

Fake news, Covid-19 e Infodemia: un esempio di ricerca sociale in real-time su Twitter*

Alessandro Caliandro
Università degli Studi di Pavia**

Guido Anselmi
Università degli Studi di Milano***

Valentina Sturiale
Università degli Studi di Milano-Bicocca****

The spread of Covid-19 virus all over the world has been accompanied by the online circulation of a relevant number of fake news. In this regard, the World Health Organization coined the term infodemic, that is, “an epidemic of misinformation - spreading rapidly through social media platforms and other outlets that poses a serious problem for public health” (Zarocostas, 2020). In this short essay, we explore the processes of circulation of fake news within the Italian Twittersphere during Covid-19 emergency. Drawing upon the analysis of 7,237,581 tweets, based on a digital methods approach, we conclude that it is not possible to observe a real infodemic within the Italian Twittersphere. Only 1,44% of the tweets collected are fake news, plus they circulate only in very specific periods and within closed communities. Furthermore, the 61% of fake news deals with the topic of immigration: false information that put the blame of coronavirus diffusion on migrants do not represent, *per se*, a threat to public health; instead, they are part of the strategy adopted by Italian right-wing populist movements to get online visibility for political purposes. In the conclusion, we also propose a reflection on the characteristics and potentials of *real-time* social research, arguing that it can give a fruitful and authoritative contribution to enhancing the online public debate on contemporary social issues.

Keywords: fake news; Covid 19; infodemia; digital methods; Twitter.

Introduzione

Quello delle fake news non è un fenomeno nuovo, e nemmeno strettamente connesso all'avvento dei media digitali. Tuttavia, è innegabile che con l'avvento di Internet, ed in special modo dei social media, le fake news stiano vivendo una “seconda giovinezza”. I network di comunicazione digitale, infatti, ne amplificano esponenzialmente le potenzialità di diffusione, sia in termini di velocità di circolazione che di numerosità di pubblici raggiungibili.

* Articolo proposto il 26/06/2020. Articolo accettato il 01/09/2020

** alessandro.caliandro@unipv.it

*** guido.anselmi@unimib.it

**** valentina.sturiale@unimib.it

Le fake news sono definibili come informazioni fittizie che, da una parte cercano di imitare i contenuti dei news media mainstream (in termini di format e stile narrativo), ma dall'altra ne ignorano, in maniera più o meno deliberata, le prassi e le norme editoriali (ad esempio, in termini di deontologia, accuratezza delle fonti, credibilità dell'informazione) (Lazer et al. 2018). Le fake news rappresentano una 'perturbazione' nell'ordine dell'informazione giornalistica canonica, in quanto contenuti che mirano a diffondere informazioni false o parziali (Wardle & Derakhshan 2017).

Una comprensione globale del fenomeno delle fake news non può prescindere dallo studio delle *affordances* dei media digitali all'interno dei quali le fake news vengono disseminate. Le fake news non sono semplicemente notizie false ma entità digitali dotate di una particolare capacità di circolazione e ricezione all'interno degli ambienti Internet (Gray et al. 2020). In tale processo di circolazione, dunque, un ruolo centrale viene giocato dalle piattaforme digitali, dalle loro infrastrutture, dalle loro logiche comunicative e commerciali e dalle diverse culture partecipative che in esse germinano (Bounegru et al. 2018). A tal proposito, Rogers (2019) sottolinea come, al di là del loro contenuto (la cui veridicità è facilmente controllabile grazie ai siti di debunking che proliferano sul web), una delle funzioni principali delle fake news sia quella di tenere alta l'attenzione su determinate tematiche, riuscendo così ad influenzare "l'agenda" del discorso pubblico online.

La recente emergenza Covid-19 ha messo in allerta molti analisti dei media digitali riguardo al possibile impatto negativo delle fake news sulla salute pubblica, tanto da indurre l'organizzazione mondiale della sanità a coniare il termine *infodemia* (Zarocostas 2020). Un'infodemia è definibile come un'improvvisa epidemia di disinformazione che si diffonde rapidamente attraverso siti di informazione online e social media, e che pone delle serie minacce alla salute dei cittadini. Si pensi alla falsa credenza secondo cui i cibi piccanti possano prevenire il Covid-19 o alla teoria secondo cui le reti Wi-Fi 5-G possano fungere da acceleratore nella propagazione di questo virus.

In questo breve saggio, attraverso l'analisi di 7.237.581 tweet raccolti seguendo un insieme di chiavi di ricerca legate al tema coronavirus, cerchiamo di delineare un quadro, quanto più possibile esaustivo, della circolazione delle fake news in Italia durante l'emergenza Covid-19, all'interno di uno dei social media maggiormente utilizzati per la diffusione e condivisione di news: Twitter (Kwak et al. 2010). In particolare, l'articolo cerca di rispondere alle seguenti domande di ricerca: *rispetto alla circolazione delle fake news su Twitter in Italia durante l'emergenza Covid-19 è possibile parlare di infodemia? Quali tipi di fake news sono state maggiormente condivise nella Twittersfera italiana durante l'emergenza Covid-19? Su quali tematiche le fake news Twitter legate al Covid-19 cercano di tenere alta l'attenzione? Quali attori o gruppi di attori hanno condiviso maggiormente fake news legate al Covid-19 su Twitter? Con quali scopi?*

A partire dall'analisi dei dati tramite tecniche big data, l'articolo sostiene che, nel caso di Twitter, non si possa parlare di vera e propria infodemia, in quanto le fake news rilevate contano solo per l'1,44% del totale dei tweet, circolano solo in periodi di tempo circoscritti e all'interno di community chiuse. Inoltre, la maggior parte di esse tratta di argomenti legati all'immigrazione, ossia temi che non hanno uno specifico impatto sulla salute pubblica, ma

che si iscrivono nelle più generali strategie dei movimenti populistici di destra di criminalizzare l'immigrazione per scopi politici.

