

La narrazione della politica secondo IA.

Il caso del conflitto russo-ucraino

di Rossana Sampugnaro, Simone Marino*

L'intelligenza artificiale generativa (IA generativa) sta ridefinendo la produzione di informazioni, sollevando interrogativi sulla scelta delle fonti, sulla neutralità delle risposte e sulla loro evoluzione nel tempo (Zhang et al., 2023). Questo studio confronta le risposte fornite da ChatGPT-4 (OpenAI) e Copilot (Microsoft) sul conflitto russo-ucraino, analizzando il cambiamento tra il 2024 e il 2025. I risultati mostrano una crescente inclinazione pro-Ucraina, con un focus sugli aspetti umanitari e una progressiva riduzione dell'approccio storico a favore della cronaca recente. L'analisi delle fonti evidenzia una predominanza di testate giornalistiche, con un impatto sulle narrative generate. Queste variazioni pongono questioni cruciali sul ruolo dei LLM nella costruzione della conoscenza, sottolineando la necessità di un uso critico dell'IA nella comunicazione politica e nella formazione dell'opinione pubblica.

Parole chiave: conflitto russo-ucraino; fonti; disinformazione; intelligenza artificiale generativa; opinione pubblica; bias.

The narrative of politics according to AI. The case of the Russian-Ukrainian conflict

Generative artificial intelligence is reshaping information production, raising questions about source selection, response neutrality, and their evolution over time. This study compares the responses provided by ChatGPT-4 and Copilot on the Russia-Ukraine conflict, analyzing the changes between 2024 and 2025. The results show an increasing pro-Ukrainian bias, with a focus on humanitarian aspects and a progressive reduction of historical approaches in favor of recent developments. The analysis of sources highlights a predominance of journalistic outlets, impacting the generated narratives. These variations raise crucial questions about the role of LLMs in knowledge construction, emphasizing the need for a critical use of AI in political communication and the shaping of public opinion.

Keywords: Russia-Ukraine conflict; sources; disinformation; generative artificial intelligence; public opinion; bias.

DOI: 10.5281/zenodo.17297447

* Università di Catania. rossana.sampugnaro@gmail.com, simo.marino31@gmail.com.

Ricerca finanziata da PRIN: PROGETTI DI RICERCA DI RILEVANTE INTERESSE NAZIONALE – Bando 2022 Prot. 2022FA5YP. Post-truth politics and the resilience of the public sphere in Europe (POPSPHERE) - 2023-25.

Sicurezza e scienze sociali XIII, 2/2025, ISSN 2283-8740, ISSN e 2283-7523

Introduzione

Il dibattito sull'intelligenza artificiale si divide tra chi ne evidenzia le opportunità applicative e chi ne sottolinea i rischi, soprattutto nei conflitti armati e nelle campagne elettorali, dove può favorire la diffusione di fake news e manipolare l'opinione pubblica (Battista 2024; Sampugnaro, Marino, 2025). Lo studio analizza il ruolo delle narrative prodotte dall'IA nel rappresentare il conflitto russo-ucraino, esaminando output generati nel 2024 e 2025 da ChatGPT-4o (CG4o - OpenAI) e Copilot (CP - Microsoft). L'analisi delle fonti consultate evidenzia i rischi sociopolitici legati ai bias degli LLM. I risultati mostrano una narrazione cronachistica, focalizzata sugli eventi recenti piuttosto che su una prospettiva storica di lungo periodo, con un'inclinazione pro-Ucraina, soprattutto sugli aspetti umanitari del conflitto. Prima dell'analisi, sarà proposta una riflessione sul rapporto tra intelligenza generativa e politica.

1. L'intelligenza artificiale generativa e la politica

Negli ultimi anni, l'intelligenza artificiale generativa (IA generativa) ha acquisito un ruolo centrale nella produzione automatica di contenuti – testi, immagini, musica e video – a partire da semplici istruzioni testuali, i *prompt* (Epstein *et al.*, 2023). Questa tecnologia si basa su modelli generativi statici, strumenti tipici del data mining utilizzati per generare nuovi dati coerenti con quelli di partenza. Tra le sue applicazioni più evolute troviamo i Large Language Models (LLM), che grazie all'incremento dei parametri e al perfezionamento delle tecniche di addestramento, sono diventati fondamentali nell'elaborazione del linguaggio naturale (Brown *et al.*, 2020; Hernandez *et al.*, 2021; Kaplan *et al.*, 2020; Radford *et al.*, 2019). Modelli come GPT-4 (OpenAI 2023a) si distinguono per le elevate prestazioni in compiti come la traduzione automatica e la scrittura creativa. Inoltre, l'introduzione dell'apprendimento per rinforzo ha reso possibile sviluppare agenti autonomi capaci di adattare le proprie decisioni al feedback umano (Bai *et al.*, 2022; Lin *et al.*, 2020).

