Dipartimento di Scienze della Comunicazione, Studi Umanistici e Internazionali

Dottorato in Studi Umanistici

L'impatto delle fake news e della media manipulation sulla comunicazione di marca e d'impresa

Area di ricerca: SPS/08

Abstract

Fake news e media manipulation sono questioni ampiamente trattate dalla letteratura in relazione al loro impatto sulla democrazia e sulla salute. Tuttavia, le stesse tecniche sono sempre più spesso usate per manipolare il mercato ed alterare la reputazione di singole imprese, sfruttando le affordance delle piattaforme sociali del web e le sottostanti dinamiche sociali. Questo progetto indaga tale aspetto, meno studiato ma non per questo meno importante, attraverso le tecniche della big data analytics e della Open Source Intelligence.

Presentazione generale del progetto e stato dell'arte

Il presente progetto di ricerca triennale, ideato per il XXXVII ciclo di Dottorati di ricerca (A.A. 2021/2022), prende ispirazione dalla tesi di Laurea Magistrale in Web Marketing presentata da Massimo Terenzi (A.A. 2019/2020), dal titolo *Il ruolo dell'OSINT nel monitoraggio di fake news e media manipulation*.

Scopo di tale lavoro era la messa a punto di un sistema di monitoraggio e di analisi della disinformazione e delle fake news circolanti in rete, per individuare e reagire in modo tempestivo a contenuti potenzialmente lesivi per imprese, leader politici o redazioni giornalistiche. Le fake news, notizie ingannevoli che hanno le sembianze di notizie reali, circolano oggi prevalentemente sul web, servendosi della grande

amplificazione offerta loro dai social media. Tali contenuti hanno spesso come obiettivo il mondo della politica, ma anche le imprese e le marche possono diventare facili bersagli. Infatti, in un contesto fortemente politicizzato, le imprese rischiano di finire nel mirino di processi di viralizzazione su questioni quali delocalizzazioni, decisioni di business, personale, utilizzo di particolari materie prime o processi produttivi e così via. Inoltre, la disinformazione risulta tipicamente connessa a temi e questioni divisive, su cui sempre più spesso vengono costruite le narrazioni dei brand.

Con *media manipulation* si intende quell'insieme di pratiche accomunate dalla sistematica creazione e diffusione di disinformazione, fake news o propaganda. Gli scopi di queste pratiche possono essere ricondotti a motivazioni ideologiche, che hanno a che vedere coi temi del multiculturalismo, dell'immigrazione, del politically correct, dell'antiglobalismo e del nazionalismo, o con la costruzione di teorie cospirazionistiche (Marwick e Lewis 2017). Altre motivazioni possono essere quelle monetarie che hanno come fine la generazione di traffico web e introiti pubblicitari, da parte dei cosiddetti *cloaked website* (Daniels 2009), ossia siti «pubblicati da individui o gruppi che nascondono la paternità o fingono legittimità al fine di mascherare deliberatamente un'agenda politica nascosta»[1].

L'uso delle piattaforme social e di sistemi automatici al fine di manipolare l'opinione pubblica viene definito *computational propaganda* (Woolley 2020). I manipolatori sarebbero fortemente interessati al framing di una news story, al fine di diffondere nuove narrative di cui il pubblico si approprierebbe, formandosi opinioni difficilmente modificabili. Lo scopo ultimo dei manipolatori risulta quello di nascondere le fonti di questa cosiddetta "informazione problematica", per far sì che essa venga assorbita dai media mainstream (Marwick e Lewis 2017). Si parla a tal proposito di "fabbrica del consenso"[2] nel caso in cui vengano artificialmente gonfiate le metriche sui social media, al fine di produrre un'illusoria apparenza di popolarità di un candidato politico o di altri temi specifici (Woolley 2020). Ciò avviene attraverso ingegnose tecniche di produzione e distribuzione di informazione falsa o inesatta, che ha come target i

[1] Traduzione nostra: «published by individuals or groups that conceal authorship or feign legitimacy in order to deliberately disguise a hidden political agenda»

[2] Woolley (2020) parla di «*manufacturing consensus*». La traduzione è nostra.

giornalisti dei media mainstream ed è congegnata in modo che questi vi si imbattano (Donovan e Friedberg 2019). Queste tecniche prendono il nome di *source hacking* (ivi).

