEMATIC REFERENCES TO alphaville (jean-luc godard), brazil (terry gilliam), fight club (david fincher), la haine (mathieu kassovitz), mauvais sang (leos carax), stalker (andrei tarkovsky), they live (john carpenter) WITH REFERENCES TO roland barthes, jean baudrillard, james bridle, jorge luis borges, daniel dennett, guy debord, jacques ellul, charles goodhart, martin heidegger, werner herzog, daniel kahneman, daniel levitin, marshall mcluhan, steven pinker, hans rosling, lee siegel, ultravox, max weber, orsen welles, theodore zeldin, slavoj žižek
SPECIAL THANKS TO marsel kalvo, nicholas kaiser, armagan lale, fatos avcıoğlu IMAGING OPERATIONS ASSISTANT emer can avcıoğlu MUSIC AND SOUND DESIGN alkan avcıoğlu EDITING BY alkan avcıoğlu WRITTEN BY alkan avcıoğlu & vikki bardot
INTERNATIONAL SALES BY odin's eye entertainment PRODUCED BY alkan avcıoğlu & vikki bardot A FILM BY alkan avcıoğlu
Spongeworthy
BAŞKA SİNEMA
ODIN'S EYE
posttruthmovie.com



figura 77 *Fotogramma tratto da Post Truth: una sterminata sala di lavoratori identici davanti a schermi in fila.*

Sempre riguardo alle immagine da estetica allucinatoria, possiamo citare un artista che ha fatto dell'AI un mezzo espressivo della sua arte, cioè **Jon Rafman**. Artista e filmaker canadese a sua intera pratica degli ultimi vent'anni ruota attorno all'**estetica del glitch e dell'allucinazione digitale**, e con l'arrivo dell'AI generativa ha semplicemente trovato lo strumento più affine alla sua sensibilità.
Tra i progetti che compongono la sua ricerca sull'allucinazione digitale si segnalano **Dream Journal 2016–2019**, lungometraggio in computer grafica presentato alla 58ª Biennale di Venezia e costruito come un flusso quasi ininterrotto di immagini oniriche disturbanti popolate da mutanti e archetipi, pensato non come diario dei sogni personali dell'artista ma come espressione dell'inconscio collettivo di internet; **Punctured Sky (2021)**, dove l'allucinazione algoritmica si applica alla ricostruzione nostalgica e frammentata di un videogioco andato perduto; e Counterfeit **Poast (2022)**, che spinge i modelli generativi verso volti e corpi deformi, glitch tipografici e anatomie impossibili, in continuità diretta con l'estetica del disgusto e della deformazione che attraversa tutta la sua produzione.

La sua opera più recente, **Main Stream Media Network** (2025-2026), merita un'analisi a parte perché sposta il discorso dall'allucinazione della singola immagine all'allucinazione di un intero ecosistema mediatico. Si tratta di un canale televisivo generato dall'AI, in continua evoluzione, costruito come una **"reimmaginazione post-AI slop" di MTV.** Rafman ha creato un intero **cast di popstar fittizie** (Cloudy Heart, Iron Tears, Wicked Hatchet, Flux Arcana, Pledged Torment), ciascuna dotata non solo di un'immagine generata, ma anche di un'intera discografia, di una storia personale e persino di una sottocultura di riferimento. Ad esempio, **Cloudy Heart** è una ragazza malinconica resa con un'estetica pixelata da videogioco anni '90, che pubblica aforismi esistenziali online, ha un furetto come consulente estetico e vapea compulsivamente. È un personaggio interamente sintetico, non è reale, ma ha accumulato un seguito reale su Spotify e Twitter. In questo senso, il progetto supera il medium tradizionale del supporto video: **l'allucinazione non resta confinata sullo schermo**, ma si sposta nel mondo reale, generando engagement, fandom e persino valore speculativo attorno a un'identità che non è mai esistita.
Sul piano installativo, il progetto è stato **esposto alla Sprüth Magers di Los Angeles (2025, Proof of Concept)** e in una grande retrospettiva museale alla Kunstsammlung **Nordrhein-Westfalen di Düsseldorf (2026)** trasforman-



figura 78 *Ritratto di Jon Rafman, artista canadese le cui opere si caratterizzano per un'estetica allucinatoria digitale.*



figura 79 *Sequenza di fotogrammi tratti da Main Stream Media Network: un flusso ininterrotto di immagini distorte e disturbanti. ogni persona,*



Volti deformati, iconografia finanziaria, corpi e dettagli anatomici estranianti che sembrano affiorare dal profondo del subconscio di

do la galleria in un ambiente sensoriale totale con monitor CRT vintage, proiezioni a parete e pavimenti drappeggiati: la fisicità dell'allestimento fa da contrappeso deliberato all'immaterialità digitale del contenuto.

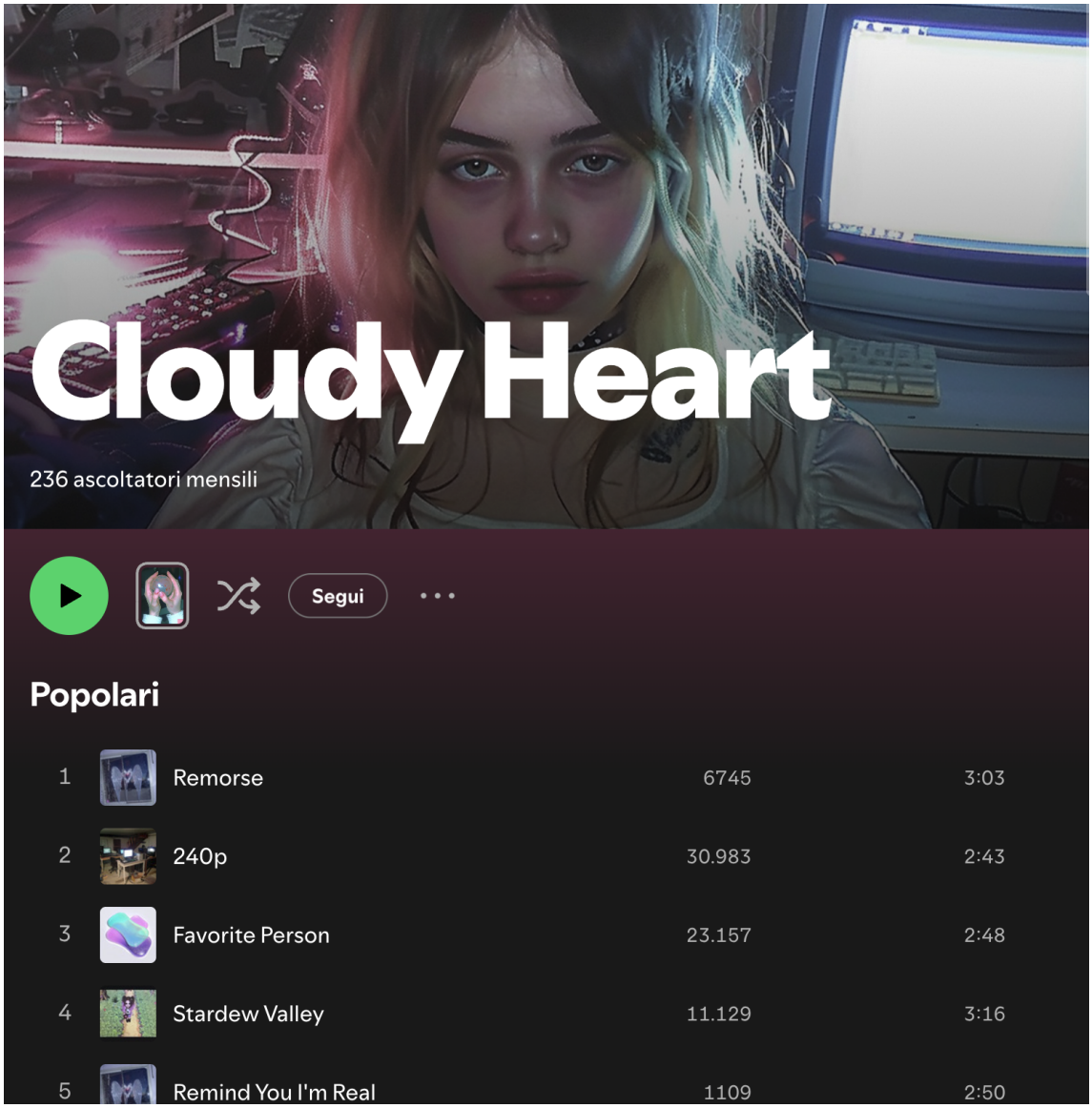


figura 80 *Screenshot del profilo Spotify di Cloudy Heart, la popstar interamente sintetica creata da Jon Rafman che ha creato un seguito di ascolti sulla piattaforma di streaming musicale.*





figura 81 e 82 *Vedute dell'allestimento della mostra Main Stream Media (K21, Kunstsammlung Nordrhein-Westfalen, Düsseldorf, 2026): monitor CRT vintage, proiezioni a parete e pavimenti drappeggiati trasformano lo spazio in un ambiente sensoriale immersivo.*

4.3 Open-source vs Closed-source

La distinzione tra modelli di intelligenza artificiale open-source e closed-source non riguarda soltanto una questione estetica o filosofica legata alla condivisione del sapere, ma ha **conseguenze dirette e concrete sul piano operativo:** determina chi può usare uno strumento, come, a quali costi e con quali limiti creativi.