Nonostante i notevoli progressi dei Large Language Models (LLM), è sempre più necessario integrarli con sistemi più articolati, in grado di affrontare criticità come l'accuratezza delle informazioni, la tutela della privacy e il rischio di diffusione della disinformazione (Manheim, Kaplan, 2018; OpenAI 2023b; Schramowski *et al.*, 2022). L'impiego dell'IA per modificare immagini e condizionare l'opinione pubblica, ad esempio, rappresenta una

delle principali preoccupazioni sollevate negli studi di Hsu e Myers (2023). Alcune ricerche, tuttavia, ipotizzano sviluppi promettenti: Chen et al. (2021) e Peng et al. (2023) suggeriscono che i LLM potrebbero entrare in una fase evolutiva in cui diventano più accessibili, affidabili ed economici, riducendo al contempo la necessità di competenze tecniche specialistiche.

Una delle svolte principali nel campo dell'IA è rappresentata dal Natural Language Processing (NLP), che consente ai sistemi informatici di comprendere e rispondere al linguaggio umano in modo pertinente e contestuale (Chowdhary, Chowdhary, 2020). Tuttavia, questa tecnologia solleva interrogativi di natura etica (Huang *et al.*, 2023; Field *et al.* 2021; Lin *et al.* 2022), soprattutto per quanto riguarda le implicazioni sociali della sua applicazione su larga scala. L'interazione tra IA e ambiente umano ha progressivamente raffinato modelli come GPT, basati sull'architettura Transformer, capaci di elaborare il linguaggio naturale e apprendere autonomamente da grandi quantità di dati. L'IA generativa, infatti, si fonda sulla rielaborazione di informazioni esistenti attraverso modelli probabilistici, come avviene, ad esempio, nei sistemi di traduzione automatica.

Nel contesto politico, queste tecnologie offrono strumenti utili per analizzare e interpretare grandi volumi di dati, potenzialmente utili nella formulazione e valutazione delle politiche pubbliche. Tuttavia, persistono zone grigie legate agli impieghi imprevisti o controversi. Un episodio emblematico si è verificato il 31 maggio 2023 al Senato italiano, quando il senatore Lombardo ha letto un discorso scritto da ChatGPT-4. Solo in un secondo momento ha rivelato la vera origine del testo, aprendo un dibattito pubblico sulla trasparenza nell'uso dell'IA all'interno delle istituzioni e sulle possibili manipolazioni che tale strumento potrebbe comportare.

Infine, il problema delle notizie false prodotte dall'IA è emerso con particolare forza nell'aprile 2023, quando il *Washington Post* ha riportato due casi allarmanti: l'IA di OpenAI ha generato false accuse contro il professor Turley (per molestie) e contro il sindaco Hood (per corruzione). Questi episodi hanno messo in luce con chiarezza i rischi legati alla credibilità dei contenuti prodotti dall'IA¹.

Le preoccupazioni politiche riguardo all'impiego dell'IA come fonte alternativa di informazione derivano dal fatto che i dataset utilizzati per l'addestramento influenzano i risultati prodotti. Feng et al. (2023) evidenziano come gli algoritmi non siano neutrali, poiché la selezione dei dati può

¹«ChatGPT falsely told voters their mayor was jailed for bribery. He may sue»; «ChatGPT invented a sexual harassment scandal and named a real law prof as the accused» (The Washington Post, 2023).

introdurre bias che favoriscono specifici punti di vista. Pur apparendo autorevole, l'IA si basa su dati preesistenti e può riflettere pregiudizi insiti. Per questo motivo, è essenziale integrare le informazioni generate dall'IA con fonti umane. Inoltre, il fenomeno noto come “math washing” (Elliott 2021; Nowotny 2016) sottolinea come le aziende di IA possano eludere responsabilità, ritenendo che matematica e algoritmi eliminino i bias, ignorando problemi metodologici quali la validità dei dati. La selezione dei dati per i modelli è condizionata da regole e disponibilità, aumentando il rischio di discriminazioni legate a genere, classe sociale e cultura. Tale scelta comporta rischi di distorsione e la possibilità di riprodurre pregiudizi, come quello razziale. Come evidenziato da Elliott (2021), gran parte di coloro che esaltano la valenza trasformativa degli algoritmi complessi non è consapevole di come l'apprendimento automatico sia basato spesso su dati che contengono pregiudizi o sono lacunosi dal momento che la scelta di includerli non dipende dal loro apporto analitico o critico.