Un fenomeno strettamente connesso agli algoritmi di raccomandazione è inoltre quello delle *filter bubble* (o bolle di filtraggio) e delle *echo chamber*, entrambe riconducibili al concetto dell'omofilia: la tendenza degli individui di ricercare connessioni con altri individui considerati simili, sulla base di determinate caratteristiche comuni. Tali bolle tendono a confinare gli individui all'interno di un microcosmo informativo, ovvero concorrono al rafforzare loro ideologie o credenze, attraverso l'isolamento degli stessi individui da una rete sociale più ampia (Bruns 2019). Il ruolo delle *filter bubble* sarebbe stato recentemente ridimensionato, riconducendo il fenomeno ad un'attitudine degli individui, non predeterminata dalle affordance delle piattaforme (ivi).

A tal proposito, in letteratura emergono nuovi security framework per rispondere a alle nuove minacce che investono le community online (Goerzen et al 2019). Tali minacce non riguardano più solo un piano puramente tecnologico, ma riguardando tecnologie partecipative e sistemi sociali complessi, emergono nuove sfide. Il *sociotechnical security framework (STsec)* sostiene che, al di là dell'integrità dei dati e della sicurezza informatica propriamente detta, occorra mettere in sicurezza gli utenti connessi in rete, come target primario della manipolazione. Gli agenti malintenzionati del web sfrutterebbero le debolezze o le falle dei sistemi tecnologici, ma anche le dinamiche sociali o le variabili psicologiche in gioco, le particolari affordance delle piattaforme, e così via. La dimensione tecnologica e la dimensione sociale andrebbero pertanto analizzate come sistemi interdipendenti e co-costituiti e la sicurezza intesa come *security*[3] diventerebbe un costrutto multidimensionale e dinamico (ivi).

Le tematiche più frequentemente oggetto di disinformazione e *media manipulation* rispondono ad una serie di ansie attivate deliberatamente dai manipolatori, come la paura dell'immigrazione, la paura dell'abbandono di valori tradizionali, la paura del globalismo o teorie cospirazionistiche di vario genere (Marwick e Lewis, 2017). Naturalmente, ciò produce un impatto diretto sulla sfera pubblica e sul dibattito politico, ma gli stessi temi sono questioni su cui sempre più spesso brand e aziende si trovano a prendere posizione, come nel caso della difesa dell'ambiente, della parità di

[3] Ossia come protezione da eventi delittuosi, reati, criminalità.

genere, della pluralità e l'inclusione, del rispetto dei diritti degli LGBT e delle minoranze, del supporto dell'immigrazione (Manfredi-Sánchez 2019). La letteratura definisce la tendenza delle aziende di prendere posizione su temi politici, *corporate political shift,* e parla di *brand activism* per definire «una strategia di comunicazione che ha l'obiettivo di influenzare il cittadino-consumatore attraverso messaggi e campagne create e sostenute da valori politici»[4] (ivi, p. 348).

Inoltre, Parsons (2020) sostiene che ci sarebbero numerosi casi in cui le fake news causano un impatto rilevante sui mercati finanziari – con particolare riferimento ai titoli azionari delle imprese quotate, in quanto sarebbero in grado di provocarne l'aumento della volatilità e in quanto i mercati potrebbero subire pesanti alterazioni se una notizia falsa godesse di una rapida ed estesa diffusione (ivi).