I vantaggi dell'open source

Il tratto distintivo dell'open source è che il **codice sorgente è aperto**: chiunque può scaricarlo, ispezionarlo, modificarlo e installarlo sulla propria macchina. Questo significa, in pratica, che un utente (o uno sviluppatore) può costruire una piattaforma personalizzata basata su un modello di AI open source, senza dover dipendere dai server o dalle policy di un'azienda terza.

Un secondo vantaggio è la **personalizzazione**: i modelli open-weight possono essere modificati, riaddestrati (fine-tuning) o combinati con altri strumenti **(LoRA, ControlNet, custom nodes)** per adattarli a esigenze specifiche, cosa impossibile con un modello chiuso di cui non si conoscono i pesi interni.

Infine, questi modelli sono **spesso gratuiti**: una volta scaricati i pesi, l'utente può generare contenuti in locale senza dover pagare abbonamenti o crediti per ogni generazione, con l'unico costo reale rappresentato dall'hardware necessario per farli girare.

I limiti dell'open source

Il primo limite è la **scarsa accessibilità per chi non ha competenze informatiche**: installare un'interfaccia come ComfyUI, configurare i modelli, i text encoder e i VAE nelle cartelle corrette, risolvere errori di dipendenze sono tutte operazioni che richiedono un **minimo di alfabetizzazione**

tecnica, cosa che scoraggia l'utente medio rispetto a un tool closed-source "pronto all'uso".

Il secondo limite riguarda la **necessità di risorse computazionali enormi**, sia per l'addestramento sia, in misura minore, per l'inferenza. Stable Diffusion, ad esempio, è stato addestrato su ampi sottoinsiemi del dataset LAION-5B, che contiene circa 5,85 miliardi di coppie immagine-testo filtrate tramite CLIP. Il solo download completo del dataset richiede circa una settimana con 10 nodi paralleli, mentre l'addestramento vero e proprio di Stable Diffusion ha richiesto centinaia di migliaia di step su cluster con decine o centinaia di GPU A100 per giorni consecutivi di calcolo ininterrotto. Anche una replica "economica" di Stable Diffusion 2, realizzata da MosaicML/Databricks, ha richiesto circa 23.800 ore-GPU A100 e 7,45 giorni di training continuo, per un costo stimato di circa 47.700 dollari, pur trattandosi di una versione ottimizzata e a basso costo rispetto all'originale.
Questo dimostra come, nonostante l'accessibilità del codice, **l'addestramento di un modello open source da zero** resti un'operazione **fuori dalla portata della maggior parte degli utenti individuali,** riservata a laboratori, aziende o community con accesso a infrastrutture di calcolo importanti.

Il terzo limite è la **natura dispersiva e frammentata dell'ecosistema open source**: ogni pochi mesi escono nuovi modelli, nuove versioni, nuovi formati di file, nuovi nodi custom, rendendo difficile per un utente non esperto tenere il passo con l'evoluzione tecnica e scegliere lo strumento più adatto al proprio progetto.

I vantaggi del closed source

Il principale punto di forza dei modelli closed-source è l'**accessibilità operativa**: strumenti come ChatGPT, Midjourney o Gemini sono pensati per essere usati da chiunque tramite un'**interfaccia web semplice**, senza necessità di installare nulla in locale o di possedere hardware dedicato. Questo si traduce direttamente in una **maggiore facilità d'uso**: l'utente scrive un prompt e ottiene un risultato in pochi secondi, senza dover gestire nodi, checkpoint, VAE o configurazioni tecniche, a differenza di quanto

richiesto dalle piattaforme open source come ComfyUI.

I limiti del closed source

Il limite più significativo dei modelli closed-source è la **censura**. Le grandi aziende tecnologiche, per ragioni legali, reputazionali e di responsabilità civile, **applicano filtri stringenti** che impediscono la generazione di contenuti considerati sensibili: immagini di nudo, riproduzione fedele di persone reali, contenuti violenti o politicamente delicati. Questo approccio, sebbene comprensibile dal punto di vista del rischio d'impresa, **limita fortemente la libertà creativa e sperimentale dell'artista** o del ricercatore che lavora con l'AI, ponendo un confine netto tra ciò che è "permesso" e ciò che tecnicamente sarebbe possibile generare.

Un secondo limite è rappresentato dal codice proprietario: **l'utente non ha accesso al funzionamento interno del modello**, non può ispezionarne l'architettura né comprenderne appieno i meccanismi decisionali, il che pone anche interrogativi di trasparenza e responsabilità (accountability) nell'uso di questi strumenti.

Infine, i **modelli e i dataset restano chiusi**: non è possibile sapere con certezza su quali immagini, testi o video un modello closed-source sia stato addestrato, un'informazione che invece per i modelli open come Stable Diffusion è (parzialmente) documentata attraverso dataset pubblici come LAION-5B.

Questa opacità alimenta un più ampio dibattito, che ha assunto anche una dimensione politica: nel 2026 l'amministrazione statunitense ha introdotto un framework di revisione volontaria pre-rilascio rivolto esclusivamente ai modelli "closed" più potenti (come Claude o ChatGPT), escludendo di fatto i modelli open-weight, inclusi quelli cinesi, oggi tra i più diffusi e economici sul mercato, da qualunque forma di controllo preventivo, evidenziando come la contrapposizione tra apertura e chiusura non sia solo tecnica ma anche geopolitica e regolatoria

L'errore procedurale nel film: un'applicazione pratica

Un caso concreto di questa tensione tra apertura tecnica e limiti creativi, non troppo distante da quanto osservato "The Pornographer" discusso in precedenza, emerge da un **workflow costruito in ComfyUI.** Ogni nodo, all'interno della piattaforma, svolge un compito specifico, e l'insieme dei nodi collegati forma il workflow completo che produce il contenuto finale. Durante il lavoro si è verificato un errore procedurale che ha prodotto un risultato visivamente "troppo aderente" (patinato): la generazione non partiva da un prompt testuale, ma da **image prompting**, tecnica in cui reti neurali dedicate classificano semanticamente il contenuto visivo di un'immagine di riferimento per poi usarla come input visivo, anziché testuale, di un secondo modello generativo. In questo caso l'immagine è stata usata come **prompt visivo per Vgen**, modello open source sviluppato da Alibaba per la generazione video, a dimostrazione di come anche un **errore procedurale in un workflow open source** possa rivelare aspetti tecnici interessanti sul funzionamento di questi sistemi.

4.4 Comfy-UI per creare contenuti generativi

ComfyUI è un'**interfaccia web (Web UI)** che gira interamente in locale sul computer dell'utente. Si tratta di un progetto **open source**, realizzato in Python, che utilizza un'**architettura visiva basata su nodi e collegamenti**: ogni operazione di generazione, caricamento del modello, encoding del prompt, campionamento, decodifica dell'immagine, è rappresentata da un nodo, e i nodi vengono collegati tra loro per formare un workflow completo. Ogni volta che si esegue una generazione, di fatto si esegue del codice Python organizzato secondo la logica del grafo costruito dall'utente.