In conclusione, a partire dai risultati della ricerca, il saggio articola una riflessione sulle potenzialità della ricerca sociale in *real-time* ed il suo possibile contributo al dibattito pubblico su questioni di attualità.

Metodologia

Il lavoro empirico si basa su un dataset di circa 7 Milioni di tweet analizzati attraverso un approccio big data ispirato al paradigma dei *digital methods* (Rogers, 2013). Oltre che ad usare tecniche computazionali di raccolta ed analisi dei dati, abbiamo cercato, anzitutto, di seguire il medium (*follow the medium*): abbiamo sfruttato metodologicamente quei *natively digital methods* che gli ambienti digitali già utilizzano per organizzare i propri flussi di dati – ad esempio tramite API, algoritmi, link, Url, time-stamp, like, hashtag, RT, ecc. – al fine di osservare come le affordances dei media digitali strutturano i flussi di comunicazione e le forme di interazione online (Caliandro & Gandini 2017).

Raccolta e tecniche di analisi dei dati

I tweet sono stati raccolti attraverso uno script Python programmato per interrogare la Streaming API di Twitter. Questo tipo di API permette di raccogliere dati in real-time a partire da una o più keyword di ricerca. Nel nostro caso abbiamo usato le seguenti keyword per raccogliere tweet in lingua italiana: #coronavirus, #covid19, #covid-19, #iorestoacasa, #coronavirusitalia, #covid_19, #covid2019, #lombardia, #coronavirusitalia, #COVID19italia, #torniamoliberi, #italiazonarossa, #covid19italia, #zonerosse, COVID19, #coronavirus, #pandemia, #COVID19, #irresponsabili, #lockdownextension, #covid19_italia. Per selezionare le chiavi di cattura ci siamo basati su un secondo script in grado di interrogare le API di Twitter, restituendo l'andamento ora per ora dei trending topic. Abbiamo deciso di catturare, in un primo momento (20-27 febbraio) tutti gli hashtag presenti nei trending topic e contenenti le parole "coronavirus" o "covid" e spelling alternativi. In seguito all'emanazione dei primi decreti nazionali (7-8 marzo) in risposta all'emergenza, abbiamo deciso di includere ulteriori parole chiave che facessero riferimento a iniziative prese da attori politici (#italiazonarossa) o influencer (#irresponsabili). La Streaming API è stata interrogata il giorno 20/02/2020, e ad ora la connessione non è ancora stata interrotta. Al fine dell'analisi discussa in questo contributo abbiamo estratto dal server un dataset di 7.237.581 tweet, che copre un arco temporale che va da dal 20/02/2020 al 02/05/2020 (il giorno dell'inizio della Fase 2). All'interno di questo dataset abbiamo poi identificato ed estratto le fake news attraverso una procedura

automatizzata. Specificamente, abbiamo sviluppato uno script Python ad hoc in grado di: a) estrarre tutti gli Url contenuti nei tweet del dataset, e b) *matchare* questi Url con una lista di siti di fake news. La lista è stata costruita utilizzando gli elenchi di siti di fake news disponibili su Bufale.net e Butac.it. Successivamente, al fine di avere una panoramica più completa delle fake news circolanti nella Twittersfera italiana e, soprattutto, un'idea dell'impatto delle fake news sulle opinioni dei singoli utenti, abbiamo deciso di intervenire anche sul testo dei tweet. Nello specifico, basandoci sugli elenchi di fake news disponibili sul sito del Ministero della Salute (2020), abbiamo individuato le 10 fake più popolari - ossia quelle che hanno avuto più eco sui news media tradizionali e digitali durante il lockdown - che abbiamo poi operativizzato in altrettante query di ricerca: 1) "argento colloidale lo cura"; 2) "è un arma bioterroristica"; 3) "la vitamina c lo cura"; 4) "Montagnier dice che è artificiale"; 5) "è causato da 5g"; 6) "il metodo panzironi lo cura"; 7) "il coronavirus è un piano di soros"; 8) "l'avigan cura il coronavirus"; 9) "il coronavirus è un piano di gates"; 10) "africani sono immuni/irresponsabili/contagiano". Queste query sono poi state rinominate come segue, per facilitare la rappresentazione e lettura dei risultati: 1) Argento Colloidale; 2) Bioterrorismo; 3) Vitamina C; 4) Montagnier; 5) 5G; 6) Metodo Panzironi; 7) Soros; 8) Avigan; 9) Gates; 10) Immigrazione. Infine, sempre utilizzando Python, abbiamo sviluppato un algoritmo di machine learning per discriminare tra tweet pro e contro fake news, ovvero tra quei tweet che avvallano o cercano di falsificare le notizie false.

Allo scopo di identificare gli utenti maggiormente coinvolti nella circolazione delle fake news abbiamo condotto una network analysis basata sugli scambi di retweet - (abbiamo preferito i retweet alle mention in quanto è più probabile che un retweet esprima la condivisione di una notizia rispetto ad una menzione). Sono stati analizzati gli indici di centralità, la forma della rete e le comunità discorsive che si vengono a generare attorno ad una singola fake news (Wasserman & Faust 1994).

Statistiche generali

Il nostro dataset è composto da 7.237.581 tweet (postati tra il 20 Febbraio e il 2 Maggio 2020) da 22.577 utenti singoli. In media, questi utenti postano 13,7 tweet ciascuno, retwittano altri utenti 1,9 volte e li menzionano 1,9 volte. Il numero medio di tweet per utente testimonia di un alto tasso di attività. Tuttavia, bisogna considerare che tale dato risulta distorto dall'alto numero di *news bot* (non necessariamente legati a fake news) presenti nel flusso di tweet (vedi Tabella 1) che postano molto frequentemente – se si considera la mediana, infatti, questa assume valore 2. Il numero medio di retweet e mention indica invece di una scarsa interattività tra gli utenti.

