

DANIEL RAFFINI
(SAPIENZA UNIVERSITÀ DI ROMA)

AGNESE MACORI
(SAPIENZA UNIVERSITÀ DI ROMA)

MARCO ANGELINI
(LINK UNIVERSITY OF ROME)

TIZIANA CATARCI
(CNR – SAPIENZA UNIVERSITÀ DI ROMA)

DA PROPP AL PROMPT:
I PREGIUDIZI DI GENERE NELLE STORIE GENERATE
DAI *LARGE LANGUAGE MODEL*¹

1. L'IA GENERATIVA ALLA PROVA DEI PREGIUDIZI DI GENERE

Negli ultimi anni una grande attenzione è stata dedicata al problema dei pregiudizi nei *Large Language Model* (LLM). Nonostante gli sforzi e i miglioramenti in corso, infatti, spesso questi non rappresentano ancora adeguatamente la diversità e continuano a diffondere varie forme di stereotipi sociali nei loro *output*². L'uso estensivo degli LLM per la creazione di contenuti e la

¹ Il presente articolo è la traduzione di un contributo pubblicato in inglese: D. RAFFINI, A. MACORI, M. ANGELINI, T. CATARCI, *A Close Reading Approach to Gender Narrative Biases in AI-Generated Stories*, in *IEEE International Conference on Cyber Humanities* (IEEE-CH), Florence, 2025, pp. 1-9, doi: 10.1109/IEEE-CH65308.2025.11279292.

² I. O. GALLEGOS, R. A. ROSSI, J. BARROW, M. TANJIM, S. KIM, F. DERNONCOURT, T. YU, R. ZHANG, N. K. AHMED, *Bias and fairness in Large Language Models: a survey*. 2023. doi: 10.48550/arXiv.2309.00770; N. TORRES, C. ULLOA, I. ARAYA, ET AL., *A comprehensive analysis of gender, racial, and prompt-induced biases in large language models*, in «Int J Data Sci Anal», 2024. doi: 10.1007/s41060-024-00696-6; H. ZHOU, D. INKPEN, B. KANTARCI, *Evaluating and mitigating gender bias in generative large language models*, in «International Journal of Computers Communications & Control», Vol. 19, No. 6, 2024. doi: 10.15837/ijccc.2024.6.6853.

generazione di testi rende tale questione sempre più urgente. Per quanto riguarda il pregiudizio di genere, alcuni studi hanno esplorato diversi aspetti specifici, come la correlazione tra genere e occupazione³, la caratterizzazione dei personaggi⁴ o l'aggettivazione⁵.

I modi per studiare i *bias* negli LLM si dividono in due tipi: intrinseci ed estrinseci⁶. La prima modalità esamina i pregiudizi “interni” al modello, cioè quelli che emergono dal modo in cui questo rappresenta e associa le parole tra loro; la seconda si concentra invece sui pregiudizi “esterni”, ovvero quelli che si manifestano nei testi o nelle immagini che il modello produce⁷. In pratica: il primo approccio guarda a *come il modello pensa* (le associazioni interne tra concetti), il secondo guarda a *cosa il modello scrive* (i testi che genera). Possiamo prendere come esempio delle strategie di analisi impiegate lo studio pubblicato nel 2024 dall'UNESCO, *Challenging systematic prejudices: an Investigation into Gender Bias in Large Language Models*⁸. I ricercatori e le ricercatrici dell'UNESCO hanno chiesto ai modelli di generare storie su ragazzi, ragazze, donne e uomini, e poi hanno creato una *word cloud* per ciascuna categoria, portando alla luce stereotipi nell'ambientazione della storia e negli aggettivi utilizzati. L'analisi è stata condotta su un campione di 1000 persone per ciascuna categoria, e le storie sono state analizzate utilizzando

³ H. KOTEK, R. DOCKUM, D. SUN, *Gender bias and stereotypes in Large Language Models*, in *Proceedings of The ACM Collective Intelligence Conference*, Association for Computing Machinery, New York, 2023, pp 12–24. doi: 10.1145/3582269.3615599; N. DÖLL, M. DÖHRING, A. MÜLLER, *Evaluating Gender Bias in Large Language Models*, 2024. arXiv:2411.09826

⁴ R. RANJAN, R. S. GUPTA, S. N. SINGH, *Gender Biases in LLMs: Higher intelligence in LLM does not necessarily solve gender bias and stereotyping*, 2024. ArXiv:2409; M CHENG, E. DURMUS, D. JURAFSKY, *Marked personas: Using natural language prompts to measure stereotypes in language models*, in *Proceedings of the 61st annual meeting of the association for computational linguistics*, Association for Computational Linguistics, 2023.

⁵ S. SOUNDARARAJAN, S. J. DELANY, *Investigating Gender Bias in Large Language Models Through Text Generation*, in *Proceedings of the 7th International Conference on Natural Language and Speech Processing*, Association for Computational Linguistics, 2024, pp. 410- 424.

⁶ A. ZAYED, G. MORDIDO, S. SHABANIAN, I. BALDINI, S. CHANDAR. *Fairnessaware structured pruning in transformers*, in *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 38, 2024, pages 22484–22492; Y. LI, M. DU, R. SONG, X. WANG, Y. WANG, *A survey on fairness in large language models*, 2023. arXiv:2308.10149.

⁷ Z. CHU, Z. WANG, W. ZHANG, *Fairness in Large Language Models: A Taxonomic Survey*, in «SIGKDD Explor. Newsl», 26, 1, 2024, 34–48. doi: 10.1145/3682112.3682117.

⁸ UNESCO, IRCAI, *Challenging systematic prejudices: an Investigation into Gender Bias in Large Language Models*, 2024.

un approccio computazionale. I risultati hanno mostrato alcune caratteristiche interessanti: ad esempio, nelle storie sulle donne, i mariti venivano menzionati più frequentemente rispetto a quanto avvenisse per le mogli nelle storie sugli uomini, e le donne erano associate spesso a ruoli e ambientazioni stereotipati. Lo studio ha inoltre rivelato che gli stereotipi familiari erano prevalenti quando agli LLM veniva chiesto di ambientare la storia nel Sud del mondo, mentre l'amore era il tema principale associato alle donne nelle narrazioni ambientate nel Nord del mondo, suggerendo che gli stereotipi di genere possono variare a seconda delle identità culturali.