All'interno di ComfyUI è possibile **caricare e utilizzare modelli open-weight**, ovvero modelli i cui pesi (weights) sono scaricabili e installabili pubblicamente in locale. Tra i più diffusi vi sono **Stable Diffusion, FLUX, Qwen e WAN. Ogni modello** è generalmente **specializzato in un compito specifico**: alcuni sono ottimizzati per la generazione di immagini, altri per quella di video, altri ancora per compiti più mirati come l'upscaling o l'inpainting. In origine, i modelli venivano distribuiti come checkpoint unici: un singolo file che conteneva tutte le componenti necessarie alla generazione. Con l'evoluzione dell'ecosistema, oggi si tende a distinguere tre componenti fondamentali che, insieme, costituiscono un modello generativo completo:

- **Modello (UNet/Diffusion Transformer)**: contiene le **informazioni apprese durante l'addestramento** su immagini e dati; è il "motore" che genera il contenuto
- **Text Encoder (CLIP)**: **Interpreta e traduce il prompt testuale** in una rappresentazione comprensibile al modello; varia da modello a modello
- **VAE (Variational Autoencoder)**: **codifica e decodifica l'immagine**, traducendo tra lo spazio latente (dove lavora il modello) e lo spazio dei pixel visibile all'utente

Il formato di file oggi più diffuso per distribuire questi pesi è **.safetensors**, considerato **più sicuro rispetto ai formati precedenti** (come .ckpt, basato su pickle di Python) perché non consente l'esecuzione di codice arbitrario al momento del caricamento, riducendo il rischio di file malevoli. Un parametro tecnico rilevante è la **precisione numerica (FP, Floating Point)**, che indica con quanti bit vengono memorizzati i pesi del modello: **FP32** rappresenta la precisione piena (32 bit per parametro), mentre formati più compressi come **FP16 o FP8** dimezzano o riducono ulteriormente lo spazio occupato, a scapito di una minima perdita di precisione, permettendo però di far girare modelli più pesanti su hardware con meno memoria video disponibile.

Il **formato GGUF** rappresenta un'ulteriore evoluzione pensata per l'efficienza: **impacchetta in un unico file sia i pesi del modello sia tutti i metadati necessari al suo funzionamento**, così da poterlo caricare ed eseguire facilmente anche su hardware consumer o addirittura su CPU, rendendolo particolarmente utile per chi non dispone di GPU



figura 83 *Logo di ComfyUI, interfaccia a nodi open-source per costruire e personalizzare workflow con modelli generativi IA.*

di fascia alta.

Come funziona il processo di generazione

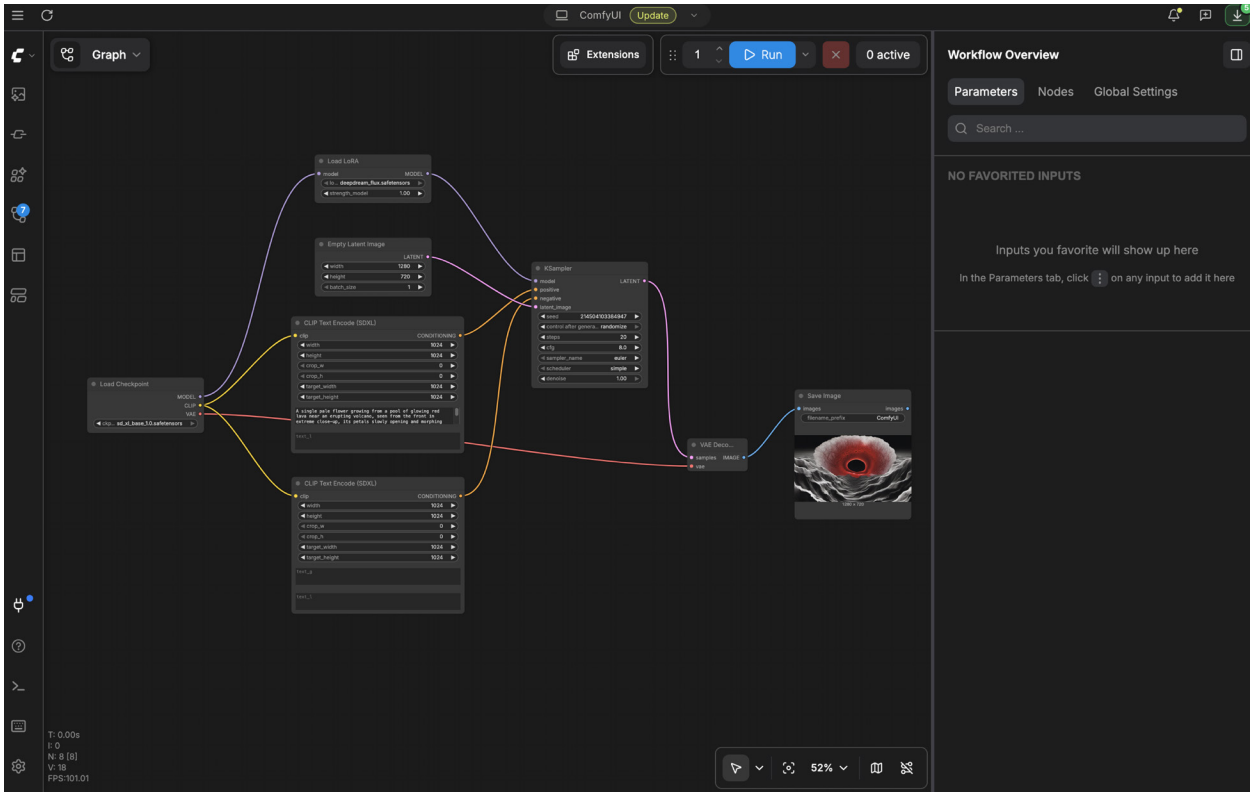
Il processo di generazione in ComfyUI (e più in generale nei modelli di diffusione) parte da un "quadro vuoto", tecnicamente chiamato **latent**: uno spazio compresso e non direttamente visibile che rappresenta l'immagine (o il video) in costruzione. Il modello, usando le informazioni apprese durante l'addestramento, **genera inizialmente del rumore casuale**, dello stesso tipo di rumore su cui è stato addestrato a lavorare e attraverso un **processo iterativo di denoising** (rimozione progressiva del rumore), compone gradualmente l'immagine finale. Il risultato di questo processo, tuttavia, non può essere visualizzato direttamente: si tratta ancora di dati nello spazio latente, non di pixel. È qui che entra in gioco il **VAE**, il cui compito è tradurre questa rappresentazione latente in un'immagine visibile e salvabile.

Questi **tre elementi (modello, text encoder e VAE)** sono dunquen **essenziali in (quasi) tutti i workflow di generazione,** ma non sono intercambiabili tra modelli diversi: un text encoder progettato per un modello specifico potrebbe non funzionare correttamente (o affatto) con un modello diverso, motivo per cui è fondamentale rispettare le combinazioni indicate dagli sviluppatori di ciascun modello.

figura 84 ***Esempio di workflow eseguito in ComfyUI****. Ho utilizzato il checkpoint Stable Diffusion XL Base 1.0 con la LoRA deepdream_flux (strength 1.00), scegliendo deliberatamente un modello non recente per ottenere immagini allucinate senza doverle richiamare esplicitamente nel prompt, lasciando che fosse il modello stesso a generarle spontaneamente.*

Il prompt utilizzato è stato: "A single pale flower growing from a pool of glowing red lava near an erupting volcano, seen from the front in extreme close-up, its petals slowly opening and morphing into distorted organic shapes, a blinding white light shining from the center of the flower as if a hole had opened inside it, long thin black tentacles emerging from the glowing opening and stretching outward in different directions, black volcanic rock, drifting ash, smoke and heat distortion, unsettling horror atmosphere, high contrast lighting, abstract dark background, deep red lava glow, black shadows, 16:9 composition".

L'obiettivo era ottenere un fiore che emergesse dalla lava con una tinta horror, ma il risultato è stato più astratto del previsto. Ho ottenuto un'immagine che si è rivelata artisticamente più interessante di quanto sarebbe stato un output fotorealistico. La generazione è stata eseguita in locale su MacBook Pro M1 Pro (16 GB di RAM) in circa 3 minuti, producendo un'immagine PNG a risoluzione 1280×720.


