

**SCUOLA DI SCIENZE**

Corso di Laurea in Informatica per il Management

**La responsabilità dei social network nella diffusione  
della disinformazione online**

**Relatore:**

Chiar.mo Prof.  
Edoardo Mollona

**Presentata da:**

Dozza Davide

Sessione II

Anno Accademico 2020/2021



# INDICE

<b>INTRODUZIONE</b>	<i>pag. 1</i>
<b>CAPITOLO 1 - LA NEUTRALITÀ DELLE PIATTAFORME ONLINE</b>	<i>pag. 3</i>
1.1 La neutralità della rete	<i>pag. 5</i>
1.2 La neutralità nella normativa europea: La direttiva europea sul commercio elettronico.	<i>pag. 6</i>
1.3 La normativa statunitense: la Sezione 230	<i>pag. 8</i>
<b>CAPITOLO 2 - LA NASCITA DELLA DISINFORMAZIONE ONLINE</b>	<i>pag. 11</i>
2.1 La disinformazione tramite i social media	<i>pag. 11</i>
2.2 Il News Feed e il suo ruolo nella diffusione di informazioni	<i>pag. 13</i>
2.3 Non è disinformazione. È propaganda amplificata	<i>pag. 15</i>
2.4 Twitter assume un ruolo attivo nella lotta alla disinformazione	<i>pag. 20</i>
2.4.1 Twitter sospende l'account di Trump	<i>pag. 21</i>
2.5 La propagazione di disinformazione inerente al Covid-19	<i>pag. 24</i>
2.5.1 L'enorme portata dei social network sulla disinformazione sanitaria globale	<i>pag. 25</i>
2.5.2 I post che hanno minato la risposta al Covid-19	<i>pag. 27</i>
2.5.3 Come Facebook può mettere in quarantena la disinformazione sulla salute	<i>pag. 29</i>
2.5.4 Correct the Record: tenere informati gli utenti	<i>pag. 30</i>
2.5.5 Detox the Algorithm: proteggere gli utenti	<i>pag. 32</i>
<b>CAPITOLO 3 - L'INTERVENTO DELLE ISTITUZIONI PER CONTRASTARE LA DIFFUSIONE DI INFORMAZIONI ILLECITE</b>	<i>pag. 36</i>
3.1 L'intervento delle Istituzioni Europee	<i>pag. 36</i>
3.2 Il Dibattito Statunitense sulla Sezione 230	<i>pag. 40</i>
<b>CONCLUSIONI</b>	<i>pag. 42</i>



# Introduzione

La tesi in questione ha come obiettivo l'analisi e lo studio della responsabilità delle piattaforme social, in particolare Facebook e Twitter, nel propagare e diffondere disinformazione, dimostrando una loro parziale colpevolezza nella diffusione non controllata di notizie false e discorsi d'odio.

Prima di vedere da vicino il ruolo dei social network, viene effettuata una panoramica sulla normative vigenti che disciplinano questi *networks*.

Etichettati dalle regolamentazioni come Internet Service Provider, viene analizzato il loro fittizio ruolo neutrale online e le loro responsabilità nel momento in cui un utente carica all'interno di questi informazioni illecite, distinguendo la legislazione europea con quella statunitense.

Anche se con qualche divergenza al loro interno, entrambe le visioni che disciplinano gli IPS poggiano sullo stesso principio, ovvero promuovere la crescita e la diffusione di aziende digitali senza caricarle di forti oneri.

Queste norme, come verrà poi approfondito, non hanno però seguito la crescita esponenziale delle piattaforme, rimanendo non aggiornate e poco coerenti con il nuovo ruolo che questi media hanno nella quotidianità.

Per comprendere dove nasce il problema della disinformazione online viene poi analizzato il primo caso dove un utente estremamente influente come Donald Trump, durante la sua corsa alla Casa Bianca nel 2015, lancia un messaggio estremamente razzista e disinformante su Facebook.

Questo episodio mise a dura prova il social di Zuckerberg, che mai prima di allora si era visto affrontare una situazione così delicata.

Partendo da questo caso, si fa un'analisi più specifica del News Feed delle piattaforme social, e di come disinformatori sistematici ne utilizzano l'algoritmo a proprio vantaggio per amplificare messaggi e cambiare la percezione della realtà.

Verrà poi fatta un'analisi più profonda sul ruolo dei Social Network nella diffusione di notizie riguardanti il Covid-19, e partendo da questo, viene visionato uno studio su come combatterne la propagazione. Studio che se applicato, potrebbe essere espanso per diminuire in generale tutte le fake news e i discorsi d'odio, senza però limitare la libertà di parola degli utenti.

Infine, viene affrontato i possibili interventi che le istituzioni potrebbero attuare in un futuro prossimo per salvaguardare gli utenti dalla diffusione di contenuti illeciti.

Concludendo la tesi, si evince come ad oggi ci si affidi sostanzialmente ad una auto regolamentazione da parte di queste piattaforme, non essendo le normative aggiornate ed in grado di compiere atti decisi nei loro confronti.

Ad ogni caso, qualsiasi sia la fonte della disinformazione, online o offline, questa va affrontata non solo con le regolamentazioni, ma soprattutto attraverso politiche volte all'educazione e responsabilizzazione dei cittadini, nonché con un'azione tesa ad alimentare il pluralismo informativo.

La scelta di sviluppare la tesi su queste tematiche nasce e segue il principio di quanto appena citato: la miglior arma di difesa nei confronti della disinformazione è l'educazione.

# Capitolo 1

## La Neutralità Delle Piattaforme Online

Le piattaforme online costituiscono il principale punto di accesso a informazioni e ad altri contenuti su internet, tramite motori di ricerca, social network, siti di micro-blogging e molto altro. Internet non solo ha aumentato il volume e la varietà di notizie a disposizione dei cittadini, ma ha anche trasformato le modalità in cui questi ultimi si relazionano ad esse. I social network, in particolare, da passatempo per lo svago, la socializzazione e la condivisione di materiale tra “amici”, sono divenuti un canale di informazione e, dunque, strumenti essenziali per operare valutazioni e pervenire a convincimenti personali. Inoltre, sul web ogni utente può diventare un attore della comunicazione, a differenza dei media tradizionali.

Internet rompe gli schemi e consente, per la prima volta, a chiunque di dire la propria opinione, addirittura di dirla nel modo che preferisce.

Questo è il motivo per cui può affermarsi che tramite la rete si attua in modo pieno e concreto la libertà di informazione prevista dall’art. 21 della Costituzione: sia dal punto di vista attivo (informare, divulgare notizie, esprimere commenti), sia dal punto di vista passivo (essere informati, scegliendo tra una pluralità di fonti, espressione di posizioni diverse). Sotto questo profilo, va anche riconosciuto al web di aver consentito una maggiore partecipazione alla vita politica del Paese: la facilità con cui informazioni di vario tipo sono messe a disposizione di chiunque rende il processo democratico più partecipativo e inclusivo.

Tuttavia, questa evoluzione positiva della rete si è accompagnata a fenomeni di degenerazione del suo utilizzo, mediante la divulgazione di manifestazioni di odio mirate e di propagazione di disinformazione su vasta scala. La progressiva crescita di questi fenomeni ha portato a una sempre maggiore attenzione al ruolo degli intermediari digitali – soprattutto social network – come “moltiplicatori” dei discorsi di odio, che minano le basi della convivenza civile, e di disinformazione, che influenzano dinamiche decisionali e distorcono eventi elettorali. Ciò ha determinato nel tempo iniziative politiche d’intervento anche normativo.

## 1.1 La neutralità della rete

“La tecnologia non è né buona, né cattiva, ma neanche neutrale” - Melvin Kranzberg<sup>1</sup>

(Solo) Questa frase, di gran lunga antecedente alla nascita di piattaforme come Facebook e Google, mette in luce quello che è il reale potere della tecnologia nel mondo.

La natura dei motori di ricerca e dei social network tende a suggerirci il loro ruolo di neutralità, infatti essi non prendono molto spesso posizione in relazione ai contenuti e alle informazioni caricate dagli utenti. Senza alcuna responsabilità lasciano che, entro certi limiti (molto labili), tutto li attraversi.

La tecnologia non è mai neutrale, ma è sempre veicolata da una scelta: dell'organizzazione dei profitti, della struttura della piattaforma, della direzione in cui si vuol far dirigere l'algoritmo.

Tali decisioni, di chiara influenza politica ed economica, hanno lo scopo ben preciso di tenere l'utente ancorato il più possibile alla piattaforma.

Una delle scelte fondamentali che operano le piattaforme risiede, come già detto, proprio nella struttura dell'algoritmo che opera in esse. Ad oggi, la sua implementazione, predilige il contenuto che genera movimento, in questo modo ciò che diventa virale viene premiato.

Se riprendiamo quindi la frase di apertura, la non neutralità dell'algoritmo è evidente, dato che esso tenderà sempre a premiare la quantità indipendentemente dalla qualità.

La natura psicologica e sociologica di questa scelta l'ha spiegata meglio di tutti Roger McNamee, uno dei venture capitalist di maggior successo d'America e tra i primi investitori in Facebook, il quale riporta come:

“Quando gli utenti sono arrabbiati, consumano e condividono contenuti. Se rimangono imparziali e calmi hanno relativamente poco valore per Facebook, che fa di tutto per attivare il cervello rettile”<sup>2</sup>

Il cervello rettile è quella parte legata all'istinto atavico e quello che si occupa della “prima impressione”, che ha pochissimo tempo per decidere e per dare una risposta immediata.

---

<sup>1</sup> R. McNamee, Zucked. Come aprire gli occhi sulla catastrofe di Facebook, Nutrimenti, 2018.

<sup>2</sup> R. McNamee, Zucked. Come aprire gli occhi sulla catastrofe di Facebook, Nutrimenti, 2018.



Notiamo perciò come il tempo diventi antagonista della viralizzazione. Avere una visione imparziale delle cose, e prendersi del tempo per capire e comprendere richiede tempo. E questo tempo è visto di cattiva luce da queste realtà, le quali mirano ad ottenere una profilazione dell'utente più precisa possibile per chiari motivi di business.

Da un punto di vista legislativo, è necessario evidenziare come queste piattaforme nascano in realtà con una sorta di esenzione di responsabilità. Infatti, storicamente tutti gli Internet Service Provider (ISP) sono nati sotto normative dal carattere estremamente liberale, create al fine di agevolare la libera circolazione e la promozione dei servizi della società dell'informazione.

La neutralità e l'immunità di cui le piattaforme godono, ha avuto il lato negativo di creare amplificatori di dinamiche quali disinformazione e hate speech. Ci si chiede, pertanto, se servano nuove norme per arginare i fenomeni descritti o se quelle già esistenti possano essere sufficienti a contrastare le condotte che, on line così come off line, violano prescrizioni dell'ordinamento.