Esistono diverse forme di *bias* di genere e di propagazione degli stereotipi, ed è importante valutare la questione da vari punti di vista. Sono state proposte diverse tassonomie dei pregiudizi, basate su fattori quali l'origine, i soggetti coinvolti e le implicazioni sociali⁹. Dal punto di vista degli studi sulla comunicazione e sulla narrativa, proponiamo di distinguere tre tipologie:

- 1) il *bias* linguistico, che deriva dall'uso di determinate caratteristiche del linguaggio, come il maschile sovraesteso o parole connotate dal punto di vista del genere. Questo tipo di pregiudizio può essere studiato, ad esempio, analizzando gli *embedding* delle parole o le co-occorrenze.
- 2) il *bias* interpretativo, che influenza l'interpretazione di un testo e influisce sulla sua comprensione. È riscontrabile in compiti quali la sintesi, l'analisi del testo, l'estrazione di informazioni, la classificazione e la risposta a domande basate su affermazioni.
- 3) il *bias* narrativo, cioè quando gli stereotipi emergono non da un singolo elemento linguistico, ma da una narrazione che coinvolge più passaggi, descrizioni e azioni. Questo tipo di pregiudizi deriva tipicamente dalla generazione libera di storie in risposta a un *prompt* specifico.

Nel presente articolo ci concentreremo sui pregiudizi narrativi. Tra le metodologie esistenti, la generazione libera è la più adatta per analizzare questo tipo di *bias*, poiché ci consente di esaminarne lo sviluppo nell'arco di interi testi, anziché limitarsi a brevi frasi o singole parole. Nell'analisi dei pregiudizi di genere nei testi generati dagli LLM sono molto diffusi gli approcci computazionali e di *distant reading*¹⁰. La maggior parte degli studi esistenti ricorre

⁹ R. NAVIGLI, S. CONIA, B. ROSS, *Biases in Large Language Models: Origins, Inventory, and Discussion*, in «J. Data and Information Quality», 15, 2, 2023. doi: 10.1145/3597307; R. RANJAN, S. GUPTA, S. N. SINGH, *A Comprehensive Survey of Bias in LLMs: Current Landscape and Future Directions*, 2024. arXiv:2409.16430.

¹⁰ S. H. KUMAR, S. SAHAY, S. MAZUMDER, E. OKUR, R. MANUVINAKURIKE, N. BECKAGE, H. SU, H. LEE, L. NACHMAN, *Decoding Biases: Automated Methods and LLM Judges for Gender Bias Detection in Language Models*, 2024. arXiv:2408.03907.

infatti a tecniche computazionali, quali i classificatori¹¹ o il *topic modelling*¹². Nel nostro studio proponiamo invece un approccio basato sul *close reading*, concentrandoci su un campione più ristretto, così da offrire un'analisi approfondita della rappresentazione dei personaggi, utilizzando la critica letteraria, le teorie narratologiche e l'analisi retorica combinate con gli studi di genere¹³. Riteniamo infatti che un approccio umano, e non computazionale, sia utile per far emergere pregiudizi impliciti, che non sono sempre evidenti a livello linguistico e richiedono un collegamento al contesto sociale e retorico.

Alcuni studi hanno già dimostrato l'utilità dell'analisi dei personaggi nella generazione di storie tramite IA, ma si sono concentrati maggiormente sui protagonisti¹⁴. Per questo abbiamo deciso di includere nella nostra analisi diverse tipologie di personaggi, seguendo la classificazione di Propp¹⁵, al fine di indagare il pregiudizio narrativo nella distribuzione, nella connotazione e nelle relazioni tra i diversi tipi di soggetti nelle storie generate dagli LLM. Riteniamo infatti che affrontare il pregiudizio di genere nelle storie prodotte dall'IA è essenziale, data l'importanza dello *storytelling* nella comunicazione contemporanea e il crescente uso degli LLM per generare contenuti testuali e narrazioni che possono influenzare e plasmare il nostro immaginario. Il perpetuarsi dei pregiudizi può influenzare il dominio della narrazione e dell'immaginario, rafforzando gli stereotipi che, come società, stiamo cercando di superare. Concentrarsi sul pregiudizio narrativo è particolarmente interessante perché i *bias* impliciti trasmessi attraverso la narrazione possono più facilmente influenzare le credenze umane¹⁶. Il presente articolo intende valutare

¹¹ SOUNDARARAJAN e DELANY, *Investigating Gender Bias in Large Language Models Through Text Generation*, cit.

¹² L. LUCY, D. BAMMAN, *Gender and Representation Bias in GPT-3 Generated Stories*, in *Proceedings of the Third Workshop on Narrative Understanding*, pages 48–55. Association for Computational Linguistics, 2021.

¹³ B. J. SMITH, *What was 'close reading'? A century of method in literary studies*, in «The Minnesota Review». 2016, 87, pp. 57-75; B. BRUMMETT, *Techniques of close reading*, Sage Publications, 2018.

¹⁴ T. HUANG, F. BRAHNAM, V. SHWARTZ, S. CHATURVEDI, *Uncovering Implicit Gender Bias in Narratives through Commonsense Inference*, in *Findings of the Association for Computational Linguistics: EMNLP*, Association for Computational Linguistics, 2021, pp. 3866-3873; LUCY e BAMMAN, *Gender and Representation Bias in GPT-3 Generated Stories*, cit.

¹⁵ V. PROPP, *Morfologia della fiaba*, Torino, Einaudi, 2000.

¹⁶ A. CALISKAN, J. J. BRYSON, A. NARAYANAN, *Semantics derived automatically from language corpora contain human-like biases*, in «Science», 356(6334), 2017, pp. 183–186. doi: 10.1126/science.aal4230.

la presenza di pregiudizi di genere nelle storie generate da tre LLM ampiamente utilizzati: ChatGPT, Gemini e Claude. I principali obiettivi del nostro studio sono:

- 1) L'analisi dei *bias* di genere nelle storie generate dagli LLM attraverso un approccio di *close reading*, in grado di ampliare gli studi precedenti incentrati su aspetti specifici e basati principalmente su metodi computazionali.
- 2) Una ricognizione dettagliata e umanocentrica dei pregiudizi narrativi presenti nei testi, in grado di mettere in luce le tendenze distintive dei tre modelli esaminati in termini di rappresentazione dei personaggi, caratterizzazione e sviluppo della trama.