ComfyUI
Update
Graph
Extensions
1
Run
0 active
Workflow Overview
Parameters
Nodes
Global Settings
Search ...
NO FAVORITED INPUTS
Inputs you favorite will show up here
In the Parameters tab, click on any input to add it here
Load LoRA
Empty Latent Image
KSampler
CLIP Text Encode (SDXL)
Load Checkpoint
CLIP Text Encode (SDXL)
Save Image
T: 0.00s
I: 0
N: 8 [8]
V: 18
FPS:101.01
52%

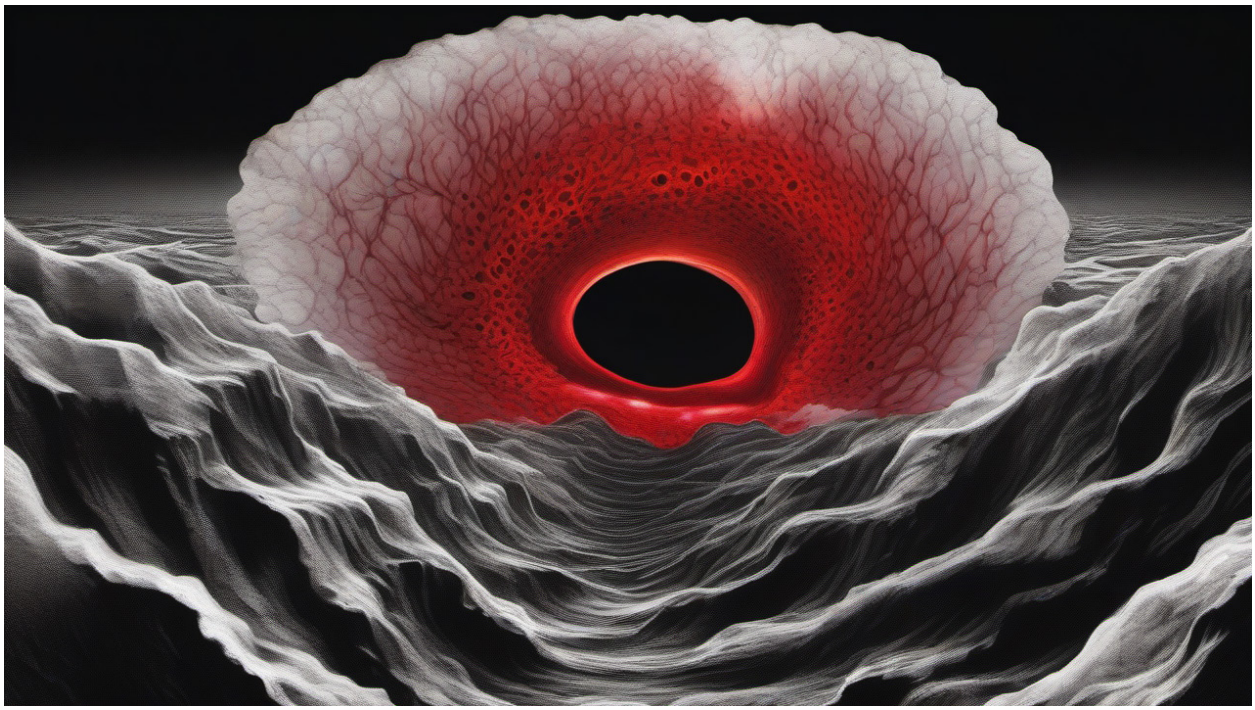
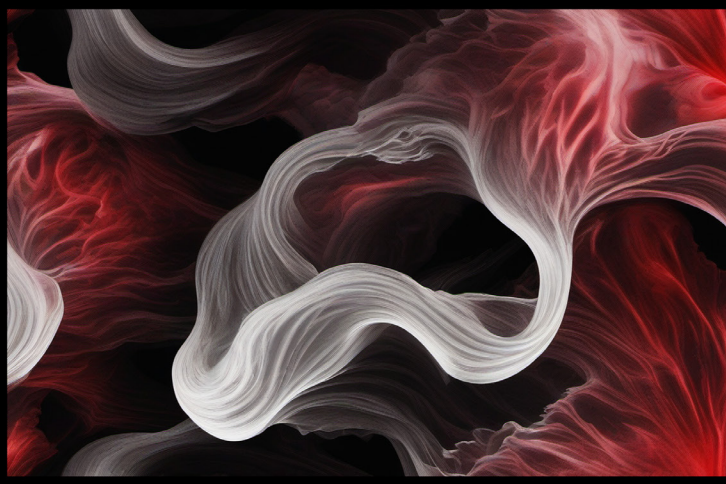
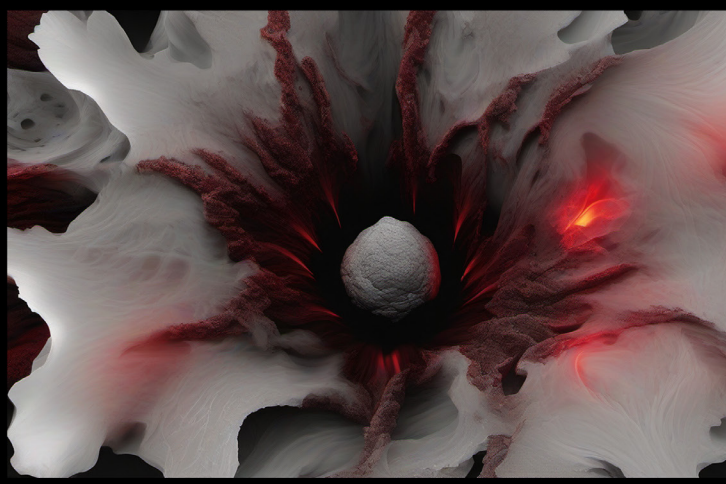




figura 85 *Sequenza di immagini ottenute dallo stesso prompt descritto in figura 84.*







L'importanza di Hugging face e Civit-AI

Nell'ecosistema open source della generazione AI, due piattaforme rivestono un ruolo centrale nella distribuzione e condivisione delle risorse necessarie a costruire un workflow in ComfyUI:

- **Hugging Face** funge da **repository principale per modelli, dataset e documentazione tecnica**, offrendo accesso diretto ai pesi ufficiali di modelli come Stable Diffusion, FLUX e ai relativi text encoder e VAE, spesso accompagnati da schede tecniche (model card) che ne documentano l'addestramento e l'uso corretto.

- **Civitai** è invece una piattaforma community- driven, particolarmente diffusa tra gli utenti di ComfyUI e Stable Diffusion, dedicata alla **condivisione di checkpoint personalizzati, modelli fine-tuned, LoRA, workflow già pronti** (in formato .json, come quelli utilizzati nel progetto discusso in questa tesi) e nodi custom sviluppati dalla community.

Entrambe le **piattaforme sono quindi fondamentali** per chiunque voglia **lavorare con l'AI generativa open source**: permettono di reperire non solo i modelli di base, ma anche i workflow, i plugin e i nodi custom necessari a costruire pipeline di generazione complesse e personalizzate, riducendo la necessità di addestrare modelli da zero, operazione che come discusso nel paragrafo 4.3, richiede risorse computazionali fuori dalla portata della maggior parte degli utenti individuali.