L'evoluzione del ruolo dei social network sta inducendo a porsi problemi circa la suddetta esenzione da responsabilità. I gestori di tali piattaforme traggono profitto dalla raccolta dei dati degli utenti (il cosiddetto user data profiting), poiché questi dati vengono usati dagli inserzionisti allo scopo di individuare i profili cui destinare pubblicità mirata in base a gusti, preferenze, attitudini (behavioural advertising). Inoltre, attraverso l'utilizzo di algoritmi, i gestori sono in grado di evidenziare alcuni contenuti (trending feeds) che generano con maggior frequenza interazione fra gli utenti, i quali in questo modo tendono a trascorrere un tempo maggiore sulla piattaforma, producendo di conseguenza un forte volume di informazioni su di loro, le quali supporteranno gli inserzionisti nell'attività pubblicitaria.

Dunque, i social network provider non si limitano a svolgere un'attività di hosting neutrale, ma intervengono direttamente nell'organizzazione, nella gestione e talvolta anche nell'editing dei contenuti, al fine di aumentare i ricavi derivanti dalla raccolta pubblicitaria.

Oltre a ciò, i social network stanno sviluppando sinergie con editori tradizionali i quali, considerati il volume di traffico e la raccolta pubblicitaria sui social media, stringono con i gestori di questi ultimi alleanze per raggiungere un pubblico più ampio e meglio profilato.

Potremmo quindi affermare con una forte certezza la non neutralità di queste piattaforme.

## **1.2 La neutralità nella normativa europea: La direttiva europea sul commercio elettronico**

Mediante i contenuti che circolano sulle piattaforme di social networking si possono riscontrare le più ampie tipologie di illeciti: alcuni più seri, come propaganda terroristica e pedopornografia; altri relativi alla dimensione individuale, legati alla violazione di diritti della persona, come il trattamento illecito dei dati personali, diffamazione, cyberbullismo; altri ancora relativi ad illecite manifestazioni del pensiero in termini di discorso d'odio o propagazione di notizie infondate.

È naturale chiedersi se le piattaforme non abbiano delle responsabilità, essendo l'intermediario con cui un atto illecito si compie.

A rispondere a questo interrogativo troviamo la direttiva europea per il commercio elettronico (n 2000/31/CE).

Pur trattandosi di una disciplina oramai datata ha fino ad oggi permesso di attribuire, in qualche caso, ai provider la responsabilità civile per gli illeciti commessi dai loro utenti.

Questa direttiva è stata emanata con l'obiettivo di favorire la libera circolazione e la promozione dei servizi della società dell'informazione, eliminando gli ostacoli allo sviluppo del commercio elettronico.

Da questo punto di vista, come per la legislazione statunitense, era necessario favorire l'attività degli "internet service providers" (ISP) senza caricarli di oneri eccessivi - come, ad esempio, la sorveglianza ex ante o post ante sui contenuti diffusi dagli utenti in rete.

Questa direttiva esula quindi i providers, in quanto fornitori di un servizio di hosting neutrale, da obblighi di sorveglianza sulle informazioni che trasmettono o memorizzano e di ricerca di fatti o circostanze che configurino attività illegali: dato che ciò rappresenterebbe per essi un impegno spropositato.

Il presupposto alla base della regolamentazione - certamente valido alla stesura di questa, ma discutibile oggi visto l'evoluzione delle piattaforme - è che tutti gli intermediari mantengano una posizione neutrale.

Infatti, a favore degli ISP opera una presunzione di non-conoscenza di eventuali illiceità dei contenuti user-generated, così che essi non siano gravati da responsabilità derivanti da questi ultimi.

La direttiva citata, tuttavia, stabilisce a loro carico l'obbligo di informare senza indugio l'autorità giudiziaria o amministrativa qualora vengano a conoscenza di presunte attività o informazioni non lecite riguardanti un utente.

Deve essere inoltre loro premura il fornire senza requie, a richiesta delle autorità competenti, le informazioni in loro possesso che consentano l'identificazione dell'utente stesso, al fine di individuare e prevenire attività illegittime.

L'ISP è civilmente responsabile dei contenuti pubblicati solo nel caso in cui, a seguito di una richiesta dell'autorità giudiziaria o amministrativa, non abbia agito prontamente per impedire l'accesso a tali contenuti ovvero se, avendo avuto conoscenza del loro carattere illecito o pregiudizievole per un terzo, non abbia provveduto ad informarne l'autorità competente.

Nel corso degli anni, fra il 2010 e il 2015, la direttiva ha mostrato alcuni limiti e la Corte di Giustizia Europea ne ha messo in evidenza due aspetti in particolare.

Il primo riguarda la neutralità degli ISP, la quale non può essere invocata qualora il prestatore del servizio non si limita al trattamento automatico dei dati, ma svolge un ruolo attivo atto a dare una connotazione diversa, per esempio attraverso un'attività di organizzazione o ottimizzazione dei contenuti.

Il secondo riguarda l'obbligo di predisporre un filtro preventivo nei contenuti caricati da parte degli utenti, il quale essendo un sistema complesso e costoso, causerebbe una grave violazione della libertà d'impresa da parte dei servizi di hosting.

### 1.3 La normativa statunitense: la Sezione 230

Nell'universo statunitense invece, a garante immunità alle piattaforme di siti web rispetto ai contenuti di terze parti troviamo la Sezione 230.

La sezione 230 è una disposizione che fa parte dello Communications Decency Act, atto sulle telecomunicazioni del 1996.

Il primo comma di questa norma specifica l'assenza di responsabilità da parte di un "servizio informatico interattivo" sulle informazioni pubblicati da terze parti.

Questa sezione, ritenuta come *le 26 parole che hanno cambiato internet* riporta infatti:

"Nessun fornitore di servizi internet e nessun utilizzatore di tali servizi può esser ritenuto responsabile quale editore o quale autore di una qualsiasi informazione che sia stata fornita da terzi."<sup>3</sup>

Con un esempio, questo significa che un player come Facebook non è responsabile di ciò che i suoi utenti postano, a differenza però di ciò che accade con i media tradizionali, che sono invece citabili in giudizio per le informazioni pubblicate.

Il secondo comma prevede invece l'immunità dalla responsabilità civile per i fornitori di servizi di informazione che rimuovono o limitano contenuti. I provider hanno la possibilità di limitare l'accesso a materiale che questi considerino "osceno, lascivo, sporco, eccessivamente violento, molesto o altrimenti discutibile, indipendente dal fatto che tale materiale è costituzionalmente tutelato", purché agiscano "in buona fede" in tale azione.

Prima di Internet, la giurisprudenza era chiara sul fatto che fosse stata tracciata una linea di responsabilità tra editori di contenuti e distributori di contenuti; un editore dovrebbe essere a conoscenza del materiale che sta pubblicando e quindi dovrebbe essere ritenuto responsabile per qualsiasi contenuto illegale che ha pubblicato, mentre un distributore probabilmente non ne sarebbe a conoscenza e quindi ne sarebbe immune.

La Corte Suprema ha stabilito che mettere la responsabilità su un utente avrebbe "un effetto collaterale di inibire la libertà di espressione, rendendo l'individuo più riluttante a esercitarlo."

---

<sup>3</sup> uscode.house.gov, fonte: [https://uscode.house.gov/view.xhtml?req=\(title:47%20section:230%20edition:prelim\)](https://uscode.house.gov/view.xhtml?req=(title:47%20section:230%20edition:prelim))

La Sezione 230 ha avuto origine da due cause intentate nei confronti di due servizi attivi all'epoca: CompuServe e Prodigy. Entrambi offrivano ai loro utenti forum in cui potevano esprimersi liberamente, ma mentre CompuServe aveva deciso di non moderarne i contenuti, Prodigy aveva un team di moderatori che dovevano approvare o meno i messaggi pubblicati.

Entrambe le società furono querelate per alcuni contenuti pubblicati dagli utenti, ma mentre CompuServe non fu ritenuta responsabile per non aver moderato i messaggi, Prodigy fu considerata al pari di un giornale, data la moderazione che effettuava, e per questo ritenuta direttamente responsabile. Dopo questa decisione, le società di Internet fecero pressioni per avere una legge che riempisse il vuoto normativo, e da lì nel 1996 il Congresso inserì la Sezione 230 nel Communications Decency Act, che faceva parte di una più ampia legislazione sulle telecomunicazioni.

Questa regolamentazione aveva il semplice intento di favorire il progresso e lo sviluppo di internet, nacque infatti in un momento storico che lo vedeva ancora agli albori, e non era chiara la possibile evoluzione del fenomeno.

“Volevamo che le piccole imprese iniziassero a concentrarsi sulle assunzioni di ingegneri, sviluppatori, designer piuttosto che preoccuparsi di come dovessero assumere un team di avvocati.”<sup>4</sup>

Come disse Ron Wyden, co-autore della sezione 230.

L'aver sostenuto e in qualche modo tutelato la Sezione 230 è stato essenziale per la crescita di Internet nella prima parte del 21° secolo. Insieme al Digital Millennium Copyright Act (DMCA) del 1998, la Sezione 230 fornisce ai fornitori di servizi Internet un porto sicuro per operare come intermediari di contenuti senza timore di essere responsabili per le informazioni al loro interno, purché adottino misure ragionevoli per eliminare o impedire l'accesso a quel contenuto. Queste protezioni hanno permesso nuove e sperimentali applicazioni nell'universo Internet senza l'interferenza di ramificazioni legali, creando le basi dei moderni servizi Internet come i motori di ricerca avanzati, social media, streaming video e cloud computing.

---

<sup>4</sup> C-Span, pronunciato durante un discorso in diretta. 21 Marzo 2018. <https://www.c-span.org/>

NERA Economic Consulting ha stimato nel 2017 che la Sezione 230 e il DMCA, insieme, hanno contribuito a circa 425.000 posti di lavoro negli Stati Uniti nel 2017 e hanno rappresentato un fatturato totale di 44 miliardi di dollari all'anno.<sup>5</sup>

Queste enorme crescita dei provider ha portato a numerose chiamate per aggiornare la sezione, provenienti da ogni player coinvolto.

Il mondo repubblicano ha sollevato più volte la necessità di una modifica alla disposizione, preoccupato che le *internet companies* stiano eliminando o oscurando svariati post di natura conservativa.

Nello schieramento opposto, i Democratici sono al contrario allarmati che questi fornitori possano lasciare online un numero eccessivo di post riguardanti principalmente disinformazioni su elezioni o altre controversie.

Tutti gli attori coinvolti, anche le piattaforme, convergono la necessità di un intervento legislativo che aggiorni la disposizione al periodo corrente.

Tuttavia, un intervento di questo genere risulta ad oggi particolarmente complesso, vista la non omogeneità dei pensieri in sede di Congresso.

---

<sup>5</sup> Economic Value of Internet Intermediaries and the Role of Liability Protections, 2017 - <https://internetassociation.org/wp-content/uploads/2017/06/Economic-Value-of-Internet-Intermediaries-the-Role-of-Liability-Protections.pdf>

# Capitolo 2

## La Nascita della Disinformazione Online

Quando nasce il problema della non neutralità delle piattaforme e la relativa propagazione di disinformazioni? È infatti lecito chiedersi il perché questo problema sia emerso in maniera così evidente solo di recente e non già in tempi più lontani.

### 2.1 La disinformazione tramite i social media

#### Il caso Donald Trump

Il 7 dicembre 2015 Donald Trump caricava su Facebook un video di un discorso tenuto a Mount Pleasant, in South Carolina.

