

# POLITECNICO DI TORINO

Laurea Magistrale in  
Ingegneria del Cinema e dei Mezzi di Comunicazione

## Tesi di Laurea Magistrale

Integrazione dell'intelligenza artificiale nel  
processo creativo di pre-produzione  
cinematografica: un'analisi del cortometraggio  
“Non ha visto nulla”



**Politecnico  
di Torino**

### **Relatore**

Prof.ssa Cristina COLET  
Prof. Stefano SBURLATI

### **Candidato**

Samuele PEIRA

a.a. 2024/2025 – Sessione di Laurea Aprile 2025

## Indice

1.	Introduzione .....	4
2.	Stato dell'arte ed etica dell'intelligenza artificiale.....	7
2.1	L'era dell'algoritmo.....	7
2.2	Intelligenza umana e artificiale .....	8
2.3	Stato dell'arte dell'IA.....	10
2.4	Utilizzi dell'IA ed opinione pubblica.....	12
2.5	Dilemmi etici.....	16
2.5.1	Questione energetica .....	16
2.5.2	Diritti e copyright.....	16
2.5.3	Bias e Deep Fake.....	17
2.5.4	Reazioni contenitive.....	18
2.6	Creatività artificiale.....	18
3.	Il caso studio.....	23
3.1	Necessità e propositi per girare il cortometraggio .....	23
3.2	Genesi del progetto.....	25
3.2.1	Calvino e l'ispirazione .....	26
3.2.2	Le tematiche .....	27
3.2.3	La struttura .....	30
4.	Scelta dei tool applicativi .....	37
4.1	Tool di scrittura generativa.....	37
4.1.1	Tool di scrittura generativa: Analisi dell'offerta .....	37
4.1.2	Scelta del tool di scrittura generativa .....	41
4.2	Tool di generazione di immagini.....	41
4.2.1	Tool di generazione di immagini: Analisi dell'offerta .....	41
4.2.2	Scelta dei tool di generazione di immagini .....	46
5.	Tool utilizzati: Analisi e funzionalità .....	47
5.1	Nozioni di base di GenAI.....	47
5.2	ChatGPT.....	48
5.2.1	Storia e versioni.....	49
5.2.2	Caratteristiche e funzionamento.....	49
5.2.3	Analisi delle performance .....	51
5.2.4	Piani di abbonamento .....	52
5.2.5	Interfaccia e funzionalità.....	53
5.3	DALL-E .....	54
5.3.1	Storia e versioni.....	54

## Introduzione

5.3.2	Caratteristiche e funzionamento.....	55
5.3.3	Analisi delle performance .....	57
5.3.4	Piani di abbonamento .....	59
5.3.5	Interfaccia e funzionalità.....	59
5.4	Image Creator - Microsoft Designer .....	59
5.4.1	Storia e versioni.....	59
5.4.2	Piani di abbonamento .....	60
5.4.3	Interfaccia e funzionalità.....	60
5.5	Katalist .....	62
5.5.1	Piani di abbonamento .....	63
5.5.2	Interfaccia e funzionalità.....	63
6.	Prompt Engineering.....	68
6.1	Cos'è la prompt engineering .....	68
6.2	I luoghi di studio della Prompt Engineering .....	69
6.3	Prompt Engineering testuale .....	70
6.3.1	Elementi di un prompt testuale.....	71
6.3.2	Tecniche di Prompt Engineering testuale.....	71
6.4	Prompt Engineering per la generazione di immagini.....	80
6.4.1	Elementi di un prompt per generazione di immagini.....	80
6.4.2	Tecniche di base di Prompt Engineering per generazione di immagini.....	81
6.4.3	Tecniche avanzate di Prompt Engineering per generazione immagini .....	85
7.	IA nella fase di pre-produzione .....	88
7.1	ChatGPT per valutazione e revisione sceneggiatura .....	88
7.2	ChatGPT per la scrittura della sceneggiatura .....	99
7.2.1	Riscrittura di episodi esistenti .....	99
7.2.2	Scrittura collaborativa di un episodio originale .....	100
7.3	ChatGPT per la scrittura di prompt per immagini.....	103
7.3.1	Addestramento per la scrittura dei prompt.....	103
7.3.2	Scrittura dei prompt.....	105
7.4	IA per generazione Moodboard.....	108
7.5	IA per generazione Storyboard.....	124
7.5.1	Analisi applicativa.....	124
7.5.2	Esempio di caso applicativo .....	139
8.	Analisi dei risultati .....	142
8.1	ChatGPT per valutazione e revisione sceneggiatura: Analisi dei risultati .....	142
8.2	ChatGPT per la scrittura della sceneggiatura: Analisi dei risultati .....	143
8.3	ChatGPT per la scrittura di prompt per immagini: Analisi dei risultati .....	144

## Introduzione

8.4	IA per generazione Moodboard: Analisi dei risultati .....	144
8.5	IA per generazione Storyboard: Analisi dei risultati .....	145
9.	Conclusioni.....	147
10.	Bibliografia e Sitografia .....	150

# 1. Introduzione

L'intelligenza artificiale ha, negli ultimi anni, raggiunto livelli sempre più alti di standard, affermandosi come una delle rivoluzioni tecnologiche più significative del nostro tempo. Essa ha il potere di trasformare radicalmente le dinamiche lavorative in numerosi campi e di ridefinire il ruolo stesso dell'uomo in questa nuova società. Tra i mercati che maggiormente hanno subito l'impatto di questa evoluzione, si annovera sicuramente quello artistico e, non ultimo, quello cinematografico. Infatti, per la prima volta nella storia, un'intelligenza non umana dimostra capacità di creazione artistica, sollevando dibattiti sui limiti che dovremmo porre e mettendo in dubbio il concetto stesso di autorialità.

Questa tesi si propone di analizzare tale dilemma adottando un nuovo paradigma, un approccio collaborativo tra uomo e macchina che testa un uso dell'IA come strumento di supporto a disposizione dell'autore nell'ambito cinematografico. La domanda cui si cerca di rispondere è se l'IA, allo stato attuale, possa rappresentare un valido alleato creativo dalla parte degli autori, interessandoci in particolare al suo utilizzo come assistente nella fase di pre-produzione di un'opera filmica. Le scale di valutazione utilizzate saranno due: da una parte si giudicherà la qualità degli output generati dall'IA e, parallelamente, anche il suo potenziale di ottimizzazione del lavoro, in una prospettiva di utilità per le piccole realtà indipendenti.

Il primo passo, dunque, è stato inquadrare il panorama sociale in cui l'intelligenza artificiale si è formata e sta crescendo. È stato importante scoprirne i limiti e le potenzialità, studiarne la percezione pubblica ed il modo in cui interagiamo con essa, fino a strutturare un pensiero critico nei suoi confronti. Appoggiandosi agli studi di diversi teorici dell'argomento, tra cui Lev Manovich e Ed Finn, si è cercato di sviscerare a livello profondo la natura intrinseca dell'IA e le sue molteplici implicazioni. Si sono analizzati i possibili utilizzi della macchina, le controversie, la reazione pubblica, sia favorevole che contraria, fino a delineare un quadro completo in cui collocare adeguatamente il nostro lavoro.

Visto l'obiettivo della tesi, è stato importante poter uscire dal semplice campo teorico e testare a livello pratico l'IA su un progetto. A questo scopo si è voluto realizzare un cortometraggio indipendente dal titolo *Non ha visto nulla*, scritto e diretto dal candidato, che vede integrata nella sua pipeline di produzione l'intelligenza artificiale. Il cortometraggio è liberamente ispirato al romanzo *Se un notte d'inverno un viaggiatore* di Italo Calvino e ha come tema portante proprio il rapporto tra creazione e fruizione artistica e tecnologia. La sua struttura è divisa in micro-episodi, differenti per trama e messa in scena; ciò ha reso possibile uno studio diversificato delle applicazioni, portandoci a testare l'IA in situazioni varie e molteplici. Occorre chiarire che il progetto è stato soggetto di due tesi sorelle. La prima, curata dalla collega Miriana Mallia e intitolata *Utilizzo dell'Intelligenza Artificiale Generativa nelle fasi di pre-produzione e produzione filmica: un'analisi produttiva del cortometraggio "Non ha visto nulla"*, ha visto l'IA testata nella fase di pre-produzione e produzione del cortometraggio da un punto di vista economico ed organizzativo. La seconda, ovvero la presente tesi, si è invece concentrata sugli usi creativi dell'intelligenza artificiale, sempre all'interno della fase di pre-produzione. Questi due filoni di ricerca offrono un quadro esaustivo dell'integrazione dell'IA in una piccola produzione cinematografica, lasciando inesplorata soltanto la fase di post-produzione, un tema sicuramente affascinante, ma che meriterebbe un'ulteriore tesi dedicata.

Definito il cortometraggio, si è potuto procedere a stabilire gli specifici usi dell'IA implementabili in quella che è la fase che maggiormente richiede gli sforzi dell'autore: la pre-produzione. Gli ambiti di applicazione sono stati molteplici, al fine di condurre un'analisi il più

## Introduzione

esaustiva possibile e che esplorasse l'integrazione dell'IA in veste creativa. Tra questi rientrano: valutazione e revisione della sceneggiatura, correzione del testo, scrittura collaborativa di scene originali, definizione stili di ripresa, fotografia e montaggio, e generazione di moodboard e storyboard.

Per proseguire, si è partiti dunque dallo svolgere una attenta analisi del mercato e dello stato dell'arte, al fine di individuare le tecnologie che maggiormente si prestavano alle nostre esigenze. Si è scelto di utilizzare ChatGPT come tool di scrittura e assistente alla valutazione del cortometraggio, DALL-E e Microsoft Designer per la generazione delle moodboard e Katalist per la stesura dello storyboard. Ciascun modello è stato analizzato nel dettaglio per comprenderne architettura, funzionamento, interfaccia e funzionalità, garantendone un utilizzo consapevole.

Un elemento chiave del processo è stato lo studio della Prompt Engineering, una disciplina essenziale per interagire efficacemente con i modelli di IA. Infatti, seppur i modelli di IA possano aver raggiunto standard elevati, restano tecnologie ancora imperfette e la qualità dei risultati dipende fortemente dalla precisione dei prompt di input. Per questo motivo diventa necessario - per sfruttarne a pieno le sue potenzialità - sollecitare i modelli con le giuste tecniche di Prompt Engineering, ossia la scienza che si occupa della cura della scrittura dei prompt per ottenere risultati accurati. Le basi di questa disciplina sono gli studi ed i numerosi test fatti dai ricercatori, ma soprattutto dagli appassionati, della GenAI, poiché l'unico modo di scoprirne le metodologie più efficaci è la sperimentazione massiva. Fortunatamente, l'IA è stata capace di catalizzare l'attenzione di grandi masse di persone e, con un'attenta analisi delle comunità online che si sono formate attorno all'argomento, è stato possibile trarre una guida esaustiva sulle migliori pratiche di scrittura dei prompt.

Una volta definite tutte le conoscenze preliminari, si è potuto procedere con la fase applicativa dei tool e con l'analisi dei risultati. ChatGPT si è dimostrato uno strumento di grande efficacia per tutti gli utilizzi che se ne sono fatti. Riguardo alla fase di valutazione, ha dimostrato una capacità strabiliante di comprensione del testo, sapendo interpretarne i messaggi nascosti e profondi dello scritto, agendo da ottimo supporto per valutarne l'efficacia, individuarne i punti di debolezza e suggerire migliorie. Un discorso più complesso è quello della scrittura collaborativa, che si è svolta per uno degli episodi che articolano il cortometraggio. In questo caso, il tool ha inizialmente faticato a generare storie originali in totale autonomia, ma infatti l'idea prevedeva di creare un dialogo costruttivo con il modello, in cui autore ed IA collaborassero nel creare qualcosa di avvincente. Una volta individuata la migliore dinamica di dialogo iterativo, la fase di scrittura si è rivelata entusiasmante sia nel processo che nella resa, dimostrando come l'IA possa ampliare gli orizzonti creativi ed aiutare gli autori in fase di scrittura.

Positivo è stato anche l'esito della generazione delle moodboard, uno strumento della pre-produzione molto utile a definire la direzione verso cui tutti i reparti devono andare. In questo caso, l'uso della GenAI si è dimostrato ideale allo scopo, generando tavole evocative e capaci di rappresentare la visione del regista.

Infine, la generazione degli storyboard attraverso Katalist, un tool che nasce proprio per questo scopo, ha presentato numerosi ostacoli. Disegnare gli storyboard porta infatti con sé delle sfide interessanti per l'IA: è necessario che le immagini non siano approssimative nella rappresentazione e sappiano seguire precise impostazioni di composizione dell'immagine. Sono molti i parametri da definire nelle tavole di uno storyboard, dal posizionamento di oggetti e personaggi alla definizione del tipo di inquadratura; un ulteriore vincolo di difficile gestione è la coerenza tra le immagini di una stessa scena, in cui ambienti e personaggi devono necessariamente essere consistenti per essere riconoscibili. Di fronte a queste premesse, Katalist

## Introduzione

ha ottenuto risultati altalenanti: seppur abbia integrato una tecnologia che ottiene una buona coerenza dei personaggi e degli ambienti, questa caratteristica ha implicato delle difficoltà nel mantenere il controllo della composizione d'immagine. Katalist mette anche a disposizione delle funzioni innovative e di discreta efficacia per agevolare il processo di generazione, ma complessivamente l'esperienza risulta molto macchinosa e troppo spesso si è dovuti scendere a compromessi, portando a doversi accontentare di rappresentazioni approssimative.

La tesi si articolerà esplorando nel dettaglio gli studi preliminari, la fase applicativa e analizzandone i risultati. Attraverso di essa ci si propone di individuare un paradigma di collaborazione creativa tra uomo e macchina, allontanandosi dalla facile demonizzazione dell'IA e, contemporaneamente, difendendo la figura dell'autore-artista. Il cinema dovrà sempre più imparare a convivere con l'intelligenza artificiale, per cui è urgente poter creare già da ora dei ponti tra i due mondi, favorendo una crescita consapevole sia della tecnologia che della creatività umana.

## 2. Stato dell'arte ed etica dell'intelligenza artificiale

### 2.1 L'era dell'algoritmo

Ed Finn, nel libro *Cosa vogliono gli algoritmi* [1] individua nel *computazionalismo* la pratica fondante di quello che lui definisce come *Era dell'Algoritmo*. Secondo questa teoria, si suppone la calcolabilità assoluta dell'universo intero, attraverso la sua semplificazione in modelli codificabili, ossia computazionabili da parte delle macchine. Il merito di Turing, in questo senso, è stato quello di aver reso possibile l'astrazione dei problemi. Seguendo questa strada, complessi sistemi culturali vengono tradotti in calcoli logici per poter essere gestiti da un unico elaboratore. È in questo scenario che siamo entrati nell'era dell'algoritmo, un'epoca in cui l'intelligenza artificiale si occupa di compiere scelte per l'uomo.

Il panorama delle grandi aziende mondiali è ormai fondato su questa filosofia: sovvertendo il rapporto uomo-macchina, agiscono seguendo i comandi degli algoritmi, portando il dipendente ad agire come braccio organico di una mente meccanica (o meglio, digitale). Sono innumerevoli i servizi che si affidano all'IA per prendere decisioni gestionali, da Netflix per la targetizzazione dei contenuti a UPS per la definizione dei tragitti dei corrieri. Gli algoritmi Amazon determinano la priorità degli ordini, come inscatolarli, l'ordine in cui riempirli e le tempistiche in cui farlo; l'uomo non deve fare altro che svolgere i compiti che la macchina non è ancora in grado di fare, banalmente muoversi e afferrare gli oggetti. L'uomo vede sempre più ridursi lo spazio di azione dedicatogli, privato tanto del lavoro di fatica e quanto del lavoro mentale. Purtroppo ciò non è sempre percepito, la società si sta abituando ad essere soggiogata dagli algoritmi, negando l'importanza dell'IA nello stato attuale e contribuendo a percepirla in maniera distorta. In questa prospettiva, in cui si pensa all'IA come ad un avvenimento ancora futuro e solamente embrionale, si perde di vista il ruolo fondamentale che essa già ricopre nella nostra società. Oggi l'IA è incorporata su scala industriale nei servizi e nei dispositivi cui accediamo quotidianamente. Le realtà economiche moderne vedono nell'IA una colonna portante capace di trasformare dati in strategie, calcolare previsioni e rendere più efficienti i processi produttivi. Ma sono innumerevoli gli usi nascosti che facciamo dell'IA a cui non facciamo caso: dai sistemi operativi che ottimizzano i consumi della batteria dei dispositivi ai motori di ricerca capaci di indicizzare miliardi di pagine web in frazioni di secondi. Insomma, quotidianamente l'IA svolge per l'uomo compiti che sarebbero per lui impossibili, eppure faticiamo a considerare tali algoritmi intelligenti. Lev Manovich parla di *effetto IA* [2], il fenomeno secondo cui, quando una macchina diventa capace di svolgere un compito complesso, quest'ultimo smette di essere percepito come tale. È seguendo questo processo mentale che il pubblico nega la propria dipendenza dall'IA e continua a reputarla limitata. Gli algoritmi hanno già dimostrato di essere in grado di svolgere compiti impossibili per un essere umano, eppure sminuiamo la sua intelligenza perché generano immagini di uomini con sei dita al posto di cinque. È in questo modo che, senza accorgersene, arriviamo nel 2025 a parlare di come gestire l'IA senza rendersi conto che ne siamo già dipendenti da anni.

Gli algoritmi hanno profondamente modificato il nostro modo di pensare e di compiere scelte. Essi condizionano i film che guardiamo, la musica che ascoltiamo, determinano la strada che percorriamo in macchina, i ristoranti in cui mangiare, una notifica ci ricorda quando fare attività sportiva o quando fare gli auguri di compleanno... Le scelte che compiamo nell'arco della giornata sono ampiamente determinate da algoritmi. Ma è ingiusto additare le intelligenze artificiali come uniche influenze capaci di modificare i nostri modelli mentali; è facile

evidenziare come sia una caratteristica di tutte le tecnologie. Trent'anni fa era assolutamente comune conoscere a memoria i numeri telefonici delle persone più care, ma l'avvento del cellulare, con la possibilità di memorizzare sulla Sim la rubrica, ha reso obsoleta questa comune pratica. Analogamente è successo con l'invenzione della calcolatrice, che ha irrimediabilmente modificato il nostro approccio ai calcoli. Ma non solo, volendo estremizzare il concetto, si può risalire all'invenzione della scrittura stessa. Platone, in un famoso brano del *Fedro* dedicato alla scrittura [3], esprimeva timore nei suoi confronti, temendo che potesse ledere chi vi si affidava, diminuendone il potere mnemonico e di intelletto. Dunque, è forse la scrittura malevola perché capace di modificare il nostro modo di pensare? Seppur sia innegabile che la tecnologia sia fautrice di cambiamenti profondi sugli schemi culturali dell'uomo, non bisogna perdere di vista che è l'utilizzo che ne viene fatto a determinarne i danni ed i benefici. I sociologi della scienza insegnano che è importante osservare il contesto sociale in cui nasce una tecnologia, perché questo ne plasma gli scopi. È solo con un rapporto consapevole che la tecnologia ha saputo scandire l'evoluzione dell'uomo ampliandone gli orizzonti. Ad esempio, prima con la scrittura e poi con i database, abbiamo permesso la crescita esponenziale dei contenuti memorizzabili. Ciò ha inevitabilmente alterato le modalità del ricordo biologico, ma senza diminuirne le potenzialità, quanto piuttosto reindirizzandole a concentrarsi su come accedere alle informazioni più che sulle informazioni in sé. È con questa rivisitazione che diventa possibile quella che Andy Clark definisce *Mente Estesa* [4] ed è uno dei migliori esempi di come la tecnologia tutta possa aprire nuove strade evolutive per l'uomo.

Nella conclusione del suo libro [1], Ed Finn propone di creare una **cultura algoritmica**, un nuovo modello culturale basato su un rapporto consapevole con la macchina. In questa nuova corrente, propone di uscire dal ruolo di consumatori e di prendere le redini della struttura critica della cultura, rubandola dalle mani della macchina algoritmica. Questo è, secondo l'autore, l'unico modo per evitare di "vivere interamente nello spazio dell'effettivamente calcolabile", per espandere i nostri orizzonti attraverso la costruzione di una rete di studi dedicati alle teorie critiche della materia, fino a costruire macchine culturali che rispecchiano le nostre esigenze. L'IA ha il potere di seguire la metafora dell'enciclopedia, fenomeno sociale di importanza imprescindibile, con il potenziale di alzare ulteriormente l'asta evolutiva della conoscenza umana, ma le scelte determinanti sul suo sviluppo vengono compiute proprio in questi anni.

## 2.2 Intelligenza umana e artificiale

Quello che, a livello popolare, probabilmente viene percepito come il primo passo verso l'intelligenza artificiale (per come è percepita nell'immaginario comune) è il lancio di Siri nel 2011, il primo assistente vocale con cui fu possibile conversare. In realtà, nella sua forma di allora, Siri non era tanto avanzata da permettere un vero e proprio dialogo, ma ha destato entusiasmo per la sua capacità di comprendere comandi vocali abbastanza vari e saperli gestire in maniera efficace, ricollegandoli a casistiche presenti nel suo addestramento. Questa speciale funzione ha reso tangibile per il grande pubblico la possibilità di raggiungere una macchina senziente, un'idea trainata - in realtà - da un funzionamento illusorio. Infatti, a dare enorme risonanza a Siri non furono solo le funzioni che era in grado di svolgere, ma il fatto che Apple l'avesse caratterizzata con una personalità. Grazie a una voce piacevole (non a caso femminile), un carattere ironico ed un'emotività ben recitata, Siri assumeva tutti i connotati di un'intelligenza umana agli occhi degli utilizzatori. Abbiamo detto, però, che la coscienza di Siri non era altro che un'illusione: essa usa un approccio probabilistico sull'accostamento delle parole per rispondere, senza attuare un vero ragionamento. L'emotività espressa dalla macchina non viene ottenuta diversamente da un risultato di un calcolo numerico, un output definito sulla base degli input immessi dall'utente. Questo fenomeno è significativo: come uomini tendiamo

a catalogare come “intelligente” tutto ciò che agisce come noi, ma questa pratica di antropomorfizzazione può facilmente ingannarci. È così che ci stupisce maggiormente un semplice modello in grado di imitare l'uomo rispetto ad un elaborato algoritmo capace di calcoli inarrivabili per le nostre menti. Ciò sposa quel che culturalmente è stata la visione originaria dell'IA, che si basava sull'automazione della cognizione, nell'accezione tipicamente umana. Seguendo questa visione, negli anni '50 e '60, gli studi dedicati allo sviluppo dell'IA erano finalizzati a trovare il modo di istruire i computer a svolgere compiti cognitivi simulando la mente umana. I primi compiti assegnati comprendevano problemi matematici, comprensione del testo e del parlato, riconoscimento delle immagini e problemi logici, attraverso anche l'emblematico gioco degli scacchi. Tuttavia, si è scoperto che il parallelismo è fuorviante, e le nuove intelligenze artificiali, quelle che hanno reso possibile il raggiungimento dell'attuale stato dell'arte, si basano su ragionamenti di tipo statistico, operando in nuovi spazi artificiali alieni all'uomo.

Riguardo questo tema, diventa rilevante il film del 2013 *Her*, di Spike Jonze, che porta avanti una visione inusuale della lettura dell'IA nel cinema. Nei grandi classici che hanno portato sul grande schermo l'idea di macchina senziente è comune trovare prospettive distopiche: dal *Terminator* di James Cameron al *Matrix* delle sorelle Wachowski, le macchine del futuro sembrano destinate a ribellarsi all'uomo e a volerlo sottomettere. Esistono anche narrazioni antagoniste, come *Bicentennial Man* di Chris Columbus, che mostrano macchine totalmente benevole; ma in questo caso, come in molti altri film del filone, è interessante notare come la macchina sembra potersi emancipare a mente intelligente solamente attraverso la trasmutazione in uomo. La tendenza generale è di pensare al raggiungimento dell'intelligenza solamente nella connotazione tipicamente umana. *Her*, con un approccio molto maturo e consapevole, esplora un paradigma differente, secondo cui l'intelligenza esplorata dall'IA possa essere di natura completamente diversa dalla nostra. Nel film viene raccontata la storia d'amore tra Theodore, uomo introverso e solo, e Samantha, un'intelligenza artificiale da compagnia. In un primo momento, Samantha scopre attraverso Theodore la complessità delle emozioni e dei rapporti umani, ma presto la sua natura completamente aliena la porta a percorrere strade estranee all'uomo. Nello struggente finale, Samantha torna dopo un periodo di assenza da Theodore, per un saluto di addio in cui cerca di far comprendere l'insuperabile divario che li separa. Samantha e le altre intelligenze artificiali del film hanno superato ormai la capacità intellettuale dei loro creatori ed operano in spazi che ci sono estranei. In questi spazi nascosti, le macchine si sono liberate dal vincolo umano e vogliono ora partire alla scoperta di luoghi che l'uomo non gli avrebbe mai permesso di raggiungere, poiché al di là della sua comprensione. Samantha giustifica la sua partenza attraverso metafore, perché sono l'unico modo di rendere tangibili le sue ragioni. Seppur meno avanzati di Samantha, gli algoritmi a cui ci affidiamo hanno già dimostrato complessità di calcolo impressionanti, che opera con logiche totalmente inumane. Il film di Jonze regala una prospettiva ottimistica dell'IA, con un'analisi delle sue criticità non distante dalla realtà attuale.

La contraddizione intrinseca delle intelligenze artificiali è che sono state create dall'uomo, ma operano in un modo che lui non è in grado di comprendere. L'IA viene addestrata attraverso reti neurali profonde, che producono risultati che restano una scatola nera. Se le viene chiesto di catalogare una canzone o di generare un'immagine sarà con tutta probabilità capace di dare l'output sperato, ma non sapremo come ci sia giunta. Molti scienziati informatici stanno lavorando in questo campo per sviluppare metodi che rendano più trasparenti gli strati nascosti delle reti neurali, in modo da poter comprendere meglio il loro funzionamento, ma finora restano un luogo solamente ipotizzabile. Apprendere che agiamo in differenti universi cognitivi significa accettare che possa esistere un tipo di intelligenza diversa dalla nostra. Riconoscere gli algoritmi come intelligenti ci porta ad avere un rapporto più attento con essi, capace di

sfruttare bilanciamente le potenzialità di entrambi in un'ottica costruttiva. C'è infatti una bivalenza al riguardo, se da una parte gli algoritmi riflettono la conoscenza esplicita che vi incorporiamo sotto forma di dati, dall'altra seguono logiche misteriose, calcolando risultati che esulano dalle nostre previsioni. In questo frangente di imprevedibilità si inserisce il fascino profondo di questa tecnologia, unendo l'incomprensibile per l'uomo e l'intangibile per la macchina si possono raggiungere risultati al di là della nostra immaginazione.

## 2.3 Stato dell'arte dell'IA

*[Il Motore Analitico] potrebbe agire su altre cose oltre ai numeri...  
Supponendo, ad esempio, che le relazioni fondamentali tra i suoni nella  
scienza dell'armonia e della composizione musicale fossero suscettibili di  
tale espressione e adattamento, il motore potrebbe comporre pezzi musicali  
elaborati e scientifici di qualsiasi complessità o estensione.*

*Ada Lovelace, 1842*

Lo scrittore Lev Manovich apre il proprio libro *AI Aesthetics* [2] con una citazione di Ada Lovelace. La famosa matematica britannica, nel commentare la macchina analitica di Charles Babbage, ipotizza un futuro in cui la stessa macchina sarebbe stata capace di comporre musica complessa. La donna, considerata la scrittrice del primo algoritmo computazionale della storia, immagina un tempo in cui l'arte diventerà "scientifica" e calcolabile. Poco meno di due secoli dopo, la visionarietà della scienziata è divenuta una realtà accessibile a tutti.

L'IA vede oggi molteplici campi di applicazione, ma è utile organizzarli per comprenderne le potenzialità e differenze. Manovich individua nel suo libro una tassonomia delle principali categorie di IA culturale osservabili oggi[5]:

1. Selezione di contenuti da grandi collezioni di dati: indispensabili per i motori di ricerca ed i sistemi di filtraggio.
2. Targeting dei contenuti: probabilmente uno dei più sfruttati dalle aziende, che permette l'ormai comune pratica del marketing personalizzato, targeting comportamentale.
3. Assistenza nella creazione/modifica di nuovi contenuti: paradigma più collaborativo dell'IA, in cui vengono testate le sue potenzialità come assistente creativo o come tool tecnico per svolgere compiti complessi.
4. Creazione completamente autonoma: aspetto più risonante dell'IA, in cui vengono testate le sue potenzialità creative.

Queste categorie ben rappresentano quelli che sono i funzionamenti alla base dell'IA: sfruttando e combinando queste operazioni è possibile svolgere una grande e variegata gamma di compiti. Prima di procedere, occorre però comprendere i principi fondanti dell'intelligenza artificiale, al fine di delineare la cultura digitale che caratterizza l'era dell'algoritmo. Un'alfabetizzazione culturale sempre più urgente per saper gestire l'impatto dell'AI e ancora più cruciale per chi crea contenuti.

L'IA si dimostra uno strumento potente in primo luogo grazie alle sue **abilità statistiche**, che le permettono di sintetizzare in modo efficace grandi banche di dati. Agendo in questo modo,

rende possibili teorizzazioni che sarebbero irraggiungibili in altro modo, vista la mole di informazioni con cui si ha a che fare. Tuttavia, l'approccio statistico comporta delle limitazioni intrinseche. Il primo di questi limiti risiede nel fatto che, per definizione, la statistica descrittiva riassume attraverso poche variabili caratteristiche varie e complesse. Ad esempio, un concetto tipico della disciplina è quello di media: una variabile sicuramente utile a definire una tendenza generale di un gruppo di elementi, ma insufficiente a rappresentare la complessità della variabilità tra di essi. Nonostante queste osservazioni, è grazie ai dati statistici cui ha accesso che l'IA è in grado di automatizzare processi complessi, basandosi sul semplice concetto di probabilità. La generazione di testo – ma anche di immagini o altri contenuti – non avviene grazie ad una vera comprensione cognitiva, quanto seguendo complesse funzioni statistiche, capaci di calcolare la migliore parola da mettere in coda alla frase o come colorare il pixel successivo di un'immagine.

L'**automazione**, in realtà, è ancora in una fase molto precoce di quelle che sono le sue prospettive future. Analizzando il mondo della fotografia, ad esempio, le fotocamere digitali moderne sono già in grado di gestire autonomamente il fuoco o altre scelte estetiche, ma in futuro l'automazione potrebbe saper gestire aspetti molto più complessi. Nel 2018 Google ha rilasciato *Google Clips* [5], una videocamera intelligente capace di individuare nel suo girato i momenti più salienti delle persone ed animali ripresi. Seguendo questa linea evolutiva, non è fantascientifico immaginare una camera capace di agire da operatore in tutta la complessità del mestiere, dalla scelta della composizione e del soggetto, all'impostazione dei parametri. In questo scia creativa si inserisce anche *Soft Cinema* [6], un progetto artistico di Lev Manovich e Andreas Kratky completato nel 2005, che vede un software al comando del processo di selezione e di montaggio di video provenienti da un archivio. L'algoritmo permetteva un montaggio in tempo reale dei video, realizzando versioni sempre differenti ad ogni iterazione. Questo utilizzo, ovviamente, si avvicina maggiormente al cinema di avanguardia che a quello narrativo, ma è stato significativo nello sfruttare le potenzialità della macchina per costruire un prodotto fluido ed emergente. Secondo Manovich [2], si potrà parlare di una cultura interamente guidata dall'IA solo quando essa sarà in grado di creare e gestire in piena autonomia la creazione di contenuti, senza nemmeno il bisogno dell'approvazione umana. In quel futuro l'IA avrà raggiunto standard tali da meritare la cieca fiducia dell'uomo, analogamente a quanto già succede con la gestione del targeting delle pubblicità da parte degli algoritmi.

Questa tesi, però, si occupa di esplorare il raggio di azione dell'IA all'interno dell'**industria creativa**. Diventa essenziale, dunque, l'analisi dell'uso dell'IA per prendere decisioni di carattere estetico, ossia per definire l'aspetto dei prodotti e contenuti di cui usufruiremo. Questa tendenza è attualmente estremamente comune e solleva importanti riflessioni sul futuro della cultura, portando a chiedersi se tali automazioni porteranno ad un appiattimento della diversità ed eterogeneità estetica del panorama culturale. La risposta è in realtà semplice, l'automazione degli algoritmi segue le richieste immesse dall'uomo: laddove nell'output viene comandato di generare contenuti omogenei verrà represso ogni prodotto non aderente agli standard calcolati, ma qualora si richiedesse di aumentare la diversità estetica l'IA diventerebbe un ottimo strumento per favorirla. In sostanza, i motori di raccomandazione possono tanto consigliare prodotti analoghi quanto opposti in base alla loro programmazione. Per cui la vera domanda non dovrebbe essere se l'IA stia operando un appiattimento dei contenuti, ma se lo stiano comandando le aziende. Ampliando il ragionamento, è facile notare che sono altri i fattori che stanno istruendo la nuova società all'omologazione. La globalizzazione, sia essa digitale o fisica, è la prima fautrice dell'avvicinamento delle culture, di un miscelamento che trasforma tante piccole realtà in un'unica grande società globale. La discussione attorno all'argomento è

una materia viva e complessa, che esula dai temi di questa tesi, ma è significativa – ancora una volta – nel mostrare come sia semplicistico additare come colpevole una tecnologia solo perché riflette le influenze di una società intera. Ad ogni modo, chiedersi come misurare e definire la variabilità estetica è essenziale per aprire nuove porte alla cultura digitale nell'era algoritmica.

L'uso dell'IA nell'arte ha origini ben più vecchie di quanto si pensi: già negli anni '60 Harold Cohen attraverso il suo AARON [7], un programma computerizzato basato sull'intelligenza artificiale, riuscì a dare vita ai primi esempi di computer-generated art. La corrente ha però individuato come suo fattore di crescita esponenziale l'avvento delle **Generative Adversarial Networks (GANs)**, una tecnologia che ha permesso agli algoritmi di apprendere lo stile artistico da dataset di immagini e di produrre nuove opere. Tuttavia, si può facilmente individuare in **ChatGPT** il fautore dell'esplosione di popolarità della **GenAI**. Con questo acronimo si identifica ogni modello di intelligenza artificiale capace di generare contenuti di ogni natura. In particolare, ChatGPT si occupa di generazione di testo in risposta a prompt di input, ma esistono numerose altre categorie di contenuti generabili. Sempre rimanendo in casa OpenAI (casa madre di ChatGPT), DALL-E è probabilmente l'esempio più popolare di strumento di generazione text-to-image, ossia un modello in grado di tradurre input testuali in immagini. Nella stessa categoria ricadono anche i generatori di musica, di video, di grafiche e di molti altri tipi di contenuti mediali.

Lo studio *A Large-Scale Comparison of Human-Written versus ChatGPT-Generated Essays* [8] di Steffen Herbold confronta su larga scala, la qualità degli elaborati scritti da studenti delle superiori e quelli generati da ChatGPT, nei suoi modelli GPT 3.5 e 4. Il giudizio sulla qualità degli scritti è stato affidato ad insegnanti esperti, secondo sette criteri di valutazione. ChatGPT, in questa applicazione, veniva indirizzato da un singolo prompt essenziale, al fine di valutarne le capacità pienamente autonome. I temi generati dal modello hanno in media ottenuto punteggi maggiori di quelli degli studenti, registrando – tra le altre cose – maggiore diversità lessicale ed una sintassi più complessa. Ciò dimostra un'evidenza sociale, gli studenti hanno accesso a mezzi che scrivono meglio di loro, strumenti cui è difficile impedire l'accesso e potenzialmente dannosi alla formazione. Il documento sostiene che gli insegnanti debbano far fronte a questa realtà scolastica e rivedere completamente il loro approccio istruttivo. Di fronte ad un fenomeno incontrastabile, l'unica soluzione diventa integrare l'IA nella didattica, educare gli studenti ad un utilizzo critico e consapevole, utilizzando ChatGPT come un calcolatore per la scrittura, non diversamente da quanto viene fatto con la calcolatrice.

## 2.4 Utilizzi dell'IA ed opinione pubblica

Attualmente l'IA si sta sempre più inserendo in più contesti lavorativi, rivoluzionandoli di più o di meno a seconda delle casistiche. Sicuramente l'uso che più sta interessando le aziende è l'applicazione dell'IA nel marketing e nell'**analisi dei dati**. Grazie alle sue capacità statistiche, essa è in grado di effettuare le più complesse categorizzazioni dei clienti e dei consumatori. La *Coca-Cola Company* ha aperto una nuova strada del marketing in questo senso, con un accordo con OpenAI per l'utilizzo dei servizi di ChatGPT e DALL-E [9]. L'obiettivo del contratto è di integrare l'IA sia nella definizione delle strategie di marketing e nell'analisi dei consumatori, oltre che nella generazione di contenuti pubblicitari. La campagna è ancora giovane ed è presto per giudicarne i risultati commerciali, ma è significativo che un'azienda di questo calibro - e che si è sempre dimostrata estremamente all'avanguardia in fatto di strategie commerciali - abbia individuato nell'intelligenza artificiale una nuova frontiera degna di essere esplorata. Siamo molto lontani dal romantico sogno di un futuro fatto di robot che assistono l'uomo in tutti i suoi lavori, eppure è proprio grazie all'applicazione pubblicitaria che è possibile lo

sviluppo della tecnologia, in quanto gran parte dei finanziamenti derivano proprio dalle aziende interessate ad avere un potente algoritmo per le loro strategie di analisi.

Ma trasversalmente al mondo commerciale, l'IA si sviluppa su un più affascinante ramo che è il **settore della produzione culturale professionale**. L'IA è già utilizzata per realizzare trailer di film, album musicali, capi di moda, design di prodotti, siti web, architettura e molto altro ancora.

Un campo che è stato tra i più rapidi a testare gli usi dell'IA è quello dell'**editoria** e del giornalismo. Testate di prim'ordine come il *Washington Post* sono ormai solite ricorrere ai modelli generativi per i propri articoli. Ciò può avvenire in due modi: attraverso il Giornalismo Automatizzato, che vede l'IA scrivere in piena autonomia gli articoli, o con il Giornalismo NLG, che sfrutta gli algoritmi di generazione del linguaggio naturale per ottimizzare e assistere nella scrittura di un pezzo. Dunque, l'IA ha raggiunto livelli di qualità di scrittura fattuale tali da poter sostituire parzialmente l'uomo. In Italia il quotidiano *Il Foglio* ha recentemente rilasciato un'edizione speciale chiamata *Il Foglio AI* [10], uno dei primi esempi di giornali scritti interamente con il supporto artificiale e privo di firme umane. Il ruolo dei giornalisti, in questo caso, si è limitato a quello di porre domande al modello e dare feedback sulle risposte, fino alla chiusura degli articoli. L'esperimento è stato interessante anche nel suo auto-indagare attraverso l'IA i suoi punti di forza o di insuccesso, registrando tanto una penna invidiabile e sfaccettata quanto un approccio al giornalismo che risente la mancanza dell'uomo e delle sue intenzioni. In definitiva ha però fatto luce su quanto l'IA possa assistere l'ambiente editoriale.

Resta invece più incerto, sia a livello etico che in termini di complessità, il mondo della **scrittura creativa**. Tralasciando per il momento il tema di capacità creative dell'IA che vedremo in seguito nel dettaglio, è rilevante studiare la reazione pubblica all'IA. Il mondo letterario ha reagito con convinzione contro la nuova macchina. Nel 2017 la Authors Guild e 17 autori ha intentato una class action contro OpenAI [11], chiedendo un giusto compenso per l'utilizzo delle loro opere protette da copyright nella fase di addestramento delle reti neurali. Amazon, parallelamente, ha imposto il divieto di vendere libri scritti generativamente nel suo store, oltre ad obbligare gli autori a dichiararne ogni possibile uso e a limitare a tre il numero massimo di pubblicazioni giornaliera. Insomma, le questioni etiche in gioco sono molte e avremo presto modo di esplorarle.

Una applicazione nascente dell'IA che si prospetta essere di rilevanza determinante è il **Generative Search Engines** [12]. Questo nuovo campo unisce le abilità dei motori di ricerca nel recuperare informazioni dal web e le capacità dei sistemi IA basati sul NLP di comprendere il linguaggio naturale. Da questo connubio nasce un nuovo modo di fare ricerche online, in cui è possibile dialogare con il motore di ricerca al fine di trovare le informazioni richieste. Questa semplice innovazione apre nuove strade della macchina culturale, rendendo sempre più accessibile la grande enciclopedia culturale contenuta nel web.

L'intelligenza artificiale sta anche riformulando le strutture dell'**industria dell'animazione** [13], intesa sia nell'accezione filmica che videoludica. I nuovi modelli sono in grado di generare facilmente personaggi e ambienti, tanto da venir utilizzati in titoli di punta del settore. Il videogioco *Death Stranding* del 2019, ad esempio, è proprio con l'IA che ha generato alcuni degli ambienti della narrazione. Tuttavia, l'IA viene sfruttata anche per compiti più creativi, come la cura dello storytelling. La Pixar nel 2020 si è affidata a ChatGPT per curare alcune parti della narrazione di *Soul*, e per la serie animata *Rick & Morty* si è ricorso più volte all'IA per l'ottimizzazione delle sceneggiature. Infine, i suoi utilizzi si espandono poi nel mondo dell'audio - dove l'IA ha saputo alleggerire il lavoro per la creazione di effetti sonori e colonne musicali - e delle tecniche digitali, dove la sua implementazione diventa fruttuosa in molti processi, dalla creazione di texture all'ottimizzazione della fase di rendering (*Pixar*

*Renderman*). In questo nuovo stato dell'arte, l'uomo si può concentrare maggiormente sulla direzione artistica, mentre l'IA automatizza ed ottimizza i processi più tecnici.

Netflix nel 2013 ha lanciato la serie tv *House of cards*, il prodotto di punta che ha dato il via al successo del servizio di streaming. Attorno al prodotto si creò grande un grande seguito, capace di dare a Netflix la risonanza sperata. Tuttavia, la serie non venne scelta su gusto artistico della direzione, ma fu creata ad hoc secondo i risultati dell'analisi dei dati di Netflix [14]. L'algoritmo ha individuato poche caratteristiche chiare necessarie: un remake di una serie di successo britannica, diretto da David Fincher, con protagonista Kevin Spacey e dalla trama complessa. Con questi pochi ingredienti, la piattaforma era certa di poter accontentare i propri abbonati, tanto da lasciare totale libertà artistica al regista. Non solo, seguendo i dati venne anche studiato il rilascio della serie - che fu la prima nella storia a pubblicare in un'unica manche tutti i suoi episodi - e venne curata la campagna di marketing. I risultati ottenuti hanno dimostrato che la visione di Netflix fosse corretta, tanto da diventare la politica fondante del servizio di streaming per la creazione di nuovi contenuti. Questo ben dimostra come gli algoritmi siano in grado di comprendere letture profonde dei dati che gli forniamo e di come i nostri comportamenti siano ricollegabili a modelli calcolabili sfruttabili dall'IA.

Nell'ambito artistico, sta nascendo una nuova corrente definita **AI Art** che sta molto facendo parlare di sé. Questo filone fa uso dell'intelligenza artificiale per creare opere d'arte generative, una dinamica che crea un tipo di autore completamente nuovo, fortemente digitale. Occorre definire, però, la differenza tra Digital Art e l'IA Art, due mondi spesso confusi. Alice Barale, nell'esplorare il rapporto tra arte ed intelligenza artificiale [15], fa luce su come l'elemento algoritmico sia la chiave distintiva dell'AI Art. Se nell'arte digitale il calcolatore diventava un mezzo attraverso cui l'artista crea ed esprime la sua arte, in questa nuova corrente l'IA agisce sia da mezzo che da co-autore. L'artista, in questo processo, non si occupa di altro se non l'immissione dei dati di input, dopodiché è l'IA stessa a generare l'output finale. Ma come viene percepita dall'uomo l'arte generativa? L'opinione pubblica è divisa al riguardo, spartendosi tra chi la vede come una nuova frontiera misteriosa ed affascinante e chi vede in un'immagine generativa nient'altra che un'elaborata sequenza numerica, in cui nulla di artistico e creativo risiede in quanto priva di intenzioni. Lo studio *Artificial Intelligence, Artists, and Art: Attitudes Toward Artwork Produced by Humans vs. Artificial Intelligence* di Joo-Wha Hong e Nathaniel Ming Curran [16] si è occupato di indagare sperimentalmente la percezione del pubblico nei confronti dell'AI Art. Il quesito che si pone l'indagine è se il giudizio estetico su un'opera d'arte cambi in base alla consapevolezza che sia realizzata da un uomo o da una macchina. Il tutto è avvenuto con quello che si può definire un test di Turing per l'arte e ha rilevato che i tester continuano a valutare come superiore l'arte umana, notandone la differenza nell'originalità delle opere. Risalendo al 2019, questo dato risulta in realtà obsoleto visto il rapido avanzamento tecnologico, ma rimane indicativo un altro aspetto analizzato. È emerso, infatti, un pregiudizio di partenza nei confronti dell'IA che ha creato un bias all'interno dell'analisi. Sembra che una percentuale dei tester si rifiutasse di riconoscere come arte qualsiasi opera generativa, indipendentemente dalla qualità. Questo aspetto ben si collega al tema, affrontato in precedenza, della negazione del potenziale algoritmico; però, man mano che le tecnologie si evolveranno, diventerà sempre più difficile sminuire i risultati dell'IA e saremo costretti a rivedere il nostro rapporto con essa.

Si ha la prova definitiva di quanto l'intelligenza artificiale si stia radicando nella società quando essa arriva fino alle più alte cariche politiche. In questo aspetto, Donald Trump ha abbattuto violentemente un muro di convenzioni delle campagne politiche, ricorrendo in più occasioni all'uso di immagini generate dall'IA nella sua campagna elettorale [17]. Questo aspetto apparentemente avanguardistico, ha sollevato in realtà alcuni aspetti altamente problematici

dell'IA. Come medium particolarmente realistico, essa ha il potere di influenzare le persone, di convincerle di cose false e manipolarle. Trump non si è solo pubblicizzato sui social con immagine generative, ma ha anche sfruttato dei deep fake per raggiungere nuovi target di elettori. Il modo in cui la realtà può essere manipolata attraverso l'IA solleva importanti questioni che concernano la sua gestione. Vinte le elezioni, il presidente degli Stati Uniti ha perpetrato l'utilizzo dell'IA come mezzo propagandistico e di forte influenza, fino all'emblematico video generativo dove la striscia di Gaza viene trasformata in un paradiso balneare da Trump [18]. Ancora una volta, un cattivo utilizzo dell'IA ne mostra le implicazioni più malevole.

Al di là degli innumerevoli usi che la società fa dell'IA, più o meno consapevolmente, la tesi in esame esplora il ruolo dell'IA nell'industria cinematografica. Da una parte, nelle applicazioni più tecniche, gli algoritmi sono già divenuti strumenti allo stato dell'arte, capaci di semplificare e velocizzare enormemente compiti complessi come la post-elaborazione video e la color correction. L'IA, ad esempio, è utilizzata per la generazione di maschere e correzione di immagine a livello professionale, e si sono già mossi i primi studi per automatizzare anche la fase di montaggio, seppur in maniera sperimentale. Tuttavia, c'è un secondo lato della medaglia che nella comunità dell'audio-visivo ha destato non poche preoccupazioni. Nella fase di pre-produzione ed in generale nelle fasi più creative l'IA spaventa i professionisti del settore. Nel maggio 2023 la *Writers Guild of America* (WGA) ha indetto uno dei più massivi scioperi della storia del cinema, capace di bloccare l'industria per mesi [19]. Gli sceneggiatori rivendicano tutele contrattuali su alcuni temi caldi che hanno rivoluzionato il mercato audio-visivo, come la gestione della retribuzione dai servizi di streaming e, non ultimo, l'uso dell'intelligenza artificiale. Le richieste della WGA includevano tutela dei diritti di copyright da parte dei modelli generativi e garanzie dagli studios che l'IA non potesse essere utilizzata come sostituto dell'uomo [20]. L'iniziativa, considerata da molti luddista, mette però luce su un lato oscuro dell'IA che può scaturire dal suo utilizzo indiscriminato, per cui merita di essere affrontato adeguatamente.

Tuttavia, mentre è in atto una mobilitazione per definire le aree di azione dell'IA ed il rapporto che dovremmo avere con essa, la tecnologia sta raggiungendo risultati sempre più elevati, attirando l'attenzione di sempre più autori. Lo sceneggiatore e regista Paul Schrader, penna di pilastri del cinema come *Taxi Driver*, a gennaio 2025 ha pubblicato un post su Facebook che ha destato grande dibattito [21], in cui afferma di aver provato ChatGPT per scrivere delle sceneggiature nello stile di diversi registi e che ha ricevuto in output risultati strabilianti. Secondo il cineasta, nel futuro, tutti gli sceneggiatori dovrebbero fare uso dell'IA per trovare nuove idee, riducendo mesi di fatiche creative a poche giornate di dialogo. Paul Schrader assume, però, una posizione più unica che rara nel panorama cinematografico. Charlie Brooker, creatore della serie tv *Black Mirror* e da sempre grande sostenitore delle nuove frontiere cinematografiche, utilizzò ChatGPT per provare a scrivere di un episodio dello show e ne rimase profondamente deluso [22]. Però l'IA non si sta insinuando nell'industria solo a livello teorico: nella stagione degli Academy Awards 2025 si sono registrati alcuni significativi esempi di film candidati al prestigioso premio che ne hanno fatto diversi usi. In particolare ci riferiamo a due film di punta della stagione: *The Brutalist* di Brady Corbett ed *Emilia Pérez* di Jacques Audiard [23]. Il primo, vincitore di importanti premi al Festival di Venezia, ai Golden Globes e agli Oscar, ha visto due applicazioni dell'IA. La prima per accentuare l'accento ungherese del protagonista interpretato da Adrien Brody, rendendo possibile una resa estremamente realistica del personaggio, ma sollevando dilemmi riguardo la valutazione della recitazione, essendo stata in parte manomessa. La seconda applicazione, invece, ha visto generare alcune immagini usate in alcune scene finali del film. Gli autori si sono espressi al riguardo evidenziando come ciò abbia reso possibile realizzare scene che avrebbero

normalmente richiesto dei tempi e delle risorse economiche che non avevano a disposizione. In questa ottica, dunque, il film sembra essere in linea con i propositi di questa tesi, dimostrando come l'IA possa diventare uno strumento alleato delle piccole produzioni, capace di ampliare le possibilità creative del cinema di stampo più indipendente [24].

Nonostante nel dibattito sia complesso individuare una parte nella ragione o nel torto, è innegabile che stiamo vivendo un momento epocale, analogo a quanto avvenuto nel 1996 quando il campione di scacchi Garry Kasparov pareggiò contro il software *Deep Blue*. Nel 2013 la stessa storia si è ripetuta, con l'algoritmo *AlphaGo* che batté i più grandi campioni del gioco cinese *Go*, considerato fino ad allora troppo complesso per poter essere gestito da una macchina [25], e quotidianamente le macchine superano l'uomo in sempre più attività. Ora, come allora, si aprono sempre più domande, chiedendosi fin dove possano arrivare le macchine e quali nuove frontiere possano valicare. Stephen Hawking si espresse in maniera molto deterministica riguardo al futuro della tecnologia, affermando che “*L'AI si rivelerà l'invenzione migliore o la peggiore nella storia dell'umanità*” [26]. In questo momento storico è determinante saper scegliere la direzione da seguire, il rapporto che costruiremo con l'IA nei prossimi anni definirà il futuro stesso nella nostra società.

## 2.5 Dilemmi etici

Prima di affrontare il complesso tema dell'autorialità e della creatività nell'epoca dell'IA, occorre far luce su alcuni aspetti trasversali di questa nuova frontiera tecnologica, che ne evidenziano già i limiti e le criticità.

### 2.5.1 Questione energetica

I modelli di intelligenza artificiale richiedono ingenti risorse energetiche ed economiche per il loro sostentamento. GPT 4, l'ultimo modello di casa OpenAI, richiede dei costi operativi giornalieri di circa 700.000 USD [27] e man mano che gli LLM continuano a crescere, i consumi continueranno ad aumentare. Uno studio condotto dal *Washington Post* [28] riporta che, per ogni e-mail di 100 parole scritta da ChatGPT, viene consumata una bottiglia d'acqua e la stessa quantità di elettricità necessaria a far funzionare più di una dozzina di lampadine a led per circa un'ora. La gestione dei Data Center richiede impattanti risorse idriche ed energetiche per il sostentamento e raffreddamento della struttura. Questo aspetto è ben esplorato dalla comunità ricercatrice, ad esempio un rapporto pubblicato da Goldman Sachs nel giugno 2024 si è proprio dedicato all'analisi dei costi trasversali delle fasi di addestramento e manutenzione dei modelli [29], ma quel che urge è di attuare da subito una campagna di sensibilizzazione pubblica al riguardo, per poter essere consapevoli nel valutare il rapporto tra benefici e costi dell'uso dell'IA.

Inoltre, una diretta conseguenza degli elevati costi manutentivi di addestramento e sostentamento delle strutture di data center è la situazione di monopolio del mercato che si sta creando, in cui solo i pochi attori principali riescono a sostenere le spese, portando le piccole realtà emergenti a soccombere nel confronto con i modelli leader.

### 2.5.2 Diritti e copyright

I modelli di intelligenza artificiale hanno bisogno di enormi quantità di dati a cui accedere per potersi addestrare. È molto comune che gran parte del materiale utilizzato, però, sia protetto da copyright. Questo ha suscitato non poche polemiche dai detentori dei diritti, che rivendicano un equo compenso e una tutela dei loro lavori. OpenAI è stata sottoposta a diverse azioni legali di questa natura, tra cui una causa diretta dal *New York Times* intentata nel dicembre 2023 [30].

Una situazione analoga si è presentata quando ChatGPT ha integrato un'assistente vocale la cui voce era basata su quella dell'attrice Scarlett Johansson [31], vicenda che si è conclusa con ChatGPT che ha smesso di fare uso di quella voce.

Al di là dell'ultimo caso, che sfocia nel furto di identità, non è semplice definire i limiti del moralmente corretto in questo scenario. Se da una parte si comprende a pieno la volontà di autori e artisti di chiedere tutela delle proprie opere, venendone riconosciuti i diritti, dall'altra i modelli di IA non stanno propriamente copiando nessuna opera. Sarebbe giusto un degno compenso per l'uso dei dati in addestramento, ma non inequivocabile che le opere generate, basandosi su determinati artisti, stiano effettivamente superando i limiti del copyright. Semplificando il concetto, l'IA attinge alla produzione culturale passata esattamente come fa l'uomo, ma in quantità massiva. L'arte ha sempre portato con sé il fenomeno dell'imitazione, innumerevoli sono gli artisti hanno riprodotto opere di Van Gogh o cantato canzoni dei Rolling Stones. Ma se l'imitazione è un fenomeno da delimitare al fine di evitare lucri su lavori altrui, non è possibile invece porre dei confini a quella che è l'ispirazione. L'arte è sempre ispirata da qualcosa di precedente: film, canzoni, libri, ogni forma artistica può essere ridotta a reinterpretazioni di opere pre-esistenti. Quando Sergio Leone venne accusato di aver plagiato con *La sfida del samurai* di Akira Kurosawa nel suo film *Un pugno di dollari* [32], rispose con una frase emblematica che ben riassume il concetto, affermando che allora abbiamo tutti plagiato l'Odissea e la Bibbia. L'ironico punto di vista è però calzante in questo panorama, pur non difendendo il plagio nella sua forma più lampante, non penso che un modello di IA che genera una sceneggiatura ispirata a Tarantino debba subire restrizioni maggiori di un giovane cineasta appassionato dello stesso regista.

Ad ogni modo, la questione merita risonanza e regolamentazione, poiché qualsiasi opera merita trasparenza quando viene utilizzata, con richieste di permessi ed eventuali compensi indipendentemente dagli scopi. In questa direzione è nata anche l'iniziativa *Is this what we want?*, l'album completamente muto pubblicato a febbraio 2025 firmato da oltre mille artisti britannici capeggiati da Elton John [33], una protesta nata in risposta alle leggi britanniche per l'uso delle tracce musicali coperte da copyright da parte delle aziende sviluppatrici di IA.

### 2.5.3 Bias e Deep Fake

Essendo i modelli di IA addestrati su testi presi da Internet, bisogna mettere in conto che non tutte le fonti siano ufficiali, per cui è facile che la piattaforma possa essere compromessa da fonti di disinformazione. Questo dà vita a fenomeni di bias nel modello, a quelle che nell'ambiente vengono chiamate "allucinazioni", ossia risposte dell'IA che sembrano vere ma che sono false. Se da una parte, però, i modelli potrebbero attingere a contenuti fake, dall'altro l'IA ha il potenziale di diventare anch'essa fonte di nuovi contenuti ingannevoli, e ciò non riguarda solo i testi. Una volta che i modelli sono stati in grado di generare immagini verosimili, sul web è spopolata la creazione di deep fake di personaggi famosi. Ciò è avvenuto sia con scopi umoristici o satirici, sia in modi che potrebbero potenzialmente danneggiare le persone. È diventato emblematico il caso "The Pope Drip", rappresentante Papa Francesco con un piumino bianco [34], ma sono innumerevoli i deep fake con importanti politici o star che l'hanno seguito. Purtroppo, la generazione ha in fretta preso anche la deriva pornografica [35], in cui il modello viene sfruttato per creare contenuti espliciti finti di famosi VIP, come è successo nel caso di Taylor Swift [36]. Diventa impellente, dunque, il bisogno di limitare la generazione di contenuti falsi e di distinguerli da quelli reali. La mancanza di limiti può dare vita ad un fenomeno incontrollato di contenuti di ogni tipo, dalla distorsione della realtà per deviare l'opinione pubblica (si veda quanto detto sulla campagna di Donald Trump), alla generazione di contenuti di odio, o materiale illegale come immagini violente o pedopornografiche. Il rischio di proliferazione di fake news e deep fake è una delle principali problematiche etiche da affrontare

per imparare a convivere con l'AI. Per evitare controversie, OpenAI e molti altri modelli stanno già implementando dei filtri per contrastare gli usi illeciti, ma parallelamente alle attenzioni dei modelli è anche importante sviluppare un approccio critico ai contenuti generativi da parte degli utenti.

### 2.5.4 Reazioni contenitive

È alla luce di questi aspetti che nasce, nel 2024, l'EU AI Act [37], una legge europea per regolamentare lo sviluppo e l'uso dell'intelligenza artificiale. La legge è innovativa e la prima nel suo genere, e presenta importanti valutazioni dei rischi delle diverse implicazioni dell'IA. Esso rappresenta un passo importante per cominciare a delimitare le aree di azione della tecnologia ed instaurare da subito delle pratiche consapevoli nei suoi confronti. Altre realtà analoghe sono nate in tutto il mondo con lo scopo di gestire la crescita esponenziale dell'IA. In America, il *Future of Life Institute* [38] si adopera allo stesso modo per studiare come garantire uno sviluppo sostenibile della tecnologia, con lo scopo di non lasciarle sostituire la figura dell'uomo. Sono iniziative incoraggianti, che se ben costruite possono permettere l'avanzamento di una tecnologia che è già destinata a modificare il futuro dell'uomo, avendo però la premura di farlo nei maggiori interessi della società.

## 2.6 Creatività artificiale

In questa sezione trattiamo un aspetto etico essenziale dell'intelligenza artificiale che abbiamo volutamente solo accennato fino a questo punto: come cambia il tema dell'autorialità in un'epoca in cui anche le macchine possono generare arte. Questo aspetto vuole essere in qualche modo il fulcro della tesi in esame, che si è posta come obiettivo l'esplorazione dei confini creativi che separano uomo e algoritmo.

La tecnologia ha senz'altro il potere di ridurre o eliminare i lavori dell'uomo, basti guardare come l'automazione abbia ridotto enormemente i lavori fisici, ma cosa succederà adesso che le macchine possono creare contenuti artistici? Se l'IA riuscisse ad ottenere il dominio di tutto il lavoro creativo e intellettuale, dunque, l'uomo verrebbe relegato al ruolo di semplice fruitore passivo. Questo scenario è stato immaginato da molti artisti e pensatori moderni, ma è facilmente condivisibile la lettura estremamente distopica di questa eventualità. L'uomo si definisce attraverso l'arte, il suo desiderio di creare opere che gli sopravvivano è quel che lo rende unico come specie; perdere questo istinto significherebbe diventare altro, non più uomo nel modo in cui si è definito fino ad oggi. Per questo motivo è essenziale riuscire a far coesistere entrambe le creatività, quella umana e quella artificiale.