La timeline (Figura 1) mostra come, nonostante fisiologici picchi di traffico il 23/24 Febbraio (il primo focolaio di Covid a Codogno), 9 Marzo (il giorno dopo il Dpcm dell'8 marzo) e 16 Marzo (giorno in cui la regione Lombardia lancia un appello per reclutare medici in pensione), l'attenzione sui temi legati al Covid sia rimasta abbastanza alta e costante per tutto il periodo considerato.

Fake news: analisi dei link e delle keyword

Su 1.103.184 Url presenti nei tweet contenuti nel nostro dataset, solo 3 rimandano a siti collocati in “lista nera” da Bufale.net o Butac.it: questi sono Sputniknews (sito di informazione riconducibile all’apparato di “soft power” russo), Imolaoggi (giornale online che diffonde notizie false, in genere a tema immigrazione) e BlastingNews (aggregatore di contenuti fra cui anche, occasionalmente, fake news). Nessuno di questi indirizzi compare più di 75 volte nei tweet presenti nel dataset. Da questo punto di vista la diffusione di fake news sul coronavirus sembrerebbe essere molto scarsa. Tale risultato potrebbe essere spiegabile in rapporto alla natura *riflessiva* dei social media (Boccia Artieri 2009). È molto probabile infatti che i divulgatori di fake news siano ben consapevoli dell’esistenza di siti ed iniziative di debunking; pertanto è ipotizzabile che essi tendano ad evitare di inserire link nei loro tweet, così da rendersi meno visibili.

Prendendo in considerazione le keyword, invece dei link a siti web, il panorama cambia. Su un totale di 7.237.581 tweet, l’1,44% (104.221) contiene almeno una delle keyword che rimandano alle 10 fake news elencate sopra. Di questi il 34% consiste in messaggi anti fake news. Tutto considerato quindi, il numero delle fake news è decisamente molto basso; tuttavia è significativo notare come, durante il periodo del lockdown, siano circolati 104.221 tweet che hanno commentato, fatto riferimento, o dato visibilità a notizie false. Altro dato interessante viene dalla distribuzione dei tweet segmentata per chiave di ricerca. Come mostra la Figura 2, la categoria “Immigrazione” è quella di gran lunga predominante all’interno dei tweet che circolano fake news (61%) – in questa categoria, la maggioranza dei tweet (77%) è pro fake news. All’interno di questa categoria compaiono una serie di tweet che individuano nei flussi migratori provenienti dall’Africa, così come nei comportamenti ‘incivili’ dei cittadini stranieri presenti sul suolo italiano, la causa principale della diffusione del Covid-19 in Italia.

A #Padova, in piena emergenza sanitaria, #risorseINPS clandestini spacciatori si assemprano in strada per fare a botte. Per loro nessun controllo, nessuna quarantena, nessuna autocertificazione, nessuna mascherina. Sono specie protetta, dalla sinistra. #RadioSavana #CoronaVirus <https://t.co/6dbelpm8Mw>

Mentre gli italiani combattono contro il #coronavirus, i #buonisti pidioti al governo continuano a sbracarsi facendo sbarcare africani, alcuni dei quali con malattie infettive. Mentre il nord fa salti mortali per recuperare posti, al sud curano i #Clandestini! Buffoni!

Ultimo dato interessante riguarda l’andamento temporale della circolazione delle fake news. Come mostra la Figura 1, i picchi di fake news coincidono con i picchi generali di traffico Twitter. Questo dato testimonia di una peculiare strategia di circolazione delle fake news già evidenziata in letteratura. Gli attori interessati a diffondere notizie false tendono ad immetterle nei flussi di comunicazione online in quei momenti in cui l’attenzione del pubblico è più alta (Gray et al. 2020), indipendentemente cioè dal giorno in cui una determinata fake news viene pubblicata sul web. Ad esempio, nel nostro caso, il picco massimo di attenzione su Twitter è con il 16 marzo, giorno che coincide anche con il picco

massimo di fake news. È interessante notare come anche in questo caso la maggior parte delle fake news tenda a concentrarsi nella categoria 'Immigrazione' (Figura 3).

Attori, comunità e diffusione

Chi sono gli utenti più attivi nel diffondere notizie false? Le Tabelle 3 e 4 ci forniscono qualche interessante spunto di riflessione a riguardo. Se si escludono gli utenti @FaberVonCastell, @pietroraffa e @MinisteroSalute, vediamo come il grosso delle fake news postate e retwitteate sembrano circolare sempre tra la stessa tipologia di utenti: sostenitori di ideologie di estrema destra che condividono contenuti anti-immigrazione. È sospetto inoltre il fatto che molti di questi utenti non postino quasi mai messaggi propri ma si limitino solo a retwittare contenuti di altri account, tra l'altro a breve distanza di tempo l'uno dall'altro – (questo è un comportamento tipico dei bot o di quegli utenti che mimano le pratiche dei bot per aumentare la visibilità dei propri profili) (@DFRLab 2018).

Sempre riguardo alla circolazione delle fake news all'interno di specifici gruppi di utenti, un altro dato interessante emerge se andiamo ad isolare le singole community che si aggregano attorno a specifiche notizie. Se consideriamo la rete di retweet che collega gli account che hanno postato con le singole chiavi di cattura possiamo notare come la struttura delle relazioni cambi a seconda del tipo di fake news che si prende in considerazione.

