

AI : LINGUAGGIO E INGANNO COME STRUMENTO TECNICO

a cura di Walter Petese

Dedicato alle mie figlie Carla e Ida



Pasolini ed il suo discorso “Lo scandalo radicale”

Ipotizzando di guardare il mondo attuale da una prospettiva privilegiata come quella di colui che non è compromesso col mondo ed ha ancora la capacità di analisi e di critica disinteressata, potremmo certamente osservare che *il grande inganno parte dal linguaggio; chi è padrone del linguaggio, è padrone del discorso; chi è padrone del discorso, è padrone del mondo, perché il mondo crede a ciò che viene detto e ripetuto, non a ciò che è vero.*

Il mistero del linguaggio umano <https://www.youtube.com/watch?v=oogYrYvCewc>

Linguaggio come strumento tecnico e biologico: fondamenti e analogie

Il linguaggio umano, inteso come logos, va considerato nella sua dimensione originaria di strumento tecnico e biologico. Non è semplicemente un veicolo di comunicazione, né una rappresentazione fedele del reale, ma un mezzo attraverso cui l'uomo estende e organizza le proprie capacità cognitive e pratiche. Carlo Sini sottolinea come, analogamente ai primi strumenti dell'uomo primitivo, il linguaggio emerga dalla necessità di gestire problemi concreti di sopravvivenza, adattamento e cooperazione sociale. Così come il bastone prolunga il braccio e la pietra affilata supplisce alla fragilità dell'unghia, il linguaggio consente all'essere umano di strutturare l'esperienza, renderla condivisibile e agire in maniera coordinata all'interno del mondo.

Questa prospettiva si radica nella filosofia, nell'epistemologia e nella sociologia del linguaggio: il logos non è mai neutro o puramente convenzionale. La sua efficacia deriva dall'interazione tra organismi e ambiente, dalle condizioni corporee, percettive e sociali in cui si sviluppa. Il linguaggio emerge come una tecnica cognitiva adattiva, capace di stabilizzare fenomeni complessi in forme simboliche operabili. Tuttavia, questa stabilizzazione implica inevitabilmente una riduzione della complessità: il linguaggio seleziona, evidenzia e fissa aspetti dell'esperienza, mentre ne oscura altri. Tale operazione non è un difetto, bensì la condizione stessa della conoscenza; senza di essa, l'esperienza rimarrebbe inarticolabile e inaccessibile all'azione.

Per rendere più evidente la natura selettiva e funzionale del linguaggio, si possono considerare analogie con la biologia. Il DNA, ad esempio, non coincide con l'organismo, ma ne rappresenta una codifica selettiva, compressa e funzionale, sufficiente a generare vita. Allo stesso modo, i concetti scientifici — come specie, geni o modelli matematici e computazionali — rappresentano strumenti potentissimi per orientarsi e intervenire sul mondo, ma non esaudiscono la totalità della realtà ontologica. Essi sono storicamente situati, rivedibili e correttivi di precedenti formulazioni, a testimonianza della natura non definitiva del linguaggio umano.

La filosofia del linguaggio e l'epistemologia storica confermano questa visione. Wittgenstein, nel suo periodo maturo, evidenzia come il significato emerga dall'uso, e non dalla corrispondenza con una realtà oggettiva. Kuhn e Lakatos mostrano che i modelli scientifici evolvono, si correggono e vengono sostituiti; Jonas avverte dei limiti etici del potere tecnico senza consapevolezza. In questo quadro, la selettività e la parzialità del linguaggio non solo sono inevitabili, ma costituiscono la condizione necessaria per la costruzione della conoscenza e per l'azione efficace.

Questa comprensione prepara il terreno per l'analisi delle tecnologie contemporanee, in particolare dell'intelligenza artificiale. L'AI non introduce un principio radicalmente nuovo, ma estende in maniera tecnica le dinamiche originarie del logos. Essa eredita la parzialità, la selettività e le potenzialità del linguaggio umano, e con esse i rischi epistemici e operativi. Il linguaggio e le tecnologie ad esso correlate possono dunque essere pensati come mappe adattive: strumenti sufficientemente affidabili per orientare l'azione e sostenere la sopravvivenza, ma mai sovrapponibili al territorio che rappresentano. Riconoscere questa distanza costituisce un esercizio di responsabilità epistemica, fondamentale per guidare un uso critico e consapevole dell'AI.