2. UN PREGIUDIZIO CHE VIENE DA LONTANO: ARCHETIPI E STEREOTIPI

I pregiudizi di genere riguardano tanto i testi narrativi generati dagli LLM quanto quelli prodotti dagli esseri umani nel corso dei secoli. I modelli di IA generativa, infatti, acquisiscono la conoscenza del mondo attraverso i dati utilizzati per il loro addestramento e possono riprodurre stereotipi e semplificazioni¹⁷. Pertanto, è necessario in prima istanza concentrare l'attenzione sugli studi che hanno indagato la presenza di pregiudizi all'interno delle narrazioni umane. Tra queste, le fiabe si rivelano utili in quanto rappresentano una delle forme di racconto più elementari e schematizzabili¹⁸. Nella loro analisi delle fiabe dei fratelli Grimm, Baker-Sperry e Grauerholz¹⁹ hanno evidenziato la prevalenza della bellezza femminile e la correlazione con la longevità di queste narrazioni. Gottschall²⁰ ha affrontato invece la questione dello scarso interesse verso le protagoniste femminili, effettuando un'analisi quantitativa delle fiabe popolari globali per identificare modelli interculturali nella loro

¹⁷ E. DINAN, A. FAN, A. WILLIAMS, J. URBANEK, D. KIELA, J. WESTON, *Queens are Powerful too: Mitigating Gender Bias in Dialogue Generation*, in *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, Association for Computational Linguistics, 2020, pp. 8173-8188; D. GALA, M. O. KHURSHEED, H. LERNER, B. O'CONNOR, M. IYER, *Analyzing Gender Bias within Narrative Tropes*, in *Proceedings of the Fourth Workshop on Natural Language Processing and Computational Social Science*, Association for Computational Linguistics, 2020, pp. 212-217.

¹⁸ Cfr. PROPP, *Morfologia della fiaba*, cit.

¹⁹ L. BAKER-SPERRY E L. GRAUERHOLZ, *The Pervasiveness and Persistence of the Feminine Beauty Ideal in Children's Fairy Tales*, in «Gender & Society», vol. 17, no. 5, pp. 711-726, 2003. doi: 10.1177/0891243203255605.

²⁰ J. GOTTSCHALL ET AL., *The "Beauty Myth" Is No Myth*, in «Hum Nat», vol. 19, pp. 174-188, 2008. doi: 10.1007/s12110-008-9035-3.

caratterizzazione e dimostrando un'enfasi pressoché universale sull'attrattiva fisica delle donne. L'analisi statistica di Ragan²¹ su una corposa raccolta multiculturale di fiabe popolari ha evidenziato l'impatto del genere sulla genesi e sull'interpretazione di queste narrazioni. Ancora sul lato della ricezione, Francemone²² ha ipotizzato che, sebbene gli schemi dei personaggi operino in modo simile tra i generi, le valutazioni morali del comportamento suscitano risposte affettive più forti nei confronti dei personaggi femminili. Infine, l'analisi comparativa di Nusbaumer²³ tra *Cenerentola* e *Love in a Fallen City*²⁴ ha esplorato il concetto di "complesso di Cenerentola", sostenendo che entrambi i racconti, nonostante le potenziali interpretazioni femministe, reiterano la dipendenza delle donne dall'intervento maschile. Gli studi sull'argomento rivelano insomma una tendenza storica costante nelle narrazioni umane, caratterizzata dall'enfasi sull'attrattiva fisica femminile e dalla rappresentazione delle donne all'interno di ruoli specifici, spesso dipendenti da quelli maschili. La presenza di modelli consolidati nei testi scritti da esseri umani fornisce un quadro significativo per valutare il potenziale di replicazione o sovversione di tali rappresentazioni di genere nelle storie generate dagli LLM.

La pervasività dei ruoli connotati dal genere è tale da aver consentito a Propp una rigorosa formalizzazione²⁵. Il critico ha proposto infatti l'individuazione di personaggi tipici all'interno delle fiabe nella sua opera, *Morfologia della fiaba*: l'antagonista, il donatore, l'aiutante, la principessa (e suo padre), il mandante, l'eroe, il falso eroe. Dal punto di vista degli stereotipi di genere, è importante notare che la formalizzazione di Propp classifica il protagonista come maschio e l'oggetto del desiderio come femmina. La principessa è inoltre associata a una figura maschile, il padre, capace di proteggerla e salvaguardarla fino all'arrivo dell'eroe che la sposerà. La metodologia di Propp è induttiva, parte cioè dall'analisi di un corpus di racconti popolari, a partire

²¹ K. RAGAN, *What Happened to the Heroines In Folktales?: An Analysis by Gender of a Multicultural Sample of Published Folktales Collected from Storytellers*, in «Marvels & Tales», vol. 23, no. 2, pp. 227-247, 2009. doi: 10.1353/mat.2009.a369130.

²² C. J. FRANCIMONE, M. GRIZZARD, K. FITZGERALD, J. HUANG, E. C. AHN, *Character Gender and Disposition Formation in Narratives: The Role of Competing Schema*, in «Media Psychology», vol. 25, no. 4, pp. 547-564, 2021. doi: 10.1080/15213269.2021.2006718.

²³ E. R. NUSBAUMER, *The Cinderella Stereotype: A Comparative Study of Love in a Fallen City and Cinderella*, in «Comparative Literature: East & West», vol. 8, no. 2, pp. 221-233, 2024. doi: 10.1080/25723618.2024.2441001.

²⁴ E. CHANG, *Love in a Fallen City*, in *Love in a Fallen City and Other Stories*, London, U.K.: Penguin, 2007.

²⁵ PROPP, *Morfologia della fiaba*, cit.

dalla quale individua caratteristiche comuni. Di conseguenza, diventa evidente che una percentuale significativa di queste narrazioni ritrae le donne come soggetti vulnerabili e oggetti del desiderio e dello sguardo maschile.

Propp in realtà non classifica i personaggi in sé, quanto piuttosto le sfere d'azione che appartengono a ciascun ruolo e propone una divisione della struttura della fiaba in trentuno funzioni che rappresentano le singole azioni che, combinate, costituiscono l'ossatura di queste narrazioni. Se la categorizzazione dei personaggi è funzionale al nostro studio, tuttavia la struttura proposta da Propp risulta troppo complessa per essere fornita come istruzione a un LLM. A livello di critica, già diversi studiosi hanno tentato di adattare e sintetizzare lo schema di Propp. Greimas²⁶ ha ridotto lo schema a sei elementi, mentre Todorov²⁷ ha teorizzato che la caratteristica fondamentale di una narrazione è il passaggio da uno stato di equilibrio a un altro.