L'allora candidato Presidente prometteva di adottare una linea decisamente più dura nei confronti dei terroristi, legando poi il terrorismo all'immigrazione.

“Donald Trump chiede il blocco totale e completo degli ingressi dei musulmani negli Stati Uniti finché i rappresentanti del nostro Paese non capiscono cosa diavolo stia succedendo.”<sup>6</sup>

Sulla piattaforma il video generò rapidamente oltre 100mila like e 14mila condivisioni.<sup>7</sup> Fu uno dei primi contenuti caricati che mise davvero in difficoltà la piattaforma, la quale si trovava totalmente impreparata ad un utente come Trump, capace contemporaneamente di ottenere immenso seguito e di dividere sia utenti ma anche dipendenti della stessa piattaforma.

I vertici di Facebook discutendo, usarono un escamotage per non essere costretti a difendere il linguaggio politico particolarmente al limite di Trump: sostennero che quello che atteneva alla politica potesse godere di protezione in quanto “meritevole di diffusione”.

L'idea alla base era quella che al discorso politico spettasse una particolare protezione, in modo da consentire al pubblico di farsi una propria opinione fondata sui punti di vista non censurati dei candidati.

---

<sup>6</sup> NBC News: At South Carolina Rally, Donald Trump Defiant on Muslim Ban - nbcnews.com

<sup>7</sup> S. Frenkel, C. Kang: “Facebook: L'inchiesta Finale” - Edito da Einaudi

In questo contesto Facebook, e in generale le piattaforme da dove Trump aveva rilanciato la notizia, erano rimaste neutrali attraverso un ritornello ripetuto quasi come un mantra: l'assoluta necessità di proteggere la libertà di parola così come espressa nel Primo emendamento della Carta dei diritti.

Questa libertà di parola adottata nel 1791 era però stata pensata specificamente per favorire una democrazia sana, assicurando una pluralità di idee senza alcun tipo di restrizione imposta dal governo, il tutto con perseguendo lo scopo di proteggere la società.

L'indirizzamento pubblicitario basato su clic e contenuti volgari e lo sfruttamento dei dati degli utenti erano però antitetici rispetto agli ideali di una società sana.

E se da una parte creava a Facebook diversi problemi, dall'altra Trump era anche un utente potente e un inserzionista importante.

Dall'inizio della sua campagna il suo comitato elettorale aveva investito nei social network la maggior parte dei fondi destinati ai mezzi di comunicazione.

Si erano concentrati su Facebook per i suoi strumenti di indirizzamento semplici ed economici così da amplificare gli spot elettorali.

Brad Parscale, digital manager di Trump, sfruttava gli strumenti di microtargeting di Facebook per raggiungere gli elettori. Lavorava con dipendenti di Facebook interni per monitorare quotidianamente i discorsi di Hillary Clinton e indirizzare annunci negativi a un determinato pubblico.

Trump divenne una presenza imprescindibile sulla piattaforma. Le elezioni presidenziali del 2016 avrebbero cancellato ogni dubbio in merito all'importanza dei social media nelle campagne elettorali. All'inizio di quell'anno il 44% degli americani diceva ormai di informarsi sui candidati tramite Facebook, Twitter, Instagram e YouTube.

Il problema della non neutralità delle piattaforme comincia a crearsi nella storia di internet da lì a poco quando un partito, che non è quello democratico, assunse la presidenza degli Stati Uniti d'America.

Alla mancata vittoria di Hilary Clinton sorse infatti il problema di come gestire le relazioni delle piattaforme con il nascente sovranismo populista rappresentato da Donald Trump che avrebbe poi contagiato l'Europa e le altre realtà.



## 2.2 Il News Feed e il suo ruolo nella diffusione di informazioni

Lo strumento con il quale Facebook, ma in generale comune a tutti i social network, trasmette agli utenti le notizie per loro più pertinenti.

Questa sezione è infatti una raccolta personalizzata in continuo aggiornamento di foto, video, link e notizie di amici, familiari, aziende e organi di informazioni che un utente segue sui Social.

Uno dei motivi che hanno permesso a Donald Trump di vincere le elezioni è stata la sua abilità (e quella del suo staff) nel comprendere e sfruttare al meglio tutte le potenzialità di questo strumento.

### **Come funziona il News Feed**

Vengono considerati principalmente tre fattori per stimare quanto un contenuto sia pertinente per l'utente:

- La fonte di pubblicazione: amici, familiari, organi di informazione, aziende e personaggi pubblici con cui un utente interagisce maggiormente hanno la priorità nel Feed.
- Il tipo di contenuto: i tipi di contenuti con cui un utente interagisce con maggior frequenza, ad esempio foto, video o link, hanno la priorità.
- Le interazioni con i post: i post con un elevato numero di interazioni (mi piace, reazioni, commenti, condivisioni).

I post vengono ordinati in modo che le notizie principali siano visualizzate per prime.

In base a questi fattori, ad ogni post è assegnato un punteggio che ne determina la posizione e la priorità nel Feed.

Quando le campagne presidenziali del 2016 entrarono nel vivo, il News Feed mostrava come gli americani fossero più divisi che mai.

La corsa tra Hillary Clinton e Donald Trump fu una frenesia di notizie altamente faziose, con le quali ognuna delle due parti non vedeva l'ora di demonizzare l'altra.

Lo staff di Trump sfruttava l'algoritmo di Facebook a proprio favore: per gli utenti i messaggi provocatori dei politici populistici erano irresistibili. D'accordo o meno con i contenuti, erano in grado di attirare l'attenzione.



**Figura 2.1** Immagini utilizzate come creatività all'interno della Campagna Pubblicitaria "Defeat Crooked Hillary", create da CA per supportare Trump durante le elezioni 2016.

Fonte: campaignlegal.org

Gli algoritmi di Facebook interpretavano quell'attrazione come interesse e silenziosamente attribuivano dei punti a quel determinato contenuto, spingendolo sempre più in alto nel Feed degli utenti.

Nonostante la chiara pericolosità di uno strumento del feed così come ottimizzata, Zuckerberg ha ribadito per anni lo stesso concetto. Facebook doveva bilanciare la libertà di espressione e la sicurezza. Non riteneva di dover eliminare le false informazioni. Lasciò infatti agli stessi utenti la possibilità di segnalare i contenuti come fake news, consentendo alla piattaforma di rendere più difficile il reperimento di quel determinato contenuto.

## **2.3 Non è disinformazione. È propaganda amplificata**

Prendiamo ora ad esempio un caso molto interessante che spiega la forza delle piattaforme nel diffondere disinformazioni tramite l'utilizzo di #hashtag e come questi possano essere usati per influenzare il news feed "organico" degli utenti.

### **Il caso Shahid Buttar e Nancy Pelosi**

L'esempio in questione vede in luce come protagonista Shahid Buttar, l'attivista di San Francisco che sfidava la presidente della Camera Nancy Pelosi al ballottaggio del Congresso degli Stati Uniti del 2020.

Buttar necessitava di attirare l'attenzione su di se, poiché i media tradizionali ignoravano in gran parte la campagna del sottovalutato partito socialista democratico.

Per attirare l'attenzione mediatica, Buttar e il suo stato avevano costruito una strategia di comunicazione Twitter-centrica. L'idea era quella di disegnare alcuni tweet con l'hashtag #PelosiMustGo per fare in modo che questo entrasse nell'elenco delle tendenze a livello nazionali, dando così massima visibilità al messaggio.

È importante ricordare come nelle piattaforme ci siano spesso attività di questo genere. Tentativi di diffondere messaggi tramite il gioco degli algoritmi che Twitter, Facebook e altri social utilizzano per identificare i contenuti popolari e farli conoscere agli utenti.

Di fronte a campagne che mirano a far sembrare determinate idee più diffuse di quanto in realtà non siano, si è iniziato ad utilizzare etichette come "misinformazione" e "disinformazione".

Il termine misinformazione indica la diffusione di notizie false in modo involontario o a causa di semplice irresponsabilità, mentre disinformazione è la divulgazione intenzionale di notizie o informazioni inesatte o distorte allo scopo di influenzare le azioni e le scelte di qualcuno.

Sentiero comune che trovano queste due espressioni è senz'altro quello di avere un intento di propaganda.

## L'evoluzione della propaganda

La propaganda è "l'attività di disseminazione di idee e informazioni con lo scopo di indurre a specifici atteggiamenti e azioni" ovvero il "conscio, metodico e pianificato utilizzo di tecniche di persuasione per raggiungere specifici obiettivi atti a beneficiare coloro che organizzano il processo"<sup>8</sup>. In antitesi alla propaganda dovrebbe essere la pura e semplice esposizione dei fatti nella loro completezza ovvero la descrizione della realtà nella sua interezza.

Questo termine nel corso della storia ha subito radicali trasformazioni.

La parola propaganda nasce in un contesto religioso quando la Chiesa cattolica nel XVI secolo organizzò, come contrattacco alla diffusione del protestantesimo, una *Congregatio de propaganda fide*, un dipartimento preposto alla propagazione della fede cattolica.

È importante sottolineare come originariamente questo termine non intendesse riferirsi a informazioni fuorvianti.

Il moderno significato, invece, risale all'uso che se ne fece a partire dalla prima guerra mondiale, dove l'intento divenne quello di eludere la ragione di un cittadino, di spingerlo con l'inganno o l'imbroglione verso specifiche direzioni atte a beneficiare coloro che organizzano il processo.

Questo modello di trasmissione di informazioni ha un approccio top-down: governi, autorità, istituzioni, mass media comunicano unidirezionalmente all'audience di riferimento, la quale ha un ruolo passivo potendo solo ricevere quanto comunicato.

Con l'approccio dei social il pubblico non ha più un ruolo unicamente ricettivo: lungi dall'essere semplicemente un bersaglio, questo diventa un partecipante attivo nella creazione e nell'amplificazione selettiva di narrazioni che modellano la realtà.

Una parola che forse descrive al meglio questa dinamica emergente bottom-up è quella che ancora non esiste: **ampliganda**<sup>9</sup>, ossia la formazione della percezione attraverso l'amplificazione.

I social media hanno così posto fine al monopolio della propaganda dei media tradizionali, introducendo un nuovo modo di comunicare, creando un sistema in cui la fiducia è stata

---

<sup>8</sup> Propaganda, Wikipedia.

<sup>9</sup> The Atlantic: It's Not Misinformation. It's Amplified Propaganda.

riallocata da figure di autorità e media legacy a individui carismatici capaci di fare appello agli aspetti dell'identità personale o ideologica che il loro pubblico ha più a cuore.

Il successo di un'ampliganda può scaturire in due modi: in primo luogo, quando un utente particolarmente influente riesce ad ottenere una folla di seguaci ideologicamente allineamenti per diffonderne il messaggio. In alternativa, può accadere che un'idea emerga spontaneamente da qualche utente online, che incontri poi una buona fetta di *supporter* che la sostengano e le diano la giusta visibilità sufficiente per arrivare ad alcuni influencer che la amplificano definitivamente, scatenando una cascata di azioni da fazioni adiacenti.

Sono proprio questi i prerequisiti chiave che hanno permesso a Buttar di creare un momento virale: una community forte, composta da una base di sostenitori online con esperienza nel conflitto sulla piattaforma ed uno slogan che esprimesse la giusta indignazione.

