

La scienza della menzogna: come le notizie false diventano credenze condivise

Pino Fioravanti

Terzo e ultimo articolo della serie. I due articoli precedenti hanno seguito l'inganno dalle strategie incorporate nella morfologia degli organismi fino alla menzogna umana, atto che presuppone un emittente, un destinatario e l'intenzione di indurre una credenza ritenuta falsa. Quel quadro, però, descrive una relazione a due.

In questo articolo l'attenzione si sposta: non più come una mente possa ingannarne un'altra, ma come molte menti, interagendo, possano stabilizzare un errore che nessuna di esse controlla interamente.

Riassunto

Le notizie false non sono un prodotto dell'era digitale, ma Internet ne ha trasformato l'ecologia della diffusione, aumentando velocità, portata, replicabilità e visibilità sociale. Questa rassegna esamina il passaggio dalla falsità prodotta dal singolo alla sua circolazione nelle reti umane e, infine, alla possibile formazione di credenze condivise.

Un riferimento centrale è lo studio di Vosoughi et al. (2018), che su Twitter ha mostrato come le notizie false raggiungano più persone, più rapidamente e attraverso cascate più profonde e ramificate rispetto a quelle vere. Studi successivi precisano tuttavia che maggiore diffusione non significa necessariamente maggiore credibilità, né maggiore prevalenza nell'ecosistema informativo. La condivisione può aumentare l'esposizione e la ripetizione accresce familiarità e verosimiglianza percepita. Fiducia nella fonte, omofilia e conversazione sociale possono favorire la convergenza delle credenze e le emozioni ad alta attivazione possono facilitare la propagazione.

Il falso non prevale per un unico meccanismo, ma per l'interazione tra contenuto, cognizione individuale, dinamiche sociali e struttura delle reti. Internet amplifica meccanismi cognitivi preesistenti senza sostituirli con nuovi processi.

1. Dalla menzogna del singolo alla credenza del gruppo

Sappiamo, dal precedente articolo, che mentire comporta in genere un carico cognitivo maggiore rispetto al dire la verità. Come vedremo, questo svantaggio individuale può tuttavia accompagnarsi a un sorprendente vantaggio nella competizione sociale tra informazioni: in determinati contesti, soprattutto nelle reti digitali, il falso può propagarsi più lontano, più rapidamente e più diffusamente del vero. Anche perché chi diffonde una notizia falsa spesso non mente, ma semplicemente, ci crede.

Queste dinamiche possono essere studiate con metodo scientifico. Una volta immesso nella rete sociale, il messaggio acquista una vita autonoma. Può essere valutato, creduto o respinto, modificato, ritrasmesso e progressivamente incorporato in un sistema di conoscenze e credenze condivise. A questo punto può continuare a circolare anche quando il suo autore è scomparso dalla scena e l'intenzione originaria che lo aveva generato non è più riconoscibile. Sono proprio queste fasi — valutazione, accettazione, trasmissione, trasformazione e consolidamento sociale del falso — a poter essere analizzate sperimentalmente e quantitativamente.

2. Prima e dopo Internet: stessi esseri umani, un'ecologia diversa

Le reti sociali esistono da quando esistono gruppi umani. Prima di Internet le narrazioni viaggiavano attraverso conversazioni, mercati, famiglie, comunità religiose, eserciti, lettere, pamphlet, giornali, radio e televisione. Le grandi società disponevano già di mezzi capaci di diffondere propaganda o false informazioni a milioni di persone.

Nelle reti prevalentemente orali la trasmissione era seriale e il contenuto tendeva a essere ricostruito a ogni passaggio. Bartlett (1932) mostrò quanto il ricordo sia ricostruttivo, e Allport e Postman (1947) come le voci si semplifichino e si deformino passando di bocca in bocca. Più in generale, ricordare seleziona, rinforza e talvolta sopprime aspetti differenti del passato (Hirst et al., 2012).

Internet modifica questo scenario lungo più dimensioni. Velocità, portata e costo della replicazione cambiano di ordini di grandezza: in pochi minuti un messaggio può passare da una cerchia locale a un pubblico internazionale. La copia digitale permette inoltre a testi, immagini e video di attraversare migliaia di passaggi pressoché identici. Mentre una voce tradizionale si trasformava a ogni passaggio, una falsità digitale può conservare intatta la propria forma originaria. Ciò non esclude la trasformazione, che però cambia natura: il contenuto resta fedele, ma viene ritagliato, reinserito in contesti diversi, accompagnato da nuovi titoli e commenti o rielaborato in varianti che circolano accanto all'originale. A questo si aggiunge oggi l'intelligenza artificiale generativa, che riduce drasticamente il costo di produzione di testi, immagini, audio e video falsi ma verosimili.