Si consideri, ad esempio, il caso di Monsanto Company, azienda multinazionale statunitense di biotecnologie agrarie, che in un periodo compreso tra il 2017 e il 2019 è stata protagonista della fake news con il numero di interazioni più elevato. L'articolo, che risulta ora archiviato[5], era relativo alle tracce di glifosato, principio attivo dell'erbacida di Monsanto, Roundup, in numerosi cereali e altri alimenti a base di avena venduti ai bambini. La notizia, giudicata falsa dai *Third-Party-Fact-Checker*[6] di Facebook, ha superato il milione di interazioni[7]. Altro caso emblematico è quello del rivenditore di videogiochi GameStop. L'azienda, finita di recente nel mirino di alcuni fondi di investimento che scommettevano sul crollo del valore delle azioni, è stata oggetto di una manovra di "salvataggio", organizzata dagli utenti di una gruppo Reddit. La community di r/WallStreetBets, al fine di contrastare le logiche speculative dei fondi di investimento che miravano ad affossare GameStop, hanno saputo dimostrare il potere delle culture partecipative, facendo risalire apprezzabilmente il valore delle azioni dell'azienda (Umar et al. 2021).

---

[4] Traduzione nostra: «brand activism is defined as a communication strategy whose aim is to influence the citizen-consumer by means of messages and campaigns created and sustained by political values»

[5] Si veda http://archive.vn/1fZt7#selection-873.116-873.134, ultimo accesso eseguito il 20 agosto 2021

[6] Programma di fact-checking di Facebook, che riunisce fact-checker certificati da tutto il mondo. Per maggiori informazioni sul programma, si veda:
https://www.facebook.com/journalismproject/programs/third-party-fact-checking

[7] Si veda https://datastudio.google.com/s/mIO6PPGI79w, ultimo accesso eseguito il 20 agosto 2021

Al fine di prevenire e di monitorare eventi di simile natura, la metodologia dell'Open Source Intelligence (OSINT), risulta frequentemente utilizzata dagli analisti addetti alla *Business e Competitive Intelligence* (B/CI) nelle aziende (Fleisher 2008), quell'attività svolta in maniera sistematica finalizzata all'analisi di ambiente esterno e ambiente competitivo, per produrre insight utili ai decision maker (ivi).

La metodologia OSINT consiste nella raccolta, nell'elaborazione e nella messa in relazione di informazioni estratte da fonti aperte o di pubblico accesso, al fine di conseguire una conoscenza approfondita – o intelligence – su un tema od un obiettivo specifico (Pastor-Galindo et al. 2020). Una definizione piuttosto ampia, come ampi sono i contesti applicativi dell'OSINT: essa viene utilizzata da governi, forze dell'ordine o servizi di intelligence, aziende private, nonché nell'ambito dell'investigazione contro i crimini informatici (Pastor-Galindo 2020; Hassan e Hijazi 2018; Hassan 2019). Le principali applicazioni spaziano infatti dall'analisi dell'opinione pubblica alla *sentiment analysis*, resa possibile dal forte sviluppo del social media e finalizzata ad analisi di marketing o di comunicazione politica; fino alla vasta area della *cybersecurity* e *cyberdefence*, in cui l'OSINT ha lo scopo prevenire o limitare i danni di azioni malevole operate sui sistemi informatici. In ambito aziendale, l'OSINT verrebbe invece già utilizzata per investigare nuovi mercati, carpire informazioni riguardanti l'attività dei competitor, pianificare attività di marketing e predire eventi o scenari futuri (Hassan 2019). Se in passato l'uso dell'Open Source Intelligence era appannaggio di aziende dai grandi budget, oggi, con la diffusione del web e dei suoi strumenti, anche piccole aziende possono integrare in modo efficace l'analisi OSINT nei loro piani di business. Contemporaneamente, un'area nella quale l'OSINT si dimostra potenzialmente promettente sarebbe proprio quella delle fake news (ivi).