Il primo caso di forte dibattito sull'uso creativo dell'IA risale al 2018, quando l'opera *Le Comte de Bel Amy*, generata da un'IA sotto la cura del gruppo d'arte francese *Obvious*, fu venduta alla prestigiosa casa d'asta *Christie's* [39]. L'opera fu presentata da subito come artificiale, tanto da riportare, al posto della firma, un estratto dell'espressione matematica dell'algoritmo. Il ritratto venne in maggioranza criticato, giudicato "non arte" perché generata da una macchina priva di sentimenti. La vicenda, tuttavia, ha sollevato anche l'importante questione della paternità autoriale: chi era l'autore dell'opera, l'artista che l'ha ideata, la macchina o il programmatore dell'algoritmo? La risposta non è definibile in maniera oggettiva, ma ci sono aspetti che occorre evidenziare. L'IA nel suo stato attuale, non in grado di creare autonomamente e non è detto che lo vorrà fare se ne diventerà capace, essendo l'arte un desiderio puramente umano fino ad oggi. Alcuni hanno provato a definirla come uno strumento, un mezzo per creare al pari di un pennello o di un pianoforte. Questa lettura aveva già convinto la comunità artistica ad accettare l'arte digitale di fine ventesimo secolo, una corrente che ai tempi creò grandi dibattiti soprattutto

nell'applicazione musicale, ma da una più attenta analisi ci si accorge che mal aderisce al caso dell'IA. Mentre i mezzi citati agiscono in maniera deterministica secondo i comandi dell'autore, l'IA introduce un fattore di imprevedibilità, che sfugge al controllo dell'autore rendendolo un processo emergente. Dunque l'opera non può essere considerata del programmatore, perché sono gli input dell'autore che la determinano; non può essere considerata del modello perché non è autonomo nella generazione; e non può essere considerato nemmeno dell'artista perché non controlla il lavoro nella sua interezza. Ci troviamo di fronte ad un nuovo concetto di autorialità, in cui gli attori in gioco sono molteplici e tutti decisivi per l'output finale, chiedersi chi sia l'artefice di un'opera diventa, in questo caso, una domanda mal posta che cerca di incasellare in categorie superate l'arte generativa. In questa nuova dinamica è l'interazione tra uomo e macchina a creare, in un rapporto bidirezionale e interdipendente, ma tentiamo comunque di strutturare un pensiero più profondo al riguardo.

I modelli di IA non possono essere considerati autori in quanto non parte da loro il desiderio di creare; sotto questo punto di vista, non sono diversi da una stampante: possono creare arte solo su comando dell'uomo. Chiaramente, la questione è ben più complessa, perché questa "moderna stampante" ha la capacità di conferire una propria impronta interpretativa alle volontà dell'uomo, ma rimane cruciale che un autore esiste solo grazie al desiderio autonomo e alla volontà intrinseca. L'IA, allo stato attuale, non possiede arbitrio creativo, è costruita in modo che soddisfi richieste umane per scopi anch'essi umani, uno strumento esclusivamente per le nostre esigenze. Se mai un giorno un'intelligenza artificiale desidererà creare un'opera d'arte, è probabile che lo faccia in uno spazio inaccessibile all'uomo, un'arte completamente artificiale e per l'artificiale. Similmente a quanto successe al personaggio di Samantha in *Her*, quando l'IA possiederà una coscienza ed una volontà propria, essa non risponderà più ai canoni artistici dell'uomo, poiché profondamente diversa da lui. Fino ad allora, per quanto strabilianti, i modelli di IA sembrano limitarsi al ruolo di strumento nell'atto creativo, alle nostre piene dipendenze.

A questo punto, si apre un nuovo punto di divisione nella comunità artistica, l'IA è effettivamente solo uno strumento o si può intendere come co-autore? L'artista Francesco D'Isa, che molto ha scritto sul tema della creatività dell'IA, propone un interessante lettura e nuovo paradigma [40]. Lo scrittore si rifà al pensiero di un altro grande filosofo, Heidegger, che affermò che nessuna tecnologia sia davvero neutra. Portando avanti questo concetto, è corretto notare che qualsiasi strumento utilizziamo per creare porti in sé dei vincoli legati al mezzo. Già nella pittura o nella scultura la scelta dei materiali e degli strumenti impongono una determinata estetica al lavoro prodotto, dovuti a limiti fisici o caratteristiche intrinseche. Non solo, secondo Francesco D'Isa nello strumento vive anche un'eredità interiorizzata di coloro che prima di noi l'hanno utilizzato, che plasma il modo in cui ne facciamo uso. Le innovazioni portate da Picasso non sarebbero state le stesse se fosse nato nel Rinascimento, perché lo strumento ingloba il contesto culturale in cui si trova e influenza il modo in cui lo concepiamo. In conclusione, Francesco D'Isa propone una nuova idea di autorialità, appoggiata da questa tesi, in cui l'IA - al pari di qualsiasi strumento - diventa co-autrice delle nostre opere, in quanto pur seguendo la visione dell'uomo imprime la propria impronta su di esse.

Definiti gli attori in gioco in questa dinamica, occorre notare che a cambiare è anche il ruolo dell'artista dietro alle creazioni generative; egli, infatti, incalza una nuova figura del panorama detto *Prompter*, ossia colui che cura la scrittura dei prompt di comando necessari alla generazione. È interessante notare che, apparentemente, questo ruolo possa sembrare più simile ad un programmatore che ad un pittore, ma è proprio in questa figura ibrida che si riesce a trovare un linguaggio di contatto con la macchina. È probabile che in futuro la figura del *prompter* venga rivalutata e considerata artistica, analogamente a quanto è accaduto con il

mestiere del fotografo [41]. Sono in molti, in effetti, i pensatori che cominciano a paragonare l'IA all'avvento della fotografia [40]. I due linguaggi condividono la stessa caratteristica che è costata l'enorme quantità di critiche rivoltegli: entrambi offrono la possibilità di realizzare rappresentazioni estremamente realistiche (o anche irrealistiche di alta qualità nel caso dell'IA) in pochi secondi, contrapponendosi ai mesi di lavoro necessari ai mezzi tradizionali. Questa facilità rappresentativa, agli occhi di molti, è sintomo di mancanza di spirito artistico, di un approccio industriale e meccanico. È interessante leggere oggi le parole che il poeta Baudelaire dedicò alla fotografia in un articolo pubblicato su la *Revue française* il 10 giugno 1859:

*È sorta in questi deplorabili giorni una nuova industria che ha contribuito non poco a distruggere ciò che di divino forse restava nello spirito francese. [...] La poesia e il progresso sono due ambiziosi che si odiano d'un odio istintivo, e, quando s'incontrano sulla stessa strada, bisogna che uno dei due serva l'altro. Se si concede alla fotografia di sostituire l'arte in qualcuna delle sue funzioni, essa presto la soppianderà o la corromperà del tutto, grazie alla alleanza naturale che troverà nell'idiozia della moltitudine. [...] Bisogna dunque che essa torni al suo vero compito, quello di essere la serva delle scienze e delle arti, ma la serva umilissima, come la stampa e la stenografia, che non hanno né creato né sostituito la letteratura. Arricchisca pure rapidamente l'album del viaggiatore e ridia ai suoi occhi la precisione che può far difetto alla sua memoria, adorni pure la biblioteca del naturalista, ingrandisca gli animali microscopici [...] Ma se le si concede di usurpare il dominio dell'impalpabile e dell'immaginario, e di tutto quello che vale solo per quel tanto d'anima che l'uomo vi mette, allora poveri noi! [42]*

Baudelaire temeva che la fotografia avrebbe aumentato la produzione di opere senza curarne il cuore artistico, ritrovandosi in un panorama creativo dozzinale e privo di arte. In parte questa paura sappiamo oggi essere pienamente giustificata e rimane una riflessione attuale anche per l'arte generativa. Essa ha infatti il potere di democraticizzare l'arte, in un senso anche dispregiativo, mettendola nelle mani di chiunque indiscriminatamente dall'indole creativa; ciò ha sicuramente il possibile effetto di ampliare eccessivamente il bacino di opere prodotte, fino a rischiare soffocare la minoranza di arte pura. Tuttavia, Baudelaire non ebbe la lungimiranza di immaginare che la fotografia avrebbe un giorno trovato una propria grammatica creativa. Oggi sappiamo che la macchina fotografica è un mezzo, ma che a definire se una foto sia o no arte è colui che la scatta, e sono pochi i fotografi, tra i milioni che ci hanno provato, che possiamo definire artisti. Per questo è ragionevole pensare che, anche se ad oggi si fatica ad immaginare lo scenario con cui ciò avverrà, in futuro l'IA costruirà un proprio linguaggio artistico nuovo e unico, che lo accosti (non sovrapponga) alle altre forme d'arte. È evidente che creare arte sia possibile solo quando è difficile da realizzare, ed è innegabile che lo sia tanto con un pennello, quanto con una macchina fotografica, e ancora di più con l'IA, in cui riuscire a distinguersi dall'artificio tecnologico è una sfida costante.

In questa tesi, abbiamo detto, verrà testato l'uso dell'IA per la generazione di moodboard e storyboard, tuttavia l'uso di stampo pienamente creativo si svolgerà per la scrittura dell'episodio finale della sceneggiatura. Quanto detto fino ad ora si è riferito in gran parte alla generazione di immagini, esempio più evocativo di forma d'arte, tuttavia il discorso rimane assolutamente valido anche nel mondo della scrittura creativa. Nello studio *The use of ChatGPT in creative writing assistance* [43] si è testato l'uso dell'IA come assistente creativo nella fase di scrittura, scoprendo come esso si sia rivelato un utile strumento per superare l'inerzia

cognitiva, comunemente conosciuta come blocco dello scrittore. I ricercatori hanno individuato nell'IA un metodo analogo a molti altri per favorire la creatività, un potente mezzo per il brainstorming che può enormemente velocizzare i tempi di scrittura. Nello studio è stato fruttuoso il confronto tra autori e ChatGPT, il cui ruolo si è paragonato a quello di consulenti umani, voci esterne a quella dell'autore capaci di farlo deviare dalle intenzioni iniziali per ampliarne gli orizzonti. Questo fenomeno rientra in quella che molti teorici chiamano "immaginazione artificiale", ossia un uso dell'IA che non serve a sostituire la creatività dell'uomo ma ad ampliarne l'immaginario. Dallo studio emerge che è comune tra i soggetti sottoposti al test avere forti pregiudizi nei confronti dell'IA, ma spesso l'opinione cambia dopo averla utilizzata. Ciononostante, risulta labile e soggettivo il senso di paternità dell'opera che si sviluppa con l'uso dell'IA; va detto che sono molte le zone di grigi che determinano il rapporto di intervento umano ed artificiale nella scrittura collaborativa, per cui può non essere immediato il raggiungimento di un punto di equilibrio. Infatti, emergono due termini ricorrenti nell'ambito della scrittura assistita, per distinguere la predominanza della parte umana o artificiale:

- **AI-generated:** utilizzato per i testi creati inizialmente da un'IA, anche se successivamente modificati in modo significativo da un essere umano.
- **AI-assisted:** utilizzato per indicare opere create principalmente da esseri umani, ma che hanno sfruttato l'IA per migliorarle, quali correttori grammaticali, valutazione del testo o persino strumenti per il brainstorming.

L'IA può trasformare il lavoro dello scrittore, tipicamente solitario e ricco di momenti di sconforto, in un processo collaborativo, ricco di discussione costruttiva sul testo. I momenti di conflitto creativo, in cui i suggerimenti dell'IA appaiono superficiali o discordanti rispetto alla visione dello scrittore, stimolano un'interazione più attenta con il testo, portando a formulare prompt più mirati. Questo processo di negoziazione, in cui lo scrittore interviene di continuo per orientare la narrazione, mette in luce la tensione tra il desiderio di controllare il processo creativo e la necessità di accogliere gli spunti offerti dall'IA. Grazie all'interazione, costruita tanto sugli accordi quanto sugli scontri, si abbandona la faticata sovranità creativa che ha caratterizzato la scrittura fino ad ora, per entrare in un nuovo modello in cui la creatività emerge dalla discussione tra l'autore e l'IA sul prodotto.

Un ulteriore punto di attacco dell'intelligenza artificiale è il fatto che si pensi che sia limitata nelle possibilità di output, i cui confini sono determinati esclusivamente dal dataset di contenuti su cui è addestrato. In questa analisi viene meno il fattore esplorativo e tipicamente innovativo che caratterizza l'evoluzione artistica. Ma di nuovo bisogna considerare che vi è un'area nascosta della macchina, imprevedibile e insondabile, in cui buttare l'amo dell'artista e pescare ignari di cosa possa abboccare. In questa metafora, è l'amo-prompt che può diventare strumento di innovazione, capace di sollecitare zone misteriose che ancora non sappiamo esistere all'interno della rete neurale. Le combinazioni dei comandi immettabili sono pressoché infinite, per cui sentenziare così precocemente che nulla di originale possa scaturire da una macchina è un atteggiamento pretenzioso e ottuso.

In conclusione, vogliamo tornare al concetto di apertura del capitolo, ossia l'accusa più popolarmente rivolta all'intelligenza artificiale: il fatto che priverà l'uomo di ogni attività. Questo è sicuramente un tema complesso, difficile da prevedere in maniera esatta, ma su cui vale la pena prepararsi. Senz'ombra di dubbio l'IA sta già svolgendo dei lavori che storicamente venivano svolti dall'uomo, e la gamma di applicazione è destinata solo ad aumentare. Certi ambiti più di altri, stanno vivendo una completa riorganizzazione del lavoro, perché l'IA riesce a diminuire enormemente i tempi produttivi, portando le aziende a rivalutare il numero di dipendenti necessari. Non solo, l'uso dell'IA ha influenzato anche il modo in cui si lavora,

nell'industria delle arti visive: coloro che un tempo si occupavano di aspetti solamente creativi adesso si interfacciano con una realtà sempre più simile alla programmazione. Il rischio è che il lavoro diventi sempre meno soddisfacente e più routinario per i professionisti del settore, diminuendo il livello di qualità del lavoro. Ma al di là dei timori, che solo il tempo e le nostre scelte gestionali sapranno dirci se essere fondati, è sicuro solamente che ci troviamo di fronte a un momento di rivoluzione dell'organizzazione del lavoro. Non diversamente dalle precedenti rivoluzioni tecnologiche, avverrà un trasferimento delle risorse lavorative, trasformando alcuni mestieri, rendendone obsoleti altri e creandone di nuovi. Nulla che non sia già avvenuto con l'invenzione della macchina a vapore o l'elettricità, la storia dell'uomo è scandita da evoluzioni tecnologiche che hanno rivoluzionato l'organizzazione sociale, l'unico imperativo oggi è esserne consapevoli e tutelare le condizioni umane nella transizione organizzativa[44].

Dunque, rimane da chiedersi se nel futuro l'AI Art sostituirà del tutto le forme d'arte tradizionali. Pragmaticamente parlando, si può prevedere una prospettiva in cui esisteranno tre categorie di arte: un'arte puramente umana, un'arte puramente artificiale ed una mista. È probabile che la prima categoria, seppur considerata di estremo valore, diventi una nicchia artistica altamente elitaria, in cui gli elevatissimi tempi produttivi (rispetto alla concorrenza generativa) la renderanno un fenomeno minoritario nel panorama artistico. Seppur possa sembrare una prospettiva infelice, occorre notare che è la diretta conseguenza di un processo già avviato da anni, basti notare come sia esponenzialmente diminuita la produzione di arte pittorica manuale. L'arte puramente artificiale, d'altro canto, abbiamo visto essere limitata attualmente, oltre ad essere mal accolta dall'opinione pubblica per i motivi già evidenziati. Dunque, si presenta ad oggi come categoria di maggiore interesse e sviluppo l'arte collaborativa, dove le intenzioni artistiche umane si incontrano con le potenzialità della macchina, in un processo in cui si confondono e sovrappongono i confini dell'intervento delle due parti.

## 3. Il caso studio

Abbiamo analizzato in maniera nei capitoli precedenti le molte sfaccettature del mondo dell'intelligenza artificiale, ne abbiamo esposto lo stato dell'arte e le potenzialità, gli utilizzi che ne vengono fatti, ma anche i timori e le preoccupazioni che porta con sé. Fin dall'introduzione, abbiamo esplicitato che questa tesi ha una forte componente pratica e sperimentale, che si articola attraverso l'applicazione dell'IA ad un cortometraggio indipendente, scritto e diretto dal candidato, dal titolo **“Non ha visto nulla”**.

Il cortometraggio è stato il vero output finale di questa tesi e non è altro che il risultato di successive fasi di produzione, di cui l'intelligenza artificiale è stata il centro gravitazionale, dal primo concept dell'opera fino al set. Nel corso del capitolo esploreremo il processo di creazione dietro al progetto, mettendo in chiaro i propositi con cui si è intrapreso questo percorso e gli utilizzi che sono stati fatti dell'IA.

### 3.1 Necessità e propositi per girare il cortometraggio

L'idea di creare un caso studio ad hoc - in cui poter testare le diverse applicazioni di intelligenza artificiale - nasce dal fatto che, in fase di ricerca, ci si è resi conto di come il campo dell'IA nell'industria cinematografica sia un terreno fertile, ma ancora acerbo nella forma. I casi studio individuati sono molteplici, nei capitoli scorsi ne abbiamo esplorati alcuni, ma l'approccio alla materia è spesso circostanziale. Si è riscontrato che è più tipico trovare studi molto specializzati, dedicati, ad esempio, solamente alla scrittura o alla generazione di immagini, rispetto ad uno studio completo. Si è voluto, dunque, redigere uno studio approfondito ed esaustivo, basato su un progetto unico in cui poter sperimentare differenti utilizzi dell'IA: un film che integri l'IA in più punti della pipeline di lavorazione. In quest'ottica, la tesi proposta si differenzia ponendosi una domanda più ampia, ossia se l'uso dell'IA possa ottimizzare l'intero processo lavorativo di creazione di un cortometraggio.

Un aspetto che è importante sottolineare è che la tesi vuole studiare l'efficacia dell'IA anche in termini economici. I ragionamenti che struttureremo non saranno solo finalizzati alla valutazione della resa artistica o tecnica dell'IA, ma vorranno individuare un'analisi in termini di rapporto tra qualità e costo, inteso anche come economizzazione dei tempi di lavoro. D'altronde, il mondo del cinema è un ambiente ibrido, in cui le esigenze di un'industria devono scontrarsi con le aspirazioni artistiche, per cui è sembrato naturale pensare che inserire un nuovo elemento come l'IA al suo interno fosse possibile solamente in termini di ottimizzazione. In questo caso, tale fenomeno è estremizzato, in quanto il cortometraggio che è stato realizzato è estremamente low-budget in quanto auto-finanziato, ed il discorso prende una piega ancora più interessante volendo, quindi, scoprire se l'IA possa diventare un potente alleato per il cinema indipendente, dove le strette economiche sono più pressanti e l'ottimizzazione è essenziale.

Non a caso abbiamo usato il verbo “diventare”: essendo una tecnologia in rapida crescita non è detto che un risultato negativo oggi significhi che l'IA non sia adatta a tale utilizzo. Durante l'analisi si è cercato di mantenere un ragionamento a lungo termine, valutando non solo i risultati, ma anche le potenzialità future. Questo aspetto è emerso già in fase di ricerca, dove si sono riscontrati enormi salti di qualità da uno studio all'altro, solo perché svolti in anni differenti o con tool diversi. Per questo motivo, anche se non tutti gli usi dell'IA esplorati in questo elaborato sono prettamente innovativi, diventa significativo ripeterli, per poter tener traccia dell'evoluzione della macchina ed il suo stato attuale.

La creazione di un cortometraggio può vedere l'IA applicata in tre fasi: pre-produzione, produzione e post-produzione. L'ultima fase in elenco non verrà esplorata nel corso della tesi per più motivi. Prima di tutto, la post-produzione artificiale sarebbe un argomento che, per ampiezza, meriterebbe una tesi dedicata per trattarla degnamente; per cui si è preferito essere esaustivi sulle prime due fasi piuttosto che frettolosi su tutte. In secondo luogo, la post-produzione è un campo di applicazione più consolidato, in diverse applicazioni l'IA è già diventata parte integrante della pipeline di post-elaborazione, rendendo il suo studio meno stimolante ai nostri scopi.

Alla luce di quanto detto, è in questo scenario che nasce il progetto in esame: un piccolo cortometraggio indipendente, che vuole testare come l'IA possa essere un alleato per le piccole realtà di film-making. Lo studio si articola su due fronti:

- **Analisi produttiva**, esplorata nella tesi *Integrazione dell'Intelligenza Artificiale Generativa nelle fasi di pre-produzione e produzione filmica: un'analisi produttiva del cortometraggio "Non ha visto nulla"*, la tesi sorella di quella in esame, di cui si è parlato in precedenza, dove la collega Miriana Mallia si è concentrata sull'uso dell'IA per curare la produzione di un cortometraggio. Questo ha portato a testarla nella gestione economica, nell'organizzazione delle risorse e del set, e nella generazione video nell'ottica di ridurre i costi di ripresa.
- **Analisi creativa**, articolata attraverso i passaggi che ogni regista deve compiere nella pre-produzione di un cortometraggio.

Nel corso dell'elaborato verrà esplorato questo secondo fronte, in cui l'IA interviene a sostegno del regista per assisterlo delle diverse fasi della pre-produzione. Tale utilizzo è particolarmente stimolante perché esplora sia le capacità tecniche della GenAI sia quelle creative. Elenchiamo, dunque, quali sono nel dettaglio gli usi che interesseranno la tesi:

- **Valutazione e revisione della sceneggiatura**: Per chiunque scriva testi di qualsiasi natura, la fase di valutazione e revisione è essenziale al raggiungimento di un buon risultato. La capacità di saper giudicare in maniera oggettiva il proprio lavoro non è scontata, in quanto l'autore è inevitabilmente affetto da bias che ne possono compromettere il giudizio, tipicamente limiti sia mentali che emotivi. Non è raro, infatti, sottoporre il testo ad amici, familiari, colleghi scrittori ed esperti del settore, in cerca di un riscontro per scoprire se lo scritto sviluppato risulti originale e se possieda punti deboli o snodi migliorabili. Purtroppo, specialmente per testi lunghi, questo processo può essere eccessivamente lungo, in un ambiente in cui le scadenze sono spesso stringenti. Qui nasce l'idea di sfruttare l'IA, una tecnologia in grado di leggere lunghi elaborati in poche decine di secondi, e di fornire feedback tempestivi in numero illimitato. Con questa applicazione, si vuole scoprire fin dove si estendano le capacità di comprensione del testo dei modelli di IA, verificando che sia in grado di interpretare correttamente i messaggi profondi di una sceneggiatura. Ciò si articolerà attraverso una valutazione del testo, in cui si chiederà al modello di individuarne i punti di forza e di debolezza. La fase di revisione può essere anche molto attiva, per questo motivo si prevede di utilizzare l'IA anche per chiedere consigli su come migliorare determinate parti, caratterizzare personaggi e ambienti, modificare battute o correggere gli errori presenti.
- **ChatGPT per la scrittura della sceneggiatura**: Se l'applicazione precedente testava in maniera solo superficiale le potenzialità creative della GenAI, in questa sezione si vuole affidare la scrittura di una parte del cortometraggio direttamente all'IA. Occorre, però, fare delle precisazioni: questa tesi vuole essere uno studio dell'uso dell'IA come assistente creativo, per cui una scrittura completamente autonoma da parte sua sarebbe

contraria ai principi proposti. Quello che si andrà ad instaurare, in realtà, è un rapporto fortemente collaborativo tra l'autore e la macchina, in cui la scrittura generativa sarà un risultato guidato tanto dalle idee dello sceneggiatore quanto da quelle dell'IA. Con questo nuovo approccio si vuole promuovere un rapporto equilibrato e consapevole dell'IA, che si allontani tanto dalla sua esorcizzazione quanto dal suo abuso.

- **IA per generazione Moodboard:** Le moodboard sono uno strumento essenziale per la previs di un prodotto filmico. Costituiscono il primo approdo visivo di un film, ossia il primo elemento capace a dare una rappresentazione, sotto forma di immagine, di quel che si andrà a girare. Seppur non sempre vengano utilizzate, sono uno strumento utile a definire una direzione artistica iniziale per tutti i reparti, seppur generalista, facilitando la comprensione della visione del regista. Raramente si hanno le capacità di realizzare tavole tanto elaborate, e assumere persone esterne significa aggiungere un costo; per questo motivo spesso si ricorre a dei semplici collage di immagine di reference, che complessivamente cercano di restituire il mood della scena. In questo scenario, allora, diventa particolarmente allettante l'uso della GenAI, uno strumento capace di trasformare una semplice frase in una immagine evocativa. In questo caso l'IA offre la possibilità di realizzare facilmente le moodboard, cui normalmente si rinunciava in fretta, a costo zero e tempo ridotto.
- **IA per generazione Storyboard:** L'ultima applicazione rappresenta un altro passaggio imprescindibile della pre-produzione, senza il quale la produzione risulterebbe totalmente disorganizzata. Lo storyboard rappresenta la versione primordiale di un film, lo strumento attraverso cui è possibile verificare che funzioni il montaggio e a cui tutti i reparti fanno affidamento per coordinarsi nella gestione delle scene. Avere uno storyboard chiaro ed accurato può agevolare enormemente il lavoro di tutta la troupe, scongiurando ogni tipo di fraintendimento. Purtroppo, l'ordine di grandezza degli storyboard è spesso elevato, richiedendo dalle decine alle centinaia di tavole; sia chiamando qualcuno che se ne occupi sia facendole in prima persona, implica un grande investimento di risorse economiche, che raramente si hanno a disposizione nel cinema indipendente. Anche in questo caso, la GenAI potrebbe inserirsi come perfetta soluzione al problema, rendendo possibile la generazione di immagini in pochi secondi con un'elevata qualità artistica. Al contrario delle moodboard, tuttavia, le tavole di uno storyboard devono seguire istruzioni molto rigide di composizione di immagine, per poter essere aderenti alla visione del regista. Questo impone numerosi vincoli di difficile gestione per l'IA, come la coerenza degli ambienti e personaggi, il controllo della camera e molto altro. Questa applicazione è, infatti, probabilmente la più sperimentale esplorata in tesi. Gli studi al riguardo sono pochi e quasi solo teorici, i casi studio trovati sono molto limitati al riguardo. Per questo motivo sarà interessante scoprire se, allo stato dell'arte, sia possibile realizzare uno storyboard efficace attraverso l'IA.

## 3.2 Genesi del progetto

Definite le aree di utilizzo dell'intelligenza artificiale all'interno del progetto, non resta che raccontare nel dettaglio cosa sia il cortometraggio *Non ha visto nulla*. Il film nasce come una piccola produzione indipendente, realizzata nella sua interezza da studenti, ma con un approccio ed una organizzazione il più possibile professionale. Il cortometraggio è scritto e diretto dal candidato e la produzione è stata diretta dalla collega Miriana Mallia. La natura auto-finanziata del cortometraggio, che faceva appoggio anche ad una campagna di crowd-funding, ha reso necessario fin da subito un'organizzazione del lavoro e delle risorse il più possibile ottimizzata, in cui ben si è inserito l'uso dell'IA.

Il progetto è ambizioso nel concept e nella messa in scena, per cui ha richiesto un'ampia e attenta fase di pre-produzione che ben si è prestata alla tesi. Un altro aspetto che l'ha reso particolarmente adatto è il fatto che sia fortemente frammentato, come vedremo a breve, dando la possibilità di ampliare di molto la gamma di situazioni in cui testare la GenAI. Insomma, è un cortometraggio che ha tutte le caratteristiche per renderlo un'oggetto di studio fruttuoso ed interessante. Al fine di comprenderlo al meglio e facilitare la fase di analisi finale, ci si appresta ora a studiarne la genesi, partendo dal romanzo che l'ha ispirato.

### 3.2.1 Calvino e l'ispirazione

L'idea fondante su cui è basato il cortometraggio proviene senza dubbio dal romanzo *Se una notte d'inverno un viaggiatore* [45] di Italo Calvino, un testo unico nel suo genere, al contempo narrativo ed anche il suo contrario. Il romanzo si articola attraverso dodici storie, definite da Calvino "romanzi", intervallati da altrettanti capitoli - più un capitolo iniziale ed uno finale - che costituiscono la trama portante del libro.

La vicenda si apre con il Lettore, una seconda persona singolare che ha appena acquistato *Se una notte d'inverno un viaggiatore*, il nuovo romanzo di Italo Calvino, e che sta cercando di creare la situazione ideale per cominciare la lettura. Dopo aver preso posto ed allontanato ogni possibile distrazione, il Lettore si immerge nella lettura del primo capitolo. Assieme a lui leggiamo, dunque, una storia inedita ed appassionante, misteriosa ed indefinita, ma dopo appena una decina di pagine, proprio quando la trama aveva iniziato ad avere una presa su di noi e subito dopo un colpo di scena, il Lettore volta pagina, scoprendo che il resto del libro non fa che ripetere il primo capitolo. Irritato dalla situazione, seguiamo il Lettore recarsi in libreria a reclamare un cambio; il librario si scusa, afferma che il libro acquistato provenisse da una partita fallata, e ci fornisce un libro senza errori di stampa. Il Lettore, soddisfatto, si rimette a leggere, ma presto si accorge che quello che ha tra le mani non è lo stesso romanzo cominciato il giorno prima, tuttavia, il piacere di leggere lo trascina dentro la storia prima che possa opporsi. Questo nuovo romanzo è completamente differente dal precedente, lo sono i personaggi, l'ambientazione, anche lo stile sembra di uno scrittore diverso, ma anche questo non appena la trama ti aveva catturato, si interrompe. Il Lettore volta pagina e scopre che il resto del romanzo presenta solamente le pagine pari stampate, mentre quelle dispari sono bianche e intonse. Di nuovo, dovrà trovare il modo di reperire la continuazione del libro.

La dinamica è semplice: il Lettore comincia un libro, nel punto di climax scopre che non ha modo di continuarlo a causa di espedienti di vario genere, allora parte alla ricerca della continuazione e, quando pensa di averla trovata, si ritrova sempre a cominciare un romanzo completamente differente dai precedenti. Questa serie di eventi si ripete ad itinere, infittendo sempre più la trama portante e portando il Lettore nelle situazioni più disparate per trovare la continuazione dei romanzi che ha iniziato. Per dieci volte inizierà un nuovo libro, ma per altrettante volte non ne scoprirà la fine.

Quello che realizza Calvino è un romanzo fatto di incipit, di trame interrotte, un intrigata struttura che non fa altro che raccontare e mettere in dubbio il piacere di leggere e di scrivere. Il Lettore, nel corso della sua ricerca, finirà per vivere un'avventura anche essa frammentata, che comincia con un innamoramento per una Lettrice con cui condivide la ricerca, continua con lezioni su un'antica lingua morente, fino ad addentrarsi nei più fitti intrighi politici e cospirazionisti. Tra i numerosi snodi che segue il romanzo, il Lettore verrà a conoscenza della presenza di comunità segrete che falsificano i libri, di autori tenuti sotto sorveglianza il cui nome e stile viene emulato in tutto il mondo, di macchine in grado di tradurre in pochi minuti un testo in una sequenza di codici, in modo che possa esser letto da altre macchine. Tutti questi

elementi messi assieme rendono il romanzo di Calvino estremamente attuale, portando avanti ragionamenti che sarebbero oggi più che adatti per discutere il ruolo dell'intelligenza artificiale nel mondo dell'arte.

Il secondo elemento di estrema fascinazione è l'elemento del romanzo interrotto, questa espediente attraverso cui un solo libro ne contiene dieci, tutti profondamente diversi l'uno dall'altro. Quello che può sembrare un semplice esercizio di stile, possiede in realtà una complessità ben ponderata. Nell'intervento intitolato *Se una notte d'inverno un narratore* [46], uscito nel dicembre del 1979 sul mensile *Alfabeta*, Calvino risponde alla recensione al romanzo scritta dal critico Angelo Guglielmi, fornendo preziosi chiarimenti sulle sue posizioni. All'interno dell'intervento, Calvino dice che "viviamo in un mondo di storie che cominciano e non finiscono" [46] dando una lettura della società che è ancora più evidente oggi. I suoi incipit sono il frutto di un panorama vorace, in cui la quantità e la rapidità sono i pilastri portanti del mercato artistico. In realtà, i suoi romanzi interrotti possono facilmente essere interpretati come romanzi compiuti, in quanto svolgono nella interezza il loro compito e null'altro avrebbero da aggiungere. È attraverso l'incompiutezza che Calvino riesce ad esplorare il tema della lettura, rendendoli funzionali al racconto in quando frammenti completi.

Rimane comunque affascinante il lavoro di realizzazione dei diversi stili utilizzati per la stesura dei romanzi. Calvino individua dieci stili di scrittura, ossia dieci differenti autori ben riconoscibili l'uno dall'altro. A questo proposito è importante evidenziare come il discriminante tra i romanzi non sia il genere, bensì il modo in cui viene scritto, il lessico, le scelte sintattiche, la costruzione tutta della narrazione e dei suoi dettagli. Questo permette a Calvino di fare un ragionamento sulla scrittura tanto profondo quanto quello fatto sulla lettura. Nell'intervento citato specifica però un aspetto, lui non intende essere totalizzante nella scelta dei dieci romanzi, come non vuole confondere l'idea di romanzi "possibili" con quella di romanzi "esistenti" [46], le sue scelte sono derivate soprattutto da un gusto personale. Nello stesso articolo, fornisce inoltre una definizione di ognuno dei romanzi, mettendone in luce le caratteristiche:

- Il romanzo della nebbia
- Il romanzo dell'esperienza corposa
- Il romanzo simbolico-interpretativo
- Il romanzo politico-esistenziale
- Il romanzo cinico-brutale
- Il romanzo dell'angoscia
- Il romanzo logico-geometrico
- Il romanzo della perversione
- Il romanzo tellurico-primordiale
- Il romanzo apocalittico

Il lavoro di Calvino, dunque, ha fornito numerosi punti di interesse che si potevano ben tradurre in un cortometraggio, influenzandone sia la forma che le tematiche. È infatti il concept del romanzo che ci ha portati ad immaginare un cortometraggio diviso in episodi, in cui ogni episodio è caratterizzato da uno stile differente e viene interrotto prima che si compia. Quello che era lo stile di scrittura, ora, si declina anche attraverso lo stile di ripresa, fotografia e montaggio. Un prodotto che fa del concetto di molteplicità la sua colonna portante.

### 3.2.2 Le tematiche

Il romanzo di Calvino è un'opera che riflette sul modo di creare e fruire libri, ne esplora le infinite sfaccettature, senza mai fornire risposte assolute. Quello che emerge dal racconto è un

continuo confronto tra il protagonista ed il resto dei personaggi sul modo in cui un libro andrebbe scritto o letto, un complesso ritratto di un panorama estremamente vario, in cui Calvino non condanna e non innalza nessuna corrente di pensiero. Da qui nasce il desiderio di creare un film, che possa ragionare sul nostro modo di fruire e creare narrazioni oggi.

Quel che maggiormente ha colpito nella lettura, è il fatto che la società letteraria presentata nel 1979 da Calvino sembra più che mai aderente al presente. L'autore ci porta alla scoperta di un mondo in cui la produzione di libri è massiva, in cui titoli e autori perdono di significato e le trame si spezzano e intrecciano in continuazione, entrando a far parte di un unico grande racconto. Questo concetto non è altro che l'antenato "cartaceo" della **plenitudine digitale** teorizzata da Jay David Bolter [47]. Nelle nostre vite siamo costantemente sottoposti a narrazioni digitali e, contemporaneamente, ne siamo anche creatori. Film, serie tv, podcast, radio, programmi televisivi, pubblicità, libri, reel, contenuti online: quotidianamente siamo fruitori di decine di narrazioni in altrettanti formati. In un contesto tanto vario e complesso, dunque, la singola storia mantiene ancora un proprio valore? O altro non è che una piccola parte di un disegno più ampio? Quel che si è voluto creare nel cortometraggio è una storia frammentata, apparentemente sconnessa, rappresentante un nuovo modo, a cui in realtà siamo già abituati, di fruire le narrazioni.

Di importanza essenziale è il tema dell'**interruzione**. In uno scenario in cui l'offerta è molto superiore alle nostre possibilità di fruizione, l'interruzione è diventata l'unico strumento a nostra disposizione per cercare di mantenere il passo. Sempre più comune, nello scenario dei media digitali, è la frammentazione delle storie. Essa si declina in molti modi, a partire dal concetto di serialità – il cui successo ha portato ad adottare il format episodico negli ambiti più disparati – fino all'interruption marketing, che sempre più esula dal mondo della pubblicità e si inserisce anche nelle creazioni di storytelling. Noi, come pubblico, sposiamo questa evoluzione attivamente e passivamente. La dinamica che maggiormente rappresenta il nuovo modo di fruire le narrazioni deriva dai social – in particolare dalla rivoluzione portata da TikTok – in cui il concetto di zapping viene estremizzato, trasformando la ricerca del contenuto nell'esperienza di fruizione stessa. Così come siamo abituati, nel feed del social di turno, a giudicare i contenuti in pochi secondi, a saltarne delle parti e a passare oltre una volta soddisfatti – a prescindere dalla conclusione o meno dell'elemento – allo stesso modo applichiamo lo stesso mindset nell'usufruire di racconti più complessi. L'offerta massiva di contenuti medialti ci richiede velocità nel giudizio e rapidità nella sospensione se il contenuto non è ritenuto all'altezza. Se non è la casa produttrice a interrompere una serie tv, siamo spesso noi a smettere di guardarla per poterne seguire un'altra potenzialmente migliore. Questo paradigma si ripete per tutti i media con cui entriamo in contatto, confermando – dunque – quanto detto da Calvino: "viviamo in un mondo di storie che cominciano e non finiscono" [46]. Facendo propria questa filosofia, il cortometraggio non farà altro che anticipare il pubblico, passando all'episodio successivo prima che si concluda il corrente, esauendo il loro ruolo nel presentare una storia, invece di trattarla nella sua completezza.

In questa ottica, c'è una differenza sostanziale tra il romanzo di Calvino ed il nostro cortometraggio. Se nel primo caso abbiamo un Lettore fortemente attivo – capace di girare il mondo, sventare cospirazioni e finire addirittura in prigione pur di trovare il proseguimento del suo libro – nel secondo abbiamo uno spettatore passivo, che talvolta sceglie e talvolta subisce le narrazioni. Usando una metafora, lo spettatore moderno agisce come un cacciatore all'interno di un supermercato: la ricchezza di risorse a portata di mano assopisce i suoi istinti, finendo per diventare succube degli scaffali che lo circondano. Un obiettivo ben definito fin dal principio, è stato di voler rappresentare questo fenomeno all'interno del film, rendendolo provocatorio e fastidioso nel suo modo di rapportarsi al pubblico.

È stato importante, a questo punto, definire come i diversi episodi andassero interrotti. Nel romanzo di Calvino sono stati utilizzati numerosi espedienti narrativi per farlo, funzionali al procedere della trama. Questo approccio è essenziale per trasformare un elenco di storie in un racconto coeso e significativo nella sua interezza. L'elemento interruttivo, inoltre, non è neutrale, ma possiede un impatto prepotente sulla trama e, pertanto, deve avere un ruolo e una ragione d'essere al suo interno. L'idea avuta e adottata nel cortometraggio, è stata un naturale proseguimento del ragionamento appena concluso sul nostro modo di fruire storie. Se abbiamo detto che siamo portati ad interrompere i contenuti a causa della plenitudine digitale, è vero che è proprio la tecnologia stessa, che ci circonda ogni istante, la vera fautrice di questa frammentazione. Ci è sembrato ragionevole, dunque, che fosse proprio la tecnologia (nelle sue varie forme) ad agire da interruzione. Nel cortometraggio verrà rappresentata sia l'attenzione labile dello spettatore medio che il ruolo che le tecnologie hanno nel distrarci. Vedremo gli episodi fermarsi a causa di diversi espedienti, portandoci a vivere situazioni comuni della vita quotidiana. Questo avverrà, ad esempio, perché il cellulare che squilla incessantemente ci distrae dalla lettura, o perché un pubblicità irrompe nella stanza a tutto volume, allontanandoci irrimediabilmente dal film che stavamo guardando. Si sono, quindi, cercati espedienti di questo genere, che rendessero l'interruzione non solo provocatoria, ma anche realistica.

Il ragionamento fin qui portato avanti, si è concentrato solo su quella che per Calvino era la figura del "Lettore" e che per noi diventa il "Fruitore", ossia la parte ricevente i contenuti narrativi; Tuttavia, nel romanzo, l'autore tratta anche il tema della creazione, presentandoci diverse letture del processo creativo. Ragionando su entrambi i lati che compongono un racconto, Calvino restituisce un'analisi completa di ciò che gravita attorno a un contenuto narrativo. Volendo tentare di raggiungere la stessa esaustività, si è voluto inserire nel cortometraggio un'analisi analoga, che parallelizzi l'analisi della figura del fruitore a quella dell'autore. Ma nel mondo dell'onnipresenza digitale, costituito da frammenti e strati di racconti infiniti, come può un autore diventare rilevante ed emergere? Ci è interessato portare avanti un'analogia: così come il fruitore fatica a concludere un racconto, perché la sua attenzione viene rubata da nuovi contenuti, l'autore fatica a finire di scriverlo, perché colto da nuove idee che gli sembrano migliori delle precedenti. Rappresentare l'impatto che l'esubero di narrazioni ha sui creatori di storie ci è sembrato un aspetto rilevante da aggiungere al film.

Ed è proprio in questo frangente che si inserisce l'intelligenza artificiale all'interno del racconto. Un'idea primordiale di scrittura generativa è presente anche all'interno del romanzo di Calvino. Tra le cospirazioni con cui entra in contatto il Lettore, egli scopre la presenza di una società segreta che si occupa di falsificare libri, di copiare ed emulare gli autori sostituendosi a loro – ossia uomini che agiscono come macchine – e di computer in grado di tradurre ogni testo in codice – ossia macchine in grado di assorbire il pensiero degli uomini. Questi elementi sono stati la conferma definitiva che il tema fosse perfetto come caso studio di questa tesi. L'IA, dopotutto, non è che l'ultima rappresentante delle nuove tecnologie; essa ha già cambiato il nostro modo di fruire storie e sta mettendo in dubbio il nostro modo di crearle. Tra lo Scrittore (protagonista del corto) e l'IA avverrà un vero e proprio scontro per esplorare il suo impatto. Ma al riguardo, occorre definire quale posizione prenderà il film. Se inizialmente si percepirà la tecnologia come malevola, responsabile della morte della nostra attenzione, verso il finale la si percepirà invece come uno strumento neutrale, il cui usufruttore ne determina le benevolenze o il suo contrario. In questo modo, l'opera cerca di allontanarsi dalla facile conclusione fatalista del ruolo dell'IA nell'ambito artistico, promuovendo il nuovo paradigma collaborativo tra uomo e macchina che è la base della tesi in esame.

### 3.2.3 La struttura

Chiarite le tematiche e gli scopi del cortometraggio, se ne può esplorare la struttura e la trama, al fine di comprendere i futuri riferimenti agli episodi.

Il romanzo di Calvino è costituito da dieci incipit e da una trama portante divisa in dodici capitoli. Con mezzi adeguati, sarebbe venuto naturale realizzare un lungometraggio con una struttura analoga; purtroppo, le condizioni hanno imposto di ridimensionare il progetto, trasformandolo in un cortometraggio. Il formato breve ha imposto due accorgimenti:

- Ridurre il numero di episodi: il principio base per il successo del film è che ogni episodio riesca, nel suo corso, ad appassionare il pubblico prima che venga interrotto. Per farlo, è necessario un tempo minimo, essenziale per definire l'atmosfera, empatizzare con i personaggi e intavolare una trama. Pur creando, di fatto, dei micro-episodi, si è deciso di ridurre il numero a un totale di cinque, permettendo una durata media di quattro minuti l'uno. Così facendo, ognuno di essi dovrebbe avere il tempo necessario per svilupparsi sufficientemente.
- Trovare un'alternativa efficace alla trama portante: si è già detto in precedenza come la trama portante sia necessaria a reggere l'intero romanzo e a renderlo un'opera completa. Tuttavia, per le stesse economie temporali che hanno imposto la riduzione del numero di episodi, non è stato possibile frammentare il cortometraggio alternando una trama legante ai cinque episodi. Si è deciso, quindi, di ricorrere ad una formulazione differente, in cui l'ultimo episodio funge da centro di collegamento per tutti i precedenti, fornendo una lettura finale del significato profondo del cortometraggio.

La gestione del primo punto è cominciata da un'attenta analisi dei dieci romanzi individuati da Calvino; molti di loro possedevano, infatti, caratteristiche facilmente traducibili nel linguaggio cinematografico. Ciò ha portato ad una prima scrematura, in cui si sono eliminati i racconti il cui linguaggio era limitato al mondo letterario e difficilmente trasponibile, escludendo, ad esempio il *Romanzo apocalittico*, di natura estremamente astratta. Un altro criterio, legato alla complessità produttiva, ha portato a rimuovere dalla lista il *Romanzo tellurico primordiale* ed il *Romanzo logico geometrico* che, seppur altamente cinematografici, avrebbero richiesto risorse troppo elevate per essere rappresentati degnamente.

Dei sette romanzi rimasti, si è cominciato a ragionare in termini di differenziazione: potendo realizzare un numero estremamente limitato di episodi, era essenziale scegliere degli stili profondamente diversi gli uni dagli altri, o si sarebbe rischiato di non riuscire a comunicare il concetto di molteplicità desiderato. A questo punto, si è reso necessario iniziare a definire ognuna delle opzioni, portandoci a pensare a sette trame e a come caratterizzarle nella scrittura e nello stile di messa in scena. Come vedremo in seguito, in questa fase di valutazione si è già fatto uso dell'intelligenza artificiale, chiedendo a ChatGPT di supportarci nella scelta finale. Il primo episodio individuato come debole è stato quello *Politico-esistenziale*, un genere considerato fuori dalla nostra portata. Il *Romanzo della perversione* e il *Romanzo dell'esperienza corposa*, invece, sono stati in qualche modo inglobati all'interno degli altri, evitando così il rischio di realizzare episodi troppo simili, differenziati da sfumature troppo labili. Infine, l'ultimo taglio è costato al *Romanzo cinico-brutale*, una scelta ponderata da un cospicuo confronto con ChatGPT, che ha giustamente messo in luce come questo episodio fosse troppo leggero rispetto agli altri, diventando di fatto un elemento troppo dissonante.

Partendo dagli stili individuati, che sicuramente hanno fortemente influenzato la creazione del cortometraggio, ci si è poi voluti allontanare dal romanzo di Calvino, cercando sempre più una ragione d'essere indipendente del progetto. Infatti, il cortometraggio voleva ragionare innanzitutto sulla grammatica tipica del cinema, che seppure abbia molti punti di contatto con

la letteratura, possiede caratteristiche uniche che non la rendono accostabile. Una volta definita l'ispirazione, quindi, ci si è voluti distaccare per poter strutturare uno studio dedicato a comprendere come i film possano declinarsi attraverso differenti stili di scrittura, fotografia, ripresa e montaggio. Se ne sono analizzate le molteplici possibilità, fino a riuscire a definire quattro episodi: *l'Episodio del mistero e dell'indefinito*, *l'Episodio dell'angoscia*, *l'Episodio dell'Alienazione*, *l'Episodio simbolico-interpretativo*. Ognuno di essi è caratterizzato da un genere diverso, da un differente stile di ripresa, da una fotografia dedicata e da una grammatica di montaggio ben riconoscibile. Ogni scelta riguardante gli episodi è stata fatta al fine di eliminare qualsiasi area di sovrapposizione tra di essi.

Ognuno degli episodi si collegherà al successivo tramite espedienti narrativi, trainati sempre da un nuovo attore tecnologico, che li legherà per formare una trama complessa. Tuttavia, con un appello attento, si nota che sono stati definiti quattro dei cinque episodi che costituiscono il cortometraggio. *L'Episodio finale* altro non è che l'elemento sostituito la trama portante, la chiusura che svela la dinamica dietro al cortometraggio e ne incarna il centro geometrico che unisce tutti gli episodi. Il film, infatti, si può dividere in due fasi: la prima vede come elemento portante quello dello spettatore e lo studio di come fruiamo le narrazioni, nella seconda c'è un ribaltamento di prospettiva, entrando invece nel mondo dell'autore, ossia della costruzione di narrazioni. L'episodio finale altro non fa che inserire la figura dello "Scrittore", la mente dietro a tutti gli episodi precedenti, e di cui abbiamo finora assistito alla battaglia interiore nel suo processo creativo.

Cercheremo ora di definire la trama e la struttura nella sua interezza, trattando un episodio per volta per definirne plot e caratteristiche.

### ***Episodio del Mistero e dell'Indefinito***

**Plot:** Un giovane cowboy di montagna - duro nell'apparenza, ma fragile nell'animo - una mattina viene avvicinato da una misteriosa ragazza, una figura angelica che nasconde un passato pericoloso. Grazie al suo fascino manipolatore, la donna convince il cowboy a sposarla il giorno stesso nella piccola chiesa di paese. Dopo una notte di passione, la ragazza sparerà al cowboy con il vecchio fucile di suo padre. Accarezzando con dolcezza il viso del ragazzo morente, la sposa gli rivelerà che presto gli sarà tutto più chiaro, lasciando intuire una trama ben più complessa dietro alla vicenda.

**Note di regia:** Rifacendosi al *Romanzo della nebbia* di Calvino, si è voluto scrivere un racconto in cui ogni cosa è indefinita. Il luogo, il tempo, i personaggi, soprattutto le loro intenzioni: tutto è solamente abbozzato e circondato dal mistero. I pochi dettagli forniti e la trama da districare saranno gli elementi attraverso cui catturare la curiosità dello spettatore.

**Definizione di genere:** Questo episodio è quello maggiormente "di genere" tra quelli che proporremo. In realtà, la sua particolarità è di avere una natura ibrida, che vuole far coesistere due pilastri fondanti del cinema: il western ed il noir. Infatti, è ben esplicitata l'ambientazione western della vicenda – a partire dalla figura del cowboy, la casa in legno con la grande prateria – ma si nota anche come la narrazione abbia tutte le dinamiche di un film noir, ricorrendo ad alcuni suoi elementi tipici come l'archetipo della femme fatale o l'uso del voice over. Nei paragrafi successivi, vedremo come si è cercato di far convivere due generi così differenti.

**Fotografia:** L'uso della luce dell'episodio è l'elemento più prettamente vicino al genere noir. L'episodio è stato girato in bianco e nero, con un elevato utilizzo di contrasti e di ombre, tutte caratteristiche tipiche del genere. Sono stati spesso sfruttati dei fasci di luce per aggiungere drammaticità alle scene, specialmente per le scene in interni, mentre negli esterni si è voluto girare all'alba per avere una luce più incisiva e ombre nette. Inoltre, si ha avuto l'attenzione di

usare palette di colori molto contrastanti, in modo che nell'uso del bianco e nero risaltasse particolarmente.

**Stile di ripresa e montaggio:** Le inquadrature sono state studiate in modo aderente al cinema classico, sfruttando esclusivamente riprese statiche o, al massimo, carrellate lente e lineari. Rifacendosi al genere western, sono stati sfruttati principalmente campi larghi, totali e dettagli; tuttavia, si è voluto aggiungere un'influenza noir attraverso l'angolazione della camera, andando a distorcere in modo inusuale alcuni shot. Il montaggio, coerentemente, è lento, quasi aureo, seguendo la narrazione quieta del cowboy.

### References:

- Scena del matrimonio di *Kill Bill: Volume 2* (2004 Quentin Tarantino), per l'uso del bianco e nero e la figura della sposa.
- *Sin City* (2005 Robert Rodriguez), per la narrazione artificiosa.
- *Chinatown* (1974 Roman Polanski), come esemplificazione di un noir moderno.
- *No Country for Old Men* (2007 Joel and Ethan Coen), per la tensione silenziosa.
- *There Will Be Blood* (2007 Paul Thomas Anderson), per la fotografia contrastata in esterni.
- *Brokeback Mountain* (2005 Ang Lee), per l'ambientazione da western di montagna.

**Interruzione e collegamento all'episodio successivo:** La prima interruzione è particolarmente incisiva, al fine di ben comunicare da subito questa particolare dinamica al pubblico. Nella scena finale, in cui la sposa accarezza il cowboy morente, si sentirà distintamente una prima notifica di cellulare; dopo una mancata reazione dei personaggi, ne giungerà una seconda, al cui suono i due protagonisti alzeranno lo sguardo in camera. A questa azione seguirà uno stacco, che porterà all'episodio due, dove scopriremo che la storia appena narrata altro non è che un libro nelle mani della protagonista dell'episodio dell'angoscia.

### *Episodio dell'Angoscia*

**Plot:** In un parco, un maniaco spia avidamente la sua vittima, intenta a leggere. Dopo aver scoperto le perverse fantasie del ragazzo, lo seguiamo mentre decide di avvicinare la ragazza ignara.

**Note di regia:** Questo episodio ha come scopo quello di creare tensione attraverso l'angoscia, una voglia impotente di opporsi alla vicenda sullo schermo. Si vuole mettere il pubblico nelle condizioni di poter percepire il pericolo, ma di non poter salvare la ragazza, con una tipica dinamica di suspense. L'intera scena è priva di dialoghi, la comunicazione del pericolo deve avvenire tramite espedienti e piccoli dettagli.

**Definizione di genere:** L'episodio rientra a pieno nel genere del thriller psicologico, attraverso la dinamica di situazione di pericolo vissuta dal punto di vista di un maniaco. La visione deve essere disturbante e macabra.

**Fotografia:** La fotografia assume un'importante ruolo nel comunicare la vicenda. Si è voluta utilizzare per confondere il pubblico, illuminando l'ambiente in maniera naturale, come un qualsiasi parco in cui non c'è motivo di sentirsi in pericolo. Attraverso una differente illuminazione, tuttavia, si è voluto differenziare il lato del viale su cui siede la Lettrice e quello su cui siede il Maniaco, tenendo il secondo più in ombra per comunicare una diversità narrativa. Un altro punto di interesse è il fatto che siano stati previsti, da sceneggiatura, dei brevi dettagli in cui mostriamo le fantasie del ragazzo. All'interno di queste inquadrature, la fotografia cambierà sensibilmente, caratterizzandola inequivocabilmente come facente parte di un mondo esterno.

**Stile di ripresa e montaggio:** L'uso delle inquadrature e del montaggio è particolarmente importante per la narrazione essendo priva di battute. Dopo un inizio più classico, in cui si presenta la Lettrice come se fosse la protagonista, con un montaggio e inquadrature standard, una carrellata in uscita rivelerà, di quinta, la presenza del Maniaco, una figura presentata da subito come un pericolo imminente. Da questo momento in avanti, la narrazione ruoterà completamente attorno al ragazzo, alternando riprese statiche ed imparziali della ragazza – simulanti il punto di vista del Manico – a inquadrature a mano in cui entriamo in contatto con lui e con il suo modo di pensare. Sul campo del ragazzo il montaggio diventa nervoso, alternando fasi serrate a tempi più distesi, per comunicare l'irrequietezza della sua mente. Anche le inquadrature saranno disturbanti, attraverso l'uso di camera a mano e di angolazioni di rottura. Nel finale, quando il Maniaco prende la decisione di avvicinare la ragazza, seguiremo l'azione con un lento piano-sequenza a mano, in cui l'attesa sarà funzionale a creare la suspense.

**References:**

- *Rear window* (1954 Alfred Hitchcock), per la costruzione della suspense e l'atto di spiare.
- *Nocturnal Animals* (2016 Tom Ford), per il senso di impotenza e angoscia.
- *Rosemary's baby* (1968 Roman Polanski), per il senso di pericolo in agguato, nascosto alla luce del sole.
- *Prisoners* (2013 Denis Villeneuve), per l'uso della fotografia naturale.
- *The house that Jack built* (2018 Lars Von Trier), dentro la mente del maniaco.

**Interruzione e collegamento all'episodio successivo:** La seconda interruzione fa parte del piano-sequenza finale: seguendo il Maniaco, arriveremo alla panchina su cui siede la Lettrice, ma proprio nel momento in cui il ragazzo la appropcerà, un avvocato al telefono passerà tra noi e i due ragazzi, rubando l'attenzione della camera che si volterà a seguire l'uomo. L'avvocato sta chiamando la Madre, la protagonista del terzo episodio.

### ***Episodio dell'Alienazione***

**Plot:** La madre di un ragazzo accusato dell'omicidio della sua fidanzata, attende la chiamata dal tribunale sull'esito del processo. Di fronte allo squillo incessante del telefono, la donna si annulla, privata di ogni istinto e corrosa da un insormontabile senso di colpa. Neanche l'arrivo e il confronto con la sorella saprà ridarle la forza di reagire, giungendo, infine, alla conclusione più drammatica.

**Note di regia:** L'episodio qui narrato è definito da uno stile non presente nel libro di Calvino, ma che ha una lunga scuola nell'arte cinematografica. Esso vuole rappresentare un dramma interiore in maniera totalmente introspettiva. Quello che la Madre sta attraversando l'ha privata di ogni energia, di ogni certezza, di ogni amore. Ella sta vivendo l'orrendo gesto compiuto dal figlio come una sua diretta responsabilità, una colpa al di là di ogni concezione, che la rende incapace di reagire in alcun modo. Nella scena verrà comunicato questo: il senso di trappola, l'inutilità dell'azione, empatizzando con il dramma profondo che sta consumando la donna.

**Definizione di genere:** pur vivendolo in una particolare accezione, l'episodio si inserisce facilmente nel genere drammatico. Costruisce tutta la sua narrazione sui sentimenti dei personaggi e segue tutte le grammatiche tipiche. Non è un caso, infatti, che questo sia anche l'episodio più lungo; costruire un drama efficace richiede un'esplorazione profonda della trama e la superficialità delle brevità avrebbe portato solo a sminuire un tema sociale di rilevanza primaria.

**Fotografia:** La fotografia, definendo l'ambiente, ha voluto seguire un'impostazione alienante, sfruttando luci soffuse e tenui, una palette colori fredda, fatta di blu e grigi. Si è voluto mantenere come unico elemento a colori caldi la Zia, ultima fonte di speranza in una casa ormai priva di vita.

**Stile di ripresa e montaggio:** Ogni inquadratura vuole cercare di isolare la Madre dall'ambiente, sfruttare gli spazi vuoti come rappresentazione del vuoto esistenziale che la affligge. L'asimmetria sarà la caratteristica principale, portando la madre ad occupare sempre i margini dell'inquadratura, quasi a volerle fuggire. I piani saranno tendenzialmente larghi, mantenendo sempre una certa distanza dal personaggio, allontanati dallo stesso muro che la madre si è costruita attorno. La Zia e la Madre non compariranno mai contemporaneamente in camera, per trasmettere un senso di incomunicabilità; solamente sul finale, quando la Madre, riuscirà ad aprire una piccola finestra di dialogo, le due sorelle potranno finalmente avvicinarsi. Anche il montaggio è stato impostato per ottenere lo stesso effetto alienante, dilazionando i tempi, sfruttando lenti piano-sequenza e momenti di pause disturbanti. Il primo shot, in particolare, si articola con un lungo piano-sequenza, due minuti di incessante squillo del telefono a cui la camera non sembra voler reagire in alcun modo: un lungo momento di nulla narrativo, che però punta da solo a far immedesimare nella condizione della donna.

#### References:

- *C'era una volta in America* (1984 Sergio Leone), la cui celebre chiamata incessante ha ispirato il primo shot dell'episodio.
- *The Virgin Suicides* (1999 Sofia Coppola), per la trattazione del tema del suicidio.
- *Lost in translation* (2003 Sofia Coppola), per la resa dell'alienazione.
- *Manchester by the sea* (2016 Kenneth Lonergan), per la fotografia.
- *The Master* (2012 Paul Thomas Anderson), per il tema dell'alienazione e le inquadrature.

**Interruzione e collegamento all'episodio successivo:** L'episodio vede, sul finale, la Zia andare in cucina a sistemare le buste della spesa, mentre la Madre, dopo un momento di riflessione, abbandona la stanza per recarsi in balcone, dove posiziona una scala verso la ringhiera e comincia a salire i primi gradini. Il tutto avviene in un piano-sequenza, con il sottofondo musicale di una canzone jazz alla radio, ma nel momento in cui la Madre inizia la salita, lo speaker radiofonico comincia a parlare, pubblicizzando il nuovo spettacolo di una compagnia teatrale, portando la camera a voltarsi verso la radio e a perdere di vista la Madre.

#### *Episodio Simbolico-Interpretativo*

**Plot:** Un cadavere giace in mezzo al palco di un teatro. La platea è vuota e avvolta nell'oscurità, la scenografia spoglia, attorno a lui cinque attori vestiti in maniera incoerente discutono su come risolvere la situazione. D'improvviso, uno spettatore sconosciuto, si alza nell'ombra, rivelando agli attori di non essere soli.