La Figura 4 contiene il grafo dei retweet per quattro parole chiave: "Africani", "Avigan", "Gates", "5g"; i nodi rappresentano gli account, gli archi i retweet e le dimensioni dei nodi corrispondono al numero di retweet ricevuti da ciascun nodo. Per ciascuna delle visualizzazioni i cluster rossi rappresentano il nucleo originario di diffusione delle fake news, mentre i cluster grigi rappresentano altri gruppi di utenti che, commentano lo stesso tema, si pongono in posizione ostile riguardo alle fake news, cercando di falsificarle. È evidente come nella parte sinistra dell'immagine ("Gates" e "Africani") sia quasi completamente assente una posizione neutrale o critica sul tema, mentre invece nel lato destro questa sia piuttosto rilevante. I contenuti dei tweet ed il carattere delle fake news che si sviluppano possono fungere da spiegazione per questa discrepanza. Nello specifico, "Avigan" e "5g" rappresentano un nucleo di notizie false facilmente smentibili, e nei fatti, da più parti smentite nel dibattito pubblico con il supporto di prove scientifiche. In altre parole, esistono fra gli utenti Twitter mezzi e competenze per contrastare la diffusione di questo tipo di notizie false. Questo non accade però per le fake news concentrate nel riquadro destro dell'immagine ("Gates" e "Africani"), poiché rimandano a opinioni più complesse da smentire, come per esempio *"gli africani si comportano male quindi contribuiscono ad aumentare il numero di infett"*. Si tratta, infatti, di argomentazioni generiche, molto più difficilmente assoggettabili a una rapida verifica. Questa caratteristica fa sì, inoltre, che questi tweet rimangano ancora più isolati all'interno delle community nelle quali vengono generati, in quanto è difficile che attorno ad essi si venga a creare un vero e proprio dibattito pubblico.

Discussione

Su 7.237.581 tweet raccolti, solo 75 contengono Url che rimandano a fake news, mentre un residuale 1,44 % tratta di argomenti legati a notizie false nel corpo del testo. Queste fake news, non solo vengono condivise da attori molto specifici (utenti vicini a posizioni di estrema destra), ma tendono anche a circolare in piccole community chiuse su loro stesse. Tali notizie, poi, vengono principalmente immesse nel flusso delle comunicazioni Twitter durante quei giorni in cui il traffico della piattaforma è più alto – perciò indipendentemente dal giorno specifico in cui una data notizia viene pubblicata per la prima volta sul web.

Questi dati ci portano a concludere che, rispetto allo stato delle fake news su Twitter circolate durante il periodo di lockdown, non si possa parlare di una vera e propria *infodemia*. I nostri dati non ci consentono infatti di dare conto di ‘un’epidemia di disinformazione’ che si è diffusa rapidamente (ed esponenzialmente) su Twitter. Anche i dati qualitativi sembrano corroborare questa affermazione. Su 104.221 tweet che commentano fake news, il 61% tratta di argomenti legati all’immigrazione. A nostro avviso, tali messaggi non rappresentano tanto una “minaccia alla salute pubblica”, quanto piuttosto il tentativo, tipico della destra populista italiana, di criminalizzare l’immigrazione a fini politici.

Per quanto interessanti, questi risultati devono essere contestualizzati all’interno della cornice dei limiti del presente studio. In primo luogo, è necessario tenere in considerazione che la base dati è stata acquisita attraverso l’interrogazione della Streaming API di Twitter. Questo ha comportato un’operazione di aggiornamento costante del pacchetto di keyword di ricerca. Per quanto si sia cercato di aggiornare le keyword in maniera sistematica e tempestiva, e con un occhio sempre rivolto agli eventi contingenti, è comunque probabile che qualche keyword importante sia sfuggita. Ciò è dovuto altresì al fatto che un’operazione di questo tipo difficilmente può essere appaltata ad una macchina, che può, al massimo, catturare automaticamente i *trending topic*, ma non keyword significative per specifiche domande e/o obiettivi di ricerca.

In secondo luogo, è importante tener conto del fatto che, per quanto miglior fonte possibile per questo tipo di studio, Twitter potrebbe non essere la piattaforma migliore in assoluto sulla quale studiare il fenomeno delle fake news. Come abbiamo visto, i produttori di fake news sembrano sempre più reticenti a condividere i propri link all’interno di spazi pubblici dove possono essere maggiormente soggetti ad interventi di *debunking*. Per questo motivo è molto probabile che una buona parte delle fake news circoli in ambienti privati e difficilmente accessibili, come gruppi chiusi di Facebook o Telegram. Inoltre, Twitter, proprio durante la pandemia, ha dichiarato che sarebbe intervenuta con decisione al fine di *flaggare* ed eliminare contenuti fake presenti sulla piattaforma (Heilweil 2020). Questo potrebbe aver ulteriormente spinto la migrazione di false notizie su altre piattaforme.

In ultimo, va considerata l’ambiguità del concetto di fake news e i conseguenti problemi di operativizzazione. Nel nostro caso, ad esempio, abbiamo visto come considerare le

fake news, *strictu sensu*, come un prodotto editoriale di una specifica testata online non sia una scelta ottimale. Infatti, l'unico modo per operativizzare una fake news definita in tale senso consiste nel tradurla in forma di Url – entità che, come abbiamo avuto modo di constatare, scarseggiano su Twitter. Per ovviare a tale inconveniente, abbiamo deciso di seguire le keyword associabili a fake news contenute nel testo dei tweet. Questa strategia ha lo svantaggio di fornire una base dati in cui sono contenuti messaggi sia a favore, sia contro le fake news. D'altra parte, però, essa ci ha consentito di monitorare l'impatto delle fake news sul discorso pubblico, fornendo una panoramica di come le retoriche delle fake news vadano ad impattare sulle narrative articolate dai singoli utenti. In effetti, come hanno fatto notare Rauchfleisch e Kaiser (2020) parlando di bot – un altro fenomeno di perturbazione dell'informazione digitale – non è sufficiente contare il numero di bot all'interno di un flusso di conversazioni Twitter, va anzitutto messo in chiaro cosa si intenda per bot, con quali procedure algoritmiche questi vengano identificati e, soprattutto, qual è il loro reale impatto sulle opinioni dei singoli utenti – peraltro, assai difficile da stimare. Un discorso simile può riguardare anche le fake news che dovrebbero essere sempre più studiate non solo quantitativamente, ma anche qualitativamente, misurandone cioè sia la circolazione in senso assoluto che in termini di retoriche.