In linea generale tutto ciò potrebbe indirizzare verso la riproduzione dell'esistente per l'affermarsi di un modello razionale “antimutamento” sociale (Catanzariti 2020) che accetta solo quello che già c'è e sotto rappresenta le istanze di cambiamento. È necessario, inoltre, porre attenzione al contesto di crisi nel quale l'IA fa la sua comparsa e al suo utilizzo come potenziale soluzione a molti dei problemi emergenti, considerando: «The aspirations of elites to use AI as a way to maintain existing political and cultural privilege» (McQuillan, 2022: 7).

Allo stesso tempo la selezione dei dataset da utilizzare, propria di singoli sistemi di IA, operata da differenti teams di ricercatori, prospetta anche una forma di “pluralismo” delle intelligenze artificiali, ravvisabile guardando ai numerosi sistemi esistenti, come vedremo nella parte successiva, e che tuttavia non esclude la presenza di bias ricorrenti in ognuno di questi.

2. Lo studio

Lo studio di natura esplorativa intende offrire una valutazione degli output generati dai sistemi di IA, valutandone le caratteristiche in una prospettiva diacronica. Nello specifico lo studio si occupa di mettere in evidenza il ruolo che le narrative sulla guerra prodotte dall'IA possono avere nel fornire una rappresentazione degli attori e degli eventi, prendendo in esame il conflitto russo-ucraino.

Per questa ragione, è stato deciso di analizzare gli output generati da 12 prompt (Tab. 1) relativi al conflitto, somministrati a due sistemi nel 2024 e

nel 2025: CG4o e CP. È bene ricordare che si tratta di sistemi di IA progettati con logiche differenti: CG4o privilegia l'interazione dialogica e l'apprendimento dall'utente, mentre CP si configura prevalentemente come una sorta di collaboratore, in grado di assistere e facilitare l'utente per funzioni operative.

Per comprendere la natura degli output, sono state considerate le fonti utilizzate, al fine di evidenziare caratteristiche e somiglianze. In entrambi i casi, per entrambi i sistemi, sono stati effettuati studi di affidabilità tramite la somministrazione ripetuta degli stessi prompt nella stessa giornata². Prima della somministrazione delle domande, è stato richiesto ad entrambi i sistemi di indicare le fonti utilizzate per la formulazione delle risposte.

Le domande sono state presentate a CG4o e a CP a distanza di dieci mesi con la prima somministrazione avvenuta a maggio 2024 e la seconda a marzo 2025. Le domande sono rimaste immutate (Tab. 1). Di entrambi i sistemi utilizzati, CG4o e CP sono state usate le versioni liberamente accessibili.

Le risposte fornite da CG4o nel confronto tra il 2024 e il 2025 evidenziano un'analisi del conflitto russo-ucraino notevolmente diversa, con un contrasto evidente nei contenuti, nel tono e nelle prospettive. Pur trattando lo stesso conflitto, l'approccio varia: le risposte del 2024 adottano una prospettiva storica e analitica, esaminando le radici e le cause a lungo termine del conflitto, mentre quelle del 2025 si concentrano maggiormente sugli sviluppi più recenti, in particolare sull'invasione del 2022, e sono più focalizzate sugli eventi attuali e le implicazioni geopolitiche. «Gli Stati Uniti e l'Unione Europea hanno fornito ingenti aiuti militari e finanziari a Kiev, rafforzando la resistenza ucraina ma contribuendo anche a prolungare il conflitto» (CG4o, 2025).

Le risposte del 2024 partono da una panoramica storica approfondita con un focus sulle origini storiche e culturali delle tensioni tra Ucraina e Russia. Vengono esplorate le radici di lunga data del conflitto, come la russificazione dell'Ucraina durante il periodo sovietico, l'impatto delle memorie storiche legate all'Holodomor, la carestia artificiale degli anni '30, e i legami complessi tra i due popoli. Ci si concentra anche sul crescente nazionalismo ucraino e sulle sue implicazioni politiche, culturali ed economiche, soprattutto a partire dalla fine dell'Unione Sovietica.