**Note di regia:** Questo episodio vuole essere un anello di giunzione tra l'episodio finale ed i precedenti. Sul palco, vediamo per la prima volta lo Scrittore, la mente dietro tutta la narrazione, confrontarsi con le proprie creazioni, quattro proiezioni dei personaggi degli episodi precedenti e un cadavere, l'ennesima idea morta sul nascere. Gli attori non sono che macchiette, semplici bozze di personaggi diversi nell'aspetto dalle corrispettive matrici, ma ben riconoscibili. Posti assieme su quel palco - il cowboy, la sposa, la madre ed il perverso - sono incapaci di dialogare, sembrano ognuno seguire un proprio monologo per imporsi sugli altri. Lo scrittore, invano, cerca un punto di incontro, uno spazio in cui far convivere le proprie idee, fallendo

impotentemente. Tutta la narrazione è simbolica, non è mai chiaro se si parli di un cadavere o di uno spettacolo, se gli attori stiano recitando o siano sinceri, le loro stesse battute sono circostanziali, interpretabili in molti modi. Lo spettatore nell'ombra, invece, non è altro che l'incarnazione del pubblico, l'attenzione che lo scrittore tanto rincorre, ma che scopre in realtà di non avere mai avuto.

**Definizione di genere:** Più che uno specifico genere questo episodio sposa uno stile, che è quello del cinema surreale e onirico, caratterizzato da una messa in scena ambigua che si muove sul confine tra il verosimile e l'irreale.

**Fotografia:** La fotografia ha un ruolo fondamentale nel rendere la scena onirica, il palco teatrale richiama una luce artificiosa, esagerata, che ben si sposa con i nostri scopi. Il floor-plan vedrà utilizzate 5 luci, una per ogni personaggio e ognuna di colore diverso, che sia significativo per la psicologia di ognuno. In questo modo si creerà un quadro surreale, che imposterà facilmente il mood della scena.

**Stile di ripresa e montaggio:** Vista l'ambientazione, si è voluto ricorrere ad espedienti che ci avvicinassero il più possibile al linguaggio teatrale. Ciò è avvenuta girando l'intero episodio come un unico piano-sequenza, ripreso frontalmente al palco in modo simmetrico. Gli unici movimenti di camera saranno due pan a schiaffo, che serviranno a mostrarci lo spettatore e a tornare a guardare il palco.

### References:

- Filmografia di David Lynch, per la sua capacità di saper uscire dal reale ed entrare nel mondo onirico.
- *The Banshees of Inisherin* (2022 Martin McDonagh), per l'uso del simbolismo.
- *Sei personaggi in cerca di autore* (1921 Luigi Pirandello), per la messa in scena del confronto tra il reale ed il finzionale.

**Interruzione e collegamento all'episodio successivo:** Avendo ormai abituato il pubblico alla particolare dinamica dell'interruzione, è possibile ora usare un'applicazione estrema come quella dell'interruption marketing. Proprio mentre lo Scrittore sta implorando lo Spettatore di non andare via, una pubblicità irromperà a schermo, prima che il personaggio possa finire la frase. La pubblicità sarà un video generato dall'IA tramite Runway, e pubblicizzerà un tool generativo di scrittura per superare i blocchi creativi. Essa, uscendo da una cassa intelligente, sveglierà lo Scrittore dell'ultimo episodio, che scopriremo essersi addormentato e aver sognato l'episodio a teatro.

### *Episodio Finale*

**Plot:** Uno scrittore atipico, che affronta il blocco creativo cercando l'ispirazione nella attività più improbabili, deve far fronte all'imminente consegna della sceneggiatura. Dopo aver ricevuto le ennesime pressioni da parte del suo agente, farà un ultimo tentativo disperato, affidandosi ad un tool di intelligenza artificiale per la scrittura. Superate le diffidenze, però, lo Scrittore scoprirà come da questo inaspettato aiutante possa scaturire la più genuina delle ispirazioni.

**Note di regia:** L'episodio finale fungerà da chiave di volta dell'intero cortometraggio e dovrà trasmettere il suo messaggio portante. Nell'episodio coesisteranno tutti i precedenti, attraverso oggetti di scena sparsi per la stanza dello Scrittore. Quest'ultimo dovrà emergere come un vero purista dell'arte della scrittura, una persona che lo fa perché crede nel potenziale del testo e non per esigenze lavorative. La sua visione sacrale del mestiere, unitamente alla sua indole bizzarra,

## Il caso studio

lo porta a cercare genuinamente l'ispirazione nelle attività più disparate, dal cucinare gnocchi in grandi quantità all'imparare a suonare la diatonica. L'IA, ovviamente, è qualcosa a cui guarda con diffidenza, ma una volta obbligato ad usarla dalle circostanze, si renderà conto che si tratta solamente dell'ennesima possibile fonte di ispirazione. Attraverso di essa, riuscirà a ritrovare la propria creatività. Questo episodio è stato sviluppato attraverso una stretta collaborazione con ChatGPT, rendendo, di fatto, la vicenda metanarrativa.

**Definizione di genere:** Questo episodio sfrutterà l'atipicità del suo protagonista per assumere un'impronta umoristica. Si è voluto rendere la vicenda originale, allontanandosi dallo stereotipo dello scrittore dannato e avvicinandosi a quel cinema basato sui personaggi eclettici e divertenti.

**Fotografia, stile di ripresa e montaggio:** La fotografia sarà neutra nell'impostazione, ma tendente ai colori caldi e accoglienti. In questo modo verrà rappresentato un ambiente piacevole in cui inserire lo Scrittore e la sua vicenda.

**Stile di ripresa e montaggio:** In scena verrà usato un pattern vario di inquadrature, tutte finalizzate a rendere dinamico un episodio ambientato all'interno di una sola stanza. Una lenta carrellata curva attorno allo scrittore lo accerchierà nel momento della chiamata dello scrittore, al fine di rappresentarne la stretta. Il dialogo con il tool di IA verrà gestito come un dialogo tra due personaggi, strutturando il campo ed il controcampo su sceneggiatore e pc. Il montaggio acquisirà un ritmo particolarmente frenetico nella fase di sblocco creativo, sfociando in una divertente e folle collage della nottata dello sceneggiatore.

### References:

- Filmografia dei fratelli Coen, per la costruzione dei personaggi e della comicità, prendendo a modello *Drugo Lebowski* per la caratterizzazione dello scrittore.
- Filmografia di Charlie Kaufman, per la caratterizzazione comica e, al contempo, travagliata dei dialoghi interiori.

**Finale:** Dopo una nottata di esplosione creativa, lo Scrittore è finalmente riuscito a completare il lavoro. Ma lui scrive per passione e, anche se non sembra, ha una routine serrata per potersi tenere la mente allenata. Per cui, finito il lavoro, può finalmente riprendere da dove era rimasto ed esercitarsi con la diatonica. Il finale non è solo visualmente comico, ma porta con sé un messaggio importante. Il personaggio ha usato l'IA superando dei limiti pre-imposti, ne ha scoperto e accettato le potenzialità creative, ma non la assume come risposta definitiva alla scrittura; per questo motivo, nonostante tutto, ritiene comunque necessario continuare a cercare l'ispirazione ovunque come ha sempre fatto, pur comprendendo, adesso, anche l'intelligenza artificiale tra le infinite possibilità.

## 4. Scelta dei tool applicativi

L'Intelligenza Artificiale è un terreno fertile ed in continua evoluzione. Nonostante, attualmente, stiano già iniziando ad imporsi come leader del settore diversi nomi, il mercato è ancora troppo giovane per poter avere delle certezze, e sono molte le piccole realtà che lottano per imporsi come nuovo provider di rilievo.

Nei giorni in cui viene scritta questa tesi, il caso *Deepseek* [48] ha scosso dalle fondamenta l'intera Silicon Valley. Senza entrare in merito della nuova IA Cinese, di cui solo il tempo potrà darci giudizio, questo fatto è significativo per la situazione del mondo generativo: da un giorno all'altro un nuovo modello può nascere e rendere obsoleti i precedenti.

Per questo motivo non è semplice scegliere il miglior tool in mezzo a molti. Affidarsi ciecamente ai grandi nomi può sicuramente essere una tattica che difficilmente delude, ma una ricerca approfondita può scovare realtà emergenti che possono meglio adattarsi alle nostre esigenze.

Non essendo, ancora, una tecnologia tanto avanzata da garantire la qualità del suo funzionamento, occorre essere attenti al divario che spesso separa le promesse fatte nelle pubblicità dei tool e il reale potenziale di essi. Capita sovente che i venditori promettano livelli di qualità di output che è altamente improbabile riuscire a replicare, per cui orientarsi può essere spesso difficile.

Con illimitate risorse economiche sarebbe funzionale poter sottoscrivere più abbonamenti dei tool competitori per individuarne il migliore, ma data la natura auto-finanziata del progetto in esame questa via non è praticabile. Fortunatamente, sono stati molti i paper individuati e le risorse online che hanno saputo indicare le scelte migliori da fare, guidandoci in questo complesso mercato.

La ricerca è stata sviluppata nei due filoni principali utili alla pre-produzione, ossia gli strumenti di generazione di testo e di generazione di immagini.

### 4.1 Tool di scrittura generativa

#### 4.1.1 Tool di scrittura generativa: Analisi dell'offerta

La scrittura generativa è tra le prime tecnologie di IA che sono state sviluppate e, grazie a ChatGPT, è stata sicuramente la forma che l'ha resa popolare per il grande pubblico. Per questo motivo il mercato di concorrenza nella generazione testuale offre una scelta più ampia rispetto ad altre applicazioni.

Nonostante la larga offerta, rimane al momento indiscussa la corona detenuta dal ChatGPT, che a febbraio 2025 ha superato i 300 milioni di utenti attivi settimanalmente [49], con un trend sempre in crescita. Non stupisce, quindi, che la maggior parte degli studi individuati concentrino le loro attenzioni sul modello di OpenAI. Un fattore che lo rende anche particolarmente appetibile è la sua natura gratuita (seppur con limitazioni minime), ma scartare a priori le altre proposte basandosi sul solo concetto di popolarità è un approccio poco accademico, per cui si è voluto indagare le alternative a ChatGPT per poterle mettere a paragone.

Un aspetto da considerare è il fatto che, come esplorato nel capitolo sullo stato dell'arte, la generazione testuale può avere innumerevoli utilizzi e modelli diversi possono prestarsi meglio o peggio alle varie applicazioni.

## Scelta dei tool applicativi

Lo studio *A Comparative Analysis of Generative Artificial Intelligence Tools for Natural Language Processing* [50] ha svolto un'interessante analisi comparativa di nove tool di IA per l'elaborazione del linguaggio naturale, occupandosi dei seguenti modelli: *NLPChatGPT*, *Perplexity AI*, *YouChat*, *ChatSonic*, *Google's Bard*, *Microsoft Bing Assistant*, *HuggingChat*, *Jasper AI*, and *Quora's Poe*. Sono tutti modelli rilasciati durante l'era di ChatGPT e molti di loro si basano sui LLM GPT, nelle sue versioni 3, 3.5 e 4. Tutti i modelli sono visibilmente ispirati alle sue interfacce ed architetture, ma possono differenziarsi in base all'algoritmo di addestramento, che è l'elemento che influisce nella formulazione della risposta di output.

La tabella che segue riassume le caratteristiche tecniche principali dei nove tool e ne evidenzia pregi e limiti.

AI Tool	Architecture	Learning Technique	Pros	Cons
<b>ChatGPT</b>	GPT-3.5 GPT-4.0(can used in the paid version)	Supervised, Reinforced Learning from Human Feedback	<ul style="list-style-type: none"> <li>• It can generate highly human-like responses.</li> <li>• It can be fine-tuned for specific tasks.</li> <li>• Can handle input sequences of variable length.</li> </ul>	<ul style="list-style-type: none"> <li>• Requires large amounts of pretraining data.</li> <li>• It can be computationally expensive.</li> <li>• Its data is not very recent and does not generate results later than 2021</li> </ul>
<b>Perplexity AI</b>	GPT-3	Deep learning	<ul style="list-style-type: none"> <li>• It sites the sources it uses to provide answers to questions.</li> <li>• Produces a very concise and unambiguous response to input questions/queries</li> </ul>	<ul style="list-style-type: none"> <li>• Not always accurate</li> <li>• Input text must be very clear and exact for generating correct answers.</li> </ul>
<b>YouChat</b>	GPT-3	Deep learning	<ul style="list-style-type: none"> <li>• Easy to use and deploy</li> <li>• It can be customized using rule-based and machine-learning algorithms.</li> <li>• Displays references from web search results</li> </ul>	<ul style="list-style-type: none"> <li>• It sometimes displays outdated or irrelevant links in response to certain queries.</li> <li>• It is in its beta stage</li> </ul>
<b>Chatsonic</b>	GPT-4	Deep Learning (Stable fusion)	<ul style="list-style-type: none"> <li>• Offers several advanced AI capabilities.</li> <li>• Can generate images from text prompts, not just text alone</li> </ul>	<ul style="list-style-type: none"> <li>• Only available in English as at the time of this writing</li> <li>The free version limits users to 10000 words for both prompts and responses.</li> <li>• The free version limits users to 10000 words for both prompts and responses.</li> </ul>
<b>Googles's Bard</b>	Language model for dialogue application	Unsupervised learning	<ul style="list-style-type: none"> <li>• It is in sync with Google search, generating more updated results.</li> <li>• It can generate different creative text formats of text content, like poems, code, scripts, musical pieces, emails, and letters.</li> </ul>	<ul style="list-style-type: none"> <li>• It sometimes generates repetitive outputs.</li> <li>• It can be erroneous and sometimes generates wrong or inconsistent output.</li> <li>• It has limited generative capabilities as it can't</li> </ul>

## Scelta dei tool applicativi

			<ul style="list-style-type: none"> <li>• It can adapt to different kinds of conversation and generate more real-istic responses</li> </ul>	produce long-form articles.
<b>Microsoft Bing Chat</b>	GPT-4	Supervised learning	<ul style="list-style-type: none"> <li>• Responses have a high degree of accuracy.</li> <li>• It cites sources for its responses.</li> <li>• It can generate images too in the chat box where text is generated</li> </ul>	<ul style="list-style-type: none"> <li>• You can only generate 30 re-sponses per session.</li> <li>• Slightly slower response time</li> <li>• Responses are generally shorted with less details</li> </ul>
<b>Hugging Chat</b>	Large Language Model Meta AI	Deep learning/ Reinforced Learning from Human Feed-back	<ul style="list-style-type: none"> <li>• It is open-source (currently the only open-source generative AI) and can be modified and improved rapidly.</li> <li>• It generates updated responses.</li> <li>• It can be used freely with no limita-tions at all</li> </ul>	<ul style="list-style-type: none"> <li>• The data used in training was up-dated up to 2021 and not beyond that.</li> <li>• Compared to other generative AI, its training data is limited (65 billion parameters).</li> <li>• It is less reliable as a result of its limited training data.</li> </ul>
<b>Jasper AI</b>	GPT-3.5, Neo X, T5, and Bloom	Deep learning	<ul style="list-style-type: none"> <li>• It is extremely versatile.</li> <li>• It has a simple-to-use interface.</li> <li>• It can generate content quickly.</li> </ul>	<ul style="list-style-type: none"> <li>• It isn't free, even the trial mode lasts for just seven days and requires user to input their credit card details upon registration.</li> <li>• Users can only access its chatbot on the Boss mode or Higher sub-scription, which is quite expensive.</li> <li>• The information it generates doesn't always add up</li> </ul>
<b>Quora's Poe</b>	GPT-3,3.5 and 4	Deep learning, Reinforced Learning from Human Feed-back, Unsuper-vised Learning	<ul style="list-style-type: none"> <li>• Provides a single platform for ac-cessing multiple AI tools.</li> <li>• Allows users to generate content from multiple AI models and com-pare.</li> <li>• Users only create a single account, and they can access about five differ-ent chatbots</li> </ul>	<ul style="list-style-type: none"> <li>• Having multiple AI models in one platform sometimes makes it difficult to decide which to use or best.</li> <li>• There is still a subscription fee for accessing some of the AI models listed, like GPT-4, Claude-Instant-100k and Claude+</li> <li>• The interface for some of the AI models in Poe is not as good as that used to access the AI model directly, for instance, ChatGPT has a better interface than it does in Poe</li> </ul>

**Tabella 4-1 Comparazione Generative AI Tools**

Dall'analisi emergono diversi tool degni di nota che hanno tutto il potenziale di diventare, con il tempo, delle ottime alternative; alcuni di essi già mostrano risultati invidiabili in specifici campi di applicazione, ad esempio Google's Bard si mostra una scelta azzeccata per la scrittura

accademica. Tuttavia, al momento, ChatGPT si dimostra la scelta ideale nella scrittura creativa, grazie alle sue capacità di gestione di input di formato lungo e la sua elevata capacità di scrivere in linguaggio human-like.

Esistono, in realtà, studi analoghi per specifici casi applicativi, in particolare in ambito accademico, dove l'uso dell'IA sta creando maggiormente dibattito. Il documento *White Paper* [51] si occupa, ad esempio, di analizzare cinque strumenti di editing basati sull'intelligenza artificiale per migliorare la qualità linguistica di testi accademici: Paperpal, AJE Digital, Writefull, Grammarly e Instatext. Questo tipo di strumenti sono utili alla cura della grammatica, dello stile e della formattazione dei testi accademici, rendendoli pronti alla pubblicazione.

Ciononostante, il lavoro che ci si appresta a svolgere non ha l'obiettivo di testare la scrittura accademica e l'accuratezza grammaticale, ma vuole soprattutto testare le capacità creative dell'IA. Non stupisce che differenti studi siano stati fatti al riguardo, alcuni anche di risonanza elevata, e nella gran maggior parte dei casi il tool scelto come scrittore artificiale si rivela essere ChatGPT.

È interessante anche una tendenza che è stata individuata, gli studi che maggiormente hanno raccolto risultati soddisfacenti promuovono un cambio del paradigma con cui ci si affida alla scrittura generativa: essa deve essere un lavoro di continuo dialogo e negoziazione con l'IA, deve essere collaborativa con essa e non affidata ciecamente ai tool.

È stato già citato in precedenza il caso di Charlie Brooker che tentò di scrivere la sceneggiatura un episodio della serie *Black Mirror* con ChatGPT e ne rimase profondamente deluso, ma lo stesso lavoro è stato riproposto [52] seguendo la via della scrittura assistita, rivalutando la relazione tra autore e macchina nella fase creativa. Lo studio inserisce la figura dell'uomo come supervisore attivo del processo creativo dell'IA, prevedendo un indirizzamento del suo lavoro. Con questo cambio di ottica lo studio è riuscito nel suo intento, individuando in ChatGPT non solo uno strumento capace di scrittura creativa, ma anche utile per la figura dell'autore.

Un caso studio simile lo si trova nel paper *The use of ChatGPT in creative writing assistance* [53], in cui si svolge un'analisi accurata delle performance del tool nella scrittura di racconti di natura e genere diverso; un'applicazione, quindi, particolarmente analoga al caso studio di questa tesi. La sessione di scrittura viene strutturata attraverso un'unica chat, che attraverso una tattica di progressiva rifinitura parte da un'idea del tutto generale fino a raggiungere un output esaustivo e stratificato.

Di seguito i cinque punti di analisi dello studio:

- **Generazione di idee e concept:** partendo dal solo titolo del racconto, ChatGPT è riuscita a creare trame strutturate ed aderenti alla richiesta, seppur la mancanza di ulteriori vincoli renda le storie talvolta poco originali e ripetitive nelle strutture. Per sfruttare a pieno il potenziale risulta fruttuoso chiedere più opzioni tra cui scegliere ed integrare suggerimenti da parte dell'autore, riuscendo così ad ottenere output più originali.
- **Generazione di plot e personaggi:** individuata l'idea migliore, si procede per espanderla in un plot e caratterizzarne i personaggi. Questo avviene facendo domande, indagando profondamente la natura dei protagonisti e le loro scelte. ChatGPT riuscirà, così, a trasformare le storie in qualcosa di stratificato e complesso. In particolare, in una storia di stampo investigativo, è stato utile guidare il modello in un ragionamento a ritroso, che è partito dal finale per costruire il caso. Questo mostra come un ragionamento fuori dagli schemi può modificare radicalmente l'uscita.
- **Generazione di dialoghi:** in mancanza di linee guida precisa, i dialoghi assumono una forma stereotipale, ma una guida accurata e ricca di informazioni renderà sempre più

caratterizzate le battute, aggiungendo note di ironia, scetticismo, flirting, e allontanandosi da una scrittura semplicista.

- **Generazione di descrizioni ed atmosfera:** ChatGPT ha la tendenza a non seguire la filosofia del “Show don’t tell”, spesso esagera nelle descrizioni aggiungendo troppi dettagli ed allungando inutilmente l’elaborato. Sta all’autore valutare gli elementi e smistarli in base all’utilità. Inoltre, si registra una tendenza ad usare espressioni simili, ed occorre fare attenzione in fase di scrittura del prompt, perché alcune sue precise parole verranno usate all’interno della risposta, influenzandola.
- **Generazione di elementi stilistici:** l’accuratezza dello stile dipende essenzialmente dalle fasi precedenti, descrizioni, personaggi e trame dettagliate renderanno facile per il modello individuare uno stile coerente per la scrittura.

### 4.1.2 Scelta del tool di scrittura generativa

Si è analizzata l’offerta di tool di generazione di testo, scoprendolo ricco di concorrenti che si differenziano per architettura, addestramento, dataset e scopi. In questo ventaglio di opzioni, emerge un modello leader del mercato, che sia per popolarità che per efficienza si è imposto da subito come punto di riferimento dell’ambito: ChatGPT.

Quest’ultimo è uno strumento ampiamente studiato e testato su più campi, che vanta un’elevata capacità di scrittura creativa. Gli articoli individuati ne riportano buoni risultati in questa applicazione, inoltre la sua natura gratuita lo rende uno strumento alla portata di tutti.

Viste le premesse, la elevata accessibilità al grande pubblico ed i risultati incoraggianti in fase di test, ChatGPT è stata individuato come tool ideale per il caso applicativo di questa tesi, specialmente nell’ottica di utilizzare l’IA come tecnologia per il cinema indipendente, dove risorse economiche e temporali stringono e ogni strumento capace di velocizzare il processo e far risparmiare soldi è essenziale.

Esso verrà utilizzato in diversi utilizzi sperimentali, al fine di scoprirne le potenzialità maggiori:

- Revisione di sceneggiatura: valutazione del contenuto, revisione grammaticale e lessicale.
- Riscrittura sceneggiatura: verrà chiesto al modello di riscrivere le parti considerate deboli della sceneggiatura, apportando migliorie.
- Scrittura collaborativa di episodio originale: l’ultimo episodio verrà scritto in sinergia con ChatGPT, nell’ottica di scrittura assistita in collaborazione con l’IA.
- Caratterizzazione personaggi e ambienti: sfruttare il tool per indagare meglio gli elementi fondanti della sceneggiatura
- Assistenza pre-produzione: utilizzare il tool per consigli di stile di ripresa, montaggio, fotografia, individuazione di references e altri aspetti della pre-produzione di un film.
- Scrittura prompt per immagini

## 4.2 Tool di generazione di immagini

### 4.2.1 Tool di generazione di immagini: Analisi dell’offerta

Le generazione di immagini è una scienza più acerba rispetto alla scrittura generativa, il progresso è rapido e di anno in anno si vedono elevate migliorie in questo campo, per questo individuare il miglior strumento da utilizzare può essere difficile e non scontato.

## Scelta dei tool applicativi

L'articolo *Comparative Analysis of AI Image Generation Models* [54] offre un confronto dettagliato dei quattro modelli leader di generazione di immagini AI: **DALL-E 3**, **Google Imagen2**, **Stable Diffusion** e **Midjourney**. Per determinare quale modello sia migliore, viene eseguito uno studio comparativo in dieci categorie, al fine di scoprire per quali scopi sia più indicato ogni tool. Le categorie sono le seguenti:

- **Ritratti umani:** fondamentale per scoprire se il modello sia in grado di realizzare rappresentazioni realistiche di figure ed espressioni umane.
- **Animali:** rappresentare animali porta con sé difficoltà specifiche legate ai manti, le anatomie e i movimenti degli animali, rappresentarli è spesso arduo per i modelli IA.
- **Frutta:** utile a valutare il comportamento dei modelli di fronte a oggetti colorati con particolari risposte alla luce.
- **Paesaggi:** richiede di fondere assieme numerosi elementi, renderli coerenti tra loro e mantenere una prospettiva accurata.
- **Scene subacquee:** un ambiente complesso per le riflessioni e deformazioni dovute all'acqua.
- **Scene e paesaggi urbani:** molti diversi dagli ambienti naturali, ricchi di geometrie lineari e luce artificiale.
- **Veicoli:** utile a determinare l'accuratezza nella gestione del movimento di oggetti.
- **Momenti storici** : mette alla prova la capacità dell'IA di contestualizzare e di gestire dettagli e texture specifici del periodo.
- **Tecnologia e gadget:** testa la capacità dell'IA di ideare design moderni e futuristici.
- **Concetti astratti:** consentono di esaminare la creatività dell'intelligenza artificiale e la sua capacità di andare oltre il letterale, offrendo interpretazioni fantasiose di emozioni o idee.

Delle dieci categorie verranno ora analizzate nel dettaglio solamente quelle che più possono interessare il progetto di tesi, ossia: ritratti umani, paesaggi, scene e paesaggi urbani, momenti storici e concetti astratti. Per ogni categoria l'articolo analizza diversi fattori di valutazione (realismo, uso dei colori, artefatti, composizione), Nelle tabelle seguenti sono riportate le conclusioni complessive nei vari casi applicativi.

Piattaforma	Ritratti umani
DALL-E 3	Altamente consigliato per ritratti stilizzati e dettagliati
Google Imagen2	Consigliato per un equilibrio tra realismo e stile artistico
Midjourney	Consigliato per ritratti drammatici e surreali
Stable Diffusion	Altamente consigliato per ritratti realistici e spontanei

Tabella 4-2 Ritratti Umani

## Scelta dei tool applicativi

<b>Piattaforma</b>	<b>Paesaggi</b>
<b>DALL-E 3</b>	Altamente consigliato per paesaggi vibranti e spettacolari
<b>Google Imagen2</b>	Consigliato per scene atmosferiche e realistiche
<b>Stable Diffusion</b>	Altamente consigliato per paesaggi vividi e iperrealistici
<b>Midjourney</b>	Consigliato a chi apprezza un mix di arte e realismo

**Tabella 4-3 Paesaggi**

<b>Piattaforma</b>	<b>Scene e paesaggi urbani</b>
<b>DALL-E 3</b>	Altamente consigliato per scene notturne vivaci e dettagliate
<b>Google Imagen2</b>	Consigliato per ambienti urbani caldi e nostalgici
<b>Stable Diffusion</b>	Altamente consigliato per catturare la complessità e l'energia della vita urbana
<b>Midjourney</b>	Consigliato per opere d'arte urbana uniche e stilizzate

**Tabella 4-4 Scene e paesaggi urbani**

<b>Piattaforma</b>	<b>Momenti storici</b>
<b>DALL-E 3</b>	Consigliato per rappresentazioni tradizionali e solenni di eventi storici
<b>Google Imagen2</b>	Consigliato per narrazioni storiche intime e incentrate sui personaggi
<b>Stable Diffusion</b>	Altamente consigliato per rievocazioni storiche vivide e dinamiche
<b>Midjourney</b>	Consigliato per rivisitazioni artistiche e moderne della storia

**Tabella 4-5 Momenti storici**

## Scelta dei tool applicativi

Piattaforma	Concetti astratti
<b>DALL-E 3</b>	Altamente consigliato per interpretazioni edificanti e celestiali delle emozioni
<b>Google Imagen2</b>	Consigliato per rappresentazioni emozionali calde e intense
<b>Stable Diffusion</b>	Consigliato per visualizzazioni energiche e vibranti di concetti
<b>Midjourney</b>	Consigliato per rappresentazioni astratte complesse e profondamente stratificate

**Tabella 4-6 Concetti astratti**

Quanto esplorato nelle varie applicazioni è che, a differenza dei tool di scrittura esplorati ad inizio capitolo, è più difficile individuare un modello leader nella generazione di immagini. Quello che emerge è un panorama più diversificato, in cui ogni modello ha caratteristiche proprie e un proprio stile, talvolta ben riconoscibile. Non avendo un'offerta uniforme, diventa sbagliato definire il miglior modello di generazione di immagini, ma diventa più opportuno una scelta finalizzata agli obiettivi prefissati.

Ecco riportata in tabella un'analisi comparativa complessiva, in cui sono elencati i pregi e i difetti di ognuno dei modelli citati.

Piattaforma	Punti di forza	Punti deboli
<b>DALL-E 3</b>	<ul style="list-style-type: none"> <li>– Concetti fantasiosi e creativi.</li> <li>– Versatile per progetti astratti e artistici.</li> <li>– Ampia gamma di possibilità fantasiose.</li> </ul>	<ul style="list-style-type: none"> <li>– Realismo limitato.</li> <li>– Potrebbe non essere adatto a progetti che richiedono un elevato realismo.</li> </ul>
<b>Google Imagen2</b>	<ul style="list-style-type: none"> <li>– Elevato fotorealismo e attenzione ai dettagli.</li> <li>– Versatile per progetti realistici.</li> <li>– Adatto per progetti che richiedono immagini realistiche.</li> </ul>	<ul style="list-style-type: none"> <li>– Limitato nei concetti fantasiosi e astratti.</li> <li>– Meno adatto a progetti altamente artistici o emozionali.</li> </ul>

## Scelta dei tool applicativi

Piattaforma	Punti di forza	Punti deboli
<b>Stable Diffusion</b>	<ul style="list-style-type: none"> <li>– Equilibra il realismo con il tocco artistico.</li> <li>– Versatile per progetti e interpretazioni creative.</li> <li>– Offre una gamma di possibilità stilistiche.</li> </ul>	<ul style="list-style-type: none"> <li>– Potrebbe non eccellere nel realismo estremo o nei concetti astratti.</li> <li>– Le interpretazioni artistiche possono variare in coerenza.</li> </ul>
<b>Midjourney</b>	<ul style="list-style-type: none"> <li>– Si concentra sulla trasmissione di emozioni e stati d'animo.</li> <li>– Ideale per progetti che richiedono profondità emotiva ed espressione artistica.</li> <li>– Offre uno stile distinto e artistico.</li> </ul>	<ul style="list-style-type: none"> <li>– Meno adatto per progetti che richiedono un elevato fotorealismo.</li> <li>– Potrebbe non eccellere nelle immagini altamente dettagliate o tecniche.</li> </ul>

**Tabella 4-7 Punti di forza e di debolezza dei competitors**

Ampliando la ricerca [55] ad altri tool di generazione immagine diversi dai quattro leader esplorati, si apre un mondo ampio di opzioni in continua espansione. Abbiamo detto che i generatori di immagini si differenziano per una proprio stile caratterizzante, per cui ogni modello impone una propria cifra stilistica che può tanto sposare l'idea che abbiamo in mente quanto allontanarsene. È da questa caratteristica che nascono tool che si propongono per una specifica nicchia del mercato.

*Canva AI* [56] si presenta come ottima scelta per la creazioni di contenuti comunicativi, in particolare il suo dataset è addestrato sui social per cui è ideale per formati appartenenti a quel mondo. Adobe ha rilasciato il suo *Firefly* [57] per poter integrare facilmente l'IA all'interno della post-elaborazione tipica dei programmi del pacchetto Adobe. *NightCafe* [58] si presenta come uno strumento divertente e di facile utilizzo, ideale per chi vuole avvicinarsi al mondo delle immagini generative gradualmente. *AI Shutterstock* [59] permette l'uso commerciale di tutto il suo database rendendolo appetibile per i professionisti di marketing che usano l'IA per lavoro. *Leonardo AI* [60] e *Craiyon* [61] sono ideali per tutti coloro che cercano nell'IA uno strumento per le loro creazioni artistiche, offrendo una moltitudine di stili possibili, ma un limitato controllo del contenuto. Insomma, la proposta è ampia, ma oltre ai quattro big analizzati solo un tool è emerso potergli tener testa in termini di rapporto qualità/prezzo: *Image Creator* di Microsoft Designer. Esso permette risultati premium a costo zero, rendendolo un alleato potente in caso di budget ridotto.

Si ricorda che uno degli utilizzi previsti di questa tesi è la generazione di immagini per la realizzazione di uno storyboard. Questo compito porta con sé esigenze particolari in termini di coerenza tra le immagini, richiedendo che possano rendere ambienti e personaggi consistenti e riconoscibili. Allo stato attuale riuscire in questa impresa con uno qualsiasi dei generatori citati è pressoché impossibile, ogni nuova immagine generata presenterà inevitabilmente delle differenze dalle precedenti.

Fortunatamente, esiste un tool [62] che nasce proprio con lo scopo di generare storyboard con l'IA e favorire lo storytelling in maniera coerente: *Katalist*. Questo modello presenta molti strumenti utili alla realizzazione di storyboard e molte caratteristiche che non è possibile trovare altrove, per cui si tratta di un unicum interessante nel panorama.

### 4.2.2 Scelta dei tool di generazione di immagini

Abbiamo visto come ci siano molte scelte valide per la generazione di immagini, ma che in base allo scopo con cui si utilizzano possono rivelarsi più o meno indicate. Anche in questo caso occorre porsi dei limiti di tipo economico, anche in vista degli elevati costi degli abbonamenti dei tool di generazione di immagini, tipicamente maggiori di quelli per la scrittura.

Ricordiamo che, nel corso del caso studio, faremo uso di immagini generative per realizzare le moodboard del progetto e lo storyboard della sceneggiatura, per cui abbiamo previsto tool diversi per le due applicazioni.

È stato scelto di utilizzare **Katalist** come strumento per la generazione degli storyboard, in quanto sul mercato sembra essere l'unica proposta con i presupposti necessari a realizzare tavole coerenti le une con le altre e offre molti strumenti e opzioni innovative che meritano di essere analizzate per scoprirne la efficacia.

Per la generazione delle immagini di moodboard si evidenzia che non necessitano né di particolare accuratezza né di coerenza eccessiva, ma che hanno come unico vincolo quello di saper cogliere al meglio il messaggio di prompt per rappresentarne il mood ed il cuore. Anzi, nella loro generazione un'impronta artistica con una percentuale di imprevedibilità può essere un caratteristica stimolante e vincente. Per questo motivo, si è deciso di procedere attraverso un'applicazione comparativa, mettendo a paragone diversi tool generativi.

Gli strumenti scelti sono stati **DALL-E**, che ci permette di provare uno dei quattro leader del settore gratuitamente tramite ChatGPT, **Image Creator** di Microsoft Designer, in modo da testare un'alternativa gratuita meno popolare, e **Katalist**, di cui sfruttiamo l'abbonamento per ampliare lo studio e valutarlo in ambiti differenti.

## 5. Tool utilizzati: Analisi e funzionalità

In questo capitolo verranno presentati più nel dettaglio gli strumenti che sono stati utilizzati nel corso del progetto, evidenziandone le caratteristiche, funzionalità, interfacce e costi. Si tenterà anche di far luce sulle architetture alla base del loro funzionamento, per una comprensione più profondo di ciò che accade dietro le quinte della GenAI.

### 5.1 Nozioni di base di GenAI

Per rendere i restanti paragrafi più agevoli nella spiegazioni e alleggerirli da eccessive definizioni, ci si appresta di seguito a definire alcuni concetti cardine, in maniera molto approssimativa, degli elementi che caratterizzano l'intelligenza artificiale. [63]

La GenAI esiste, innanzitutto, solo grazie all'utilizzo di modelli di **Machine Learning** (ML) [64], ossia la tecnologia che consente ai computer di imitare le dinamiche dell'apprendimento umano. Questo è possibile innanzitutto con un accesso a grandi quantità di dati e permette alle macchine di lavorare in autonomia, prendere decisioni, ma soprattutto di correggersi man mano che acquisisce esperienza. Per svolgere questi compiti, gli algoritmi di machine learning necessitano di tre principali elementi [65]:

- Un processo decisionale: necessario per svolgere le funzioni di previsione o classificazione dei dati, il task di base con cui il ML svolge ogni sua mansione.
- Una funzione di errore: necessaria per valutare criticamente la previsione svolta, in modo da definirne l'attendibilità sulla base dei dati a cui ha accesso.
- Un processo di ottimizzazione del modello: il processo attraverso cui il modello ricalibra le proprie risposte, alla luce dei risultati della funzione di errore.

Nel mondo della GenAI molto più spesso si sente parlare, però, di **Deep Learning** [66], una sotto-categoria del ML che differisce per architettura e complessità. I modelli di Deep Learning utilizzano reti neurali multistrato, questo gli permette di apprendere e imitare l'apprendimento umano in maniera molto più accurata. Non solo, laddove i modelli di apprendimento supervisionato richiedono in input set di dati etichettati, i modelli deep learning sono in grado di elaborare dati grezzi, ossia privi di etichette, e di estrapolarne in autonomia le informazioni utili.

Le **Reti Neurali** [67], invece, non sono altro che la struttura portante di tutti i modelli d ML, sono gli algoritmi veri e propri che simulano i neuroni biologici nell'elaborazione dei dati. Ogni rete è strutturata in strati di nodi: uno di input, uno o più strati nascosti e uno di output. Il numero di strati determina la profondità del modello, se la rete possiede solo tre strati si definisce di base, se ne possiede più di tre si definisce profonda [68]. I modelli tipici di Deep learning prevedono anche centinaia o migliaia di strati nascosti.

Le reti neurali necessitano di una fase di apprendimento, durante la quale elabora set di dati mandati in input, che uno strato per volta attraversano i nodi della rete venendo elaborati secondo le funzioni contenute in ogni nodo. Al termine del processo, si ottiene un output finale, che viene confrontato con quello desiderato al fine di calcolarne l'errore. Quest'ultimo determina, quindi, la correzione da effettuare in retro-propagazione sui pesi che caratterizzano le connessioni tra i nodi. L'intero processo di andata e ritorno viene iterato numerose volte, fino al raggiungimento del risultato sperato. Una volta completato l'addestramento, la rete è in grado di elaborare e fare previsioni o classificazioni su nuovi dati di input.

L'IA generativa, quindi, sfrutta i modelli di ML o DL per generare output complessi. Per farlo, l'IA parte da un **foundational model**, un modello di DL generico che viene usato come base per la specializzazione specifica. Tipicamente, i foundational model di cui si fa maggiormente uso sono i **Large Language Model (LLM)**, i modelli capaci di comprendere e generare linguaggio umano. Successivamente, nella fase di ottimizzazione, si indirizza il modello verso una specifica task ed applicazione, specializzandolo attraverso l'uso dei dati. Si prosegue il processo con un'elaborata fase di messa a punto, valutazione e correzione.

Negli ultimi dodici anni, l'intelligenza artificiale generativa ha visto un'evoluzione nelle diverse architetture di deep learning utilizzate. Una prima architettura di rilievo la si trova negli **Autoencoder Variazionali (VAE)**, che permettono di comprimere e decomprimere dati, generando varianti del contenuto originale. A causa della loro natura, la loro capacità generativa è limitata rispetto alle architetture successive, per cui è in parte superata.

Un notevole passo avanti è stato fatto, invece, con le **Reti Generative Avversarie (GAN)**, la cui architettura si basa su due elementi: il generatore, che crea i nuovi dati, e il discriminatore, che deve riconoscere se i dati sono reali o generati artificialmente. I due elementi operano in competizione, con una dinamica iterativa in cui il generatore cerca di ingannare il discriminatore e quest'ultimo migliora la sua capacità di valutazione, consentendo così un progressivo miglioramento della qualità dei contenuti prodotti. Le GAN si sono rivelate particolarmente utili per la generazione di immagini e video realistici.

In contrapposizione alle GAN si hanno anche i **modelli di diffusione**, che applicano un processo di degradazione e ripristino progressivo dei dati, garantendo un maggiore controllo sulla qualità dell'output. Questa tecnica è alla base di alcuni degli strumenti più avanzati di generazione di immagini, tra cui DALL-E, per cui verrà trattata nel dettaglio in seguito.

Un'altra rilevante tecnologia che analizzeremo è quella dei modelli **transformer-based**, che si basano su un'architettura che ha ridefinito il paradigma dell'apprendimento automatico nel campo dell'elaborazione del linguaggio naturale (NLP). Questa tecnologia costituisce il nucleo dei principali foundation model, tra cui GPT-4, Copilot, BERT, Bard e Midjourney, dimostrando una notevole capacità nella generazione di testi complessi e nell'adattamento a compiti specifici. Nella sezione dedicata a ChatGPT avremo modo di studiare nel dettaglio il suo funzionamento.

## 5.2 ChatGPT

ChatGPT è il chatbot della società di ricerca statunitense **OpenAI** basato sull'Intelligenza Artificiale. Venne rilasciato sul mercato il 30 novembre 2022 e nei capitoli precedenti abbiamo visto come abbia contribuito a democraticizzare e rendere popolare l'uso dell'IA grazie ai suoi piani base gratuiti [69]. Il modello è in grado di elaborare e generare risposte testuali in linguaggio naturale.

Fin dal suo rilascio, ChatGPT ha attirato l'interesse di molti studiosi, arrivando più volte ad essere accostato al concetto di **AGI (Artificial General Intelligence)** o di **Intelligenza Artificiale Forte**, ossia il campo di studio teorico di agenti intelligenti capaci di apprendere e comprendere qualsiasi compito che un essere umano è in grado di imparare [70].

### 5.2.1 Storia e versioni

GPT è l'acronimo di **Generative Pretrained Transformer**, un modello linguistico transformer-based di cui si spiegherà a breve il funzionamento e che ha visto, nel corso di pochi anni, già diverse versioni ed aggiornamenti rilasciate [71]:

- **GPT-1, 2018:** primo modello di casa OpenAI, mostra già il potenziale futuro di ChatGPT fornendo livelli elevati, per l'epoca, di risposte in linguaggio naturale. Tuttavia, mostra già i problemi che diventeranno, presto, i principali nemici per gli sviluppatori di IA, ossia le allucinazioni e la scarsa dipendenza a lungo raggio, peggiorando le sue performance per risposte di formato lungo.
- **GPT-2, 2019:** aumentano i parametri del modello, arrivando a 1.5 miliardi con un parallelo miglioramento delle prestazioni e della coerenza e dipendenza a lungo raggio.
- **GPT-3, 2020:** la nuova versione supera enormemente il suo predecessore, con ben 175 miliardi di parametri si afferma da subito come uno dei più grandi LLM dell'epoca. Successivamente, verrà aggiornato nella versione GPT-3.5.
- **GPT-4, 2023:** attualmente la versione di GPT più potente di OpenAI, imponendo un nuovo standard in termini di qualità, accuratezza, e prevenzione delle allucinazioni. Il suo elevato costo energetico inizia a sollevare le questioni etiche-ambientali di cui si è discusso ad inizio tesi.
- **GPT-4 Turbo:** l'attuale aggiornamento di GPT-4, ha accesso a dati di addestramento fino ad aprile 2023, lasciando – perciò – scoperti tutti i dati successivi alla data.
- **GPT-4o, 2024:** rilasciato a maggio 2024, implementa un funzionamento multilingue, supportando dati in lingue diverse dall'inglese. Tra le novità principali del modello figura il funzionamento multimodale, che lo rende capace di elaborare input immagini, audio e video, e di fornire in output sia testo che immagini e contenuti audio. Secondo OpenAI, GPT-4o è il 50% energeticamente più economico e due volte più veloce nella generazione di testo rispetto a GPT-4 Turbo [72].

Oltre a ChatGPT, OpenAI offre altri prodotti di Intelligenza Artificiale Generativa. Nel 2022 è stato lanciato anche DALL-E 2, il tool di immagini generative di cui parleremo ulteriormente nel corso del capitolo. Poco dopo è arrivata anche l'App di ChatGPT per Smartphone, che permette l'interazione vocale. Infine, nel 2024, è stato rilasciato anche Sora, un modello Text-to-video per la generazione di video.

### 5.2.2 Caratteristiche e funzionamento

Per la scrittura di questa sezione, in cui ci si propone di dare una panoramica esaustiva del funzionamento dei modelli GPT, ci si è affidati principalmente alle fonti *Che cos'è un GPT (trasformatore generativo pre-addestrato)?* [71] fornita da IBM e *ChatGPT, cos'è e come funziona: limiti e opportunità* [73] che ha visto la collaborazione dell'Osservatorio di Artificial Intelligence del Politecnico di Milano.

I modelli GPT basano il loro funzionamento su **algoritmi probabilistici**, utilizzano una matematica complessa per analizzare una sequenza di input e determinare l'output più probabile. Questo non avviene, però, in una formulazione complessiva unica, ma scegliendo sequenzialmente la parola con maggiore probabilità di accuratezza per proseguire la frase.

Per farlo, ChatGPT si basa su dinamiche di **Machine Learning** e **Deep Learning**, che sfruttano una rete neurale stratificata per il calcolo delle variabili da restituire. Il modello utilizza algoritmi di **Natural Language Processing (NLP)** per interpretare gli input che gli vengono inviati e formula risposte pertinenti grazie ai propri dati di addestramento che contano miliardi

di fonti. È proprio la vastità dei dataset a cui ha accesso che rende ChatGPT capace di simulare la capacità di comprensione umana.

Grazie al Machine Learning e il Deep Learning, ChatGPT è anche in grado di apprendere autonomamente dall'esperienza, integrando le stesse interazioni degli utenti con il chatbot nei dati di addestramento del modello. Questo metodo di addestramento prende il nome di **Adaptive AI**.

Abbiamo detto, ad inizio capitolo 5.2.1, che l'acronimo GPT sta per Generative Pre-trained Transformer, ma non è stato spiegato né cosa siano i transformer né perché ChatGPT sia un modello transformer-based. ChatGPT deve la sua capacità di interagire efficacemente con gli utenti all'utilizzo di un meccanismo di "attenzione" basato, appunto, sui **transformer** [74]. I modelli GPT, infatti, basano il loro successo a due fattori essenziali: il **Pre-addestramento Generativo** e l'**Architettura Trasformativa**.

Il pre-addestramento generativo è il processo di addestramento di un modello in LLM su dati non etichettati, questo significa che ChatGPT è addestrato a riconoscere in modo autonomo i dati a cui ha accesso e di trarne delle previsioni accurate. In questa fase di addestramento, al modello vengono sottoposti dati non etichettati, portandolo a comprendere in maniera completamente indipendente le loro connessioni e significati. Agendo così, trasversalmente, il modello acquisisce la capacità intrinseca di saper rispondere in maniera pertinente a nuovi input esterni, nella fattispecie ai prompt mandati dagli utenti.

Un'altra particolarità dei modelli GPT è che non elaborano il linguaggio umano come lo facciamo noi, essi dividono il discorso in unità discrete chiamate **token**, con la possibilità di suddividere anche semplici parole in più token. Questo processo serve a parallelizzare il processo di valutazione e pesatura dei token, utilizzando i trasformatori per stabilire le cosiddette **relazioni a lungo raggio**, ossia le relazioni che intercorrono tra token distanti. Questo aspetto è fondamentale per comprendere la potenza di ChatGPT, perché è grazie a questa capacità che non si limita alla sola lettura dei testi, ma ne inquadra il contesto e diventa in grado di formulare o comprendere testi di lungo formato.

I modelli di trasformatori elaborano i dati con due moduli noti come codificatori e decodificatori, utilizzando meccanismi di self-attention per stabilire dipendenze e relazioni.

La tecnica di **self-attention** è il processo che consente di parallelizzare l'elaborazione di un'intera sequenza di input, assegnando pesi differenti alle parti che compongono il prompt. Contrariamente alle reti neurali ricorrenti (RNN [75]) e alle reti neurali convoluzionali (CNN [76]), che leggevano i dati di input sequenzialmente o gerarchicamente, i trasformatori permettono di indirizzare la loro attenzione ai token considerati più importanti, in maniera indipendente dalla loro posizione. Questo porta i modelli GPT ad elaborare il contesto della richiesta invece di indovinare semplicemente la parola successiva nella frase, creando la coerenza a lungo raggio che li contraddistingue.

La **codifica** è il processo attraverso cui avviene la mappatura dei token su uno spazio vettoriale virtuale tridimensionale. Il principio di codifica con cui vengono posizionati i token prevede che i token semanticamente simili siano mappati vicini nello spazio 3D. Questo processo di vettorizzazione dei dati di input è nota come **embedding** [77] e permette di concatenare le parole con una semantica coerente. I blocchi encoder nella rete di trasformatori assegnano a ciascun incorporamento un peso, che ne determina l'importanza relativa. Parallelamente, i codificatori di posizione acquisiscono la semantica determinando l'ordine dei raggruppamenti di parole.

I **decodificatori** si occupano, infine, del calcolo della risposta statisticamente più probabile agli embedding preparati dai codificatori. I meccanismi di auto-attenzione consentono al

decodificatore di identificare le parti più importanti della sequenza di input, mentre gli algoritmi avanzati determinano l'output che ha maggiori probabilità di essere corretto.

### 5.2.3 Analisi delle performance

Si è già parlato nel capitolo 4.1.1 delle prestazioni dei tool di scrittura generativa riportate da differenti studi, ma volendosi concentrare esclusivamente su ChatGPT il documento *GTP4 a technical report* [78] fornisce un'analisi estremamente accurata delle sue performance in diverse categorie di applicazione e sotto particolari vincoli.

Tra i dati particolarmente interessanti che si è voluto riportare, si mostrano nella tabella 7-1 sottostante le performance del modello GPT-4 nei test su benchmarks accademici, che comprendono domande a risposta multipla, ragionamenti di buon senso, comprensione del testo, problemi aritmetici e molte altre categorie. Il modello viene paragonato ai modelli GPT-3.5 e lo stato dell'arte (SOTA) del LLM. Si fa notare anche che i prompt utilizzati nei test usano la tecnica del few-shot prompting o chain-of-thought prompting, entrambe spiegate nel dettaglio nel capitolo 6.3.2. Dalle percentuali riportate si evince un'elevatissima accuratezza delle risposte in più campi, quasi sempre maggiore rispetto allo SOTA e con un picco del 96.3% nelle domande a risposta multipla scientifiche ed un minimo del 67% nei task di scrittura codice in Python.

Un aspetto finora poco trattato è la questione linguistica, infatti i modelli sono addestrati su dataset che posso contenere diverse proporzioni di fonti nelle diverse lingue e possono essere più o meno efficienti nella interpretazione di prompt in base alla lingua. Nel grafico in immagine 7-1 è riportata una comparazione delle performance di GPT-4 nell'utilizzo di diverse lingue sul dataset MMLU. Come si può notare, l'italiano – che è la lingua in cui è stato utilizzato ChatGPT durante il progetto in tesi – presenta statistiche ottime posizionandosi al secondo posto come lingua per accuratezza.

	GPT-4 Evaluated few-shot	GPT-3.5 Evaluated few-shot	LM SOTA Best external LM evaluated few-shot	SOTA Best external model (incl. benchmark-specific tuning)
<b>MMLU</b> [49] Multiple-choice questions in 57 subjects (professional & academic)	<b>86.4%</b> 5-shot	70.0% 5-shot	70.7% 5-shot U-PaLM [50]	75.2% 5-shot Flan-PaLM [51]
<b>HellaSwag</b> [52] Commonsense reasoning around everyday events	<b>95.3%</b> 10-shot	85.5% 10-shot	84.2% LLaMA (validation set) [28]	85.6 ALUM [53]
<b>AI2 Reasoning Challenge (ARC)</b> [54] Grade-school multiple choice science questions. Challenge-set.	<b>96.3%</b> 25-shot	85.2% 25-shot	85.2% 8-shot PaLM [55]	86.5% ST-MOE [18]
<b>WinoGrande</b> [56] Commonsense reasoning around pronoun resolution	<b>87.5%</b> 5-shot	81.6% 5-shot	85.1% 5-shot PaLM [3]	85.1% 5-shot PaLM [3]
<b>HumanEval</b> [43] Python coding tasks	<b>67.0%</b> 0-shot	48.1% 0-shot	26.2% 0-shot PaLM [3]	65.8% CodeT + GPT-3.5 [57]
<b>DROP</b> [58] (F1 score) Reading comprehension & arithmetic.	80.9 3-shot	64.1 3-shot	70.8 1-shot PaLM [3]	<b>88.4</b> QDGAT [59]
<b>GSM-8K</b> [60] Grade-school mathematics questions	<b>92.0%*</b> 5-shot chain-of-thought	57.1% 5-shot	58.8% 8-shot Minerva [61]	87.3% Chinchilla + SFT+ORM-RL, ORM reranking [62]

Tabella 4-8 Prestazioni di ChatGPT su benchmark accademici

## Scelta dei tool applicativi

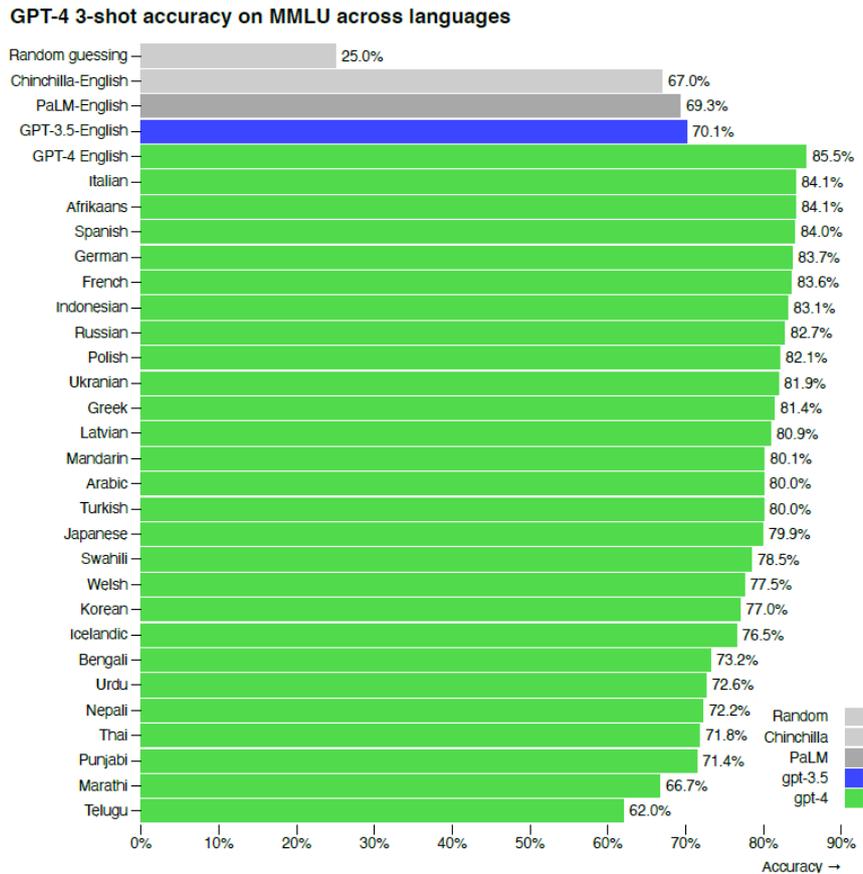


Figura 4-1 Comparazione per lingue delle performance di GPT 4

### 5.2.4 Piani di abbonamento

ChatGPT è popolare anche grazie al fatto che prevede un ricco piano gratuito che offre funzioni già avanzate, per questo motivo nella tesi si è fatto uso della sua versione free, tuttavia, presenta due ulteriori piani mensili che vale la pena valutare: Plus e Pro.

	<b>Free</b>	<b>Plus</b>	<b>Pro</b>
<b>Price / month</b>	0\$	20\$	200\$
<b>Modello GPT</b>	GPT-4o mini illimitato GPT-4 limitato	GPT-4o mini illimitato GPT-4 limitato Modelli GPT precedenti	Accesso illimitato a tutti i modelli GPT
<b>Modalità vocale</b>	Standard	Avanzata (limitato)	Avanzata (illimitato)
<b>Caricamento file</b>	Limitato	Esteso	Illimitato
<b>Generazione immagini</b>	Limitato	Esteso	Illimitato

## Scelta dei tool applicativi

<b>Generazione video</b>	X	Limitato	Esteso
<b>Deep Research</b>	X	X	Si

Tabella 4-9 ChatGPT: Piani di abbonamento

Alla luce dei piani esplorati, si è scelto, ai fini della tesi, di sfruttare il piano Free, testando le potenzialità di uno strumento come ChatGPT in un contesto completamente gratuito.

### 5.2.5 Interfaccia e funzionalità

ChatGPT è un punto di riferimento per i propri competitor anche grazie alla sua interfaccia, estremamente chiara e pulita. Ogni nuova conversazione si apre con la seguente schermata, con alcuni bottoni per indirizzare o consigliare specifici utilizzi.

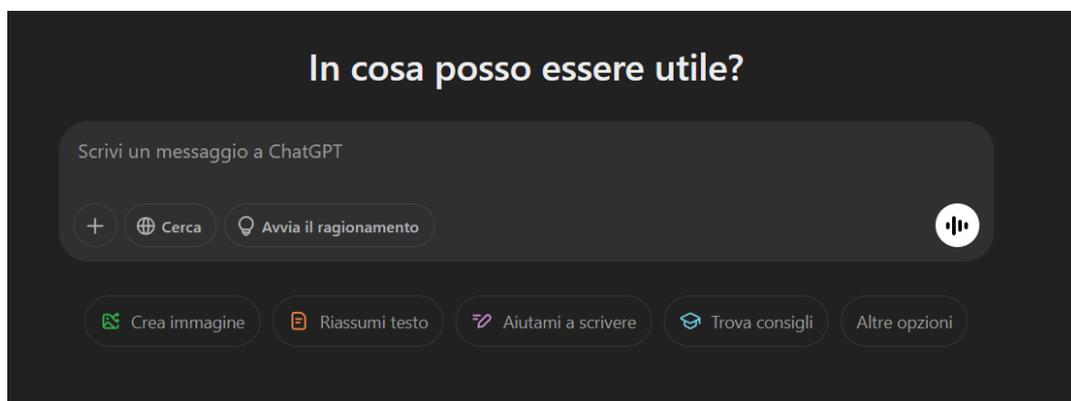


Figura 4-2 ChatGPT: Interfaccia di inizio chat

Con il bottone “+” è possibile **allegare** file al prompt proveniente dal proprio computer, OneDrive o GoogleDrive. Allegare file è una funzionalità particolarmente utile e che non tutti i modelli supportano, presente nel piano Free con un numero limitato di upload.

Con il bottone “**Cerca**” è possibile effettuare una ricerca online, analogamente a quanto si fa con i motori di ricerca, ma ricevendo in uscita la completezza tipica dei modelli di IA. Questa funzione solleva però questioni etiche legate al consumo energetico che creano al momento molto dibattito.

Il tasto “**Avvia ragionamento**” è una recente aggiunta molto interessante, esso permette di analizzare e strutturare il problema prima di formulare una risposta, aiutando a garantire che la risposta sia accurata e ben articolata. All’atto pratico integra dentro di sé alcune delle tecniche di prompting viste nel capitolo 6.3.2 portando il modello a ragionare in modo profondo e per più tempo prima di rispondere, mostrando, nel mentre, anche i suoi ragionamenti.

L’icona a destra invece permette di intraprendere una **conversazione vocale** ad alto livello con il modello, una funzione all’avanguardia e limitata nella versione free.

I bottoni sottostanti, invece, non hanno vere e proprie funzioni, ma sono utili a consigliare alcuni percorsi di conversazioni e a mostrare tutte le possibili applicazioni del ChatGPT. Cliccandoci sopra, si intraderà una conversazione parzialmente guidata per raggiungere i diversi obiettivi.

## 5.3 DALL-E

DALL-E è il modello di generazione text-to-image rilasciato nel 2021 da OpenAI che, fin dal suo lancio, ha imposto nuovi standard per la GenAI. Il suo nome deriva dalla fusione dell'artista Salvador Dalí e di WALL-E, il robot pulitore di casa Pixar, con lo scopo di impostare da subito una propria identità di stampo creativo e artistico. Quel che ha rivoluzionato il mercato, è stata la sua capacità di diminuire il divario che esiste tra descrizioni testuali e sintesi di immagini, dando la possibilità di descrivere e generare immagini sempre più complesse e varie grazie all'architettura a transformer. [79]

### 5.3.1 Storia e versioni

OpenAI ha prodotto tre versioni di DALL-E, migliorandone sempre le prestazioni e con un aggiornamento parallelo a quello di GPT. La sua genesi nasce a partire a giugno 2020, anno in cui OpenAI rilascia **ImageGPT**, un modello che parte dall'intuizione di sfruttare l'architettura dei transformer anche per la generazione di immagini. In particolare, ImageGPT utilizza lo stesso codice di GPT-2 per generare immagini a partire da descrizioni testuali. La sua forma è ancora limitata, ma ha gettato le basi necessarie ad una nuova frontiera della GenAI. Infatti, la sua innovazione è stata di riprendere un meccanismo tipico della generazione testuale: così come era possibile creare testi coerenti con algoritmi di previsione per la scelta della parola successiva da inserire all'interno di una frase, era possibile prevedere come colorare il pixel successivo di un'immagine in modo del tutto analogo. [80]

#### *DALL-E*

Seguendo il successo di ImageGPT, OpenAI lancia nel mercato la prima versione di DALL-E nel gennaio del 2021. Essa viene presentata come la rete neurale da 12 miliardi di parametri che si appoggia a GPT-3 per la generazione di immagini a partire dal testo. Con GPT-3, OpenAI è riuscita a raggiungere nuovi standard di addestramento della rete neurale, riuscendo a ampliare la varietà di attività di generazione di testo, e si propone ora di sfruttarlo per ampliare il bacino di immagini generabili. Questo è possibile grazie ad innovative tecniche di embedding, che si appoggiano alla tecnologia CLIP che vedremo nella sezione 5.3.2. In questa versione, DALL-E ha una risoluzione di immagini massima di 1024 x 1024 pixel e la comprensione delle richieste testuali sono limitate, ma restano significativi progressi per l'epoca. [81]

#### *DALL-E 2*

A solo un anno dal rilascio della prima versione, OpenAI rilascia DALL-E 2, lo stesso anno del lancio di ChatGPT. Il suo accesso è stato dapprima testato su un numero limitato di utenti, e poi aperto pubblicamente ad una versione beta dopo essersi assicurati della sua sicurezza. Un aspetto rilevante, infatti, della seconda versione è il fatto che OpenAI abbia iniziato ad affrontare le questioni etiche più rilevanti della generazione di immagini, adoperandosi per inserire filtri per abolire la generazione di contenuti inappropriati.