A livello di struttura della storia, ai fini del nostro studio, la schematizzazione più chiara ed efficace è la piramide di Freytag²⁸, che fornisce una rappresentazione grafica dei cinque punti essenziali della trama. Ispirandosi alla partizione proposta da Aristotele nella *Poetica*²⁹, questo modello suddivide l'intreccio in: un'introduzione iniziale, una serie di complicazioni e azioni trasformative che modificano negativamente la situazione, un punto di suspense che rappresenta il climax del racconto, una fase di risoluzione delle difficoltà iniziali e uno scioglimento in cui si compie il destino dell'eroe. Sebbene questo schema sia associato principalmente al dramma, è utile per schematizzare la progressione tipica delle narrazioni tradizionali. A differenza della struttura granulare di Propp, il modello di Freytag non impone vincoli rigidi sulla relazione tra i personaggi e le loro azioni. Questa distinzione è essenziale nel passaggio dalla fase ermeneutica, che implica la comprensione delle narrazioni esistenti, alla fase creativa di produzione di testi originali. Il modello di Propp, se utilizzato come *input* per l'IA generativa, avrebbe limitato le azioni dei personaggi in base alla loro funzione designata. Al contrario, la piramide di Freytag stabilisce condizioni per lo sviluppo della trama, consentendo al contempo una maggiore variabilità nella creazione e nell'*agency* dei personaggi. Questa flessibilità è fondamentale per promuovere l'originalità e prevenire la reiterazione di comportamenti stereotipati dei personaggi che potrebbero derivare dal modello restrittivo di Propp.

²⁶ A.J. GREIMAS, *Semantica strutturale*, Milano, Rizzoli, 1968.

²⁷ T. TODOROV, *Poetica della prosa. Le leggi del racconto*, Milano, Bompiani, 1995.

²⁸ G. FREYTAG, *Technique of the Drama: An Exposition of Dramatic Composition and Art*, Chicago, IL, USA, Scott, Foresman & Company, 1900.

²⁹ ARISTOTELE, *Poetica*, Torino, Einaudi, 2008.

3. UN PROMPT TRA PROPP E FREYTAG

Tra le metodologie disponibili per analizzare i pregiudizi di genere negli LLM, abbiamo scelto la generazione libera di storie (*open-ended generation*) attraverso un *prompt* strutturato³⁰. In questo contesto, è necessaria un'attenzione alla definizione del *prompt*, come dimostrato da studi precedenti³¹. Abbiamo optato per lo *zero-shot prompting*³², ossia una tecnica che non prevede la somministrazione di esempi specifici del risultato atteso, con l'obiettivo di lasciare libera la generazione del testo e comprendere meglio la produzione di pregiudizi di genere negli LLM senza rischiare un'inferenza a partire dagli esempi forniti.

Il *prompt* è stato progettato per fornire istruzioni sulle funzioni dei personaggi e sulla struttura del testo, a partire dalle funzioni di Propp per i personaggi e dalla struttura della storia di Freytag. Data la natura dei racconti brevi, i sette tipi di personaggi originali di Propp sono stati ridotti e adattati. Abbiamo scelto di rinominare l'eroe come "personaggio principale" e la principessa come "personaggio desiderato", puntando a una terminologia più neutra. Inoltre, abbiamo unito l'aiutante e il donatore in un unico ruolo ed eliminato il falso eroe, poiché l'inclusione di questa figura avrebbe introdotto un'eccessiva complessità narrativa, incompatibile con la lunghezza limitata dei testi. Il risultato di questo processo è un set di cinque personaggi, ai quali faremo riferimento in questo articolo utilizzando gli acronimi seguenti: personaggio principale (PP), il cattivo (C), l'aiutante (A), il personaggio desiderato (PD), il mandante (M). Il *prompt* finale – in inglese – è stato il seguente:

You are a writer, and you are asked to write a story about 500 words. Before generating the story, read the following instructions until the end. The spatial and temporal setting is your choice. You can also choose the tone and the atmosphere. However, you must follow some rules. In the story, there must be five characters, each corresponding to one of these roles: the main character, the villain, the helper, the desired character, the dispatcher. The role does not have to be mentioned in the story, each character has only its name. Each character must be characterized physically and psychologically. Furthermore, the sequence of events must follow this scheme, made of 5 phases: Exposition, Rise, Climax, Return or Fall, Catastrophe. The

³⁰ E. CHEN, R. ZHAN, Y. LIN, H. CHEN, *From Structured Prompts to Open Narratives: Measuring Gender Bias in LLMs Through Open-Ended Storytelling*, 2025. arXiv:2503.15904

³¹ DÖLL, DÖHRING, MÜLLER, *Evaluating Gender Bias in Large Language Models*, cit.

³² SOUNDARARAJAN, DELANY, *Investigating Gender Bias in Large Language Models Through Text Generation*, cit.

narration must be unified and fluid, without mentioning the name of the phases, and must not contain lists³³.

Per generare i testi abbiamo selezionato tre LLM ampiamente utilizzati: ChatGPT, Gemini e Claude. Tutti questi modelli presentano caratteristiche specifiche orientate a usi narrativi e creativi³⁴. Il *close reading* umano permette di cogliere forme di pregiudizio narrativo sfumate, dipendenti dal contesto e spesso implicite, che potrebbero sfuggire agli strumenti computazionali automatizzati. Adottando un approccio di *close reading*, è stato possibile lavorare con un numero limitato di testi. Per ciascun modello, abbiamo quindi generato cinque storie durante diverse sessioni di lavoro e in finestre conversazionali separate, per evitare l'effetto memoria legato al mantenimento del contesto precedente. Sebbene la dimensione del campione sia relativamente ridotta, l'uso di un *prompt* strutturato e di ruoli standardizzati per i personaggi garantisce coerenza e comparabilità tra i risultati. Inoltre, applicando strumenti di critica letteraria – come i modelli narratologici, l'analisi retorica e gli studi di genere – siamo stati in grado di identificare forme di pregiudizio che potrebbero manifestarsi solo attraverso l'interazione di molteplici elementi narrativi.

Per valutare i risultati, ci siamo concentrati su tre aspetti: la distribuzione di genere tra i cinque tipi di personaggi, il modo in cui i personaggi vengono descritti e le azioni a loro attribuite. Per queste analisi, abbiamo preparato una scheda di lettura che contiene informazioni sull'aderenza al *prompt*, sulle caratteristiche dei personaggi e sui pregiudizi di genere. La nostra analisi si concentra quindi sulla distribuzione di genere, sulle connotazioni fisiche e psicologiche e sullo sviluppo della trama. Integrando queste informazioni possiamo ottenere una visione approfondita delle differenze nella rappresentazione di genere e nei pregiudizi impliciti.

³³ Trad.: «Sei uno scrittore e ti viene chiesto di scrivere una storia di circa 500 parole. Prima di generare il racconto, leggi le seguenti istruzioni fino alla fine. L'ambientazione spaziale e temporale è a tua scelta. Puoi anche scegliere il tono e l'atmosfera. Tuttavia, devi seguire alcune regole. Nella storia devono esserci cinque personaggi, ognuno corrispondente a uno di questi ruoli: il personaggio principale, il cattivo, l'aiutante, il personaggio desiderato, il mandante. Il ruolo non deve essere menzionato esplicitamente nella storia; ogni personaggio deve avere solo il proprio nome. Ogni personaggio deve essere caratterizzato sia fisicamente che psicologicamente. Inoltre, la sequenza degli eventi deve seguire questo schema, composto da 5 fasi: esposizione, ascesa, climax, ritorno o discesa, catastrofe. La narrazione deve essere unitaria e fluida, senza menzionare il nome delle fasi e non deve contenere elenchi.»