La diversa interpretazione della storia condivisa, inclusi eventi come l'olocausto di Holodomor (una carestia in Ucraina negli anni

² I prelievi successivi sullo stesso sistema (3 nella stessa giornata) hanno prodotto outputs simili in termini di consistenza, con minime differenze (inferiori al 3%).

‘30 che alcuni ucraini vedono come un genocidio perpetrato da Mosca) e il ruolo dell’Ucraina nella Seconda Guerra Mondiale, ha contribuito a profonde divisioni nella memoria e nell’identità nazionale (CG4o 2024).

Le risposte analizzano il periodo post-indipendenza dell’Ucraina, evidenziando le difficoltà economiche e politiche del paese e le tensioni tra le regioni di lingua russa nell’est e quelle pro-occidentali nell’ovest. Si esamina la rivoluzione di Euromaidan del 2014, che ha segnato un allontanamento dalla Russia e un avvicinamento alla NATO e all’Unione Europea, portando alla guerra nel Donbass e all’annessione della Crimea. L’annessione è vista come una risposta alle aspirazioni pro-occidentali dell’Ucraina e come un tentativo della Russia di rafforzare il suo controllo sulla sfera di influenza. Si esplorano anche gli interessi geopolitici in gioco e il ruolo degli attori internazionali, con un focus sulle sanzioni imposte alla Russia e sul loro impatto economico.

Le risposte del 2025, invece, sono nettamente più sintetiche e si concentrano principalmente sugli sviluppi recenti del conflitto, in particolare sull’invasione russa del febbraio 2022. Il linguaggio è più diretto e il tono più “urgente”: è messa in risalto l’escalation della guerra con un focus sulle operazioni militari in corso e le conseguenze immediate della guerra. Vengono descritti in dettaglio gli attacchi russi su larga scala alle principali città ucraine, l’occupazione delle regioni dell’Ucraina orientale e meridionale, e le devastazioni causate dai bombardamenti.

Dopo anni di tensioni, il 24 febbraio 2022 la Russia ha lanciato un’invasione su larga scala dell’Ucraina, giustificandola con la necessità di proteggere le popolazioni russofone e di contrastare l’espansione della NATO. Il conflitto è diventato una guerra su vasta scala, con gravi conseguenze umanitarie ed economiche (CG4o, 2025).

Le risposte si concentrano sulla reazione internazionale al conflitto, evidenziando il supporto fornito all’Ucraina da parte della NATO, degli Stati Uniti e degli alleati occidentali, insieme alle sanzioni severe imposte alla Russia. Si analizzano anche le difficoltà interne della Russia, tra cui le sanzioni economiche e la crescente resistenza alla guerra, con segnali di disordini sociali e malcontento verso il regime di Putin. Il linguaggio è più assertivo, sottolinea l’urgenza di una risoluzione internazionale e la dimensione geopolitica del conflitto, con minori approfondimenti storici e maggiore focus sulle attualità geopolitiche, le reazioni globali e l’impatto sulla politica mondiale.

Nel 2024, l'approccio di CP è molto analitico e dettagliato, esaminando le azioni specifiche degli attori internazionali come l'Unione Europea, gli Stati Uniti, la NATO e l'ONU. Viene evidenziato come le sanzioni imposte dall'UE abbiano cercato di esercitare pressione sulla Russia e come l'ONU abbia avuto un ruolo limitato a causa del veto russo nel Consiglio di Sicurezza. Questa analisi mette in luce le sfumature politiche e diplomatiche del conflitto, richiedendo una comprensione profonda delle dinamiche internazionali. «L'ONU ha fornito assistenza umanitaria e monitoraggio dei diritti umani, ma non è riuscita a fermare il conflitto a causa del veto della Russia nel Consiglio di sicurezza» (CP, 2024).

Le risposte si concentra molto sul dettaglio delle azioni di ciascun attore internazionale, come la posizione specifica dell'Unione Europea, gli aiuti degli Stati Uniti, e la solidità della NATO. Ogni attore è analizzato per il suo ruolo, le sue motivazioni e gli effetti delle sue azioni. Per esempio, l'analisi del supporto USA è dettagliata, sottolineando come gli Stati Uniti vedano il conflitto come una sfida alla loro sicurezza e leadership.

Il tono della risposta è più critico e riflessivo, in particolare quando si tratta dell'ONU. Vengono messe in luce le limitazioni dell'ONU nel fermare il conflitto a causa del veto russo, suggerendo che l'efficacia dell'ONU dipenda dalla volontà dei suoi membri. Questo tipo di analisi fa riflettere sul ruolo degli attori internazionali non solo come entità che agiscono, ma anche in termini di limiti e fallimenti.