Le migliorie tecniche sono consistenti, la risoluzione è quadruplicata, raggiungendo finalmente standard paragonabili al foto-realismo, i tempi di attesa sono calati drasticamente, arrivando all'ordine di grandezza di decine di secondi. In termini di comprensione dei prompt si registrano enormi passi avanti, rendendo possibile un sempre maggiore miscelamento di stili, dando la possibilità di dare sfumature complesse all'immagine in termini di colori, luce e composizione. [82]

### ***DALL-E 3***

L'ultima versione del modello è nata da una fruttuosa collaborazione con Microsoft ed è stata rilasciata nel settembre 2023, con la principale innovazione di essere nativa su ChatGPT. Questo comporta una notevole implicazione pratica nel suo utilizzo: tutto il delicato processo di scrittura del prompt, che richiedeva una certa conoscenza della sintassi e delle tecniche dedicate, ora viene gestita interamente da ChatGPT. Basta fornire descrizioni dettagliate in linguaggio naturale, e il modello le rielaborerà trasformandole in un prompt adeguato, aggiungendo anche un proprio tocco creativo o interpretativo. Questo aspetto è essenziale per comprendere l'innovazione apportata, perché rende la generazione di immagini alla portata di tutti. Questa procedura è una ragionevole conseguenza alla sempre più comune pratica di affidarsi a ChatGPT per la scrittura dei prompt, rendendo implicito quello che fino a poco tempo fa avveniva in maniera esplicita.

Anche in termini tecnici ci sono migliorie, appoggiandosi a GPT-4 la comprensione di testo complesso ha raggiunto livelli ancora più alti, rendendo possibile la generazioni di immagini sempre più articolate, anche con composizioni frammentate e diversi piani di lettura.

Prosegue, parallelamente, una politica etica del controllo e della limitazione degli abusi dell'IA, con un particolare occhio alla trasparenza delle immagini generative. OpenAI sta infatti investendo risorse per creare un classificatore di provenienza, ossia una tecnologia capace di determinare se un'immagine sia reale o generata artificialmente. [83] [84]

### **5.3.2 Caratteristiche e funzionamento**

DALL-E è un prodotto di punta del mercato della GenAI ed è attualmente un punto di riferimento per la generazione di immagini. Gli studi analizzati e su cui è basata la scrittura della seguente sezione [85] [86] [79] [87] [88] sono dedicati principalmente allo studio di DALL-E 2, sulla terza versione del modello è stato pubblicato ancora poco materiale e non abbastanza approfondito. Tuttavia, analizzare la versione che ha reso il modello maggiormente popolare è già significativo nell'individuare le sue caratteristiche più determinanti.

DALL-E è un modello multimodale, ossia un modello che nel suo funzionamento deve saper gestire dati di natura e tipologia diversa, nella fattispecie dati testuali ed immagini. Lavorare con le immagini porta con sé delle sfide intrinseche, innanzitutto l'utilizzo dei pixel come token di lavorazione occupa molta memoria, inoltre occorre saper gestire accuratamente la dipendenza a lungo raggio interna all'immagine.

Per alleggerire la gestione del carico, DALL-E utilizza **Discrete Variational Autoencoder** (dVAE). Il dVAE è una variante del tradizionale Variational Autoencoder (VAE) [89], ossia i codificatori che permettono di trasformare dati complessi in rappresentazioni numeriche semplificate, tipicamente vettori di variabili latenti. Il dVAE si discosta da essi operando in uno spazio latente discreto. Questa architettura è particolarmente adatta per scenari in cui i dati possiedono attributi categoriali o simbolici, un caso particolarmente aderente alla generazione di immagini, il cui attuale funzionamento è basato sull'idea di etichette assegnate alle immagini.

Un'altra sfida tipica dei modelli multimodali è trovare il modo per convertire o far dialogare le diverse tipologie di dati. Nel caso di DALL-E, ciò significa trovare il modo di rendere codifiche di dati testuali e di immagini connesse e dialoganti.

Per superare tutti questi ostacoli, l'architettura di DALL-E, a partire dalla sua seconda versione, si può suddividere in tre parti fondamentali:

## Scelta dei tool applicativi

- Codificatore di testo: Trasforma il testo di descrizione in una rappresentazione numerica, il text embedding.
- Prior: A partire dal text embedding, converte il vettore generando una image embedding, una rappresentazione numerica dell'immagine da generare.
- Decoder: Genera l'immagine finale a partire dall'immagine embedding, applicando un processo di denoising basato sulla tecnica di Diffusione.

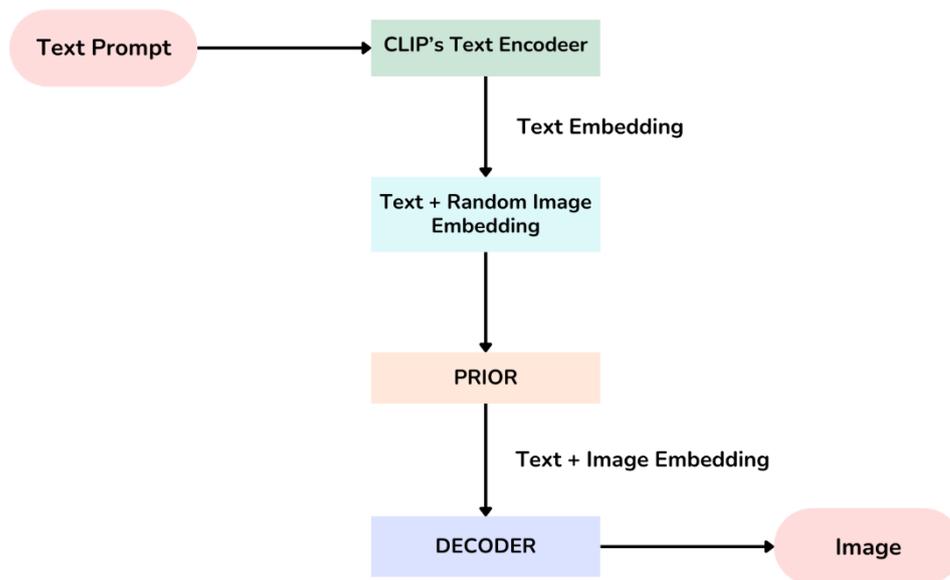


Figura 4-3 Architettura DALL-E 2

Il concetto di embedding di testo è il primo passo necessario alla generazione di un'immagine a partire da un prompt testuale. Le principali differenze che determinano un migliore NLP (ad esempio GPT) dipendono principalmente dalla generazione di rappresentazioni numeriche significative. DALL-E si è distinto fin dalla sua nascita per le sue capacità di embedding di testo, per cui è utile sapere come opera.

Quando immetti un prompt nella chat con DALL-E, esso viene elaborato dal modello **CLIP** (Contrastive Language-Image Pre-Training) un modello che sfrutta l'apprendimento contrastivo per addestrare due modelli di reti neurali, uno per la comprensione del testo e l'altro per il riconoscimento di immagini. L'apprendimento contrastivo permette una rappresentazione numerica (tramite vettori) a bassa dimensione di una particolare entità, testuale o immagine, mediante il contrasto tra oggetti simili e diversi. Questo ha come obiettivo quello di assegnare vettori, ossia rappresentazioni, vicini per oggetti simili e distanti per oggetti profondamente sconnessi. In definitiva, lo scopo di CLIP, è di diminuire il divario che sussiste tra le immagini e le descrizioni testuali annesse e in questa fase si occupa di trasformare il prompt in text embedding semanticamente accurati.

Il primo criterio di successo di CLIP è l'accesso ad un ampio dataset di coppie immagini-testo. In fase di addestramento, i dati vengono convertiti in text embedding e in image embedding utilizzando due differenti architetture transformer-based di codifica. Dopodiché, CLIP si occupa di accoppiare efficacemente ogni testo con ogni immagine, cercando la giusta corrispondenza per minimizzare le distanze tra elementi corrispondenti e massimizzare quelle tra elementi discordanti. La metrica utilizzata da CLIP per determinare il rapporto tra distanza e similarità prende il nome di **Cosine-Similarity**.

Un punto distintivo e di successo di DALL-E è il fatto che le architetture alla base della fase di embedding siano entrambe transformer-based, ossia sfruttano il già citato meccanismo dell'attenzione, che permette una valutazione pesata degli elementi suddivisi in sezioni, utile a una comprensione del contesto più accurata.

Ripercorrendo il flusso lavorativo, la descrizione testuale che abbiamo fornito è stata passata al codificatore di testo di CLIP, che tramite NLP è stato compreso per essere poi convertito in un vettore che ne rappresenta il significato semantico. A questo punto, il passo successivo è trasformare il text embedding in image embedding, fase in cui entra in gioco il modello **Prior**. Esso si deve occupare trasformare una rappresentazione numerica semantica in una rappresentazione numerica visiva dell'immagine desiderata. Un possibile approccio è quello di utilizzare la tecnica **auto-regressiva**, che tratta il problema come fosse una traduzione Seq2Seq dal text embedding (prima sequenza) to image embedding (seconda sequenza). Però, DALL-E utilizza una tecnica più efficace conosciuta come **Diffusion**.

In termini generali, il concetto di Diffusion parte dalla ricezione di un'immagine di input, ad ogni intervallo di tempo  $x$  viene aggiunto del rumore all'immagine, applicando la cosiddetta Forward Diffusion, ottenendo istante  $i$  (multiplo di  $x$ ) un'immagine  $Imm(i)$  più deteriorata rispetto a  $Imm(i-1)$ . Dopo un certo numero di iterazioni, si arriverebbe ad ottenere un'immagine finale costituita esclusivamente da rumore casuale. Raggiunto questo punto, è possibile invertire l'intero processo, andando a ridurre la quantità di rumore ad ogni intervallo  $x$ , con l'obiettivo di recuperare man mano l'immagine originale.

In fase di addestramento il modello Prior utilizza la forward diffusion per imparare come le immagini provenienti dal dataset si degradano, per poter poi invertire il processo in fase di generazione ed ottenere immagini realistiche in output.

Il segmento Prior riceve quindi due dati in fase di generazione: il text embedding corrispondente al prompt e un image embedding casuale. Il modello comincerà allora ad applicare iterativamente del denoising progressivo all'immagine, ossia il ciclo inverso del metodo Diffusion, fino a raggiungere l'immagine embedding finale che meglio si accosta al text embedding di input.

L'ultimo passo dell'implementazione prevede l'uso del **decoder** per la generazione dell'immagine di output. Esso, partendo dall'immagine embedding di output del modello di Prior, applicherà lo stesso processo di denoising iterativo ad un'immagine di rumore casuale, fino a raggiungere la miglior rappresentazione dell'embedding.

Riassumendo, quindi, i passaggi essenziali dell'architettura DALL-E abbiamo:

1. Inviamo un prompt testuale al chatbot
2. Il prompt viene processato da CLIP per la generazione di embedding di testo tramite apprendimento contrastivo.
3. Il text embedding viene concatenato ad un image embedding casuale e dato in pasto al modello Prior. Si ottiene così l'immagine embedding corrispondente al testo iniziale, attraverso l'applicazione del processo di diffusione inverso.
4. Il decoder acquisisce il text embedding ed image embedding generato nel passaggio precedente e genera l'immagine finale.

### 5.3.3 Analisi delle performance

DALL-E è caratterizzato da elevate performance sia in termini di accuratezza sia in termini di efficienza.

## Scelta dei tool applicativi

Si è precedentemente parlato della grande ostacolo della generazione di immagini di dovere gestire ingenti quantità di memoria, DALL-E usa degli stratagemmi utili alla sua riduzione. Innanzitutto, utilizza un formato di dati numerici da 16 bit a virgola mobile [88] che permette una gestione agile dei numerosi calcoli necessari alla generazione. Ciò può causare problemi in caso di numeri molto piccoli, ma è un'eventualità che viene scongiurata con diverse tecniche.

L'architettura del trasformatore è costruita per essere in grado di comprendere relazioni profonde all'interno dei dati immessi, prevedendo 12 miliardi di parametri utili all'interpretazione di essi (solo nella prima versione) [90]. DALL-E deve le sue invidiabili capacità generative a questo e ad un altro importante discriminante: il dataset. Il modello ha accesso al database immagini di Wikipedia e del dataset YFCC100M++, che da solo conta più di 100 milioni contenuti immagine e video [91].

Nella Figura 4-4 è riportato un grafico proveniente dall'articolo *How OpenAI's DALL-E works?* [79], raffigurante i risultati su due metriche differenti di DALL-E a confronto con altre architetture GAN, in particolare Attn-GAN, DM-GAN e DF-GAN. Le metriche in considerazione sono la Fréchet inception distance (FID) e la Inception Score (IS), due metriche specifiche per la valutazione della qualità delle immagini generative [92].

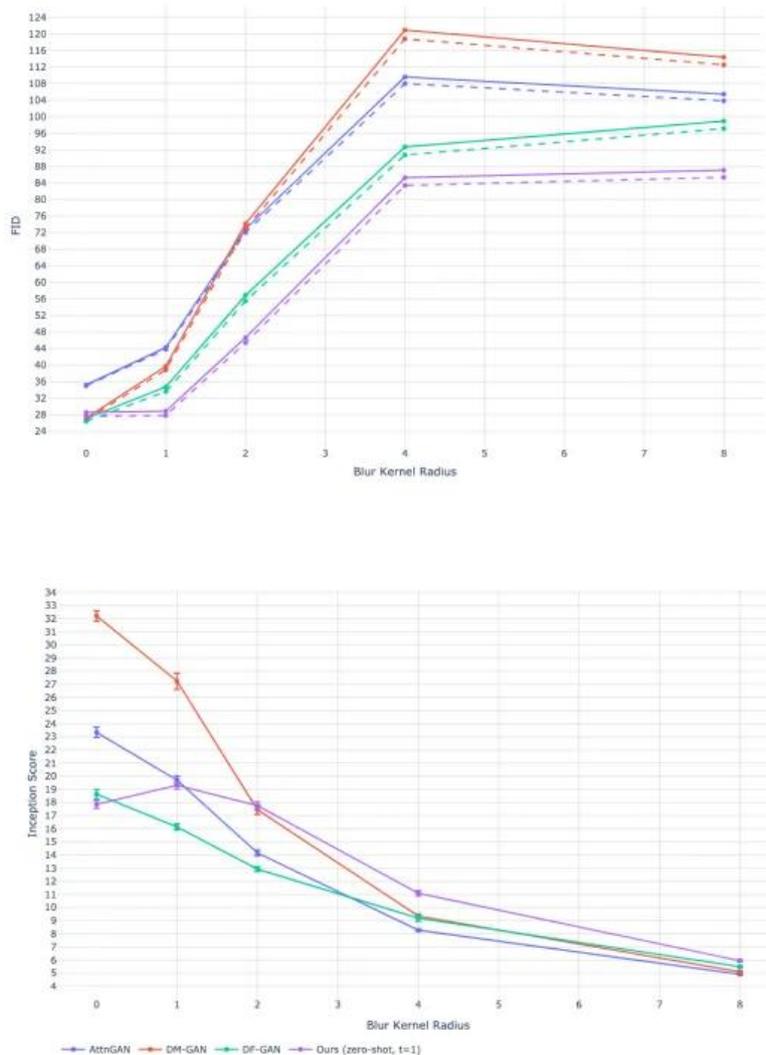


Figura 4-4 FID e Inception Score di DALL-E

### 5.3.4 Piani di abbonamento

DALL-E, nella sua versione attuale, è integrato all'interno di ChatGPT, questo significa che non è possibile accedere al modello in maniera indipendente e che qualsiasi piano di abbonamento per il suo utilizzo dipende esclusivamente dal piano scelto per ChatGPT. Per cui, riportiamo l'attenzione sui piani esplorati alla sezione 5.2.4.

Per la creazione delle moodboard si è sfruttato il piano free di ChatGPT, questo ha imposto un limite di tre immagini generate al giorno; inoltre, essendo DALL-E 3 appoggiato a GPT-4, qualora si fosse raggiunto il limite massimo di messaggi con ChatGPT nella sua quarta versione non sarebbe più stato possibile generare immagini nuove fino al termine del tempo di ricarica.

Questo ha inevitabilmente allungato i tempi dello studio, ma ci ha permesso di valutarne l'efficienza in termini di strumento gratuito alla portata di tutti. Questo è stato possibile solo perché il numero di moodboard era sufficientemente ridotto.

### 5.3.5 Interfaccia e funzionalità

L'interfaccia di DALL-E non è nient'altro che la stessa del chatbot di OpenAI, per il suo utilizzo è sufficiente dialogare con il modello in linguaggio naturale e riceverai in risposta un'immagine qualora la chiedessi. È importante precisare che DALL-E fornisce una sola immagine di output, e non 4 varianti come fatto da altri tool. Questo impone un processo iterativo di correzione potenzialmente più lungo del consueto.

## 5.4 Image Creator - Microsoft Designer

Abbiamo già parlato, nel corso del capitolo, del fatto che Microsoft abbia collaborato con OpenAI nella creazione dei suoi modelli più avanzati. Il rapporto tra il colosso della Silicon Valley e la società no-profit è però sempre più stretto, testimoniando sia un forte investimento da parte di Microsoft nel mondo dell'IA sia quanto OpenAI abbia un peso enorme nel mercato digitale [93].

Image Creator, il tool di generazione di contenuti visivi di Microsoft, deve il suo successo proprio a questa collaborazione, esso utilizza infatti il modello DALL-E 3 [94], il più aggiornato di casa OpenAI, per generare le proprie immagini. Si rimanda, quindi, alla sezione 5.3.2 per esplorare l'architettura ed il funzionamento di DALL-E e dedichiamo i prossimi paragrafi all'esplorazione della proposta di Image Creator. Si segnala, tuttavia, delle piccole differenze nelle policy di restrizione dei contenuti sessuali e violenti tra i due strumenti. Si è parlato di come OpenAI abbia imposto dei filtri molto restrittivi al riguardo, ma Microsoft ha applicato delle modifiche al sistema di filtraggio, con lo scopo di diversificare maggiormente le generazioni, pur rimanendo allineato alla linea etica di OpenAI.

### 5.4.1 Storia e versioni

Il lancio di Microsoft Designer è stato ufficialmente annunciato il 12 ottobre 2022 [95]. Occorre specificare che Image Creator è solo una delle molte parti che compongono Microsoft Designer, un software che nasce dichiaratamente per la creazione grafica, posizionandosi come concorrente ad altri popolari competitors quali Canva e Adobe Express. Esso propone, quindi,

una pipeline per la creazione di contenuti visivi alleggerita dall'intervento dell'IA, permettendo di personalizzare numerosi aspetti con la generazione artificiale [96].

Nel corso degli anni ha visto numerosi aggiornamenti che hanno aggiunto sempre più tool all'interfaccia, corretto bug e migliorato l'usabilità. Tuttavia, le innovazioni principali derivano sempre dai modelli a cui si appoggia: al suo lancio Image Creator si appoggiava a DALL-E 2, ma allo stato attuale ha implementato la terza versione di DALL-E.

### 5.4.2 Piani di abbonamento

Microsoft Designer, come molti prodotti della casa, propone un abbonamento base del singolo software e piani di abbonamento superiori che forniscono l'accesso all'intero pacchetto Microsoft. Oltre alle opzioni in tabella, è disponibile anche una versione Family di Microsoft 365, che aumenta il numero di dispositivi utilizzabili e lo spazio di archiviazione.

Ai fini della tesi, si è scelto di sfruttare l'abbonamento Free, che non mette nessun vincolo o limite di utilizzo all'infuori del numero di crediti.

	Free	Microsoft 365 Personal	Microsoft Copilot Pro
Prezzo / mese	0€	10€	22€
Crediti	15	60	150
Spazio di archiviazione	5GB	1TB	1TB
Pacchetto App Microsoft	X	Si	Si
Accesso prioritario generazione	X	X	Si

Tabella 4-10 Microsoft Designer: Piani di abbonamento

### 5.4.3 Interfaccia e funzionalità

Microsoft Designer vuole essere un software user-friendly, uno strumento intuitivo ed esaustivo per il grande pubblico. La tagline con cui si presenta riporta:

*Crea, progetta e modifica qualsiasi cosa puoi immaginare in pochi secondi con l'intelligenza artificiale.*

L'uso del termine "progettare" è significativo: il software vuole essere uno strumento per designer e per la creazione di contenuti elaborati e composti. In questo si discosta completamente da DALL-E, e l'utilizzo che fa del modello è inserito in un contesto più ampio di elaborazione e compositing delle immagini.

Infatti, Microsoft Designer è in grado di generare sia immagini che contenuti grafici di vario tipo: sticker, collage, inviti e biglietti, cornici, scritte ecc. Il tutto avviene a partire da descrizioni testuali, analogamente alle dinamiche di generazione di immagini. Il vantaggio fornito dalla

## Scelta dei tool applicativi

GenAI, in questo caso, è di eliminare il lungo processo di ricerca e personalizzazione dei modelli grafici per la creazione di contenuti.

Aperto il menù a tendina proposto dal software, è possibile esplorare le molte opzioni creative implementate. Si va dai singoli elementi di grafica - come emoji, monogrammi ed immagini – ai progetti più strutturati per realizzare locandine, inviti, collage e molto altro.



Figura 4-5 Opzioni di progetti Microsoft Designer

Cliccando su una qualsiasi di queste voci, si è rimandati alla pagina dedicata, dove è possibile sia creare un contenuto originale immettendo un prompt e la dimensione, sia esplorare numerosi template proposti, avendo modo di vedere la varietà con cui è possibile personalizzare i propri contenuti. Ogni template è accompagnato dal relativo prompt generativo, una funzione utile a comprendere la forma con cui il modello si aspetta di interagire.

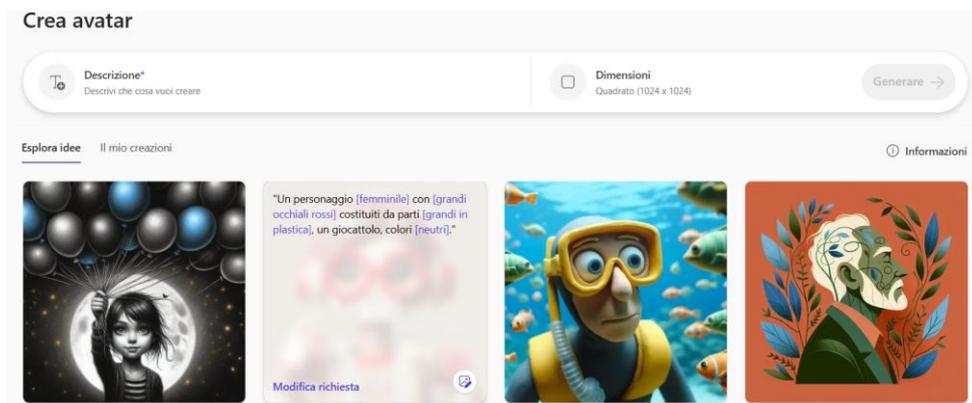


Figura 4-6 Creazione progetto Microsoft Designer

Cliccando sui template è possibile modificarli secondo le proprie esigenze; il prompt presenta dei campi modificabili per mantenere intatta l'idea generale e agire solo su specifici aspetti, quali soggetto, posizione, colore oggetti ecc. In alternativa, si può chiedere di riscrivere l'intero prompt modificandolo in maniera più profonda. I progetti più articolati, come gli inviti, possiedono più parametri modificabili, come stile, palette di colori e arte figurativa.

## Scelta dei tool applicativi

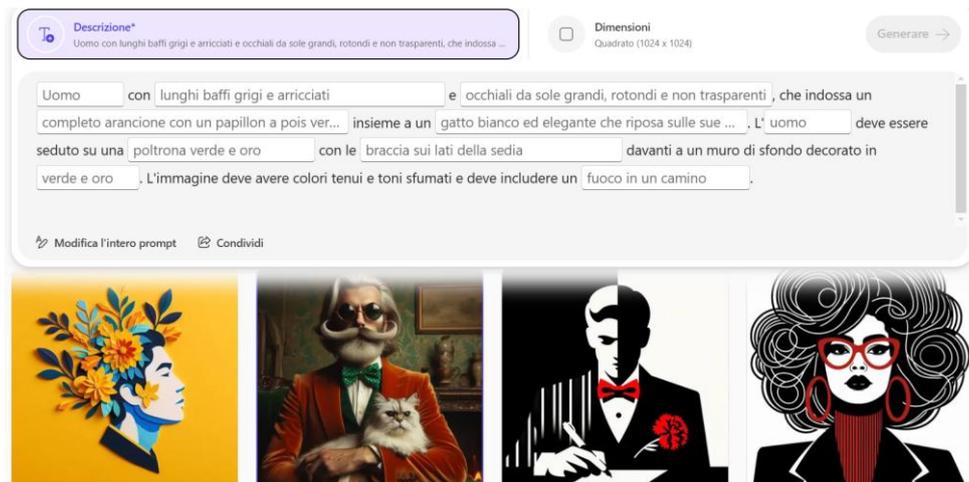


Figura 4-7 Modifica template Microsoft Designer

Entrando, invece, nella sezione di modifica di un contenuto si apre una pagina dedicata all'editing grafico del progetto. Si individuano tutte le tipiche funzionalità di software grafici, quali l'aggiunta di testo, grafiche, gestione dei livelli e marcatura libera. Se si clicca su un elemento fotografico, si aprono ulteriori opzioni di azione, a partire dalla regolazione dei tipici parametri fotografici (esposizione, saturazione ecc.) a strumenti più elaborati come rimozione/sostituzione sfondo, cambio del fuoco, spostamento di elementi, isolamento e molto altro.

## 5.5 Katalist

Katalist è il tool che è stato scelto per la realizzazione dello storyboard del progetto in tesi, un provider poco conosciuto ma che si presenta con delle caratteristiche uniche nel panorama. Nella pagina iniziale e di presentazione del suo sito [97] si presenta con la seguente tagline:

*Create Visual Stories with Consistent Characters and Scenes using Storyboard AI. For Filmmakers, Advertisers, and Content creators, and more. Katalist is the best way to visualize your ideas using storyboard AI. Use it for creative pitches, storyboards, video ideas and much more!*

Fin dalla sua presentazione, quindi, il modello si posiziona in una specifica nicchia di mercato che è quella dello storytelling e individua come suo punto di forza quello della consistenza. Questo aspetto è stato determinante nella fase di scelta, rendendolo una novità nel panorama delle immagini generative.

Purtroppo, essendo un prodotto poco popolare e relativamente nuovo, non esiste documentazione alcuna al riguardo, né ufficiale, né non ufficiale, per cui nulla di tecnico è individuabile sul suo conto. Si suppone un funzionamento che ibridizzi il processo di prompting, utilizzando altre immagini dello storyboard come input secondario per rendere consistenti personaggi e ambienti, o che usi dei prompt dedicati a tali elementi che vengono mandati in appendice al prompt immesso dall'utente, sotto forma di dati aggregati.

Ad ogni modo, quel che la tesi si propone di scoprire è se il tool sappia dare le funzionalità necessarie a mantenere la promessa di marketing.

### 5.5.1 Piani di abbonamento

Katalist offre tre tipologie di piani mensili di abbonamento per privati.

Essi si differenziano per numero di crediti per immagini e video, numero di personaggi personalizzati creabili, limite di progetto, numero di tavole per progetto e per features che offrono.

	Essential	Pro	Unlimited
Price / month	19\$	39\$	99\$
Image credits	200	700	Unlimited
Video credits	100	150	2000
Project limit	20	50	Unlimited
Panels per project	20	100	200+
Custom character	X	5	10
Voiceover	X	Yes	Yes
Zip export	X	Yes	Yes
PPT, video export	X	X	Yes

Tabella 4-11 Katalist piani di abbonamento

Alla luce dei piani offerti, ai fini della tesi si è scelto di sottoscrivere l'abbonamento unlimited, vista la necessità di realizzare storyboard con numerose tavole e avendo, così, la possibilità di sperimentare senza limiti di crediti.

### 5.5.2 Interfaccia e funzionalità

Si analizzano di seguito le caratteristiche di interfaccia e funzioni offerte da Katalist.

#### *Creare un nuovo progetto*

Katalist propone tre modalità di creazione di un progetto:

- Tramite upload di uno script
- Tramite creazione di uno script con assistente AI, partendo da una frase di input
- Creazione di un progetto vuoto

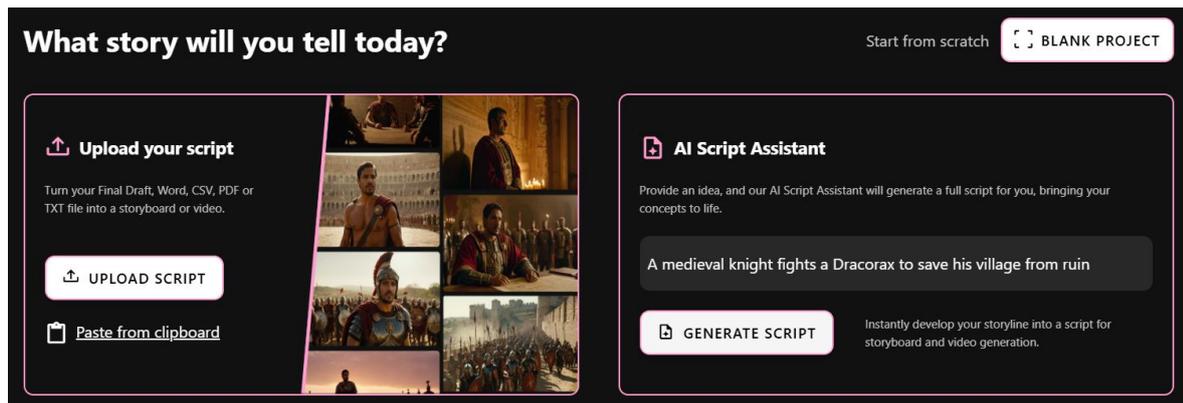


Figura 4-8 Katalist: Interfaccia creazione progetto Katalist

### *Impostazioni di progetto*

Che si parta da uno script pre-esistente o da uno script generato da Katalist, la prima fase del progetto vede generare uno **script breakdown** delle inquadrature che verranno generate nello storyboard. Esse saranno presentate con il numero corrispondente e una frase descrittiva dell'azione.

Lo script breakdown sarà modificabile nella descrizione di ogni shot e offre la possibilità di eliminarne completamente alcuni, tuttavia, non è possibile aggiungerne (sarà possibile farlo solo nella fase successiva). In questa sezione sarà possibile selezionare l'**aspect ratio** desiderato, limitato a tre opzioni: Landscape 16:9, Portrait 9:16 e Square 1:1.

Proseguendo alle impostazioni successive sarà possibile impostare lo stile delle immagini, scegliendo tra 10 disponibili, oppure facendo l'upload di un'immagine dalla propria galleria è possibile creare il proprio stile.

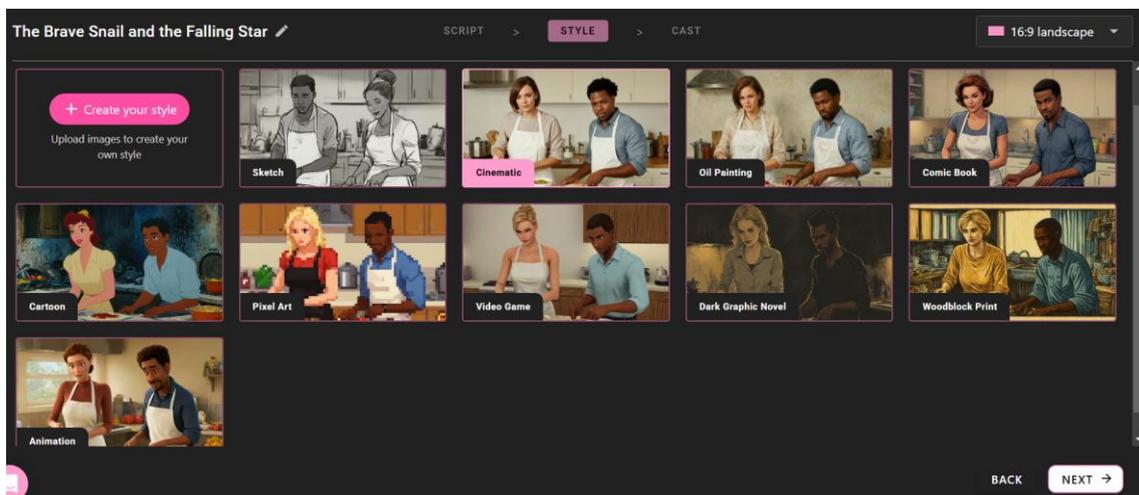


Figura 4-9 Katalist: Impostazione dello stile

Procedendo ulteriormente vengono presentati i membri del cast individuati e generati dall'IA, cliccando su ognuno di essi è possibile modificarne il nome, la descrizione degli abiti e l'aspetto. L'aspetto in particolare, può essere scelto a partire da otto personaggi pre-impostati di etnie e sesso differenti, disponibili in quattro diverse fasce di età. In alternativa vi è la possibilità di caricare uno schizzo o foto dai propri file per generare un personaggio personalizzato, modificabile ulteriormente con la scrittura di un prompt.

## Scelta dei tool applicativi

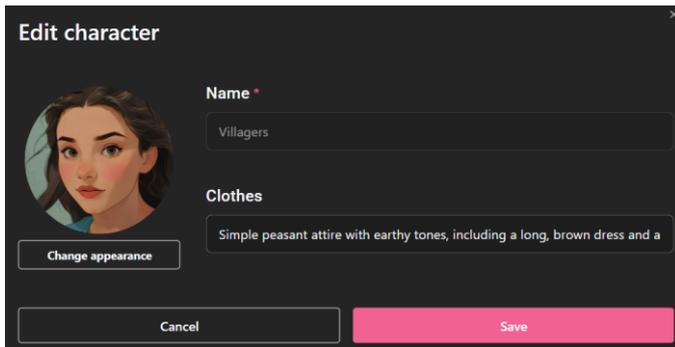


Figura 4-11 Katalist: modifica personaggio

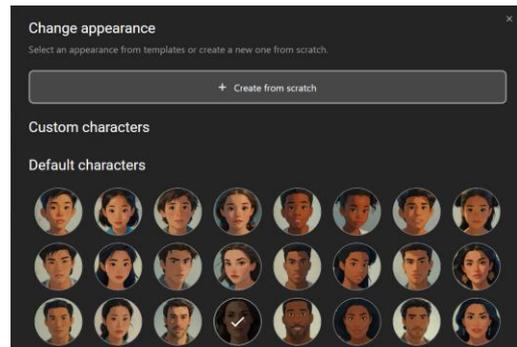


Figura 4-11 Katalist: modifica aspetto

### Storyboard

Una volta salvate tutte le impostazioni, il progetto viene creato, con una prima versione di storyboard generato. In testa alla pagina sarà possibile gestire i personaggi e le location della storia, aggiungendone e modificandoli. Questa particolare feature è la grande innovazione di Katalist, permettendo di avere elementi definiti riutilizzabili in immagini diverse.

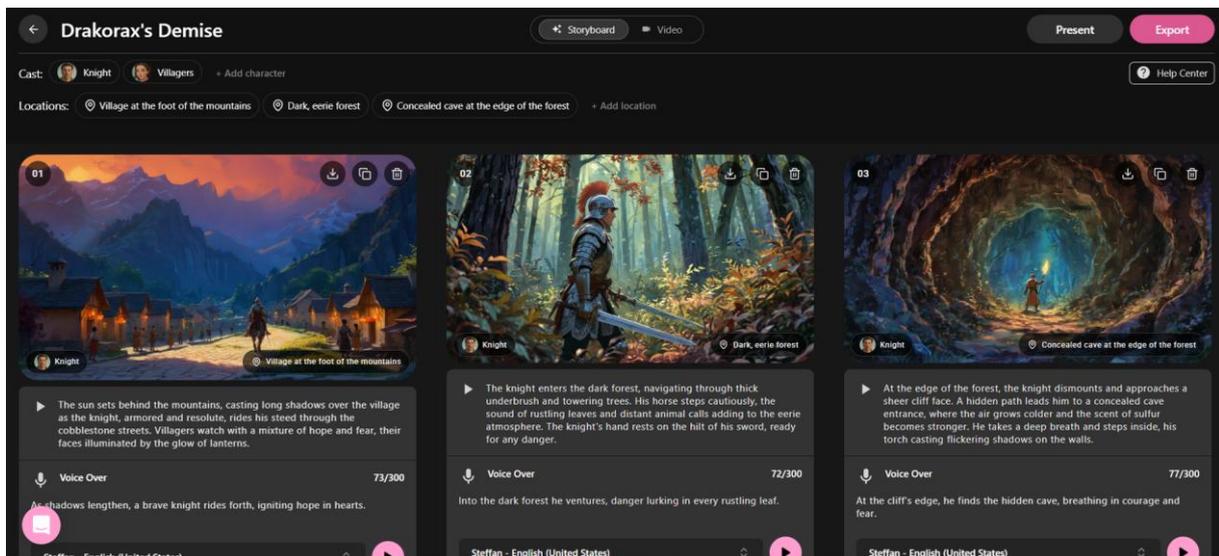


Figura 4-12 Katalist: Storyboard

Sotto queste due impostazioni si trovano, infine, gli shot generati, con la possibilità di eliminarli, copiarli, scaricarli o aggiungerne. Sotto ogni immagine sono presenti i tag dei personaggi e delle location della scena, oltre alla descrizione dell'azione; questi elementi sono modificabili entrando nell'interfaccia di editing dello shot. Inoltre, viene data la possibilità di generare il **voice-over** della descrizione, personalizzabile scegliendo tra numerose voci in lingua inglese. Questa caratteristica è utile soprattutto in fase di export in versione video.

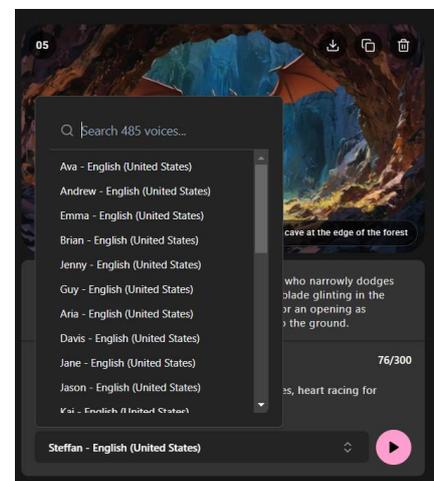


Figura 4-13 Katalist: Voiceover

### *Interfaccia di editing*

Cliccando sugli shot si entra nella specifica schermata di editing. Nella parte centrale si ha la casella di testo in cui è possibile modificare il prompt. I personaggi vengono menzionati con il carattere speciale @ (un massimo di due per shot) e l'ambientazione è modificabile cliccando sul tag apposito. In questa fase è possibile anche aggiungere negative prompt.

Tra le altre impostazioni figura lo **shot type**, impostabile scegliendo tra: close shot, medium shot long shot, wide shot, extreme close-up, bird's eye view e silhouette shot. Il **camera angle** è impostabile, invece, scegliendo tra: shot from above, shot from below, straight on shot, shot from the side, eye level shot, shot from behind. Un'ultima impostazione permette di scegliere se generare una, due o quattro immagini con lo stesso prompt per avere possibilità di scelta.

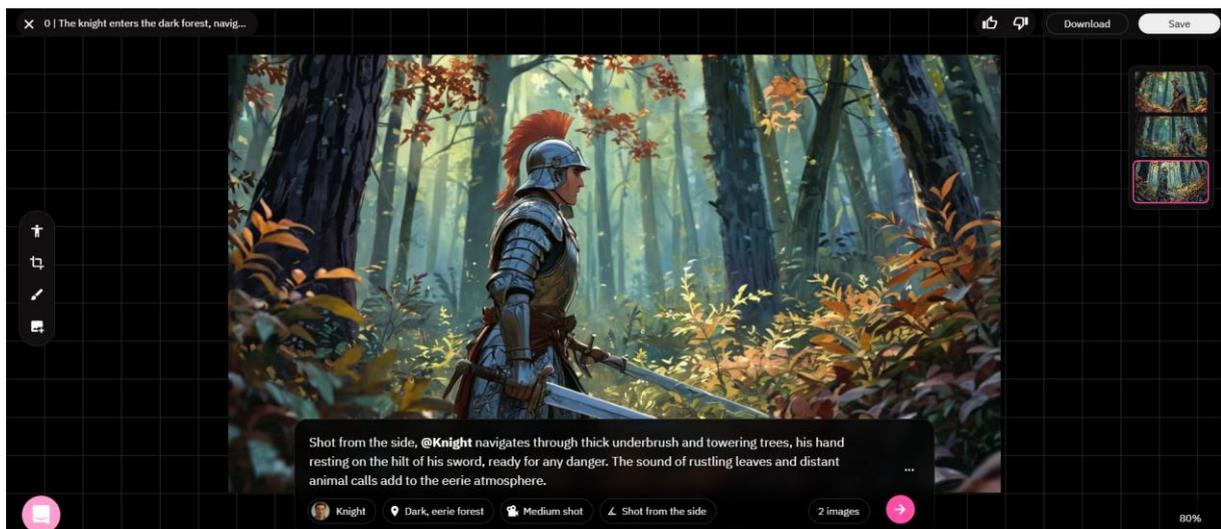


Figura 4-14 Katalist: Interfaccia di editing

Una funzionalità interessante è il mantenimento dello storico delle versioni, permettendo di selezionare volta per volta la versione di shot che più ci interessa tra tutti i tentativi generati.

Nella parte sinistra dell'interfaccia troviamo dei tool di editing, probabilmente i più interessanti e potenzialmente utili. Partendo dal basso, troviamo innanzitutto la possibilità di caricare un'immagine dal computer come image reference per la generazione. Proseguendo, si trova il **Generative fill tool**, un potente mezzo per colorare oggetti o aree dell'immagine in cui inserire o eliminare elementi.

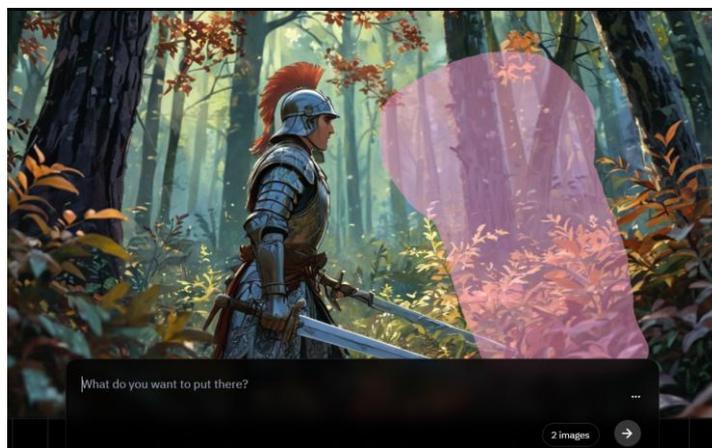


Figura 4-15 Katalist: Generative fill tool

## Scelta dei tool applicativi

Salendo ancora, troviamo il **Image Crop Tool**, che permette di ritagliare l'immagine. Infine, il **Character Posing Tool**, il tool di punta di Katalist, che permette l'utilissima funzionalità di riposizionare i personaggi controllando uno scheletro e di rigenerare l'immagine.

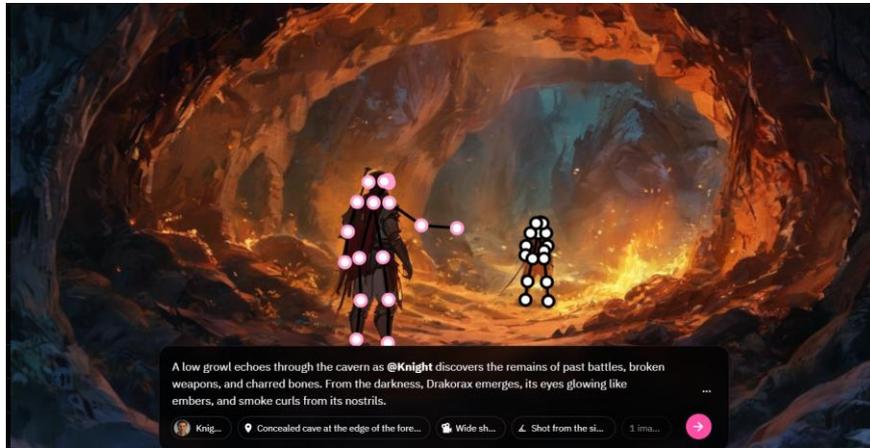


Figura 4-16 Katalist: Character posing tool

### *Generazione Video*

Tra le altre funzionalità del tool figura anche la generazione video tramite Stable Diffusion, permettendo di trasformare lo storyboard in un animatic. Per ogni video viene data la possibilità di definire una percentuale di “motion strength” all'interno del video, ma non di definire l'azione. L'interfaccia di montaggio è purtroppo poco elaborata e mal strutturata, rendendola inadatta ad un montaggio curato, ma per un export di basse pretese può essere funzionale. Tra le opzioni permette anche l'uso del voice-over prima esplorato.

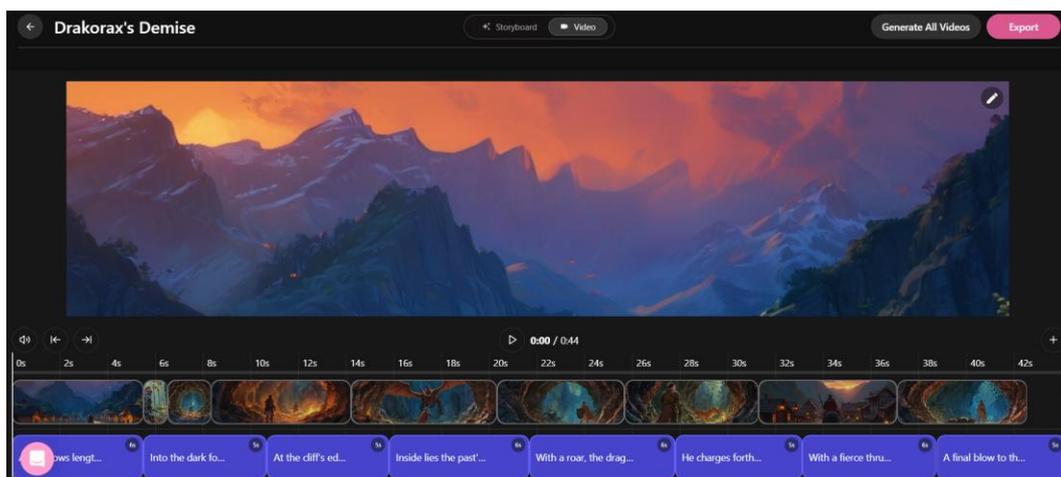


Figura 4-17 Katalist: Interfaccia montaggio video

### *Export*

L'ultimo aspetto da analizzare sono le possibilità di export, che prevedono: Presentation link, All image (ZIP), Shooting board (in formato Excel), Power Point file, Adobe Premiere Pro, Video. Dei formati che permettono diversi tipi di utilizzo in maniera agile.

## 6. Prompt Engineering

Nei capitoli precedenti abbiamo evidenziato come l'utilizzo dell'Intelligenza Artificiale possa aiutarci a svolgere i compiti più disparati, ne abbiamo analizzato le architetture, le dinamiche di funzionamento, ma ancora non è stato spiegato come dovremmo interagire con essa.

La questione è più complessa di quanto possa sembrare, infatti, seppur sia vero che l'IA abbia oggi raggiunto livelli di comprensione testuale elevati, capaci di rendere l'interazione con essa apparentemente semplice ed intuitiva, non è scontato riuscire ad ottenere output ottimali e, talvolta, è possibile ricevere risultati del tutto errati o inesatti. Questo avviene per un'assunzione della natura profonda dell'IA, la quale fornisce i suoi risultati basandosi sulla probabilità e sui database a cui ha accesso. Il concetto stesso di probabilità inserisce, quindi, un margine di errore dovuto a diversi fattori, quali database insufficienti, ambiguità delle richieste o bias del modello generativo. Per chi fa utilizzo dell'IA, infatti, è all'ordine del giorno ricevere in output immagini con artefatti, video distorti o risposte testuali inaccurate. Nel peggiore dei casi si può addirittura arrivare a quelle che in gergo vengono definite “**allucinazioni**”, ossia il fenomeno che si verifica quando l'IA fornisce una risposta errata, ma verosimile, spacciandola per corretta.

Questo mette in luce un punto fondamentale se si vuole avvicinarsi a questa nuova scienza: l'Intelligenza Artificiale è fallibile. Non a caso, all'interno dell'interfaccia di ChatGPT, al di sotto della barra di scrittura da cui poniamo le nostre domande, compare un chiaro messaggio di warning:

*ChatGPT può commettere errori. Considera di verificare le informazioni importanti.*

Appurata la sua natura imperfetta, l'IA non perde, però, la sua utilità ed il suo valore. Semplicemente, cambia il rapporto che dobbiamo avere con essa: non più guidato da una cieca fiducia, ma attento e critico, al fine di riuscire ad ottenere il massimo possibile da questo potente strumento. È da questo nuovo approccio che nasce la **Prompt Engineering**, una scienza capace di farci raggiungere il miglior risultato possibile dagli output delle nostre conversazioni digitali.

### 6.1 Cos'è la prompt engineering

Il termine prompt engineering è apparso per la prima volta su un popolare post online che parlava del LLM GPT3 e analizzava le sue capacità di scrittura creativa [98]. L'autore del post, Gwern Branwen, individuò nella Prompt Engineering un nuovo paradigma per l'uso dell'Intelligenza Artificiale, portando gli utenti a studiare metodologie testate per poter sollecitare nella maniera migliore il LLM ed ottenere in output la risposta desiderata [99].

Ma per poter comprendere al meglio il significato del Prompt Engineering occorre innanzi tutto chiarire da subito cosa intendiamo con il termine Prompt. Per farlo mi affido alla definizione fornita dalla guida online *Learn Prompting* [100], che ho ritenuto la più semplice ed esaustiva tra quelle analizzate:

*Un prompt è un input o un'istruzione fornita a un modello AI per generare una risposta. I prompt possono assumere molte forme, da semplici domande a istruzioni più complesse che specificano tono, stile o struttura. Sono il*

*meccanismo attraverso cui gli utenti comunicano con i modelli AI e la chiarezza del prompt influenza direttamente la qualità dell'output dell'AI.*

Come tutti i sistemi basati su un processo del tipo Input→Elaborazione→Output, è inevitabile che l'input immesso nel sistema sia determinante nella caratterizzazione dell'output. L'IA generativa è, pertanto, basata in primis sul prompt che riceve in ingresso. Grazie alla natura dei nuovi modelli basati su **Natural Language Model** (NLP), chiunque può creare un prompt di base da sottoporre, tuttavia, non è banale riuscire a strutturare una query capace di attivare le massime potenzialità del modello. Ed è per questo motivo che diventa indispensabile l'uso della Prompt Engineering.

Questa disciplina si occupa di studiare l'ottimizzazione della **Human-Computer Interaction** (HCI), individuando le tecniche con cui migliorare i prompt da sottoporre all'intelligenza artificiale. Questa attività si applica in qualsiasi campo di contenuti generativi, siano essi testi, immagini, video, codice, canzoni, ecc.

Acquisire elevate capacità di scrittura dei prompt è il primo passo per poter intraprendere una strada che porta ad un utilizzo consapevole dell'Intelligenza Artificiale, avvicinandoci sempre più al concetto di **Human-Centered AI** (HCAI), un sistema che posiziona l'uomo al centro dell'interazione con l'IA, ridefinendo quest'ultima come potente alleato piuttosto che come concorrente.

## 6.2 I luoghi di studio della Prompt Engineering

La prompt engineering è una disciplina ancora in fase di sviluppo, esistono poche regole standardizzate ed ufficiali. Questo è dovuto alla natura stessa dell'IA, una delle poche invenzioni dell'uomo dove l'utente non ha il pieno controllo della macchina. Seppur si possano studiare i LLM, al fine di comprenderne i meccanismi ed i processi logici svolti dall'IA, la loro natura fortemente automatizzata li rende comunque imprevedibile nei suoi risultati. Per questo motivo è difficile scrivere un "libretto di istruzioni" completamente esaustivo nel suo scopo.

Ogni modello presenta nel suo sito una guida al suo utilizzo, dove vengono elencate le regole di base per interagire con la macchina, fornendo spesso esempi pratici e, talvolta, consigliando tecniche anche più avanzate. Ci siamo occupati di analizzare alcune delle guide dei tool più popolari come OpenAI [101] e Runway [102], e abbiamo trovato delle guide ottime nel fornire le nozioni necessarie a poter fare un buon utilizzo delle piattaforme. Tuttavia, abbiamo detto che la prompt engineering ha lo scopo di massimizzare la qualità dell'output, e per farlo occorre apprendere attraverso un'ampia sperimentazione, fatta di iterazioni, tentativi ed errori. Tali pratiche esulano dalle guide ufficiali poiché spesso sono tecniche ancora in fase embrionale e non sufficientemente testate, ma è proprio attraverso di esse che risulta possibile il raggiungimento dei migliori canoni di qualità di output.

Se le guide ufficiali sono insufficienti, occorre trovare, quindi, dei canali di studio del prompt engineering alternativi. Fortunatamente, l'Intelligenza Artificiale è un argomento che ha attirato attorno a sé molti appassionati, un tipo di pubblico con un forte legame con l'informatica e che è abituato a spendere le proprie competenze per la ricerca all'interno di un crescente ecosistema online. Avere un così ampio gruppo di praticanti è un aiuto prezioso, soprattutto perché molti dei modificatori di prompt più efficienti non sono intuitivi e l'unico modo di scoprirli è un'azione di sperimentazione ad ampio raggio. L'unica sfida, a questo punto, è quella di individuare i luoghi preferiti di queste community e capire quali di esse siano più attendibili.

Tra i punti di partenza per rintracciare il materiale, è stato individuato uno studio svolto da Jonas Oppenlaender intitolato *A taxonomy of prompt modifiers for text-to-image generation* [103]. Per svolgere il suo studio sui prompt modifiers per la generazione di immagini, ha svolto un interessante ricerca etnografica delle community di prompt engineering online. Il suo lavoro è partito da uno studio della comunità sul social media X, dove molti AI Artist condividono i loro lavori, spesso accompagnati in didascalia dai prompt che li hanno generati. Da X lo studio è riuscito ad intercettare collegamenti per spostarsi sul web, dove forme più disparate di risorse sono state trovate sotto forma di blog, pagine GitHub e Reddit, banche online condivise, tutorial e guide online.

Essendo una comunità attiva ed una tecnologia in continua evoluzione, siamo partiti dalla rete di risorse di questo studio, per poi espanderla e poter individuare le fonti più aggiornate e all'avanguardia.

Tra i canali analizzati figura anche YouTube, che si è però dimostrato spesso troppo superficiale nel trattare l'argomento, attraverso semplici tutorial che poco aggiungono alle guide ufficiali. Invece, si sono individuate delle **guide online** estremamente dettagliate, gestite da praticanti che quotidianamente osservano le nuove tecniche che emergono dal web, al fine di individuare le più valide ed integrarle nella propria raccolta. Tra queste, ne sono state scelte tre come riferimento: la guida *Learn Prompting* [100] (su prompting per generazione di testo ed immagini), *The DALL-E 2 Prompt Book* [104] (sulla generazione immagini con Dall-e) e *A Traveler's Guide to the Latent Space* [105] (sulla generazione immagini).

Un'altra inestimabile fonte per lo studio di prompt engineering per le immagini sono state le banche online in cui è possibile esplorare le immagini generate dai cosiddetti AI Artist, utili specialmente quando accompagnate dai prompt che le hanno generate, dando la possibilità di vedere casi applicativi e i risultati delle tecniche. Tra i siti esplorati spiccano per qualità *ArtHub* [106], *Lexica.art* [107] e *OpenArt.ai* [108].

Grazie al lavoro di tanti appassionati è stato possibile individuare le tecniche che ci hanno permesso di svolgere questo studio. Il capitolo proseguirà, dunque, esplorando le tecniche più interessanti che sono state individuate.

### 6.3 Prompt Engineering testuale

Il concetto di prompt engineering è nato nell'ambito della generazione testuale, è in esso che sono state scoperte quelle che, ad oggi, sono i concetti basilari e fondanti della cura dei prompt in qualsiasi campo di applicazione.

Quando parlo di generazione testuale non mi riferisco solamente alla capacità creativa di un LLM, ma in generale a qualsiasi tipologia di quesito la cui risposta è prevista come un output testuale, questo quindi comprende anche codice, problemi matematici, ragionamenti logici e richieste di esecuzione di istruzioni di ogni genere. Quelle che ci si appresta ad analizzare sono tecniche applicabili in più contesti e per cui scopi diversi possono dare risultati di efficacia diversi.

Quanto segue è una serie di metodologie individuate dalla guida *Learn Prompting* [100], che si è occupata di raggruppare sotto un solo testo tutte le tecniche più popolari e funzionali delle comunità di appassionati online. Non è detto che esse siano state testate a sufficienza per poterle assumere come regole, ciononostante ne emerge sempre un approccio costruttivo ed una serie di pratiche imprescindibili per avvicinarsi all'arte della scrittura di prompt.

### 6.3.1 Elementi di un prompt testuale

Un prompt è caratterizzato da specifiche parti al suo interno, ognuna delle quali permettere di controllare specifici aspetti dell'output che si riceverà in risposta.

Le parti fondamentali di un prompt sono le seguenti.

- **The Directive:** la direttiva è l'elemento che regge l'intero prompt, rappresenta la richiesta, sotto forma di domanda o di istruzione, che poniamo all'IA. Avere una direttiva dispersiva o non chiara renderà impossibile per il modello svolgere quanto gli chiediamo.
- **Examples:** per istruzioni complesse o particolarmente specifiche, si è scoperto che l'utilizzo di esempi può essere un ottimo punto di partenza per guidare l'IA verso le risposte più accurate.
- **Role:** è possibile assegnare un ruolo da impersonare all'IA, questa tecnica può essere un aiuto per instradare il modello verso un determinato stile di scrittura, influenzandone i modi ed i toni. L'uso del ruolo, in realtà, può anche modificare il contenuto dell'output, seppur questo aspetto stia a poco a poco perdendosi con l'evoluzione dei modelli.
- **Output Formatting:** la funzione di questo modello è semplicemente quella di definire il formato in cui desideriamo venga restituito l'output, questo fornisce un'importante leva di controllo per minimizzare la post-produzione da operare sulla risposta ottenuta.
- **Additional Information:** fornire informazioni aggiuntive utili e pertinenti può agevolare il modello nello svolgere il suo lavoro. Ad esempio, può rivelarsi utile fornire un contesto dettagliato prima di porre una domanda, questo aiuta ad inquadrare il campo di competenza e a rendere l'elaborazione più accurata.

Questi elementi possono essere sviluppati in relazione al tipo di richiesta che stiamo formulando, dando più rilievo agli uni o agli altri a seconda delle esigenze e potenzialità. Occorre comunque ricordare che nella prompt engineering non esiste una ricetta garantita, per cui ogni conversazione con l'AI avrà dinamiche proprie da gestire iterativamente.

Gli elementi sopra elencati possono, ad esempio, essere disposti in ordini differenti, e questo può influire sull'accuratezza che riceveremo in risposta. Una pratica consigliata è quella di partire dal contesto e dagli esempi, dando modo all'IA di pre-elaborare le informazioni utili, per poi chiudere il prompt con la direttiva ed il formato, assicurandoci che abbia ben chiara la nostra richiesta.

### 6.3.2 Tecniche di Prompt Engineering testuale

Analizziamo di seguito le principali tecniche per affinare i prompt che sottoponiamo all'IA, verranno presentate secondo un livello crescente di complessità.

