

19 Settembre 2025

La sottile linea rossa

Di Lorella Gatto



Abstract

Nella società dell'informazione i dati non servono solo a descrivere la realtà, ma determinano il funzionamento dell'intera struttura sociale. In tale contesto, la qualità dei dati incide direttamente sulla qualità dei sistemi e non è più solo un dettaglio tecnico. In che modo la produzione di dati sintetici ad opera delle AI generative e la loro immissione silenziosa nel sistema sta degradando la qualità del sistema stesso? In che modo e in che misura l'intelligenza biologica, umana, può aiutare a garantire la qualità del dato e del sistema? Una sfida non solo tecnologica, ma profondamente culturale in una società che delega sempre più ai modelli generativi.

Indice

- La centralità del dato e la qualità come principio regolativo
- L'errore come limite intrinseco
- Degrado culturale
- L'autofagia dell'AI
- Conclusioni

La centralità del dato e la qualità come principio regolativo

A seconda di quale fonte consultiamo, le misurazioni e le stime variano un po' ma si assestano intorno a valori significativamente simili, secondo i quali i dati che abbiamo oggi a disposizione hanno raggiunto circa i 180 zettabyte. Un zettabyte corrisponde a un miliardo di terabyte, misura che, per i non addetti ai lavori, significa poco, ma che diventa impattante se pensiamo che **l'insieme dei dati generati dall'uomo, dal primo momento della sua storia fino a circa vent'anni fa, si assesta intorno ai 2 zettabyte**.

Si tratta di una crescita esplosiva, trainata dalla trasformazione tecnologica, dalla diffusione dei dispositivi IoT e dalla proliferazione di servizi digitali per i quali **il dato non è soltanto uno strumento descrittivo, ma la materia prima** che determina il funzionamento delle infrastrutture sociali, economiche e relazionali.

I dati alimentano algoritmi di intelligenza artificiale, definiscono profili comportamentali, guidano decisioni pubbliche e private e, nell'era dei "Big Data", sono passati dall'essere una risorsa all'essere parte del tessuto stesso su cui si basa la nostra società dell'informazione.

Appare evidente, quindi, che **stiamo vivendo una rivoluzione quantitativa** nell'uso e nella produzione dei dati. Tuttavia, considerando che l'aumento dei dati è sia causa che conseguenza della trasformazione digitale e che una parte crescente dei contenuti disponibili è generata dall'intelligenza artificiale (AI), **si impone anche un'importante riflessione qualitativa**.

In un ecosistema digitale profondamente interconnesso, la qualità del singolo dato non ha più solo un valore locale, diventa un fattore sistemico.

Ogni informazione immessa nei flussi digitali modella algoritmi, condiziona decisioni, ridefinisce scenari economici, sociali e culturali. Per questo, un degrado della qualità dei dati non resta confinato, viene amplificato da modelli generativi, piattaforme di ricerca e sistemi automatizzati di produzione di contenuti e la questione della qualità si trasforma da problema tecnico a rischio strutturale, con impatti potenzialmente sistemici sull'intero ecosistema informativo.

E se l'importanza del dato, e la capacità di raccogliarlo e produrlo in modo coerente e certo, era già un tema critico nella società tecnologica pre-AI, oggi il tema della qualità diventa cruciale, e la sua salvaguardia è sottoposta a nuovi rischi.

La “sottile linea rossa” rappresenta il delicato confine oltre il quale questa spinta all'adozione acritica dell'AI rischia di compromettere l'intero sistema socio-economico e informativo in cui viviamo. L'articolo analizza tre aspetti critici che, mettendo a rischio la qualità dei dati nell'era dell'AI, possono degradare la qualità dell'intero sistema.

L'errore come limite intrinseco

A differenza dei software tradizionali deterministici, **i modelli di apprendimento automatico**, specialmente le reti neurali e i modelli generativi di grandi dimensioni, **producono risultati su base statistica e probabilistica**. Ciò significa che **un certo margine di errore**, sotto forma di imprecisioni, allucinazioni o bias, **è inevitabilmente insito nei loro output**, perché sono strumenti statistici che forniscono la risposta più probabile o verosimile secondo i dati in loro possesso. Si tratta, quindi, di un limite strutturale, non di un bug accidentale.