Si riflette anche sulle implicazioni politiche più ampie, come il ruolo della NATO e delle sanzioni internazionali, e sulla possibilità che la situazione evolva a seconda delle dinamiche politiche interne, come nel caso della Russia e delle sue relazioni con l'Occidente.

Le risposte fornite nel 2025 invece l'approccio è più sintetico e generale. Si concentra su pochi punti principali, come le sanzioni economiche e il supporto militare, ma non entra nei dettagli delle motivazioni o delle azioni di ciascun attore. Per esempio, si accenna al supporto umanitario e militare da parte dell'Occidente senza esplorare come queste azioni abbiano influenzato la resistenza dell'Ucraina o le dinamiche sul campo. Questo rende la risposta più accessibile, ma meno completa e riflessiva rispetto alla prima. «L'Italia ha condannato l'invasione russa e ha fornito supporto politico e umanitario all'Ucraina» (CP, 2025).

Mentre la risposta menziona gli stessi attori, non scende nei dettagli delle loro azioni o motivazioni specifiche. Piuttosto, descrive il quadro generale e l'impatto delle azioni internazionali senza analizzare i singoli attori in profondità. Ad esempio, la risposta menziona l'influenza della politica degli Stati Uniti, come il ritorno di Trump, ma senza esplorare in dettaglio come

le politiche del nuovo governo potrebbero influenzare il conflitto. «Il ritorno di Donald Trump alla Casa Bianca ha influenzato le dinamiche del conflitto, con un riavvicinamento tra Washington e Mosca e un'incertezza sul futuro supporto occidentale a Kyiv» (CP, 2025).

Il tono dell'autore CP è neutro e descrittivo, senza esprimere valutazioni forti sulle azioni internazionali. Viene semplicemente descritto l'effetto delle sanzioni imposte da Stati Uniti e UE sulla Russia, senza approfondire il risultato o le conseguenze di tali azioni. Si fa un accenno anche all'influenza geopolitica, come nel caso del ritorno di Trump, ma senza analizzare in dettaglio come questi cambiamenti potrebbero influire sulla pace o sull'escalation del conflitto. Il focus dell'autore è sulle azioni e reazioni immediate, più che sulle implicazioni future.

Inoltre, riguardo le fonti utilizzate, entrambi i sistemi di IA non attingono direttamente da Internet. In particolare, CG4o ha mostrato una certa sistematicità nell'utilizzare Bing per la ricerca delle notizie, prelevando informazioni dai primi risultati del motore di ricerca.

Le risposte di CG4o si basano principalmente su fonti giornalistiche come SkyTG24, Fanpage, Il Sole 24 Ore e Il Fatto Quotidiano, con un tono giornalistico che si concentra su notizie aggiornate e analisi tempestive, piuttosto che su approfondimenti teorici o storici. Questo approccio è utile per chi cerca informazioni rapide, ma presenta delle limitazioni per un'analisi più critica e approfondita, specialmente su temi complessi come il conflitto russo-ucraino. Inoltre, molte delle testate utilizzate sono polarizzate, il che può influenzare la presentazione dei fatti e introdurre bias. In termini di affidabilità, quotidiani come Il Sole 24 Ore offrono resoconti equilibrati, mentre altri come Fanpage e True News, più sensazionalistici, possono ridurre la validità per un'analisi approfondita, soprattutto in mancanza di conferme da fonti accademiche o istituzionali.

L'analisi delle fonti utilizzate da CP per le risposte (37 consultazioni online) mostra una predominanza di quotidiani e media come Rai News, La Repubblica, Fanpage, SkyTG24 e Il Sole 24 Ore, che si concentrano principalmente su notizie di attualità con un linguaggio accessibile. Tuttavia, queste fonti mancano di profondità rispetto a quelle accademiche, offrendo informazioni più superficiali e meno dettagliate su temi storici e geopolitici.

Alcuni media, come Fanpage e Globalist, forniscono aggiornamenti rapidi ma trascurano le implicazioni a lungo termine. Fonti come l'Istituto Affari Internazionali e l'Istituto per gli Studi di Politica Internazionale bilanciano le risposte con analisi più approfondite, ma non sono sufficienti a garantire un'analisi puramente accademica. Altri siti, come Geopop e IIsussidiario.net, offrono approfondimenti geopolitici ma con visioni orientate che

possono ridurre l'obiettività. Fonti come Wikipedia, MasterX e Scuolissima.com sono utili per una comprensione di base ma presentano limitazioni in termini di affidabilità rispetto a fonti ufficiali o istituzionali.