Come detto, non sono nuove le cosiddette "campagne #hashtag", ma quella di Buttar è stata particolarmente efficace e ha presto illuminato il moderno sistema di propaganda.

Questa catena di amplificazione è uno strumento incredibilmente potente; fa emergere violazioni dei diritti civili, movimenti di protesta ed eventi di rotture indipendentemente dal fatto che i media tradizionali scelgano di coprire tali eventi o meno. Ma non solo, è anche il modo in cui le affermazioni mediche ciarlatane una sfilata quotidiana di teorie cospirative vengono rese di tendenza.



Figura 2.2 Il Tweet di Shahid Buttar che ha dato il via al movimento.

Fonte: Twitter

## **La risposta delle piattaforme**

Negli anni precedenti al caso in esame, sfruttare funzionalità come le *Trending List* era diventato più impegnativo dato che le società di social media erano diventate consapevoli della possibile manipolazione.

Nel 2018, Twitter aveva già iniziato a scartare i post di bot e “account marionette” per determinare quali argomenti stavano diventando popolari.

Facebook dal lato suo aveva già avuto a che fare con una fabbrica di troll russi che avevano destabilizzato la piattaforma.

Sulle piattaforme furono di conseguenza istituiti team di integrità per cercare "comportamenti non autentici coordinati", ovvero attività sospette da parte di reti di account che, in molti casi, consistevano in falsi personaggi.

Per le piattaforme tecnologiche, era facile giustificare la repressione di account falsi, reti di bot e troll istituzionali, dato che al pubblico in generale non importava molto dei diritti di libertà di parola di persone false, ovviamente.

La forza di #PelosiMustGo non derivava da account illeciti, si trovava proprio nell'amplificazione organica del messaggio. Di fatti, sempre più account cosiddetti “reali” si sono fatti avanti per diffondere messaggi che, in passato, avrebbero potuto essere amplificati dai bot.

Nel corso della giornata l'hashtag raggiunse rapidamente il 7° posto nella *Trending List*.

Da quel momento, diversi personaggi politici (a favore o contrari) hanno contribuito alla diffusione dell'hashtag citandolo in diversi tweet.

Qualsiasi fosse lo schieramento nei confronti di Buttar o Pelosi, la ricondivisione di questo ha solo contribuito all'amplificazione del messaggio.

Il motivo di tutto questo è presto spiegato: l'indignazione genera coinvolgimento, che algoritmicamente genera più interazione, e anche coloro che non vogliono fare a pezzi il tessuto della società americana sono comunque incoraggiati a rispettare queste regole nel loro sforzo di richiamare l'attenzione sulla loro causa.

Di tutti i cambiamenti operati dai social network, questa capacità delle folle online di influenzarsi a vicenda è tra le più importanti e sottovalutate.

Cinque ore dopo l'inizio della campagna, erano stati postati più di 100.000 tweet e l'hashtag #PelosiMustGo aveva raggiunto il primo posto in tendenza.

La vittoria finale, ovviamente, sarebbe stata di Pelosi; lei rimane il presidente della Camera. Ma un fatto non indifferente è emerso: un politico e poco più di 100 account si sono mossi di concerto con decine di migliaia di altri utenti per richiamare l'attenzione del pubblico sul dodicesimo distretto della California, dove Shahid Buttar, socialista, attivista e femminista - stava facendo una campagna per spodestare Nancy Pelosi, l'allora democratico più potente d'America.

Le norme politiche e civiche tardano ad adattarsi a situazioni come questa.

Online, vi è la sincera sensazione di essere circondati in ogni momento dall'urgenza, dalla richiesta di agire.

Anche se l'utente è spesso ignaro del perché qualcosa è apparso sul suo feed, ciò non elimina quella sensazione di necessità di prestare attenzione.

Comprendere gli incentivi degli influencer, riconoscere le tecniche retoriche molto comuni che scatenano l'indignazione, sviluppare la consapevolezza di come le folle online ora partecipano alla cristallizzazione dell'opinione pubblica: questa è un'educazione di cui si ha bisogno.

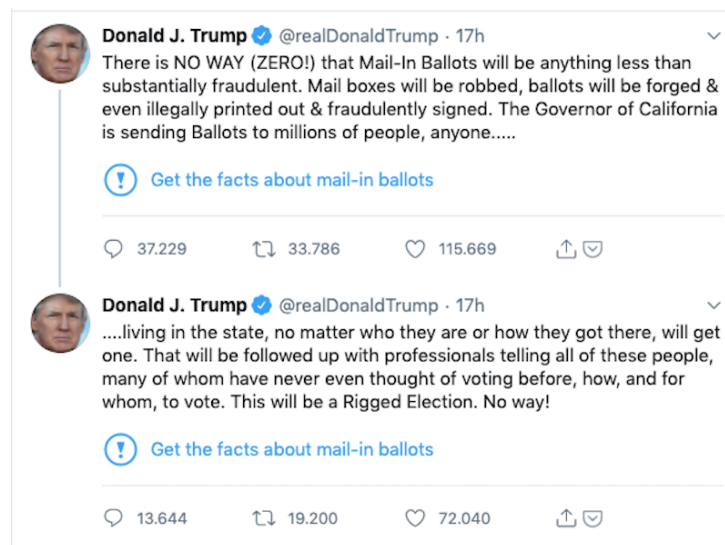
I regolatori stanno tentando di stabilire quali guardrail potrebbero richiedere l'infrastruttura di comunicazione e le piattaforme che hanno progettato l'architettura modificano incessantemente le loro politiche in risposta all'ultima esposizione mediatica di conseguenze indesiderate.

## 2.4 Twitter assume un ruolo attivo nella lotta alla disinformazione

Questa libertà di poter utilizzare le news feed delle piattaforme utenti politici influenti a proprio piacimento senza ne curarsi della veridicità delle informazioni espresse ne dei movimenti che queste generano trova finalmente un ostacolo.

Porta la data del 28 Maggio 2020 il primo intervento che potremmo definire “editoriale” da parte di una piattaforma, in questo caso Twitter, nei confronti di un organo di Stato.

Infatti, sotto l’ennesimo post populista dell’allora presidente americano, il social ha segnalato come fuorviante il contenuto del post ed aggiunto in coda un link per apprendere i fatti. È stata la prima volta che Twitter ha preso un provvedimento del genere, dopo che per anni esperti e osservatori avevano criticato il fatto che non avesse un protocollo per occuparsi delle notizie false o fuorvianti pubblicate da personaggi con milioni di follower.



**Figura 2.3** . I Tweet fuorvianti di Donald Trump riguardanti il voto via posta.

**Fonte:** Twitter

Il tweet in questione riguardava il voto via posta, oggetto di acceso dibattito durante la appena conclusa campagna elettorale per via della pandemia da coronavirus. Trump sosteneva con sicurezza che il voto postale sarebbe stato «falsato» perché un sacco di schede sarebbero finite per essere rubate, e che il governatore della California (un Democratico)



stava inviando «milioni» di schede, anche a chi non dovrebbe riceverla, e tutto ciò avrebbe condizionato gli elettori nello spiegare loro come votare.

Tutte queste informazioni erano e sono ovviamente false: diversi stati americani usano già il voto per posta con vari strumenti di controllo e non esistono prove di brogli; il governatore della California non ha inviato nessuna scheda elettorale ai cittadini.

In fondo ai due tweet che contengono il testo, Twitter ha inserito una frase in blu – «leggi come stanno le cose sul voto postale» – preceduta da un punto esclamativo: due elementi che segnalano chiaramente il contenuto come fuorviante.

Il testo rimanda ad un articolo di CNN che spiega le molte imprecisioni e forzature del tweet di Trump.

Questa azione del social network, anche se possiamo interpretarla come eticamente lecita, è in forte contrasto con il ruolo neutrale che le normative conferiscono a queste piattaforme. Infatti, è legittimo definire a tutti gli effetti questo intervento come un'intervento editoriale, operazione che alla piattaforma in realtà non spetterebbe.

#### **2.4.1 Twitter sospende l'account di Trump**

Quanto appena visto non è il solo atto di moderazione compiuto da Twitter nei confronti dell'ex presidente Donald Trump.

L'8 gennaio 2021, Trump è stato “Sospeso in maniera permanente” da Twitter, con la motivazione che l'ex presidente avesse violato più volte le linee guida che vietano di condividere contenuti che promuovano la violenza e le notizie false.

L'8 gennaio 2021, il presidente Donald J. Trump ha twittato:

“I 75.000.000 di grandi patrioti americani che hanno votato per me, AMERICA FIRST e MAKE AMERICA GRAND AGAIN, avranno una VOCE GIGANTE nel futuro. Non saranno mancati di rispetto o trattati ingiustamente in alcun modo, forma o forma!!!”

Poco dopo, il presidente ha twittato:

"A tutti quelli che me lo hanno chiesto, non andrò all'inaugurazione il 20 gennaio".

A seguito delle tensioni allora in corso negli Stati Uniti, legate anche all'aumento della conversazione globale riguardo le persone che hanno preso d'assalto violentemente il Campidoglio il 6 gennaio 2021, i due Tweet in questione devono essere letti nel contesto più ampio di questi recenti avvenimenti, nel quale i comunicati dell'ex Presidente possono essere interpretati in diversi modi, fra cui anche l'incitamento alla violenza.

A causa delle tensioni in corso negli Stati Uniti e dell'aumento della conversazione globale riguardo alle persone che hanno preso d'assalto violentemente il Campidoglio il 6 gennaio 2021, questi due Tweet devono essere letti nel contesto di eventi più ampi nel paese e nei modi in cui le dichiarazioni del Presidente possono essere mobilitate da un pubblico diverso, anche per incitare alla violenza, nonché nel contesto del modello di comportamento di questo resoconto nelle ultime settimane.



**Figura 2.4** La comunicazione di Twitter Safety che annuncia la sospensione permanente dell'account di Donald Trump. **Fonte:** Twitter

“Abbiamo valutato i due Tweet sopra menzionati nell'ambito della nostra politica sulla glorificazione della violenza , che mira a prevenire la glorificazione della violenza che potrebbe ispirare altri a replicare atti violenti e abbiamo stabilito che è molto probabile che incoraggino e ispirino le persone a replicare gli atti criminali che hanno avuto luogo al Campidoglio degli Stati Uniti il 6 gennaio 2021.”<sup>10</sup>

Sono queste le dichiarazioni di Twitter, che divenne così il primo importante social network a rimuovere l'account di una figura così importante come Trump.

L'approccio all'interesse pubblico della piattaforma, come spiegato da Twitter stesso, esiste per consentire al pubblico di ascoltare direttamente i funzionari eletti e i leader mondiali.

Questo approccio è costruito sul principio secondo il quale le persone hanno il diritto di avere il potere di rendere conto allo scoperto.

Il social network tende inoltre a sottolineare come “da anni abbiamo chiarito che questi account non sono completamente al di sopra delle nostre regole e non possono utilizzare Twitter per incitare alla violenza, tra le altre cose.”

É con questo motivo che viene motivata la sospensione permanente dell'account di Trump.