Le conseguenze di tali errori possono essere rilevanti quando l'output dell'AI viene impiegato in contesti reali senza un'adeguata supervisione. È nota, ad esempio, la vicenda dell'avvocato Steven Schwartz che, nel 2023, aveva citato sentenze inesistenti nel corso di un procedimento davanti alla Corte distrettuale di New York e quello, più recente (ordinanza del 14 marzo 2025), del Tribunale di Firenze che ha rilevato la citazione di una sentenza inesistente nell'atto difensivo di un avvocato. In entrambi i casi, **i legali avevano usato l'AI per svolgere ricerche giurisprudenziali e avevano incluso nei documenti riferimenti a sentenze che, in realtà, non esistevano**. Nei due esempi citati, l'errore è stato rilevato tempestivamente, ma non sempre questo accade e **l'allucinazione** di informazioni da parte dell'AI, che può “inventare” fatti credibili, rischia di riversarsi nel sistema.

L'AI è parte del problema anche quando non genera l'errore, ma semplicemente lo raccoglie e lo diffonde, rendendone la correzione quasi impossibile. Un esempio emblematico è quello della cosiddetta “microscopia elettronica vegetativa”, concetto inesistente frutto di un errore nella digitalizzazione di due articoli degli anni '50, durante la quale il termine “vegetativa” è stato erroneamente associato a “elettronica”. Decenni dopo, l'espressione è riemersa in diverse pubblicazioni accademiche, complici anche alcune traduzioni imperfette e, ad oggi, secondo Google Scholar, la “microscopia elettronica vegetativa” appare in 22 articoli, alcuni dei quali poi ritrattati o corretti. Sembra, inoltre, che tutti i principali modelli di intelligenza artificiale conoscano (e usino) questa frase sbagliata. Questo perché i sistemi di intelligenza artificiale, che attingono a enormi dataset come CommonCrawl, [possono perpetuare e amplificare errori preesistenti nei dati, che diventano poi difficili da trovare e quasi impossibili da risolvere](#).

L'errore nell'AI è quindi un fattore endemico e l'adozione dell'AI senza adeguate cautele può introdurre errori pervasivi nelle nostre infrastrutture informative e decisionali, a scapito sia della qualità del singolo dato e della correttezza delle

decisioni che ne conseguono, sia, una volta che l'informazione errata circola nel sistema, della qualità complessiva del sistema stesso. **L'errore deve essere riconosciuto e gestito.**

In alcuni contesti, la supervisione affidata a un'altra AI può ridurlo, ma non sempre è sufficiente, perché modelli addestrati sugli stessi dati o che condividono gli stessi bias rischiano di replicarlo o non vederlo.

Per questo l'intelligenza umana non è un elemento accessorio, ma una condizione necessaria per bilanciare i limiti strutturali dei modelli generativi e mantenere affidabili i processi decisionali che ne dipendono, soprattutto lì dove l'output sintetico diventa presupposto per scelte e azioni.

Degrado culturale

Un secondo aspetto critico dipende dal fatto che **un sistema di AI generativa non può creare informazione di qualità superiore a quella contenuta nei suoi dati di addestramento**. La regola del machine learning può essere riassunta nel **principio informatico "garbage in, garbage out" (GIGO)**, secondo cui, se vengono forniti dati errati, incompleti o di scarsa qualità (garbage in), il sistema produrrà inevitabilmente risultati altrettanto errati, inutili o insensati (garbage out), indipendentemente dall'accuratezza del processo.

Ma, anche ammettendo che l'AI venga addestrata su dati di alta qualità, resta il limite fondamentale costituito dal fatto che essa non è in grado di superare quei dati.

Il prof. Luciano Floridi sostiene che la spiegazione di questo fenomeno risiede nel fatto che non si può creare una macchina a moto perpetuo, per cui **qualsiasi apprendimento automatico è buono quanto, o solo una piccola frazione meno buona, dei dati di input e l'output sintetico è buono quanto, o leggermente meno buono, dei dati di partenza**. Moltiplicando questa "piccola frazione" meno buona per ogni iterazione, il sistema complessivo si degrada. Se ciascun ciclo generativo introduce anche una minima degradazione qualitativa (perdita di dettaglio, semplificazione eccessiva, omissione di eccezioni, ecc.), col tempo il sistema si impoverisce e le implicazioni sono rilevanti in un mondo in cui una quota sempre maggiore di contenuti online (articoli, saggi, post, documenti) viene prodotta da AI.