Obiettivi e metodologia della ricerca

Le piattaforme social vengono correntemente ritenute responsabili di offrire ai *media manipulator* ampio spazio di azione per diffondere fake news e tentare di influenzare l'opinione pubblica. Nel corso degli anni, le piattaforme sarebbero corse ai ripari, predisponendo dei meccanismi difensivi in grado di limitare il cosiddetto

comportamento coordinato inautentico. Esso consiste nell'attività di alcuni individui che sui social media tentano di manipolare la pubblica opinione, servendosi di bot e sistemi automatici, pagine e gruppi create ad hoc, account falsi o compromessi (Giglietto et al. 2018). I cosiddetti *social bot* avrebbero lo scopo precipuo di generare automaticamente contenuti ed imitare comportamenti umani sui social media e su altre community online. Essi, per mezzo di sofisticati algoritmi di *machine-learning* o *deep-learning*, starebbero peraltro diventando sempre più efficaci (Woolley 2020). Le piattaforme social si sarebbero attrezzate di strumenti che consentono loro di distinguere l'attività di utenti reali da quella di utenti-fantoccio (ad esempio di quei sistemi automatici che simulano l'attività di persone reali), utilizzando i cosiddetti *metadata*: delle informazioni che servono a descrivere come un qualsiasi oggetto digitale sia stato generato e archiviato e possa circolare in rete (Acker 2018). Nel contempo, i manipolatori più esperti starebbero diventando sempre più scaltri a falsificare quello che agli occhi dei social media appare come un comportamento perfettamente reale sui social media, in particolare affiancando l'intervento umano all'operato dei bot (Woolley 2020).

Sulla scorta del lavoro di tesi già menzionato, questo progetto di ricerca mira a studiare il ruolo fake news e della media manipulation in relazione alla loro influenza sulle imprese, con particolare riferimento alla *brand reputation* ed al loro valore di mercato. Riteniamo infatti che sia possibile per analisti d'impresa e dipartimenti di Business/Competitive Intelligence sfruttare i tool dell'OSINT per monitorare e meglio comprendere come originano e circolano le fake news, nonché il modo in cui esse impattino in maniera diretta o indiretta sui propri stakeholder. Un sistema di intelligence come quello qui proposto, permetterebbe di limitare il rischio connesso al potenziale lesivo delle fake news, equipaggiando le aziende con dei potenti strumenti di difesa, in grado di rilevare tempestivamente potenziali minacce a cui far fronte adottando adeguate contromisure.

Metodologia e risultati attesi

Per documentare i casi di studio seguiremo il modello proposto dal framework "ABC" proposto dal Transatlantic Work Group (François 2019), che individua 3 vettori

della manipolazione in rete: A per *Actors*, ossia gli attori della manipolazione, B come *Behavior*, cioè il comportamento ingannevole degli stessi e C come contenuti (*Content*) di cui si serve la manipolazione. Verrà inoltre impiegato il modello *media manipulation life cicle (MMLC)*, proposto dal Shorenstein Center on Media Politics and Public Policy dell'Università di Harvard (The Media Manipulation Casebook, 2021). Il modello propone di studiare una campagna di manipolazione in 4 stadi, partendo dall'analisi dell'origine e della pianificazione della campagna, all'esecuzione della campagna attraverso le tattiche e tecnologie impiegate, all'analisi di come eventuali attori istituzionali (politici, organizzazioni della società civile, media mainstream, ecc.) amplifichino, adottino o estendano la campagna. Infine, il modello MMLC mira a documentare come tech company, governi, giornalisti o società civile tentino di mitigare gli effetti e la diffusione dei contenuti manipolatori, nonché eventuali tentativi di aggiustamento dei manipolatori alle nuove condizioni ambientali.

Il lavoro sarà supportato dagli strumenti della *big data analytics*, finalizzati all'individuazione del comportamento coordinato inautentico operato in rete (livello *behavior*), partendo dal lavoro di Giglietto e colleghi (2020). In particolare, grazie al pacchetto R CooRnet (sviluppato da Fabio Giglietto, Nicola Righetti e Luca Rossi)[8], sarà possibile analizzare grandi volumi di notizie, estratte da MediaCloud[9], accertando, qualora queste vengano condivise in maniera coordinata su social media quali Facebook o Instagram, eventuali tentativi di manipolazione.