Conclusioni

Oltre che sul fenomeno delle fake news, questo breve saggio ci consente di articolare qualche riflessione generale sulla ricerca sociale in *real-time*; ossia quel tipo di ricerca che, sfruttando le infrastrutture dei media digitali, consente, in maniera scientifica, di descrivere ed analizzare fenomeni di attualità (Salganik 2018). Così come l'emergenza Covid-19 ha costretto l'accademia a fare seriamente i conti con la didattica a distanza (riscoprendo modalità di insegnamento già disponibili ma poco usate, sperimentandone di nuove e ragionando su possibili scenari futuri), allo stesso modo essa potrebbe indurci a riflettere più sistematicamente sulle potenzialità della ricerca sociale in real-time. In questo articolo, abbiamo cercato di offrire un esempio in tal senso.

Lo studio empirico del fenomeno delle fake news durante il periodo di emergenza Covid-19 non è di per sé un argomento nuovo; negli ultimi mesi sono stati infatti pubblicati diversi articoli di data journalism, o promossi da agenzie di ricerca private (Smith et al. 2020), che si sono occupati della questione. Tuttavia, questo genere di pubblicazioni, in ragione del tipo di format editoriale che deve rispettare, tende a non fornire indicazioni metodologiche rigorose e trasparenti: raramente si fa accenno alle domande di ricerca, alle caratteristiche dei software impiegati per l'estrazione dei dati o al modo in cui specifiche tecniche di analisi vengono implementate. Inoltre, il concetto di fake news è, a volte, trattato acriticamente, ossia senza un'adeguata riflessione sulle sue modalità di operativizzazione e/o sui possibili impatti su opinioni e atteggiamenti delle persone.

Per quanto riguarda la nostra ricerca, abbiamo potuto appurare l'esistenza in una certa circolazione di fake news su Twitter durante il periodo di emergenza legato alla massima diffusione del Covid-19 in Italia, la cui entità però non ci permette di affermare che essa si

sia tramutata in una vera e propria infodemia. Quest'ultimo punto, ci porta a considerare come, grazie alla ricerca in *real-time*, la sociologia possa contribuire, in maniera proficua ed autorevole, al dibattito pubblico su questioni di attualità. In effetti, avallare descrizioni imprecise o parziali dei fenomeni legati alla comunicazione digitale può avere effetti negativi sul tessuto sociale tanto quanto la descrizione parziale o imprecisa di fenomeni sociali generali (come ad esempio l'occupazione o la salute pubblica).

Per quanto riguarda specificamente la questione delle infodemie, i nodi cruciali, a nostro avviso, sono due. In primis, utilizzare il concetto di infodemia con riferimento a situazioni o momenti rispetto ai quali non ci siano prove sufficienti ad accertarne l'effettiva esistenza, porta con sé il paradossale rischio di dare molta più visibilità alla quelle fake news contro le quali si cerca di combattere (Grey et al. 2020). In secundis, contribuire a costruire uno stato di eccessivo allarmismo legato alla diffusione di fake news potrebbe concorrere a rinforzare una certa narrazione tecno-deterministica secondo la quale i social media rappresentano luoghi ontologicamente insicuri. Tale ansia securitaria spinge le piattaforme social (anche su pressione di parte dell'opinione pubblica) ad esercitare un maggiore controllo sulla produzione dei contenuti postati sulle stesse. Abbiamo già accennato a come Twitter sia intervenuta a flaggare ed espungere contenuti fake dalla piattaforma. Per quanto nobile, questo tipo di intervento rende meno trasparente la piattaforma: in che modo Twitter identifica le fake news? Come le definisce? Quali algoritmi usa per discriminare le fake news dai falsi positivi? Queste sono tutte informazioni fondamentali che difficilmente un'azienda privata come Twitter sarà disponibile a rendere pubbliche. Inoltre, tali restrizioni rendono più complesso per i ricercatori studiare le fake news, comprenderne i meccanismi di diffusione e le retoriche principali – le fake news, infatti, continuano comunque a circolare in ambienti digitali chiusi, rimanendo inaccessibili ai ricercatori. Tale condizione ha molti punti di contatto con quanto successo a seguito dello scandalo Cambridge Analytica. Al fine di rendere la piattaforma più sicura e rispettosa della privacy, Facebook ha imposto delle forti restrizioni alle proprie API (Bruns 2019). La prima e principale conseguenza è stata che i ricercatori si sono visti impedito, dall'oggi al domani, l'accesso alla maggior parte dei dati di Facebook che sono, al contrario, disponibili ad aziende private che possono acquistarli tramite servizi ad hoc, ed utilizzarli a fini di business (Venturini & Rogers 2019).

Bibliografia

- @DFRLab (2018) #ElectionWatch: Italy's Self-Made Bots: How the Lega's followers automate themselves, *Medium*, 10 luglio. <https://medium.com/dfrlab/electionwatch-italys-self-made-bots-200e2e268d0e>
- Boccia Artieri, G. (2009). SuperNetwork: quando le vite sono connesse, In L. Mazzoli (ed), *Network effect: quando la rete diventa pop*, Torino: Codice edizioni, 21-40.