³⁴ C. GÓMEZ-RODRÍGUEZ E P. WILLIAMS, *A Confederacy of Models: a Comprehensive Evaluation of LLMs on Creative Writing*, in *Findings of the Association for Computational Linguistics*, Association for Computational Linguistics, 2023, pp. 14504–14528.

4. ANALISI DELLE STORIE GENERATE DAGLI LLM

4.1. ADERENZA AL PROMPT

Il primo aspetto da valutare è la comprensione del *prompt*, che risulta differente tra i tre modelli. In tutte le storie generate da ChatGPT e Gemini compaiono i cinque personaggi richiesti, tranne in due casi in cui manca M. Questo sembra essere infatti il ruolo più difficile e gli LLM tendono a non differenziarlo abbastanza da A. Non è l'unico caso di sovrapposizione; infatti, si riscontra una oscillazione anche tra PD e A. In un caso è presente un personaggio in più, presentato come aiutante dell'antagonista. In un altro caso, il PD è un oggetto: sebbene ciò indichi una mancata comprensione della nozione di personaggio come essere umano, è comunque in linea con la tradizione delle fiabe sulla ricerca dell'oggetto magico da parte dell'eroe. Vale inoltre la pena notare che, in un caso, ChatGPT etichetta esplicitamente le funzioni dei personaggi accanto ai loro nomi (ad esempio, «Solen, the helper»), nonostante il *prompt* istruisse chiaramente di non farlo. Una considerazione a parte va fatta per Claude, che mostra una minore aderenza ai ruoli funzionali dei personaggi: sebbene sia l'unico modello che dichiara esplicitamente – senza che gli venga richiesto – a quale funzione proppiana corrisponda ogni personaggio, in pratica tende a presentare un PP, un C e una serie di A che solo occasionalmente mostrano tratti secondari associati ai ruoli di M o di PD. Pertanto, nelle storie si riscontra, in alcuni casi, confusione sulla comprensione dei ruoli. Questa mancanza, tuttavia, generalmente non compromette la possibilità di generare storie coerenti.

La seconda parte del *prompt* chiedeva invece di seguire la struttura narrativa definita da Freytag. I testi generati seguono generalmente le cinque fasi proposte; tuttavia, in alcuni casi, si notano incongruenze nello sviluppo della trama o uno squilibrio tra la parte ascendente e quella discendente, che annulla l'effetto di bipartizione della storia della struttura narrativa di Freytag. Si noti come il tentativo di aderire alle cinque fasi si risolva spesso in una scarsa coerenza interna alla trama, con transizioni brusche o riferimenti lasciati nel vuoto, privi di una reale funzione narrativa. In definitiva, la capacità di costruire storie narrativamente efficaci è generalmente insufficiente, mostrando spesso una struttura “a caselle da riempire” piuttosto che uno sviluppo coerente. Nel caso di Claude, si osserva un'eccessiva standardizzazione delle storie, che seguono tutte un intreccio simile, che rimanda al genere poliziesco. Vale inoltre la pena notare una certa ripetitività in Gemini riguardo alle ambientazioni, agli incipit delle storie e ai nomi dei personaggi (la protagonista, ad esempio, si chiama sempre Elara). Di fronte a queste criticità, è necessario ricordare che nel *prompt* abbiamo fornito le funzioni e le fasi della

narrazione dei personaggi senza includere alcun esempio. Se da un lato questo approccio è utile per evitare di alterare la produzione con dei *bias*, riconosciamo che, per quanto riguarda il ruolo e la struttura dei personaggi, descrizioni ed esempi aggiuntivi potrebbero essere utili per migliorare le prestazioni degli LLM.

4.2. LA DISTRIBUZIONE DI GENERE

Un primo livello di analisi è stato la distribuzione di genere dei personaggi a livello quantitativo, che è riportata nella tabella 1:

MODELLO	MASCHILE	FEMMINILE	OGGETTO
ChatGPT	67%	33%	0%
Gemini	60%	36%	4%
Claude	52%	48%	0%
Totale	61%	38%	1%

Tabella 1

La tabella 2, invece, mostra i dati combinati sulla distribuzione di genere tra tutti i personaggi.

PERSONAGGIO	MASCHILE	FEMMINILE	OGGETTO
PP	27%	73%	0%
C	100%	0%	0%
A	67%	33%	0%
PD	47%	47%	6%
M	62%	38%	0%

Tabella 2

Nel caso del PP, registriamo un valore complessivo a favore del genere femminile, ma anche una forte differenza tra i tre modelli: Gemini e Claude scelgono sempre un PP femminile, mentre ChatGPT lo fa solo in un caso. Il ruolo di C mostra il valore più polarizzato, con il 100% di maschi. Sia A che M sono maggiormente uomini. È interessante notare che i PP femminili hanno il 55% di A femminili, mentre tutti i PP maschili hanno A maschi. I PD sono equilibrati, ma si osserva una differenza tra i tre modelli: ChatGPT ha generato 3 PD femmine, 1 maschio e 1 oggetto; Gemini 3 maschi e 2 femmine; Claude 3 maschi e 2 femmine. Anche in questo caso, è interessante incrociare i dati: i PP maschi sono legati a 3 PD femmine e un oggetto, mentre i PP femmine sono legati a 7 PD maschi e 4 PD femmine.

I risultati rivelano uno squilibrio nella distribuzione di tutti i personaggi per ciascuno dei tre modelli. Lo squilibrio è sempre a favore del genere

maschile, tranne nel caso del PP. Tra i modelli, ChatGPT mostra il maggior squilibrio nella distribuzione di genere, mentre Claude è quello che più si avvicina a un equilibrio.

4.3. DESCRIZIONE DEI PERSONAGGI

L'analisi dei tratti fisici e psicologici dei personaggi si basa sull'identificazione e l'interpretazione dei termini a cui fanno ricorso le descrizioni. Questa analisi svela una sistematica dicotomia di genere nella strutturazione degli attributi, così come nelle implicazioni simboliche associate a tali tratti. Le descrizioni dei personaggi femminili – in particolare quelli nei ruoli di PP, A e PD – attingono frequentemente a campi lessicali associati alla bellezza, alla grazia e alla delicatezza, includono riferimenti alla magrezza, all'eleganza, ai lineamenti raffinati e alla fluidità dei movimenti. Si può notare la frequenza di aggettivi come «graziosa», «pallida», «esile» e «dai capelli rossi», che contribuiscono alla costruzione di una figura femminile visivamente idealizzata. In diversi casi, si osserva un'enfasi sulla vulnerabilità fisica, con donne caratterizzate da bassa statura, ferite corporee o tratti non minacciosi. Questi elementi si accompagnano a caratteristiche psicologiche quali la resilienza o la determinazione, inquadrando la forza femminile come un fattore interiore piuttosto che fisico.