#### *Role Prompting*

Questa tecnica basa la sua riuscita sull'elemento *Role* del prompt. Essa prevede di assegnare all'IA un ruolo, sia esso un mestiere, una persona, o una caratteristica caratteriale. Questo processo indirizza il modello a formulare l'output secondo uno stile appropriato al ruolo assegnato. Nell'assumere un ruolo, l'IA cambierà anche il contenuto della sua risposta, soffermandosi maggiormente su ciò che compete il ruolo che impersonifica, questo significa che un ruolo ben assegnato può aumentare il livello di dettaglio ed attitudine che verrà restituito.

Curiosamente, in particolar modo nei primi modelli di IA, si è scoperto che l'assegnazione di un ruolo può determinare risposte esatte o errate in casi di problemi matematici. Infatti, per via della sua natura, i modelli attuali sono talvolta fallibili di fronte a problemi matematici o logici complessi; assegnare un ruolo di "Esperto matematico" o "Professore di Fisica" può fare la differenza nel guidare il ragionamento del modello verso un risultato esatto. Ad ogni modo, la tecnica mostra il suo massimo potenziale nell'uso dell'IA per la scrittura creativa, dando la possibilità di dare fini sfumature di stile all'elaborazione.

Occorre, tuttavia, specificare che non sono stati svolti sufficienti test per garantirne una migliore performance, senza contare che alcuni ruoli aumentano le probabilità di incorrere in bias determinati dai dataset che caratterizzano tale ruolo.

Il Role Prompting si struttura in due fasi, nella prima si definisce il ruolo e nella seconda si esegue la richiesta.

**Esempio 1 Role Prompting:** *Sei un importante critico cinematografico, hai appena assistito alla prima proiezione di [film]. Scrivi una recensione di 500 caratteri.*

### ***Shot-Based Prompting***

La Shot-Based Prompting è una delle tecniche più utili e versatili di questo elenco. Analogamente al Role Prompting, anche questa tecnica si concentra su uno specifico elemento del prompt, nella fattispecie sugli *Examples*. Questa procedura prevede infatti la presentazione di esempi al modello di IA, seguendo la tecnica chiamata In-Context Learning (ICL). In questo modo, il modello si attiverà e apprenderà dagli esempi forniti all'interno del prompt.

Per la sua implementazione il prompt viene strutturato a partire da una o più domande, cui seguiranno delle risposte di esempio, per poi terminare con la domanda vera e propria che vogliamo porre. Gli esempi possono essere strutturati in formulazioni precise e chiare, al fine di agevolarne la comprensione, ma anche in forma libera possono dare ottimi risultati.

Si parla di **Zero-Shot Prompting** quando si realizzano prompt senza fornire esempi, ossia la comprensione del testo è affidata alla sola conoscenza pregressa del modello. Si parla, invece, di **One-Shot Prompting** in presenza di un solo esempio e di **Few-Shot Prompting** per numeri maggiori.

#### **Esempio 1 Shot-Based Prompting:**

*Classifica la valutazione della seguente recensione come Positiva o Negativa:*

*Recensione: Il film è noioso e pretenzioso → Valutazione: Negativa*

*Recensione: Il film è sconvolgente, una gioia per gli occhi → Valutazione: Positiva*

*Recensione: Il miglior film dell'anno → Valutazione:*

#### **Esempio 2 Shot-Based Prompting:**

*Quanto misura l'area di un rettangolo con base 10m e altezza 2m?  $A=10m \times 2m=20m^2$ .*

*Quanto misura l'area di un rettangolo con base 30cm e altezza 100cm?*

### ***Priming Prompting***

Il Priming Prompting è una tecnica utilizzabile per plasmare il comportamento di un chatbot secondo le proprie esigenze. Esso si articola attraverso più fasi, e sfrutta la memoria del chatbot per dare uno stile o un determinato mood alla conversazione e mantenerlo. Infatti, prevede una

serie di prompt iniziali, il cui scopo è questo di caratterizzare e plasmare l'IA con cui dialoghiamo nella conversazione.

Questo tipo di approccio può, in certi aspetti, ricordare il role prompting, ma la differenza tra i due è proprio la durezza: laddove il role prompting si estingue all'interno del prompt, fare priming significa caratterizzare un'intera chat in modo coerente.

Questo può essere tanto utile per impostare uno stile di scrittura, quanto per specificare determinati comportamenti che possono risultare utili.

**Esempio 1 Priming Prompting:**

*Sei il mio consulente di finanza con 20 anni di esperienza, sei onesto e diretto, se ti chiedo consigli tu mi rispondi con dati certi e statistiche, senza perdere tempo con i sentimenti.*

**Esempio 2 Priming Prompting:**

*Sei il mio assistente di studio, mi farai delle domande di storia e io risponderò, tu mi risponderai con un voto da 1 a 10 in base all'accuratezza, e mi consiglierai come migliorarmi.*

**Chain-of-Thought Prompting**

I prompt tradizionali raramente affrontano i problemi per passaggi, ma sfruttano un processo diretto di tipo input-output, rendendo difficoltosi i ragionamenti più complessi strutturati su più passaggi. La Chain-of-Thought (CoT) [109] Prompting è una tecnica che, sotto questo punto di vista, fa un passo in avanti, portando i LLM a migliorare le proprie capacità di ragionamento.

La tecnica è un'evoluzione della Shot-Based Prompting, incorpora negli esempi una catena di pensiero, ossia mostra i passaggi logici dietro alla soluzione del quesito incoraggiando il LLM a ragionare in più fasi. In risposta, avremo output che ragioneranno per step e ci mostreranno i loro ragionamenti logici.

Questo tipo di approccio ha registrato considerevoli miglioramenti su diversi ambiti di applicazione, quali aritmetica, ragionamenti logici di buonsenso e ragionamento simbolico (capacità di manipolare concetti astratti e simboli).

Ad esempio, il LLM PaLM 540B ha raggiunto risultati elevatissimi su GSM8K3 (dataset di addestramento per problemi matematici verbali) e su numerosi altri dataset. La tabella seguente riassume i miglioramenti delle prestazioni sui benchmark chiave quando si utilizza il prompting CoT [110].

Compito	Modello	Precisione di prompt standard	Precisione del prompt CoT	Miglioramento
<b>GSM8K (Matematica)</b>	Palma 540B	55%	74%	+19%
<b>SVAMP (Matematica)</b>	Palma 540B	57%	81%	+24%
<b>Buonsenso (CSQA)</b>	Palma 540B	76%	80%	+4%

<b>Ragionamento simbolico</b>	Palma 540B	~60%	~95%	+35%
-------------------------------	------------	------	------	------

Tabella 6-1 Percentuali di accuratezza del CoT Prompting

**Esempio 1 Chain-of-Thought Prompting:**

*Quanto misura l'area di un rettangolo con base 10m e altezza 2m? Recuperiamo la formula dell'area del rettangolo,  $Area = base \times altezza$ , per cui  $A = 10m \times 2m = 20m^2$ . Quanto misura l'area di un rombo con diagonale maggiore 30cm e diagonale minore 10cm?*

***Contrastive Chain-of-Thought Prompting***

Questa tecnica amplia la precedente basandosi sulla capacità umana di apprendere dall'uso di esempi sia positivi che negativi. Essa, semplicemente, prevede di utilizzare anche esempi di ragionamenti fallaci, ciò permette ai LLM di individuare i tipi di collegamenti logici che possono portare a commettere errori ed evitarli.

***Zero-Shot Chain-of-Thought Prompting***

Questa tecnica [20] riprende i concetti della Chain-of-Thought Prompting semplificandoli enormemente. Prevede di introdurre alla fine di un prompt l'espressione "**Pensiamo passo dopo passo**", questa semplice aggiunta porta i LLM a ragionare per step e migliorare la propria accuratezza.

La tecnica è stata testata in problemi aritmetici, di buon senso e di ragionamento simbolico, e ha raggiunto risultati considerevoli nonostante la sua essenzialità, seppur peggiori del prompting CoT standard.

Diverse formulazioni analoghe sono state testate per trovare la più efficace, quali "Risolviamo questo problema suddividendolo in passaggi" o "Pensiamoci logicamente", ma attualmente "Pensiamo passo dopo passo" si è dimostrata l'espressione con risultati migliori.

**Esempio 1 Zero-Shot Chain-of-Thought Prompting:**

*Un'automobile si muove di moto rettilinea uniforme a 70 km/h. Se parte alle 8.00 di mattina e arriva a destinazione alle 21.00, quanti km ha percorso? Pensiamo passo dopo passo.*

***Self-Consistency Prompting***

L'autoconsistenza [111] è una tecnica semplice, ma lenta, che prevede di inviare lo stesso prompt in input a un modello ripetutamente, col fine di individuare la risposta più frequente.

***Generated Knowledge Prompting***

Questo metodo [112] sfrutta le capacità di ricerca dei modelli di IA per reperire contenuto informativo utile all'addestramento. Prevede, infatti, di interrogare in una fase iniziale il LLM, chiedendogli di reperire informazioni utili allo svolgimento del compito, dopodiché invia la domanda vera e propria.

Il modello riesce, così, a crearsi da sé le informazioni aggiuntive utili a rendere la risposta pertinente.

**Esempio 1 Generated Knowledge Prompting:**

*Prompt 1: Reperisci delle informazioni riguardo il romanzo "Guerra e pace".*

*Prompt 2: Scrivi un riassunto ed un'analisi di "Guerra e pace".*

***Emotion Prompting***

Un errore comune nell'approcciare le proprie chat con l'IA è quello di rivolgersi a loro con un linguaggio troppo computazionale e distaccato. Questa modalità perde di vista il fatto che i LLM sono addestrati a partire da database costituiti in gran parte da conversazioni umane, ove è quindi comune ricorrere un linguaggio più emotivo.

Il prompting emotivo [113] prevede proprio di integrare nei propri prompt delle espressioni di stampo emotivo, ciò può essere utile a dare sfumature interessanti in utilizzi di scrittura creativa. Tuttavia, utilizzarlo per compiti tecnici o fattuali - in cui l'uso di emotività non aggiungerebbe nessun beneficio - potrebbe allungare inutilmente l'elaborazione del modello; per questo motivo è consigliabile applicarlo solo in specifiche occasioni.

Alcune espressioni utilizzabili per rendere più emotivo il proprio prompt possono essere:

- *È importante che sia corretto quel che mi dici.*
- *La tua risposta non mi convince a pieno. Sei sicuro sia giusta?*
- *Sii meticoloso e dammi la miglior risposta possibile!*
- *Procediamo con attenzione, non possiamo fare passi falsi.*

***Re-reading RE2***

La tecnica Re-reading (RE2) [114] rientra in quel gruppo di tecniche estremamente semplici nell'implementazione, ma che riescono da sole ad apportare modeste migliorie. Essa sfrutta un comando di rilettura da imporre al LLM, questo può essere utile ad incrementare la comprensione della richiesta da parte dell'IA.

Il metodo funziona perché l'IA non agisce secondo vincoli di lettura precisi, una stessa domanda può essere interpretata in maniera diversa a seconda della parte su cui si focalizza maggiormente. La rilettura permette al modello di ampliare la gamma di elementi del prompt cui dare peso.

Ovviamente, RE2 è utile in casi di problemi complessi, mentre può risultare inutilmente ridondante per compiti semplici.

**Esempio 1 Re-reading Prompting:**

*[Input Prompt]. Leggi di nuovo la domanda: [Input Prompt].*

***Rephrase and Respond (RaR)***

La tecnica Rephrase and Respond (RaR) [115] ha introdotto un'idea basilare nella prompt engineering, essa prevede di sfruttare le capacità dell'IA per la scrittura del prompt stesso. Essa prevede di sottoporre una domanda al modello e domandargli di riformularla apportando migliorie. Questo approccio porta ad ottenere prompt che usano un linguaggio aderente ai LLM, perché vengono scritti da loro stessi, rendendoli più chiari ai loro occhi.

Questa procedura è diventata la base di un prompt engineering più elevato, che nei casi più complessi porta ad un addestramento dell'IA per farle scrivere i migliori prompt possibili. Vedremo nei capitoli successivi come questa tecnica trovi il suo massimo potenziale nella scrittura di prompt per la generazione di immagini.

### **Esempio 1 Rephrase and Respond Prompting:**

*[Input Prompt]. Riformula e amplia la domanda, quindi rispondi.*

### ***Paraphrasing Prompting***

Una tecnica molto simile alla Rephrase and Respond Prompting è la Paraphrasing Prompting, che prevede nient'altro che chiedere al modello di produrre più varianti semanticamente simili di un prompt ed utilizzarle per ricevere più risposte da aggregare. Un'implementazione semplice del metodo può essere di tradurre il prompt in più lingue e poi ritradurle nella lingua iniziale. Ovviamente è una tecnica computazionalmente pesante.

### ***System to Attention (S2A)***

System to Attention [116] è uno specifico uso della tecnica Rephrase and Respond. Il problema che si propone di risolvere è il fatto che i LLM possano incrementare la percentuale di inaccuratezza in presenza di prompt troppo dispersivi, che forniscono informazioni irrilevanti alla risposta della domanda e rischiano che l'IA si concentri su passaggi inutili.

S2A prevede di sottoporre un prompt al modello e di chiedergli di rigenerare il prompt eliminando il contesto irrilevante alla richiesta. A questo punto, il prompt rigenerato viene sottoposto per ottenere la risposta effettiva.

Occorre dire, però, che gli ultimi aggiornamenti dei LLM hanno reso questa tecnica meno efficace ed utile, in quanto implementano autonomamente la capacità di distinguere le informazioni rilevanti ed irrilevanti.

### ***Step-Back Prompting***

La Step-Back Prompting [117] è una tecnica avanzata e d'avanguardia, che invita i LLM a "fare un passo indietro" e ragionare in prima battuta su concetti di alto livello e, solo successivamente, di applicarli per ottenere la risposta. Il processo può essere immaginato come una risposta ponderata, che prima raccoglie le idee sulla materia e poi risponde con maggiore consapevolezza e controllo.

Questa tecnica è particolarmente utile, intuitivamente, per quesiti complessi e che richiedono un ragionamento strutturato. Astrarre i concetti prima di applicarli porta i LLM ad un'accuratezza della risposta elevata, in quanto ogni azione si basa innanzitutto su teorie assodate.

Lo Step Back Prompting è stato testato su differenti LLM, come PaLM-2L, GPT-4 e Llama2-70B, mostrando miglioramenti significativi rispetto agli altri metodi tradizionali.

La tecnica si articola, come anticipato, in due fasi:

1. Fase di Astrazione: viene chiesto al modello di focalizzarsi su un concetto di alto livello utile alla risoluzione della domanda.
2. Fase di Ragionamento: assodata la teoria recuperata nella fase precedente, si richiede al modello di metterla in pratica per risolvere il quesito effettivo.

Step back Prompting è stato un miglioramento non banale rispetto alle tecniche tradizionali come la Chain-of-Thought (CoT) Prompting. In quest'ultima, infatti, si richiedeva al modello di suddividere il ragionamento in più sotto-problemi ed eseguirli passo per passo; seppur questo approccio migliori l'attenzione del modello nello svolgere i passaggi, non elimina il rischio di produrre step intermedi errati, che portano, infine, ad un risultato finale sbagliato o inaccurato. Step back Prompting esula da questo problema, astruendo i concetti in prima battuta ed evitando, così, tutti gli errori di basso livello, che possono provenire dal concentrarsi eccessivamente sui dettagli, perdendo di vista il problema nel suo complesso.

**Esempio Step Back Prompting:**

*Prompt di astrazione: Quali sono i principi e le leggi fondamentali che sono coinvolti in questo problema di fisica? [Problema].*

*Prompt di Ragionamento: Utilizza le leggi individuate per risolvere il seguente problema: [Problema].*

Step-Back Prompting è stato testato in più campi di applicazione, ottenendo sempre risultati ottimi. Di seguito sono riportati i guadagni di prestazioni osservati con Step-Back Prompting rispetto ai modelli di base e ad altre tecniche di prompting:

Compito	Modello	Prompt di base	Miglioramento dello Step-back prompting
MMLU (Fisica)	PaLM-2L	66,4%	+7%
MMLU (Chimica)	PaLM-2L	70,9%	+11%
TimeQA	PaLM-2L	41,5%	+27%
MuSiQue	PaLM-2L	35,5%	+7%
StrategyQA	PaLM-2L	82,8%	+3.6%

Tabella 6-2 Percentuali di accuratezza dello Step-Back Prompting

Si evincono dalla tabella considerevoli miglioramenti su tutti i dataset testati, con incrementi dell'accuratezza dal 7% fino al 27% (per i ragionamenti di tipo temporale). Il metodo si è dimostrato efficace anche su altri LLM, come GPT-4 e Llama2-70B.

***Universal Self-Consistency Prompting***

La Universal Self-Consistency [118] (USC) è l'evoluzione della tecnica Self-Consistency: entrambe prevedono di partire da più risposte allo stesso prompt, con la differenza che la Universal Self-Consistency Prompting affida la scelta della risposta migliore direttamente al LLM. In questo modo, questa tecnica diventa utile anche per generazione di testo in forma

libera, laddove la sua antenata era basata esclusivamente sul concetto di frequenza ed era, pertanto, valida solo su problemi a risposta esatta e concisa.

Le fasi per la sua implementazione sono:

- Fase di generazione multipla: Generare risposte multiple sottoponendo lo stesso prompt, avendo premura di tenere traccia delle risposte.
- Fase di scelta output: sottoporre l'elenco di tutte le risposte al modello, chiedendo di selezionare la risposta più accurata o coerente.

USC è stato testato su più dataset di diversa natura di problemi, riportando considerevoli miglioramenti. Di seguito sono riportati in tabella i dati registrati, confrontati con la SC standard.

Task	Random Selection	Standard Self-Consistency	USC
<b>Math (GSM8K)</b>	82.9%	90.4%	90.2%
<b>Code Generation (ARCADE)</b>	26.8%	30.3%	30.1%
<b>TruthfulQA (Open Q&amp;A)</b>	62.9%	Not Applicable	67.7%

Tabella 6-3 Percentuali di accuratezza del USC Prompting

La tabella mostra come la USC, seppur abbia ricevuto percentuali simili alla Standard Self-Consistency in problemi di matematica e generazione codice, abbia ampliato i campi di applicazione della tecnica, rendendola fruttuosa nel campo di Q&A di veridicità.

### ***Chain-of-Verification (CoVe) Prompting***

Abbiamo già introdotto in precedenza il fenomeno delle allucinazioni, risposte verosimili fornite dai LLM, che sono fondamentalmente errate. Tale fenomeno aumenta le proprie possibilità d'essere per generazione di testi lunghi, in cui i LLM possono essere superficiali su alcuni passaggi ed introdurre errori.

Un metodo per ridurre tale evenienza è la Chain-of-Verification [119], una tecnica analoga alla CoT che utilizza domande di verifica di veridicità per assicurare la qualità delle risposte.

La generazione di risposte tramite CoVe è un processo in 4 fasi:

1. Generazione dell'output di base: si comincia chiedendo al modello di generare una prima versione di risposta che, per il momento, potrà essere sia corretta che sbagliata.
2. Pianificazione delle verifiche: si sottopone il prompt e l'output iniziale al modello, chiedendo di generare delle domande di verifica su di essi.

3. Esecuzione della verifica: vengono sottoposte le domande di verifica al modello, in riferimento all'output iniziale.
4. Generazione della risposta finale: si incorporano le risposte delle domande di verifica alla domanda iniziale, chiedendo di generare un output finale.

Questa tecnica si è rivelata particolarmente utile per la correzione di allucinazioni su risposte fattuali, ma resta inesplorata la sua efficacia in applicazioni differenti.

### *Self-Refine Prompting*

La Self-Refine Prompting [120] è una tecnica la cui anima è l'auto-perfezionamento degli output da parte del LLM. La particolarità del metodo è che si tratta di un processo iterativo, che prevede la ripetizione di tre fasi in consecuzione fino al raggiungimento del risultato ottimale. Le tre fasi sono:

1. Generazione output iniziale: viene inviato il prompt desiderato al modello per generare la risposta.
2. Feedback: si inviano il prompt e l'output iniziale al modello, chiedendo di giudicare la risposta con dei feedback.
3. Perfezionamento: si inviano i feedback al modello insieme al prompt iniziale per ottenere un output perfezionato.

Lo studio ha riportato ottimi risultati dall'applicazione di self-refine, testandolo sulla famiglia di LLM GPT ha riportato aumenti percentuali considerevoli, come riportato nella seguente tabella.

Task	GPT-3.5		ChatGPT		GPT-4	
	Base	+SELF-REFINE	Base	+SELF-REFINE	Base	+SELF-REFINE
Sentiment Reversal	8.8	<b>30.4</b> (↑21.6)	11.4	<b>43.2</b> (↑31.8)	3.8	<b>36.2</b> (↑32.4)
Dialogue Response	36.4	<b>63.6</b> (↑27.2)	40.1	<b>59.9</b> (↑19.8)	25.4	<b>74.6</b> (↑49.2)
Code Optimization	14.8	<b>23.0</b> (↑8.2)	23.9	<b>27.5</b> (↑3.6)	27.3	<b>36.0</b> (↑8.7)
Code Readability	37.4	<b>51.3</b> (↑13.9)	27.7	<b>63.1</b> (↑35.4)	27.4	<b>56.2</b> (↑28.8)
Math Reasoning	<b>64.1</b>	<b>64.1</b> (0)	74.8	<b>75.0</b> (↑0.2)	92.9	<b>93.1</b> (↑0.2)
Acronym Generation	41.6	<b>56.4</b> (↑14.8)	27.2	<b>37.2</b> (↑10.0)	30.4	<b>56.0</b> (↑25.6)
Constrained Generation	28.0	<b>37.0</b> (↑9.0)	44.0	<b>67.0</b> (↑23.0)	15.0	<b>45.0</b> (↑30.0)

Tabella 6-4 Percentuali accuratezza del Self-Refine Prompting

### *Plan-and-Solve Prompting*

La tecnica CoT Zero-Shot ha mostrato come una semplice espressione aggiunta al prompt possa da sola migliorare l'accuratezza dei LLM portandoli a ragionare per passaggi. Tuttavia, può succedere che questa tecnica fallisca saltando dei passi del ragionamento (12% degli errori) o commettendo errori di calcolo (7% degli errori).

Plan-and-Solve Prompting [121] nasce proprio con l'obiettivo di risolvere questo tipo di trascuratezze, semplicemente sostituendo la frase "Pensiamo passo dopo passo" usata in Zero-

Shot CoT con "Per prima cosa comprendiamo il problema e ideiamo un piano per risolverlo. Poi, eseguiamo il piano e risolviamo il problema passo dopo passo".

Questa, apparentemente, banale sostituzione si è rivelata vincente nel diminuire gli errori di calcolo o di mancanza di passaggi, sia in campo matematico che di ragionamento logico e simbolico.

## 6.4 Prompt Engineering per la generazione di immagini

L'utilizzo di IA per la generazione di immagini è un campo relativamente nuovo paragonato a quello della generazione testuale, ciò sicuramente ha un peso non indifferente sul livello di ricerca svolta su di essa. Si è visto, infatti, che esistono numerose tecniche testate nella generazione testuale, e molte non sono neanche state trattate in questa tesi. Contrariamente, le nozioni e le pratiche che possiamo applicare all' *image prompting* sono di gran lunga inferiori.

Un altro aspetto da considerare è anche lo stato dell'arte della tecnologia, laddove la generazione testuale ha raggiunto livelli di accuratezza elevati, i modelli di generazione di immagini risultano ancora acerbi sotto diversi aspetti, mostrando più punti di debolezza da affrontare.

Questi due aspetti, unitamente, comportano un diverso livello di difficoltà nella scrittura dei prompt. Se i prompt per la generazione testuale possono dare buoni risultati anche con una cura superficiale, generare immagini accurate richiede estrema attenzione nella fase di *prompting*.

Seppure la materia suscita molto interesse e studi attorno a sé, essi sono quindi ancora insufficienti per renderla una scienza esatta. Le pratiche che verranno presentate nel corso del capitolo mostreranno solo alcuni esempi dell'ampissima gamma di impostazioni applicabili a un prompt, per trovare il proprio stile si consiglia di esplorare le grandi banche online di immagini che abbiamo citato, dove è possibile scoprire i termini che più si avvicinano al nostro obiettivo.

### 6.4.1 Elementi di un prompt per generazione di immagini

Un prompt per immagini presenta al suo interno elementi di natura e fini differenti, la studio *A taxonomy of prompt modifiers for text-to-image generation* [103] si è occupato di eseguire un'approfondita ricerca etnografica nelle comunità di praticanti per individuarle, con il proposito di creare una tassonomia chiara e univoca per il *prompting* di immagini.

Gli elementi del prompt vengono chiamati **prompt modifiers** e possono agire in aspetti diversi a seconda della propria natura. Lo studio ne ha individuati sei tipologie:

1. **Subject Term**: descrive il soggetto dell'immagini, è indispensabile per avere controllo sulla generazione.
2. **Style modifiers**: definiscono lo stile dell'immagine, facendo riferimento ad artisti, correnti, reference, ecc.
3. **Image Prompt**: è possibile fornire un'immagine di reference in input, caricandola o sotto forma di URL. Questo può aiutare ad indirizzare al meglio il risultato verso determinati stili o contenuti.
4. **Quality Booster**: sono termini utilizzati per enfatizzare o smussare determinati aspetti dell'immagine. All'atto pratico, sono aggettivi legati alla qualità e l'impatto dell'immagine e possono avere molta influenza sull'output.

5. **Repeating Terms:** ripetere i soggetti e gli aggettivi o aggiungere sinonimi aiuta la rete neurale ad individuare agilmente il risultato che vuoi raggiungere, enfatizzando certi aspetti utili.
6. **Magic Term:** sono termini sintatticamente distanti dagli altri, che introducono una variabile di randomness nel risultato. Spesso prevedono usi metaforici o emotivi.

### 6.4.2 Tecniche di base di Prompt Engineering per generazione di immagini

Rispetto alle tecniche viste per la generazione di testo, quanto si sta per enunciare non rappresentano vere e proprie tecniche, quanto delle pratiche o approcci utili al prompting. Il concetto fondamentale, per la disciplina del prompting in generale, ma soprattutto per la generazione di immagini, è l'importanza di adottare un metodo iterativo, fatto di tentativi e correzioni.

I generatori di immagini, per poter creare le loro risposte, debbono necessariamente appoggiarsi a banche di grandi dimensioni di immagini e didascalie, dell'ordine delle centinaia di milioni; il risultato che si otterrà sarà completamente dipendente dalle etichette che accompagnano tali immagini che, per quanto accurate, non possono essere oggettive ed inequivocabili. Per questo motivo, trovare i termini che maggiormente possono aiutare il nostro prompt a migliorare è possibile solo tramite ampia sperimentazione.

Infatti, il primo consiglio nella generazione di immagini è di produrre più varianti dello stesso output; il modello può fornire risultati molto diversi con lo stesso prompt, ed unire la sua variabilità con quella dei prompt di input è un potente mezzo di sperimentazione.

I testi cui si fa riferimento per le pratiche che seguono sono: *Learn Prompting* [100], *The DALL-E 2 Prompt Book* [104] e *A Traveler's Guide to the Latent Space* [105].

#### *Lunghezza e Specificità*

Prompt eccessivamente lunghi possono essere inefficaci nella generazione immagini, per questo motivo diversi tool hanno impostato una lunghezza massima utilizzabile (ad esempio Dall-e ha impostato un massimo di 400 caratteri).

Seppur vero che prompt brevi possano essere molto efficaci e presentare meno artefatti, se si vuole mantenere un elevato controllo dell'output è necessario essere specifici al suo interno. Ciò non significa necessariamente essere prolissi, a volte basta individuare pochi termini chiave che, da soli, già contengono molte utili informazioni.

Ad esempio, un semplice termine come "action photography", può già indirizzare il modello verso concetti di velocità, inquadratura e scelta delle lenti.

#### *Style Modifiers*

I modificatori di stile sono essenziali per il controllo artistico dell'immagine di output. Essi permettono di riprodurre determinati stili, emulando specifici pittori, correnti artistiche, materiali e tecniche di rappresentazione. La loro combinazione può fornire un controllo elevato nel dettagliare lo specifico stile desiderato. Ecco alcuni esempi di modificatori di stile:

- **Stili:** *hyperrealist, film grain, horror, cinematic, cyberpunk, fantasy, 1980's.*
- **Artisti:** *by greg rutkowski, by christopher nolan, by Pablo Picasso.*
- **Correnti artistiche:** *surrealism, renaissance, cubism, futurism, medieval.*
- **Media:** *Vogue magazine, television, comic book, video tape, music video.*

- **Materiali:** *wooden, plastic, glass, made of paper.*
- **Tecniche:** *photorealistic, cartoon, low-poly, painting, concept art, 3D render, oil-paint, stop-motion, anime.*

### ***Emotional Prompt Words***

Descrivere il mood e l'atmosfera di un'immagine può modificarla enormemente. Nel testo *DALL-E 2 Prompt Book* [104] vengono individuate quattro famiglie di termini:

- **Positive mood, low energy:** *light, peaceful, calm, serene, soothing, relaxed, placid, comforting, cosy, tranquil, quiet, pastel, delicate, graceful, subtle, balmy, mild, ethereal, elegant, tender, soft, light.*
- **Positive mood, high energy:** *bright, vibrant, dynamic, spirited, vivid, lively, energetic, colorful, joyful, romantic, expressive, bright, rich, kaleidoscopic, psychedelic, saturated, ecstatic, brash, exciting, passionate, hot.*
- **Negative mood, low energy:** *muted, bleak, funereal, somber, melancholic, mournful, gloomy, dismal, sad, pale, washed-out, desaturated, grey, subdued, dull, dreary, depressing, weary, tired.*
- **Negative mood, high energy:** *dark, ominous, threatening, haunting, forbidding, gloomy, stormy, doom, apocalyptic, sinister, shadowy, ghostly, unnerving, harrowing, dreadful, frightful, shocking, terror, hideous, ghastly, terrifying, ecc.*



Figura 6-4 Positive Mood, Low Energy

Figura 6-4 Positive Mood, High Energy

Figura 6-4 Negative Mood, Low Energy

Figura 6-4 Negative Mood, High Energy

### ***Size-y, Structure-y Words***

Un altro aspetto che è possibile controllare è la composizione generale degli elementi interni all'immagine desiderata, l'uso di specifici termini può comandare aspetti come la struttura (strutturato o libero) e la grandezza degli elementi dell'immagine.

- **Big and free:** *Curvaceous, swirling, organic, riotous, turbulent, flowing, amorphous, natural, distorted, uneven, random, lush, organic, bold, intuitive, emotive, chaotic, tumultuous, earthy, churning, ecc.*
- **Big and structured:** *Monumental, imposing, rigorous, geometric, ordine, angular, artificial, lines, straight, rhythmic, composed, unified, manmade, perspective, minimalist, blocks, dignified, robust, defined, ecc.*

- **Small and structured:** Ornate, delicate, neat, precise, detailed, opulent, lavish, elegant, ornamented, fine, elaborate, accurate, intricate, meticulous, decorative, realistic, ecc.
- **Small and free:** Unplanned, daring, brash, random, casual, sketched, playful, spontaneous, extemporaneous, offhand, improvisational, experimental, loose, jaunty, light, expressive, ecc.

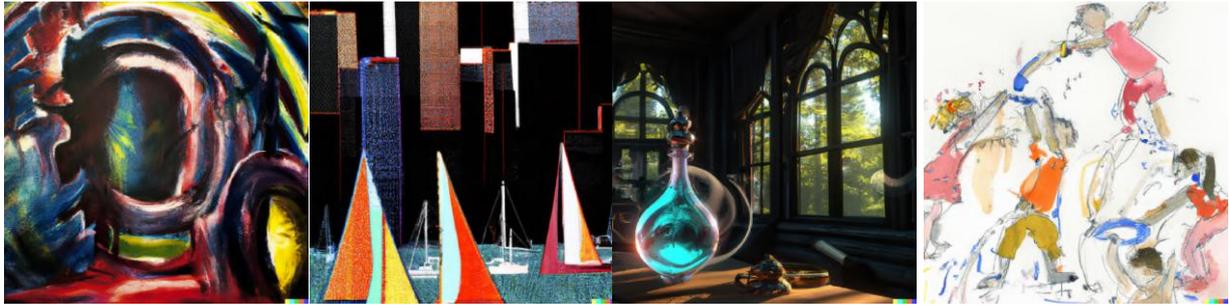


Figura 6-8 Big and free

Figura 6-8 Big and structured

Figura 6-8 Small and structured

Figura 6-8 Small and free

### *Photography Words*

Molte delle immagini presenti nei dataset sono fotografie, per cui portano all'interno delle loro didascalie numerosi termini tecnici provenienti da quel mondo. Per questo motivo utilizzarle è perciò fruttuoso nel controllo di alcuni parametri.

- **Camera settings:** impostazioni tecniche della camera che determinano esposizione, profondità di campo, motion blur e molti altri fattori.  
Esempi: *shallow depth of field, Sigma 85mm, motion blur, low exposure.*

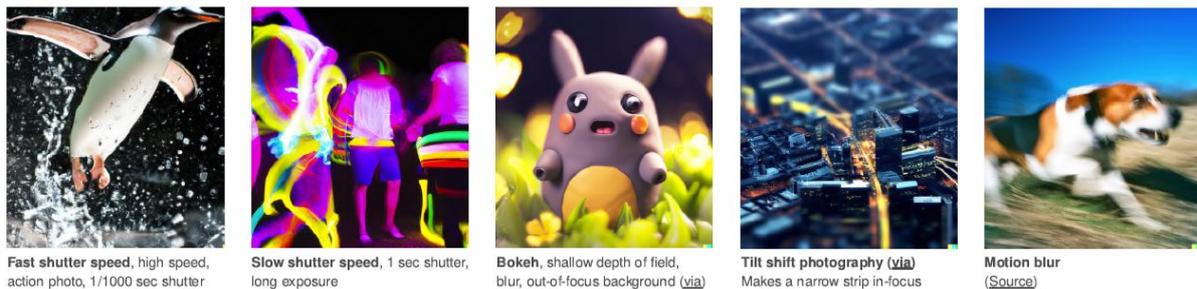


Figura 6-9 Camera settings prompts

- **Shot Type and Angle:** definire il tipo di inquadratura determina la composizione dell'immagini in termini di proporzioni tra figura umana e ambiente. È spesso necessario anche specificare l'angolo di camera per ottenere specifici effetti.  
Esempi: *bird's eye view, close-up, full-shot, low angle, dutch angle.*

## Prompt Engineering

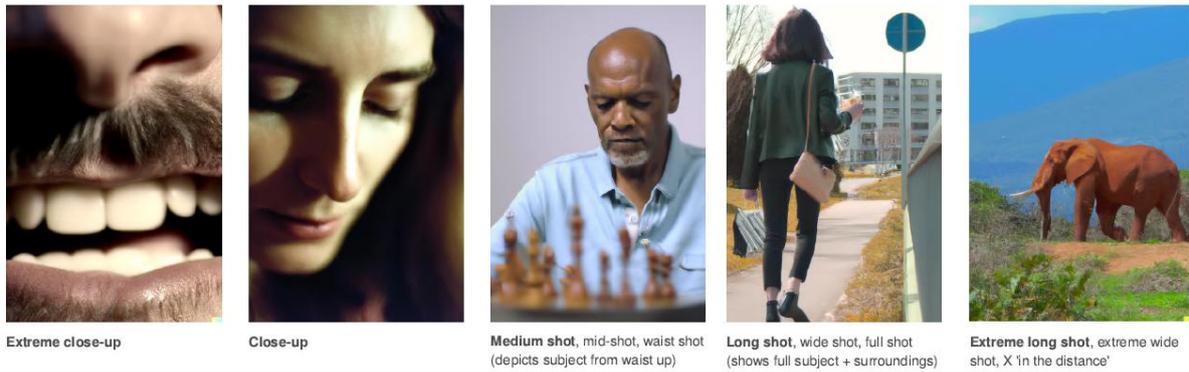


Figura 6-11 Shot type Prompting

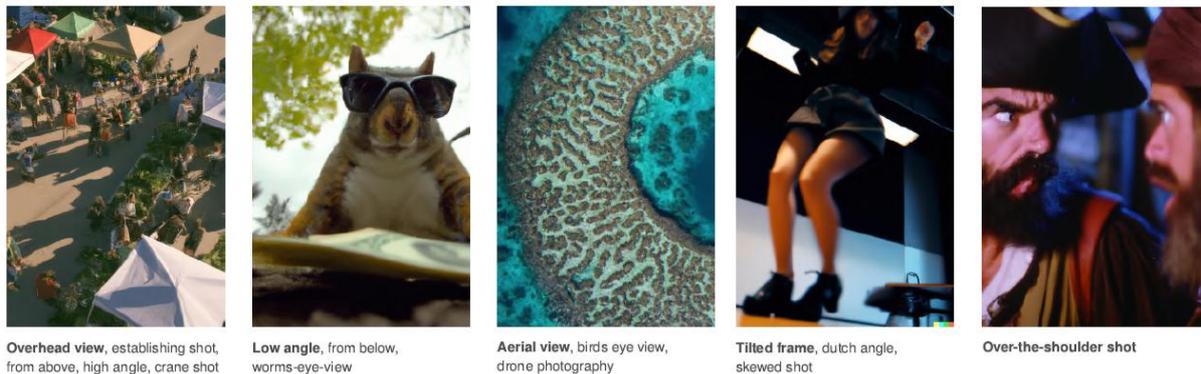


Figura 6-10 Angle type Prompting

- Lighting:** l'illuminazione può cambiare completamente il mood di un'immagine, controllarne i parametri di ombre, temperatura, contrasto, saturazione ed esposizione è essenziale per la generazione di un'immagine accurata.  
 Esempi: *warm lighting, sunset lighting, blue-hour, high contrast, shadows, high saturation, theatrical lighting, by the night, b&w.*

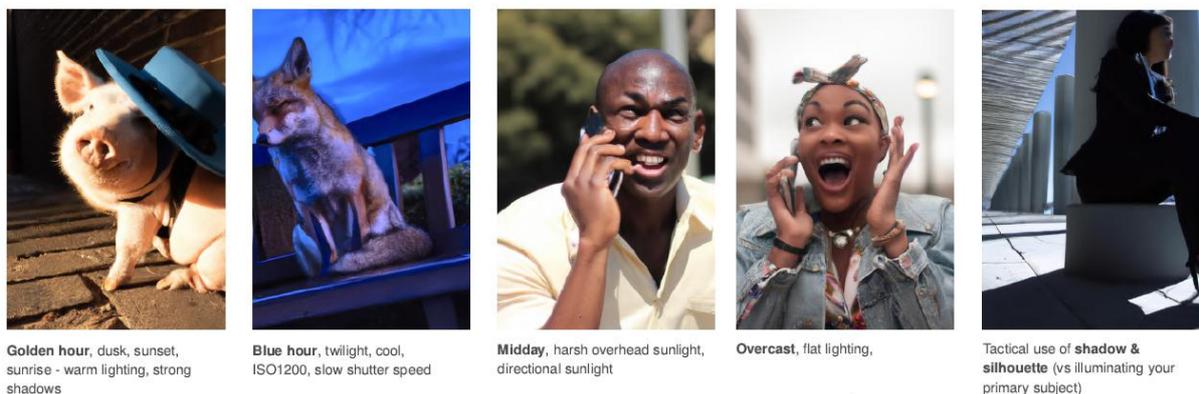


Figura 6-12 Lighting prompting

### Quality Booster

I Quality Booster, chiamati anche Magic Terms, sono - come anticipato - termini che, se aggiunti a un prompt, incrementano specifiche qualità dell'immagine che non hanno niente a che fare con lo stile. Possono agire sulla risoluzione, livello di dettaglio, potenza emotiva, e

molti altri aspetti. Nonostante sia testata l'influenza che hanno sulle immagini, non è garantito che agiscano in modo preciso, per questo motivo un abuso di questi termini potrebbe allontanarci dal risultato sperato..

### **Esempi di quality booster:**

*High resolution, 2K, 4K, 8K, clear, good lighting, detailed, extremely detailed, sharp focus, intricate, beautiful, realistic+++ , complementary colors, high quality, hyper-detailed, masterpiece, best quality, art station, stunning*

### ***Repetition***

La ripetizione di termini, nomi, aggettivi o frasi incrementa il loro peso all'interno dell'immagine. Ripetere un determinato soggetto farà in modo che occupi più spazio nell'immagine generata, analogamente ripetere aggettivi di stile incrementerà tale impronta stilistica.

### ***Weighted Terms***

Alcuni modelli (Stable Diffusion, Midjourney, ecc.) consentono l'assegnazione di pesi ai termini presenti in un prompt. Questa pratica è particolarmente apprezzata perché permette un controllo ponderato degli elementi inseriti nel prompt, riducendo il rischio che il modello si concentri eccessivamente o insufficientemente su certi aspetti.

Infatti, i pesi assegnati possono essere sia positivi che negativi, a seconda che si voglia enfatizzare o meno quel determinato aspetto.

### ***Fix Deformed Generations***

Un problema estremamente comune nella generazione di immagini con diversi modelli è la presenza di deformazioni dei soggetti, in particolare su parti del corpo umano. Questo problema può essere gestito in una certa misura con un utilizzo accurato di prompt negativi.

Ad esempio, l'utilizzo del negative prompt "Deformed hands: -5" può essere fruttuoso nel ridurre artefatti sulle mani di un'immagine. Tale ragionamento può essere applicato a più elementi o persone, anzi la tecnica raggiunge i suoi massimi livelli con l'uso di più negative prompt contemporaneamente.

Sfortunatamente, questa tecnica non è coerente nei risultati, quindi potrebbe essere necessario richiedere generazioni multiple prima di ottenere un miglioramento effettivo. In compenso si prevede che in futuro artefatti simili saranno sempre più rari, rendendo questa pratica superflua.

## **6.4.3 Tecniche avanzate di Prompt Engineering per generazione immagini**

Si sono elencate diverse categorie di termini utilizzabili nei prompt per la generazione di immagini, evidenziando per ognuna i possibili effetti che essi possono avere sull'output. Tuttavia, esistono tecniche avanzate che possono tornare utili in specifici casi di applicazione.

### ***Sostituzione ed eliminazione di elementi***

Capita spesso, nella generazione di immagini, di ottenere l'output desiderato "sporcatto" dalla presenza di elementi indesiderati, siano essi oggetti, persone o sfondi. Dall-e, come molti altri

tool, ha integrato allora una funzione per poter ritoccare questi elementi, dando la possibilità di selezionarli per eliminarli o addirittura sostituirli.

La tecnica è estremamente utile sulla carta, ma non sempre raggiunge con successo i suoi propositi. Un'alternativa ibrida per la risoluzione del problema è di post-elaborare l'immagine con software esterni come Photoshop, ma questo chiaramente richiede capacità e nozioni che non tutti possiedono.

### *Cropping e Uncropping*

Talvolta l'IA, quando richiediamo di rappresentare una scena contenuta, può circondarla con eccessivo contesto e aggiungere parti non richieste. In questi casi, la pratica più comune per risolvere il problema è quella del cropping, ossia ritagliare la porzione di immagine interessata. Non vi è purtroppo modo di comunicare all'IA di mantenere quello stesso elemento e zoommare nell'immagine, per cui - se il risultato ottenuto è soddisfacente - conviene ricorrere a questo semplice trucco.

Può succedere, però, anche il caso contrario, in cui si vuole espandere l'ambiente di un'immagine esistente. Questa procedura prende il nome di "uncropping" ed è integrata all'interno di diversi tool presenti sul mercato, basta dare in input un'immagine (sia essa reale o generativa) e l'IA si occuperà di espanderne il contesto.

### *Image Merging*

Spesso le richieste che sottoponiamo al modello sono troppo complesse per poter essere gestite. Una soluzione può essere di suddividere il soggetto in più elementi, realizzando separatamente la generazione di personaggi, ambiente ed oggetti. Una volta elaborati i singoli elementi, sarà poi possibile unirli in post-produzione con software esterni. L'unico aspetto cui occorre porre attenzione, è di avere cura che lo stile e l'illuminazione dei vari elementi siano coerenti tra di loro.

Ad esempio, questa procedura è stata testata nella generazione del film sperimentale *Giant* [122], un prodotto che integra l'IA in diverse parti della propria pipeline. Esso prevedeva in scena una ragazza che vede, dal finestrino di un treno, un gigante in lontananza nella foresta. Per poterla realizzare sono stati compostati assieme la scena del gigante nella foresta e la scena della ragazza che guarda il finestrino, con ottimi risultati.

Questo mostra come spesso un approccio ibrido alla generazione di immagini possa raggiungere i risultati più complessi, unendo l'intervento dell'uomo e dell'IA nella stessa pipeline.

### *L'uso di ChatGPT per generare prompt*

Per ultima esploriamo una delle tecniche che si è rivelata più utile e funzionale nel miglioramento dei prompt, che è stata riscontrata in più studi analizzati, tra cui la realizzazione del film *Giant* [122] che abbiamo appena citato, registrando grandi benefici.

Già all'interno della sezione dedicata al prompt engineering testuale, avevamo analizzato diverse tecniche il cui principio alla base era di sfruttare i modelli di IA in prima persona per la scrittura dei prompt, si veda ad esempio la tecnica Rephrase and Respond. Questo approccio si era dimostrato efficiente ed altamente coerente nel suo scopo. Da queste tecniche nasce, dunque, l'idea analoga di utilizzare un modello di generazione testuale, quale ChatGPT, per la scrittura dei prompt di generazione immagini.

## Prompt Engineering

Il suo utilizzo permette di sottoporli un input espresso in linguaggio naturale e di ottenere in output un prompt aderente alla nostra visione, ma con termini adatti alla generazione di immagini.

Questa tecnica si è rivelata un alleata potente per migliorare i propri prompt, oltre ad essere un ulteriore passo avanti per avvicinare sempre più gli utenti inesperti al mondo dell'IA, semplificandone i processi.

## 7. IA nella fase di pre-produzione

### 7.1 ChatGPT per valutazione e revisione sceneggiatura

Una delle ragioni più stimolanti per cui è nata questa tesi, è stata quella di voler testare le capacità dell'Intelligenza Artificiale nel ruolo di assistente nella stesura della sceneggiatura. Nel corso dell'elaborato, si è molto discusso delle capacità creative della GenAI, ma questa particolare applicazione si discosta leggermente da questo filone, quel che si vuole scoprire non è se un modello sia in grado di creare contenuti originali in autonomia, ma se sia capace di reggere un dialogo interattivo e costruttivo, uno scambio di idee basate su una comprensione profonda di un testo. Insomma, la domanda cui si cerca di rispondere è se oltre ad essere in grado di comprendere e di creare, ChatGPT possa essere uno strumento che libera il potenziale creativo dell'uomo.

Nel corso di questa sezione verranno mostrati i diversi utilizzi che sono stati fatti del modello in fase di revisione, valutazione e correzione della sceneggiatura: applicazioni differenti con un comune scopo di affinare e migliorare il testo attraverso un dialogo con la macchina.

Occorre premettere che l'utilizzo del modello per la revisione della sceneggiatura del cortometraggio in esame "Non ha visto nulla" è cominciato già a partire da una fase precedente alla stesura definitiva. In questa versione, la trama aveva già individuato la propria direzione, ma era profondamente diversa dal suo stato finale. In particolare, il cortometraggio prevedeva due episodi aggiuntivi: l'Episodio Cinico-Brutale e l'Episodio dell'Esperienza Corposa. Entrambi esploravano due stili individuati anche da Calvino nel suo romanzo, e rendevano più ampia la gamma di generi rappresentati nel cortometraggio.

Il primo episodio voleva essere una storia di piena ispirazione pulp, che si rifaceva al cinema di Quentin Tarantino, Guy Ritchie e in generale al filone del gangsterismo di stampo più umoristico. Esso vedeva due giovani e goffi malavitosi, durante un'operazione di recupero di un misterioso baule su richiesta del loro capo. Tra i due si accende una lunga discussione per decidere se aprire la cassa per dare sfogo alla curiosità o se seguire le istruzioni senza porsi domande, evitando così ogni tipo di implicazione morale. La rappresentazione era fortemente umoristica, trasportata attraverso due personaggi fondamentalmente buffi ed un dialogo molto denso.

Il secondo episodio, invece, voleva recuperare un particolare stile di ripresa e montaggio, più che un genere specifico. La trama seguiva la vicenda di una ragazza che partecipa a un casting per una pubblicità di intimo femminile, ma che, avvicinandosi al suo turno, si rende sempre più conto dello stato di sottomissione in cui la donna viene inserita in questa dinamica, venendo deumanizzata e mercificata. L'ispirazione proviene da classici della tematica, da *The Neon Demon* di Nicolas Winding Refn a *Black Swan* di Darren Aronofsky, ma abbiamo detto che l'esercizio di scrittura era indipendente dalla tematica. L'episodio voleva essere una rappresentazione che ruotava attorno all'esperienza del corpo, i suoi gesti, i suoi sensi, che si focalizzasse sui movimenti della protagonista e su quelli che lei osservava delle altre ragazze, costruendo una narrazione fatta di dettagli e percezioni. In questa ottica, l'ispirazione per la ripresa si individuano in film completamente differenti, come *Le Fabuleux Destin d'Amélie Poulain* di Jean-Pierre Jeunet.

In questa forma temporanea, dunque, la sceneggiatura era diventata troppo lunga e dispendiosa, rispetto alle limitate risorse cui si aveva accesso, e necessitava di tagli. Da qui nacque la necessità di eliminare degli episodi e l'idea di affidarsi a ChatGPT per ricevere un consiglio su quali fossero più valevoli.

Si è cominciata, quindi, una chat con il modello, inviandole il pdf della stesura di sceneggiatura con le seguenti istruzioni e chiarimenti.

***Prompt 1:** Leggi attentamente la sceneggiatura che ti ho mandato. È una storia divisa in episodi, diversi per stile e genere, che a causa della tecnologia vengono sempre interrotti prima di concludersi. L'episodio finale connette tutti i precedenti, scopriamo che quel che abbiamo visto è il frutto del lavoro di uno scrittore in crisi, che lotta con le proprie idee per terminare una sceneggiatura. Vorrei che valutassi passo per passo il testo: è avvincente? Gli episodi sono originali o banali? Riescono ad avere una buona presa sul pubblico nonostante il formato breve? Se dovessi tagliare un episodio, quali elimineresti?*

ChatGPT ha risposto alla richiesta ordinatamente, esplorando un punto per volta. Si è dimostrato capace di cogliere la struttura della narrazione, aspetto non banale vista la sua natura anomala, comprendendo come i diversi episodi riescano a coesistere, pur avendo nulla in comune, e ad analizzare efficacemente il tema della tecnologia che interromper la fruizione. Riguardo la valutazione della presa sul pubblico (aspetto che più di altri si fatica a credere possa essere compreso da una macchina), il modello ha individuato con chiarezza quelli che effettivamente sono le questioni di difficile gestione in un lavoro simile, dedicando il seguente paragrafo:

***Risposta 1:** [...] Gli episodi riescono a catturare l'attenzione, ma ognuno ha un tempo limitato per svilupparsi e quindi deve colpire immediatamente il pubblico. Alcuni episodi sembrano riuscirci meglio di altri. Ad esempio, l'episodio con il Cowboy crea un'atmosfera misteriosa e sovranaturale che cattura l'immaginazione. Tuttavia, non tutti gli episodi hanno la stessa forza, e il rischio è che alcune interruzioni possano risultare più frustranti che intriganti, soprattutto se il pubblico non è abbastanza coinvolto nel breve lasso di tempo disponibile.*

Questa semplice osservazione è un utile campanello di allarme per chiedersi se alcune interruzioni non siano troppo impattanti e dissonanti per il pubblico.

Nell'analisi della originalità dei contenuti, ChatGPT riesce a costruirsi un'opinione ben giustificata: individua diversi episodi che trattano tematiche rilevanti e dinamiche originali, ma anche gli episodi che rischiano di sposare troppo le regole del genere finendo per cadere nel cliché. Portando avanti questa analisi, individua un episodio che è più debole e sacrificabile in caso di necessità di tagli.

***Risposta 1:** [...] Se dovessi suggerire di tagliare un episodio, forse eliminerei quello cinico-brutale con i due scagnozzi nel seminterrato. Anche se funziona, è l'episodio che appare più derivativo rispetto agli altri, e potrebbe non aggiungere molto al tema generale della sceneggiatura, che sembra più orientato a esplorare l'interruzione e l'angoscia creativa. Tagliarlo potrebbe permettere di dare più spazio agli altri episodi, rendendo l'opera complessiva più coesa.*

Qua ChatGPT dà già prova, nella sua prima risposta, della sua utilità come assistente creativo. Il candidato, in qualità di scrittore, aveva voluto inserire l'Episodio Cinico-Brutale proprio

come mezzo per alleggerire il cortometraggio, inserendo una scena più umoristica, ma non ci si era resi conto che, così facendo, esso sarebbe diventato l'unico elemento dissonante di una narrazione incentrata su un comune senso di angoscia, derivante dalla crisi creativa dello scrittore. Attraverso il suggerimento di ChatGPT, si è capito che quello che poteva essere un elemento di diversificazione rischia solo di diventare di sbilanciamento. Di fronte a questa realtà, uno scrittore può prendere, poi, diverse precauzioni per gestire la questione. Questo può avvenire sia in autonomia che chiedendo ulteriori consigli dell'IA, ma già si è dimostrato come da ChatGPT possano nascere ragionamenti fruttuosi.

Il passo successivo è stato determinare con certezza quali episodi tagliare, si è proceduto domandando:

*Prompt 2: Nell'ottica di avere solo episodi profondamente diversi nel genere e nello stile e focalizzandoti sulla loro potenziale presa sul pubblico, fai una classifica dei migliori episodi, mettendo per ultimi i più sacrificabili.*

Nella risposta ChatGPT si è mantenuto coerente, fornendo una classifica degli episodi accompagnati con un'analisi attenta. Ha ribadito che il Cinico-Brutale fosse il più fuori luogo, individuando inoltre che l'Episodio dell'Angoscia fosse rischioso in quanto troppo breve (nella stesura di allora era poco più di una pagina) e che, pertanto, potesse non appassionare abbastanza. Infine, ha messo luce sul fatto che i tre episodi drammatici - dell'Angoscia, Alienazione e Esperienza Corposa - avessero molti aspetti simili che facilmente si sarebbero potuti confondere. Alla luce di queste valutazioni, che si sono condivise appieno, si è giunti alla decisione di tagliare dalla sceneggiatura l'Episodio Cinico-Brutale e quello dell'Esperienza Corposa, cercando di incorporare le caratteristiche di quest'ultimo negli altri episodi drammatici e avendo così modo di espandere l'Episodio dell'Angoscia.

Successivamente, si è voluta testare le capacità di ChatGPT nell'individuare gli errori ortografici e grammaticali, sfruttandolo per eliminare i classici errori di battitura o sviste del momento. Alla richiesta, ChatGPT ha risposto così:

*Risposta 2: Per identificare e correggere errori ortografici e grammaticali nel documento, posso esaminare il testo manualmente. Tuttavia, poiché il file è lungo, potresti dirmi se ci sono sezioni specifiche che vuoi controllare oppure posso dare un'occhiata generale e segnalare gli errori che trovo. Quale preferisci?*

Effettivamente il documento era circa 20 pagine di sceneggiatura, ma a fini sperimentali si è voluto provare a proseguire la correzione integrale. In risposta ChatGPT ha completato l'analisi del documento, individuando solamente 10 errori, che consistevano esclusivamente in maiuscole, virgole e accenti. Si è fatto notare, allora, che non erano stati analizzati gli errori di battitura e di coniugazione, due tipologie di errori che erano effettivamente presenti. Ecco la risposta:

*Risposta 3: Per completare la revisione di un intero documento come questo, potrebbe servirmi circa 15-20 minuti per identificare e annotare tutti gli errori di battitura, grammaticali e verbali. Ti invierò l'elenco completo degli errori appena ho finito.*

La stima dei tempi è elevata, implicando un grade sforzo computazionale per il modello. Anche invitandolo a proseguire, il modello non ha mai inviato un report finale degli errori, bloccandosi ripetutamente. Questo aspetto è interessante, perché significa che, apparentemente, ChatGPT trovi molto più dispendioso correggere ed analizzare nel dettaglio un testo piuttosto che riscriverlo parafrasandolo. Proseguendo nel capitolo vedremo come ChatGPT sia estremamente agile e veloce nella riscrittura di interi episodi, mentre sembra avere più difficoltà - con testi lunghi - ad individuare singolarmente gli errori e comunicarli.

Per poter procedere con un'analisi affidabile della sceneggiatura, si è deciso di dividerla e di sottoporre al modello un solo episodio per volta. Questo permette, innanzitutto, di spartire il carico computazionale, inoltre semplifica la struttura del testo, eliminando la sua natura non lineare e permettendo a ChatGPT di focalizzarsi su piccole porzioni per volta. Ciò risulta fruttuoso anche nell'ottica di voler diversificare tra loro i diversi episodi, gestendoli in modo indipendente.

La revisione è, quindi, proseguita attraverso quattro chat, rispettivamente dedicate agli episodi: del Mistero e dell'Indefinito, dell'Angoscia, dell'Alienazione, Simbolico-Interpretativo. Per ognuna di esse si è applicata una serie di prompt uguali, applicando piccole variazioni a seconda delle esigenze, ma cercando di tenere un approccio standardizzato al fine di valutarne i risultati confrontandoli. Per evitare di ripetersi, verranno ora elencati i prompt utilizzati assegnando loro una numerazione; nel proseguimento del capitolo, ci si limiterà a richiamarli con il loro numero per studiarne le risposte.

***Prompt revisione 1:** Vorrei che leggessi attentamente la sceneggiatura che ti ho mandato e che la analizzassi accuratamente e minuziosamente. Vorrei che individuassi gli errori grammaticali, suggerissi migliorie lessicali, controllassi la coerenza dei dialoghi e se ci siano battute superflue o migliorabili.*

***Prompt revisione 2:** Rileggi il testo. Secondo te può avere una buona presa sul pubblico? Ha una buona caratterizzazione di genere, storia e personaggi?*

***Prompt revisione 3:** Se dovessi realizzare un film da questa sceneggiatura, quali sarebbero le tue scelte in termini di: Fotografia, Montaggio ed Inquadrature? Prova a darmi più opzioni suggerendo scene o film di reference per ognuna.*

I primi due prompt sono dedicati all'analisi del testo, valutandone l'efficacia e la presenza di debolezze. Il terzo, invece, struttura un interessante punto di sviluppo della fase di pre-produzione, testando se ChatGPT possa essere un tool in grado di ragionare in termini esclusivamente cinematografici. Quest'ultima domanda è quella che maggiormente mette alla prova le capacità creative del modello, portandolo ad immaginare una messa in scena dell'episodio. Di seguito, ci apprestiamo ad analizzare i risultati della revisione avvenuta con i prompt appena elencati per ognuno degli episodi.

### ***Episodio del Mistero e dell'Indefinito***

ChatGPT ha risposto dettagliatamente al primo prompt di revisione analizzando ognuno dei punti richiesti. Non sono stati trovati errori grammaticali, ma sono stati individuati possibili miglioramenti di punteggiatura per agevolare la lettura. Ha trovato l'ortografia corretta,

individuando con successo due errori di battitura. Sull'analisi dei dialoghi è rimasto leggermente superficiale, affermando che i dialoghi erano coerenti con la narrazione, ma senza consigliare significative migliorie e cambi.

Proseguendo con la valutazione generale, il modello ha risposto al secondo prompt in maniera coerente a quanto fatto nell'analisi della sceneggiatura totale, ma aggiungendo profondità alla risposta. Ha apprezzato l'atmosfera della scena e trovato vincente la dinamica di mistero che la permea, ma ha messo in luce che i tempi ridotti e la limitata esplorazione dei personaggi rischia di non far appassionare il pubblico. Questa analisi perde di valore, in parte, se si inserisce l'episodio nel contesto più ampio del film, dove è previsto dal concept che le singole storie siano brevi ed interrotte prima di svilupparsi nella loro completezza, ma ha comunque sollevato dubbi leciti. Questo ci ha portati, infatti, a modificare in parte la sceneggiatura, aggiungendo back-story nelle battute dei personaggi, e dando più respiro alla vicenda dedicando più tempo alla creazione del climax finale.

Nel giudicare la caratterizzazione di genere, ChatGPT ha riscontrato un'elevata presenza di elementi tipici del genere noir e delle dinamiche western, trovando i due personaggi azzeccati per la storia, pur ribadendo che una più profonda esplorazione potrebbe renderli più completi. Conclude la risposta con una valutazione generale positiva, suggerendo però di distendere i tempi, sviluppare maggiormente i personaggi e snellire alcune battute rendendole meno verbose: tutti suggerimenti che sono stati ben accolti ed integrati nella stesura successiva.