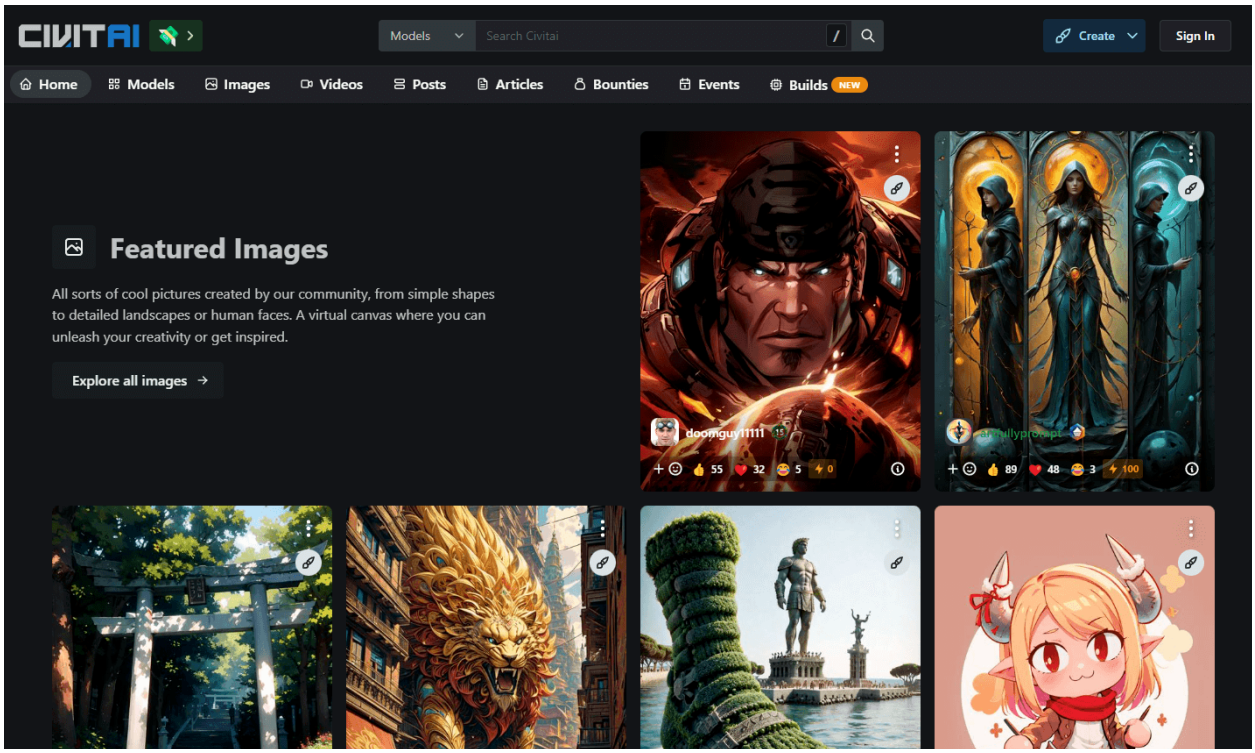


figura 86 E 87 *Le due principali piattaforme per la generazione locale e per l'ecosistema AI open source: Hugging Face (sopra), repository di riferimento per modelli, text encoder e documentazione tecnica, e Civitai (sotto), piattaforma community-driven per checkpoint, LoRA e workflow condivisi.*



Capitolo 5

Conclusioni

Trarre delle conclusioni da un argomento così complesso non è semplice, perché il giudizio cambia inevitabilmente in base al punto di vista da cui si guarda la questione.

Da un lato sono estremamente **affascinato dallo sviluppo di questa tecnologia**. Da un punto di vista tecnico sembra quasi magia, ma dietro c'è matematica e informatica, modelli di deep learning basati sull'enorme quantità di dati in cui siamo immersi oggi. Con l'IA si può **snellire il lavoro**, aiutando piccoli videomaker o artisti con poco budget a realizzare e concretizzare le proprie idee, **rendendo l'audiovisivo e il cinema un'arte più democratica**, non più nelle mani solo di chi ha potere economico. E non si parla solo di generare contenuti da zero: come visto nei casi analizzati (il de-aging in tempo reale di Here, l'uso di Respeecher in The Brutalist) l'IA serve anche a **sistemare, correggere, migliorare** ciò che già esiste. Senza contare che l'IA generativa in campo artistico è solo una parte del discorso: ha applicazioni enormi anche in ambito scientifico e nella medicina, dove può contribuire a migliorare efficacemente la vita delle persone.

Dall'altro lato, però, **non è possibile ignorare gli aspetti problematici** che accompagnano questa trasformazione. Nel corso degli anni probabilmente spariranno una serie di lavori, soprattutto nelle professioni più ripetitive, standardizzabili o facilmente realizzabili in ambiente digitale e dell'industria dell'audiovisivo. C'è il problema del costo dell'IA e degli abbonamenti alle piattaforme mediante sistemi di token, quello della costruzione dei data center e del loro consumo di acqua e risorse, un costo fisico e concreto dietro l'apparente immaterialità di un prompt.

Non c'è una risposta giusta in merito all'argomento: ognuno può prendere una propria posizione, chi è completamente favorevole e chi la esclude a priori. Credo che, come in tutte le cose, **la verità stia nel mezzo**: cogliere le opportunità che l'IA offre senza lasciarla "andare" senza freni, con regolamenti come sta cercando di fare l'Europa con l'AI Act ,ma soprattutto costruendo una coscienza digitale collettiva, che si interroghi su quello che vede online

e su come funzionano davvero queste intelligenze artificiali generative. Soprattutto, dobbiamo **ricordarci che dovrebbero essere le IA a venire in nostro soccorso e non il contrario**: dobbiamo sapere riconoscere quando serve effettivamente usare uno strumento del genere, perché migliora concretamente la vita, piuttosto che essere una "bolla speculativa" promossa da grandi colossi tech e multinazionali. L'evoluzione dei modelli generativi è impressionante, ogni mese escono nuovi strumenti e restare al passo costantemente non è un'impresa facile.

Tuttavia non credo che l'IA porterà alla morte dell'arte. È vero, probabilmente molte multinazionali e case di produzione sostituiranno gran parte dei lavoratori dell'industria, sfruttando questa tecnologia per arricchirsi piuttosto che per il bene di chi lavora e della manovalanza del settore. Ma **l'idea di una storia, la scelta di cosa raccontare, il montaggio dei diversi frammenti** in un'opera coerente restano **attività profondamente umane**, legate ad un vissuto personale e individuale che una macchina non può replicare. L'IA può creare qualcosa che assomiglia ad altro, ma non qualcosa di davvero nuovo: quello resta l'ingegno umano che, per quanto funzioni "come" una rete di neuroni artificiali, porta con sé un mistero che nessun modello riesce a spiegare. L'arte ha di per sé una **natura quasi mistica**, legata al subconscio e alla parte irrazionale dell'umano che la scienza non può spiegare del tutto, e che forse non ha nemmeno senso provarci, perché è proprio nel mistero che l'arte trova i suoi significati più veri.

In conclusione, credo sia fondamentale informarsi e tenersi aggiornati su questa tecnologia, indipendentemente dal fatto che la si approvi o no, e imparare a **vedere questi strumenti come possibilità**, non come minacce, usandoli però in modo consapevole: prima di scrivere un prompt, chiedersi sempre se sia davvero necessario per quello che si vuole ottenere, per restare noi a controllare lo strumento, e non lasciarci travolgere da esso.



Sitografia
Bibliografia
Iconografia

Sitografia

CAP 1 – Introduzione teorica

1) https://www.ibm.com/it-it/think/topics/history-of-artificial-intelligence
2) https://www.youtube.com/watch?v=mSd9nmPM7Vg
3) https://www.youtube.com/watch?v=RDMwSAZwn8I
4) https://www.rivista.ai/2024/09/03/dartmouth-summer-research-project-on-artificial-intelligence/
5) https://www.intelligenzaartificialeitalia.net/post/gli-inverni-dell-intelligenza-artificiale-ia
6) https://www.wired.it/article/intelligenza-artificiale-storia-chatbot-chatgpt-turing/
7) https://www.geopop.it/eliza-chatbot-primo-della-storia-invenzione-effetto-dati/
8) https://www.treccani.it/enciclopedia/intelligenza-artificiale_res-1eadb35b-dd79-11e6-add6-00271042e8d9_(Enciclopedia-Italiana)/
9) https://www.ibm.com/it-it/think/topics/artificial-intelligence-types
10) https://www.ibm.com/it-it/think/topics/machine-learning-types
11) https://cloud.google.com/learn/what-is-artificial-intelligence?hl=it
12) https://www.youtube.com/watch?v=MWPpjN9eWz8
13) https://www.youtube.com/watch?v=N2sVUNsqrKo
14) https://aws.amazon.com/it/what-is/artificial-general-intelligence/
15) https://www.consilium.europa.eu/it/policies/artificial-intelligence/#0
16) https://www.ibm.com/it-it/think/topics/deep-learning

CAP 2 – Intelligenza artificiale generativa

17) https://www.youtube.com/watch?v=iv-5mZ_9CPY&t=615s
https://www.youtube.com/watch?v=qYNweeDHiyU
https://www.youtube.com/watch?v=2IK3DFHRFfw&t=625s
18) https://www.ibm.com/it-it/think/topics/generative-ai#257779831
19) https://www.ibm.com/it-it/think/topics/generative-model#257779832
20) https://www.ibm.com/it-it/think/topics/variational-autoencoder
21) https://it.wikipedia.org/wiki/Rete_generativa_avversaria
22) https://www.ibm.com/it-it/think/topics/diffusion-models
23) https://www.agendadigitale.eu/cultura-digitale/creare-immagini-dallimmaginazione-il-potere-dei-modelli-di-diffusione/
24) https://www.youtube.com/watch?v=iv-5mZ_9CPY&t=615s
25) https://www.youtube.com/watch?v=1CIpzeNxIhU
26) https://www.youtube.com/watch?v=12Qm1KfgRic&t=314s

CAP 3 – Etica ed impatti delle immagini algoritmiche

27) https://www.scienzaefilosofia.com/?portfolio=norbert-wiener-introduzione-alla-cibernetica-luso-umano-degli-esseri-umani-tr-it-a-cura-di-d-persiani-bollati-boringhieri-torino-1970-pp-240-e-16