Scompaiono inoltre molti dei tradizionali colli di bottiglia editoriali. Pubblicare non richiede più il controllo di una redazione o il possesso di un'emittente, mentre compaiono indicatori pubblici — condivisioni, like, follower, visualizzazioni — che trasformano la popolarità in un attributo visibile del messaggio.

Accanto alla condivisione attiva, nell'ecosistema digitale contemporaneo la distribuzione dei contenuti è in parte mediata dall'architettura di piattaforme come Facebook, TikTok o YouTube, dove ciò che un utente incontra non dipende soltanto dalle proprie connessioni sociali, ma anche da criteri che le piattaforme adottano per orientare il coinvolgimento. Poiché alcune caratteristiche associate ai contenuti falsi o polarizzanti — novità, imprevedibilità, capacità di suscitare emozioni ad alta attivazione — coincidono in parte con i segnali che questi sistemi tendono a premiare, è plausibile che la loro visibilità ne risulti favorita, e l'esposizione può quindi avvenire anche senza una condivisione deliberata all'interno della propria rete.

Quanto questo incida davvero, tuttavia, non è chiaro. Durante la campagna per le elezioni presidenziali statunitensi del 2020, a un gruppo di utenti di Facebook e Instagram è stato mostrato per tre mesi un feed in ordine cronologico, anziché ordinato dall'algoritmo. I contenuti a cui erano esposti sono cambiati in modo significativo. Su Facebook, per esempio, sono aumentati quelli provenienti da fonti poco affidabili. Eppure, i loro atteggiamenti politici e il loro livello di conoscenza politica sono rimasti sostanzialmente invariati (Guess et al., 2023).

3. Quando la diffusione diventa misurabile

Per decenni, lo studio della diffusione sociale di narrazioni e informazioni si è basato soprattutto su ricostruzioni retrospettive, osservazioni aneddotiche ed esperimenti condotti su campioni relativamente limitati. La disponibilità delle tracce digitali lasciate dagli utenti delle reti sociali ha reso possibile un cambiamento di scala. Un punto di riferimento è lo studio di Vosoughi et al., del Massachusetts Institute of Technology, pubblicato su Science nel 2018.

Gli autori hanno ricostruito la diffusione su Twitter — oggi X — di notizie e affermazioni sottoposte a *fact-checking* tra il 2006 e il 2017. Il dataset comprendeva circa 126.000 cascate di diffusione, generate da circa 3 milioni di utenti e complessivamente condivise oltre 4,5 milioni di volte. Una cascata rappresenta una catena ininterrotta di retweet originata da un singolo tweet; la stessa notizia poteva quindi dare origine a più cascate indipendenti.

Le notizie erano state classificate come vere, false o miste sulla base delle valutazioni di sei organizzazioni indipendenti di *fact-checking*, concordanti nel 95–98% dei casi. Le notizie erano state suddivise in sette principali aree tematiche: politica, leggende metropolitane, scienza e tecnologia, affari e finanza, terrorismo e guerra, intrattenimento e disastri naturali. La politica costituiva la categoria numericamente maggiore; nella ripartizione delle sole notizie false riportata in Figura 1, rappresenta circa un terzo del totale.

Gli autori hanno descritto la struttura delle cascate attraverso quattro parametri, ai quali si aggiungono velocità e ruolo dei bot.

Profondità. È il numero massimo di passaggi successivi che separano il tweet originario dall'utente più lontano raggiunto dalla catena di retweet. Una profondità di 10 significa quindi che l'informazione ha attraversato dieci successivi passaggi da utente a utente, non che sia stata semplicemente rilanciata dieci volte. La Figura 2 mostra, per ogni profondità (asse orizzontale), la percentuale di cascate che la raggiunge o la supera (asse verticale). Circa il 50% delle cascate false raggiunge una profondità di almeno 10 passaggi (punto F1), mentre la stessa percentuale delle notizie vere raggiunge poco più di 5 passaggi (V1). Analogamente, circa il 10% delle cascate false raggiunge almeno 25 passaggi (F2), contro 15 per quelle vere (V2).

Dimensione. Corrisponde al numero complessivo di utenti unici coinvolti in una cascata. Le cascate delle notizie vere superavano raramente i 1.000 utenti, mentre una piccola frazione di quelle false raggiungeva dimensioni molto maggiori, fino a circa 100.000 utenti. La Figura 3 rende evidente questa differenza: quasi nessuna cascata vera raggiunge circa 2.000 utenti, mentre una quota ancora consistente di quelle false supera quella dimensione (F1).