### **La decisione delle altre piattaforme**

A seguito dell'intervento di Twitter anche altre piattaforme hanno agito in direzioni simili.

L'amministratore delegato di Facebook Mark Zuckerberg ha annunciato con un post di aver deciso di sospendere l'account di Trump «indefinitamente» e almeno per due settimane, da Facebook e Instagram. Nel suo messaggio Zuckerberg ha detto che “il rischio nel continuare a permettergli di usare i nostri servizi in questo momento è semplicemente troppo grande”.<sup>11</sup>

YouTube ha spiegato di aver rimosso il video in cui Trump incoraggia i propri sostenitori perché violava le regole sulla comunicazione delle notizie legate alle elezioni presidenziali.<sup>12</sup>

Anche le linee guida di TikTok, popolarissimo soprattutto tra gli adolescenti, prevedono che i contenuti che promuovono odio e violenza siano vietati. Per questa ragione, cliccando su

---

<sup>10</sup> Twitter, Permanent suspension of @realDonaldTrump . fonte: [blog.twitter.com/en\\_us/topics/company/2020/suspension](https://blog.twitter.com/en_us/topics/company/2020/suspension)

<sup>11</sup> Il Post: Facebook ha sospeso l'account di Donald Trump per almeno due settimane. [ilpost.it](https://www.ilpost.it/2021/01/06/facebook-sospeso-donald-trump/)

<sup>12</sup> TechCrunch: Facebook and YouTube remove Trump video calling extremists 'special'

hashtag come #stormthecapitol (“invadiamo il congresso”) e #patriotparty (“partito dei patrioti”) si viene reindirizzati alla pagina delle norme di sicurezza della piattaforma.

La rimozione degli account di Trump pone questioni molto rilevanti sulla libertà di opinione di un leader politico, e sull’affidare ai social network il compito di risolverle da soli. Diversi esperti ritengono<sup>13</sup> comunque che negli anni Trump abbia utilizzato i social network soprattutto per avvelenare il dibattito pubblico e incoraggiare odio e violenza nei confronti dei propri avversari, come del resto molti leader di estrema destra in giro per il mondo.

## 2.5 La propagazione di disinformazione inerente al Covid-19

L’epidemia di COVID-19 ha colpito anche i social, affetti da una eccessiva abbondanza di informazioni – alcune accurate e altre no – che ha reso e ancora rende difficile per le persone trovare fonti affidabili e una guida sicura al bisogno: “We’re not just fighting an epidemic; we’re fighting an *infodemic*”, ha detto il direttore generale dell’OMS Tedros Adhanom Ghebreyesus alla Conferenza sulla sicurezza tenutasi a Monaco il 15 febbraio scorso.

Molto prima che i funzionari del governo statunitense e gran parte del mondo ne fossero consapevoli, Mark Zuckerberg conosceva la pericolosa rapidità con cui il Covid-19 si stava diffondendo.

A metà gennaio avevo già cominciato a ricevere i resoconti di intelligence sulla diffusione globale del virus da parte di esperti di malattie infettive che lavoravano per la Chan Zuckerberg Initiative, società dello stesso CEO di Facebook e della moglie che persegue l’obiettivo di risolvere alcune delle sfide più difficili della società, dall’eradicazione delle malattie e il miglioramento dell’istruzione, fino alla risposta ai bisogni delle nostre comunità locali.<sup>14</sup> Facebook aveva il potenziale per svolgere un ruolo critico in quella che sembrava la prima pandemia globale dell’era di internet.

Il 26 gennaio Zuckerberg diede l’ordine ai suoi capi di dipartimento di mollare tutto il lavoro non necessario per prepararsi all’emergenza.<sup>15</sup>

Il team di partecipazione civica, che prima delle elezioni presidenziali si era occupato di

---

<sup>13</sup> NYTimes: Twitter Permanently Bans Trump, Capping Online Revolt

<sup>14</sup> Chanzuckerberg: società filantropica - fonte: chanzuckerberg.com/about/our-approach/

<sup>15</sup> S. Frenkel, C. Kang: “Facebook: L’inchiesta Finale” - Editore da Einaudi

sviluppare una serie di nuove funzionalità tra cui lo snodo centralizzato per condividere le informazioni sul voto, si impegnò ad organizzare la raccolta di informazione dal Centro per il controllo e la prevenzione delle malattie e dall'Organizzazione mondiale della sanità.

Per smentire teorie complottistiche sul Covid-19 sarebbero stati usati strumenti di fact-checking che avrebbero individuato le notizie false.

Quando il 30 gennaio l'OSM definì il Covid-19 un'emergenza sanitaria globale, da lì a poche ore la società di Palo Alto rese pubblico il suo piano per rimuovere notizie false dannose, comunicare informazioni autorevoli sul virus e fornire spazio pubblicitario illimitato affinché l'OSM e Centri per il controllo e la prevenzione delle malattie pubblicassero annunci di interesse pubblico sul nuovo coronavirus.

Facebook fu la prima tra le piattaforme a dare una risposta.

### **2.5.1 L'enorme portata dei social network sulla disinformazione sanitaria globale**

All'inizio della crisi, proprio l'Organizzazione mondiale della sanità (OMS) sollevò immediatamente il possibile scaturire di una "infodemia" di disinformazione durante la crisi sanitaria.<sup>16</sup> Gli stessi medici e infermieri in prima linea rimbalzarono l'allarme, evidenziando come la disinformazione sui social media avrebbe potuto rappresentare una minaccia per le vite di tutto il mondo.<sup>17</sup>

#### **Il ruolo dell'algoritmo di Facebook nell'amplificare la disinformazione sulla salute**

(Come già affrontato) L'algoritmo di Facebook decide il contenuto che gli utenti vedono nei propri feed di notizie e dove tale contenuto viene posizionato in base a un'ampia gamma di variabili e calcoli, come la quantità di reazioni e commenti ricevuti da un post, se l'utente ha mostrato interesse per un contenuto proveniente da specifici di gruppi o pagine e una serie di altri segnali.<sup>18</sup>

---

<sup>16</sup> un.org: UN tackles 'infodemic' of misinformation and cybercrime in COVID-19 crisis , United Nations, March 31, 2020.

<sup>17</sup> avaaz.org: Health Professionals Sound Alarm Over Social Media Infodemic , Avaaz, July 2020.

<sup>18</sup> facebook.com: How News Feed Works , Facebook.

Questo posizionamento nel feed delle notizie, che Facebook definisce “classifica”(ranking), è determinato dall' algoritmo e può fornire ad un contenuto un'amplificazione significativa.

Come ha evidenziato Mark Zuckerberg: "Uno dei maggiori problemi a cui i social network devono far fronte è che, se non controllate, le persone interagiranno in modo sproporzionato con contenuti più sensazionalistici e provocatori... Su larga scala tutto ciò può minare la qualità del discorso pubblico e portare alla polarizzazione".<sup>19</sup>

I contenuti disinformativi tema salute e Covid-19 sono appositamente sensazionalistici e provocatori, costruiti unicamente per ottenere un engagement significativo.

Questo, a sua volta, viene interpretato dall'algoritmo come un motivo per potenziare ulteriormente tale contenuto nel News Feed, creando un circolo vizioso in cui l'algoritmo fornisce in modo coerente e artificiale disinformazione sulla salute, ad esempio, ostacolando la diffusione di contenuti sanitari autorevoli all'interno dell'ecosistema di informazioni che presenta agli utenti di Facebook.

Non da sottovalutare poi l'incredibile numero di circa 125 milioni di account falsi<sup>20</sup> che Facebook ammette ancora attivi sulla piattaforma, i quali hanno un ruolo non indifferente nella distorsione dell'algoritmo in modi non rappresentativi degli utenti reali.<sup>21</sup>

Facebook ha dichiarato pubblicamente nel 2018, dopo lo scandalo Cambridge Analytica e la crisi della disinformazione che ha dovuto affrontare nel 2016, di aver ridisegnato il calcolo del ranking che seleziona cosa far vedere prima all'interno nel News Feed “in modo che le persone abbiano più opportunità di interagire con le persone a loro care”.<sup>22</sup>

Come passo in quella direzione, (e ammettendo indirettamente il ruolo dell'algoritmo nell'amplificare la disinformazione), Mark Zuckerberg ha anche annunciato che la piattaforma si sarebbe mossa per moderare l'algoritmo assicurando che "i post classificati come falsi vengano retrocessi e perdano in media l'80% delle loro visioni future".<sup>23</sup>

---

<sup>19</sup> facebook.com: Zuckerberg M, A Blueprint for Content Governance and Enforcement , Facebook, November 15, 2018.

<sup>20</sup> technologyreview.com: Hao K, How Facebook uses machine learning to detect fake accounts , MIT Technology Review, March 4, 2020.

<sup>21</sup> snopes.com: Matthews J, How Fake Accounts Constantly Manipulate What You See on Social Media , Snopes, June 24, 2020 (republished with permission from The Conversation).

<sup>22</sup> facebook.com: Zuckerberg M, Facebook post , January 12, 2018 ( <http://archive.vn/CyIzG> ).

<sup>23</sup> facebook.com Zuckerberg M, Preparing for Elections , Facebook, September 13, 2018.

## 2.5.2 I post che hanno minato la risposta al Covid-19

Di seguito viene riportato come esempio uno di quei post come rappresentate delle migliaia di contenuti virali falsi o fuorvianti che hanno minato i sistemi sanitari, gli ospedali e le linee guida di protezione durante la pandemia di COVID-19.

**“L’American Medical Association sta ora "incoraggiando" i medici a sovrastimare i decessi per coronavirus in tutto il paese” 160,5 milioni di visualizzazioni.**

Questo primo esempio è un articolo dal titolo: "Ospedali statunitensi pagati di più per etichettare la causa della morte come 'Coronavirus'".<sup>24</sup>

L'articolo presenta un video dell'intervista della conduttrice di Fox News Laura Ingraham con il medico del Minnesota e senatore dello stato repubblicano, il dottor Scott Jensen, e ha mal estrapolato quanto scaturito dall'intervista, affermando che "l'American Medical Association sta 'incoraggiando' i medici a sovrastimare i decessi per coronavirus in tutto il paese".



**Figura 2.5:** L'articolo di Global Research, pubblicato sulla loro pagina Facebook. **Fonte:** Facebook

<sup>24</sup> globalresearch.ca: Jensen S, Dupree W, US Hospitals Getting Paid More to Label Cause of Death as 'Coronavirus', Global Research, July 11, 2020. (<https://archive.md/UGdDh>)

In realtà il dottor Jensen non menziona mai l'American Medical Association nell'intervista.<sup>25</sup> Secondo un giornalista di FactCheck.org, "Jensen ha affermato di non ritenere che gli ospedali classificassero intenzionalmente in modo errato i casi per motivi finanziari ... Ma è così che le sue dichiarazioni sono state ampiamente interpretate e diffuse sui social media".<sup>26</sup>

L'articolo in esame è stato pubblicato da GlobalResearch.ca, un sito etichettato da NewsGuard<sup>27</sup> come assiduo pubblicante e promotore di contenuti falsi, teorie del complotto e propaganda pro-Russia, e poi ri-condiviso sulla stessa pagina Facebook di Global Research e su altre della stessa caratura.