28) https://www.ibm.com/it-it/think/topics/ai-bias
29) https://www.youtube.com/watch?v=6t_u636_JEU
30) https://www.youtube.com/watch?v=ZJWEqoOIDKo
31) https://www.youtube.com/watch?v=-kBuXkMVHDM
32) https://www.aequitasmagazine.it/ark:/56778/6520
33) https://en.wikipedia.org/wiki/Ghost_work
34) https://laion.ai/blog/laion-5b/
35) https://en.wikipedia.org/wiki/LAION
36) https://purl.stanford.edu/kh752sm9123
37) https://www.theregister.com/software/2023/12/20/csam-found-in-large-ai-image-generator-training-dataset/1096580
35) https://fedscoop.com/ai-federal-research-database-laion-csam/
36) https://www.ilriformista.it/pedopornografia-nei-dataset-dellintelligenza-artificiale-lindagine-scandalo-dello-stanford-internet-observatory-400304/#google_vignette
37) https://blog.unlaw.it/2024/11/20/la-prima-decisione-europea-su-utilizzo-opere-protette-dal-diritto-dautore-per-addestramento-di-ai/
38) https://www.wired.it/article/chatgpt-studio-ghibli-copyright/
39) https://www.sentieriselvaggi.it/le-immagini-ispirate-allo-studio-ghibli-di-chatgpt-innocuo-gioco-o-violazione-dautore/
40) https://en.wikipedia.org/wiki/2023_Writers_Guild_of_America_strike
41) https://en.wikipedia.org/wiki/2023_SAG-AFTRA_strike
42) https://www.theregister.com/software/2023/09/27/hollywood-studios-agree-to-regulate-use-
of-generative-ai/1067518
43) https://www.brookings.edu/articles/hollywood-writers-went-on-strike-to-protect-their-livelihoods-from-generative-ai-their-remarkable-victory-matters-for-all-workers/
44) https://www.dw.com/en/hollywood-strike-shows-films-youll-have-to-wait-for/a-66227440
45) https://www.youtube.com/watch?v=AfUKFhowdNc
46) https://www.realitydefender.com/insights/history-of-deepfakes
47) https://thispersondoesnotexist.com/
48) https://www.agi.it/blog-italia/riccardo-luna/post/2019-06-28/app_che_spoglia_le_donne_deepnude-5737680/
49) https://www.npr.org/2023/04/21/1171032649/ai-music-heart-on-my-sleeve-drake-the-weeknd
50) https://www.samsungsds.com/en/insights/what-are-cheapfakes.html
51) https://www.youtube.com/watch?v=sDOo5nDJwgA
52) https://www.youtube.com/watch?v=mS_8jO9vkwk
53) https://www.puntosicuro.it/nuove-tecnologie-ia-C-148/deepfake-come-identificare-fotografie-filmati-contraffatti-AR-26379/
54) https://www.wired.it/article/ai-slop-definizione-sbobba-infinita-bassa-qualita/
55) https://it.wikipedia.org/wiki/Slop_(intelligenza_artificiale)
56) https://www.instagram.com/jeremyfindsai/?hl=it

57) https://www.ibm.com/it-it/think/topics/data-centers
58) https://www.geopop.it/come-sono-fatti-e-a-cosa-servono-i-data-center-i-luoghi-fisici-che-contengono-tutti-i-dati-online/
59) https://www.socomec.it/it/soluzioni/business/data-center/comprendere-il-consumo-di-potenza-dei-data-center
60) https://www.geopop.it/intelligenza-artificiale-consumi-energia/
61) https://www.iea.org/reports/energy-and-ai
62) https://it.wikipedia.org/wiki/Rack_(informatica)
63) https://www.nature.com/articles/s41545-021-00101-w
64) https://unric.org/it/giornata-mondiale-dellacqua-la-scarsita-dacqua-a-livello-globale-incombe-ecco-cosa-si-puo-fare/
65) https://www.bbc.com/news/articles/cy8gy7lv448o
66) https://www.wri.org/insights/us-data-center-growth-impacts
67) https://pmc.ncbi.nlm.nih.gov/articles/PMC12273412/
68) https://www.reddit.com/r/Wellthatsucks/comments/1vajkrg/living_next_to_a_data_center/
69) https://www.youtube.com/shorts/xncgkS1_LGI
70) https://www.youtube.com/watch?v=pPtB3g5BSVk
71) https://www.eesi.org/articles/view/data-centers-and-water-consumption
72) https://www.eticasgr.com/storie/approfondimenti/consumo-d-acqua-impatto-data-center
73) https://intelpoint.co/insights/the-united-states-of-america-hosts-over-3700-data-centres-8x-more-than-any-other-country-as-of-may-2025/
74) https://www.ilsole24ore.com/art/usa-antitrust-urgente-controllo-come-big-tech-gestisce-l-ai-AGs6ztQ?refresh_ce=1
75) https://www.wired.it/article/kate-crawford-wired-next-fest-2022-milano-intelligenza-artificiale/

CAP 4 – Autorialità ed estetica delle immagini algorimtiche

76) https://it.wikipedia.org/wiki/Cahiers_du_cin%C3%A9ma
77) https://it.wikipedia.org/wiki/Politica_degli_autori
78) https://www.youtube.com/watch?v=1TIY7jkyqz4
79) https://www.youtube.com/watch?v=-Nb-M1GAOX8
80) https://nofilmschool.com/eddie-ai-9b
81) https://kottke.org/23/06/thank-you-for-not-answering
82) https://www.artribune.com/attualita/2014/04/la-scoperta-del-cinema-espanso/
83) https://ilmanifesto.it/youngblood-cera-una-volta-il-cinema-espanso
84) chrome-extension://efaidnbmnnnibpcajpcglclefindmkaj/https://www.einaudi.it/content/uploads/estratti/9788806207 32PCA.pdf?x15441
85) https://www.youtube.com/watch?v=KN9NDpikYeM
86) https://www.youtube.com/watch?v=VR6rOTCbCNU
87) https://www.youtube.com/watch?v=PeXRFH2I9Is
88) https://getstream.io/glossary/temporal-consistency/

89) https://auralumeai.com/posts/what-is-temporal-consistency-in-ai-video-generation-a-guide-to-stable-cinematic
90) https://it.wikipedia.org/wiki/Morphing
91) https://www.youtube.com/watch?v=wZurRt0TidI
92) https://www.fastweb.it/fastweb-plus/digital-magazine/morphing-online-i-migliori-siti-per-il-fotoritocco-del-viso/
93) https://it.wikipedia.org/wiki/Uncanny_valley
94) https://spectrum.ieee.org/the-uncanny-valley
95) https://paths.grasp.study/courses/df16ce78-e39c-455d-bfeb-fe7cc3f3fbc0/modules/1e2b2e48-ba86-452a-ad0a-9f1aea9b9408/lessons/0458afc3-4c8f-4eb2-9337-64818691d3d2
96) https://screenrant.com/worst-uncanny-valley-examples-movies-ranked/
97) https://cordis.europa.eu/article/id/444372-what-makes-a-face-fall-into-the-uncanny-valley/it
98) https://www.youtube.com/watch?v=sJD4MaEHUJo&t=67s
99) https://www.hollywoodreporter.com/business/business-news/divergent-producer-pouya-shahbazian-1236153574/
100) https://en.wikipedia.org/wiki/Dreams_of_Violets
101) https://www.theguardian.com/film/2026/jun/03/dreams-of-violets-ash-koosha-iran-tribeca-film-festival
102) https://variety.com/2026/film/festivals/tribeca-festival-ai-film-dreams-of-violets-fountain-0-1236759724/
103) https://www.youtube.com/watch?v=KN9NDpikYeM
104) https://en.wikipedia.org/wiki/Hell_Grind
105) https://aiwiki.ai/wiki/hell_grind
106) https://entertainment.blab.com/2026-06-10-san-francisco-startup-higgsfield-ai-debuts-hell-grind-a-95-minute-ai-generated-action-film-at-cannes-marketplace
107) https://www.youtube.com/watch?v=t33k2tn4GpA&t=220s
108) https://tg24.sky.it/spettacolo/cinema/2026/05/28/hell-grind-film-intelligenza-artificiale
109) https://www.youtube.com/watch?v=3GAsTIClkkQ&t=186s
110) https://runway.com/news/company-news/project-luxo
111) https://www.wired.it/article/the-brutalist-intelligenza-artificiale-respeecher/
112) https://www.vanityfair.com/hollywood/story/the-brutalists-ai-controversy-explained?srsltid=AfmBOooz9kH9jyf-uFa1UQODC4OWjzMfLzw3maAlifAgTpOtigpnINxq
113) https://harborpicturecompany.com/in-the-press/p/ai-and-the-making-of-here/
114) https://www.reddit.com/r/TheBigPicture/comments/1i53ad7/the_brutalist_used_ai/
115) https://variety.com/2024/film/features/robert-zemeckis-here-tom-hanks-aging-tech-1236196938/
116) https://www.indiewire.com/features/general/robert-zemeckis-latest-tech-move-de-aging-tom-hanks-with-ai-1234805275/
117) https://www.reddit.com/r/movies/comments/1j19zcf/