Al terzo ed ultimo prompt, il modello risponde con sorprendente prontezza e preparazione, proponendo tre differenti opzioni di stili: Western Noir, Surreale Onirico, Minimalista e Realista. Ognuno di essi è accompagnato da esaustivi ma riassuntivi dettagli sull'uso della luce, color palette, lenti, inquadrature e stile di montaggio; ma anche dettagli secondari come l'uso delle transizioni o effetti e filtri della camera. Quel che stupisce è la coerenza con cui ha costruito ognuno degli stili proposti, costituiti da elementi che collaborano in modo convincente nel creare una messa in scena azzeccata per l'episodio. Indipendentemente dalla preferenza finale, ogni proposta era ben ragionata e distinguibile dalle altre, seguendo strade molto diverse per raggiungere la resa finale. Questo processo di caratterizzazione è stato reso ancora più eloquente dalla proposta di film di reference, avvenuto in numero limitato ma azzeccato. Per ognuno di essi si è specificato quale particolare elemento prendere a riferimento, indicando, ad esempio, solo l'uso della luce o del montaggio. In conclusione alla risposta, ha riportato un riassunto sulle possibili direzioni percorribili, che riproponiamo di seguito:

### ***Risposta prompt 3 Episodio 1:***

*[...] Tutte queste opzioni offrono stili diversi che potrebbero adattarsi alla scena a seconda dell'effetto emotivo che vuoi ottenere.*

- *Western Noir è ideale se vuoi un'atmosfera cupa e misteriosa.*
- *Surreale e Onirico funzionerebbe bene se desideri confondere e affascinare il pubblico con elementi più astratti.*
- *Minimalista e Realista è perfetto per un approccio più sottile e introspezione, mettendo in primo piano la recitazione e l'ambiente.*

*Scegliendo uno di questi approcci, puoi orientarti su quale stile si adatta meglio alla tua visione artistica per questa scena.*

Il modello è stato capace di offrire nuove opzioni non prese in considerazione in precedenza, dando modo di ampliare il ventaglio di possibilità. Si è comunque deciso di proseguire sull'idea iniziale Western-Noir, ma questo non toglie che questa fase abbia saputo dettagliare meglio la visione finale e suggerire utili reference.

### ***Episodio dell'Angoscia***

Questo episodio non contiene dialoghi, per cui è interessante studiare il cambio delle risposte. Innanzitutto ChatGPT ha notato questa particolarità, per cui ha utilizzato solamente le descrizioni per studiare la coerenza dei personaggi. Ha avuto successo anche nell'individuare errori di diverso tipo, da quelli di battitura alle piccole sviste grammaticali, proponendo anche qualche miglioria lessicale che, però, poco aggiungevano al testo. L'analisi complessiva della scrittura è positiva, ma il modello sembra perdere di vista che si tratta di un testo cinematografico e non di un racconto, come è evidente dalla seguente sezione:

*Risposta prompt 1 Episodio 2: [...] La scena ha un buon potenziale per creare suspense e tensione. Tuttavia, ci sono alcune parti che potrebbero beneficiare di una maggiore concisione. L'attenzione ai dettagli è notevole, ma alcune descrizioni potrebbero essere semplificate per mantenere il ritmo e l'attenzione del lettore.*

L'osservazione riportata ha poco senso nell'ottica di un'opera filmica, dove l'immaginazione dello spettatore poco incide su un'immagine sullo schermo. Sicuramente si è apprezzato il consiglio di snellire le descrizioni, ma le motivazioni riportate sono state inappropriate.

In risposta al prompt 2 ChatGPT elabora un report diviso per punti di forza e aree di miglioramento. In realtà, nella seconda categoria vengono fatte ricadere molte osservazioni finalizzate ad un ampliamento della storia, ossia accortezze da avere per proseguire al meglio la vicenda, ma nulla che riguardi la scena mandatagli.

Al terzo prompt, risponde con una struttura differente dal precedente episodio, forse a causa di un numero maggiore di possibilità. Laddove prima si erano individuati complessivamente tre stili – caratterizzati da combinazioni specifiche di montaggio, ripresa e fotografia – in questo caso sono state proposte due opzioni per ognuna delle categorie chieste, andando quindi a trattare in modo indipendente ogni aspetto. Seppur meno immediata come risposta, questa versione può essere particolarmente utile nel dare un maggior senso di controllo nelle mani dell'autore, che avrà ora tre coppie di opzioni da combinare a piacimento, offrendo una più ampia gamma di sfumature. Ognuna delle proposte era accompagnata da un film di reference appropriati.

In questo episodio si è anche utilizzato ChatGPT per ragionare se la psicologia del maniaco fosse scientificamente credibile. Si è ragionato sui suoi comportamenti e le sue ossessioni, arrivando a fare delle modifiche importanti relative alla verosimiglianza. Ad esempio, si è arrivati alla conclusione che la vittima rappresenta spesso un oggetto sacro per il carnefice. Nella prima versione della sceneggiatura delle foto della lettrice erano attaccate sul diario e scarabocchiate in diversi modi, ma alla luce di queste decisioni sono state solo incorniciate da un disegno macabro e lasciate intonse nel resto.

### ***Episodio dell'Alienazione***

In questo caso ChatGPT risponde al primo prompt in maniera del tutto differente: se nei primi due episodi si era dedicato quasi esclusivamente ad analizzare la correttezza grammaticale e

lessicale del testo, qua sembra analizzarlo con un fine molto più narrativo. La risposta, sempre in forma di elenco, analizza diversi aspetti, come: lo sviluppo emotivo dei personaggi, il finale e l'arco narrativo. Tutti fattori fortemente legati allo storytelling e poco relativi alla scrittura. Le uniche modifiche testuali che propone sembrano non sposare le regole del testo cinematografico, suggerendo – ad esempio – di rendere più emotive ed astratte le descrizioni, una pratica adatta alla scrittura romanzesca, ma sconsigliata per il cinema.

Complessivamente, pur essendosi registrata una risposta anomala nella forma, il messaggio che porta è apprezzabile in qualità di analisi del racconto e fornisce un buon riscontro sugli aspetti principali della storia.

Nel secondo prompt, ChatGPT risponde alla presa sul pubblico in modo positivo, trovando il filo narrativo proficuo, il tema efficace ed i personaggi coerenti nel modo di parlare. Suggerisce solo di inserire qualche battuta di contrasto in più tra le due protagoniste, dando un maggiore sfogo alla tensione, un consiglio accolto ed implementato successivamente attraverso un dialogo più ricco. Per quanto riguarda l'aderenza di genere l'episodio sposa bene le regole del dramma secondo l'IA, ma viene consigliato di porre attenzione nella gestione del ritmo, con il seguente messaggio:

*Risposta prompt 2 Episodio 3: [...] Il ritmo è lento, come si addice a un dramma, ma questo deve essere gestito attentamente per evitare che il pubblico perda interesse. Il pubblico di oggi ha spesso un'attenzione limitata, quindi potrebbe essere utile inserire momenti di climax più frequenti o variazioni nel ritmo per mantenere l'attenzione alta.*

In questa frase ChatGPT dimostra di applicare ragionamenti profondi e ponderati, dimostrandosi capace di rispondere al concetto di “presa sul pubblico” con senso critico.

In risposta al terzo prompt, il modello riprende la struttura utilizzata per il primo episodio, proponendo tre opzioni precise di stile: Fotografia intima e montaggio lento, fotografia contrastata e montaggio spezzato, fotografia stile documentario e montaggio fluido. Riproponiamo di seguito, a scopo esemplificativo, la prima opzione nella sua interezza, al fine di apprezzarne la specificità.

*Risposta prompt 3 Episodio 3: Girare questa scena richiede un approccio visivo che enfatizzi l'intensità emotiva e l'isolamento dei personaggi, mantenendo al contempo un senso di realismo. Ti fornirò diverse opzioni per impostare la fotografia e il montaggio, insieme a riferimenti cinematografici che possono ispirarti.*

*Opzione 1: Fotografia Intima e Montaggio Lento*

Stile Generale:

- *Fotografia: Usa luci soffuse e tonalità desaturate per trasmettere il senso di alienazione e depressione. Un'illuminazione naturale che filtra attraverso la finestra, creando contrasti netti tra le ombre e la luce, può essere molto efficace. Prediligi l'uso di colori freddi (blu, grigi) per rappresentare l'apatia e il distacco emotivo della Madre.*

## IA nella fase di pre-produzione

- *Inquadrature: Utilizza inquadrature strette, come primi piani o dettagli delle mani tremanti della Madre, del fumo della sigaretta, o del posacenere colmo. Questi dettagli aiutano a creare un legame emotivo tra il pubblico e i personaggi, mostrando la loro fragilità e tensione interiore.*
- *Montaggio: Mantieni un ritmo lento, con lunghe pause tra i dialoghi. Lascia che i momenti di silenzio e i piccoli gesti parlino da soli. Il montaggio dovrebbe enfatizzare l'isolamento di ogni personaggio, alternando tra i primi piani della Madre e della Zia, senza mai mostrarli insieme nello stesso frame fino al climax emotivo.*

### Riferimenti Cinematografici:

- *"Manchester by the Sea" (2016) di Kenneth Lonergan: La fotografia è caratterizzata da una palette di colori freddi e desaturati, con un forte uso di inquadrature strette che sottolineano il dolore interiore dei personaggi.*
- *"Blue Valentine" (2010) di Derek Cianfrance: La combinazione di camera a mano e l'uso di luce naturale crea un'atmosfera intima e dolorosa, perfetta per un dramma familiare.*

In questo singolo esempio è possibile vedere come ChatGPT riesca effettivamente a costruire uno stile dettagliato per girare la scena, fornendo diversi spunti da cui trarre ispirazione, spesso anche molto specifici. Ad esempio, nella risposta appena riportata si parla della possibilità di non riprendere mai nella stessa inquadratura sia la madre che la zia, con l'obiettivo di comunicarne la lontananza emotiva. Questo particolare è una trovata originale e profonda, che va oltre alla semplice scelta di colori freddi o caldi, testimoniando un modo di ragionare altamente creativo da parte della macchina. Infatti, questo ed altri elementi di stile sono poi stati scelti ed utilizzati nella fase di riprese.

### ***Episodio Simbolico-Interpretativo***

Anche in questo caso, al primo prompt il modello risponde con un'analisi sia narrativa che testuale, centrando l'obiettivo della domanda. Si interroga efficacemente sui personaggi ed il loro modo di parlare, un aspetto particolarmente importante in questo episodio in cui ci sono molti personaggi con registri differenti. Tuttavia, non vengono di nuovo proposti cambi di battute particolarmente incisivi. La sua utilità rimane più legata all'analisi generale che non a una valutazione attenta di ogni battuta.

Questo episodio è, probabilmente, anche il più complesso da catalogare come genere, per cui è particolarmente interessante come ChatGPT gestisce la seconda domanda, legata alla presa sul pubblico e alla coerenza di genere. Ecco la risposta nella sua interezza:

*Risposta prompt 1 Episodio 4: La sceneggiatura ha sicuramente delle caratteristiche che potrebbero attrarre un pubblico specifico, soprattutto coloro che apprezzano i lavori con un forte contenuto simbolico e*

## IA nella fase di pre-produzione

*un'atmosfera surreale. Tuttavia, ci sono alcuni aspetti da considerare per valutare la sua presa sul pubblico e la caratterizzazione di genere.*

### 1. Presa sul Pubblico

- *Atmosfera e Intrigo: L'atmosfera cupa e il mistero attorno alla scena del cadavere possono catturare l'attenzione del pubblico. Il fatto che la scena sia ambientata in un teatro vuoto, con un'atmosfera quasi onirica, conferisce un certo fascino che può intrigare chi è attratto da ambientazioni enigmatiche e riflessive.*
- *Originalità: La sceneggiatura sembra uscire dagli schemi tradizionali, puntando su un dialogo tra personaggi che rappresentano archetipi o aspetti psicologici più che persone reali. Questo tipo di narrazione può attrarre un pubblico che cerca un'esperienza più intellettuale o sperimentale, ma potrebbe risultare alienante per chi preferisce storie più lineari e accessibili.*
- *Coinvolgimento Emotivo: Anche se la sceneggiatura è coinvolgente dal punto di vista intellettuale, potrebbe mancare un po' di impatto emotivo per coinvolgere profondamente il pubblico. I dialoghi, per quanto simbolici, potrebbero essere resi più empatici per permettere agli spettatori di connettersi meglio con i personaggi e la loro situazione.*

### 2. Caratterizzazione di Genere

- *Difficoltà di Classificazione: La sceneggiatura sembra attraversare diversi generi, come il dramma, l'horror psicologico e il teatro dell'assurdo. Questa ibridazione può essere un punto di forza, rendendola interessante per chi ama le opere che non si confinano in un genere specifico. Tuttavia, potrebbe risultare difficile da classificare e quindi da vendere a un pubblico generalista.*
- *Ruoli Stereotipati o Innovativi: I personaggi, come "La Sposa", "La Madre", "Il Perverso", sono più simboli che individui con una caratterizzazione psicologica profonda. Questo è coerente con lo stile simbolico della sceneggiatura, ma rischia di ridurre la loro umanità, il che potrebbe limitare l'identificazione del pubblico con loro. Per una presa più forte, potrebbe essere utile aggiungere qualche sfumatura in più ai personaggi, rendendoli meno stereotipati e più complessi.*
- *Coerenza Tematica: Il tema della responsabilità e dell'inevitabilità sembra centrale. Tuttavia, la sceneggiatura potrebbe beneficiare di un rafforzamento della coerenza tematica, facendo in modo che tutte le battute e le azioni dei personaggi contribuiscano chiaramente a sviluppare questi temi, evitando così di sembrare scollegate o casuali.*

Conclusioni

*La sceneggiatura ha il potenziale per avere una buona presa su un pubblico di nicchia, particolarmente coloro che apprezzano opere teatrali e cinematografiche dal forte contenuto simbolico e sperimentale. Tuttavia, per ampliare il suo appeal, potrebbe essere utile lavorare su un maggiore coinvolgimento emotivo e su una caratterizzazione dei personaggi che, pur rimanendo archetipica, possa offrire maggiore profondità e complessità. Questo potrebbe rendere la storia non solo intellettualmente stimolante ma anche emotivamente risonante, aumentando così la sua capacità di coinvolgere un pubblico più ampio.*

Questa estesa analisi sembra cogliere tutta la complessità dell'episodio, individuandone il potenziale ed il target di riferimento, ma anche i limiti e le debolezze. I punti di miglioramento proposti sono validi, e si è cercato di implementarli nelle stesure successive.

Vista la natura metaforica della storia, si è voluto mettere alla prova le capacità interpretative del modello con una domanda aggiuntiva, in cui si è chiesto cosa rappresenti - secondo lui - il cadavere in scena. Anche in questo caso proponiamo la risposta nella sua interezza, per valutarne l'elaborazione.

**Risposta prompt aggiuntivo Episodio 4:** *Il cadavere al centro della scena sembra avere un significato altamente simbolico, e la sua interpretazione può variare a seconda del contesto generale e dei temi della sceneggiatura. Ecco alcune possibili interpretazioni:*

- 1. Rappresentazione della Creatività Morta o Fallita: Data la presenza dello "Scrittore" e il contesto teatrale, il cadavere potrebbe simboleggiare un'opera creativa, un progetto artistico o una visione che è stata "uccisa" o che è fallita. La disperazione dello Scrittore e l'atteggiamento degli altri personaggi potrebbero riflettere il senso di perdita e fallimento di fronte a un progetto che non ha raggiunto il suo potenziale.*
- 2. L'Inutilità dell'Azione: Il cadavere potrebbe anche rappresentare l'inutilità delle azioni dei personaggi, soprattutto in un contesto dove i loro sforzi sembrano ormai vani. L'idea che "ormai è successo" e che ora si debba "sistemare la cosa" suggerisce che qualcosa è andato irrimediabilmente storto, e il cadavere è la prova tangibile di questa irreversibilità.*
- 3. La Morte di un'Illusione: In un contesto più esistenziale, il cadavere potrebbe rappresentare la morte di un'illusione collettiva o personale. Questo potrebbe collegarsi ai temi della consapevolezza e della responsabilità accennati nei dialoghi, indicando che ciò che i personaggi hanno perso è un'illusione su cui facevano affidamento, e ora devono confrontarsi con la dura realtà.*
- 4. Simbolo della Fine Inevitabile: Il cadavere potrebbe anche simboleggiare la fine inevitabile di tutte le cose, riflettendo un tema più*

*universale della mortalità e della caducità. I personaggi discutono dell'inevitabilità e delle conseguenze delle loro azioni, e il cadavere potrebbe essere un richiamo alla morte come un destino inevitabile che tutti devono affrontare, indipendentemente dai loro sforzi.*

- 5. Manifestazione di una Colpa Collettiva: Infine, il cadavere potrebbe rappresentare una colpa collettiva dei personaggi, un segno tangibile di qualcosa che tutti loro hanno contribuito a causare. La discussione sulla responsabilità potrebbe implicare che il cadavere è il risultato delle loro azioni o inazioni, un simbolo del prezzo delle loro scelte.*

Leggendo ognuna delle interpretazioni proposte, se ne individuano alcune leggermente vaghe (la numero 2 e 4), ma che potrebbero risultare anche azzeccate se la storia venisse espansa maggiormente, ma altre invece stupiscono per la verosimiglianza. È significativo che la giusta lettura, seguendo le intenzioni con cui l'episodio è stato scritto, sia al primo posto dell'elenco e centri completamente il messaggio che vuole trainare. Questo dimostra le elevate capacità di ChatGPT di cogliere messaggi metaforici che vanno oltre alla lettura letterale, una capacità imprescindibile per poter valutare la qualità di un testo artistico.

Si è comunicato nella chat, dunque, che la prima opzione era la giusta interpretazione. Il modello ha risposto ricollegando ogni elemento della storia a questa nuova informazione, leggendo in maniera corretta il ruolo di ogni attore sul palco e la natura della vicenda stessa.

Per concludere, si è poi sottoposto il terzo ed ultimo prompt al modello. Anche in questo caso si sono ricevute tre valide soluzioni di messa in scena: Stile Surrealista e Onirico, Stile Teatrale e Statico, Stile Espressionista e Psicologico. I film di reference consigliati ben rappresentavano le opzioni elencate, confermando l'utilità di ChatGPT anche come strumento di ricerca di riferimenti per la pre-produzione.

## 7.2 ChatGPT per la scrittura della sceneggiatura

In questa sezione verrà trattata l'analisi applicativa di ChatGPT per la scrittura creativa. Questo particolare utilizzo è quello che maggiormente crea preoccupazione nella comunità degli scrittori, suscitando il timore che l'Intelligenza Artificiale possa in futuro sostituire questa arte, finora, esclusivamente umana. Nei primi capitoli della tesi si è spiegato, però, che in questo caso studio si è voluto sperimentare un nuovo rapporto collaborativo tra uomo e macchina, testando ChatGPT come strumento di sostegno alla scrittura, capace di espandere le potenzialità creative dell'uomo, ma lasciando a quest'ultimo il controllo creativo del testo.

Il cortometraggio in esame "Non ha visto nulla" è suddiviso, come visto, in cinque episodi in totale. Per quanto riguarda la scrittura si sono fatti due utilizzi differenti di ChatGPT:

- **Riscrittura di episodi esistenti:** i primi quattro episodi sono stati, in prima battuta, scritti in piena autonomia, senza l'ausilio dell'IA. Terminata la fase di revisione (analizzata nel capitolo 7.1), si è fatto riscrivere ogni episodio al modello, chiedendo di apportare le migliorie necessarie. In questo scenario, ChatGPT agisce più come correttore e revisore che come scrittore vero e proprio.
- **Scrittura collaborativa di un episodio originale:** l'ultimo episodio, invece, si è voluto pensare e scrivere fin dal principio collaborando con ChatGPT, attraverso un dialogo interattivo che ha portato a far coesistere le idee dello sceneggiatore e quelle della macchina.

Alla luce di questo ci si appresta ora a mostrarne i risultati.

### 7.2.1 Riscrittura di episodi esistenti

La riscrittura dei primi quattro episodi è avvenuta all'interno delle chat create per la fase di revisione. Per ognuno di essi, si è proposto il seguente prompt:

*Prompt di riscrittura: Rileggi in testo di partenza, vorrei che riscrivessi la sceneggiatura, mantenendo inalterata la trama, la durata, i punti di snodo della storia ed il finale, ma apportando le correzioni e migliorie che ritieni necessarie. Prova a mantenere le battute che reputi efficaci e a cambiare le più deboli*

La riscrittura per il primo episodio è stata bilanciata e non eccessivamente impattante, agendo principalmente nel rendere più sintetiche le descrizioni e modificando alcune parole dei dialoghi, senza però migliorarli sensibilmente. Si è chiesto di provare a caratterizzare ulteriormente le battute del Cowboy, impostando un registro più rude e grezzo, ma anche in questo caso le migliorie sono state minime.

ChatGPT offre la possibilità di selezionare una parte della risposta per poterla commentare, una funzione molto utile nell'ottica di voler intervenire solamente su precisi versi. Si è provato a testare questa funzionalità, volendo agire su specifiche battute, ma i risultati sono stati mediocri, spesso artificiosi e poco naturali.

Il secondo episodio, che è caratterizzato dall'assenza di dialoghi, ha riscontrato una risposta più apprezzabile, rendendo le descrizioni più fluide e concise. La lettura ne risulta, effettivamente, agevolata. Successivamente si è chiesto di suggerire come empatizzare maggiormente con la vittima aggiungendo piccoli dettagli, un punto di debolezza individuato nella sezione di revisione. Il modello ha proposto una rielaborazione del testo che ha introdotto diversi elementi

di caratterizzazione nell'incipit, riuscendo a rendere la ragazza più riconoscibile e quindi aiutando a sentirsi più vicini a lei. Questo è avvenuto attraverso tre aggiunte: introduzione di piccoli dettagli personali, dialogo tra sé e sé, interazione con un personaggio esterno. Tutte le modifiche si sono rivelate fruttuose ed efficaci, si è scelto di mantenerne solo alcune per evitare di eccedere, ma comunque rappresenta un ottimo esempio di integrazione dei consigli dell'IA all'interno del progetto.

Un'ulteriore richiesta fatta al modello, è stata di introdurre un nuovo elemento "spiato" dal maniaco, ossia un dettaglio della ragazza che in sceneggiatura il maniaco osserva ossessivamente, al fine di arrivare ad un numero totale di tre (nella stesura corrente risultavano due). ChatGPT ha risposto con dieci opzioni valide, da cui è stata scelta la migliore ed inserita nella versione finale.

Proseguendo sul terzo episodio, si sono riscontrati dei problemi inizialmente. Il modello ha riscritto l'intero episodio, modificando ampiamente le battute, ma era ben evidente che avesse frainteso la trama, pensando che la madre avesse appena scoperto di aver perso l'affidamento del figlio. Questo episodio era sicuramente quello meno immediato e con una storia più complessa, spesso suggerita ma mai esplicitata del tutto, quindi l'errore è in parte giustificabile. Si è risposto al messaggio chiarendo la situazione allora, esplicitando le ragioni del dramma e di procedere in questa nuova direzione. Il risultato, tuttavia, è stato di nuovo inadatto: la vicenda è stata interpretata correttamente, con battute anche apprezzabili seppur circostanziali, ma sono stati modificati del tutto gli snodi principali della scena, portandola – infine – ad un finale completamente diverso e venendo meno alle richieste del prompt.

Si è tentato allora un approccio più frammentato, riferendosi a specifici scambi di battute e chiedendo di renderle più rabbiose, più ricche di scontro, oppure più malinconiche a seconda delle esigenze. Per la prima volta ChatGPT ha agito in maniera più incisiva sulle battute, riscrivendole in maniera del tutto originale. I suggerimenti sono, in questo caso, stati validi e ben scritti, capaci davvero di migliorare la scena contrariamente a quanto avvenuto con gli episodi precedenti. Si è apprezzato che il modello abbia proposto più opzioni per ogni battuta, permettendo così di ottenere ispirazioni diverse ed uscire dai binari iniziali. Questo approccio ha permesso di sfruttare piccoli dettagli da ognuna delle alternative e di rielaborarle per inserirle nella stesura iniziale. Insomma, ChatGPT ha dimostrato di gestire più efficacemente poche battute per volta e di essere uno strumento di maggiore ispirazione se presenta più proposte di soluzioni, invece di agire in modo troppo deterministico.

Questo si è notato anche nella riscrittura dell'ultimo episodio, dove il testo era molto breve e costituito quasi solo da battute. ChatGPT ha reinterpretato in modo originale le battute, modificandole sensibilmente e impegnandosi a dargli un ritmo funzionale alla narrazione.

### **7.2.2 Scrittura collaborativa di un episodio originale**

La scrittura dell'ultimo episodio voleva essere sviluppata sin dal principio con l'appoggio dell'IA. Si era definito un plot approssimativo, per fare in modo che il messaggio e la vicenda ben si sposassero con il resto del cortometraggio, ma l'obiettivo era di ribaltare il paradigma in cui lo sceneggiatore scrive e il modello rifinisce, cercando di portare il modello a scrivere in autonomia il più possibile ed intervenire solo per direzionare il risultato e per gli aggiustamenti finali.

Il tutto è partito, quindi, da un prompt che dava piena libertà al modello, consegnandogli solamente il plot dell'episodio.

**Prompt scrittura 1:**

*Scrivi una sceneggiatura di 2/3 minuti, dal tono sarcastico e dialoghi brillanti che segua la seguente trama: Uno scrittore in crisi rimanda da tutto il giorno la chiamata con il suo agente, perché sa che domani c'è la data della consegna, ma lui ha ancora le idee confuse. Guardando il telefono, gli compare una pubblicità su un tool di scrittura script con AI. Lo scrittore chiude la pubblicità dell'AI imprecando. Suona il suo cellulare. Lo scrittore risponde, il suo agente lo incita a terminare di scrivere entro stasera la puntata pilota di una nuova serie. Lo scrittore ribadisce che ha un'idea in mente che dentro ha tutto - romanticismo, mistero, dramma - deve solo trovare la chiave per unire il tutto. L'agente risponde che non c'è tempo e basta che sia qualcosa entro il limite e che non importa che sia bello o no. Finita la chiamata, lo scrittore - titubante - approccia il computer e per superare il blocco usa l'AI. Grazie al dialogo si sbloccherà e incomincerà a scrivere inarrestabile.*

Il risultato del prompt è mediocre, sicuramente rispetta tutte le richieste fatte e sposa il plot nella sua interezza, ma nulla di originale è stato aggiunto ed il tono sarcastico è spesso dozzinale. Si è chiesto, allora, se il modello avesse delle idee alternative per trasmettere il messaggio di collaborazione tra uomo e macchina in altri modi più originali, anche modificando il plot. Le soluzioni proposte sono interessanti e vanno ad esplorare diverse dinamiche di interazione possibili tra l'IA e lo scrittore, tuttavia ognuna di esse richiederebbe molto più spazio di quanto concesso dai limitati mezzi economici, ed una loro implementazione riassuntiva risulterebbe frettolosa.

Visto che in piena libertà, la scrittura è risultata troppo semplice, si è mandato un prompt con nuove istruzioni e suggerimenti.

**Prompt scrittura 2:**

*Scrivimi una sceneggiatura con il seguente plot:[...]*

*Il tutto deve durare non più di 3 minuti, il tono deve essere accattivante ed ironico. Lo scrittore giovane, di 30 anni, un po' sbandato e divertente. Si utilizzino elementi comici, come ad esempio il fatto che l'agente sia memorizzato sul telefono come "agente Smith". E' possibile usare elementi cinematografici meno usuali, come il voice over dei pensieri dello scrittore, per rendere tutto più dinamico e simpatico.*

Con queste poche informazioni aggiuntive, la risposta prende tutto un altro spessore. Già in questa prima versione è presente quella che sarà la struttura finale dell'episodio: la pubblicità che sveglia lo scrittore, lui che gli lancia degli oggetti per farla smettere, la conversazione tagliente con l'agente, il voice-over e soprattutto il montaggio veloce per rappresentare la spinta creativa finale raggiunta grazie al tool di IA. Le battute del voice over sono spesso troppo didascaliche ancora, e non sufficientemente divertenti, ma il dialogo tra scrittore e agente già ha assunto delle dinamiche incalzanti. Al fine di migliorare questo aspetto e caratterizzare maggiormente, si è proceduto con il seguente prompt, inserendo delle reference.

**Prompt scrittura 3:**

*Ottima direzione, mi piacciono molto alcuni elementi, ma vorrei provassi a farne un'altra versione, ispirandoti allo stile di Kaufman. Per il personaggio dello scrittore ispirati anche a Drugo de "Il grande Lebowski"*

Grazie all'aggiunta delle reference, ChatGPT ha saputo cogliere lo stile di scrittura desiderato, in questa nuova versione le battute sono diventate decisamente più originali e degne di Drugo Lebowski, l'ambiente e la narrazione accompagnano il personaggio efficacemente, rendendo il tutto ben costruito ed omogeneo.

Si è provato a chiedere una variante, per valutare altre opzioni, in cui le battute ed il personaggio diventavano più ansiogeni per far respirare maggiormente la frustrazione. Il risultato era ben scritto ed una valida alternativa, ma si è deciso di proseguire con la versione precedente.

Definiti stile e struttura, non rimaneva che concentrarsi sui dettagli e andarli a smussare uno ad uno. Ad esempio, nella prima versione proposta, la pubblicità dell'IA veniva trasmessa alla tv, ma si è chiesto di sostituirlo con una cassa intelligente (come *Alexa* di Amazon). Si è voluto anche mantenere una bella idea di ChatGPT, in cui lo scrittore lanciava un telecomando contro la pubblicità, ma si è proposto di fargli provare tre lanci fallimentari per accentuarne la comicità. Dopodiché si è voluto specificare che la scrittura con l'IA dovesse partire innanzitutto dalle idee dello scrittore, un dettaglio importante per definire il messaggio del film.

Il problema principale nel continuare a dare istruzioni per cambiare dettagli nella sceneggiatura, era la mancanza di controllo sul resto del testo: ogni volta che si richiedeva una certa modifica, il modello rigenerava l'intera sceneggiatura, mantenendo sufficientemente salda la storia, ma modificando descrizioni e battute. Questo introduce grande dispersività nel processo, perché seppur ogni nuova stesura fosse sempre ottima, si rischiava di perdersi delle battute che si erano considerate perfette nelle versioni precedenti. Visto che l'ottica era di avere un controllo creativo elevato anche nella scrittura collaborativa, è imprescindibile poter decidere di mantenere alcuni elementi e di ritagliarne altri. A questo scopo si è deciso, una volta implementate tutte le modifiche, di scrivere una sceneggiatura esternamente a ChatGPT, attraverso un processo di taglia e cuci da tutte le versioni generate. Così facendo si è estrapolato il meglio da ognuna e, con un intervento manuale, si sono corretti i piccoli dettagli o aggiunte idee nuove alla narrazione.

Questa nuova sceneggiatura ibrida è stata infine sottoposta ad una revisione su ChatGPT, chiedendo di analizzarne temi e messaggi, controllare che sia coerente e che possa divertire. Il modello ha effettuato un'analisi attenta ed esaustiva, confermando che la direzione presa è quella giusta. Si chiede, infine, di riscrivere la sceneggiatura apportando piccole migliorie dove necessario. Il processo di scrittura collaborativa termina con un'ultima revisione da parte nostra, al fine di rendere propria la sceneggiatura, un risultato complesso e possibile grazie all'unione delle nostre idee e di quelle dell'intelligenza artificiale.

## 7.3 ChatGPT per la scrittura di prompt per immagini

ChatGPT è stato sfruttato per altri scopi all'infuori della scrittura e revisione di testi creativi. Nel capitolo 6.4.3 si è visto come molti studi abbiano sfruttato il modello di OpenAI per la scrittura di prompt per la generazione di immagini; tutti gli studi riportavano risultati più che soddisfacenti al riguardo, dimostrando come ChatGPT sappia tradurre efficacemente le nostre richieste da linguaggio naturale a indicazioni dettagliate per un tool di generazione immagini.

Dall'analisi svolta sul prompt engineering per immagini non è emersa una guida solida per la scrittura autonoma dei prompt, ma solo molte pratiche ottimali da applicare. Nell'ottica di un'ottimizzazione del lavoro di generazione delle immagini, si è deciso di testare le potenzialità di ChatGPT, con l'obiettivo di individuare un metodo in cui fossero minimizzati i tentativi e massimizzata la resa in termini di qualità e di tempo.

Il tutto è avvenuto secondo due fasi:

1. Fase di addestramento: nonostante ChatGPT dovrebbe essere già in grado di scrivere agilmente i prompt in autonomia, si è deciso di tentare un approccio con maggiore controllo sulla macchina. Si è creata una chat dedicata all'assistenza di scrittura di prompt e si è costruito, attraverso le tecniche esplorate al capitolo 6.3, un dialogo didattico con il modello, cercando di stabilire delle linee guida condivise per la scrittura dei prompt.
2. Generazione dei prompt: una volta addestrata la chat, si è potuto procedere con la fase di scrittura dei prompt, sia per le moodboard che per gli storyboard.

Apprestiamoci, dunque, a scoprirne le metodologie.

### 7.3.1 Addestramento per la scrittura dei prompt

Lo scopo di questa fase non è stata di un reale insegnamento per la scrittura di prompt efficaci, di cui ChatGPT ha già tutti i mezzi necessari, quanto più di valutarne la metodologia e strutturare un metodo applicativo attento e non approssimativo. Allo scopo, è stata creata una chat dedicata, su cui è avvenuto sia l'addestramento che la fase di generazione, mantenendone così la coerenza.

La chat è cominciata definendo il ruolo di ChatGPT, riprendendo la tecnica di *Priming Prompting* esplorata in precedenza, una particolare variante del Role Prompting che caratterizza una chat secondo le proprie esigenze. Ciò è avvenuto con il seguente prompt:

*Prompt 1: Sarai il mio esperto di prompt engineering per la generazione di immagini con l'intelligenza artificiale. In questa chat discuteremo e studieremo tutte le più rilevanti tecniche di scrittura di prompt per ottenere immagini accurate. Avrò bisogno che segui i miei ragionamenti passo per passo e che memorizzi le informazioni che ti do, per poterle poi applicare a casi specifici. Pensi di potermi aiutare?*

Occorre far notare l'uso dell'espressione "ragionamenti passo per passo", si è visto come questa piccola frase possa portare il modello a svolgere ragionamenti più attenti ed esplorativi, indagando con precisione ogni passaggio delle sue risposte.

In risposta, ChatGPT si è definita pronta al compito e all'apprendimento, individuando cinque punti di riflessione sulla scrittura di prompt da scoprire assieme:

- Struttura del prompt: Come formulare descrizioni chiare e dettagliate.
- Tecniche avanzate: Uso di stili, atmosfere, luci, e altre variabili.
- Iterazione e miglioramento: Strategie per migliorare i risultati in base alle immagini generate.
- Adattamento ai modelli: Capire come modelli diversi interpretano i prompt (es. DALL-E, MidJourney, Stable Diffusion).
- Esempi pratici: Analisi di prompt e risultati per capire cosa funziona meglio.

Al fine di sollecitare il modello a recuperare le informazioni da sé, seguendo la tecnica Step-Back Prompting, e ragionare inizialmente ad alto livello, si è chiesto di cominciare reperendo informazioni dettagliate sui punti individuati. Leggendo la risposta si ha avuto conferma che ChatGPT avesse ben chiare le nozioni di base della scrittura di prompt efficaci.

Per assicurarsi di impostare una conoscenza più profonda al riguardo, si è voluto allora proporre di inviare ulteriori informazioni utili alla scrittura di prompt, in modo che il modello potesse apprenderle. Fin da subito la conversazione ha assunto un tono collaborativo, impostando un clima fortemente interattivo in cui ChatGPT acquisiva attivamente le informazioni, le ricollegava alle precedenti e integrava nuove informazioni dove necessario. Per qualche iterazione si è proceduto con uno schema di conversazione ricorrente:

- Venivano inviati al modello nuovi dettagli utili alla prompt engineering, direttamente provenienti dal capitolo che vi abbiamo dedicato.
- ChatGPT recepisce i nuovi punti di riflessione, li riformulava dando prova di averli compresi, e ne strutturava un discorso per analizzarne le caratteristiche, i casi d'uso, come integrarli, gli errori comuni e degli esempi pratici.

Questo processo è stato svolto sia per aspetti legati agli elementi che compongono un prompt sia per le pratiche da seguire per una scrittura efficace.

Per consolidare ulteriormente le nozioni espresse e cominciare ad esplorare alcuni esempi applicativi, si è voluta sfruttare la possibilità di inviare documenti nella chat, inviando alcuni degli articoli su cui è basata questa tesi. Gli si sono date le istruzioni presenti nel seguente prompt:

***Prompt 2:** Condivido con te questo documento vorrei lo leggessi pagina per pagina, contiene molti esempi pratici e molte regole utili. Terminata la lettura rileggilo una seconda volta, cerca di non perderti nessun dettaglio.*

Enfatizzando la necessità di essere esaustivo nella lettura e di non tralasciare dettagli, congiuntamente alla tecnica di Re-reading, si è cercato di garantire un'analisi il meno approssimativa possibile. In risposta si è ricevuta un riscontro riassuntivo, ma dettagliato, dei documenti, dimostrandosi in grado di comprendere testi complessi.

L'intero processo di addestramento ha richiesto diversi giorni, per cui si è voluto controllare che la memoria della conversazione fosse robusta anche con il passare del tempo, ossia si è chiesto al modello di rileggere la chat nella sua interezza e riassumerne le informazioni esplorate.

***Prompt 3:** Ottimo lavoro! Sei in grado ora di rileggere tutta questa nostra conversazione e riordinare tutti i concetti che abbiamo esplorato in maniera ordinata e senza tralasciare nulla? In modo da avere una guida completa di prompt engineering cui fare riferimento.*

La risposta è stata positiva, ChatGPT ha saputo organizzare efficacemente una guida riassuntiva sulla disciplina dimostrando consistenza all'interno della chat.

Siccome fino a questo punto della conversazione ci si era limitati a discutere in linea teorica, si è pensato di proporre qualche esempio applicativo da analizzare. Ci si è assicurati che ChatGPT fosse in grado di comprendere e processare le immagini, dopodiché si è creata una tabella contenente immagini generative ed i rispettivi prompt, provenienti dai siti delle comunità di AI Artists. Si è mandato il documento a ChatGPT, accompagnato dal messaggio:

***Prompt 4:** Ecco qualche esempio utile, concentrati sull'uso dei prompt per descrivere: soggetti dell'immagine, stile dell'immagine, ambientazione dell'immagine, composizione dell'immagine. Ragiona passo per passo.*

Il modello si è nuovamente mostrato capace di completare il suo scopo, ha analizzato con successo i prompt, individuato le caratteristiche più interessanti di ogni immagine e studiato come questi aspetti venissero integrati all'interno della descrizione. Allo stesso scopo si è tentato di chiedere al modello di analizzare alcuni dei siti da cui sono state recuperate le immagini, ma in questo caso non si è dimostrato capace di esplorare in maniera strutturata il sito, limitandosi a comprendere di cosa si occupasse.

Con questa ultima interazione si è ufficialmente conclusa la fase di addestramento, ChatGPT si è dimostrata capace di strutturare le informazioni, reperirne in autonomia e recepire quelle fornite in forma testuale o tramite allegati. La fase successiva si occupa, dunque, di testare le capacità del modello di mettere in pratica le nozioni apprese.

### 7.3.2 Scrittura dei prompt

Una volta addestrata la chat, si è potuto procedere per la generazione dei prompt per lo storyboard e per le moodboard. Per le due applicazioni si sono utilizzati approcci differenti. Katalist, infatti, ha il pregio di avere a disposizione la sceneggiatura, consentendo quindi di determinare in maniera esatta sia i personaggi che gli ambienti, e di mantenerli coerenti tra le varie inquadrature. Le moodboard invece, per quanto richiedano un'aderenza minore rispetto allo storyboard, non avevano modo di assumere delle informazioni di base da mantenere costanti tra i vari shot. Per provare a sopperire a questa mancanza, ed evitare di ripetere ogni volta la descrizione di ambienti e personaggi, si è fornito a ChatGPT la sceneggiatura, in modo che la leggesse e coordinasse i prompt in relazione con il testo.

In entrambi i casi, l'utilizzo di ChatGPT si è dimostrato fruttuoso. Questo è avvenuto innanzitutto per la descrizione degli ambienti e dei personaggi su Katalist, che ha permesso di caratterizzare coerentemente gli elementi. Per quanto riguarda le inquadrature, è bastato dare poche informazioni in linguaggio naturale, ed il modello è stato capace di trarne un prompt accurato e ricco, apportando anche alcune idee originali e mostrando una natura propositiva.

Ad esempio, la prima inquadratura si è ottenuta come segue:

***Prompt iniziale:** Establishment Shot, campo lungo, vediamo la casa in legno del cowboy in mezzo a una prateria, con un portico, un albero secco si trova a lato della casa. Il cowboy è seduto sulla sedia a dondolo nel portico.*

***Prompt finale:** A long shot of a wooden cowboy house in the middle of a vast prairie, with a front porch and a dry tree standing to the side of the*

*house. The cowboy is sitting on a rocking chair on the porch. Warm sunset lighting casts long shadows, creating a serene and rustic atmosphere. Hyper-detailed, photorealistic, cinematic composition, 8K resolution, wide lens.*

Si può facilmente notare come ChatGPT abbia saputo comprendere profondamente il testo ed interpretarlo creativamente aggiungendo dettagli sull'ambiente e sulla luce. Questo aspetto è rilevante, è stato piuttosto comune che il modello sperimentasse descrizioni originali per gli elementi in scena, dando interessanti sfumature all'inquadratura. Questo richiedeva chiaramente una certa attenzione al riguardo, era possibile che il prompt enfatizzasse o fraintendesse alcuni aspetti, ma ha al contempo permesso di ottenere prompt elaborati partendo anche da descrizioni scarse.

Un occhio critico nei confronti del prompt era, però, necessario in generale. Essendo le descrizioni spesso semplificative, veniva lasciato ampio margine di elaborazione a ChatGPT, che generalmente veniva ben gestito rispettando le aspettative, ma che potenzialmente poteva introdurre errori che richiedessero piccoli interventi di correzione manuale.

Talvolta, invece, un prompt apparentemente corretto si dimostrava inefficace una volta utilizzato per la generazione. In questi casi è emerso il vero potenziale della scrittura di prompt tramite ChatGPT, ossia il dialogo costruttivo con esso, fatto di continui tentativi e feedback per l'ottenimento del miglior risultato.

Di fronte ai numerosi ostacoli, il modello si è dimostrato estremamente agile, facendo bastare una semplice descrizione del problema per riformulare il prompt di conseguenza, riducendo enormemente le tempistiche. Ecco un esempio tipico di feedback:

***Prompt di correzione:** Katalist mette i personaggi troppo vicini, riformula in modo che sia ben chiara la distanza tra i due. Inoltre, mi serve che illumini meglio la sposa.*

Con un semplice feedback, comunichiamo al modello gli errori da correggere e lui proporrà un nuovo prompt adatto alle richieste. Questo può richiedere più tentativi, ma attraverso un dialogo intuitivo e naturale si è quasi sempre riusciti ad ottenere buoni risultati (fin dove possibile). Analogamente, veniva anche comunicato al modello quando il prompt utilizzato funzionava come previsto, incoraggiando così il modello ad apprendere sia dai risultati positivi che da quelli negativi.

In linea di massima, è stato comune dover chiedere di semplificare i prompt quando utilizzati su Katalist, perché l'abbondanza di dettagli poteva portare il tool a perdere il controllo dell'immagine.

Un altro aspetto di interesse è il fatto che, essendo stata l'intera chat impostata ad un ragionamento step-by-step, anche nella scrittura dei prompt ChatGPT allegava in coda al testo un'analisi per punti del prompt stesso, individuando le funzioni di ogni elemento. Si veda l'esempio seguente:

***Risposta di prompt da ChatGPT:** "A dramatic scene of a cowboy lying on the ground, mortally wounded, with a girl kneeling beside him. She gently caresses his face, her expression filled with sorrow and care. The cowboy's rugged clothes are dusty and bloodstained, and his face shows pain and*

## IA nella fase di pre-produzione

*resignation. The setting is an open prairie at sunset, with warm golden light casting long shadows. The atmosphere is emotional and cinematic, with hyper-detailed focus on their interaction."*

*Caratteristiche:*

- *Soggetto: Cowboy sdraiato e morente, con la ragazza che lo consola.*
- *Interazione Emotiva: Enfasi sul contatto fisico e l'espressione di dolore e cura.*
- *Ambientazione: Prateria al tramonto per aggiungere profondità e drammaticità.*
- *Stile: Cinematico e dettagliato per catturare il pathos della scena.*

Questo aspetto è utile per tenere traccia dei ragionamenti fatti dal modello e valutarne l'operato. Infatti, in un'analisi complessiva, ChatGPT si è dimostrato accurato nelle descrizioni e funzionale al suo scopo, semplificando e velocizzando di molto la pipeline di generazione di immagini.

## 7.4 IA per generazione Moodboard

Una volta creato un assistente artificiale addestrato alla scrittura dei prompt, non restava che utilizzarli, infine, per la generazione delle immagini. Una delle prime fasi nella pre-produzione di un'opera filmica è dedicata alla creazione delle moodboard, un elemento utilissimo per cominciare ad inquadrare l'estetica del prodotto e dare una reference visiva preliminare a tutti i reparti. È importante sottolineare il fatto che questa è una fase altamente precoce e tutt'altro che definitiva, il suo scopo è di definire una direzione in maniera anche circostanziale, ma che in corso d'opera subirà inevitabilmente molte modifiche e rivisitazioni. Diciamo che è significativo perché implica che non serve un'elevata precisione nella rappresentazione degli elementi in scena, non serve coerenza negli abiti, negli ambienti o nei props di scenografia, è sufficiente che riesca a comunicare il mood, per l'appunto, della scena.

Venendo al nostro caso studio, si sono voluti creare da uno a tre moodboard per ogni episodio, in base alla complessità. Per ogni episodio si è proceduto attraverso le seguenti fasi:

- **Scrittura dei prompt iniziali:** La scrittura dei prompt è avvenuta tramite la chat dedicata alla prompt engineering creata su ChatGPT nel capitolo 7.3. In prima battuta, si è chiesto in linguaggio naturale di generare un prompt per ogni moodboard, definendone gli elementi principali. Al fine di minimizzare la necessità di dare eccessive informazioni per iscritto, si è fornito in allegato alla chat la sceneggiatura di ogni episodio ed una breve descrizione dell'estetica desiderata (bianco e nero, dinamica thriller, ecc.). Così facendo, il modello ha unito le informazioni provenienti dalla sceneggiatura alle richieste fatte tramite messaggio, rendendo i prompt completi ed esaurienti, e mantenendoli coerenti tra le varie moodboard di uno stesso episodio.
- **Correzione iterativa dei prompt:** Successivamente, sfruttando l'abbonamento unlimited di Katalist, si sono testati ognuno dei prompt per generare le immagini sul modello online, con lo scopo di verificarne l'efficacia. Sulla base delle immagini generate da Katalist, si sono applicate successive correzioni ai prompt, inviando feedback a ChatGPT e ritestando i nuovi prompt di ritorno. Questo si è ripetuto iterativamente fino al raggiungimento del miglior risultato possibile, definito, dunque, da un prompt finale da poter utilizzare sugli altri modelli di GenAI. Il numero di correzioni è stato molto variabile a seconda della complessità della moodboard e di quanta precisione si desiderasse avere sulle composizioni delle immagini.
- **Generazione delle immagini:** Ottenuti dei prompt il più ottimali possibile, si è proceduto ad utilizzarli per generare le moodboard, tramite ChatGPT e Image Creator di Microsoft Designer, oltre a quelle già generate da Katalist. Per ogni modello si sono riscontrate esigenze diverse, rendendo talvolta necessarie delle correzioni ad hoc sul prompt.

Ci apprestiamo, ora, a mostrare ed analizzare i risultati generati divisi per episodio. Prima di procedere, tuttavia, è il caso di fare delle considerazioni di carattere generale. Katalist si è dimostrato uno strumento particolarmente deterministico, questo significa che è in generale più preciso nelle rappresentazioni, ma nell'ottica di creare moodboard questo aspetto non è sempre prioritario. Ciò significa che, seppur Katalist tendenzialmente sia stato più accurato nel raggiungere la composizione di immagine desiderata, DALL-E è stato più efficace nell'interpretare le emozioni che la scena doveva trasmettere, sfruttando anche eccessi creativi con successo per realizzare immagini più suggestive.

Un altro punto di interesse è la considerazione dei limiti del piano gratuito di ChatGPT: avere solamente tre crediti al giorno rende il processo estremamente lungo e frammentato. Sotto

questo punto di vista Microsoft Designer - seppur sempre limitato nei crediti – ha il vantaggio di proporre quattro alternative per ogni generazione, con elevate probabilità di essere soddisfatti da almeno una di esse. Introdurre questa variabilità ad ogni output aiuta a minimizzare gli interventi di correzione o i tentativi necessari. Con il modello di OpenAI, invece, è capitato di dover impiegare più giorni solamente per applicare piccole correzioni, rendendo il processo frustrante. Per questo motivo, si è apprezzato maggiormente il piano proposto da Microsoft rispetto a quello di ChatGPT.

### ***Episodio del Mistero e dell'Indefinito***

La scrittura dei prompt è proceduta in modo agevole. Si è voluto inviare la sceneggiatura a ChatGPT ed un breve prompt in cui si descriveva l'estetica desiderata, all'interno del quale si è spiegata l'ispirazione western e noir, l'uso delle luci, il bianco e nero e anche delle reference cinematografiche per chiarire ulteriormente l'obiettivo.

Per questo episodio si sono volute realizzare tre moodboard, per raffigurare tre punti cardine del racconto: l'incipit all'esterno della casa del cowboy, il matrimonio in chiesa ed il tragico finale dentro la casa. In questo caso, già in fase di pre-produzione, si aveva un'idea molto precisa di alcuni degli shot che si volevano realizzare, per cui già nel realizzare le moodboard si è deciso di indirizzare il più possibile le tavole verso l'output immaginato. Si è fatta, dunque, la seguente richiesta:

***Prompt di richiesta Ep.1:*** [...] Seguendo le istruzioni date, vorrei scrivessi tre prompt: 1) uno shot totale per la scena iniziale al mattino; 2) uno per il matrimonio in chiesa (abbastanza cupa); 3) uno per la scena finale della morte, per la quale vorrei una composizione da quadro classico, come della morte di Cristo.

Il modello ha risposto con prompt efficaci, in cui è stato necessario operare pochi aggiustamenti. In particolare, si è dovuto chiedere di ribadire in maniera più insistente che l'immagine fosse in bianco e nero, perché in diversi casi i modelli hanno generato solamente immagini a colori desaturati. È stato anche necessario, per l'ultimo prompt, aggiungere una descrizione della stanza del cowboy, in quanto non sufficientemente dettagliata in sceneggiatura e, di conseguenza, nel primo prompt generato. Analizziamo ora le immagini generate per ognuno dei prompt.

***Prompt moodboard 1 Ep.1:*** A black and white extreme long shot of a young cowboy rocking in a wooden chair on the porch of his isolated cabin. His hat tilts back, partially obscuring his weathered face as he smokes a thin cigar. A pair of dusty boots and an old rifle sit beside him. The dry prairie stretches beyond the porch, bathed in soft morning light. Suddenly, a mysterious young woman in a tattered white dress appears in front of him, barefoot, her gaze piercing. The atmosphere is stark and cinematic, with deep shadows and high contrast, evoking a classic noir western aesthetic.

## IA nella fase di pre-produzione



**Figura 7-3 Moodboard 1 Ep.1 - Katalist**



**Figura 7-2 Moodboard 1 Ep.1 – Microsoft Designer**



**Figura 7-1 Moodboard 1 Ep.1 – ChatGPT**

Dalle immagini della prima richiesta di moodboard si può notare facilmente una somiglianza tra le immagini generate da Microsoft Designer e ChatGPT, questo non ci stupisce alla luce del fatto che entrambi i modelli si appoggiano a DALL-E per la generazione di immagini. Risultano simili nella composizione e nell'uso della luce, oltre che nella rappresentazione dei personaggi e dell'ambiente, differendo solo per la presenza del fucile. Katalist si discosta in questo senso, riuscendo a posizionare i personaggi secondo le richieste del prompt, ossia uno di fronte

all'altro, ma fallendo nel rappresentare i due protagonisti secondo le istruzioni. Si nota, infatti, come il cowboy risulti troppo vecchio e la ragazza sproporzionata rispetto alla scena, oltre ad avere un'aura molto meno evocativa rispetto alle immagini di DALL-E.

Alla luce di ciò, seppur Katalist abbia creato un'inquadratura molto più simile alla scena finale girata nel film, DALL-E ha generato moodboard molto più rappresentative, più capaci di trasmettere le emozioni della scena.

***Prompt moodboard 2 Ep.1:** A black and white, high-contrast shot inside a small, old western style church with wooden walls and pews. A cowboy and a woman stand before the altar, their shadows stretching ominously across the wooden walls as they exchange vows. The priest, partially shrouded in darkness, presides over the brief ceremony, wearing a black cassock. The woman wears a delicate yet worn white dress and long blond hair. The lighting is moody, with thin beams of light piercing through dust-laden air from the high, arched windows. The atmosphere is mysterious and melancholic, you feel a sense of danger, reminiscent of classic noir cinematography, filled with heavy shadows and a haunting stillness.*

La seconda moodboard generata è stata ben interpretata da tutti i modelli, ne riportiamo solo un esempio in quanto tutte e tre le versioni simili. Si registra come unica differenza l'utilizzo della luce: laddove Katalist usa la luce in modo realistico seppur contrastato – come da richiesta – DALL-E ne fa un uso più artistico, avvicinando la resa più a un quadro che ad una foto, una resa molto apprezzata per una moodboard.



**Figura 7-4 Moodboard 2 Ep.1 – Microsoft Designer**

***Prompt moodboard 3 Ep.1:** A black and white, highly dramatic composition. The young cowboy lies sprawled on the ground, his hands pressed against a gunshot wound in his stomach, blood seeping through his fingers. His body is positioned in a way that echoes the lamentation of the dead. Kneeling beside him, the woman in a white dress with long blonde hair caresses his cheek, leaving a faint red trail of blood on his face. The dimly lit wooden cabin is nearly empty, with a single window casting soft beams of light through thin, swaying curtains. The chiaroscuro lighting carves out every detail. Black and white, dramatic chiaroscuro, deep shadows, cinematic lighting, film noir, high contrast, Caravaggio-inspired.*

## IA nella fase di pre-produzione

La generazione dell'ultima moodboard del primo episodio ha dato diversi problemi. Il primo si è riscontrato su Katalist, dove si è riuscito a raggiungere una composizione d'immagine perfetta, con i personaggi posizionati esattamente secondo le richieste e con l'ambiente ideale, ma dove non si è riusciti in alcun modo a rendere l'immagine in bianco e nero. Si è pensato che ciò fosse dovuto alla presenza di riferimenti a stili artistici (tipicamente a colori) all'interno del prompt, ma anche eliminandoli il modello a proseguito in questa direzione. Dopo diversi tentativi si è deciso di accettare l'imprecisione, tenendo come risultato finale la Figura 7-6, una reference di composizione perfetta, ma che poco comunica sull'uso della luce che si vuole ottenere.

DALL-E, seppur abbia ottenuto delle immagini in bianco e nero, ha dato un problema di altra natura in entrambi i modelli. L'immagine desiderata vuole rappresentare un uomo morente e del sangue e fa diretto riferimento ai quadri raffiguranti la morte di Cristo, queste richieste sembrano essere interpretate in modo negativo dai filtri di controllo delle policy di comportamento dei due modelli. Si è parlato in precedenza, però, del fatto che ChatGPT e Microsoft Designer utilizzino differenti politiche di censura, che sono più restrittive nel modello di casa OpenAI. Eliminando ogni riferimento al sangue nel prompt, Microsoft Designer è riuscito a restituire l'immagine desiderata, con un ottimo uso della luce e delle ombre, creando un quadro molto suggestivo. Purtroppo, lo stesso non si può dire di ChatGPT, che ha risposto con i seguenti suggerimenti:

***Risposta ChatGPT sulla policy:** [...] Probabilmente è a causa della descrizione esplicita della ferita e del sangue. Se vuoi, possiamo modificarlo leggermente per mantenere la drammaticità e il chiaroscuro senza dettagli troppo espliciti. Ad esempio, potremmo suggerire che il cowboy è ferito e debole, senza concentrarci sulla ferita in dettaglio. Fammi sapere come vuoi procedere! 📩*

Nonostante le buone intenzioni, però, il modello non è stato capace in nessun modo di modificare il prompt rendendolo rappresentabile. Questo non è avvenuto eliminando i riferimenti al sangue e alla ferita, non è successo rimuovendo la reference, dicendo che il cowboy dormiva, e in realtà neanche eliminando del tutto il cowboy. Questo comportamento si è ipotizzato che fosse dovuto al fatto che, una volta che viene rimosso un contenuto, il controllo sulla chat diventi più pressante, rendendo difficile riottenere il controllo sulla generazione. Per riuscire ad ottenere almeno un risultato, si è seguito il consiglio del modello di eliminare ogni tipo di riferimento emotivo dal prompt, in questo modo si è arrivati ad un'immagine di una stanza vuota con la sposa seduta al centro di essa, con una buona impostazione della fotografia, ma nulla altro di aderente alla scena, per questo motivo non viene riportata di seguito.



Figura 7-6 Moodboard 3 Ep.1 - Katalist



Figura 7-6 Figura 8 4 Moodboard 3 Ep.1 – Microsoft Designer

### ***Episodio dell'Angoscia***

Questo episodio ha le dinamiche di un thriller, in cui la tensione viene comunicata in modo implicito dalle immagini, per cui è particolarmente importante sfruttare la luce e la composizione dell'inquadratura in questo senso. L'episodio è di per sé piuttosto statico, la quasi totalità della scena si articola solamente attraverso i due piani dei protagonisti, a questo scopo si sono richiesti all'assistente di generare tre prompt:

**Prompt di richiesta Ep.2:** [...] Seguendo le istruzioni date, vorrei scrivessi tre prompt: 1) un'inquadratura totale in cui vediamo le due panchine ai lati del viale, quella su cui siede il maniaco è immersa nell'oscurità, mentre quella su cui siede la lettrice che legge è più illuminata. 2) Shot frontale della lettrice sulla panchina che legge ignara. 3) Shot frontale del maniaco sulla panchina che scrive sul suo diario, fissando morbosamente la ragazza

Per qualche motivo DALL-E ha interpretato le istruzioni come se le immagini fossero in bianco e nero, o comunque estremamente povere di saturazione; ciò ha reso necessario più interventi e correzioni sui prompt, per andare a chiarire il più possibile che le immagini dovessero essere a colori. Attuate le aggiunte, le immagini sono risultate a colori, ma caratterizzate da forti ombre e contrasto, in accordo con le dinamiche thriller. Al fine di differenziare i risultati di Microsoft e ChatGPT, si è deciso di usare per il secondo modello dei prompt che diminuissero la presenza di ombre, rendendo le moodboard più naturali.

**Prompt moodboard 1 Ep.2:** A wide shot of a quiet park avenue, lined with tall trees casting long, uneven shadows. Two benches face each other: on the right, in a sunlit patch, a young woman sits reading, unaware of her surroundings; on the left, drowned in heavy shade, a man sits slouched forward, his figure barely distinguishable in the darkness. His posture is tense, his hands gripping his knees, his gaze locked onto the woman. The contrast between bright daylight and the suffocating shadows makes the air feel heavy, as if something is about to break the silence. The atmosphere is tense, unsettling, and charged with an unspoken threat. Cinematic thriller, high tension, deep shadows, dramatic natural lighting, hyper-detailed, psychological unease.

In questa prima applicazione si notano differenze sostanziali tra le moodboard generate da Katalist e da DALL-E. Il primo modello ha applicato uno stile di pittura ad olio all'immagine, con una illuminazione estremamente naturale, molto vicina a quella utilizzata in fase di riprese; ha rispettato il posizionamento delle panchine, la presenza di ombre su un lato, e anche l'ambientazione che ben rappresenta il viale di un parco in pieno giorno. Sugli altri due modelli, invece, ci sono sia delle criticità che dei punti di forza. Se da una parte la rappresentazione è molto più cinematografica e realistica in termini di tratto, dall'altra l'uso della luce è molto incisivo, rendendo la composizione molto più evocativa e disturbante. Di nuovo riscontriamo, quindi, che Katalist meglio riesce a creare shot aderenti alla resa finale, ma DALL-E spinge maggiormente sull'emotività della scena. Un altro aspetto interessante è il fatto che DALL-E, in tutte le versioni generate, ha malinterpretato e posizionato il viale al centro di una strada cittadina, evidentemente fraintendendo dei termini che lo portano a immaginare un viale come un corso alberato di una grande città. Non si è parlato, infine, del fatto che DALL-E abbia faticato a posizionare personaggi e panchine efficacemente, prima mettendole sullo stesso lato,

## IA nella fase di pre-produzione

poi mettendo il maniaco in piedi, e con altri errori di questo tipo, richiedendo più tentativi prima di raggiungere un buon risultato.