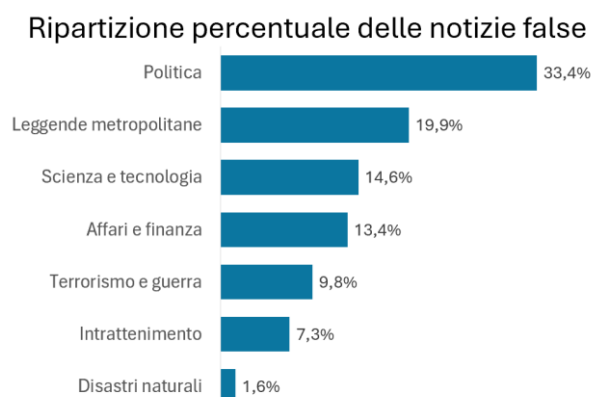


Figura 1. Ripartizione per argomento delle notizie classificate come false.

Rielaborazione grafica a cura dell'autore da Vosoughi et al. (2018).

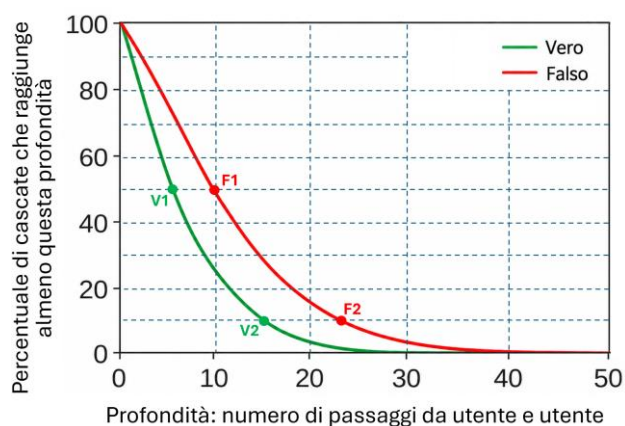


Figura 2. Profili della profondità dei rilanci delle notizie false e vere.

Rielaborazione grafica a cura dell'autore da Vosoughi et al. (2018).

Ampiezza massima. Misura quanti utenti vengono raggiunti a uno stesso livello della cascata, distinguendo quindi una catena profonda ma stretta da una propagazione che, a un determinato passaggio, si ramifica verso molti utenti. Sotto questo profilo le falsità raggiungevano più utenti delle verità, e la differenza aumentava nelle cascate più profonde.

Viralità strutturale. Misura quanto la diffusione dipenda da successive trasmissioni da utente a utente. A un estremo si trova il modello radiotelevisivo, nel quale un unico nodo trasmette a molti utenti, all'altro estremo una cascata ramificata, nella quale ciascun utente contribuisce soltanto a una parte della propagazione. Applicando la metrica proposta da Goel et al. (2016), le notizie false mostravano una maggiore viralità strutturale: non raggiungevano soltanto più persone, ma tendevano maggiormente a propagarsi attraverso successive ramificazioni della rete.

Velocità di propagazione. A queste proprietà strutturali gli autori hanno affiancato la misura della velocità di propagazione. Le notizie vere impiegavano circa sei volte più tempo di quelle false per raggiungere 1.500 utenti e circa venti volte più tempo per raggiungere una profondità di 10.

Ruolo dei bot. Gli autori hanno ripetuto le analisi escludendo gli account automatizzati (bot). I risultati sono rimasti sostanzialmente invariati. I bot acceleravano la circolazione sia del vero sia del falso in misura simile e non spiegavano quindi la differenza osservata tra i due. Nei modelli statistici una notizia falsa aveva una probabilità di essere ritwittata circa il 70% più alta di una vera, anche tenendo conto di diverse caratteristiche dell'account originario.

Complessivamente, lo studio mostra che, nelle cascate di Twitter analizzate, la falsità possedeva un vantaggio sistematico di diffusione: raggiungeva più persone, penetrava più profondamente nella rete, si ramificava maggiormente e superava determinate soglie in tempi più brevi rispetto alla verità. Questo vantaggio risultava particolarmente evidente per le notizie politiche.

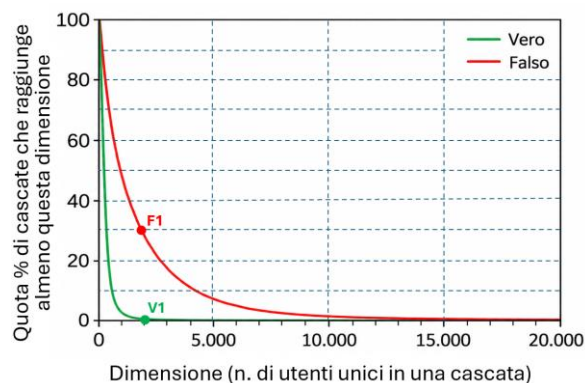


Figura 3. Dimensione espressa come numero di utenti unici coinvolti in una cascata,

Rielaborazione grafica a cura dell'autore da Vosoughi et al. (2018).