Una volta rilevata l'attività di condivisione coordinata verranno utilizzati gli strumenti OSINT, per investigare sulle fonti da cui originano i tentativi di manipolazione dell'informazione, ossia cercare di individuarne le motivazioni. Questa fase sarà dunque mirata ad individuare quegli attori (livello *actors*), la cui identità è spesso celata, che utilizzano i social media per manipolare il discorso pubblico.

Ci rifacciamo, in questo caso, al modello della *Investigative Digital Ethnography*, che associa tecnologie di investigazione forense digitale all'osservazione etnografica

[8] Per maggiori informazioni si veda: https://rdrr.io/github/fabiogiglietto/CooRnet/f/README.md, ultimo accesso eseguito il 20 agosto 2021

[9] Strumento open source di analisi che mira a mappare la copertura mediatica degli eventi attuali. Per maggiori informazioni si veda: https://mediacloud.org, ultimo accesso eseguito il 20 agosto 2021

(Friedberg 2020). Mentre gli approcci quantitativi misurano la disinformazione su larga scala, tale modello suggerisce l'uso della ricerca qualitativa per valutare i comportamenti degli individui in rete nonché l'esposizione alla disinformazione, sezionandone i contenuti, identificandone i temi ricorrenti ed osservando come questi vengano amplificati e fruiti (ivi).

Sarà dunque possibile procedere a studiare i contenuti impiegati dai manipolatori del web. Essi costituiscono certamente la parte più visibile della *media manipulation*, ma anche quella di più difficile definizione per i gestori delle piattaforme, che tuttora faticano a moderare contenuti lesivi o ingannevoli, senza incorrere in questioni legate alla libertà di espressione (François 2019). Cercheremo quindi di tracciare le narrative che vengono diffuse e fatte circolare in rete e di comprendere come le stesse influenzino le opinioni sui social media (Williams e Blum 2018), attraverso tecniche informatiche di analisi dei contenuti (Hassan e Hijazi 2018) o di social network analysis. Definiti i topic di riferimento candidati per il monitoraggio, sarà necessario testare il loro impatto sulle imprese, attraverso l'individuazione di adeguati KPI (*Key Performace Indicator*), che permettano di determinare come le narrative circolanti in rete si ripercuotono sulle narrative dei brand, ovvero sul valore di mercato delle medesime imprese. Infatti, la letteratura di riferimento raccoglie ancora pochi contributi che leghino le campagne di manipolazione rilevate in rete ad effetti diretti e concreti sui pubblici a cui sono rivolte (Woolley 2020).

Infine, una criticità ancora irrisolta della metodologia OSINT risiede nel fatto che essa risenta tuttora di una scarsa standardizzazione e che soccomba a grandi quantità di informazioni spesso contraddittorie, fra le quali occorre fare ordine con procedimenti di *truth discovery*. Non è un caso che i processi di raccolta e di analisi dei dati risultino, ancora oggi, fondati su qualità e competenze prettamente soggettive degli analisti (Fleisher 2008). Pertanto, il nostro lavoro si propone come ulteriore risultato l'intento di sistematizzare il processo di intelligence basato sulle fonti aperte, limitatamente all'oggetto dell'indagine precedentemente menzionato, allo scopo di fornire un approccio di analisi replicabile in qualunque contesto applicativo d'impresa che implichi il monitoraggio dell'informazione sul web.

Bibliografia

Acker, A. (2018), "Data Craft: The Manipulation of Social Media Metadata", in *Data & Society*, https://datasociety.net/wp-content/uploads/2018/11/DS_Data_Craft_Manipulation_of_Social_Media_Metadata_.pdf, ultimo accesso eseguito il 20 agosto 2021

Bruns, A. (2019), *Are filter bubbles real?*, Polity

Daniels, J. (2009), "Cloaked Websites: Propaganda, Cyber-Racism and Epistemology in the Digital Era", in *New Media & Society*, vol. 11, pp. 659-683

Donovan, J. e Friedberg, B. (2019), "Source Hacking: Media Manipulation in Practice", in *Data & Society*, https://datasociety.net/wp-content/uploads/2019/09/Source-Hacking_Hi-res.pdf, ultimo accesso eseguito il 20 agosto 2021