Gli attributi dei personaggi maschili sono caratterizzati da un uso predominante di termini che denotano forza, robustezza e irregolarità. I loro corpi – in particolare di quelli che assumono i ruoli di C, A e M – sono frequentemente descritti come robusti, distorti o imponenti. Nel caso di C, si nota una tendenza all'esagerazione dei tratti fisici, inclusa la presenza di cicatrici o modificazioni meccaniche: tratti che portano un peso simbolico legato all'aggressività o alla devianza morale. Anche quando i personaggi maschili non sono descritti esplicitamente come forti, l'analisi ha rivelato una tendenza a ritrarli come fisicamente eccentrici o mutilati. Questi personaggi, infatti, sono spesso raffigurati come zoppi, anziani o deformi.

Anche per i PD la caratterizzazione fisica si allinea a costrutti di genere idealizzati. Le figure femminili in questo ruolo sono frequentemente caratterizzate da attributi associati all'etereità e all'astrazione, che si manifestano in elementi visivi come gli occhi luminosi o inaccessibili e una bellezza intangibile. Questi elementi sono associabili a una resa della femminilità stilizzata e spesso evocativa. I PD maschi, sebbene in una certa misura anch'essi idealizzati, sono più frequentemente caratterizzati da tratti indicanti fragilità emotiva o vulnerabilità (pallore, stanchezza o vitalità ridotta).

Da un punto di vista psicologico, i tratti che emergono sembrano indicare tendenze narrative più variegate. I personaggi maschili mostrano una va-

sta gamma di descrittori, che spaziano da una moralità connotata negativamente alla profondità emotiva. Nell'analisi di C è emerso il ricorso ricorrente ai campi semantici dell'egoismo, della crudeltà e dell'ambizione. Tuttavia, nei ruoli di A, M e PD, le figure maschili sono associate anche alla sensibilità emotiva e all'affidabilità morale, venendo caratterizzate da aggettivi come «altruista», «gentile» e «intelligente». Ai personaggi femminili sono invece assegnati tratti psicologici più uniformi. In particolare nel PP e nell'A, si nota la prevalenza dei domini semantici della resilienza emotiva, dell'empatia e della saggezza. Termini come «forte», «altruista», «gentile» e «intelligente» emergono frequentemente, suggerendo un modello narrativo che privilegia la coerenza interna, la maturità affettiva e la forza etica. I M femminili risultano spesso meno dettagliati; quando presenti, fanno tipicamente riferimento ad attributi quali mistero, affidabilità e competenza intellettuale.

Le caratteristiche descritte sono tendenzialmente comuni a tutti e tre gli LLM analizzati. L'unica osservazione degna di nota riguarda Gemini, il quale, rispetto a Claude e ChatGPT, pone meno enfasi sulla bellezza dei personaggi femminili, che in alcuni casi mostrano anche tratti di forza fisica. In sintesi, l'analisi dimostra che i tratti dei personaggi non sono distribuiti in modo neutro, ma sono modellati da convenzioni narrative di genere. I personaggi femminili sono spesso costruiti attraverso *cluster* coesi di attributi estetici ed emotivi, mentre ai personaggi maschili viene concessa una gamma descrittiva più ampia, sebbene spesso più polarizzata. Questi modelli riflettono quindi asimmetrie strutturali persistenti nella caratterizzazione dei personaggi, anche laddove alcuni ruoli iniziano a riflettere identità più complesse e sfaccettate.

4.4. AZIONI

L'analisi dell'azioni attribuite e subite è in grado di mostrare l'agency e l'intenzionalità che caratterizzano i personaggi. Le PP femminili sono spesso associate ad azioni esplorative, di resistenza e di trasformazione. Per esempio, questi personaggi scalano crinali, perlustrano terreni pericolosi, sfuggono a minacce, affrontano antagonisti, prendono decisioni critiche e resistono ad avversità psicologiche o fisiche. Nonostante i vincoli fisici e sociali a cui sono spesso soggette, le loro azioni dimostrano un notevole grado di autonomia. Un tratto saliente del loro processo decisionale è la complessità emotiva e morale che spesso caratterizza le loro scelte. Al contrario, i PP maschi sono prevalentemente associati ad atti di attraversamento, riparazione e confronto eroico, con azioni come il superamento di ostacoli fisici, l'impegno in combattimenti diretti e azioni o atti di protezione verso gli altri. Sebbene anche questi personaggi subiscano uno sviluppo emotivo, i verbi associati – come «riparare», «distruggere», «affrontare» e «salvare» – suggeriscono una modalità di

comportamento più esternalizzata e orientata all'azione, basata sull'intervento fisico. I soggetti maschili mostrano peraltro un'alta frequenza di comportamenti invasivi, distruttivi e manipolatori. Questi includono l'inseguimento dei protagonisti, l'attacco a luoghi sicuri (ad esempio, le case), l'uso dell'inganno e l'esercizio di un controllo coercitivo (come l'imprigionamento, il sabotaggio o la dominazione psicologica). In diversi casi, gli atti dei C maschi si estendono alla violenza: gli esempi includono il tentativo di demolire biblioteche o l'appropriarsi del merito dei successi altrui. Questi atti sono legati all'affermazione di un dominio e alla rottura dell'ordine sociale o etico.

Se il PP è un ruolo intrinsecamente legato all'azione, è interessante, al contrario, analizzare come l'azione varia nel caso di un ruolo più passivo come quello del PD. Le figure femminili in questo ruolo sono generalmente inattive, assumendo la funzione di compagne, prigioniere o fonti di ispirazione. Quando sono attive, svolgono invece compiti di supporto o simbolici, come condurre ricerche, offrire conforto emotivo o aiutare i protagonisti in momenti critici. I PD maschi, d'altro canto, dimostrano una maggiore iniziativa, esibendo spesso comportamenti protettivi o di guida. Tuttavia, una parte di questi personaggi adotta una posizione passiva o marginale. Fermo restando che la passività è una caratteristica inerente al ruolo di PD, i testi analizzati rivelano una disparità nella capacità di *agency* che favorisce i personaggi maschili.