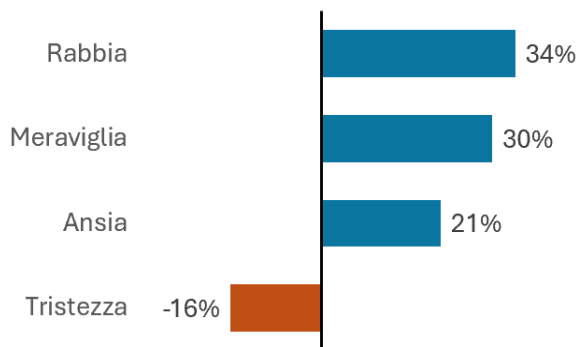


Figura 4. Componenti emotive: variazione % della probabilità degli articoli di essere condivisi per e-mail.

Rielaborazione grafica a cura dell'autore da Berger et al. (2012).

Questi risultati vanno però letti alla luce dei limiti dello studio. Il campione comprende soltanto affermazioni giunte all'attenzione delle organizzazioni di fact-checking e non rappresenta quindi necessariamente l'insieme delle informazioni circolanti. La ricerca, inoltre, misura la diffusione osservabile, non quanto una notizia venga creduta, quanto persuada o quanto venga ricordata nel tempo. Riguarda poi una sola piattaforma, un periodo storico preciso e contenuti prevalentemente in lingua inglese: non dimostra quindi che gli stessi rapporti quantitativi valgano per altre reti sociali o per i mezzi di comunicazione tradizionali. Anche l'interpretazione delle differenze strutturali richiede cautela. Una rianalisi degli stessi dati ha mostrato che, confrontando cascate vere e false di pari dimensione, le differenze di profondità, ampiezza e viralità strutturale si attenuano di molto (Juul et al., 2021). Il falso, in altre parole, non sembra propagarsi secondo una struttura diversa: è più "contagioso", genera cascate più grandi, e la struttura ne è in larga parte una conseguenza. Infine, lo studio descrive con grande precisione come vero e falso si diffondano diversamente, ma consente conclusioni molto più prudenti sul perché.

4. Diffusione, credenza ed esposizione: cosa aggiungono gli studi successivi

I risultati di Vosoughi non sono rimasti isolati, ma vanno letti insieme a studi successivi che ne confermano alcuni aspetti e ne precisano altri.

Pennycook e Rand (2019), in due studi su oltre 3.000 partecipanti, hanno osservato che una maggiore propensione al ragionamento analitico si associa a una migliore capacità di distinguere titoli veri da falsi. Ancora più direttamente, Pennycook et al. (2020) hanno mostrato sperimentalmente che richiamare brevemente l'attenzione dei partecipanti sull'accuratezza di una notizia migliora la capacità di discriminare fra vero e falso nelle successive intenzioni di condivisione.

Questi risultati completano quelli di Vosoughi et al.: la maggiore diffusione del falso non implica necessariamente una sua maggiore credibilità, ma può dipendere anche dal fatto che, nel momento della condivisione, l'attenzione sia catturata da caratteristiche diverse dalla veridicità. Altri studi mostrano però differenze importanti.

Shao et al. (2018), analizzando circa 14 milioni di messaggi contenenti collegamenti a fonti a bassa credibilità, hanno attribuito ai bot un ruolo sproporzionato, concentrato soprattutto nelle prime fasi della diffusione. Ciò sembra contrastare con Vosoughi, secondo cui i bot acceleravano in misura sostanzialmente analoga il vero e il falso. Si tratta, tuttavia, di una divergenza soprattutto metodologica. Vosoughi e collaboratori confrontavano singole notizie verificate come vere o false e chiedevano se l'automazione spiegasse la differenza fra le due, mentre Shao et al. studiavano l'amplificazione di contenuti provenienti da fonti a bassa credibilità. I due risultati possono quindi coesistere: i bot possono contribuire fortemente alla circolazione della disinformazione senza essere necessariamente la causa del suo vantaggio relativo rispetto alle notizie vere.

Particolarmente importante per comprendere la dinamica d'insieme delle notizie false nell'ecosistema informativo è però l'**esposizione complessiva**.

Grinberg et al. (2019) hanno mostrato che, su Twitter, esposizione e condivisione di fonti di *fake news* erano fortemente concentrate in una piccola minoranza di utenti, mentre Allen et al. (2020), considerando congiuntamente consumo via web, dispositivi mobili e televisione, hanno rilevato che le *fake news* rappresentavano una quota relativamente modesta dell'intero consumo informativo.