Queste mosse hanno permesso all'articolo di ricevere più di 160,5 milioni di visualizzazioni stimate su Facebook.

Lead Stories, partner ufficiale di Facebook come fact-checker esterno<sup>28</sup>, si occupato dell'analisi dell'articolo, arrivando alla chiara conclusione che fosse un contenuto disinformante: gli ospedali non possono ottenere rimborsi ulteriori per i pazienti COVID-19 e non ci sono prove che i medici siano incentivati a mentire su un certificato di morte per aumentare il numero di decessi da coronavirus.<sup>29</sup>

È degno di nota il fatto che i post "kernel of truth" come questo siano spesso utilizzati dai disinformatori per attirare le persone verso contenuti di disinformazione, intervallando false affermazioni con informazioni verificate, per rendere il post più credibile.<sup>30</sup>

Questo tipo di contenuto falso e fuorviante rappresenta una tendenza ricorrente durante questa pandemia e richiede sforzi più sofisticati da parte dei fact-checker e di Facebook per garantire che i contenuti di disinformazione ben progettati vengano rilevati e smascherati efficacemente.

---

<sup>25</sup> leadstories.com: Jessica, Ravitz. Fact Check: Doctors Are NOT 'Encouraged' By Hospitals And AMA To Overcount COVID-19 Deaths , Lead Stories, May 13, 2020.

<sup>26</sup> Fichera A, Hospital Payments and the COVID-19 Death Count , Factcheck.org, April 21, 2020.

<sup>27</sup> newsguardtech.com: NewsGuard Nutrition Label for globalresearch.ca ( <http://archive.vn/WM6CY> ).

<sup>28</sup> facebook.com: Fact-Checking on Facebook: What Publishers Should Know , Facebook.

<sup>29</sup> leadstories.com: Ravitz J, Fact Check: Doctors Are NOT 'Encouraged' By Hospitals And AMA To Overcount COVID-19 Deaths , Lead Stories, May 13, 2020.

<sup>30</sup> NYTimes: .Ellick A, Westbrook A, Kessel J, The Seven Commandments of Fake News , New York Times (video).



### 2.5.3 Come Facebook può mettere in quarantena la disinformazione sulla salute

Un importante studio su come Facebook potrebbe intervenire sulla disinformazione è portato avanti da AVAAZ, organizzazione non governativa internazionale che promuove attivismo su tematiche quali il cambiamento climatico, i diritti umani, i diritti degli animali, la corruzione, la povertà e i conflitti.<sup>31</sup>

Analizzando questo studio viene espansa una soluzione in due fasi, la quale suggerisce come poter mettere in quarantena l'infodemia oggi presente sulla piattaforma, riducendo la credenza in contenuti falsi di quasi il 50% e la loro portata dell'80%.

Facebook come già indagato, non è una piattaforma neutrale, avendo il chiaro obiettivo di massimizzare il coinvolgimento degli utenti<sup>32</sup> con i contenuti e aumentare il tempo che gli utenti trascorrono sulla piattaforma<sup>33</sup> tramite un'algoritmo che decide quali e come mostrarli.

Per raggiungere tali risultati viene suggerito come Facebook dovrebbe iniziare ad implementare soluzioni sistemiche semplici, che si sono dimostrate efficaci nel ridurre al minimo la portata e l'impatto della disinformazione sulle piattaforme di social media, ovvero:

- **Correct the Record:** fornire a TUTTI gli utenti che hanno riscontrato disinformazione correzioni verificate in modo indipendente, riducendo così la credenza nella disinformazione di una media di quasi il 50%;
- **Detox the Algorithm:** declassamento dei post di disinformazione e degli attori sistematici della disinformazione nei feed di notizie degli utenti, diminuendo le loro visualizzazioni future di circa l'80%.

---

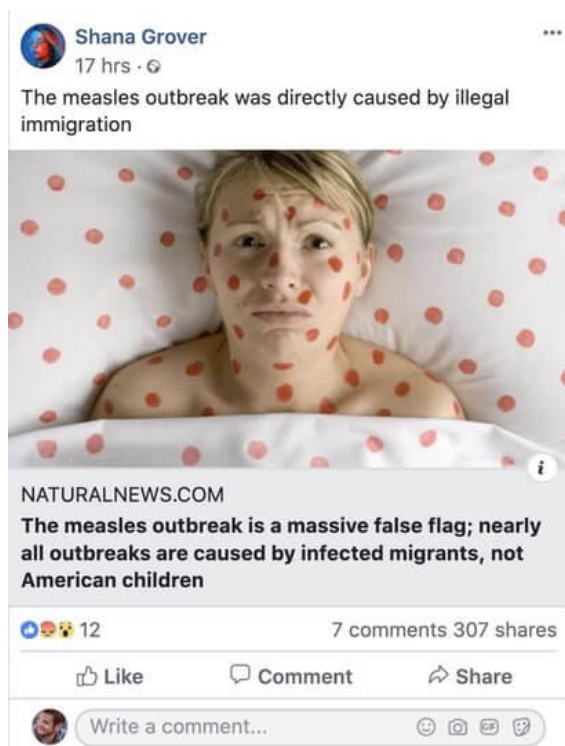
<sup>31</sup> [avaaz.org: https://secure.avaaz.org/page/it/](https://secure.avaaz.org/page/it/)

<sup>32</sup> knightfoundation: Bradshaw S, Howard P, Why Does Junk News Spread So Quickly Across Social Media? - Algorithms, Advertising And Exposure In Public Life, Knight Foundation, January 29, 2018.

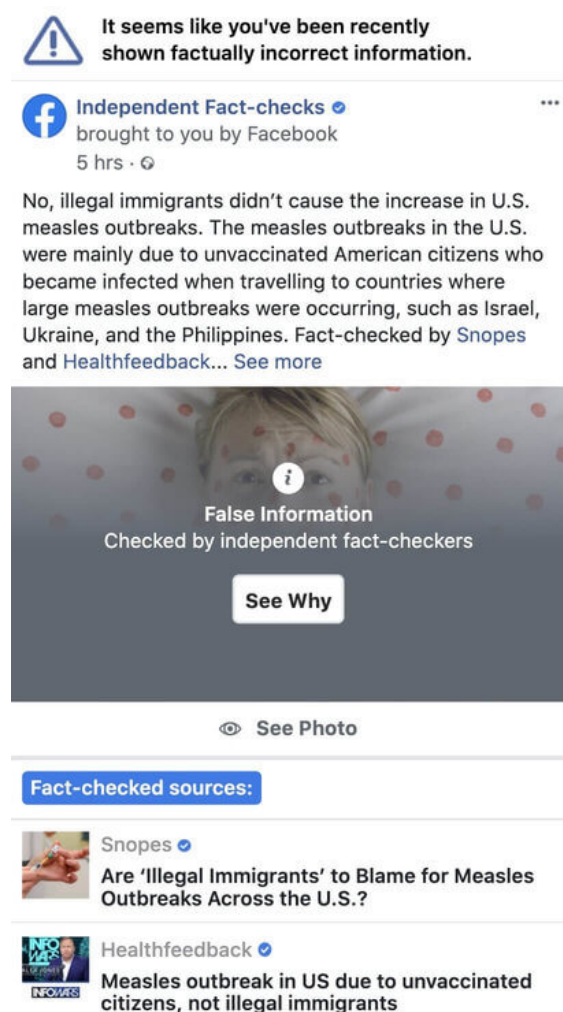
<sup>33</sup> Horwitz J, Seetharaman D, Facebook Executives Shut Down Efforts to Make the Site Less Divisive, the Wall Street Journal, May 26, 2020.

## 2.5.4 Correct the Record: tenere informati gli utenti

Il principio su cui si basa Correct the Record richiederebbe alle piattaforme di distribuire retroattivamente correzioni da fact-checker indipendenti a ogni singola persona esposta a informazioni false o fuorvianti. Facebook attualmente aggiunge solo etichette sui contenuti verificati, ma non torna indietro e fornisce correzioni retroattivamente alle centinaia di milioni di persone che hanno visto la falsità iniziale nei loro feed di notizie.



**Figura 2.6** Un esempio verosimile di post Facebook contenente disinformazioni. **Fonte:** Avaaz



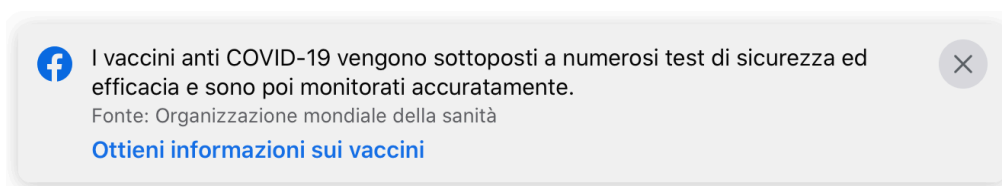
**Figura 2.6** Un contro esempio verosimile di post Facebook , dove la disinformazione viene contrassegnata con una grafica che rimanda l'utente a controllare la veridicità dei fatti su sito esterno indipendente. **Fonte:** Avaaz

## Correct the Record taglierebbe quasi della metà la credenza nella disinformazione

Un ulteriore studio accademico, commissionato da Avaaz e condotto dal dott. Ethan Porter della George Washington University e dal dott. Tom Wood della Ohio State University, ha scoperto che fornire agli utenti di social media che hanno visto false o informazioni fuorvianti con correzioni dei fact-checker possono ridurre la credenza nella disinformazione in media di quasi il 50% e fino al 61%.<sup>34</sup>

Ad aprile, in concomitanza con l'uscita del precedente rapporto di Avaaz sulla disinformazione sul COVID-19 su Facebook, la piattaforma ha annunciato che avrebbe condiviso un avviso generale che informava gli utenti che avevano interagito con una dannosa disinformazione sul COVID-19 di visitare il sito Web dell'OMS contro le fake news.<sup>35</sup>

Ma questo passaggio è limitato solo alla disinformazione dannosa sul COVID-19 e l'avviso fornito dalla piattaforma non è una correzione.



**Figura 2.7.** Avviso di Facebook quando l'argomento del contenuto è inerente al Covid-19.

**Fonte:** Facebook

Sondaggi in Germania, Francia, Spagna e Italia hanno mostrato che l'87% è favorevole a questa idea, concordando sul fatto che "piattaforme di social media come Facebook e Twitter dovrebbero lavorare con organizzazioni di verifica dei fatti indipendenti e affidabili per fornire a tutti gli utenti che sono stati esposti a un'ampia diffusione di contenuti falsi o fuorvianti con correzioni verificate".<sup>36</sup> Tra i cittadini britannici, l'84% è favorevole all'idea, così come il 68% degli americani.<sup>37</sup>

<sup>34</sup> Libro bianco: Correggere il record - Le correzioni come antidoto alla disinformazione , Avaaz, 16 aprile 2020.

<sup>35</sup> Rosen G, Un aggiornamento sul nostro lavoro per mantenere le persone informate e limitare la disinformazione su COVID-19 , Facebook, 16 aprile 2020.

<sup>36</sup> Sondaggio: l'87% delle persone in Germania, Francia, Spagna e Italia vuole che Facebook "corregga il record" per combattere le notizie false , Avaaz, 28 febbraio 2019.