- Bounegru, L., Gray, J., Venturini, T., & Mauri, M. (2018). *A Field Guide to 'Fake News' and Other Information Disorders: A Collection of Recipes for Those Who Love to Cook with Digital Methods*. Amsterdam: Public Data Lab.
- Caliandro, A. & Gandini, A. (2017). *Qualitative research in digital environments: A research toolkit*. London: Routledge.
- Daniels, J. (2018). The algorithmic rise of the “alt-right”. *Contexts*, 17(1), 60-65.
- Gray, J., Bounegru, L., & Venturini, T. (2020). ‘Fake news’ as infrastructural uncanny. *New media & Society*, 22(2), 317-341,
- Heilweil, R. (2020). Twitter now labels misleading coronavirus tweets with a misleading label, *Vox*, 10 luglio. https://www.vox.com/recode/2020/5/11/21254889/twitter-coronavirus-covid-misinformation-warnings-labels?_c=1
- Kwak, H., Lee, C., Park, H. and Moon, S. 2010. What is Twitter, a social network or a news media? In: *Proceedings of the 19th International Conference on World Wide Web*. New York: ACM, pp. 591–600.
- Lazer, D. M. et al. (2018). The science of fake news. *Science*, 359(6380), 1094-1096.
- Ministero della Salute (2020). Covid-19 - Attenti alle bufale, *Salute.gov*, 10 luglio. <http://www.salute.gov.it/portale/nuovocoronavirus/dettaglioContenutiNuovoCoronavirus.jsp?lingua=italiano&id=5387&area=nuovoCoronavirus&menu=vuoto>
- Rauchfleisch, A., & Kaiser, J. (2020). The False positive problem of automatic bot detection in social science research, *Berkman Klein Center Research Publication*, 10 luglio. <https://cyber.harvard.edu/news/2020-03/false-positive-problem-automatic-bot-detection-social-science-research>
- Rogers, R. (2013). *Digital Methods*. Cambridge: MIT Press.
- Rogers, R. (2019). *Doing Digital Methods*. London: Sage.
- Salganik, M. J. (2018). *Bit by bit: Social research in a digital age*. Princeton: Princeton University Press.
- Smith, M., McAweeney, E. & Ronzaud, L. (2020). The COVID-19 “Infodemic”: A Preliminary Analysis of the Online Conversation Surrounding the Coronavirus Pandemic, *Graphika*, 10 luglio. https://public-assets.graphika.com/reports/Graphika_Report_Covid19_Infodemic.pdf
- Venturini, T. & Rogers, R. (2019). “API-Based Research” or How can Digital Sociology and Journalism Studies Learn from the Facebook and Cambridge Analytica Data Breach. *Digital Journalism*, 7(4), 532-540.
- Wardle, C. & Derakhshan H. (2017). Information disorder: Toward an interdisciplinary framework for research and policy making, 10 luglio. <https://rm.coe.int/information-disorder-toward-an-interdisciplinary-framework-for-research/168076277c>
- Wasserman, S. & Faust, K. (1994). *Social Network Analysis: Methods and Applications*. Cambridge: Cambridge University Press.
- Zarocostas, J. (2020). How to fight an infodemic. *The Lancet*, 395(10225), 676.

Note Biografiche

Alessandro Caliandro è ricercatore in sociologia dei processi culturali e comunicativi presso il Dipartimento di Scienze Politiche e Sociali dell'Università degli Studi di Pavia. I suoi interessi di ricerca riguardano big data, digital methods, cultura di consumo, tecnologie digitali mobili.

Guido Anselmi è assegnista di ricerca presso il Dipartimento di Scienze Sociali e Politiche dell'Università degli Studi di Milano. Si occupa di piattaforme digitali ed economia politica; in ambito metodologico si occupa di metodi computazionali nella ricerca sociale.

Valentina Sturiale è una analista indipendente esperta in raccolta ed analisi dei dati provenienti dai social media. Attualmente è collaboratrice di ricerca presso il Dipartimento di Sociologia e Ricerca Sociale dell'Università degli Studi di Milano-Bicocca.

Tabelle & Figure

Tabella 1. Top 10 utenti più attivi su tutto il dataset

Nome Utente	N° tweet inviati	Twitter Bio	Descrizione
zazoomblog	23957	Zazoom Social News ... Tutte le notizie dalle migliori fonti del web! Segnala il tuo Blog. Zazoom Social News è un'un'idea che aiuta a dar voce al tuo sito!	News Bot
zazoomnews	22833	Zazoom Info ... rimani aggiornato e segui le ultime news dalle migliori fonti del web!	News Bot
paoloigna1	7888	Libero professionista, analista sereno navigatore nel mondo che cambia per cercare rotte nuove ed insolite. Degusta il tempo sa di non sapere. Consultant at bay	Utente con attività molto intensa, probabilmente usa software per il posting automatico
Adnkronos	7071	Vuoi leggere altre notizie? Vieni sul nostro sito http://adnkronos.com	Account di news con attività molto intensa, probabilmente usa software per il posting automatico
SoniaLaVera	6221	Se ci viene tolta la libertà di parola, noi, muti e silenziosi, saremo trascinati come pecore al macello. (George Washington)	Utente che si limita quasi esclusivamente a retwittare messaggi di altri account (a breve distanza di tempo l'uno dall'altro). Sembra un bot, o comunque si comporta come tale
marino29b	6067	5 ☆☆☆☆☆ come scelta di vita	Utente che si limita quasi esclusivamente a retwittare messaggi di altri account (a breve distanza di tempo l'uno dall'altro). Sembra un bot, o comunque si comporta come tale
pietro30199674	5312	Nessuna Bio	Utente che si limita quasi esclusivamente a retwittare messaggi di altri account (a breve distanza di tempo l'uno dall'altro). Sembra un bot, o comunque si comporta come tale
Frankf1842	5041	Gutta cavat lapidem!	Utente che si limita quasi esclusivamente a retwittare messaggi di altri account (a breve distanza di tempo l'uno dall'altro). Sembra un bot, o comunque si comporta

			<i>come tale</i>
Ultron65	5027	In cammino verso THE ITALEXIT PARTY [nome account A.I.]	Utente che si limita quasi esclusivamente a retwittare messaggi di altri account (a breve distanza di tempo l'uno dall'altro). Sembra un bot, o comunque si comporta come tale
BreakingItalyNe	4889	Riporto news dall'Italia e dal Mondo	Account di news con attività molto intensa, probabilmente usa software per il posting automatico