Di recente, il primo ministro svedese ha ammesso in un'intervista di utilizzare regolarmente strumenti di intelligenza artificiale per ottenere pareri preliminari su questioni legate alla gestione del governo. Impensierisce che si sia utilizzato il termine "parere" perché l'AI può analizzare dati, ordinare informazioni, creare (in questo caso) testi o audio, ma non deve formulare pareri, non può esprimere opinioni. È una macchina, fa calcoli. Siamo di fronte ad un degrado molto più subdolo e insidioso rispetto all'errore di cui abbiamo parlato nella sezione precedente. Un errore c'è, è un elemento concreto, in alcuni casi difficile da trovare, ma "esiste" e ne siamo certi. **Questo lento e costante degrado nutrito da un abbassamento qualitativo delle conclusioni, dei ragionamenti, delle relazioni, delle soluzioni, infetta inesorabilmente e in silenzio un sistema, senza necessariamente commettere errori.**

Non possiamo metterci alla ricerca, possiamo solo affidarci alla consapevolezza che questo inquinamento va arginato, che non dobbiamo lasciare che l'algoritmo tragga conclusioni, determini relazioni, scelga soluzioni senza metterne in dubbio il ragionamento, senza farci domande. Solo le domande che si pone un'intelligenza

biologica, l'uomo, possono generare un output più "ricco" dell'input ricevuto.

L'autofagia dell'AI

Il terzo aspetto di questa analisi riguarda il fenomeno definito come **"autofagia" dell'intelligenza artificiale, in altre parole il fatto che i modelli di AI possano nutrirsi dei dati sintetici, prodotti da altre AI**, sia per la loro presenza silenziosa nel tessuto informativo, sia per la scarsità di dati "naturali" disponibili e perché è più facile generare dati artificiali in grandi quantità piuttosto che raccogliere dati reali. È di marzo 2025 la notizia dell'acquisizione della startup di dati sintetici Gretel da parte di Nvidia, con l'obiettivo di potenziare la creazione di dati per l'addestramento dell'IA. Ana-Maria Cretu, ricercatrice post-dottorato all'École Polytechnique Fédérale de Lausanne, ha spiegato che **i dati sintetici possono essere impiegati come dati tabulari (demografici, medici) per risolvere carenze o creare dataset più diversificati, ma soprattutto per aumentare la quantità di dati disponibili per i modelli linguistici di grandi dimensioni (LLM)**. Un articolo pubblicato su Nature nel luglio 2024, però, ha evidenziato come **i modelli linguistici AI potrebbero "collassare" se continuamente affinati con dati generati da altri modelli**. In pratica, [alimentare la macchina solo con output generati artificialmente potrebbe portare a un degrado significativo della qualità](#).

Per preservare la salute dell'ecosistema informativo e tecnologico è importante quindi mantenere sempre una percentuale significativa di dati umani nei set di addestramento, ma l'industria AI sta investendo nella generazione di dati sintetici, più scalabile, meno impegnativa e più accessibile agli sviluppatori più piccoli o con meno risorse.

Senza considerare che la nostra infosfera è invasa da una diffusione spesso silenziosa o nascosta degli output prodotti dall'AI per cui diventa cruciale disporre di strumenti per identificare e filtrare tali contenuti evitando di riutilizzarli inavvertitamente nell'addestramento di nuovi modelli.

Conclusioni

Marshall McLuhan sosteneva che "noi plasmiamo i nostri strumenti e poi i nostri strumenti plasmano noi" e se pensiamo a come i grandi modelli linguistici stiano già iniziando a influenzare il modo in cui scriviamo, comunichiamo e perfino pensiamo, comprendiamo che questo fenomeno è in corso. **Se da un lato l'AI generativa offre opportunità straordinarie in termini di efficienza, automazione e accesso alle informazioni, dall'altro, senza opportune e necessarie attenzioni, rischia di intossicare silenziosamente il nostro ecosistema socio-informativo.**

Per evitare questa deriva, **serve prima di tutto consapevolezza del problema e volontà di difendere la qualità come valore sociale** fondamentale e imprescindibile, perché parlare di degenerazione significa richiamare l'attenzione su una traiettoria potenziale che possiamo e dobbiamo ancora influenzare attivamente. **Serve poi un impegno tecnico**, rivolto a filtrare, validare e certificare i dati, attraverso procedure di audit e standard di qualità sempre più rigorosi. E **serve un impegno culturale e normativo**, perché gli esiti dipendono anche da come la società reagisce e regola i nuovi strumenti.

Per questo è urgente trattare la qualità come una priorità strategica nelle politiche digitali, nella progettazione degli algoritmi e nella cultura della comunicazione.

Questo non significa solo sviluppare metriche affidabili per valutare la qualità degli output e l'impatto che essi hanno sul sistema, integrandole nei processi di audit e certificazione dei sistemi di AI.

Significa soprattutto adottare un approccio sistemico, che consideri la qualità non più come un fattore marginale, ma come una grandezza fondamentale da salvaguardare nell'ecosistema digitale.

La sottile linea rossa tra innovazione benefica e degenerazione del sistema informativo passa dalla nostra capacità di **elevare la qualità a principio regolativo** nella gestione e implementazione dei sistemi di AI e delle infrastrutture digitali.