L'intelligenza artificiale: estensione del linguaggio e criticità operative

L'intelligenza artificiale rappresenta oggi l'evoluzione tecnica del logos umano, estendendo capacità cognitive, di calcolo e di elaborazione del linguaggio, ma non sostituendole. La sua potenza risiede nella capacità di analizzare grandi volumi di dati, identificare pattern complessi, generare simulazioni e produrre elaborati testuali coerenti, rendendo così disponibili strumenti di supporto senza precedenti. Tuttavia, questa potenza è strettamente vincolata alle condizioni in cui l'AI opera e alla natura dei dati e del linguaggio su cui si fonda.

L'uso dell'AI diventa particolarmente efficace e sicuro in contesti a rischio gestibile, dove le decisioni umane restano centrali e la supervisione consente di correggere tempestivamente eventuali errori. In tali situazioni, l'AI agisce come una protesi cognitiva: facilita la scrittura, organizza informazioni, simula scenari complessi e supporta l'apprendimento, senza sostituire la capacità critica e valutativa dell'essere umano. Qui, la tecnologia amplia le potenzialità del linguaggio umano senza sovrapporsi completamente al giudizio e alla responsabilità.

Il quadro cambia radicalmente quando l'AI è chiamata a operare in ambiti ad alto rischio, dove le conseguenze degli errori sono irreversibili o difficili da correggere. In tali contesti, il linguaggio generato dalla macchina non è più uno strumento consultivo, ma guida direttamente l'azione, trasformando ogni elaborazione in un intervento sul mondo reale. L'errore, se presente, può manifestarsi in forma differita, cumulativa o indiretta, rendendo difficile individuare e attribuire la responsabilità. Questa frammentazione della catena decisionale tra sviluppatori, operatori e istituzioni aumenta il rischio etico e operativo, evidenziando come l'AI erediti le fragilità intrinseche del linguaggio umano senza possederne il radicamento biologico e situato.

Per comprendere l'estensione dei rischi e delle responsabilità, si possono considerare casi limite: la gestione di una centrale nucleare, decisioni cliniche in medicina intensiva, scenari militari o sistemi di controllo critico. In questi ambiti, l'AI può assistere nell'analisi dei dati e nella simulazione, ma non può sostituire il giudizio umano, che rimane l'elemento decisivo per valutare priorità, rischi e conseguenze. L'affidamento esclusivo a sistemi linguistici automatizzati in tali contesti esporrebbe a gravi conseguenze, perché la tecnologia, pur avanzata, opera secondo algoritmi privi di comprensione ontologica e morale.

Un altro aspetto critico riguarda l'opacità dell'AI e la difficoltà di comprendere come le sue decisioni vengano prodotte. Nei sistemi complessi, l'elaborazione linguistica e statistica può produrre output convincenti e coerenti, ma senza trasparenza sulle cause sottostanti. Questa caratteristica rende difficile intervenire in caso di errore, limitando la capacità umana di esercitare controllo e responsabilità. Inoltre, la naturale selettività e riduzione della complessità del linguaggio si amplifica in ambito computazionale: ciò che viene rappresentato o valutato dalla macchina non esaurisce mai la ricchezza del reale, ma ne offre solo una mappa operativa parziale.

La lezione epistemologica che emerge è chiara: l'AI funziona quando è un'estensione consapevole del linguaggio umano, supervisionata e guidata, ma diventa critica e rischiosa quando agisce senza un quadro di responsabilità chiaro e in contesti dove

l'errore ha conseguenze gravi o irreversibili. Il linguaggio, umano o artificiale, rimane uno strumento adattivo; la sua efficacia non implica coincidenza ontologica con il reale, e la sua potenza tecnica non sostituisce la capacità morale e critica dell'essere umano.

In sintesi, l'AI può essere compresa come un'estensione contemporanea del logos umano: amplifica capacità cognitive e di elaborazione, ma eredita sia le potenzialità sia le fragilità intrinseche del linguaggio. L'uso responsabile richiede giudizio critico, supervisione continua e consapevolezza dei limiti strutturali, garantendo così che la tecnologia rimanga uno strumento al servizio della conoscenza e dell'azione umana, piuttosto che un decisore autonomo.

Frontiere contemporanee dell'AI: LLM, RAG e RAGEP

Negli ultimi anni, l'intelligenza artificiale ha conosciuto una rapida evoluzione grazie allo sviluppo di modelli di apprendimento automatico (ML, **Machine Learning**) e di grandi modelli linguistici (LLM, **Large Language Models**). Questi sistemi sono progettati per elaborare grandi quantità di dati e generare testi, previsioni o risposte coerenti a partire da esempi precedenti. Tuttavia, i LLM presentano limiti intrinseci: la loro conoscenza deriva esclusivamente dai dati di training, che sono statici, e possono produrre **allucinazioni** — cioè affermazioni convincenti ma false o inesatte.