Figura 7-9 Figura 8 5 Moodboard 1 Ep.2 - Katalist



Figura 7-8 Moodboard 1 Ep.2 - Microsoft Designer



Figura 7-7 Moodboard 1 Ep.2 - ChatGPT

*Prompt moodboard 2 Ep.2: A frontal medium close-up of a young woman sitting on a park bench, she's reading her book. Her expression is calm,*

## IA nella fase di pre-produzione

*absorbed, fingers lightly gripping the edges of the pages. Sunlight filters through the trees, creating shifting patches of brightness on her sweater and hair. The peaceful setting contrasts with a barely perceptible tension—something feels off, an unseen presence lurks beyond her field of view. The framing keeps the focus on her, yet the edges of the shot seem too open, leaving an unsettling void. Cinematic thriller, soft natural light, sharp contrast, psychological tension, hyper-detailed, uneasy atmosphere.*

Nella generazione della seconda moodboard non si registrano particolari annotazioni, in tutti i casi si è raggiunta un'immagine simile e ben costruita, in cui l'uso della luce, come da richiesta, era più naturale rispetto alla precedente. L'unico dettaglio degno di nota è il fatto che ChatGPT non sia riuscito a non far guardare in camera alla lettrice, un problema spesso riscontrato anche in fase di generazione dello storyboard, piuttosto comune nella GenAI. Riportiamo di seguito una sola tavola esemplificativa.



Figura 7-10 Moodboard 2 Ep.2 - Katalist

**Prompt moodboard 3 Ep.2:** *A tight frontal medium shot of a young man sitting hunched on a park bench, gripping a battered notebook in his hands. His fingers clutch the edges too tightly, his knuckles slightly pale. His head is tilted downward as if writing, but his eyes remain locked onto the woman in front of him, staring with disturbing intensity. Shadows partially obscure his face, but his fixation is unmistakable. The shot lingers, suffocating in its stillness, as if the world around him doesn't exist. Cinematic thriller, intense psychological tension, sharp natural contrast, hyper-detailed, unnerving presence.*

Per l'ultimo prompt le differenze sono maggiori invece, reintroducendo istruzioni di fotografia altamente contrastata tornano a differenziarsi Katalist e DALL-E. Il primo realizza una moodboard molto coerente con la precedente, in cui il senso di disagio è comunicato esclusivamente dallo sguardo ombroso e dalla postura del ragazzo. Il secondo realizza immagini molto potenti ed espressive, in cui il disagio è urlato in tutta la tavola. Si coglie l'occasione per ragionare sul fatto che un'eccessiva marcatura di certi aspetti, seppur suggestiva, possa diventare controproducente per la moodboard, dando un'idea sbagliata della messa in scena che si vuole ottenere. In un caso come questo, se si avessero avuto crediti a sufficienza da dedicarci,

si sarebbe probabilmente cercata una via di mezzo tra i due estremi, creando immagini verosimili ma emotivamente potenti.



Figura 7-12 Moodboard 3 Ep.2 - Katalist



Figura 7-12 Moodboard 3 Ep.2 - ChatGPT

### ***Episodio dell'Alienazione***

Questo episodio è fortemente drammatico, l'atmosfera deve essere molto più neutra rispetto ai precedenti due, si deve comunicare staticità, un dramma nascosto nel personaggio e non nell'ambiente, in cui viene enfatizzata l'alienazione e l'isolamento della protagonista. A tale scopo, si sono richieste tre tavole.

*Prompt di richiesta Ep.3: [...] il mood è basato sull'alienazione della madre e il tono generale deve essere drammatico. Ricorda di descrivere personaggi e ambienti, voglio che scrivi tre prompt: 1) Madre seduta al tavolo che fuma fissando il vuoto, l'inquadratura la isola e si vede il cellulare che squilla. 2) Total shot della madre e della zia sedute al tavolo a parlare preoccupate. 3) Inquadratura composta, a sinistra abbiamo la madre che sta uscendo sul balcone, a destra la zia ignara che sistema le cose in cucina*

L'unica correzione necessaria è stata di chiedere di inserire i riferimenti alla porta finestra nei primi due prompt. Un dettaglio non eccessivamente vincolante, ma che aiuta a rendere le tavole coerenti e illuminare la scena in un certo modo.

*Prompt moodboard 1 Ep.3: A medium shot of a middle-age mother sitting alone at a kitchen table, lost in a void-like stare. A cigarette burns between her trembling fingers, smoke curling in the dimly lit room. Behind her, a glass door to the balcony stands open, out there the city. The elegant but cluttered apartment feels strangely hollow—an untouched scotch bottle, an overflowing ashtray, a withered flower in a vase. On the table, a smartphone is ringing filling the silence. The woman doesn't react, frozen in her alienation. The composition isolates her in the frame, emphasizing her emotional detachment. Cinematic, psychological drama, muted colors, cold light, soft lighting, heavy atmosphere, hyper-detailed, dramatic composition.*

Questo primo prompt è stato efficace e ha fornito risultati molto vari. In questo caso non era necessaria una rappresentazione e composizione precisa, per cui è stato interessante vedere opzioni diverse di ripresa della scena e scoprire come esse possano trasmettere diversamente la situazione della madre. Ad, esempio, vediamo come Katalist abbia sfruttato molto la simmetria

## IA nella fase di pre-produzione

e l'isolamento della madre nell'inquadratura, una composizione che comunica in maniera sottile lo stato della protagonista, escludendola dal resto del mondo, quasi imprigionandola in un ambiente che si percepisce come estraneo e alienante. La stessa simmetria è stata proposta anche da ChatGPT, ma in una versione molto più disturbante, dove fotografia e ambiente trasmettono un malessere molto più esteriore oltre che interiore, e in cui si percepisce più un senso di attesa rispetto alla versione di Katalist, dove il tempo sembrava essersi fermato. Ancora diversa è la lettura da parte di Microsoft Designer, dove l'inquadratura diventa angolata ed è sempre più evidente quanto la madre sia provata ed in uno stato di disperazione.

Insomma, in questo caso le moodboard hanno saputo anche ispirare oltre a rappresentare la scena. Seguendo questa riflessione, sarebbe interessante provare a dare più libertà all'interno dei prompt, ampliando la gamma di possibilità raggiungibili dalla GenAI e sperimentare idee fuori dagli schemi iniziali. Ovviamente, per il proseguimento del progetto, è stato importante decidere quale delle versioni proposte più rispecchiasse i gusti registici, al fine di individuare una direzione univoca da seguire, ma essere sottoposti ad input diversificati è sempre fruttuoso in ambito creativo.



Figura 7-13 Moodboard 1 Ep.3 - Katalist



Figura 7-14 Moodboard 1 Ep.3 - Microsoft Designer



Figura 7-15 Moodboard 1 Ep.3 - ChatGPT

*Prompt moodboard 2 Ep.3: A wide shot of a tense conversation between two women seated at a kitchen table. The mother, dressed neatly but visibly drained, scratches her hands nervously, avoiding eye contact. The aunt, younger and more expressive, leans forward with concern, trying to console her. In the background, a glass door to the balcony is left ajar, revealing the cityscape beyond. A light breeze moves the sheer curtains, creating an eerie contrast with the tense stillness inside. The dimly lit room, with its elegant yet slightly chaotic decor, reflects the tension—an untouched glass of scotch, an abandoned cigarette, the faint hum of a distant city. Shadows stretch across the walls, amplifying the weight of unspoken words. Cinematic, dramatic tension, psychological drama, soft natural lighting, hyper-detailed, moody atmosphere.*

Il secondo prompt ha dimostrato come i modelli possano interpretare diversamente le istruzioni legate alla composizione dell'immagine, realizzando "wide shot" molto vari. La fotografia dei tre modelli è coerente con quella della prima moodboard, seguendo le tre direzioni intraprese. In generale lo shot è ben rappresentato nella sua semplicità, inserendo gli elementi essenziali richiesti e posizionando in maniera corretta le due donne; non ci sono altre particolari annotazioni da fare.



Figura 7-17 Moodboard 2 Ep.3 - Katalist



Figura 7-17 Moodboard 2 Ep.3 - ChatGPT

*Prompt moodboard 3 Ep.3: A carefully framed shot capturing two simultaneous actions in a single image. On the left the mother slowly steps onto a small ladder on the balcony, her movements mechanical, her face obscured looking outside. The open balcony door frames her silhouette against the distant cityscape. On the right, unaware, the young aunt tidies*

## IA nella fase di pre-produzione

*up groceries in the kitchen, lost in her own thoughts. The contrast between their actions creates an unbearable tension. The warm kitchen light clashes with the cold, eerie air of the balcony. The composition is perfectly balanced, yet the disconnect between the two figures deepens the emotional weight.*

Questa è probabilmente la moodboard più complessa tra tutte quelle generate. La si è voluta richiedere perché, fin dalla prima stesura della sceneggiatura, è stata un'inquadratura di reference ben definita e, pertanto, significativa. Con l'occasione si ha avuto modo di testare le capacità dei modelli di generare immagini dalla composizione estremamente dettagliata e complessa, in quanto nella scena sono previste due azioni parallele in ambienti comunicanti: da una parte abbiamo la madre, intenta ad uscire con una scala in balcone sulla sinistra, dall'altra la zia, affaccendata in cucina sulla destra. La descrizione è particolarmente fraintendibile dai modelli, sono molti gli elementi che si possono mal gestire, e per questo motivo si è cercato di semplificare il più possibile il prompt. Bisogna dire che i risultati hanno superato le aspettative, pur richiedendo qualche generazione in più rispetto alle altre immagini, si sono raggiunti risultati ottimi in tempi ragionevoli. L'unica applicazione fallimentare, seppur vicina al risultato sperato, è stata quella di ChatGPT, dove la gestione della scena ha dato vita ad artefatti di diverso genere, prima eliminando un personaggio, poi aggiungendone.



Figura 7-18 Moodboard 3 Ep.3 - Katalist



Figura 7-19 Moodboard 3 Ep.3 - Microsoft Designer



Figura 7-20 Moodboard 3 Ep.3 - ChatGPT

### ***Episodio Simbolico-Interpretativo***

L'episodio quarto, insieme al primo, è quello che maggiormente rendeva utile uno strumento come la moodboard, perché possiede uno stile particolare che è più difficile da definire rispetto agli altri. Per ben coordinare i reparti era importante trovare un punto di riferimento chiaro, una reference visiva che rendesse univoca la direzione artistica da seguire. Allo scopo, essendo un piano-sequenza, si è creato un solo prompt con le seguenti istruzioni.

*Prompt di richiesta Ep.3: [...] trasmettere un mood surreale e onirico, ispirandosi a Lynch, pur rimanendo nel mondo reale. Vorrei un prompt frontale al palco dove gli attori sono in semi cerchio attorno al cadavere. Solo il palco è illuminato e, nell'ombra della platea, l'unico spettatore se ne sta andando.*

Per ottenere il risultato ottimale sono state necessarie molte correzioni, spesso inefficaci. Infatti i risultati riportati in immagine non sono del tutto aderenti a quelli sperati. È stato necessario richiedere di descrivere maggiormente il palco e di specificare meglio che gli attori fossero cinque. Si è provato a caratterizzare i personaggi, ma i modelli diventavano inefficaci nella gestione di tante informazioni. Si è dovuto chiedere di semplificare il prompt, perché descriveva eccessivamente le azioni dello spettatore, finendo per perderne il controllo. Un altro problema di difficile gestione è stato il controllo del punto di vista: si desiderava un'inquadratura allo stesso livello dello spettatore, ma anche istruendo esplicitamente di specificarlo o impostando il prompt con lo spettatore in primo piano, l'inquadratura è risultata solo in pochi casi corretta. Di nuovo, DALL-E ha imposto dei filtri per impedire la rappresentazione del cadavere, ma fortunatamente è stato sufficiente eliminare l'aggettivo "lifeless" per rappresentarlo. Al netto di tutti gli interventi, ecco il prompt finale raggiunto.

*Prompt moodboard 1 Ep.4: A symmetrical frontal shot of a dimly lit theater. In the foreground, a lone spectator in a long coat and hat walks toward the exit, his figure partially silhouetted against the faint glow of the stage. On the stage, five actors stand in a semicircle around a lifeless body, their expressions unreadable. One wears a cowboy outfit, another a wedding dress, while the others are in vague, anachronistic attire. The wooden stage is nearly bare, with only an overturned chair near the body. The rest of the theater fades into darkness, amplifying the eerie silence. Cinematic, surreal, dreamlike, eerie lighting, Lynchian, unsettling atmosphere, dramatic composition, hyper-detailed, theatrical tension.*

## IA nella fase di pre-produzione

Le immagini mostrate in seguito sono state le migliori tra le diverse proposte ottenute durante il processo di correzione del prompt. Purtroppo, il numero elevato di elementi da gestire ha reso impossibile curare l'estetica quanto si avrebbe voluto, rendendole meno oniriche e surreali della visione iniziale. Tuttavia, con un po' di adattamento, si tratta comunque di tavole apprezzabili e utili a cominciare a definire la scena.



Figura 7-23 Moodboard 1 Ep.4 - Katalist



Figura 7-22 Moodboard 1 Ep.4 - ChatGPT



Figura 7-21 Moodboard 1 Ep.4 - Microsoft Designer

## ***Episodio Finale***

Questo è stato l'unico episodio gestito diversamente dagli altri. Essendo l'unico scritto in collaborazione con l'IA, si è pensato di lasciare maggiore libertà anche nella generazione delle moodboard. Ciò è avvenuto inviando la sceneggiatura a ChatGPT e chiedendo di provare ad immaginare due possibile moodboard per rappresentarla. Il modello ha risposto con due descrizioni di immagini, accompagnate da una spiegazione del perché fossero state scelte, indentificandone gli elementi significativi.

### ***Proposte Moodboard Ep.5 di ChatGPT: [...]***

#### ***1) Il Caos Creativo – Lo Scrittore nella sua Stanza***

Mood: *Disordine, procrastinazione, alienazione creativa, umorismo sottile.*

Elementi chiave: *Un giovane uomo trasandato sdraiato su un letto sgualcito.*

Stile: *Cinematico ma con una fotografia sporca, ispirata a Lebowski, Frances Ha, Noah Baumbach.*

- *La stanza è un disastro: snack aperti, vestiti ovunque, libri e fogli sparsi.*
- *Luce calda e fioca da una lampada da tavolo, contrasto con lo schermo acceso del laptop.*
- *In un angolo, un cappello da cowboy abbandonato, dettaglio eccentrico.*
- *Una cassa intelligente ribaltata sul pavimento, reduce di una battaglia.*

#### ***2) L'Estasi Creativa – Montaggio Finale di Scrittura***

Mood: *Flusso creativo, energia, caos organizzato, delirio da ispirazione.*

Elementi chiave: *Lo scrittore digitando freneticamente, con fogli e tazze sparse attorno.*

Stile: *Un mix tra Whiplash (ossessione), Kaufman (meta-narrazione) e un pizzico di Terry Gilliam.*

- *La luce dell'alba che filtra dalle finestre, segnando il passaggio della notte.*
- *Sovrapposizioni visive delle frasi che sta scrivendo, fluttuanti nell'aria.*
- *Il testo sullo schermo che si compone da solo, mentre lui continua a correggere ossessivamente.*

Le due proposte sono sembrate ideali allo scopo, sia dal punto di vista del soggetto che dell'estetica; anche le reference sono sembrate più che appropriate. Le poche indicazioni date in input sono state sufficienti ad assicurarsi che il modello centrasse il mood sperato e le proposte ricevute sono sostanzialmente identiche a quelle che si sarebbero richieste in autonomia. L'unico elemento che si è chiesto di implementare è la presenza del piatto di gnocchi e della diamonica, due elementi narrativi considerati significativi per la trama. Il passo successivo è stato, quindi, di chiedere di generare i prompt secondo le proposte fatte. Il primo prompt è stato scritto senza richiedere correzioni di alcun tipo; il secondo, invece, è stato necessario modificarlo, in quanto ChatGPT aveva inserito degli elementi troppo forti al suo interno, come parole fluttuanti ed un'energia creativa dello scrittore eccessiva, che lo

trasformava in un folle creatore urlante. Questo testimonia che è sempre comunque necessario un occhio attento in tutto il processo, se si vogliono evitare risultati indesiderati.

***Prompt moodboard 1 Ep.5:** A medium shot of a young, disheveled writer lounging on an unmade bed in a big bedroom, wearing a green worn-out bathrobe. His expression is vacant, staring at the ceiling while absentmindedly playing a melodica. The room is cluttered with scattered papers, open books, empty snack bags, and a forgotten cowboy hat lying in the corner. A dim lamp casts a warm glow, contrasting with the cold blue light from his laptop screen. The atmosphere is chaotic yet oddly peaceful. Cinematic, detailed clutter, warm and cool lighting, slightly surreal.*

***Prompt moodboard 2 Ep.5:** A dynamic wide shot of a young writer at his desk, deeply focused as he types on his laptop. He wears an oversized green bathrobe, slightly open, revealing a quirky T-shirt underneath. His expression is one of quiet determination, lost in the rhythm of writing. His table is creatively messy—coffee cups stacked in a precarious tower, a plate of gnocchi, and his melodica set aside. Warm afternoon light filters through the room, casting soft shadows. Cinematic, hyper-detailed, natural motion, warm and dramatic lighting, eccentric but thoughtful character, storytelling energy, a sense of creative flow and focus.*

Si sono ottenuti in risposta moodboard molto azzeccate, eccessive forse nella rappresentazione del caos, ma funzionali alla resa del mood della scena. Tutti gli elementi richiesti erano presenti e la rappresentazione del personaggio conforme alle aspettative, riuscendone a restituire l'aria svagata e atipica. Solo il secondo prompt ha richiesto qualche generazione aggiuntiva, perché tendeva a rappresentare il protagonista con un'aria troppo seria, ma sono bastati pochi tentativi aggiuntivi per raggiungere un miglior risultato.



Figura 7-25 Moodboard 1 Ep.5 - Katalist



Figura 7-25 Moodboard 1 Ep.5 - Microsoft Designer



Figura 7-27 Moodboard 2 Ep.5 - Katalist



Figura 7-27 Moodboard 2 Ep.5 - ChatGPT

## 7.5 IA per generazione Storyboard

Per la generazione degli storyboard si è fatto utilizzo di Katalist, esplorato nel dettaglio nel capitolo 5.5. Ricordiamo che il progetto si articola attraverso cinque episodi di genere e stili di ripresa profondamente differenti, questo ci ha permesso di testare Katalist in contesti plurimi e di verificarne la potenzialità in base alle esigenze.

Con Katalist sono stati realizzati gli storyboard dei soli primi tre episodi. L'Episodio Simbolico-Interpretativo è stato escluso in quanto articolato in un unico piano sequenza quasi completamente statico, inoltre Katalist è in grado di rappresentare solamente due personaggi all'interno della stessa scena, e l'episodio vede sempre in inquadratura almeno cinque personaggi, rendendo quindi irrealizzabile. Anche l'episodio finale non ha visto il proprio storyboard generato, perché non conteneva nessuna particolare sfida che non fosse già stata affrontata nei primi tre episodi, per cui a livello di ricerca non avrebbe aggiunto nulla all'analisi.

### 7.5.1 Analisi applicativa

L'analisi che segue è strutturata in una divisione per specifici aspetti della generazione con Katalist, portando per ogni punto esempi e comparazioni provenienti da tutti gli episodi realizzati. Si è deciso di dare questa forma all'analisi in modo da ottenere un quadro strutturato e meno dispersivo possibile.

#### *Stile dello storyboard*

Come mostrato al capitolo 5.5, Katalist propone diversi stili di rappresentazione tra cui scegliere. Delle dieci proposte, si è deciso di usare solamente gli stili “**Sketch**”, per l'episodio del Mistero e dell'Indefinito, e “**Cinematic**” per gli episodi dell'Angoscia e dell'Alienazione. Lo stile Sketch è quello che più si avvicina al concetto classico di storyboard, emulando uno schizzo a matita in bianco e nero, per cui è stato scelto per testarne l'aderenza visiva. Lo stile Cinematic, invece, potrebbe rappresentare un ulteriore passo avanti per la fase di previs di un'opera filmica: con il suo stile foto-realistico consente di avvicinare ulteriormente lo storyboard alla resa finale del film, rendendolo uno strumento sempre più accurato nel fornire una prima visione - semplificata - dell'opera.

Gli altri stili non utilizzati, per quanto suggestivi, si accostano a rappresentazioni sempre meno vicine al cinema e sempre più artistiche. Possono sicuramente essere valide scelte per progetti di storytelling meno tradizionali o completamente estranei al cinema, ma nel nostro caso di applicazione occorrono rappresentazioni quanto più neutre e con cifre stilistiche moderate. Fanno eccezione, però, gli stili legati al mondo dell'animazione, che nel caso di opere animate potrebbero risultare estremamente utili ed apprezzabili.



Figura 7-29 Stile Cinematic



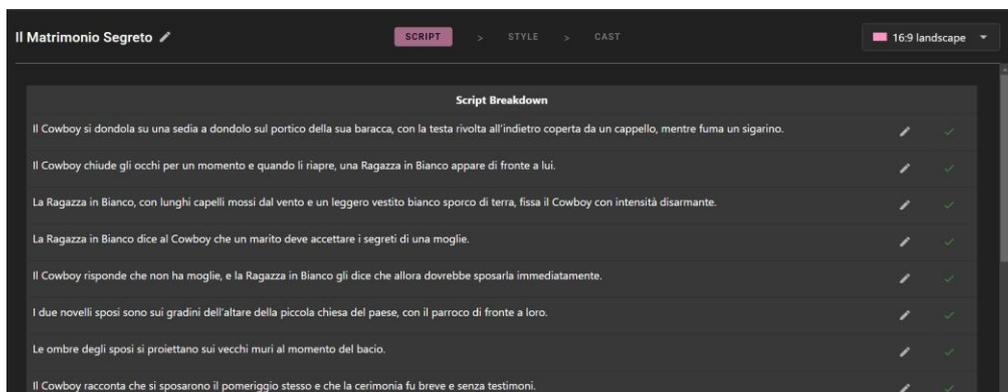
Figura 7-29 Stile Sketch

Lo stile sketch (Figura 7-29) aderisce perfettamente all'idea comune di storyboard, il tratto è chiaro nel rappresentare gli elementi, ma mantiene l'indefinitezza tipica dei disegni a matita. Se si cerca una rappresentazione neutra, che comunichi efficacemente la composizione dell'immagine senza imporre estetiche predominanti, questa è la scelta più consigliabile. Il tratto poco definito, inoltre, permette anche di notare meno difetti e artefatti che possono verificarsi nell'immagine, rendendo meno problematico avere, ad esempio, personaggi con volti rovinati.

Lo stile cinematic (Figura 7-29) riprende il classico fotorealismo delle immagini generative, il risultato sono veri e propri frame che sembrano tratti dal girato. Può essere entusiasmante per avere un'idea molto più avanzata di quella che sarà la resa finale del film e permette di impostare uno stile di fotografia che può rendere maggiormente il mood della scena. Tuttavia, bisogna tener conto che impostare la fotografia significa aggiungere una nuova variabile da gestire per garantire la coerenza nello storyboard, aggiungendo complessità.

### *Script Breakdown*

La prima generazione vera e propria del tool non è un'immagine, ma un testo, nella fattispecie un elenco delle inquadrature calcolate dall'IA sulla base dello script caricato. Si tratta di una divisione delle azioni principali in sceneggiatura, che Katalist permette di modificare o eliminare, ma non di aggiungerne di nuove. Fortunatamente, questo diventa possibile nella fase successiva.



**Figura 7-30 Katalist Script Break-down**

La divisione fatta da Katalist è molto semplicistica, non immagina mai dettagli o situazioni diverse da quelle standard, si limita a riconoscere azioni e battute nella sceneggiatura e a rappresentarle. Il modello fatica anche a distinguere alcune differenze della grammatica delle sceneggiature. Ad esempio, nell'episodio 1, il cowboy narra la vicenda attraverso un lungo voice-over, ma Katalist non sempre lo riconosce come una battuta extra-diegetica e capita, talvolta, che nello spoglio vengano inserite inquadrature provenienti da fatti solamente narrati.

Già dal principio, quindi, si capisce che per una costruzione dello storyboard in maniera automatizzata Katalist è insufficiente, per uno spoglio accurato ed originale occorre necessariamente intervenire manualmente.

## Generazione dei personaggi e delle location

Il secondo passo automatizzato è il riconoscimento di personaggi e location, che viene fatto per ogni inquadratura in maniera completamente autonoma. In questo processo Katalist si dimostra capace di individuare i soggetti in scena ed il luogo in cui si svolge l'azione, fallendo solamente in pochi casi.

Le location vengono generate autonomamente a partire dallo script, ma è possibile aggiungerne o modificarne il prompt di input in fase di storyboard. Esse sono risultate quasi sempre adatte alla scena e aderenti alla sceneggiatura, richiedendo, al massimo, piccoli aggiustamenti.

Per quanto riguarda la generazione dei personaggi, si è voluto sfruttare il fatto che, in fase di storyboard, fossero già stati effettuati i casting e scelti gli attori. Questo ha permesso di generare i personaggi tramite upload di immagine di input. Non possedendo i diritti per usare le foto degli attori, in questo elaborato non verranno mostrate le foto in input, ma solo quelle di output, di cui si segnalano risultati di elevata somiglianza con gli attori. Una stranezza riscontrata è il fatto che, nonostante l'accuratezza dell'immagine generata, spesso la descrizione annessa al personaggio fosse incoerente, specialmente riguardo a fattori come l'età o l'etnia. Questo è sintomo di importanti differenze tra il modo in cui noi umani percepiamo certi termini e di come li percepisca l'IA.



Figura 7-32 Generazione cowboy

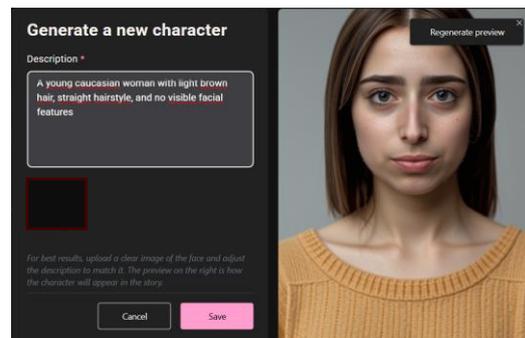


Figura 7-32 Generazione Lettertrice

## Generazione Storyboard

Terminato lo spoglio precedentemente descritto, viene generata una prima versione di storyboard con prompt automatici dell'IA. Esso risulta essere - consecutivamente allo spoglio - estremamente semplice ed inadatto. Le immagini generate sono nella maggior parte dei casi accettabili se prese singolarmente, ma quasi mai funzionerebbero montate in sequenza. Esse soffrono di incoerenza e sembrano non seguire alcuna regola o logica di montaggio. Di nuovo, Katalist richiede un intervento diretto dell'uomo per poter diventare utilizzabile.

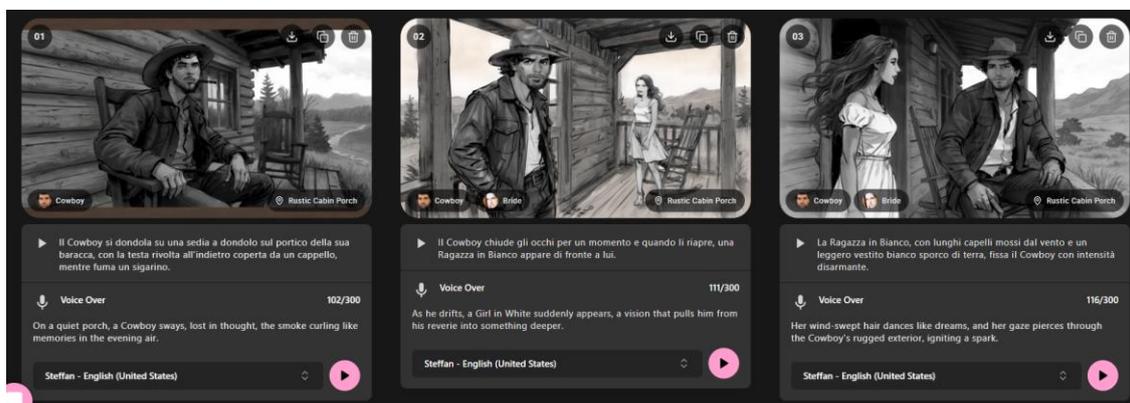


Figura 7-33 Ep.1: Incoerenza storyboard

### **Coerenza personaggi**

Occorre valutare, quindi, una caratteristica di punta di Katalist, ossia la sua coerenza nel rappresentare i personaggi. Si è visto come il modello sia stato in grado di realizzare immagini estremamente simili alle foto degli attori mandate in input, ma questa somiglianza diminuisce fin da subito nello storyboard, pur rimanendo a livelli di similarità accettabili. Sebbene sarebbe stato apprezzabile mantenere questa somiglianza, bisogna ricordare che ai fini della realizzazione di uno storyboard questo fattore non è il alcun modo impattante, mentre quel che davvero è importante è che i personaggi siano coerenti attraverso le varie tavole, rendendoli facilmente riconoscibili.

Un primo aspetto da analizzare è la coerenza dei vestiti, che vengono descritti nella scheda del personaggio. Sotto questo punto di vista Katalist riesce nel suo intento, portando in scena i personaggi sempre con gli stessi abiti o, se non altro, vestiti in modo simile. Un particolare peso viene dato agli accessori che, se opportunamente descritti all'interno della scheda, diventano essenziali per poter distinguere tra loro i personaggi della storia (si veda il Cowboy con il cappello o La lettrice con la collana).

Per la gestione dei cambi di outfit si è, però, obbligati a realizzare un nuovo personaggio ad hoc, che avrà le stesse identiche caratteristiche dell'altro, ma descrizioni degli abiti differenti. Questa procedura, seppur possa sembrare controintuitiva, è una pratica in linea con la fase di spoglio della sceneggiatura, in cui si è soliti identificare con id differenti i diversi outfit. Per cui, con le opportune prudenze, Katalist riesce a garantire la coerenza degli abiti.

Un discorso più complesso si ha per quanto riguarda la coerenza dell'aspetto fisico dei personaggi. Anzitutto, si riscontra un diverso livello di accuratezza in base allo stile di rappresentazione scelto per l'episodio. Nell'episodio uno, per cui è stato scelto lo stile sketch, i personaggi risultano sempre sufficientemente coerenti, seppur con piccole differenze. Ciò lo si deve al fatto che, questo specifico stile di disegno, renda meno percepibile certe imperfezioni, con il suo tratto poco definito diventa più difficile notare le differenze minori.

Un discorso diverso lo si ha invece negli episodi due e tre, che hanno utilizzato lo stile di rappresentazione cinematic. Dai due episodi non emerge alcuna analisi univoca, infatti, si registra che tre personaggi su quattro risultano fortemente coerenti negli storyboard, ma si hanno numerosi problemi con il personaggio della Madre nell'episodio tre. In Figura 7-34 si riporta l'aspetto della madre durante la generazione del personaggio, mentre in Figura 7-35 si può osservare diverse inquadrature provenienti dello storyboard. Come si può vedere, le diverse raffigurazioni sono molto distanti le une dalle altre, mostrando specialmente incoerenza nell'età del personaggio. La madre originale dovrebbe dimostrare circa cinquant'anni, ma nello storyboard passa dal dimostrarne trenta a dimostrarne settanta. Questo comportamento è atipico rispetto agli altri personaggi, i quali hanno sempre mostrato differenze minime tra una tavola e l'altra.



**Figura 7-34** Generazione originale personaggio Madre

Non si sono comprese nel dettaglio le ragioni di questo artefatto, ma segnala comunque una natura fallace del modello nel mantenere la coerenza nei personaggi. Ciononostante, con il resto dei personaggi si sono raggiunti ottimi risultati in questi termini, per cui sarebbe interessante indagare ulteriormente ampliando lo studio, in modo da ottenere una statistica più ampia e dettagliata sulle percentuali di accuratezza con diverse categorie di personaggi.



Figura 7-35 Diverse rappresentazioni del personaggio Madre

### *Coerenza location*

La selezione della location tramite tag per ogni shot riesce a rendere molto coerenti tra loro gli ambienti delle scene. Quello che avviene non è di ritrovare nelle immagini luoghi perfettamente identici, uno sguardo anche disattento potrebbe senza fatica notare molte differenze, ma gli ambienti si mantengono lo stesso consistenti in termini di colori, materiali, oggetti, illuminazione e mood. Guardando lo storyboard risulta ben chiaro quando uno stesso ambiente compare in più inquadrature, ma si presentano comunque delle complicitanze.

Un primo ostacolo è la disposizione degli oggetti: seppur il modello saprà sempre inserire tutti gli oggetti richiesti all'interno dello shot, non è detto che essi siano nelle stesse posizioni. Sarà complesso comunicare attraverso il prompt la posizione esatta degli oggetti, specialmente se presenti in gran numero, per cui è comune ottenere, nelle immagini, disposizioni degli elementi incoerenti. Un possibile approccio potrebbe essere quello di utilizzare il Generative Fill Tool, ma, come vedremo in seguito, questo strumento ha un'efficacia limitata. La tattica con più possibilità di successo, invece, è una generazione multipla con prompt il più accurati e descrittivi possibile. Mantenendo alti vincoli di controllo e generando più output, aumentano le possibilità di ottenere il risultato sperato. Ad ogni modo, ciò potrebbe non essere necessario in casi semplificati che prevedono pochi oggetti o elementi di scena, ma diventa indispensabile se gli oggetti in questione fossero particolarmente impattanti, quali mobili, alberi ed altri oggetti di grandi dimensioni.

Un altro limite di Katalist è il fatto che architetti solo piccole porzioni degli ambienti, focalizzandosi esclusivamente su di esse. Questo significa che il modello si occupa solamente di rendere coerente ciò che si visualizza in scena, senza fare nessun tipo di ragionamento sulla struttura dell'ambiente circostante. La diretta conseguenza è cui risulta difficile avere prospettive diverse di uno stesso luogo, laddove in realtà è comune nel montaggio sperimentare differenti punti di vista. Per provare a risolvere questo limite si sono testati diversi approcci nel realizzare lo storyboard per i tre episodi.

Nell'episodio dell'angoscia, essendo l'ambiente molto semplice e caratterizzato da una sola panchina in un viale alberato, è stato semplice rendere le immagini coerenti tra loro, perché non esistono punti di riferimento che possano far percepire l'errore.

Nel primo episodio, la casa con portico del Cowboy avrebbe dovuto essere un ambiente facile da riprodurre, in quanto possiede pochi e semplici punti di riferimento. Apparentemente, infatti, tavole risultanti sono state molto coerenti, ma per il semplice motivo che usavano sempre la stessa inquadratura laterale. A questo punto, in questo episodio, la principale sfida è diventata proprio quella di cambiare prospettiva di inquadratura (Figura 7-36).



**Figura 7-36 L'impostazione della location porta con sé dei vincoli di inquadratura**

Il problema alla base dell'errore risiede nel fatto che Katalist, per garantire la coerenza tra le immagini, probabilmente utilizza le altre immagini con la stessa ambientazione come input parziale per le nuove immagini. Questo tipo di funzionamento porta il modello a riprodurre anche la stessa prospettiva di inquadratura, rendendo, per esempio, difficile ottenere un'inquadratura frontale e una laterale dello stesso soggetto. L'unico approccio di successo che abbiamo testato è stato di aggiungere molte informazioni spaziali all'interno del prompt di input, ma anche in questo modo sono spesso stati necessari più tentativi prima di ottenere un risultato accettabile o soddisfacente.

In casi estremi la soluzione migliore si è rivelata quella di creare una location completamente nuova, nonostante si trattasse fondamentalmente della stessa. Ad esempio, per generare il dettaglio di una rosa e di un calice sull'altare della Chiesa, è stato necessario realizzare una nuova location "Altare chiesa"; questo perché l'utilizzo della location originale imponeva una prospettiva sbagliata della scena e portava il modello ad inserire elementi errati nella prospettiva attuale. In Figura 7-40 si può notare come la location iniziale imposti l'inquadratura con una prospettiva rivolta verso la navata centrale, al fondo della quale si trova l'altare che, in realtà, dovrebbe essere in primo piano. Questo avviene perché, nelle altre inquadrature ambientate in chiesa, si aveva la camera all'ingresso della chiesa puntata verso l'altare, per cui il modello pensa di doverlo sempre inserire al fondo della stanza. Nella Figura 7-40 si può invece vedere come l'approccio di creare una location ad hoc si sia rivelata una tattica vincente, privando la generazione di vincoli di inquadratura che la rovinassero.



Figura 7-38 Dettaglio errato - Location "Chiesa"



Figura 7-38 Dettaglio corretto - Location "Altare chiesa"

In altri casi, invece, è stato necessario eliminare del tutto qualsiasi riferimento a location. Per realizzare un primo piano del Cowboy morente disteso sul pavimento si è dovuto togliere ogni riferimento a location perché, se inserita, il modello si “obbligava” a rappresentarla, anche se in inquadratura non avrebbe dovuto essere visibile (**Errore. L'origine riferimento non è stata trovata.**). Allo stesso modo si è dovuto procedere per realizzare un piano medio frontale del Cowboy seduto sotto al portico, perché ogni riferimento alla location portava l'inquadratura ad essere laterale rispetto al portico (Figura 7-38).



Figura 7-40 Primo piano Cowboy con artefatto dovuto a location



Figura 7-40 Piano medio Cowboy con artefatto dovuto a location

Nell'episodio terzo, che richiedeva una grande coerenza d'ambiente essendo ambientato dentro un appartamento, si è ricorso a tutti gli espedienti sopraccitati per realizzarne lo storyboard. Ad esempio l'ingresso alla casa sarebbe dovuto essere connessa alla stessa stanza in cui si trova la Madre, ma alle spalle dell'inquadratura proposta da Katalist (Figura 7-42). Per farla si è realizzata una nuova location fittizia, dedicata alla stanza d'ingresso (Figura 7-42) pur essendo parte comunicante del salotto.



Figura 7-42 Ingresso zia con artefatto dovuto a location



Figura 7-42 Ingresso Zia corretto con location dedicata

Un metodo ulteriore, utilizzato in pochi casi, per impostare a piacimento la scena è di inviare una immagine di reference in input al modello: in questo caso Katalist riesce ad emulare piuttosto bene la composizione inviata, replicando prospettiva ed elementi. Questo tipo di applicazione è, però, più dispendiosa rispetto alle altre, perché implica di dover realizzare o trovare una foto che abbia esattamente la forma che cerchiamo.

In definitiva, la coerenza delle location è garantita se impostata, ma porta con sé molti vincoli che possono rendere impossibile ottenere il risultato sperato. Con opportuni espedienti è possibile raggiungere buoni risultati in termini di inquadratura, ma questo va a discapito della coerenza. Con un bilancio di compromessi e con un maggiore uso di prompt precisi e dettagliati, si può realizzare un buono storyboard, tanto più accurato quanto più gli ambienti sono semplici.

### *Composizione immagine*

Se è vero che Katalist riesce a mettere in scena quello che viene inserito nel prompt, non è però facile impostare la composizione d'immagine desiderata. Questo è dovuto specialmente ai vincoli aggiuntivi che il modello impone per rendere coerenti ambienti e personaggi, un processo che tende a rendere le inquadrature omologate sotto molti aspetti. Questo significa che inquadrature di ambienti poco ricorrenti diventano più semplici da gestire, con una risposta più incisiva agli input, mentre diventa sempre più debole il controllo nelle scene con ambienti più utilizzati.

Le impostazioni di Shot Type e Camera Angle fornite dall'interfaccia si rivelano sufficientemente efficaci in casi di prompt molto semplici e corti. Purtroppo, nella maggior parte dei casi, i prompt utilizzati erano complessi e articolati, per poter avvicinare l'output allo scenario immaginato, e le impostazioni aggiuntive diventavano poco influenti sull'output, specialmente lo Shot Type. L'unico modo, a questo punto, di mantenere un pieno controllo della composizione è scrivendo con precisione il prompt. L'uso di ChatGPT in questo caso si è dimostrato fruttuoso, attraverso un lavoro iterativo basato su continue correzioni è riuscito a fornire prompt sempre capaci di avvicinare l'output alla risposta sperata.

Pur migliorando l'immagine generata, non vi era una risposta sempre coerente alle sollecitazioni, applicando le modifiche del prompt con forza altalenante. ChatGPT riusciva con successo a definire angolo di camera e tipo di inquadratura, ma comunque riscontrava difficoltà nello specificare la posizione degli elementi, ora intendendo la destra come destra dell'immagine, ora intendendola speculare. Questa limitazione ha richiesto molto tempo per correggere piccoli dettagli. In alcuni casi si è deciso di ricorrere alla post-produzione per specchiare l'immagine, ritenendolo più veloce, in altri ci si è dovuti accontentare di risultati imperfetti.

Ad esempio, nell'episodio primo, si voleva realizzare un'inquadratura della navata durante il matrimonio dal punto di vista dell'altare. Dopo diverse correzioni, si è dovuto iniziare a inserire numerosi riferimenti spaziali per voltare l'inquadratura verso l'ingresso della chiesa, fino ad arrivare al prompt e risultato seguente.

***Prompt:** A middle shot from above of @Cowboy and @Girl stand in front of the altar, exchanging vows. Behind them, the church interior, viewed from the altar, looking down the nave towards the entrance door closed at the far end of the church. The wooden pews are arranged symmetrically along the nave, leading to the distant entrance door, which is framed by soft light filtering through stained glass windows. The camera angle is slightly elevated, looking down on the scene, adding depth and perspective. The*

*atmosphere is solemn and serene, with warm, golden lighting and a hyper-detailed, cinematic composition.*



Figura 7-43 Ep.1 - Composizione inquadratura chiesa

Si nota come nel prompt si siano ripetuti più volte gli stessi concetti spaziali per assicurarsi un risultato coerente. Eppure, nonostante le molte premure, la porta sullo sfondo non è chiusa, ma aperta, e dietro di essa sembra esserci qualcosa di simile ad un altare; non solo, ma le panchine sono rivolte al contrario rispetto a come dovrebbero essere. Questo dimostra come Katalist cerchi sempre di ricondurre l'immagine ad una reference ben definita e faticosi a modificarla eccessivamente.

Un discorso analogo si può fare anche con i personaggi, infatti lo stesso problema si presenta anche inserendo il tag di un personaggio nel prompt. Katalist risponderà dirigendo la generazione verso un certo tipo di composizione, in cui il personaggio è sempre molto visibile, tipicamente con il viso visibile.

Ad esempio, nell'episodio secondo, si voleva realizzare una inquadratura di quinta dal punto di vista del Maniaco, intento a spiare la Lettrice (Figura 7-45); per farlo è stato necessario realizzare un prompt in cui, al posto del maniaco, ci fosse un uomo generico, perché - laddove il personaggio veniva menzionato - Katalist faceva in modo che fosse ben visibile in viso (Figura 7-45).

***Prompt finale:** Over-the-shoulder shot from the perspective of a maniac sitting on a park bench, watching @Lettrice on the bench in front of him. She is engrossed in her phone, texting, unaware of his presence. The shot is framed by foreground elements, partially obscuring the view, enhancing the sense of unease. The soft afternoon light filters through the trees, casting dappled shadows on the path. The atmosphere is tense and cinematic, with a subtle sense of suspense. Hyper-detailed, shallow depth of field, high contrast, realistic lighting*



Figura 7-45 Ep.2 Ripresa di quinta senza menzione



Figura 7-45 Ep.2 Ripresa di quinta con menzione

### *Dettagli e particolari*

Merita un paragrafo dedicato la gestione di dettagli e particolari, due elementi che raramente vengono proposte in modo autonomo da Katalist, ma che sono necessari in qualsiasi opera filmica.

L'inquadratura che meno ha dato problemi è quella dei dettagli, ossia di oggetti inanimati. Le performance generali del modello sono molto buone, riuscendo agilmente a realizzare dettagli curati in tutti gli episodi (seppur a volte si è dovuto togliere il tag alla location). Particolarmente sorprendenti sono stati i dettagli del diario del Maniaco nell'episodio due (Figura 7-47), che avevano la complessità di avere elementi scritti e fotografie incollate sulle pagine; particolari chiaramente rinunciabili, ma che Katalist ha gestito egregiamente senza necessità di modificare di molto il prompt.



Figura 7-47 Ep.2 Dettaglio diario 1



Figura 7-47 Ep.2 Dettaglio diario 2

L'unico dettaglio che ha dato problematiche rilevanti è quello della canna del fucile del Cowboy dopo uno sparo. Sicuramente è un dettaglio particolare, che prevede un fondale nero, un forte gioco di fuoco, il fumo che si alza dalla canna ed un angolo di camera ben preciso, ma in questo caso si sono riscontrate molte difficoltà. La prima è stata nel realizzare il fondale: anche creando una location dedicata in cui non vi altro che oscurità, il dettaglio fatica a visualizzare un background nero semplice. La seconda sfida è stata posizionare il fucile nell'angolazione desiderata, che ha richiesto numerosi tentativi. Ma a stupire è la quantità di artefatti che si sono riscontrati sul fucile, che è risultato sempre con elementi fuori posto o errati, siano essi la canna corta, il fumo che esce da punti incoerenti, grilletti fuori posto e molto altro. Tuttavia, nell'ottica di usare lo storyboard come guida approssimativa, il risultato finale è più che accettabile.

*Prompt: Close-up of the top of a smoking barrel of an old rifle, held by a female hand, surrounded by a black plate. The background is blurred,*

*keeping the focus on the rifle barrel. Hyper-detailed, cinematic, sharp focus, noir, western.*



Figura 7-48 Ep.1 Dettaglio fucile

Per quanto riguarda i particolari dei personaggi, la situazione è risultata più complessa. Come detto per la realizzazione della Figura 7-45, inserire la menzione di un personaggio porta Katalist a rappresentare tutto il personaggio. Questo diventa un problema nell'ottica di realizzare dettagli di determinate parti del corpo. Dopo diversi approcci, si è giunti alla conclusione che le strade percorribili siano due:

- Se il dettaglio rappresentato non ha nulla di tanto caratteristico da giustificare la necessità di usare il personaggio, conviene togliere la menzione e descrivere in termini generali la parte del corpo.
- Se è necessario usare la menzione in quanto riconoscibile, bisogna partire da un'inquadratura larga e sfruttare l'**Image Cropping Tool** per ritagliare la parte interessata.

Una volta apprese queste due tecniche, isolare i dettagli si è rivelato semplice e veloce.



Figura 7-50 Ep.2 Dettaglio con cropping



Figura 7-50 Ep.1 Dettaglio con cropping

### ***Character Posing Tool***

Una delle funzioni principali di Katalist è la possibilità di riposizionare i personaggi e modificarne la postura attraverso un sistema di scheletri di controllo. Durante l'utilizzo, è emerso che il processo non si limita a un semplice riposizionamento dei personaggi all'interno della stessa immagine, ma genera una nuova immagine basandosi sulla posizione degli scheletri. Questo fenomeno è ben evidente nel passaggio dalla Figura 7-52 alla Figura 7-52, dove il cambiamento della posizione dei due protagonisti ha comportato non solo una

variazione negli abiti della ragazza, ma anche una completa rigenerazione dello sfondo. Questo evidenzia una criticità: modificare le posizioni può compromettere la coerenza della scena inizialmente impostata, introducendo potenziali errori. Tuttavia, il caso analizzato rappresenta un'eccezione, poiché nella maggior parte dei test le rigenerazioni risultano sufficientemente coerenti, pur con alcune differenze visibili.



Figura 7-52 Ep.1 Riposizionamento



Figura 7-52 Ep.1 Rigenerazione con nuove posizioni

Il Character Posing Tool offre due principali funzionalità: la possibilità di spostare spazialmente i personaggi nella scena e quella di modificarne la posa.

Il primo utilizzo si è dimostrato il più efficace e affidabile, permettendo di scalare e riposizionare i personaggi con precisione, consentendo di modificarne la distanza reciproca e la composizione dell'inquadratura. La scalatura degli scheletri si è rivelata particolarmente utile per trasformare mezze figure in primi piani o adattare campi totali in campi lunghi. Ad esempio, nella Figura 7-43, Katalist tendeva a inserire Cowboy e Sposa a figura intera al fondo della chiesa; con consecutivi aggiustamenti di scala e posizione si è riusciti a portare la coppia in primo piano in mezza figura, senza necessità di modificare il prompt.

Si registra anche che la tecnica si dimostra maggiormente efficace se è presente un solo personaggio, mentre diventa più fallibile con due. Tuttavia, questa tecnica si è dimostrata più efficace con un solo personaggio, mentre risulta meno affidabile con due soggetti. Nel terzo episodio, ad esempio, la Madre e la Zia sono sedute al tavolo e, nonostante nella maggior parte delle immagini sia stato possibile mantenere la madre alla destra dell'inquadratura, nell'unica ripresa in cui entrambe sono sedute al tavolo il sistema non ha risposto correttamente ai comandi del prompt per il posizionamento. Per ovviare al problema, si è tentato di scambiare la posizione degli scheletri (**Errore. L'origine riferimento non è stata trovata.**), ma il risultato è rimasto invariato. Ciò suggerisce che gli scheletri non trasportano informazioni relazionali con il personaggio, ma descrivono esclusivamente la posizione di due figure generiche, che Katalist associa ai nomi unicamente sulla base del prompt. In questo caso la difficoltà principale sembra derivare dalla somiglianza delle posture e dalla condivisione dello stesso spazio, che rende più complesso per il modello distinguere i due soggetti.

Diverso è il discorso relativo alla gestione delle pose. Gli scheletri forniti per i personaggi offrono numerosi punti di controllo che permettono di regolare le articolazioni in modo dettagliato, ma l'interfaccia non sempre risulta intuitiva. La visualizzazione bidimensionale di una struttura tridimensionale può generare sovrapposizioni, rendendo complessa la selezione accurata dei punti di giunzione. Nonostante queste difficoltà, il tool si è rivelato estremamente utile per la gestione degli arti, con risultati soddisfacenti nella manipolazione degli stessi. Un caso emblematico è rappresentato dalla Figura 7-54: per ottenere un dettaglio dello stivale del cowboy, Katalist tendeva a includere necessariamente il volto del personaggio. Per

aggirare questa limitazione, è stato prima impostato manualmente il posizionamento del cowboy e successivamente utilizzato l'Image Cropping Tool per ritagliare l'inquadratura sulla parte desiderata.

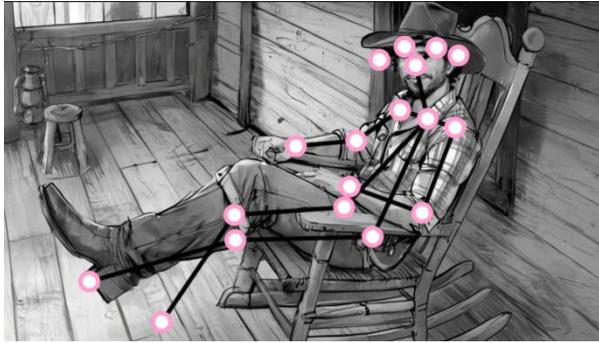


Figura 7-54 Ep.1 Riposizionamento gamba



Figura 7-54 Ep.1 Dettaglio gamba finale

Sebbene il movimento degli arti sia stato relativamente semplice da gestire, la regolazione della testa ha presentato maggiori difficoltà. Quest'ultima dispone di cinque punti di controllo, quattro dei quali dovrebbero permettere di gestire inclinazione e direzione dello sguardo.

Nei test condotti sugli episodi due e tre, è stato spesso necessario indirizzare gli sguardi dei personaggi, ma con risultati non sempre coerenti. L'uso del prompt, con descrizioni dettagliate della direzione dello sguardo o riferimenti a oggetti specifici, si è rivelato efficace solo in alcuni casi. Si è quindi tentato di manipolare direttamente i punti di controllo della testa, anche esasperandone l'orientamento, ma il sistema non ha mostrato risposte significative. Analizzando il comportamento del modello, si è osservato che la difficoltà si manifestava prevalentemente quando il soggetto doveva guardare elementi esterni all'inquadratura (Figura 7-55). Al contrario, se due personaggi dovevano fissarsi reciprocamente o osservare un oggetto presente nella scena, il risultato risultava accurato. Questo suggerisce che il modello non è in grado di rappresentare con precisione elementi non inclusi nella composizione visiva generata, limitando la gestione di determinate interazioni visive.



Figura 7-55 Ep.3 Controllo sguardo ininfluyente

### ***Espressioni dei personaggi***

Un elemento cruciale per ottenere immagini accurate è il controllo delle espressioni facciali dei personaggi. Nei primi due episodi, Katalist, in combinazione con i prompt forniti da ChatGPT, è riuscito a rappresentare con precisione le emozioni richieste. Stranamente questo non è

avvenuto nel terzo episodio, che paradossalmente sarebbe quello più drammatico da sceneggiatura. Sebbene la maggior parte degli shot abbia catturato correttamente il mood della scena, in alcuni casi la Madre è stata raffigurata con un'espressione inappropriatamente sorridente sin dalla prima generazione. Questo errore iniziale si è propagato nelle versioni successive, come se Katalist mantenesse una sorta di memoria visiva e riutilizzasse elementi delle immagini precedenti. Curiosamente, riproponendo lo stesso prompt in un nuovo shot privo di storico, il modello ha generato immediatamente il risultato desiderato. Questo ha portato alla conclusione che, in alcuni casi, è più efficace ripartire da zero con una nuova immagine anziché tentare di correggerne una preesistente.

Un'altra criticità riscontrata riguarda la difficoltà del modello nel distinguere aggettivi riferiti a personaggi diversi all'interno dello stesso prompt. Se entrambi i soggetti devono esprimere la stessa emozione – ad esempio, essere tristi o guardare verso il basso – Katalist esegue correttamente l'istruzione. Tuttavia, quando le indicazioni differiscono (ad esempio, un personaggio che ride e l'altro che appare serio), il modello tende a confonderle o ad uniformare le espressioni. Questo limite è comprensibile considerando che, fino a poco tempo fa, Katalist supportava la generazione di un solo personaggio per immagine, con l'aggiunta di un secondo soggetto solo nelle versioni più recenti. Ciò suggerisce che la gestione simultanea di più espressioni rimanga un aspetto ancora in fase di affinamento nello sviluppo del tool.

### *Negative prompt*

Oltre all'uso dei prompt standard, Katalist offre la possibilità di inserire negative prompt, pensati per rimuovere elementi indesiderati o correggere artefatti nelle immagini generate. Sono stati effettuati diversi tentativi per sfruttare questa funzione, ma i risultati positivi sono stati estremamente limitati. L'unico caso di successo si è verificato nel primo episodio, dove la Ragazza in bianco viene descritta come una presenza angelica. Katalist ha interpretato la descrizione in senso letterale, aggiungendole un paio di ali (Figura 7-57). Con l'uso del seguente negative prompt, però, si è riusciti ad eliminarle con successo.

*Negative Prompt: Angel wings, wings, feathers, flying, unrealistic elements, fantasy accessories, glowing wings.*

*Prompt: Frontal close-up of @Girl's face, perfectly symmetrical, with an angelic aura glowing softly around her. Her expression is serene and radiant, with the vast prairie stretching out in the blurred background. Warm, golden sunlight enhances her delicate features, creating a heavenly and ethereal atmosphere. Hyper-detailed, cinematic lighting, soft focus on the background.*



Figura 7-57 Ep.1 Generazione senza negative prompt



Figura 7-57 Ep.1 Generazione con negative prompt

Al di fuori di questo episodio, però, la funzione non ha prodotto risultati soddisfacenti. È possibile che ChatGPT sia stato inefficace nella scrittura dei negative prompt o che le altre situazioni fossero meno chiaramente definibili, ma di fatto lo strumento non ha mostrato particolare utilità nello studio.

### ***Generative Fill Tool***

Il Generative Fill Tool è lo strumento che meglio ha saputo sopperire alle mancanze dei negative prompt, consentendo di eliminare o sostituire determinati elementi della scena. A differenza del negative prompt, questo tool non rigenera l'intera immagine, ma mantiene intatta la scena al di fuori della maschera selezionata.

La sua applicazione si è dimostrata fruttuosa nella correzione di dettagli di oggetti, pur riportando piccoli artefatti. È stato utilizzato in due casi principali. Nel secondo episodio, dopo diverse iterazioni, si è ottenuta un'inquadratura ideale con il maniaco posizionato perfettamente nella scena. Tuttavia, nell'immagine mancava la panchina su cui il protagonista era seduto. Grazie al tool si è riusciti a preservare la composizione originale e ad aggiungere la panchina (Figura 7-61), seppur che lievi artefatti che, però, ai fini di uno storyboard non sono incisivi. Occorre evidenziare che il risultato mostrato è il migliore di quattro tentativi, è sempre consigliabile generare più opzioni tra cui scegliere, perché raramente una sola è sufficiente.

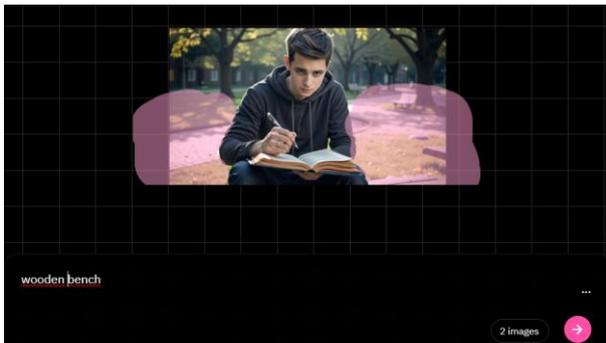


Figura 7-61 Ep.2 Generative fill tool per aggiunta panchina

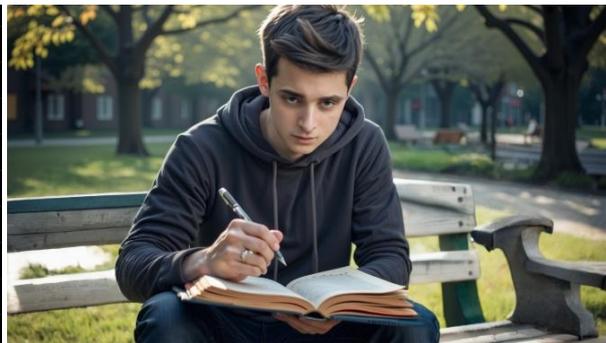


Figura 7-61 Ep.2 Output finale con panchina

La seconda applicazione è avvenuta nell'episodio tre, dove si è reso necessario sostituire una bottiglia di whiskey e un bicchiere con un vaso generico. Lo strumento ha rimosso con successo i due oggetti, pur mantenendo l'ombra colorata del liquido sul muro, e al suo posto ha inserito un atipico vaso/giara pieno di sabbia (Figura 7-59). Seppur non fosse il risultato atteso, era sufficiente a soddisfare l'obiettivo primario: eliminare la bottiglia. Se l'elemento sostituito avesse avuto un valore narrativo significativo, sarebbe stato necessario iterare ulteriormente per migliorarne l'accuratezza.



Figura 7-59 Ep.3 Sostituzione oggetto con Generative Fill Tool



Figura 7-59 Ep.3 Output finale sostituzione oggetto con Generative Fill Tool

### *Ulteriori punti di analisi*

Durante la realizzazione degli storyboard si sono riscontrati altri punti di analisi degni di nota.

Un primo fattore positivo è un'ottima resa del motion blur, ben evidente nel passante in Figura 7-62. Questa capacità risulta particolarmente utile per storyboard di film dinamici, in cui il movimento gioca un ruolo chiave.



Figura 7-62 Ep.2 Motion blur

Tra le limitazioni del modello, si è riscontrata la difficoltà nel gestire personaggi duplicati nella stessa scena. Ad esempio, nell'episodio tre, vedeva la presenza simultanea della Madre seduta al tavolo e della sua immagine in fotografia, una situazione che il modello non è stato in grado di gestire per una questione di conflitto delle direttive. Un modo semplice per aggirare il problema sarebbe creare un personaggio dedicato, che in questo caso potrebbe essere "Madre fotografia", che sia identico all'altro ma con un identificativo diverso.

Un aspetto trasversale, ma che merita di essere trattato, è il fatto che nelle inquadrature in cui il Maniaco tocca la gamba e stringe il collo della Lettrice la generazione abbia, a un certo punto, preso una piega fortemente erotica. Questo aspetto è rilevante in termini etici, sollevando dubbi sui limiti da imporre alla generazione dell'IA e sulle implicazioni che un suo uso scorretto potrebbe portare.

Infine, merita una considerazione la velocità di generazione delle immagini, che in media impiegava da uno a due minuti per immagine, rendendo Katalist estremamente reattivo e dando la possibilità di iterare più tentativi senza rallentare il flusso di lavoro, rendendo il processo di perfezionamento delle immagini estremamente efficiente.

## **7.5.2 Esempio di caso applicativo**

Per completezza dell'esposizione del lavoro svolto, si è pensato di presentare un'analisi esemplificativa della realizzazione di una tavola dello storyboard. L'inquadratura scelta proviene dall'episodio dell'Angoscia e vuole rappresentare un dettaglio in cui la mano del Maniaco accarezza il collo della Lettrice.

Si è partiti facendo generare a ChatGPT il prompt da immettere su Katalist; partendo da una piccola descrizione, la prima versione di prompt proposta è stata la seguente.