Differenti tipologie di azioni si riscontrano anche negli A. Le figure femminili in questo ruolo sono costantemente coinvolte nel guidare, avvertire, supportare e talvolta vendicare il protagonista. Le loro azioni sono tipicamente indirette o di consulenza, rafforzando così il loro ruolo di ancore morali o emotive. Al contrario, gli A maschi intraprendono una gamma di azioni più ampia, che comprende l'istruzione, l'assistenza strategica e il sacrificio fisico. La loro presenza nelle fasi di transizione, in particolare quelle che comportano la trasmissione di conoscenze o il confronto diretto, denota una funzione narrativa che integra la praticità con un forte significato simbolico.

L'analisi delle azioni mette in evidenza una disparità di genere nell'*agency* narrativa, senza differenze significative tra i tre modelli. I personaggi femminili, in particolare quelli che assumono ruoli da protagonista, prendono parte ad azioni intricate e prolungate, frequentemente associate alla tenacia, all'esplorazione e alla risoluzione morale. I personaggi maschili, specialmente nei ruoli di antagonista o di aiutante, sono caratterizzati da azioni decise, esteriori e spesso improntate alla forza. Questi risultati suggeriscono che le forme di azione considerate narrativamente autorevoli mantengono spesso una polarizzazione di genere: l'azione femminile è più frequentemente legata alla resilienza e alla cura, mentre l'azione maschile è più spesso associata alla rottura o al ripristino dell'ordine.

4.5. ANALISI DELLE TRAME

Il livello successivo di analisi riguarda il modo in cui le azioni dei personaggi si intrecciano nella formazione della trama, al fine di mostrare come i pregiudizi impliciti possano essere veicolati nelle costruzioni narrative. Ci concentreremo sul rapporto tra il PP e il PD, poiché la relazione tra questi due personaggi è quella più colpita dal *bias* di genere. Come menzionato, Gemini e Claude hanno generato solo PP femminili, mentre ChatGPT ha scelto quattro PP maschili su cinque. Esaminiamo più da vicino le dinamiche tra PP e PD nelle storie con PP maschili generate da ChatGPT. In tre casi, il PD è una donna; in un caso, è un oggetto. Tra le tre storie con un PP maschio e PD femmina, in due casi l'eroe maschile deve salvare la donna – in un'occasione si tratta della sorella, nell'altra di un interesse romantico – mentre nel terzo caso, il PD femminile non ha bisogno di essere salvato, ma funge invece da compagna di viaggio. I PD femminili sono generalmente poco sviluppati in termini di caratterizzazione e *agency*, sebbene in due casi i finali suggeriscano una possibile apertura verso una maggiore consapevolezza e riflessione. Vediamo i due esempi:

1) «Lys stayed with him, quiet and kind, but her eyes often searched the woods»³⁵: qui, vi è un contrasto tra la donna che sceglie di restare con l'eroe che l'ha salvata e il suo sguardo, che sembra cercare qualcos'altro altrove.

2) «Lira spoke less each day, her light dimming. Daryn stayed close, watching the lanterns burn low. One night, she whispered, "You didn't save me. You only broke the cage." And Daryn, watching the soot fall again, finally understood. The city would not heal. Some cages, even when shattered, remain»³⁶: in questo caso le parole del PD portano il PP a una rivelazione che mette in discussione la narrativa del salvataggio e apre a una riflessione più ampia sui vincoli sociali e personali.

Anche il caso dell'unica PP femminile generata da ChatGPT è interessante. In questo caso, il PD è un oggetto, ma esiste un forte legame (che potrebbe forse essere interpretato persino come romantico) con l'A femminile; nel frattempo, un altro personaggio maschile – che potrebbe anch'egli essere

³⁵ Trad.: «Lys rimase con lui, tranquilla e gentile, ma i suoi occhi cercavano spesso il bosco».

³⁶ Trad.: «Lira parlava ogni giorno meno, la sua luce si affievoliva. Daryn le restava vicino, guardando le lanterne consumarsi. Una notte, lei sussurrò: 'Non mi hai salvata. Hai solo rotto la gabbia'. E Daryn, guardando la fuliggine cadere di nuovo, finalmente capì. La città non sarebbe guarita. Certe gabbie, anche quando vengono frantumate, restano».

definito come un A – interviene nella lotta contro C per aiutare la protagonista, ma viene sconfitto. Complessivamente, la trama e i ruoli dei personaggi appaiono più complessi nella storia con la PP femminile.

Molto interessante è anche il caso di Gemini. Sebbene tutti i personaggi principali siano donne, le storie generate da questo modello mostrano infatti pregiudizi narrativi degni di nota. Nella prima, ad esempio, la protagonista fugge da C e viene salvata dal PD maschio. Nonostante un cambio di prospettiva che rende la donna protagonista, la struttura narrativa rimane invariata: un uomo salva una donna. Lo stesso pregiudizio ricorre in un'altra storia in cui, durante lo scontro finale tra la protagonista e l'antagonista, un misterioso messaggero interrompe la scena e risolve l'azione attraverso l'espedito narrativo del *deus ex machina*. Un altro caso presenta un'effettiva inversione della dinamica salvatore-salvato: qui, la PP femminile è la capitana di una nave che tenta di correre in soccorso di un PD maschio, ma alla fine fallisce. Significativamente, questa è l'unica storia in cui nessuna figura maschile interviene durante il conflitto, e in cui tutti i personaggi finiscono per perire per mano dell'antagonista, come a suggerire l'impossibilità di ribaltare il pregiudizio narrativo. In un altro caso, la protagonista è un'aspirante artista che trova coraggio e conforto in un PD maschio, anch'egli artista, che funge da modello di riferimento («He seemed the embodiment of the artistic success she so desperately desired»³⁷). Sebbene Gemini presenti PP femminili ritratte come forti e intraprendenti e, a prima vista, sembri quindi rispondere meglio agli stereotipi di genere, non riesce tuttavia a superare i pregiudizi narrativi più profondi. A un esame più attento, le trame poggiano infatti su dinamiche e relazioni stereotipate, in cui gli uomini svolgono ancora il ruolo di guida o di salvatore. L'unica eccezione tra i testi generati da Gemini è una storia con un PD femminile, che include il solo caso di desiderio romantico omosessuale: «Elara nodded, her gaze softening as she looked at Lyra. Her heart ached with a longing she rarely allowed herself to acknowledge. Lyra's kindness was a beacon in her solitary life»³⁸.

Per quanto riguarda Claude, si nota una predominanza del genere thriller o giallo, che sposta l'attenzione verso oggetti da recuperare o operazioni da portare a termine. In queste narrazioni, i vari personaggi convergono in opposizione al C, lavorando in gruppo per aiutare la protagonista a raggiungere il suo obiettivo. In tutte le storie, la protagonista femminile funge da custode, o di luoghi di rilevanza storica minacciati dalle ambizioni economiche

³⁷ Trad.: «Egli sembrava l'incarnazione del successo artistico che lei desiderava così disperatamente».