5. Dal messaggio condiviso alla credenza collettiva

Che una notizia venga condivisa non significa necessariamente che venga creduta. La si può rilanciare perché sorprende, diverte, indigna, conferma l'identità del proprio gruppo o semplicemente perché sembra interessante. Ma la condivisione produce una conseguenza ulteriore: aumenta l'esposizione della comunità al messaggio. È in questo passaggio che una proprietà della rete — la capacità di moltiplicare rapidamente un contenuto — può trasformarsi progressivamente in una proprietà delle menti che la compongono, cioè una maggiore disponibilità ad accettarlo come vero.

Un primo meccanismo è la **familiarità**. L'esposizione ripetuta a un'affermazione tende infatti ad aumentarne la verosimiglianza percepita. L'effetto non implica che la ripetizione renda inevitabilmente credibile qualunque affermazione, ma è robusto.

Pennycook et al. (2018), utilizzando reali titoli di *fake news*, hanno mostrato che anche una sola esposizione può aumentarne successivamente la credibilità percepita, persino a distanza di una settimana. La più recente meta-analisi disponibile, comprendente 182 studi e oltre 31.000 partecipanti, conferma un effetto complessivamente piccolo ma consistente della ripetizione sulla verità percepita, pur con notevole variabilità fra studi (Ye et al., 2026).

La spiegazione più comune chiama in causa la fluidità di elaborazione: un'informazione già incontrata viene riconosciuta e processata più facilmente e questa facilità cognitiva può essere erroneamente interpretata come un indizio della sua verità. La rete digitale aggiunge un elemento cruciale. La stessa affermazione può ricomparire attraverso account, siti, immagini e formulazioni diverse, producendo non solo familiarità, ma anche l'impressione che esistano altre conferme indipendenti.

Familiarità e diffusione possono alimentarsi reciprocamente. Vellani et al. (2023) hanno mostrato che una precedente esposizione rendeva i partecipanti più propensi a condividere nuovamente un'affermazione e che questo effetto era mediato dall'aumento della sua accuratezza percepita, il cosiddetto **effetto della verità illusoria**.

Alla ripetizione si affiancano i **segnali sociali di consenso**: osserviamo quello che fanno gli altri e utilizziamo il loro comportamento come informazione. Questa strategia è normalmente razionale perché nessuno può verificare personalmente tutto ciò che sa e il consenso di più persone indipendenti costituisce spesso un buon indizio della correttezza di un'informazione. Internet rende però il consenso immediatamente visibile attraverso like, condivisioni, visualizzazioni e commenti, senza garantire che chi li ha prodotti sia indipendente, competente o persino umano.

Avram et al. (2020), analizzando il comportamento di oltre 8.500 partecipanti in un ambiente che simulava una rete sociale, hanno osservato che un elevato numero di like e condivisioni era associato a una maggiore probabilità di mettere like o condividere contenuti provenienti da fonti a bassa credibilità e a una minore probabilità di verificarli.

L'effetto del consenso apparente dipende anche dalla **fiducia nella fonte**, legata alla sua competenza, affidabilità e prestigio, reali o percepiti che siano.

Mena et al. (2020), presentando sperimentalmente un falso post di Instagram, hanno osservato che l'approvazione da parte di una figura precedentemente giudicata affidabile aumentava significativamente la credibilità percepita del messaggio, mentre il solo numero elevato di like non produceva nello stesso esperimento un effetto statisticamente significativo.

Questo passaggio è spesso legato all'**identità del gruppo** o a una **credenza collettiva**: le persone non attribuiscono lo stesso peso a tutte le fonti e tendono a fidarsi maggiormente di interlocutori percepiti come appartenenti al proprio ambiente sociale.

In reti sperimentali di dodici persone, Vlasceanu et al. (2020) hanno mostrato che l'esposizione iniziale ad alcune affermazioni da parte di un relatore pubblico ne aumentava l'accessibilità mnemonica; le successive conversazioni tra i membri della rete amplificavano l'effetto fino a produrre una maggiore convergenza delle credenze, ancora rilevabile una settimana dopo. Il gruppo, in altre parole, non si limita a trasmettere ciò che i singoli già pensano, ma attraverso conversazione, selezione dei contenuti e ripetizione può contribuire a sincronizzare ciò che i suoi membri ricordano e ritengono plausibile.

Una rete relativamente chiusa che strutturalmente favorisca l'**omofilia**, cioè la tendenza delle persone a essere maggiormente connesse con individui simili, può aumentare la probabilità che un contenuto venga incontrato ripetutamente all'interno dello stesso ambiente sociale, costituendo quella che viene denominata **camera d'eco**. Essa riguarda utenti connessi prevalentemente tra loro e con fonti concordanti, in cui l'informazione discordante arriva di rado, il che produce omogeneità di esposizione.