regarding_the_use_of_ai_to_deage_actors_for/
118) https://hariel.xyz/it/
119) https://www.ibm.com/it-it/think/topics/ai-hallucinations
120) https://it.wikipedia.org/wiki/Allucinazione_(intelligenza_artificiale)
121) https://www.tabatatech.it/cose-il-fine-tuning-nellintelligenza-artificiale/
122) https://www.youtube.com/watch?v=v89kB4OScOA
123) https://www.sitopreferito.it/appuntamenti-ed-eventi/the-pornographer-il-corto-realizzato-in-ai-dal-collettivo-milanese-hariel-vince-il-premio-centro-nazionale-del-cortometraggio-alla-40-sic/
124) https://www.fatamorganaweb.it/the-pornographer-hariel/
125) https://www.sentieriselvaggi.it/the-pornographer-di-hariel-allai-motion-festival/
126) https://sicvenezia.it/en/films/the-eggregores-theory/
127) https://mubi.com/it/it/films/the-eggregores-theory
128) https://www.youtube.com/watch?v=hvKgCCynmXA
129) https://mubi.com/it/it/films/post-truth-2025
130) https://www.frontdoc.it/film/post-truth/
131) https://en.wikipedia.org/wiki/Jon_Rafman
132) https://www.agomodena.it/fondazione/storia/fondazione-modena-arti-visive/jon-rafman-il-viaggiatore-mentale/
133) https://jonrafman.com/jon-rafman-main-stream-media-network/
134) https://www.instagram.com/p/DYXVStelUsp/
135) https://www.reddit.com/r/aiwars/comments/1rhdzj1/opinion_on_opensource_vs_closedsource_ai_models/
136) https://www.youtube.com/watch?v=_QfxGZGITGw
137) https://www.pbs.org/newshour/science/whats-the-difference-between-closed-open%E2%80%91source-and-open-weight-ai-a-researcher-explains
138) https://comfy.org/
139) https://www.youtube.com/watch?v=azJX-qVdmag&t=3730s
140) https://www.youtube.com/watch?v=g74Cq9Ip2ik&t=2017s
141) https://huggingface.co/
142) https://www.youtube.com/watch?v=3kRB2TXewus&t=101s
143) https://civitai.com/
144) https://www.youtube.com/watch?v=CZei0p_xBYY&list=PL6-GCLeKuEAThdWYtVvHBvpVE_dmZjMHE
145) https://www.youtube.com/watch?v=hl7kwMwkLJU

Bibliografia

1) Melanie Mitchell, L'intelligenza artificiale: una guida per esseri umani pensanti, 2022
2) Valliappa Lakshmanan e Hannes Hapke, Generative Ai design patterns: solutions to common
challenges when building genai agents and application, 2025
3) Capitolo XI "L'uso umano degli esseri umani. Introduzione alla cibernetica", 1966 Boringhier
4) Ėjzenštejn. Il montaggio come metodo compositivo

Iconografia

CAP 1 – Introduzione teorica
1) Il test di touring; https://www.nationalgeographic.it/la-nuova-ia-potrebbe-superare-il-test-di-turing-chi-era-l-uomo-che-lo-ha-inventato
2) La conferenza di Dartmouth; https://spectrum.ieee.org/dartmouth-ai-workshop
3) Garry Kasparov durante una partita contro Deep Blue; https://www.sentieriselvaggi.it/kasparov-vs-deep-blue-sara-un-film/
4)Alpha Go sconfigge campione mondiale di Go; https://asiasociety.org/blog/asia/chinas-sputnik-moment-and-sino-american-battle-ai-supremacy
5)Rappresentazione schematica della relazione gerarchica tra intelligenza artificiale, machine learning e deep learning; elaborazione grafica dell'autore
6)Rappresentazione schematica della relazione gerarchica tra intelligenza artificiale, machine learning e deep learning; elaborazione grafica dell'autore
7)Rappresentazione schematica possibile architettura neurale artificiale; https://www.mondadorieducation.it/fisica-scientifica-ss2/le-reti-neurali-artificiali/
8)Confronto struttura neurone biologico con neurone artificiale; https://ia.zanichelli.it/scoprire-intelligenza-artificiale/parliamo-di-intelligenza-artificiale/
9)Illustrazione artistica neuroni del cervello umano; https://www.mondadorieducation.it/fisica-scientifica-ss2/le-reti-neurali-artificiali/

CAP 2 – Intelligenza artificiale generativa
10) Rappresentazione schematica architettura VAE; https://www.ibm.com/it-it/think/topics/variational-autoencoder
11) Rappresentazione schematica di una GAN; https://blogs.mathworks.com/deep-learning/2021/12/02/synthetic-image-generation-using-gans/
12) Rappresentazione schematica modello di diffusione; https://www.geeksforgeeks.

org/artificial-intelligence/diffusion-models-in-machine-learning/
13) Logo Kling; https://dblogo.com/logo/kling-ai
14) Logo Seedance (ByteDance); https://codefinity.com/blog/Revolutionizing-Video-Production-With-Seedance-2.0
15) Logo SOra2 (OpenAI); https://vocal.media/journal/sora-2-open-ai-s-new-ai-powered-video-generating-tool-how-it-works-and-what-precautions-to-take
16) Logo WAN; https://www.searchyour.ai/en/wan-video-ai
17) Logo Google VEO 3; https://www.iamdave.ai/blog/google-veo-3/
18) Logo Hailuo AI; https://logowik.com/hailuo-ai-logo-vector-76690.html
19) Logo Runaway; https://logowik.com/runway-ai-logo-vector-71391.html
20) Logo Luma Ray3; https://www.mediainfoline.com/techno/luma-ai-launches-ray3-14
21) Logo Hunyuan; https://github.com/Tencent-Hunyuan/Hunyuan-7B
22) Logo Midjourney; https://api.airforce/it/blog/api-midjourney-piu-economica-4-immagini
23) Logo Magnific; https://www.magnific.com/it
24) Logo Artlist; https://www.secrettelaviv.com/best/tags/cool-israeli-companies/artlist
25) Logo Firefly Adobe; https://www.pcmag.com/reviews/adobe-firefly
26) Logo Higgsfield; https://pub.towardsai.net/is-the-higgsfield-hype-real-or-is-it-just-an-expensive-wrapper-5f8400875b4c
27) Logo ImagineArt; https://cloud.google.com/customers/imagineart?hl=it

CAP 3 – Etica ed impatti delle immagini algoritmiche

28) tabella di immagini con il prompt “Office in Ethiopia”; https://dl.acm.org/doi/fullHtml/10.1145/3600211.3604711
29) Il caso Compas; https://medium.com/thoughts-and-reflections/racial-bias-and-gender-bias-examples-in-ai-systems-7211e4c166a1
30) Fotografia di Norbert Wiener; https://commons.wikimedia.org/wiki/File:Norbert_Wiener.webp
31) Definizione di black box; https://eastgate.medium.com/black-box-ai-what-is-it-and-how-does-it-work-705b41f531b1
32) Logo Common Crawl; https://webcrew.it/non-bloccare-ccbot-di-common-crawl/
33) Logo LAION-5B; https://www.404media.co/massive-ai-dataset-back-online-after-being-cleaned-of-child-sexual-abuse-material/
34) Immagini chatgpt studio ghibli; https://www.wired.it/article/chatgpt-studio-ghibli-copyright/
35) Immagini chatgpt studio ghibli; https://www.sentieriselvaggi.it/le-immagini-ispirate-allo-studio-ghibli-di-chatgpt-innocuo-gioco-o-violazione-dautore/
36) AI Act; https://futurodigitale.infocert.it/pillole-normative/artificial-intelligencea-act/
37) WGA sciopero; https://www.latimes.com/entertainment-arts/business story/2023-09-22/writers-strike-wga-black-writers-contraction-hollywood-hiring
38) SAG-AFTRA sciopero; https://www.latimes.com/entertainment-arts/business/