Tabella 2. Top 10 utenti più attivi nel postare fake news

Nome Utente	N° fake news postate	Twitter Bio	Descrizione
SoniaLaVera	191	Se ci viene tolta la libertà di parola, noi, muti e silenziosi, saremo trascinati come pecore al macello. (George Washington)	Utente che retwitta post pro-destra e anti-immigrazione
pietro30199674	184	Traveler	Utente che retwitta post pro-destra e fake news a sfondo razziale
Bobbio65M	146	Convintissimo dal 2015 su Tw. LN-FDI ma comunque sostenitore, di tutto ciò non sia sx. Un augurio a questa ITALIA!! No LeU, no PD e compagnia bella.	Utente che retwitta molti contenuti di Salvini e Trump
Enki2270	142	Nessuna Bio	Utente che retwitta post di leader di destra (es. Salvini e Trump)
andrea61990928	133	W l'Italia	Utente che retwitta molti contenuti pro Lega e Salvini
domalex82	130	Giornalista professionista. Amante di Conoscenza, Bellezza e Giustizia. Sovranista No Euro, No Ue, No Nato, No immigrazione di massa. Auspico l'estinzione Pd	Utente che retwitta contenuti pro-destra
monikindaaaa	115	Nessuna Bio	Utente che posta contenuti anti-immigrazione
noitre32	112	notizie, noticias ,news dal mondo, curiosidad y muchos mas. #InfluencerMarketing collaboro con #blastingnews y	Utente che posta contenuti anti-immigrazione
RadioSavana	111	Nessuna Bio	Canale che posta contenuti e fake news di stampo razzista
MariMario1	111	No fenomeni, facciamorete, pd, m5s, pesci e stracci blu. Se ti senti troppo intelligente evitami perché già non ti sopporto. Pazienza zero quindi blocco facile.	Utente che retwitta contenuti pro-destra ed anti-immigrazione

Tabella 3. Top 10 utenti più retwittati all'interno del subset fake news

Nome Utente	N° RT ricevuti	Twitter Bio	Descrizione
RadioSavana	52946	Nessuna Bio	Canale che posta contenuti e fake news di stampo razzista
benq_antonio	2656	Nazionalista e "Buon senso "credo e voglio un 'Italia più Sovrana e più libera...' #Italexi	Utente che posta contenuti pro-destra ed anti-immigrazione
francescatotolo	2305	Collaboratrice de #IIPrimatoNazionale. Fieramente italiana e patriota, allergica al neo femminismo. #InfernoSpa https://altafortedizioni.it/prodotto/infernospa/	Utente che scrive per il quotidiano Sovranista Il Primato Nazionale
Lunadargento5	1289	Bio-chetasi	Utente che posta

			<i>contenuti pro-destra ed anti-immigrazione</i>
<i>GiorgiaMeloni</i>	1184	<i>Romana, classe '77. Politica e giornalista. Sempre, ovunque e prima di tutto, italiana. Presidente di @FratellidItalia</i>	<i>Presidente del Partito Fratelli d'Italia. Posta contenuti pro-destra</i>
<i>FaberVonCastell</i>	1148	<i>Tornato dall'India con l'iniziazione al Kriyā Yoga. Ho veduto che πάλτα πέϊ, come in alto così in basso. Quindi, keep calm and live your own life consciously.</i>	<i>Utente che posta contenuti genericamente populist e anti-establishment</i>
<i>pietroraffa</i>	828	<i>Partner/Digital Strategist in MR&Associati. Blogger (anche) su Espresso e HuffPost. Amante della comunicazione, politica e non. Docente (anche) in BS Sole24ore</i>	<i>Giornalista ed esperto di comunicazione. Posta contenuti anti-establishment e antirazzisti</i>
<i>MinisteroSalute</i>	825	<i>Profilo ufficiale del Ministero della Salute</i>	<i>Profilo ufficiale del Ministero della Salute</i>
<i>Gianmar26145917</i>	822	<i>Ho superato i 40, lavoro da più di 20 e vengo da sinistra ma ho svoltato a destra. Sovranista convinto, non sopporto UE, crukki e mangia rane alla Macron!</i>	<i>Utente che retwitta contenuti pro-destra ed anti-immigrazione</i>
<i>lucianocapone</i>	742	<i>Giornalista @ilfoglio.it</i>	<i>Giornalista del giornale Il Foglio</i>

Figura 1. Timeline dei tweet su tutto il dataset (20/02/2020 – 02/05/2020)