<sup>37</sup> Libro bianco: Correggere il record - Le correzioni come antidoto alla disinformazione , Avaaz, 16 aprile 2020.

### **2.5.5 Detox the Algorithm: proteggere gli utenti**

Alla base della soluzione definita come “Detox the Algorithm” c’è la necessità da parte delle piattaforme di regolare in modo trasparente il proprio algoritmo, per garantire che questo riduca efficacemente la disinformazione nota, la misinformazione, nonché le pagine, i gruppi e i siti Web che diffondono sistematicamente notizie false e/o compromesse.

La sua attuazione, stando ad Avaaz, richiederebbe una trasparenza significativamente maggiore da parte delle dei social media su come funzionano i loro sistemi di raccomandazione e amplificazione.

È fondamentale per Facebook agire contro i disinformatori sistematici, lavorando con urgenza assieme ai fact-checker e alla società civile per definire politiche chiare e strumenti di trasparenza che consentano di identificare rapidamente i responsabili della diffusione della disinformazione sanitaria e non solo.

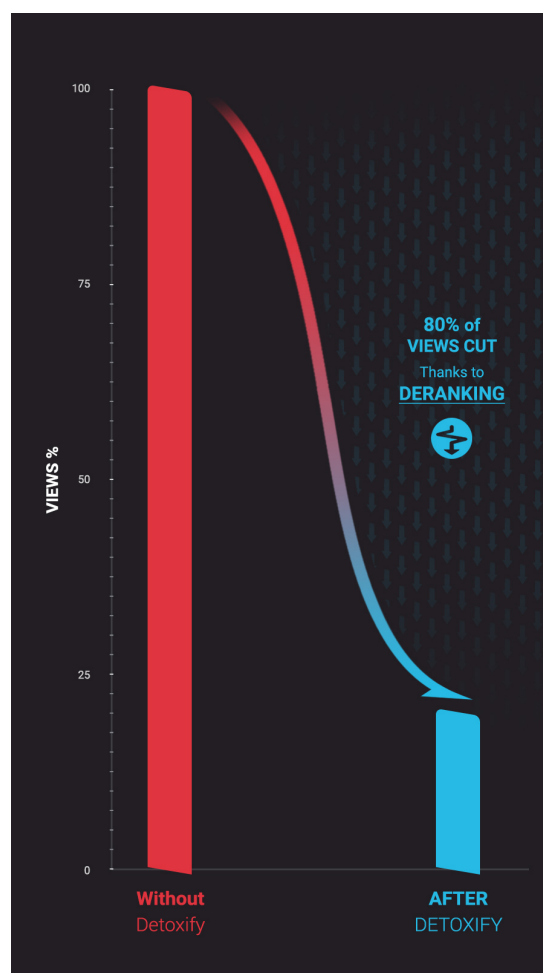
Questo sforzo deve portare ad una politica solida di declassamento della portata di questi propagatori regolari di disinformazione, dando però a coloro che sono stati segnalati per la condivisione di disinformazione un modo chiaro e rapido per fare appello, in modo da non generare disparità fra le parti.

Questo passaggio garantirà che gli attori disinformanti non continuino a diffondere contenuti falsi e fuorvianti su una scala che non può essere verificata, distorcendo la realtà sfruttando l’amplificazione permessa dall’odierna struttura dell’algoritmo.

**Detox the Algorithm (DTA)** può essere implementato in quattro fasi:

- **Rilevare e declassare i contenuti di disinformazione nota** : in primo luogo, le social media platforms devono garantire che tutti i contenuti identificati come disinformazione da fact-checker indipendenti vengano immediatamente declassati alla fine del feed delle notizie.
- **Rilevare e declassare gli attori della disinformazione sistematica** : in secondo luogo, i social media devono collaborare con verificatori di fatti indipendenti per rilevare e declassare pagine, gruppi e siti Web che diffondono ripetutamente e sistematicamente disinformazione o che hanno violato ripetutamente gli standard della comunità di Facebook nel tentativo di diffondere disinformazione, ingannare gli utenti o manipolare l'algoritmo.<sup>38</sup>

- **In situazioni di emergenza** , come una pandemia o un'elezione, Facebook dovrebbe avere team mirati che lavorano con la società civile per monitorare, rilevare e declassare gli attori sistematici della disinformazione, con soglie definite per quando una pagina/un gruppo diventa un disinformatore seriale.
- **Diritto di ricorso**: il declassamento degli attori della disinformazione è una sanzione grave e dovrebbe essere inflitta solo dopo che sono stati inviati avvisi a queste pagine, gruppi e siti web. Questi attori devono avere l'opportunità di adeguare il loro comportamento o di esercitare il loro diritto di ricorso se non sono d'accordo con una decisione di declassamento.



**Figura 2.8.** Riduzione delle visualizzazioni stimate degli attori della disinformazione sistematica attraverso l'implementazione di Detox the Algorithm.

**Fonte:** Avaaz

<sup>38</sup> Community Standards , Facebook.

- **Demonetizzare i disinformatori sistematici:** quando si scopre che un attore diffonde sistematicamente disinformazione, Facebook deve vietare a questi attori la pubblicità e la monetizzazione sulla piattaforma. Anche la demonetizzazione è una sanzione grave e dovrebbe essere imposta solo dopo che sono stati inviati avvisi a queste pagine, gruppi e siti web.
- **Informare gli utenti e tenerli al sicuro:** gli utenti che visualizzano o desiderano interagire con tali pagine, gruppi o siti Web devono essere informati tramite etichette chiare o altri mezzi che è stato riscontrato che questi attori diffondono ripetutamente e sistematicamente disinformazione e forniscono loro l'accesso a maggiori informazioni .

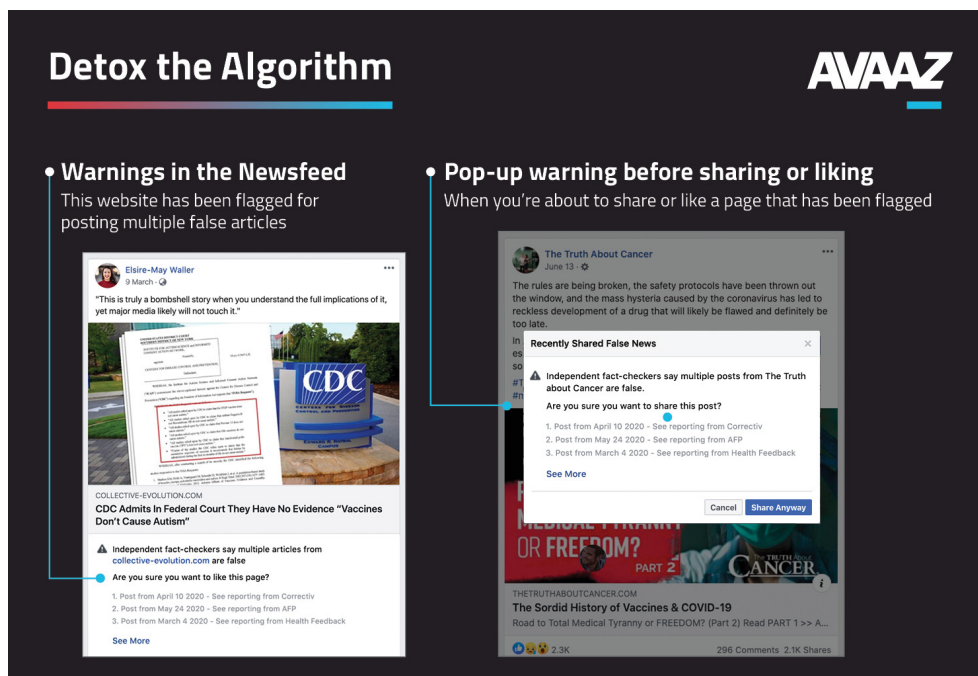


Figura 2.9. Due esempi visivi dell'implementazione di Detox the Algorithm.

Fonte: Avaaz

## **Detox the Algorithm preserva la libertà di espressione**

Con l'attuazione di Detox the Algorithm, le pagine di gruppi che diffondono disinformazione non verrebbero cancellati, né i siti web sarebbero banditi da Facebook. I contenuti di questi attori avrebbero una portata organica fortemente limitata, e non verrebbero così amplificati dall'algoritmo di Facebook.

Seguendo tale logica, con il tempo, questa politica contribuirà a garantire che il contenuto di disinformazione sia espanso in maniera molto limitata dall'algoritmo, dato che inizierebbe a imparare ad identificare e di conseguenza a non amplificare in modo massiccio i contenuti caratterizzati come disinformazione, emarginando così gli attori della disinformazione seriale invece di aiutarli a far crescere la loro lista di follower.

La soluzione preserva quindi i diritti di libertà di parola, non introducendo nessun filtro al momento della pubblicazione.

Non solo, potrebbe anche salvaguardare dalla violazione della libertà di pensiero e dei diritti di opinione date che ad oggi, il modo in cui l'algoritmo amplifica la disinformazione, inconsapevolmente intossicando e polarizzando un ecosistema informativo degli utenti, può interferire in modo significativo con la capacità degli utenti di formare opinioni indipendenti.

Nell'implementazione dell'algoritmo Detox the Algorithm, le pagine e i gruppi trovati come disinformatori seriali avrebbero l'opportunità di apportare correzioni ai loro utenti o contestare la decisione di disintossicazione se non sono d'accordo con essa.

Facebook dovrebbe anche assicurarsi che le pagine o i gruppi vengano avvisati dopo il primo e il secondo avvertimento, dando loro il tempo di cambiare comportamento o contestare la segnalazione.

Questa implementazione non sarebbe quindi in contrasto con le normative vigenti, viste nei precedenti, dato che queste tutelano sia la libertà personale degli utenti.

# Capitolo 3

## L'intervento delle Istituzioni per Contrastare la Diffusione di Informazioni Illecite

### 3.1 L'intervento delle Istituzioni Europee

Tutto ciò porta a chiedersi: Queste piattaforme sono rimaste quelle previste dalle normative o stanno assumendo un'altra configurazione?

Oltre che a livello di autoregolamentazione interna su come trattare la disinformazione, è necessario un intervento da parte delle istituzioni governative, veri garanti di principi come libertà di espressione e informazione.

A livello europeo se ne sta parlando nel Digital Services Act, una serie di proposte nelle quali la Commissione Europea si impegna ad aggiornare le normative che definiscono le responsabilità e gli obblighi dei prestatori di servizi digitali, in particolare delle piattaforme online.

Il rapido e diffuso sviluppo dei *digital services* è stato al centro dei cambiamenti digitali che influiscono sulle vite dei cittadini. Sono apparsi molti nuovi modi per comunicare, acquistare o accedere alle informazioni online, e sono in continua evoluzione. È necessario quindi che la legislazione europea evolva con loro.

Le piattaforme online hanno creato notevoli vantaggi per i consumatori, contribuendo inoltre a rendere maggiormente efficiente il mercato interno della stessa Unione.

Esse hanno inoltre facilitato gli scambi transfrontalieri all'interno e all'esterno dell'Unione. Tutto ciò ha aperto nuove opportunità ad una varietà di imprese e commercianti europei facilitandone l'espansione e l'accesso a nuovi mercati.