Per superare questi limiti, è nata la tecnica nota come **Generazione Aumentata dal Recupero (RAG, Retrieval-Augmented Generation)**. Ulteriori spessori di analisi e controllo che si introducono prima delle risposte finali. In pratica, RAG combina la capacità dei modelli linguistici di generare testo con la possibilità di **recuperare informazioni aggiornate da fonti esterne**, come documenti, articoli o database. Tradotto in termini semplici: invece di rispondere solo con ciò che ha già “memorizzato”, il modello può consultare fonti aggiuntive per fornire risposte più accurate e contestualizzate.

L'evoluzione più recente di questa tecnica è rappresentata dal **RAGEP (Retrieval-Augmented Generation with Enhanced Prompting)**. Rispetto alla RAG base, RAGEP introduce un **prompting potenziato**, ovvero strategie avanzate per guidare il modello nell'uso delle informazioni recuperate. Questo significa che il modello non solo legge i dati esterni, ma è anche istruito su **come selezionare, ordinare e integrare le informazioni** in modo più coerente e mirato, riducendo gli errori e aumentando l'affidabilità delle risposte.

Queste innovazioni rappresentano una frontiera importante perché affrontano due problemi fondamentali dei LLM tradizionali:

1. **Allucinazioni e imprecisioni**, generate dalla limitatezza dei dati di training.
2. **Dipendenza da dati statici**, che rendono i modelli rapidamente obsoleti rispetto alle informazioni correnti.

In termini concreti, RAG e RAGEP permettono di creare un “studente virtuale” che consulta continuamente fonti aggiornate, seleziona i dati rilevanti e li integra con la propria conoscenza pregressa, senza sostituire il giudizio umano. Tuttavia, queste tecniche non eliminano completamente i rischi: la qualità delle risposte dipende sempre dalla **qualità e**

dalla pertinenza delle fonti esterne, e il modello resta incapace di comprensione ontologica o di valutazione etica autonoma.

In definitiva, l'evoluzione verso RAG e RAGEP rappresenta un passo decisivo verso **modelli più affidabili e contestualizzati**, ma evidenzia anche l'importanza di un uso guidato, critico e consapevole: l'AI diventa uno strumento potente di supporto, non un sostituto del pensiero umano.

Implicazioni sociali e cognitive dell'AI

L'uso dell'intelligenza artificiale, specialmente nelle sue forme più avanzate come RAG e RAGEP, porta con sé profonde implicazioni sia sul piano sociale sia su quello cognitivo. Sul piano sociale, l'adozione acritica di queste tecnologie rischia di diffondere informazioni non verificate o parziali, favorendo una cultura della delega e diminuendo la capacità collettiva di confronto critico. Quando l'AI diventa l'unica fonte di risposta, la responsabilità decisionale e il dibattito vengono frammentati, con conseguenze sulla qualità delle scelte individuali e collettive. La società rischia così di abituarsi a un flusso continuo di contenuti pronti, senza strumenti per valutarli, interpretarli o contestualizzarli.

Sul piano cognitivo, queste tecnologie influenzano la modalità con cui apprendiamo e risolviamo problemi. Se utilizzate in maniera guidata e consapevole, possono ampliare le capacità di esplorazione concettuale, facilitare la comprensione di sistemi complessi e stimolare ragionamenti più sofisticati. Al contrario, un uso passivo o acritico tende a ridurre la memoria, il pensiero critico e la capacità di analisi autonoma, poiché l'utente si abitua a ricevere risposte pronte senza esercitare il proprio giudizio. In questo senso, l'AI funziona come una mappa adattiva del sapere: potente e utile se integrata con la supervisione umana, potenzialmente limitante se diventa il punto di riferimento esclusivo.

In ultima analisi, la sfida più importante non è tecnologica, ma educativa ed etica: è necessario sviluppare competenze cognitive e sociali che permettano di usare l'AI come strumento di supporto, evitando la delega totale del pensiero e preservando la capacità di valutazione, critica e responsabilità. Solo in questo modo la tecnologia può ampliare la conoscenza e l'azione umana senza sostituirsi al logos, rimanendo un'estensione consapevole delle nostre capacità.