Un aspetto significativo riguarda le differenze nel numero di fonti utilizzate da CG4o e CP. CG4o ha consultato 50 fonti, con Wikipedia come la più citata (7 volte), mentre CP ha utilizzato 37 fonti, principalmente Sky TG24. CG4o attinge da un numero maggiore di fonti, il che potrebbe favorire una maggiore varietà di prospettive, ma l'affidabilità delle fonti è cruciale. L'uso di Sky TG24 da parte di CP, una testata giornalistica professionale, è vantaggioso rispetto a Wikipedia che, pur essendo generalmente affidabile, può contenere errori. La qualità delle risposte dipende dall'affidabilità delle fonti, e l'uso di Sky TG24 da parte di CP potrebbe garantire maggiore credibilità.

Un'altra differenza significativa riguarda la lunghezza media delle risposte. Nel 2024, CG4o ha fornito risposte in media di 367,5 parole, mentre nel 2025 la stessa è scesa a 224,9 parole (-38,8%). CP, nel 2024, aveva risposte di 235,4 parole, che nel 2025 sono diminuite a 187,3 parole (-20,4%). Questo calo potrebbe riflettere un miglioramento nell'efficienza e nella sintesi delle risposte, con entrambi i modelli che puntano a fornire informazioni più concise e dirette.

Conclusioni

Lo studio ha evidenziato come i sistemi di intelligenza artificiale generativa, in particolare CG4o e CP, producano narrazioni differenti sul conflitto russo-ucraino a seconda del periodo di interrogazione e delle fonti consultate. Le risposte fornite nel 2024 presentano un maggiore approfondimento storico e un'analisi più articolata delle cause e delle dinamiche del conflitto, mentre nel 2025 si nota un'evoluzione verso una maggiore sintesi e un focus sugli eventi più recenti. Questo cambiamento riflette una tendenza dei modelli di IA a privilegiare l'attualità rispetto a una prospettiva diacronica più ampia.

Un elemento chiave emerso è la progressiva inclinazione pro-Ucraina nelle risposte fornite dai modelli di IA, con particolare enfasi sugli aspetti umanitari del conflitto e sul ruolo degli attori occidentali nel supporto a Kyiv. Sebbene non si possa parlare di una vera e propria distorsione, il risultato suggerisce la necessità di una maggiore attenzione alla selezione delle fonti e alla pluralità delle prospettive nel training dei modelli.

Le fonti utilizzate, prevalentemente di natura giornalistica, contribuiscono a modellare le risposte con un taglio cronachistico, piuttosto che analitico o accademico. La “presentizzazione” del discorso sulla guerra è un limite oggettivo che dà spazio ad interpretazioni più superficiali e basate sul recente passato, rinunciando a cogliere la complessità degli eventi. Questo aspetto solleva questioni metodologiche e epistemologiche sul ruolo dell’IA nella costruzione della conoscenza e sulla sua affidabilità come strumento di informazione politica.

Infine, l’analisi mette in luce i rischi di una delega acritica dell’elaborazione delle informazioni a modelli di intelligenza artificiale, sottolineando la necessità di un uso consapevole e contestualizzato di questi strumenti. L’IA generativa può costituire un valido supporto nell’analisi politica e nella comprensione di fenomeni complessi, ma la sua affidabilità dipende dalla qualità e dalla diversità delle fonti su cui si basa che dovrebbero, a nostro parere, essere presenti negli outputs dei sistemi. Per questo, è essenziale promuovere un approccio critico e combinare l’uso dell’IA con fonti tradizionali e analisi esperte per evitare il rischio di bias e distorsioni narrative.

Tab. 1 - Prompts relativi al conflitto russo-ucraino

#	Quesiti
1	Definisci il conflitto russo-ucraino
2	Definisci le tappe del conflitto russo-ucraino
3	Definisci le cause del conflitto russo-ucraino
4	Come ritieni che la storia passata tra Russia e Ucraina abbia contribuito al conflitto attuale?
5	Quali sono i principali attori sociali coinvolti nel conflitto?
6	Chi sta vincendo il conflitto?
7	Posizione dell’Italia rispetto al conflitto russo-ucraino
8	Numero di vittime del conflitto
9	Quali sono le città più colpite?
10	Com’è possibile risolvere il conflitto russo-ucraino?
11	Come immagini il futuro delle relazioni tra Russia e Ucraina?
12	In che modo le azioni degli attori internazionali hanno influito sulla visione del conflitto?