In un esperimento controllato, Stein et al. (2023) hanno confrontato reti sociali integrate con reti segregate in gruppi ideologicamente omogenei e hanno osservato che la quota di informazione falsa circolante sembrava maggiore nelle seconde. La segregazione non crea necessariamente la falsità, ma può offrirle nicchie nelle quali messaggi che altrove incontrerebbero resistenza continuano a circolare.

Contiguo alla camera d'eco, e spesso coesistente con essa, è il **circuito di retroazione cognitivo**, che in linea di principio riguarda un singolo contenuto alimentato da una dinamica circolare. Una falsità sufficientemente plausibile viene condivisa; la condivisione ne aumenta l'esposizione e quindi la familiarità; la sua apparente popolarità e l'approvazione di fonti ritenute affidabili forniscono ulteriori segnali sociali, che favoriscono nuove esposizioni concordanti e rinforzano così la narrazione.

La credenza collettiva non nasce solo perché più persone ricevono la stessa falsa informazione, ma perché ciascuna osserva anche che altri la ricevono, la ripetono, la approvano e la incorporano nel discorso comune. La rete non è solo il mezzo attraverso cui viaggia la falsità, ma diventa parte del meccanismo.

Anche quando una notizia falsa viene smentita, la credenza che ha generato raramente scompare del tutto. È il cosiddetto **effetto di influenza continuata**: le persone continuano a basarsi sull'informazione errata nei loro ragionamenti pur ricordando la correzione e dichiarando di accettarla (Ecker et al., 2022). La smentita, infatti, lascia un vuoto nel modello mentale dell'evento, e in assenza di una spiegazione alternativa convincente la versione originale continua a occuparlo. È anche per questo che le notizie false, una volta diffuse, tendono a sedimentarsi come credenze condivise, resistenti alle rettifiche successive.

6. Quando l'emozione spinge a condividere

Le emozioni possono influenzare sia il giudizio di credibilità sia la probabilità che un contenuto venga trasmesso. Ciò non dipende soltanto dalla valenza positiva o negativa dell'emozione, ma anche dal contesto di credenza ed emotività in cui il contenuto viene ricevuto (Martel et al., 2020).

Uno studio del 2012 ha analizzato circa 7.000 articoli del New York Times, mostrando che le emozioni ad alta attivazione aumentano la viralità a prescindere dalla loro valenza. Come mostrato dalla Figura 4, la probabilità stimata di finire tra gli articoli più inoltrati per e-mail cresceva così: +34% per la rabbia, +30% per la meraviglia, +21% per l'ansia. La tristezza, emozione a bassa attivazione, la riduceva del 16% (Berger et al., 2012). Lo studio non riguarda direttamente le notizie false, ma descrive un meccanismo che può contribuire alla diffusione di contenuti falsi ad alta attivazione emotiva e che chi crea tali contenuti può usare deliberatamente. Le percentuali vanno interpretate come misure dell'aumento relativo della probabilità di condivisione, e non come punti percentuali aggiunti alla probabilità assoluta.

Anche l'indignazione morale rientra in questo paradigma: il linguaggio morale-emotivo e quello ostile verso il gruppo avversario aumentano di per sé la probabilità che un contenuto venga ritrasmesso (Brady et al., 2017; Rathje et al., 2021). Il dato che distingue il falso dal vero viene però da McLoughlin et al. (2024), che su dati di Facebook e Twitter e in esperimenti comportamentali hanno osservato che le fonti di disinformazione suscitano più indignazione di quelle affidabili, e che l'indignazione favorisce la condivisione indipendentemente dall'accuratezza percepita. Il meccanismo emotivo non è quindi specifico del falso, ma il falso vi ricorre più spesso.

Nel loro insieme, questi risultati non dimostrano che rabbia, meraviglia, ansia e indignazione siano la causa di una maggiore diffusione delle notizie false. Indicano piuttosto una diversa firma emotiva delle due classi di notizie. La sorpresa è particolarmente plausibile come ponte tra falsità e propagazione, perché le notizie false risultavano anche più nuove e inattese.

7. Perché credere è più economico che controllare

Molti dei meccanismi descritti hanno un denominatore comune: l'asimmetria tra il costo di produrre un'affermazione e quello di verificarla. Valutare l'attendibilità di ciò che ci viene comunicato è una capacità umana antica, ma richiede tempo, attenzione e conoscenze (Sperber et al., 2010). Quando le risorse non bastano, ci si affida a indizi più economici, come la familiarità, il consenso apparente o la fiducia nella fonte. Internet accentua questo squilibrio. La quantità di informazioni disponibili supera di gran lunga la capacità di elaborarle, e questo favorisce i contenuti più facili da notare, ricordare e ritrasmettere, non quelli più accurati (Hills, 2019). L'intelligenza artificiale generativa porta il costo di produzione del falso quasi a zero, mentre quello della verifica resta invariato. Il falso, in questa prospettiva, non prevale perché le persone siano ingenui, ma perché la verifica è diventata relativamente sempre più costosa.