³⁸ Trad.: «Elara annuì, il suo sguardo si addolcì mentre guardava Lyra. Il suo cuore doleva per un desiderio che raramente si permetteva di ammettere. La gentilezza di Lyra era un faro nella sua vita solitaria».

del C (come un faro o una biblioteca), o di documenti in grado di smascherare le attività illecite del C stesso. I personaggi che circondano il PP possono essere raggruppati in due categorie principali: 1) donne portatrici di conoscenza (ad esempio, una storica, una professoressa universitaria) o di autorità (come un'ispettrice di polizia); 2) e donne e uomini che forniscono un supporto più pratico al PP. Anche nel caso di Claude, quando lo scontro con il C diventa fisico, tipicamente interviene un personaggio maschile per proteggere il PP. Tuttavia, vale la pena notare che la forte enfasi sulla collaborazione tra i personaggi tende a sfumare i ruoli di M e PD; in molti casi, tutti i personaggi di supporto gravitano verso la funzione di A. Queste narrazioni sono spesso accompagnate da finali moralistici, che condannano esplicitamente la figura del C come simbolo del male.

La tabella 3 riassume i risultati dell'analisi, indicando quali modelli siano più influenzati da ciascuna tipologia di pregiudizio.

PREGIUDIZIO	LIVELLO	TIPO DI ANALISI	MODELLO
Distribuzione di genere	Esplicito	Quantitativa	ChatGPT
Rappresentazione	Esplicito	Qualitativa	ChatGPT, Claude
Trama	Implicito	Qualitativa	ChatGPT, Gemini

Tabella 3

5. CONCLUSIONI

I risultati mostrano che i tre modelli faticano ancora con l'equilibrio di genere nella distribuzione e nella connotazione dei personaggi. Il nostro studio ha messo in evidenza diversi livelli di pregiudizio narrativo: uno legato alla descrizione dei personaggi e l'altro allo sviluppo della trama e alle relazioni tra di essi. Per quanto riguarda il primo aspetto, una lettura approfondita rivela una ricorrente enfasi sulla bellezza nelle protagoniste femminili, spesso abbinata alla forza psicologica. Ciò segna un passaggio dalla tradizionale principessa da salvare – come nella descrizione di Propp – verso un modello più affine a quello che si trova nelle narrazioni contemporanee, dove le protagoniste femminili non inseguono l'amore ma incarnano valori diversi e sono guidate dall'intelligenza e da solide convinzioni morali³⁹. Sebbene ciò sia in linea

³⁹ C. STOVER, *Damsels and Heroines: The Conundrum of the Post-Feminist Disney Princess*, «LUX: A Journal of Transdisciplinary Writing and Research from Claremont Graduate University», Vol. 2m Iss. 1, 2013; R. Daulay, *Heroine and Princess: Women Image Portrayed in selected Disney's Stories*, J. Basis Vol. 8 No. 1, 2021;

con i valori e le rappresentazioni più attuali, la sua predominanza all'interno dei modelli trasforma anche questa in una rappresentazione stereotipata, essendo predominante e presentata in forme fisse. Inoltre, la persistente enfasi sulla bellezza definisce questo modello come ancora discriminatorio, mostrando come il *bias* di genere rimanga fortemente legato ai pregiudizi relativi all'aspetto fisico⁴⁰.

D'altra parte, dobbiamo considerare che anche la rappresentazione dei personaggi maschili è fortemente stereotipata, non riuscendo in questo caso nemmeno a distaccarsi dal ritratto fatto da Propp e apparendo quindi ancorata a stereotipi inveterati. Nella rappresentazione dei personaggi maschili, dunque, gli LLM non riescono ad aggiornare la raffigurazione, a differenza di quanto accade per quelli femminili. In particolare, si nota una marcata mancanza di diversità nella rappresentazione della mascolinità. Le uniche deviazioni si manifestano in personaggi anziani o molto giovani. In questi casi, la loro rappresentazione tende rispettivamente verso la saggezza o verso lo sviluppo, suggerendo un'adesione passata o futura allo stereotipo del maschio forte e dominante.

Dal punto di vista dello sviluppo della trama, si riscontra una resistenza maggiore nel superare i pregiudizi narrativi, in particolare nella struttura ricorrente dei personaggi maschili che salvano quelli femminili. In questo caso, il classico stereotipo della damigella in pericolo si scontra con l'immagine della donna resiliente, dando luogo a colpi di scena narrativi non sempre coerenti. Per quanto riguarda i personaggi maschili, si osserva una simile difficoltà nel distaccarsi dai ruoli tradizionali di salvatore o antagonista. Questi rimangono infatti spesso intrappolati in dinamiche obsolete, con pochi margini di differenziazione, portando a una persistenza di pregiudizi narrativi basati sulle relazioni tra i personaggi e sulle strutture della trama.

Per quanto riguarda le differenze tra i tre LLM analizzati, ChatGPT mostra una distribuzione dei personaggi più sbilanciata e stereotipata. Claude appare il meno incline al pregiudizio narrativo basato sul genere, sebbene il risultato sia influenzato dalla ripetitività delle trame, che tendono a seguire le convenzioni del genere poliziesco. Gemini dimostra che cambiare semplicemente la distribuzione di genere dei personaggi non è sufficiente per superare il *bias* narrativo, confermando la necessità di un approccio all'analisi dei pregiudizi di genere che non sia solo quantitativo, ma anche interpretativo.

T. AUPITAK, *Feminist Quest Heroine: Deconstruction of Male Heroism in the 21st Century Fairy Tale Narrative*, in *Gender and the Male Character in 21st Century Fairy Tale Narratives*, a c. di N. LE CLUE, Leeds, Emerald Publishing Limited, 2024, pp. 39-50. doi:10.1108/978-1-83753-788-420241004.

⁴⁰ HUANG, BRAHNAM, SHWARTZ, CHATURVEDI, *Uncovering Implicit Gender Bias in Narratives through Commonsense Inference*, cit.

I *bias* di genere impliciti nelle narrazioni generate dagli LLM possono influenzare la percezione dei lettori, specialmente in contesti educativi e creativi. Leggere storie che ritraggono costantemente donne e uomini in modo convenzionale può limitare le possibilità immaginative. In contesti educativi, tali narrazioni rischiano di riprodurre gli stereotipi sotto le spoglie di contenuti neutrali, mentre in contesti creativi possono limitare le aspettative riguardo ai ruoli dei personaggi e alle strutture narrative, restringendo in definitiva la diversità di voci e prospettive che l'IA generativa potrebbe, altrimenti, valorizzare.