Fleisher, C. (2008), "OSINT: Its Implications for Business/Competitive Intelligence Analysis and Analysts", in Intelligencia y seguridad, vol. 4, pp. 115-141

François, C. (2019), "Actors, Behaviors, Content: A Disinformation ABC: Highlighting Three Vectors of Viral Deception to Guide Industry & Regulatory Responses", Transatlantic Working Group, https://science.house.gov/imo/media/doc/Francois%20Addendum%20to%20Testimony%20-%20ABC_Framework_2019_Sept_2019.pdf, ultimo accesso eseguito il 20 agosto 2021

Friedberg, B. (2020), "Investigative Digital Ethnography: Methods for Environmental Modeling, The Media Manipulation Casebook, Harvard Kennedy School, https://mediamanipulation.org/sites/default/files/2020-10/Investigative_Ethnography_v1.pdf, ultimo accesso eseguito il 20 agosto 2021

Giglietto, F. et al. (2018), "Understanding Coordinated and inauthentic link sharing behavior on Facebook in the run-up to 2018 general election and 2019 European election in Italy", Larica, Università di Urbino Carlo Bo, https://osf.io/preprints/socarxiv/3jteh/

Giglietto, F. et al. (2020)," Coordinated Link Sharing Behavior as a Signal to Surface Sources of Problematic Information on Facebook", in *International Conference on Social Media and Society*, pp. 85–91.

Goerzen, M. et al. (2019), "Entanglements and Exploits: Sociotechnical Security as an Analytics Framework", in *9th USENIX Workshop on Free and Open Communications on the Internet (FOCI '19)*, Santa Clara, CA 7

Golebiewski, M. e boyd, d. (2019), "Data Voids: When Missing Data Can Be Easily Exploited", in *Data & Society*,

https://datasociety.net/wp-content/uploads/2019/11/Data-Voids-2.0-Final.pdf, ultimo accesso eseguito il 20 agosto 2021

Hassan, N. A. (2019), *Digital Forensics Basics*, Apress

Hassan, N. A. e Hijazi, R. (2018), Open Source Intelligence Methods and Tools A Practical Guide to Online Intelligence, Apress

Manfredi-Sánchez, J. (2019), "Brand Activism", in *Communication & Society*, vol. 32. pp. 343-359

Marwick A e Rebecca Lewis R, (2017), "Media Manipulation and Disinformation Online", in *Data & Society*
https://datasociety.net/pubs/oh/DataAndSociety_MediaManipulationAndDisinformationOnline.pdf, ultimo accesso eseguito il 20 agosto 2021

Pastor-Galindo, J. et al. (2020), "The Not Yet Exploited Goldmine of OSINT: Opportunities, Open Challenges and Future Trends," in *IEEE Access*, vol. 8, pp. 10282-10304

The Media Manipulation Codebook (2021), "Code Book",
https://mediamanipulation.org/sites/default/files/media-files/Code-Book-1.3-July-9-2021.pdf, ultimo accesso eseguito il 20 agosto 2021

Umar, Z. et al. (2021), "A tale of company fundamentals vs sentiment driven pricing: The case of GameStop", in *Journal of Behavioral and Experimental Finance*, vol 30

Williams, H. J. e Blum I. (2018), *Defining Second Generation Open Source Intelligence (OSINT) for the Defense Enterprise*, RAND Corporation

Woolley, S. C. (2020), "Bots and Computational Propaganda", in Persily, N. e Tucker J. A. (a cura di), *Social Media and Democracy: The State of the Field and Prospects for Reform*, Cambridge University Press

Descrizione della ricerca nel triennio (fattibilità)

Periodo	Fase
I – II anno	Studio delle tecniche avanzate di analisi dei dati (framework OSINT)
II anno	Definizione dell'impatto delle fake news sui brand/imprese e individuazione dei KPI per la B/CI
III anno	Progettazione di un modello di monitoraggio basato sull'analisi dei dati da fonti aperte