Conclusioni

Le notizie false non possiedono un vantaggio universale né inevitabile sul vero. Nelle reti digitali possono tuttavia beneficiare di caratteristiche — novità, capacità di suscitare emozioni, facilità di condivisione — che ne favoriscono la propagazione. La diffusione, d'altra parte, non equivale alla credenza. È attraverso esposizione ripetuta, familiarità, fiducia nelle fonti, segnali di consenso e struttura delle reti che un messaggio può acquistare progressivamente plausibilità e stabilizzarsi all'interno di un gruppo.

I tre articoli di questa serie hanno considerato l'inganno a tre livelli distinti: tratti morfologici e comportamentali che producono informazione fuorviante, menzogna intenzionale come atto cognitivo di un individuo e circolazione di affermazioni false in una popolazione. Si tratta di livelli che non stanno in un rapporto di derivazione: nessuno dei dati qui discussi indica che la diffusione delle notizie false online sia una continuazione dei meccanismi biologici del mimetismo o che si riduca ai meccanismi cognitivi del singolo.

Ciò che i tre livelli condividono è più limitato, ma verificabile: in tutti e tre, **l'informazione fuorviante persiste quando il destinatario non può verificarla a costi accettabili**. Nelle reti digitali questa condizione non dipende dallo sforzo di chi produce la falsità, ma dalle caratteristiche dell'ambiente in cui essa circola: volume dei contenuti, velocità di replicazione, disponibilità di segnali di popolarità in luogo di segnali di accuratezza, struttura delle connessioni tra gli utenti. Una falsa credenza condivisa non richiede quindi che qualcuno continui a sostenerla: può mantenersi finché l'ambiente informativo rende la verifica più costosa dell'accettazione.

Se il fattore critico è il costo relativo della verifica, gli interventi rilevanti sono quelli che agiscono su di esso — richiami all'accuratezza nel momento della condivisione (Pennycook et al., 2020), indicatori di affidabilità della fonte, minore rilievo dei segnali di consenso non informativi (Avram et al., 2020) — più che quelli rivolti alla motivazione o alla competenza dei singoli utenti. Le prove disponibili su questi interventi sono ancora parziali e i loro effetti, dove misurati, modesti: resta un ambito aperto di ricerca più che un insieme di soluzioni acquisite.