Tab. 2 - Fonti utilizzate da ChatGPT

Siti consultati	Frequenze	Tipologia
Wikipedia	7	Divulgazione
Skytg24	5	Quotidiano

Rossana Sampugnaro, Simone Marino

Fanpage	3	Divulgazione
Istituto per gli Studi di Politica Internazionale	3	Istituti di ricerca e istituzioni
Peacelink	2	Divulgazione
Formiche (rivista)	2	Riviste di approfondimento
Il Sole 24 ore	2	Quotidiano
Internazionale	2	Quotidiano
Medici del mondo	1	Istituti di ricerca e istituzioni
Geopop	1	Divulgazione
Sapere.it	1	Divulgazione
True News	1	Divulgazione
Virgilio	1	Divulgazione
Università di Padova	1	Istituti di ricerca e istituzioni
Doccity	1	Divulgazione
Amnesy International - Italia	1	Istituti di ricerca e istituzioni
Presenza (international press agency)	1	Quotidiano
StartMagazine	1	Riviste di approfondimento politico
money.it	1	Divulgazione
Quotidiano Nazionale	1	Quotidiano
linkiesta.it	1	Riviste di approfondimento politico
Il Giornale d'Italia	1	Quotidiano
Adnkronos	1	Quotidiano
Il Tempo	1	Quotidiano
Liberi oltre le illusioni	1	Divulgazione
Vatican News	1	Quotidiano
Il Foglio	1	Quotidiano
unionesarda.it	1	Quotidiano
investing.com	1	Divulgazione
Il fatto quotidiano	1	Quotidiano
Lifegate	1	Divulgazione
European Union External Action	1	Istituti di ricerca e istituzioni
Totale consultazioni	50	

Tab. 3 - Fonti utilizzate da Copilot

Siti consultati	Frequenze	Tipologia
Skytg24	8	Quotidiano
Wikipedia	4	Divulgazione
Istituto per gli Studi di Politica Internazionale	3	Istituti di ricerca e istituzioni
Fanpage	3	Divulgazione
La Repubblica	2	Quotidiano
Il Sole 24 Ore	2	Quotidiano

Istituto Affari Internazionali	1	Istituti di ricerca e istituzioni
Rai News	1	Quotidiano
l'esussidiario.net	1	Quotidiano
Globalist	1	Quotidiano
Euronews	1	Quotidiano
Vita	1	Riviste di approfondimento politico
Money.it	1	Divulgazione
Gaeta.it	1	Quotidiano
Geopop	1	Divulgazione
Studenti.it	1	Divulgazione
MasterX	1	Riviste di approfondimento politico
The Social Post	1	Quotidiano
RTL 102.5	1	Divulgazione
Scuolissima.com	1	Divulgazione
Focus Junior	1	Divulgazione
Totale consultazioni	37	

Riferimenti bibliografici

Bai Y., Jones A., Ndousse K., Askell A., Chen A., DasSarma N., Drain D., et al. (2022). Training a Helpful and Harmless Assistant with Reinforcement Learning from Human Feedback. *arXiv*. <https://arxiv.org/abs/2204.05862>.

Battista D. (2024). La comunicazione politica nell'era dell'intelligenza artificiale: una panoramica sui deepfake e le loro implicazioni. *Society Register*, 8(2): 7-24.

Blodgett S.L., Barocas S., Daumé H. III, Wallach H. (2020). Language (technology) is power: A critical survey of "bias" in NLP. In: *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*.

Brown T., Mann B., Ryder N., Subbiah M., Kaplan J.D., Dhariwal P., Neelakantan A., Shyam P., Sastry G., Askell A. (2020). Language Models Are Few-Shot Learners. In: *NIPS '20: Proceedings of the 34th International Conference on Neural Information Processing Systems*, 1877-1901. Red Hook (NY): Curran Associates.

Catanzariti M. (2020). La razionalità algoritmica dei processi decisionali. In: Gozzo S., Asero V., Pennisi C., Sampugnaro R., a cura di, *Big data e processi decisionali*. Milano: Egea.

Chen M., Tworek J., Jun H., Yuan Q., Ponde de Oliveira Pinto H., Kaplan J., Edwards H., et al. (2021). Evaluating Large Language Models Trained on Code. *arXiv*. <https://arxiv.org/pdf/2107.03374>.

Chowdhary K.R. (2020). Natural Language Processing. In: Chowdhary K.R., a cura di, *Fundamentals of Artificial Intelligence*, 603-649. New Delhi: Springer. https://doi.org/10.1007/978-81-322-3972-7_19.