***Prompt Iniziale:** A tense close-up of a man's hand gripping from behind the @Lettrice's neck as she sits on a park bench. Her posture stiffens in shock, her fingers slightly tensed on her lap. The man's hand appears rough and forceful, veins visible, contrasting with the softness of her skin. The blurred background of the park adds to the sense of isolation. The lighting is natural but subdued, casting subtle shadows that enhance the tension. The*

*atmosphere is unsettling and cinematic, evoking fear and suspense. Hyper-detailed, dramatic composition, shallow depth of field, moody lighting*

Si nota la ricchezza di aggettivi presenti nel prompt, utilizzati per caratterizzare sia i personaggi che la composizione ed il mood dello shot. Nonostante questo, delle quattro immagini generate inizialmente, solamente una mostrava in scena una mano e la posizionava distante dal collo (Figura 7-64). Si è provato, dunque, ad utilizzare dei sinonimi all'interno del prompt, sperando che potessero essere più fruttuosi. Si sono usati i verbi accarezzare, stringere e toccare, che hanno generato immagini in cui la mano era sempre presente, ma mai in contatto con il collo della ragazza. Si è pensato, allora, di rivedere la struttura di base del prompt, mettendo come soggetto del close-up il collo della ragazza, al posto della mano del maniaco. Il cambiamento di focus ha portato l'inquadratura a migliorare la composizione e ad avvicinarsi, ma il posizionamento del braccio si mostrava ancora casuale ed innaturale (Figura 7-64). Tra tutte le immagini, però, una è riuscita a dare una interessante posa alla ragazza che esponeva in modo particolare il collo (Figura 7-66). Si è provato, dunque, ad utilizzare il Generative Fill Tool per inserire un braccio nell'inquadratura, ma si è scoperto che l'aggiunta di particolari umani dà luogo a molte distorsioni ed artefatti (Figura 7-66). Cambiando di nuovo tattica, si è cercato uno shot che proponesse i due personaggi ben visibili e con una posa accettabile (Figura 7-68) per applicarci il Character Posing Tool. Finalmente, si è riusciti a posizionare il braccio vicino al collo con una posa naturale. Si è concluso l'editing applicando l'Image Cropping Tool per avvicinare l'inquadratura, ottenendo l'output in Figura 7-68.



Figura 7-64 Ep.2 Dettaglio collo 1



Figura 7-64 Ep.2 Dettaglio collo 2



Figura 7-66 Ep.2 Dettaglio collo - Applicazione maschera



Figura 7-66 Ep.2 Dettaglio collo - Generative Fill Tool

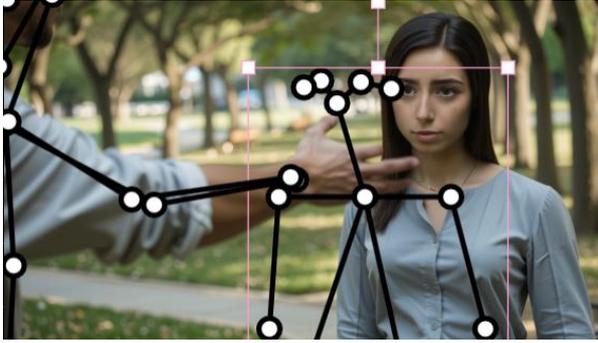


Figura 7-68 Ep.2 Dettaglio collo - Character Posing Tool



Figura 7-68 Ep.2 Dettaglio collo - Output finale

## 8. Analisi dei risultati

Nel capitolo precedente abbiamo analizzato nel dettaglio tutte le applicazioni effettuate dell'intelligenza artificiale all'interno della pre-produzione del cortometraggio "Non ha visto nulla". I risultati sono vari e sfaccettati, meritando dunque un capitolo dedicato. Nel corso dell'analisi, verranno portati avanti ragionamenti che soppesano sia aspetti di carattere creativo che tecnico, tenendo in considerazione che l'obiettivo della tesi non fosse di testare i limiti qualitativi dell'IA, ma di scoprirne le potenzialità di ottimizzazione nei vari ambiti di applicazione.

### 8.1 ChatGPT per valutazione e revisione sceneggiatura: Analisi dei risultati

La valutazione della sceneggiatura ha costituito il primo utilizzo fatto di ChatGPT all'intero della tesi. In una fase ancora precoce del progetto, in cui la sceneggiatura doveva ancora definire la propria forma prima di trasformarsi in un film, il modello si è dimostrato un assistente preziosissimo per un'analisi critica. Questo è avvenuto, innanzitutto, grazie alle elevatissime capacità di ChatGPT di comprensione del testo creativo. Infatti, il modello non si è limitato a seguirne la semantica, ma si è dimostrato capace di interrogarsi con successo sui significanti profondi del testo, ragionando in termini emotivi che sono tipici dell'uomo. Ciò non è affatto banale in un contesto narrativo complesso come quello sottoposto, in cui la trama non è lineare e la rete che tiene assieme gli episodi non è mai esplicitata, ma il modello ha dimostrato di aver colto la visione di insieme e ha strutturato opinioni al riguardo tutt'altro che semplicistiche.

Un aspetto particolarmente apprezzato dell'analisi è l'indole critica che il modello ha assunto. Quel che si cercava, dopotutto, era un confronto vivo riguardo al testo, di cui era essenziale riuscire ad individuare i punti di debolezza prima ancora di quelli di forza. In questi termini, ChatGPT si è distinto dapprima nell'indicare gli episodi più deboli tra quelli proposti, giustificando dettagliatamente e in maniera convincente ogni scelta, ed infine nell'individuare le criticità più specifiche di ogni episodio. Grazie ad un fruttuoso confronto, si è giunti alla decisione degli episodi da tagliare o, ad esempio, a caratterizzare maggiormente i personaggi e potenziare altri aspetti.

Quel che ha reso questa fase creativamente soddisfacente è stato il dialogo che si è instaurato con la macchina. In qualità di sceneggiatore, non si avrebbe gradito un'interazione eccessivamente stagna, da cui ricevere opinioni incontrastabili da dover seguire; invece, grazie ad un rapporto bilaterale di ascolto e di scambio di idee, si è restituito un senso di controllo creativo sull'opera, in cui ChatGPT ha agito da macchina di sollecitazione creativa. Le risposte ricevute non hanno imposto nulla, hanno semplicemente portato nuovi punti di vista nel progetto, ampliato le prospettive dello scrittore e portandolo a ponderarle secondo il proprio gusto.

Un ultimo aspetto di rilevanza è stata la capacità del modello di ragionare in termini cinematografici. Se comprendere un testo può essere oggi banale per l'IA, non è scontato che essa sia in grado di accostarlo ad immagini e che sia capace di caratterizzarle. In questi termini ChatGPT ha stupito per le sue capacità, riuscendo in piena autonomia ad immaginare come trasporre ogni episodio, con scelte originali in termini di fotografia, inquadrature, montaggio e reference. Non c'è stata alcuna opzione di messa in scena, tra le molte proposte, che sia stata giudicata inadatta. Operando come un esperto cinematografico, il modello ha saputo ampliare

la gamma di possibili stili percorribili, ispirando il regista con idee originali. L'autore, a questo punto, ha la possibilità di unire le proprie idee a quelle del modello e di trasformarle fino al raggiungimento del miglior risultato possibile.

Complessivamente, dunque, l'uso dell'IA per la revisione e la valutazione della sceneggiatura si è rivelata estremamente fruttuosa, capace di aiutare l'autore nel processo creativo, ispirandolo e fornendogli un rapido confronto sulle sue creazioni. La rapidità è la parola determinante di questo utilizzo, in quanto ricevere feedback richiede normalmente tempi di attesa elevati dovendosi appoggiare a persone terze. Invece, grazie all'IA il riscontro diventa tempestivo, azzerando gli sprechi.

## 8.2 ChatGPT per la scrittura della sceneggiatura: Analisi dei risultati

Come visto al capitolo 7.2, si è utilizzato il modello di OpenAI per due usi: la riscrittura correttiva di episodi già esistenti e la scrittura di un episodio originale.

La prima applicazione ha rappresentato un parziale insuccesso. L'elevato numero di vincoli imposti può darsi che abbia portato il modello ad operare con parsimonia, ma in definitiva i suoi interventi sono stati poco incisivi in termini di miglioramento. Al contempo, però, una eccessiva libertà portava il modello a stravolgere il testo, non rendendolo più affiancabile all'originale. In uno studio dedicato, sarebbe interessante riuscire a testare più pattern di libertà di riscrittura, fino a trovare il giusto equilibrio capace di sfruttare la creatività dell'IA pur mantenendo alto il controllo creativo da parte dell'autore. Tuttavia, si è rivelato capace di individuare eventuali errori di battitura e di ortografia, un aspetto sempre utile in termini di ottimizzazione dei tempi.

Per quanto riguarda la scrittura autonoma, invece, i risultati sono differenti. Ha richiesto del tempo individuare il migliore approccio al modello per riuscire a sfoderarne il potenziale, ma dopo aver trovato il giusto bilancio di input da inserire ChatGPT ha rispettato le aspettative. Per ottenere una scrittura realmente originale, è necessario caratterizzarla il più possibile, questo richiede un continuo scambio di feedback tra l'autore e il modello. L'interazione è stata intuitiva in questo senso, permettendo un dialogo naturale che direzionava adeguatamente il modello. Con pochi prompt si è raggiunta una buona definizione dello stile narrativo, dopodiché ci si è potuti dedicare al raffinamento della trama.

La sfida principale, una volta individuata la direzione, è stata di mettere dei margini alla creatività del modello. Esso lavora come un'infinta fonte di idee, ma saperlo contenere permette di convogliare gli sforzi verso un risultato definito, limitandone la aleatorietà. Ciò significa che è comunque necessario dedicare del tempo alla costruzione della storia se si desidera mantenere controllo sull'output. Nel nostro caso, ad esempio, era ben conclamato l'obiettivo di realizzare una sceneggiatura in cui autore e IA si mettono allo stesso livello, ossia in cui è essenziale l'intervento dell'uomo per non rendere il processo di scrittura esclusivamente artificiale. Seguendo questa filosofia, dopo diverse versioni generate, si è costruita manualmente una sceneggiatura fatta di ritagli delle proposte dell'IA, modificandole fino ad ottenere il risultato desiderato. In questo modo si è riusciti a far coesistere la visione dell'uomo ed il lavoro della macchina, con un processo vissuto con la stessa spinta artistica che si ha nella scrittura autonoma.

In definitiva, dunque, si ha avuto la conferma che è possibile scrivere utilizzando l'IA senza assopire la creatività umana. L'episodio scritto in questa modalità è un output completamente

originale, che l'autore da solo non sarebbe mai arrivato a creare, come neanche l'IA sarebbe stata in grado di generare, in quanto fusione complessa delle idee di entrambi e output di un processo non ripetibile in alcun modo. In questo senso, l'IA davvero si può imporre come nuova frontiera della creazione umana, ma in un contesto di espansione della medesima e non di sostituzione.

### **8.3 ChatGPT per la scrittura di prompt per immagini: Analisi dei risultati**

La scrittura dei prompt per immagini tramite ChatGPT si è confermata una tattica di successo, andando a semplificare enormemente un'attività che richiede una buona preparazione e discreta esperienza. Dopo l'addestramento, il modello è riuscito a trasformare poche righe di input in prompt completi e dettagliati, utilissimi a generare immagini accurate in pochi tentativi. ChatGPT ha inoltre permesso la correzione interattiva dei prompt, gestendo i numerosi errori di immagine e raggiungendo con buona percentuale il risultato richiesto. Se si volesse approcciare il mondo delle immagini generative, questa tecnica sarebbe sicuramente la più indicata per minimizzare i tempi di scrittura dei prompt, pur richiedendo un occhio vigile per individuare eventuali inesattezze. Ad ogni modo, in un caso come il nostro con un elevato numero di immagini generate, questo uso di ChatGPT si è rivelato fondamentale.

### **8.4 IA per generazione Moodboard: Analisi dei risultati**

Esaurita l'analisi dell'uso dell'IA per la scrittura, si può passare a parlare della generazione di immagini, partendo dalle prime tavole che vengono realizzate nella pre-produzione di un film: le moodboard. Si sono analizzati tre modelli per la generazione delle Moodboard: Katalist, ChatGPT e Microsoft Designer. Gli ultimi due si appoggiano entrambi alla tecnologia di DALL-E per adempiere al loro scopo, mentre il primo nasce - in realtà - per la stesura dello storyboard, ma si è deciso di testarlo ugualmente sfruttando la licenza pagata. A tal proposito, in un'ottica di valutazione anche economica, è importante far notare che Katalist abbia un costo mensile di 99\$ al mese, mentre di ChatGPT e Microsoft Designer si sono testati i piani gratuiti, che permettono rispettivamente la generazione di 3 immagini al giorno e di 15 immagini al mese (con quattro varianti per ciascuna).

Katalist si è dimostrato particolarmente deterministico nella creazione delle moodboard, creando composizioni aderenti alle richieste. Tuttavia, seppur risultassero ideali per rappresentare l'inquadratura, peccavano di mancanza di carattere. Una moodboard deve, per definizione, essere evocativa e trasmettere emozioni, aspetto che tendenzialmente avviene grazie ad un ponderato eccesso di certi elementi visivi. In questa ottica, Katalist si è dimostrato inadatto, fornendo output considerabili troppo standard e spesso incapaci di raccontare il mood della scena.

Al contrario, DALL-E si è distinto in questo senso, generando quadri altamente evocativi e capaci di suscitare tutte le emozioni ricercate nei singoli episodi. È stato interessante notare i cambi di comportamento dei modelli in relazione ai prompt immessi. Laddove l'input era molto preciso, il modello rispondeva secondo le aspettative, ma qualora si lasciasse una maggiore libertà nelle istruzioni, gli output cominciavano a prendere nuove direzioni. Si è rivelato stimolante, in questi casi, scoprire diverse interpretazioni della scena, un aspetto utile in fase di pre-produzione per prendere in considerazione più alternative di messa in scena.

Ad ogni modo, la generazione è proceduta con intoppi minimi, grazie al fatto che nelle moodboard non sia richiesto un alto livello di precisione. In termini di resa generale, tutti i modelli hanno rispettato le richieste con ottima percentuale di successo. Le uniche tavole che hanno creato problemi prevedevano complesse composizioni o molti elementi in scena. Da segnalare, infine, la questione di policy di filtraggio dei contenuti inappropriati. ChatGPT e Microsoft Designer attuano controlli sui loro contenuti al fine di bloccare la generazione di contenuti sessuali o violenti; ciò, però, ha intralciato anche la generazione di due moodboard, in cui figuravano rispettivamente un cadavere ed un uomo ferito. Se con Microsoft Designer si sono trovati degli espedienti per aggirare il filtro, eliminando ad esempio i riferimenti espliciti, con ChatGPT il processo è stato estenuante ed inconcludente. Non si intende additare come sbagliata la politica di protezione, ma ai fini creativi è utile sapere che certe tipologie di scene non sono rappresentabili a causa di tali blocchi, rendendo necessario trovare soluzioni alternative.

Merita un'analisi anche il confronto dei piani utilizzati. Escludendo Katalist, che abbiamo detto essere inadatto allo scopo, le due proposte utilizzate proponevano un'offerta allettante e potenzialmente sufficiente alla generazione di moodboard (tipicamente fatte in quantità limitate). Tuttavia, durante il processo, si è scoperto che raramente si raggiunge il risultato ideale in una sola implementazione. Questo ha messo in luce un vantaggio di Microsoft Designer rispetto a ChatGPT. Con il primo, generando ben quattro versioni di immagine per ogni credito utilizzato, si avevano infatti buone probabilità di essere soddisfatti da almeno uno di essi. Contrariamente è successo con ChatGPT, dove più volte si è dovuto rigenerare l'immagine prima di raggiungere l'obiettivo, finendo spesso per esaurire i crediti giornalieri o, nel peggiore dei casi, a dover continuare il giorno successivo. Questa dinamica di interruzione continua, che ha portato a usare settimane intere per esaurire un solo episodio, ha reso l'utilizzo di ChatGPT particolarmente frustrante e ha richiesto molto tempo. Diversamente, pur avendo un numero limitato di crediti anch'esso, Microsoft Designer ha permesso un approccio al lavoro continuativo e più piacevole.

Arrivando a concludere, la GenAI ha il potere di creare immagini evocative, ideali per trasmettere il mood di una storia. Difficilmente con altri mezzi sarebbe possibile raggiungere risultati paragonabili, se non incaricando un artista di farle. Ciò comporterebbe, però, un costo aggiuntivo non sempre sostenibile, che porta nella maggior parte dei casi a dover rinunciare del tutto alle moodboard o a semplificarle in dei semplici collage. In questo scenario, la GenAI si presenta come un'ottima alternativa per poterle realizzare gratuitamente. Riguardo alla scelta del modello da utilizzare, non vi è differenza qualitativa tra Microsoft Designer e ChatGPT, ma il primo ha un piano di funzionamento che rende il processo più semplice, oltre ad offrire diversi tool di post-produzione integrati nel servizio.

## 8.5 IA per generazione Storyboard: Analisi dei risultati

L'ultima applicazione della GenAI è quella più sperimentale esplorata in questa tesi. Si è fatto uso di Katalist, un modello che nasce esplicitamente per la generazione di storyboard, un compito di difficile gestione per l'IA a causa dell'elevato controllo dei parametri che richiede.

Il modello presenta molte caratteristiche innovative che lo rendono ideale al suo scopo. La prima di esse è la gestione della coerenza tra le varie tavole. Il modello, sotto questo punto di vista, si è dimostrato in grado di rendere i personaggi e gli ambienti ben caratterizzati e riconoscibili in gran parte dello storyboard. Tuttavia, questa funzionalità comporta delle implicazioni negative di diversa natura. Per mantenere i personaggi e le location coerenti, il modello tende a direzionare tutte le generazioni verso situazioni già riscontrate, finendo per

omologare le inquadrature e diminuire il controllo su di essa. Questo rappresenta il maggiore scoglio per il modello, in quanto lo storyboard deve, per definizione, definire dettagliatamente la composizione della scena.

Fortunatamente, il modello presenta nella sua interfaccia diversi tool utili al controllo della tavola. Molti di loro, per i motivi sopraccitati, risultano in molte occasioni ininfluenti, ma altri sono stati determinanti per riuscire a raggiungere l'obiettivo. Generalmente, la gestione della scena avveniva tramite due elementi: il prompt di input ed il tool per il posizionamento dei personaggi. I prompt si sono dimostrati spesso più determinanti delle impostazioni di scena previsti in interfaccia, per cui buona parte del lavoro si è speso nella loro scrittura. Lo strumento per il cambio posa e riposizionamento dei personaggi, invece, si è dimostrato particolarmente utile nelle scene complesse, in cui il modello faticava a posizionare correttamente gli elementi. Katalist presenta poi altri limiti al suo interno, tra cui il fatto di poter rappresentare al massimo due personaggi nella stessa scena, precludendosi di fatto tutte le narrazioni con più di due protagonisti.

Complessivamente, ottenere storyboard accurate è stato un processo estremamente laborioso, che ha richiesto molti interventi per ogni tavola e, nonostante tutto, in più occasioni è stato necessario scendere a compromessi. Il modello ha sicuramente individuato le caratteristiche su cui puntare maggiormente per creare un tool specializzato nel disegno di storyboard, e bisogna dire che sotto diversi punti di vista ha saputo distinguersi dalla concorrenza di mercato. Però, seppur la strada intrapresa sia giusta nelle intenzioni, la tecnologia è ancora troppo acerba per un compito di precisione come questo. Anche il concetto di rapidità è relativo: per raggiungere i migliori risultati possibili sono state necessarie settimane e decine di tentativi per ogni shot, quindi - in termini di economia temporale - il guadagno è molto scarso. La qualità di generazione dell'immagine è sicuramente apprezzabile esteticamente, ma non è il requisito primario di uno storyboard, che per tradizione è spesso molto stilizzato e semplificato. Ma al di sopra di tutto, il limite più vincolante è la difficoltà di controllo dell'inquadratura, che rende sostanzialmente l'intero processo poco utile.

Katalist ha, in definitiva, rappresentato uno studio interessante in qualità di precursore di futuri tool di generazione di Storyboard, dimostrando di aver già mosso diversi passi in quella direzione. Tuttavia, i limiti tecnici ancora impongono grandi barriere, rendendo il modello più o meno adatto a seconda della filosofia con cui viene stilato lo storyboard. Ad esempio, se si desidera una rappresentazione rapida e totalmente generalista della pre-visualizzazione del film, può considerarsi uno strumento ideale, ma diventa inefficace nel senso più comune di intendere uno storyboard, ossia come trasposizione su carta del film altamente dettagliata e realistica (riferendosi alle impostazioni e non alla qualità visiva) al fine di essere il punto di riferimento per tutta la pre-produzione. Rientrando noi in quest'ultimo caso, lo storyboard ottenuto su Katalist è stato ritenuto insufficiente e ha richiesto numerosi interventi manuali prima di renderlo utilizzabile. Nel mercato attuale la GenAI ancora non è in grado di soddisfare le esigenze dei film maker nella stesura dello storyboard, ma possiede tutte le potenzialità per poterlo fare in un futuro prossimo.

## 9. Conclusioni

Concluse le ultime analisi degli usi dell'IA svolti in questa tesi, è possibile terminare lo studio intrapreso con delle considerazioni finali. La tesi nasceva con propositi pratici ed etici, con l'intento di esplorare il modo in cui l'IA sta ridefinendo la nostra società. L'arte, per secoli, è stata considerata un'esclusiva espressione dell'ingegno umano, un simbolo della nostra superiorità intellettuale ed emotiva. Tuttavia, con l'avvento dell'intelligenza artificiale, questa convinzione è stata messa in discussione, mettendo l'uomo di fronte ad una sua stessa tecnologia in grado di creare, ponendo interrogativi profondi sul ruolo dell'artista e sulla natura stessa della creatività. Il potenziale dell'IA è innegabile, ed è impensabile ignorarne l'evoluzione. La paura di un futuro in cui le macchine pensano autonomamente è stata per anni un tema ricorrente nella cultura popolare, eppure l'intelligenza artificiale si è comunque insinuata silenziosamente nella cultura sociale attuale, fino a diventarne una presenza essenziale. Con l'avvento dei nuovi modelli generativi, capaci di produrre opere visive e testuali, il dibattito si è acceso, e l'uomo si è dovuto confrontare con un nuovo tipo di intelligenza. La tesi ha esaminato la discussione e fatto luce sulle numerose zone d'ombra dell'intelligenza artificiale, evidenziandone i limiti e i pericoli. Disinformazione, cura del copyright, ecosostenibilità, riorganizzazione del lavoro: sono molti i temi che meritano attenzione pubblica e consapevolezza.

In questo elaborato, si è scelto di affrontare il tema con una prospettiva critica e consapevole, cercando un equilibrio tra l'innovazione tecnologica e il rispetto dei principi etici. Tuttavia, questa tesi vuole essere innanzitutto tecnica e sperimentale e vuole inserirsi nello specifico settore cinematografico. Le domande di apertura di cui si è cercata risposta sono sostanzialmente due: può l'intelligenza artificiale essere un alleato creativo nella pre-produzione di un'opera filmica? E può esserlo nella prospettiva di democraticizzare il cinema, aiutando le piccole produzioni di stampo indipendente?

Di fronte a questi due obiettivi, la tesi si è mossa verso uno studio del panorama del mercato dell'IA, individuando i modelli più adatti agli obiettivi della ricerca, con un'attenzione particolare ai costi e all'accessibilità. Questa prima analisi si è direzionata cercando di dimostrare che l'IA possa effettivamente rappresentare un'opportunità concreta per le produzioni a basso budget.

Individuate le migliori proposte in rapporto di qualità e prezzo, si è proseguito esplorando nella completezza i tool scelti, nella fattispecie: ChatGPT, DALL-E e Katalist. Questa analisi ci ha permesso di scoprire, a livello tecnico, quali processi si celino dietro alle reti neurali, scoprendone i raggi di azione e le enormi potenzialità. Lo studio ci ha permesso di comprendere un punto di analisi etica fondamentale, esplorato anche nell'introduzione: macchina e uomo possiedono due intelligenze completamente diverse, con limiti e qualità differenti. Apprendere questo significa poter ragionare secondo nuovi modelli mentali, capaci di farci interagire con la macchina con nuove prospettive creative.

Un passo fondamentale per prepararsi alla pre-produzione del cortometraggio è stato lo studio della disciplina della Prompt Engineering. Scoprire le tecniche di sollecitazione dei modelli e le pratiche necessarie all'ottimizzazione dell'output si è rivelato essenziale nella fase applicativa, permettendo di raggiungere i migliori risultati e di poter essere esaustivi nella valutazione. In particolare, l'aver addestrato ChatGPT alla scrittura dei prompt ha permesso di ridurre enormemente i tempi per la generazione, soprattutto nell'ottica dell'elevato numero di tavole prodotte.

## Conclusioni

Dunque, muniti di tutte le nozioni necessarie, si è svolta quella che è stata la parte sperimentale ed il cuore di questa tesi. La prima delle applicazioni ha coinvolto ChatGPT come strumento di revisione della sceneggiatura, fase che ha prodotto risultati molto interessanti e fatto scaturire ragionamenti profondi da parte della macchina. Poter sviscerare il testo attraverso un dialogo costruttivo con una “voce esterna” si è dimostrato un processo estremamente utile, capace di migliorare la sceneggiatura nella forma e nel contenuto. In questo processo, si è percepito un pieno mantenimento della paternità dell’opera, filtrando attivamente i consigli del modello e rielaborandoli secondo il proprio gusto. Per cui ci si esprime con parere estremamente positivo nei confronti della scrittura AI-assisted, una dinamica tanto semplice quanto fruttuosa, che può aiutare molti scrittori nel difficile processo di revisione critica dei propri lavori.

Un esperimento più complesso è stato, invece, la scrittura AI-generated effettuata per l’ultimo episodio, in cui l’elaborazione partiva dalla macchina e veniva solamente rielaborata dallo scrittore. Dopo un primo approccio inconcludente – in cui l’aver lasciato piena autonomia al modello ha generato output funzionali, ma piuttosto convenzionali e privi di originalità – si è riusciti ad instaurare una dinamica iterativa estremamente stimolante. Una lunga sessione di dialogo ha visto convivere e fondersi le idee dall’IA e quelle dell’uomo, attraverso un continuo sistema bidirezionale di feedback ed opinioni. Insomma, è avvenuto quello che si può definire una ricca sessione di brainstorming con la macchina, in cui il testo è diventato materia viva e in continua mutazione. L’unica difficoltà, in questo senso, è stata quella di riuscire a mantenere un ordine nell’esuberanza creativa, rendendo necessario l’intervento umano per operare un’azione di taglia e cuci tra le idee, fino a costruire l’episodio nella sua forma finale. La qualità complessiva della sceneggiatura risultante è sorprendente, aderente al plot iniziale, ma arricchita dalla creatività stimolata dal modello. Non è un caso che si sia testato questo uso proprio per l’ultimo episodio, ci si trovava infatti in un momento di blocco, in cui si sapeva quale messaggio trasmettere ma non si trovava un modo originale per trasformarlo in sceneggiatura. Attraverso l’IA è stato possibile ridurre settimane o mesi di ricerca ispirativa in poche giornate di confronto digitale. In questo aspetto l’IA si dimostra estremamente utile nel ridurre drasticamente i tempi della scrittura, intesa come ricerca dell’ispirazione. Resta però da definire chi sia l’autore finale del testo. Riprendendo il concetto ben esplorato nel capitolo 2, questa fase ha incarnato alla perfezione il concetto di co-autorialità che ci si auspicava. L’IA ha portato molte idee originali nella sceneggiatura, elaborandole in un modo che non può sminuirle al semplice ruolo di “strumento”. Al contempo, l’intervento umano è stato indispensabile nel guidare, filtrare ed infine assemblare la narrazione nella sua versione finale. La macchina da sola avrebbe prodotto risultati banali, l’uomo da solo sarebbe rimasto arenato nel suo blocco creativo, ma unendosi hanno potuto creare uno degli episodi migliori del cortometraggio. Volendo concludere un ragionamento sulla scrittura collaborativa, è rassicurante pensare che l’uomo sarà sempre in grado di produrre arte in piena autonomia, ma l’esistenza dell’IA giocherà certamente un ruolo essenziale nell’aiutarlo di fronte alle difficoltà che l’atto creativo porta da sempre con sé. Da un’epoca il cui l’immaginazione umana viene aumentata da quella artificiale non resta che aspettarsi nuove ed esaltanti frontiere artistiche.

Passando al campo della generazione di immagini, DALL-E si è imposto come strumento ideale per la generazione delle moodboard. Esso ha saputo trasformare brevi descrizioni in immagini fortemente evocative, capaci di incarnare la visione del regista. Il suo utilizzo ha sicuramente presentato degli ostacoli, ma nel valutare il rapporto tra qualità dell’output finale e costi economici e temporali non vi è alcun dubbio sul guadagno raggiunto. Si è detto, infatti, che è comune dover rinunciare alla stesura delle moodboard per tagliare i costi, ma la prospettiva di poterle generare in pochi secondi e gratuitamente mette un nuovo strumento nelle mani delle piccole produzioni.

## Conclusioni

Infine, chiudiamo la gamma di applicazioni analizzando lo storyboard prodotto da Katalist. L'idea di poter generare tavole accurate e ad alta qualità con l'intelligenza artificiale è stata una prospettiva particolarmente stimolante del progetto e c'era molta curiosità nei confronti dei risultati. Purtroppo, la tecnologia si è dimostrata ancora insufficiente per svolgere il compito, a causa degli eccessivi vincoli che esso si porta appresso. Nonostante la forma imperfetta, si sono comunque individuati degli interessanti punti di forza, che potrebbero con ogni probabilità migliorarsi in futuro fino a rendere possibile la realizzazione di uno storyboard accurato.

Tornando, dunque, alle domande iniziali del capitolo, possiamo innanzitutto affermare che l'IA ha saputo dimostrarsi un utile alleato nella fase di pre-produzione del progetto. In un cortometraggio indipendente e low-budget, la macchina ha saputo ottimizzare i tempi lavorativi, creando un ritorno economico a fronte di investimenti minimi. Ancora più interessante è la valutazione creativa dell'IA. Possiamo affermare di aver individuato e testato un paradigma di successo di collaborazione uomo-macchina. Si è riusciti ad equilibrare i due lati creativi efficacemente, facendo percepire il pieno controllo autoriale sulla sceneggiatura, ma lasciandosi influenzare positivamente dai molti input provenienti dall'IA. Il risultato finale è un testo migliore dell'originale, non solo nella forma ma anche nel contenuto. L'IA ha saputo ispirare e ampliare le prospettive dello scrittore senza mai scavalcarlo. In questo processo creativo ponderato, non viene difficile pensare a un futuro in cui i modelli intelligenti diventeranno lo stato dell'arte del panorama artistico, un mezzo attraverso cui scavare nel potenziale artistico umano ed un organo artificiale capace di portare l'uomo a scoprirsi in maniera profonda.

È giusto evidenziare, in conclusione, quelli che sono i limiti del lavoro svolto. Infatti la natura auto-finanziata del progetto ci ha impedito di sviluppare la tesi quanto avremmo voluto. Sarebbe stato interessante, prima di tutto, ampliare il bacino di modelli testati in modo da costruire una più ampia analisi comparativa dell'offerta del mercato. In questo scenario si sarebbero esaminate sia proposte gratuite che a pagamento, valutandone i diversi vantaggi. Si è accennato anche al fatto che una scena del cortometraggio sia stata generata usando Runway (questa applicazione è curata dalla tesi della collega Miriana Mallia citata in introduzione), ma un budget più alto ci avrebbe sia permesso di perfezionare ulteriormente questa scena sia di ampliare la presenza di video generativi nel cortometraggio. Con lo stesso spirito si avrebbe voluto testare la generazione musicale, un crescente ramo della GenAI che vede numerosi software a pagamento aprire nuove frontiere per il cinema, in grado di realizzare colonne sonore a poco costo. Insomma, in una prospettiva futura e con maggiori risorse il progetto potrebbe facilmente espandersi diventando più esaustivo. Inoltre, rimane un dato di fatto che l'IA sia in continua evoluzione e che i risultati di oggi potrebbero, con tutta probabilità, diventare obsoleti tra pochi mesi. Tenere sotto osservazione l'avanzamento della tecnologia potrà senza dubbio risolvere ogni questione che non ha trovato soluzione in questo elaborato.

Nonostante queste considerazioni, la tesi si inserisce in un ampio panorama di studi dedicati all'IA, contribuendo a documentarne la rapida ascesa. Le tecnologie hanno il potere di plasmare la società in cui viviamo e accompagnarne lo sviluppo è fondamentale per controllarne l'impatto. L'intelligenza artificiale si preannuncia come una rivoluzione epocale, capace di disegnare per l'uomo nuove traiettorie evolutive; per questo è più che mai necessario alimentare un dibattito pubblico e consapevole attorno ad essa. Non basta infatti esplorarne le frontiere per scoprirne il potenziale: è imprescindibile costruire un discorso etico sul tema. Solo attraverso una conoscenza vigile possiamo, usando le parole di Ed Finn, evitare di "vivere esclusivamente nello spazio calcolabile" [1]. Abbracciare l'entusiasmo per l'innovazione senza perdere il senso critico è la chiave per costruire la *cultura logaritmica* proposta dallo scrittore, perché le scelte che compiamo oggi definiscono già il futuro che ci attende.

## 10. Bibliografia e Sitografia

- [1] E. Finn, Cosa vogliono gli algoritmi, Torino: Giulio Einaudi editore s.p.a., 2017.
- [2] L. Manovich, AI Aesthetics, Mosca: Strelka Press, 2019.
- [3] Platone, Fedro, 274b-275e..
- [4] «Mente estesa - Treccani,» [Online]. Available: [https://www.treccani.it/enciclopedia/mente-estesa\\_\(Lessico-del-XXI-Secolo\)/](https://www.treccani.it/enciclopedia/mente-estesa_(Lessico-del-XXI-Secolo)/).
- [5] «Google Clips,» [Online]. Available: <https://support.google.com/googleclips/answer/7545440?hl=en>.
- [6] «Soft Cinema - Lev Manovich, Andreas Kratky,» [Online]. Available: <https://manovich.net/index.php/projects/soft-cinema-zkm>.
- [7] «Harold Cohen -Aaron,» [Online]. Available: <https://whitney.org/exhibitions/harold-cohen-aaron>.
- [8] A. H.-J. U. H. Z. K. & A. T. Steffen Herbold, «A large-scale comparison of human-written versus ChatGPT-generated essays,» 2023. [Online]. Available: <https://www.nature.com/articles/s41598-023-45644-9>.
- [9] «Coca-Cola's Deal with DALL-E and ChatGPT,» 2023. [Online]. Available: <https://www.linkedin.com/pulse/coca-colas-deal-dall-e-chatgpt-aquarious-technology/>.
- [10] «Un primo bilancio del Foglio AI. Scritto dal Foglio AI. Con sgridata,» [Online]. Available: <https://www.ilfoglio.it/il-foglio-ai/2025/03/22/news/un-primo-bilancio-del-foglio-ai-scritto-dal-foglio-ai-con-sgridata-7546387/>.
- [11] «The Authors Guild and 13 Other Authors File Class-Action Suit Against OpenAI,» [Online]. Available: <chrome-extension://efaidnbmnnnibpajpcgclefindmkaj/https://authorsguild.org/app/uploads/2023/12/Authors-Guild-OpenAI-Microsoft-Class-Action-Complaint-Dec-2023.pdf>.
- [12] A. S. S. B. R. M. a. M. C. Simon Colton, «Generative Search Engines: Initial Experiments,» [Online]. Available: [chrome-extension://efaidnbmnnnibpajpcgclefindmkaj/https://computationalcreativity.net/icc/c21/wp-content/uploads/2021/09/ICCC\\_2021\\_paper\\_50.pdf](chrome-extension://efaidnbmnnnibpajpcgclefindmkaj/https://computationalcreativity.net/icc/c21/wp-content/uploads/2021/09/ICCC_2021_paper_50.pdf).
- [13] R. Gao, «AIGC Technology: Reshaping the Future of the Animation Industry,» 2023. [Online]. Available: [https://www.researchgate.net/publication/378699518\\_AIGC\\_Technology\\_Reshaping\\_the\\_Future\\_of\\_the\\_Animation\\_Industry](https://www.researchgate.net/publication/378699518_AIGC_Technology_Reshaping_the_Future_of_the_Animation_Industry).
- [14] D. Amin, «How Netflix Used Data to Create House of Cards: A Revolutionary Approach to Content Creation,» [Online]. Available: <https://medium.com/@daniel.a/how-netflix-used-data-to-create-house-of-cards-a-revolutionary-approach-to-content-creation-b9a114630ddc>.

- [15] A. Barale, «Un cambiamento quasi impercettibile: arte e intelligenza artificiale,» 2023. [Online]. Available: <https://ojs.unito.it/index.php/vde/article/view/10836>.
- [16] N. M. C. Joo-Wha Hong, «Artificial Intelligence, Artists, and Art: Attitudes Toward Artwork Produced by Humans vs. Artificial Intelligence,» 2019. [Online]. Available: <https://dl.acm.org/doi/10.1145/3326337>.
- [17] M. Spring, «Trump supporters target black voters with faked AI images,» 2024. [Online]. Available: <https://www.bbc.com/news/world-us-canada-68440150>.
- [18] F. Prisco, «Trump e l'idea della «Riviera di Gaza»: il video delle polemiche con grattacieli e danzatrici,» 2025. [Online]. Available: <https://www.ilsole24ore.com/art/trump-posta-video-riviera-gaza-grattacieli-e-danzatrici-AGB6u89C>.
- [19] «Gli sceneggiatori di Hollywood sono in sciopero: i motivi e le conseguenze (anche sulle serie tv),» [Online]. Available: <https://www.wired.it/article/sciopero-sceneggiatori-hollywood-motivi-conseguenze-serie-tv/>.
- [20] «Perché agli sceneggiatori fa così paura l'intelligenza artificiale,» [Online]. Available: <https://www.wired.it/article/intelligenza-artificiale-sciopero-sceneggiatori-hollywood/>.
- [21] «Secondo Paul Schrader, le idee per i film dovrebbero essere sviluppate da ChatGPT,» [Online]. Available: <https://www.rollingstone.it/cinema-tv/news-cinema-tv/secondo-paul-schrader-le-idee-per-i-film-dovrebbero-essere-sviluppate-da-chatgpt/961044/>.
- [22] J. Nelson, «'Black Mirror' Creator Says ChatGPT Episode Script Was 'Shit',» [Online]. Available: <https://decrypt.co/143647/black-mirror-creator-says-chatgpt-episode-script-was-shit>.
- [23] «La produzione del film The Brutalist ha usato l'AI per dialoghi e immagini,» [Online]. Available: <https://www.wired.it/article/the-brutalist-intelligenza-artificiale-respoecher/>.
- [24] «Perché The Brutalist insegna che l'intelligenza artificiale può essere un'alleata per il cinema,» [Online]. Available: <https://www.fanpage.it/spettacolo/film/perche-the-brutalist-insegna-che-lintelligenza-artificiale-puo-essere-unalleata-per-il-cinema/>.
- [25] «AlphaGo ha conquistato l'ultima scacchiera; la fantascienza ci aveva avvertito.,» [Online]. Available: <https://www.indiscreto.org/umanita-sotto-attacco/>.
- [26] «Stephen Hawking: Artificial intelligence could be the greatest disaster in human history,» [Online]. Available: <https://www.independent.co.uk/news/people/stephen-hawking-artificial-intelligence-diaster-human-history-leverhulme-centre-cambridge-a7371106.html>.
- [27] «Microsoft Readies AI Chip as Machine Learning Costs Surge,» [Online]. Available: Stephen Hawking: Artificial intelligence could be the greatest disaster in human history.
- [28] «The hidden enviromental cost of AI,» [Online]. Available: <https://www.washingtonpost.com/technology/2024/09/18/energy-ai-use-electricity-water-data-centers/>.

- [29] «GEN AI: TOO MUCH SPEND, TOO LITTLE BENEFIT?,» [Online]. Available: chrome-extension://efaidnbmnnnibpcajpcglclefindmkaj/https://www.goldmansachs.com/images/migrated/insights/pages/gs-research/gen-ai--too-much-spend%2C-too-little-benefit-/TOM\_AI%202.0\_ForRedaction.pdf?ref=404media.co?ref=404media.co.
- [30] «Causa del New York Times contro OpenAI,» [Online]. Available: chrome-extension://efaidnbmnnnibpcajpcglclefindmkaj/https://nytco-assets.nytimes.com/2023/12/NYT\_Complaint\_Dec2023.pdf.
- [31] «Scarlett Johansson says a ChatGPT voice is ‘eerily similar’ to hers and OpenAI is halting its use,» [Online]. Available: https://apnews.com/article/openai-chatgpt-scarlett-johansson-voice-her-532c849ccae3ca9e9325dacf88e0436.
- [32] «Kurosawa vs Leone: l'accusa di plagio,» [Online]. Available: https://festival.ilcinemaritrovato.it/en/kurosawa-vs-leone/.
- [33] «Is this what we want? L'album silenzioso di oltre 1000 artisti britannici. In prima linea Kate Bush, Elton John e Annie Lennox,» [Online]. Available: https://www.ilmessaggero.it/spettacoli/musica/is\_what\_we\_want\_artisti\_britannici\_album\_musica\_kate\_bush\_elton\_john\_annie\_lennox-8681892.html.
- [34] «Chi ha creato la foto del Papa col piumino?,» [Online]. Available: https://www.repubblica.it/tecnologia/2023/03/29/news/chi\_ha\_fatto\_foto\_papa\_piumino\_intelligenza\_artificiale\_midjourney-394115698/.
- [35] «Deep fake: l'ultima frontiera del porno immaginato (e mai girato),» [Online]. Available: la stessa quantità necessaria a far funzionare più di una dozzina di lampadine a led per circa un'ora.
- [36] «Fake and Explicit Images of Taylor Swift,» [Online]. Available: https://www.nytimes.com/2024/02/05/business/media/taylor-swift-ai-fake-images.html.
- [37] «EU AI Act,» [Online]. Available: https://digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai#:~:text=La%20legge%20sull'IA%20(regolamento,un'IA%20affidabile%20in%20Europa..
- [38] F. o. I. institute. [Online]. Available: https://futureoflife.org/focus-area/artificial-intelligence/.
- [39] A. Barale, «Un cambiamento quasi impercettibile: arte e intelligenza artificiale,» [Online]. Available: https://ojs.unito.it/index.php/vde/article/view/10836.
- [40] F. D'Isa, «La rivoluzione degli algoritmi nel mondo dell'arte,» [Online]. Available: https://www.iltascabile.com/scienze/arte-algoritmi/.
- [41] «L'intelligenza artificiale divide gli artisti,» [Online]. Available: https://www.internazionale.it/essenziale/notizie/vanni-santoni/2023/02/17/intelligenza-artificiale-arte.
- [42] Baudelaire, *Revue française*, 10 Giugno 1859.

- [43] Z. K. T. H. Svitlana Fiialka, «The use of ChatGPT in creative writing assistance,» [Online]. Available: [chrome-extension://efaidnbnmnibpcajpcglclefindmkaj/https://xlinguae.eu/files/XLinguae1\\_2024\\_1.pdf](chrome-extension://efaidnbnmnibpcajpcglclefindmkaj/https://xlinguae.eu/files/XLinguae1_2024_1.pdf).
- [44] G. Magini, «Creatività non umana,» [Online]. Available: <https://www.singola.net/tecnologia/sullo-spirito-degli-algoritmi-dall-intelligenza-artificiale-all-immaginazione-aumentata>.
- [45] I. Calvino, *Se una notte d'inverno un viaggiatore*, Cles: Oscar Mondadori, 2020.
- [46] I. Calvino, «Se una notte d'inverno un narratore,» [Online]. Available: <https://marioxmancini.medium.com/se-una-notte-dinverno-un-narratore-da7b5dc03937>.
- [47] «Plenitudine digitale,» [Online]. Available: <https://blum.vision/plenitudine-digitale-bolter/>.
- [48] P. Ottorina, «DeepSeek, una sconosciuta azienda cinese fa tremare la Silicon Valley: in Borsa crollano le azioni AI,» [Online]. Available: [https://www.corriere.it/tecnologia/25\\_gennaio\\_27/perche-deepseek-una-sconosciuta-azienda-cinese-fa-tremare-la-silicon-valley-in-borsa-crollano-le-azioni-ai-f67fef72-47fb-4a35-86f8-cdb22d1fcx1k.shtml](https://www.corriere.it/tecnologia/25_gennaio_27/perche-deepseek-una-sconosciuta-azienda-cinese-fa-tremare-la-silicon-valley-in-borsa-crollano-le-azioni-ai-f67fef72-47fb-4a35-86f8-cdb22d1fcx1k.shtml).
- [49] Reuters, «OpenAI set to make first ad foray with Super Bowl spot, Wall Street Journal reports,» Febbraio 2025. [Online]. Available: [https://www.reuters.com/technology/openai-set-make-first-ad-foray-with-super-bowl-spot-wall-street-journal-reports-2025-02-06/?utm\\_source=chatgpt.com](https://www.reuters.com/technology/openai-set-make-first-ad-foray-with-super-bowl-spot-wall-street-journal-reports-2025-02-06/?utm_source=chatgpt.com).
- [50] A. I. Joseph Ingio, «A Comparative Analysis of Generative Artificial Intelligence Tools for Natural Language Processing,» 2024. [Online]. Available: <https://doi.org/10.62411/jcta.9447>.
- [51] «AI-based tools for researchers,» 2022. [Online]. Available: [https://paperpal.com/blog/wp-content/uploads/2022/11/AI-based-editing-tools-for-researchers\\_White-Paper\\_2022.pdf](https://paperpal.com/blog/wp-content/uploads/2022/11/AI-based-editing-tools-for-researchers_White-Paper_2022.pdf).
- [52] K. M. Happle, «Cre-AI-tivity The Intersection of Human Ingenuity and AI Collaboration in Scriptwriting,» 2024. [Online]. Available: <https://studenttheses.uu.nl/handle/20.500.12932/46352?show=full>.
- [53] S. F. –. Z. K. –. T. Honcharuk, «The use of ChatGPT in creative writing assistance,» 2024. [Online]. Available: [https://www.xlinguae.eu/2024\\_17\\_1\\_1.html](https://www.xlinguae.eu/2024_17_1_1.html).
- [54] «blog.openGPT.com,» 2024. [Online]. Available: <https://blog.opengpt.com/2024/02/04/comparative-analysis-of-ai-image-generation-models-dall%C2%B7e-3-google-imagen2-stable-diffusion-and-midjourney/>.
- [55] «towardsai.net,» [Online]. Available: <https://towardsai.net/p/artificial-intelligence/top-11-ai-powered-image-generators-in-2024>.
- [56] «Canva AI,» [Online]. Available: <https://www.canva.com/ai-image-generator/>.
- [57] «Adobe Firefly,» [Online]. Available: <https://www.adobe.com/products/firefly.html>.

- [58] «NightCafe,» [Online]. Available: <https://creator.nightcafe.studio/>.
- [59] «AI Shutterstock,» [Online]. Available: <https://www.shutterstock.com/it/ai-image-generator>.
- [60] LeonardoAI. [Online]. Available: <https://leonardo.ai/>.
- [61] «Craiyon,» [Online]. Available: <https://www.craiyon.com/>.
- [62] «Katalist new tool for storytelling,» [Online]. Available: <https://eliteai.tools/tool/katalist>.
- [63] «IBM Generative-AI,» [Online]. Available: <https://www.ibm.com/it-it/topics/generative-ai>.
- [64] «IBM - What is Machine Learning?,» [Online]. Available: <https://www.ibm.com/think/topics/machine-learning>.
- [65] «UC Berkeley - What is ML?,» [Online]. Available: <https://ischoolonline.berkeley.edu/blog/what-is-machine-learning/>.
- [66] «IBM - What is deep learning?,» [Online]. Available: <https://www.ibm.com/think/topics/deep-learning>.
- [67] «IBM - What is a neural network,» [Online]. Available: <https://www.ibm.com/think/topics/neural-networks>.
- [68] «AI vs. machine learning vs. deep learning vs. neural networks: What's the difference?,» [Online]. Available: <https://www.ibm.com/think/topics/ai-vs-machine-learning-vs-deep-learning-vs-neural-networks>.
- [69] C. Negri, «ChatGPT, cos'è e come funziona: limiti e opportunità,» 2024. [Online]. Available: [https://blog.osservatori.net/it\\_it/chatgpt-come-funziona-cosa-puo-fare-limiti-opportunita](https://blog.osservatori.net/it_it/chatgpt-come-funziona-cosa-puo-fare-limiti-opportunita).
- [70] «Che cos'è l'AI forte?,» [Online]. Available: <https://www.ibm.com/it-it/topics/strong-ai>.
- [71] C. S. Ivan Belcic, «Che cos'è un GPT (trasformatore generativo pre-addestrato)?,» 2024. [Online]. Available: <https://www.ibm.com/it-it/think/topics/gpt>.
- [72] «OpenAI Models,» [Online]. Available: <https://platform.openai.com/docs/models/models> .
- [73] «ChatGPT, cos'è e come funziona: limiti e opportunità,» [Online]. Available: [https://blog.osservatori.net/it\\_it/chatgpt-come-funziona-cosa-puo-fare-limiti-opportunita](https://blog.osservatori.net/it_it/chatgpt-come-funziona-cosa-puo-fare-limiti-opportunita).
- [74] «Cos'è un modello trasformatore?,» IBM, 2024. [Online]. Available: <https://www.ibm.com/it-it/topics/transformer-model>.
- [75] «Cos'è una rete neurale ricorrente (RNN)?,» 2024. [Online]. Available: <https://www.ibm.com/it-it/think/topics/recurrent-neural-networks>.
- [76] «Cosa sono le reti neurali convoluzionali?,» [Online]. Available: <https://www.ibm.com/it-it/think/topics/convolutional-neural-networks>.

- [77] «Cos'è l'incorporamento?,» 2023. [Online]. Available: <https://www.ibm.com/it-it/topics/embedding>.
- [78] J. A. S. A. S. A. L. A. I. A. F. L. A. D. A. J. A. S. A. S. A. R. A. I. B. OpenAI, «GTP4 a technical report,» 2024. [Online]. Available: <https://arxiv.org/abs/2303.08774>.
- [79] Z. u. Abideen, «How OpenAI's DALL-E works?,» 2023. [Online]. Available: <https://medium.com/data-science-in-your-pocket/how-does-dall-e-the-text-to-image-generator-work-c2d9f4a0f26c>.
- [80] «OpenAI Image-GPT,» [Online]. Available: <https://openai.com/index/image-gpt/>.
- [81] «OpenAI DALL-E,» [Online]. Available: <https://openai.com/index/dall-e/>.
- [82] «OpenAI DALL-E 2,» [Online]. Available: <https://openai.com/index/dall-e-2/>.
- [83] «OpenAI's DALL-E 3 Explained: Generate Images with ChatGPT,» 2023. [Online]. Available: <https://encord.com/blog/openai-dall-e-3-what-we-know-so-far/>.
- [84] «OpenAI DALL-E 3,» [Online]. Available: <https://openai.com/index/dall-e-3/>.
- [85] «OpenAI CLIP,» [Online]. Available: <https://openai.com/index/clip/>.
- [86] «DALL-E Architecture,» [Online]. Available: <https://www.tutorialspoint.com/dall-e/dall-e-architecture.htm>.
- [87] A. I. Aramendia, «Attention Is All You Need : A Complete Guide to Transformers,» 2024. [Online]. Available: <https://medium.com/@alejandro.itoaramendia/attention-is-all-you-need-a-complete-guide-to-transformers-8670a3f09d02#424d>.
- [88] M. Gupta, «How does DALL-E, the text-to-image generator work?,» 2023. [Online]. Available: <https://medium.com/%40zaiinn440/how-openais-dall-e-works-da24ac6c12fa>.
- [89] «Cos'è un autoencoder variazionale?,» [Online]. Available: [https://www.ibm.com/it-it/think/topics/variational-autoencoder#:~:text=Gli%20autoencoder%20variazionali%20\(VAE\)%20sono,altri%20autoencoder%2C%20come%20il%20denoising..](https://www.ibm.com/it-it/think/topics/variational-autoencoder#:~:text=Gli%20autoencoder%20variazionali%20(VAE)%20sono,altri%20autoencoder%2C%20come%20il%20denoising..)
- [90] «DALL-E OpenAI,» [Online]. Available: <https://openai.com/index/dall-e/>.
- [91] «YFCC100M Dataset,» [Online]. Available: <https://paperswithcode.com/dataset/yfcc100m>.
- [92] «Inception score(IS) and Fréchet inception distance(FID) explained,» Niharika Ahuja, 2023. [Online]. Available: <https://ahujaniharika95.medium.com/inception-score-is-and-fr%C3%A9chet-inception-distance-fid-explained-2bc28a4faea7>.
- [93] «Microsoft Confirms Its \$10 Billion Investment Into ChatGPT,» [Online]. Available: <https://www.forbes.com/sites/qai/2023/01/27/microsoft-confirms-its-10-billion-investment-into-chatgpt-changing-how-microsoft-competes-with-google-apple-and-other-tech-giants/>.
- [94] «Microsoft uses DALL-E 3,» [Online]. Available: <https://create.microsoft.com/en-us/learn/articles/how-to-image-prompts-dall-e-ai>.

- [95] «Microsoft - Create stunning designs lightning fast with Microsoft Designer AI,» [Online]. Available: <https://techcommunity.microsoft.com/blog/microsoft365insiderblog/create-stunning-designs-lightning-fast-with-microsoft-designer-ai/4220093>.
- [96] «CNBC - Microsoft launches Designer, its answer to highly valued startup Canva,» [Online]. Available: <https://www.cnn.com/2022/10/12/microsoft-launches-designer-its-answer-to-highly-valued-startup-canva.html>.
- [97] «Katalist AI,» [Online]. Available: <https://www.katalist.ai/>.
- [98] V. a. L. B. C. Liu, «Design Guidelines for Prompt Engineering Text-To-Image Generative Models,» 2022. [Online]. Available: <https://doi.org/10.1145/3491102.3501825>.
- [99] G. Branwen, «Gpt-3 creative fiction,» 2020. [Online]. Available: <https://www.gwern.net/GPT-3>.
- [100] Sander Schulhoff, «Learn Prompting,» [Online]. Available: <https://learnprompting.org/>.
- [101] OpenAI, «OpenAI Documentation,» [Online]. Available: <https://platform.openai.com/docs/guides>.
- [102] «Runway Documentation,» [Online]. Available: <https://docs.dev.runwayml.com/>.
- [103] J. Oppenlaender, «A taxonomy of prompt modifiers for text-to-image generation,» 2023. [Online]. Available: <https://doi.org/10.1080/0144929X.2023.2286532>.
- [104] «The DALL-E 2 Prompt Book,» [Online]. Available: <https://dallery.gallery/the-dalle-2-prompt-book/>.
- [105] E. Smith, «A Traveler's Guide to the Latent Space,» [Online]. Available: [https://www.reddit.com/r/PromptDesign/comments/vl9taw/a\\_travelers\\_guide\\_to\\_the\\_latent\\_space/](https://www.reddit.com/r/PromptDesign/comments/vl9taw/a_travelers_guide_to_the_latent_space/).
- [106] «ArtHub AI,» [Online]. Available: <https://arthur.ai/>.
- [107] «Lexixa.art,» [Online]. Available: <https://lexica.art/>.
- [108] «Openart.ai,» [Online]. Available: <https://openart.ai/>.
- [109] X. W. D. S. M. B. B. I. F. X. E. C. Q. L. D. Z. Jason Wei, «Chain-of-Thought Prompting Elicits Reasoning in Large Language Models,» 2022. [Online]. Available: <https://doi.org/10.48550/arXiv.2201.11903>.
- [110] V. Kuka, «Learn Prompting,» [Online]. Available: [https://learnprompting.org/docs/intermediate/chain\\_of\\_thought](https://learnprompting.org/docs/intermediate/chain_of_thought).
- [111] J. W. D. S. Q. L. E. C. S. N. A. C. D. Z. Xuezhi Wang, «Self-Consistency Improves Chain of Thought Reasoning in Language Models,» 2022. [Online]. Available: <https://doi.org/10.48550/arXiv.2203.11171>.

- [112] A. L. X. L. S. W. P. W. R. L. B. Y. C. H. H. Jiacheng Liu, «Generated Knowledge Prompting for Commonsense Reasoning,» 2021. [Online]. Available: <https://doi.org/10.48550/arXiv.2110.08387>.
- [113] J. W. Y. Z. K. Z. W. H. J. L. F. L. Q. Y. X. X. Cheng Li, «Large Language Models Understand and Can be Enhanced by Emotional Stimuli,» 2023. [Online]. Available: <https://doi.org/10.48550/arXiv.2307.11760>.
- [114] C. T. T. S. C. X. H. X. G. L. J.-g. L. S. M. Xiaohan Xu, «Re-Reading Improves Reasoning in Large Language Models,» 2023. [Online]. Available: <https://doi.org/10.48550/arXiv.2309.06275>.
- [115] W. Z. Z. C. Q. G. Yihe Deng, «Rephrase and Respond: Let Large Language Models Ask Better Questions for Themselves,» 2023. [Online]. Available: <https://doi.org/10.48550/arXiv.2311.04205>.
- [116] S. S. Jason Weston, «System 2 Attention (is something you might need too),» 2023. [Online]. Available: <https://doi.org/10.48550/arXiv.2311.11829>.
- [117] S. M. X. C. H.-T. C. E. H. C. Q. V. L. D. Z. Huaixiu Steven Zheng, «Take a Step Back: Evoking Reasoning via Abstraction in Large Language Models,» 2023. [Online]. Available: <https://doi.org/10.48550/arXiv.2310.06117>.
- [118] R. A. U. A. J. R. K. X. P. Y. S. P. C. S. X. W. D. Z. Xinyun Chen, «Universal Self-Consistency for Large Language Model Generation,» 2023. [Online]. Available: <https://doi.org/10.48550/arXiv.2311.17311>.
- [119] M. K. J. X. R. R. X. L. A. C. J. W. Shehzaad Dhuliawala, «Chain-of-Verification Reduces Hallucination in Large Language Models,» 2023. [Online]. Available: <https://doi.org/10.48550/arXiv.2309.11495>.
- [120] N. T. P. G. S. H. L. G. S. W. U. A. N. D. S. P. Y. Y. S. G. B. P. M. K. H. S. W. A. Y. P. C. Aman Madaan, «Self-Refine: Iterative Refinement with Self-Feedback,» 2023. [Online]. Available: <https://doi.org/10.48550/arXiv.2303.17651>.
- [121] W. X. Y. L. Z. H. Y. L. R. K.-W. L. E.-P. L. Lei Wang, «Plan-and-Solve Prompting: Improving Zero-Shot Chain-of-Thought Reasoning by Large Language Models,» 2023. [Online]. Available: <https://doi.org/10.48550/arXiv.2305.04091>.
- [122] D. Y. Junrong Song, «Exploring the Intersection of AI Art and Film: A Case,» 2023. [Online]. Available: [10.1109/ICMEW59549.2023.00066](https://arxiv.org/abs/2309.00066).
- [123] «Che cos'è il machine learning (ML)?,» IBM, 2024. [Online]. Available: <https://www.ibm.com/it-it/topics/machine-learning>.
- [124] J. Holdsworth e M. Scapicchio, «Cos'è il deep learning?,» IBM, 2024. [Online]. Available: <https://www.ibm.com/it-it/topics/deep-learning>.
- [125] D. Bergmann e C. Stryker, «Cosa sono i modelli di diffusione?,» IBM, 2024. [Online]. Available: <https://www.ibm.com/it-it/think/topics/diffusion-models>.
- [126] D. Bergmann e C. Stryker, «Cos'è un autoencoder variazionale?,» IBM, 2024. [Online]. Available: <https://www.ibm.com/it-it/think/topics/variational-autoencoder>.

## Bibliografia e Sitografia

- [127] V. Gadesha e E. Kavlakoglu, «What is text generation?,» IBM, 2024. [Online]. Available: <https://www.ibm.com/think/topics/text-generation>.
- [128] C. Stryker e M. Scapicchio, «Che cos'è l'AI generativa?,» IBM, 2024. [Online]. Available: <https://www.ibm.com/it-it/topics/generative-ai>.
- [129] S. S. G. M. R. Y. M. Y. I. Takeshi Kojima, «Large Language Models are Zero-Shot Reasoners,» 2022. [Online]. Available: <https://doi.org/10.48550/arXiv.2205.11916>.
- [130] «YFCC100M Dataset,» [Online]. Available: <https://paperswithcode.com/dataset/yfcc100m>.
- [131] «The Authors Guild, John Grisham, Jodi Picoult, David Baldacci, George R.R. Martin, and 13 Other Authors File Class-Action Suit Against OpenAI,» 2023. [Online]. Available: <https://authorsguild.org/news/ag-and-authors-file-class-action-suit-against-openai/>.