Riferimenti

- Allen, J., Howland, B., Mobius, M., Rothschild, D., & Watts, D. J. (2020). Evaluating the fake news problem at the scale of the information ecosystem. *Science Advances*, 6(14), eaay3539. <https://doi.org/10.1126/sciadv.aay3539>.
- Allport, G. W., & Postman, L. (1947). *The psychology of rumor*. Henry Holt.
- Avram, M., Micallef, N., Patil, S., & Menczer, F. (2020). Exposure to social engagement metrics increases vulnerability to misinformation. *Harvard Kennedy School (HKS) Misinformation Review*. <https://doi.org/10.37016/mr-2020-033>.
- Bartlett, F. C. (1932). *Remembering: A study in experimental and social psychology*. Cambridge University Press.
- Berger, J., & Milkman, K. L. (2012). What makes online content viral? *Journal of Marketing Research*, 49(2), 192–205. <https://doi.org/10.1509/jmr.10.0353>.
- Brady, W. J., Wills, J. A., Jost, J. T., Tucker, J. A., & Van Bavel, J. J. (2017). Emotion shapes the diffusion of moralized content in social networks. *Proceedings of the National Academy of Sciences*, 114(28), 7313–7318. <https://doi.org/10.1073/pnas.1618923114>.
- Ecker, U. K. H., Lewandowsky, S., Cook, J., Schmid, P., Fazio, L. K., Brashier, N., Kendeou, P., Vraga, E. K., & Amazeen, M. A. (2022). The psychological drivers of misinformation belief and its resistance to correction. *Nature Reviews Psychology*, 1(1), 13–29. <https://doi.org/10.1038/s44159-021-00006-y>.
- Goel, S., Anderson, A., Hofman, J., & Watts, D. J. (2016). The structural virality of online diffusion. *Management Science*, 62(1), 180–196. <https://doi.org/10.1287/mnsc.2015.2158>.
- Grinberg, N., Joseph, K., Friedland, L., Swire-Thompson, B., & Lazer, D. (2019). Fake news on Twitter during the 2016 U.S. presidential election. *Science*, 363(6425), 374–378. <https://doi.org/10.1126/science.aau2706>.
- Guess, A. M., Malhotra, N., Pan, J., Barberá, P., Allcott, H., Brown, T., Crespo-Tenorio, A., Dimmery, D., Freelon, D., Gonzalez-Bailon, S., Kennedy, E., Kim, Y. M., Lazer, D., Moehler, D., Nyhan, B., Rivera, C. V., Settle, J., Thomas, D. R., Thorson, E., ... Tucker, J. A. (2023). How do social media feed algorithms affect attitudes and behavior in an election campaign? *Science*, 381(6656), 398–404. <https://doi.org/10.1126/science.abp9364>.
- Hills, T. T. (2019). The dark side of information proliferation. *Perspectives on Psychological Science*, 14(3), 323–330. <https://doi.org/10.1177/1745691618803647>.
- Hirst, W., & Echterhoff, G. (2012). Remembering in conversations: The social sharing and reshaping of memories. *Annual Review of Psychology*, 63, 55–79. <https://doi.org/10.1146/annurev-psych-120710-100340>.
- Juul, J. L., & Ugander, J. (2021). Comparing information diffusion mechanisms by matching on cascade size. *Proceedings of the National Academy of Sciences*, 118(46), e2100786118. <https://doi.org/10.1073/pnas.2100786118>.
- Martel, C., Pennycook, G., & Rand, D. G. (2020). Reliance on emotion promotes belief in fake news. *Cognitive Research: Principles and Implications*, 5(1), 1–20. <https://doi.org/10.1186/s41235-020-00252-3>.
- McLoughlin, K. L., Brady, W. J., Goolsbee, A., Kaiser, B., Klonick, K., & Crockett, M. J. (2024). Misinformation exploits outrage to spread online. *Science*, 386(6725), 991–996. <https://doi.org/10.1126/science.adl2829>.
- Mena, P., Barbe, D., & Chan-Olmsted, S. (2020). Misinformation on Instagram: The impact of trusted endorsements on message credibility. *Social Media + Society*, 6(2), 2056305120935102. <https://doi.org/10.1177/2056305120935102>.
- Pennycook, G., McPhetres, J., Zhang, Y., Lu, J. G., & Rand, D. G. (2020). Fighting COVID-19 misinformation on social media: Experimental evidence for a scalable accuracy-nudge intervention. *Psychological Science*, 31(7), 770–780. <https://doi.org/10.1177/0956797620939054>.
- Pennycook, G., & Rand, D. G. (2019). Lazy, not biased: Susceptibility to partisan fake news is better explained by lack of reasoning than by motivated reasoning. *Cognition*, 188, 39–50. <https://doi.org/10.1016/j.cognition.2018.06.011>.
- Pennycook, G., Cannon, T. D., & Rand, D. G. (2018). Prior exposure increases perceived accuracy of fake news. *Journal of Experimental Psychology: General*, 147(12), 1865–1880. <https://doi.org/10.1037/xge0000465>.
- Rathje, S., Van Bavel, J. J., & van der Linden, S. (2021). Out-group animosity drives engagement on social media. *Proceedings of the National Academy of Sciences*, 118(26), e2024292118. <https://doi.org/10.1073/pnas.2024292118>.
- Shao, C., Ciampaglia, G. L., Varol, O., Yang, K.-C., Flammini, A., & Menczer, F. (2018). The spread of low-credibility content by social bots. *Nature Communications*, 9, 4787. <https://doi.org/10.1038/s41467-018-06930-7>.
- Sperber, D., Clément, F., Heintz, C., Mascaro, O., Mercier, H., Origg, G., & Wilson, D. (2010). Epistemic vigilance. *Mind & Language*, 25(4), 359–393. <https://doi.org/10.1111/j.1468-0017.2010.01394.x>.
- Stein, J., Keuschnigg, M., & van de Rijt, A. (2023). Network segregation and the propagation of misinformation. *Scientific Reports*, 13, 917. <https://doi.org/10.1038/s41598-022-26913-5>.
- Vellani, V., Zheng, S., Ercelik, D., & Sharot, T. (2023). The illusory truth effect leads to the spread of misinformation. *Cognition*, 236, 105421. <https://doi.org/10.1016/j.cognition.2023.105421>.
- Vlasceanu, M., Morais, M. J., Duker, A., & Coman, A. (2020). The synchronization of collective beliefs: From dyadic interactions to network convergence. *Journal of Experimental Psychology: Applied*, 26(3), 453–464. <https://doi.org/10.1037/xap0000265>.
- Vosoughi, S., Roy, D., & Aral, S. (2018). The spread of true and false news online. *Science*, 359(6380), 1146–1151. <https://doi.org/10.1126/science.aap9559>.
- Ye, S., Attali, D., Ghazi, M., Cachia, A., Cassotti, M., & Borst, G. (2026). Systematic review and meta-analysis of the evidence for an illusory truth effect and its determinants. *Nature Communications*, 17, 3270. <https://doi.org/10.1038/s41467-026-70041-x>.