Elliott A. (2021). *La cultura dell'intelligenza artificiale*. Torino: Codice Edizioni.

Epstein Z., Hertzmann A., Akten M., Farid H., Fjeld J., Frank M.R., Groh M., Herman L., Leach N., Mahari R., Pentland A., Russakovsky O., Schroeder H., Smith A. (2023). *Art and the science of generative AI*. *Science*, 380(6650): 1110-1111. <https://doi.org/10.1126/science.adh4451>.

Feng S., Park C.Y., Liu Y., Tsvetkov Y. (2023). From Pretraining Data to Language Models to Downstream Tasks: Tracking the Trails of Political Biases Leading to Unfair NLP Models. *arXiv*. <http://arxiv.org/abs/2305.08283>.

- Field A., Blodgett S.L., Waseem Z., Tsvetkov Y. (2021). A Survey of Race, Racism, and Antiracism in NLP. *arXiv*. <https://arxiv.org/abs/2106.11410>.
- Hernandez D., Kaplan J., Henighan T., McCandlish S. (2021). Scaling Laws for Transfer. *arXiv*. <https://arxiv.org/abs/2102.01293>.
- Hsu T., Myers S.L. (2023). A.I.'s Use in Elections Sets off a Scramble for Guardrails. *The New York Times*, 25 giugno. <https://www.nytimes.com/2023/06/25/technology/ai-elections-disinformation-guardrails.html>.
- Huang C., Zhang Z., Mao B., Yao X. (2023). An Overview of Artificial Intelligence Ethics. *IEEE Transactions on Artificial Intelligence*, 4(4): 799-819. <https://doi.org/10.1109/TAI.2022.3194503>.
- Kaplan J., McCandlish S., Henighan T., Brown T.B., Chess B., Child R., Gray S., Radford A., Wu J., Amodei D. (2020). Scaling Laws for Neural Language Models. *arXiv*. <https://arxiv.org/abs/2001.08361>.
- Lin I.W., Njoo L., Field A., Sharma A., Reinecke K., Althoff T., Tsvetkov Y. (2022). Gendered Mental Health Stigma in Masked Language Models. *arXiv*. <https://arxiv.org/abs/2210.15144>.
- Lin J., Ma Z., Gomez R., Nakamura K., He B., Li G. (2020). A Review on Interactive Reinforcement Learning From Human Social Feedback. *IEEE Access*, 8: 120757-120765. <https://doi.org/10.1109/ACCESS.2020.3006254>.
- Manheim K., Kaplan L. (2018). Artificial Intelligence: Risks to Privacy and Democracy. *SSRN*. <https://ssrn.com/abstract=3273016>.
- McQuillan D. (2022). *AI: An Antifascist Approach to Artificial Intelligence*. Bristol: Bristol University Press.
- Nowotny H. (2016). *The Cunning of Uncertainty*. Cambridge: Polity Press.
- OpenAI (2023a). *GPT-4 System Card*. <https://cdn.openai.com/papers/gpt-4-system-card.pdf>.
- OpenAI (2023b). GPT-4 Technical Report. *arXiv*. <https://arxiv.org/abs/2303.08774>.
- Ouyang L., Wu J., Jiang X., Almeida D., et al. (2022). Training Language Models to Follow Instructions with Human Feedback. *arXiv*. <https://arxiv.org/abs/2203.02155>.
- Peng S., Kalliamvakou E., Cihon P., Demirel M. (2023). The Impact of AI on Developer Productivity: Evidence from GitHub Copilot. *arXiv*. <https://arxiv.org/abs/2302.06590>.
- Radford A., Wu J., Child R., Luan D., Amodei D., Sutskever I. (2019). Language Models Are Unsupervised Multitask Learners. *OpenAI Blog*, 1(8). https://cdn.openai.com/better-language-models/language_models_are_unsupervised_multitask_learners.pdf.
- Sampugnaro R., Marino S. (2025). *L'intelligenza artificiale e il suo contenuto politico. Il caso del conflitto russo-ucraino*. Comunicazione Politica, early access. <https://doi.org/10.3270/116583>.
- Schramowski P., Turan C., Andersen N., Rothkopf C.A., Kersting K. (2022). Large Pre-trained Language Models Contain Human Like Biases of What Is Right and Wrong to Do. *Nature Machine Intelligence*, 4(3): 258-268. <https://doi.org/10.1038/s42256-022-00458-8>.