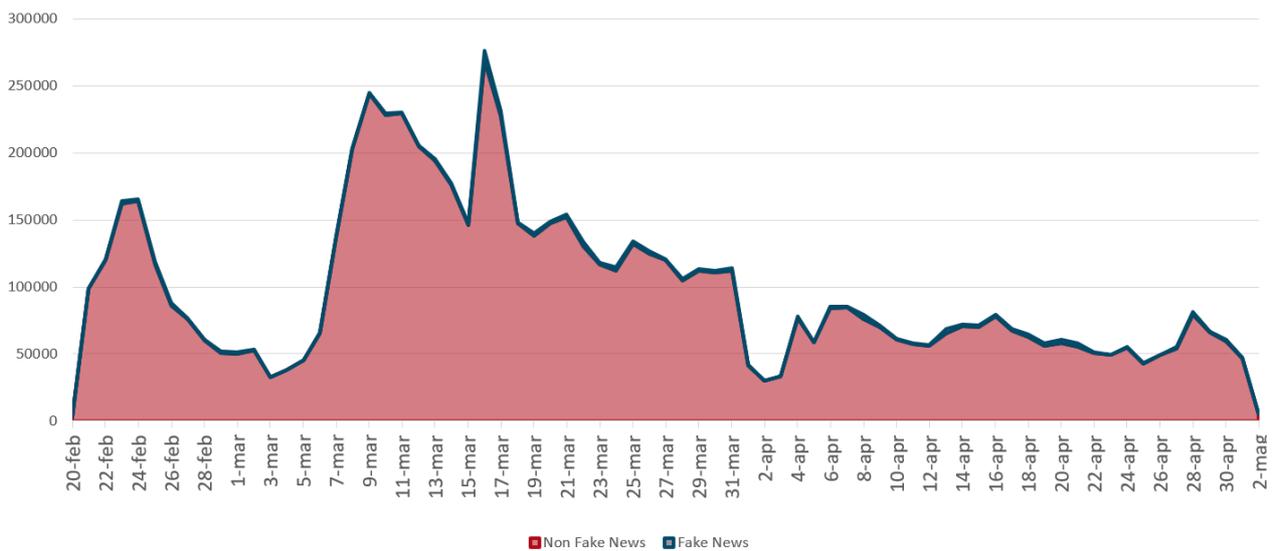
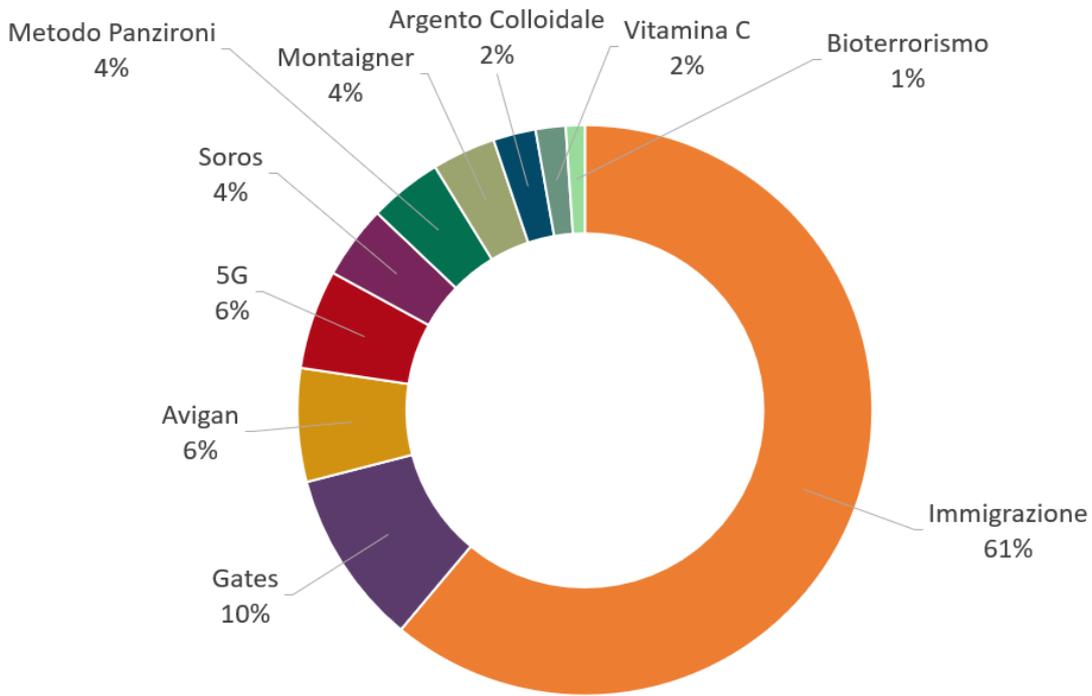


Figura 2. Distribuzione categorie fake news



Base dati: 104.221 tweet

Figura 3. Timeline delle fake news (20/02/2020 – 02/05/2020)

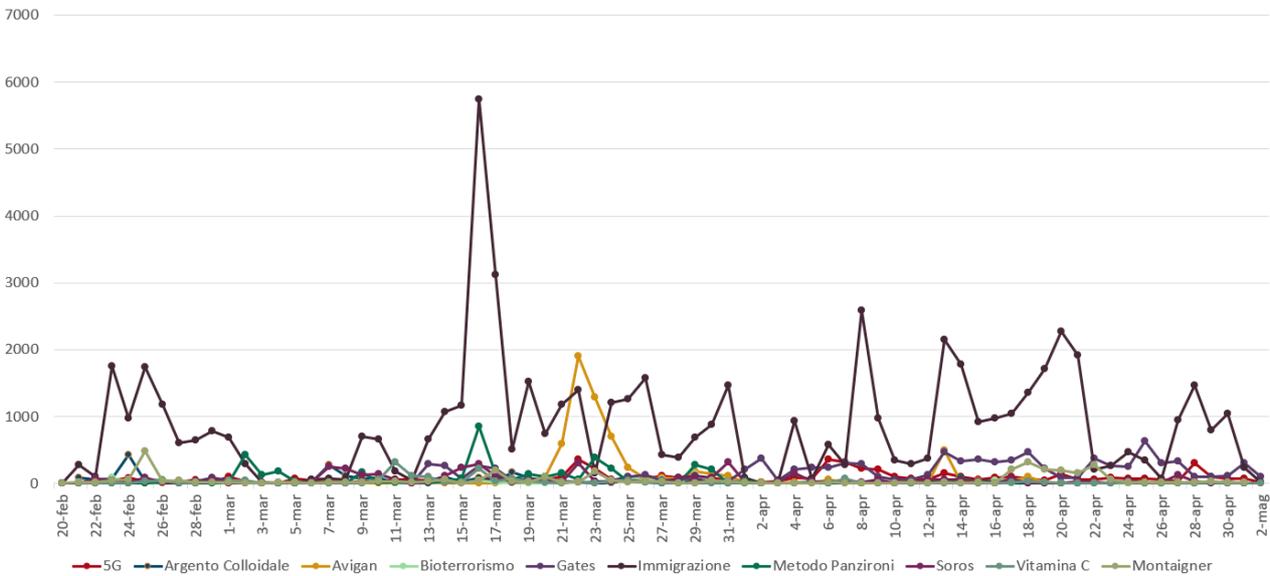


Figura 4. Visualizzazione dei network di RT relativi alle keyword 5G, Avigan, Gates, Africani