Limiti dei sistemi RAG/RAGEP e i bias intrinseci al linguaggio

Nonostante i progressi rappresentati da RAG e RAGEP, i limiti fondamentali dei sistemi basati su LLM restano. La prima ragione è **la natura stessa del linguaggio**. Ogni parola, frase o costrutto non è mai una copia perfetta della realtà, ma una **mappa selettiva e semplificata** di ciò che descrive. Il linguaggio riduce complessità, generalizza fenomeni e astrae concetti, consentendo la comunicazione e l'azione, ma allo stesso tempo introducendo inevitabilmente ambiguità, approssimazioni e lacune.

Questi limiti si riflettono direttamente nell'AI: anche quando un modello recupera informazioni aggiornate con RAG o RAGEP, **le risposte restano mediate dal linguaggio**

stesso. I dati esterni che il modello consulta sono formulati in parole e concetti umani, portando con sé gli stessi errori, imprecisioni e parzialità. In altre parole, il modello non “vede” il mondo direttamente: interpreta solo rappresentazioni linguistiche.

Un secondo limite riguarda i **bias**: pregiudizi, stereotipi e distorsioni intrinseche al linguaggio e ai dati su cui i modelli sono addestrati. I bias non sono semplicemente errori di programmazione o di training, ma **proprietà strutturali del linguaggio umano**. Ad esempio, termini culturali, storici o sociali riflettono visioni parziali della realtà; ciò significa che un modello AI può riprodurre queste distorsioni, anche in sistemi avanzati come RAGEP. Questi bias non possono essere completamente eliminati perché sono **insiti nella natura stessa del linguaggio e della comunicazione**.

Di conseguenza, anche con tecniche avanzate di recupero e prompting, l'AI resta limitata dalla qualità dei dati disponibili e dalla struttura del linguaggio: le informazioni possono essere più aggiornate o meglio integrate, ma **non diventano mai una verità assoluta o neutra**. Questo ritorna al concetto epistemologico che ogni mappa, per quanto dettagliata, resta una semplificazione del territorio che rappresenta.

In sintesi, i limiti permanenti dei sistemi RAG e RAGEP derivano da due fattori fondamentali:

1. La **parzialità intrinseca del linguaggio**, che non coincide mai completamente con la realtà.
2. I **bias inevitabili**, che appartengono al linguaggio stesso e ai dati umani da cui l'AI apprende.

Questi fattori sottolineano l'importanza di un uso guidato, consapevole e critico dell'AI: essa può essere uno strumento potentissimo di supporto, ma non può sostituire il giudizio umano né eliminare la necessità di responsabilità epistemica.

Controbattere all'obiezione dello scienziato

Uno scienziato, che lavora in biologia, fisica o matematica, potrebbe obiettare che i limiti descritti del linguaggio e dei modelli AI non si applicano al suo campo, perché lì si ottengono **risultati risolutivi** e misurabili: esperimenti, equazioni e calcoli producono risposte definitive, con verità empiriche o dimostrabili. Tuttavia, questo argomento non invalida la critica epistemica al linguaggio e all'AI, ma ne chiarisce i contesti e i limiti.

1. Differenza tra contesto chiuso e mondo aperto

- In matematica e fisica, molti problemi possono essere definiti in modo chiuso, con assiomi, regole e sistemi logici precisi. Questo permette di ottenere risultati “risolutivi” all'interno di quel sistema.
- Nel linguaggio naturale, nella comunicazione sociale o nella gestione di dati complessi del mondo reale, le variabili sono **infinite, interconnesse e spesso non misurabili**. Qui il linguaggio non può produrre verità definitive: può solo organizzare, approssimare e guidare l'azione.

2. Differenza tra formalismo e interpretazione

- Le scienze esatte operano con **simboli rigorosi e formali**, dove ogni passaggio è definito e verificabile.
- L'AI e il linguaggio umano operano con **simboli naturali e concetti ambigui**, dove le stesse parole possono avere significati diversi in contesti diversi. I risultati possono quindi essere coerenti, ma non sempre assolutamente affidabili o applicabili in contesti nuovi.

3. I risultati risolutivi non eliminano la dipendenza dai dati

- Anche nelle scienze esatte, la validità di un esperimento o di un modello dipende dalla **qualità e completezza dei dati e delle misurazioni**. Se un modello AI viene alimentato con dati incompleti o distorti, produrrà output convincenti ma potenzialmente errati.
- Tradotto: il biologo può ottenere una soluzione precisa in laboratorio, ma il linguaggio umano e l'AI si confrontano con la **complessità del mondo reale**, dove i dati sono spesso incompleti, ambigui o in continua evoluzione.