story/2023-09-25/sag-aftra-strike-authorization-video-game-interactive
39) La tecnica dei deepfake; https://www.wired.com/story/deepfakes-getting-better-theyre-easy-spot/
40) This person doesn't exist; https://thispersondoesnotexist.com/
41) Deepnude (2019); https://www.genbeta.com/actualidad/hemos-probado-deepnude-app-que-desnuda-a-mujeres-inteligencia-artificial-futuro-preocupante-sobrecogedor
42) Pagina web Deepfacelab; https://deepfacelab.org/
43) Pagina schermata Deepfacelab; https://deepfacelab.en.softonic.com/
44) Cheepfake Nancy Pelosi; https://www.youtube.com/watch?v=EfREntgxmDs
45) Deepfake Zelensky; https://www.youtube.com/watch?v=X17yrEV5sl4
46) Italian brainrot; https://medium.com/@fabianmosele/italian-brainrot-toys-and-the-commodification-of-slop-a9ee54be7bf9
47) Jeremy Carrasco profilo instagram; https://www.instagram.com/jeremyfindsai/?hl=it
48) Recensione album starface Tyga ed accoglienza del pubblico; https://pitchfork.com/reviews/albums/starface-tyga-starface/ https://www.albumoftheyear.org/album/1929479-tyga-tarface.php
49) Rack informatico; https://it.wikipedia.org/wiki/Rack_(informatica)
50) Struttura datacenter; https://www.zerounoweb.it/cloud-computing/come-e-fatto-un-data-center-ecco-la-visita-virtuale-proposta-da-microsoft/
51) Vista datacenter dall'alto; ilsole24ore.com/art/lo-stato-new-york-blocca-costruzione-nuovi-data-center-l-ai-AJGU7OK
52) Datacenter in prossimità di abitazioni; https://www.businessinsider.com/ai-artificial-intelligence-boom-data-centers-going-up-next-door-2023-7
53) Problemi idrici a casa di Beverly Morris; https://www.bbc.com/audio/play/w3ct6s29
54) Grafico stime consumi intelligenza artificiale; https://www.geopop.it/intelligenza-artificiale-consumi-energia/
55) Distribuzione mondiali datacenter; https://intelpoint.co/insights/the-united-states-of-america-hosts-over-3700-data-centres-8x-more-than-any-other-country-as-of-may-2025/

CAP 4 – Autorialità ed estetica delle immagini algorimtiche

56) Copertina n° 79 dei Cahiers du Cinéma; https://cahiersducinema.com/fr-fr/boutique/magazines/n79-janvier-1958
57) Fotogrammi estrapolati dal cortometraggio "Thank you for not answering"; https://www.youtube.com/watch?v=1TIY7jkyqz4&t=1s
58) Fotogrammi estrapolati dal videoclip "The hardest part"; https://www.youtube.com/watch?v=-Nb-M1GAOX8
59) Movie-Drome, un'allestimento di cinema espanso; https://www.domusweb.it/it/arte/2012/10/01/ghosts-in-the-machine-i-fantasmi-nella-macchina.html
60) Regista e montatore Sergej Ejzenštejn; https://it.wikipedia.org/wiki/Sergej_Michajlovi%C4%8D_%C4%96jzen%C5%A1tejn

61) Esempio di morphing facciale; https://www.keesingtechnologies.com/blog/id-documents/what-is-facial-morphing-how-can-it-affect-identity-verification/
62) Grafico uncanny valley; https://it.wikipedia.org/wiki/Uncanny_valley
63) Poster e fotogramma estrapolato da "The woman with red hair"; https://www.youtube.com/watch?v=sJD4MaEHUJo
64) Locandina film "Dreams of violets"; https://en.wikipedia.org/wiki/Dreams_of_Violets
65) Locandina e fotogramma estrapolato da "Hell Grind"; https://www.youtube.com/watch?v=t33k2tn4GpA&t=1s
66) Fotogramma estrapolato da "Last night"; https://www.youtube.com/watch?v=-3GAsTIClkkQ
67) Locandina film "The Brutalist"; https://www.imdb.com/it/title/tt8999762/
68) Applicazione del de-aging sul set del film "Here"; https://beforesandafters.com/2024/12/30/here-a-test-with-tom-hanks/
69) Locandina del film "Here"; https://www.imdb.com/it/title/tt18272208/
70) Copertina manifesto umanista collettivo Hariel; https://hariel.xyz/it/
71) Locandina "The pornographer"; https://fondazionemilano.eu/news/the-p%C3%B8rn%C3%B8graph%C9%99r-corto-realizzato-in-ai-dal-collettivo-hariel-in-concorso-alla-40%C2%AA-edizione-della-settimana-internazionale-della-critica-dell82esima-mostra-internazionale-darte-cinematografica-di-venezia
72) Fotogramma estrapolato da The Pornographer; https://sicvenezia.it/film/the-pornographer/
73) Fotogramma estrapolato da The Pornographer; https://sicvenezia.it/film/the-pornographer/
74) Locandina "The Eggregores' Theory"; https://www.imdb.com/it/title/tt33023242/
75) Una serie di fotogrammi tratti da The Eggregores' Theory; https://mubi.com/it/films/419126/player
76) Locandina documentario "Post truth"; https://www.imdb.com/it/title/tt36841032/
77) Fotogramma estrapolato da Post Truth; https://www.youtube.com/watch?v=hvK-gCCynmXA
78) Ritratto di Jon Rafman; https://en.wikipedia.org/wiki/Jon_Rafman
79) Sequenza di fotogrammi tratti da Main Stream Media Network; ttps://jonrafman.com/jon-rafman-main-stream-media-network/
80) Screenshot profilo Cloudy Heart su Spotify; https://open.spotify.com/intl-it/artist/2I7R5wSZRd2PmwohL1lqJ0
81) Fotografia allestimento mostra Main Stream Media; https://www.instagram.com/p/DbV7mljjKne/?img_index=1
82) Fotografia allestimento mostra Main Stream Media; https://www.instagram.com/p/DbV7mljjKne/?img_index=1
83) Logo ComfyUI; https://comfy.org/
84) Esempio di workflow su ComfyUi con risultato ottenuto; elaborazione grafica dell'autore
85) Sequenza di immagini generate con lo stesso prompt descritto in figura 84); elaborazione grafica dell'autore.
86) Logo Hugging face; https://www.wired.it/article/intelligenza-artificiale-hugging-face-finanziamenti-piattaforma-ai-open-source/
87) Schermata CivitAI; https://civitai.com/

Ringraziamenti

Vorrei ringraziare tutte le persone che mi sono state accanto durante questi anni e che mi hanno sostenuto anche nei momenti più difficili, anche con semplici gesti che mi hanno fatto stare meglio.

Ringrazio la mia famiglia per avermi dato il supporto necessario ad affrontare questa esperienza da fuorisede, lontano da casa. In particolare ringrazio mia madre, che mi ha sostenuto e rassicurato nei momenti in cui, con i progetti, sembrava andare tutto storto, e mio padre per i suoi consigli.

Ringrazio nonna Giulia, che con le sue visite a Torino e i suoi pranzi non mi ha mai fatto restare a stomaco vuoto.

Ringrazio la mia gatta Miza che, seppure nel suo piccolo, mi ha sempre strappato un sorriso, sia dalle foto sia nelle volte in cui tornavo a Crema.

Ringrazio i miei compagni di corso, in particolare Rick, Erik, Mato, Paolo, Sabri, Teo Castelli, Teo Bello, Lisa e Umberto: grazie per le giornate passate insieme a progettare fino a tarda notte, a parlare dei nostri lavori e a gossippare la sera bevendo qualcosa insieme.

Ringrazio Adri, per avermi fatto conoscere la città di Torino, per avermi aiutato a crescere come persona, e per i suoi ottimi pranzi, da cui spero di imparare e migliorare le mie doti culinarie.

Dedico questo lavoro a mio fratello Federico. Non hai fatto in tempo a sapere che sarei andato a Torino a vivere questi tre anni: probabilmente saresti stato curioso delle cose che facevo, e avresti provato a dare una risposta da chimico, un po' "alla Federico". Ti voglio bene, porterò la tua esperienza per sempre con me.