4. Riconoscere differenze senza svalutare la scienza

- Non si sta dicendo che le scienze esatte siano inefficaci, ma che i limiti di linguaggio e AI emergono **quando si esce da contesti formalizzati**.
- La critica riguarda l'uso dell'AI come strumento decisionale autonomo in contesti aperti, complessi e sociali, non il metodo scientifico rigoroso.

Si può controbattere ricordando che i limiti non derivano da incapacità tecniche, ma dalla **natura intrinseca del linguaggio e dall'ampiezza del mondo reale**. Nei sistemi chiusi e formalizzati, come in fisica o matematica, le regole permettono risultati risolutivi; nel linguaggio naturale e nell'AI applicata alla vita reale, questi stessi strumenti restano **mappe approssimative**, sempre soggette a parzialità, ambiguità e bias.

Wittgenstein e Gödel: limiti del linguaggio e della formalizzazione

Le riflessioni sul linguaggio e sulla possibilità di risoluzione totale trovano radici profonde nella filosofia e nella logica matematica. Wittgenstein, nel suo **Tractatus Logico-Philosophicus** e nelle opere successive, sottolinea come il linguaggio abbia dei limiti intrinseci: può descrivere la realtà fino a un certo punto, ma non può catturare ogni sfumatura o complessità del mondo. Le parole funzionano come mappe che semplificano il territorio, e ogni tentativo di estendere il linguaggio oltre i propri confini rischia di generare contraddizioni o incomprensioni.

Allo stesso modo, Kurt Gödel, con i suoi **teoremi di incompletezza**, dimostra che in ogni sistema formale sufficientemente complesso esistono proposizioni vere che non possono essere dimostrate all'interno dello stesso sistema. Questo significa che anche un linguaggio altamente formalizzato, come quello matematico, **non può mai essere completamente autosufficiente o risolutivo**: ci saranno sempre problemi o verità che sfuggono alla formalizzazione.

Collegando queste intuizioni all'AI, emerge una lezione chiara: i sistemi basati sul linguaggio naturale o formalizzato, per quanto avanzati, restano **strumenti limitati**. Possono amplificare il ragionamento umano, simulare scenari complessi e migliorare l'accesso all'informazione, ma non possono sostituire completamente il giudizio critico o ottenere soluzioni assolute in tutti i contesti. In questo senso, l'AI è potente come

estensione del logos umano, ma la sua efficacia resta sempre **condizionata dai limiti intrinseci del linguaggio e dei sistemi formali**.

AI, potere e responsabilità sociale

Non possiamo considerare l'AI come un oggetto neutro: essa è **prodotto di sistemi politici, economici e tecnologici** che ne determinano l'orientamento, le priorità e i limiti. Le decisioni su quali dati raccogliere, quali problemi affrontare e quali risorse implementare non sono casuali: riflettono interessi, strategie e vincoli sociali. L'AI, dunque, **incorpora il potere e la visione di chi la sviluppa**, così come i bias culturali e sociali che esistono nel contesto in cui viene progettata.

Per un giovane che si affida all'AI senza basi critiche, questo significa che le risposte non sono solo mediate dal linguaggio e dai dati, ma anche da **scelte di chi controlla la tecnologia**. Le informazioni apparentemente neutre possono essere influenzate da priorità economiche, politiche o commerciali, creando un rischio di manipolazione o di percezione distorta della realtà.

Inoltre, l'AI non distribuisce il potere in maniera uniforme: chi ha accesso alle risorse, ai dati e alle competenze tecnologiche possiede un vantaggio significativo. Questo accentua le disuguaglianze e rende la **consapevolezza critica ancora più necessaria**. Un uso educativo e guidato dell'AI deve quindi includere la capacità di riconoscere questi fattori: sapere **chi ha costruito il sistema, quali interessi può rappresentare e quali limiti strutturali possiede**.

In sintesi, l'AI non è solo uno strumento tecnico: è anche un prodotto sociale e politico. Usarla responsabilmente significa non solo valutare la qualità delle informazioni e i limiti del linguaggio, ma anche comprendere il **contesto di potere** in cui essa esiste e opera. Questo insegnamento è cruciale per formare cittadini consapevoli, capaci di usare la tecnologia senza subire passivamente la sua influenza.