Sebbene vi sia quindi un ampio consenso sui benefici di questa trasformazione, come ogni innovazione anche quella delle piattaforme ha fatto sorgere diversi problemi che hanno numerose conseguenze per la società e per l'economia.

Preoccupazioni fondamentali riguardo il commercio e lo scambio di beni, servizi e contenuti illegali online nonché la diffusione della disinformazione e dei discorsi d'odio.



Attraverso lo strumento del regolamento, la Commissione mira ad ottenere un quadro normativo uniforme in tutti gli Stati membri dell'Unione relativamente agli ambiti considerati dalle proposte, in modo da favorire lo sviluppo del mercato unico digitale riducendo gli ostacoli causati dalle difformità legislative.

In particolare, uno degli aspetti trattati nel Digital Services Act (d'ora innanzi DSA) riguarda l'introduzione dell'obbligo, per prestatori di servizi di hosting, di agire immediatamente per rimuovere o oscurare i contenuti illeciti diffusi dagli utenti, non appena acquisita la conoscenza di tali attività o contenuti illeciti o la consapevolezza di tali fatti o circostanze.

Ciò fugge ogni dubbio sul fatto che il fornitore di servizi di hosting, se e quando la proposta di DSA entrerà in vigore, avrà l'obbligo di attivarsi con prontezza per contrastare la diffusione di contenuti illeciti, indipendentemente dalla modalità tramite le quali abbia acquisito conoscenza o consapevolezza della loro illiceità.

Tra l'altro, proprio per facilitare l'acquisizione di conoscenza dei contenuti illeciti da parte degli hosting provider, la proposta di regolamento introdurrebbe un'apposita procedura di notice-and-take-down di facile accesso e uso per gli utenti, i quali potranno così segnalare la presenza di contenuti illeciti nella piattaforma.

L'hosting provider, nella valutazione che dovrà compiere circa l'opportunità di rendere inaccessibili taluni contenuti, dovrà tener conto di un corpus normativo assai vasto di livello tanto europeo quanto nazionale, cosa che può rivelarsi piuttosto complicata, se si considera, da un lato, la natura transfrontaliera dell'attività di gran parte degli intermediari digitali e, dall'altro, la variabilità con cui le legislazioni nazionali tipizzano le varie categorie di illeciti.

Si pensi, ad esempio, ad alcune fattispecie di reato presenti nell'ordinamento giuridico italiano, ma non in quello di altri Stati membri dell'Unione europea: l'apologia del fascismo, cioè la propaganda di organizzazioni o movimenti di stampo fascista o l'esaltazione pubblica di esponenti, principi, metodi, fatti o finalità antidemocratiche del fascismo; oppure il cosiddetto reato di "negazionismo", che comprende la propaganda, l'istigazione e l'incitamento alla discriminazione o all'odio razziale, etnico o religioso, se commessi in modo che ne derivi un concreto pericolo di diffusione e fondati in tutto o in parte sulla negazione della Shoah o dei crimini di genocidio, dei crimini contro l'umanità e dei crimini di guerra.

Il provider, dunque, dovrebbe conoscere la normativa in vigore nei diversi Stati membri in cui opera e adottare policies più o meno restrittive a seconda del contesto nazionale di riferimento. È evidente come questo possa complicare non poco l'attività degli intermediari digitali, oltre al fatto che una disparità di trattamento per il medesimo contenuto, giustificata in base alla diversità della legislazione nazionale applicabile, potrebbe determinare l'insorgere di controversie fra operatori digitali, da un lato, e utenti o autorità nazionali, dall'altro. Proprio per ovviare a questo tipo di problemi, l'ordinamento giuridico dell'Unione sta mostrando una iniziale tendenza a raggiungere un certo livello di armonizzazione normativa fra tutti gli Stati membri almeno in relazione ad alcuni tipi di contenuti illeciti online considerati particolarmente gravi o frequenti, attraverso l'emanazione di una legislazione settoriale a livello europeo che imponga ai provider l'adozione di misure di contrasto alla loro diffusione.

### **Il contrasto alla disinformazione**

Un discorso a parte va fatto per la disinformazione, che è un tema rispetto al quale l'Unione europea è particolarmente sensibile, pur mostrando notevoli incertezze e difficoltà rispetto alla sua regolamentazione. La disinformazione, infatti, non mina soltanto il diritto delle persone ad acquisire informazioni complete, veritiere ed attendibili sulla realtà che le circonda, potendo così compiere delle scelte avvedute e consapevoli, ma costituisce anche un significativo strumento di influenza politica di cui ci si può deliberatamente servire per minare le basi democratiche della società europea, tanto più che la risonanza dei contenuti di disinformazione è amplificata attraverso la rete Internet anche mediante il frequente ricorso a strumenti automatizzati (bot) per la loro propagazione.

Pur avendo acquisito negli ultimi anni piena consapevolezza della gravità del fenomeno, come dimostrano numerose comunicazioni che la Commissione europea ha adottato dal 2015 in avanti relativamente alla disinformazione online, sul piano regolatorio l'Unione Europea stenta ad adottare norme cogenti, data l'obiettivo difficoltà di trovare un efficace punto di bilanciamento fra il diritto individuale alla libera manifestazione del pensiero, caposaldo di ogni ordinamento democratico, e l'esigenza di proteggere i sistemi democratici dagli attacchi destabilizzanti provenienti dalla deliberata organizzazione di massicce campagne di disinformazione.

Nel rapporto pubblicato a gennaio 2018 dal gruppo di esperti di alto livello sulle notizie false e la disinformazione, istituito dalla Commissione europea, la disinformazione è stata definita come informazione falsa, imprecisa o fuorviante, concepita, presentata e diffusa a scopo di lucro o per ingannare intenzionalmente il pubblico, suscettibile di arrecare un pregiudizio pubblico. Quindi, gli elementi costitutivi della fattispecie sono la consapevolezza della falsità del contenuto e l'intenzionalità della condotta diffusiva, congiuntamente al fine di profitto o alternativamente di danno.

La definizione esclude così tutti i contenuti inattendibili che gli utenti di Internet diffondono, anche intenzionalmente, o con intenti giocosi o perché convinti in buona fede della verità dell'informazione. In ogni caso, si tratta di una definizione che, non essendo recepita all'interno di un atto normativo, non ha alcuna efficacia giuridica vincolante.

Come è noto, la Commissione europea per il momento ha incardinato la strategia di contrasto alla disinformazione sul sistema di autodisciplina previsto dal Code of Practice on Disinformation, un codice di buone pratiche per contrastare la disinformazione, adottato a settembre 2018, cui i prestatori di servizi via Internet aderiscono spontaneamente.

Per il momento vi hanno aderito i maggiori player della Rete, ma ne restano fuori moltissimi operatori di minori dimensioni. Certamente il Codice ha il merito di incentivare i provider a prevedere nelle proprie policies efficaci misure di contrasto alla disinformazione, ma lascia a ciascuno di essi la massima libertà di determinare i criteri in base ai quali i contenuti di disinformazione vengono qualificati e le procedure e modalità con cui essi vengono rimossi o bloccati, nonché l'eventuale possibilità per l'utente di opporsi alla decisione inhibitoria del provider. Dunque, posto che finora nessun atto normativo vincolante prodotto a livello di Unione europea ha stabilito i parametri in base ai quali i contenuti di disinformazione possano essere qualificati come illeciti, è assai dubbio che l'entrata in vigore del regolamento DSA possa comportare un obbligo di contrasto nei loro confronti in capo ai prestatori dei servizi di intermediazione digitale. Più probabilmente, nel campo della disinformazione continuerà ad applicarsi ancora per molto tempo l'attuale sistema di autodisciplina volontaria, fatta eccezione per i pochi Stati membri dell'UE che hanno adottato almeno alcune regole cogenti in tale ambito.

## 3.2 Il Dibattito Statunitense sulla Sezione 230

La situazione statunitense sente invece la mancanza di una chiara leadership e strategia per il problema della disinformazione, nonostante il suo stesso riconoscimento dell'impatto sulla salute pubblica, le elezioni, le imprese, la tecnologia e le continue campagne sulle comunità di colore, compresi gli immigrati e rifugiati.

Nel febbraio 2020 il Dipartimento di Giustizia degli Stati Uniti ha tenuto un seminario relativo alla Sezione 230 nell'ambito di un'indagine antitrust in corso sulle società "big tech".

Durante questo workshop è stato evidenziato come la sezione 230 fosse in qualche modo non aggiornata alla nuova situazione tecnologica globale.

Il procuratore generale William Barr, capo del Dipartimento di Giustizia, ha sottolineato la questione affermando come la norma fosse nata per la necessità di proteggere la crescita di Internet in un momento di incertezza economica delle sue imprese, ma di come ora queste "Non sono più le le parvenu perdenti ... Sono diventate titani dell'industria statunitense" ed ha quindi sollevato l'eventuale necessità di un aggiornamento della norma.

Per i critici dell'industria tecnologica, la Sezione 230 del Communications Decency Act è diventata il simbolo di un trattamento eccezionale da parte del governo che ha alimentato la crescita di un piccolo numero di player digitali.

Per le aziende tecnologiche invece, la legge fissa i valori fondanti di aperta e libera espressione di Internet, ed è propria grazie ad essa che queste hanno anche la possibilità di rimuovere i discorsi più insidiosi senza inciampare in un campo minato legale.

"La sezione 230 è stata interpretata in modo abbastanza ampio dai tribunali", ha affermato Barr. "Oggi, molti sono preoccupati che l'immunità della Sezione 230 sia stata estesa ben oltre quanto originariamente previsto dal Congresso. Ironia della sorte, la Sezione ha consentito alle aziende tech di assolversi completamente dalla responsabilità di controllare le proprie piattaforme, bloccando o rimuovendo il discorso di terze parti, incluso il discorso politico".<sup>39</sup>

---

<sup>39</sup> justive.gov: Attorney General William P. Barr Delivers Remarks at the National Association of Attorneys General 2019 Capital Forum, The United State Department Of Justice

Questo seminario si è concluso con l'emanazione di quattro raccomandazioni principali al Congresso per modificare la Sezione 230. Queste includono:

- I. Incentivare le piattaforme a gestire i contenuti illeciti, incluso il richiamo ai "cattivi samaritani" che sollecitano attività illecite.
- II. Rimozione delle tutele dalle cause civili promosse dal governo federale;
- III. La disabilitazione delle tutele dell'art. 230 in relazione ad azioni antitrust sulle grandi piattaforme Internet;
- IV. Promuovere il discorso e la trasparenza definendo termini esistenti nello statuto come "altrimenti discutibili" e "buona fede" con un linguaggio specifico e richiedendo alle piattaforme di documentare pubblicamente quando intraprendono azioni di moderazione contro i contenuti, a meno che ciò non possa interferire con le forze dell'ordine o rischiare di danneggiare un individuo.

In risposta a quanto scaturito dal workshop, vi è una proposta<sup>40</sup> nata proprio da parte di Mark Zuckerberg, la quale prevederebbe di aumentare, per le società di social media a cui è attualmente concessa l'immunità, la responsabilità per i contenuti che appaiono sulle loro piattaforme.

Il CEO di Facebook suggerirebbe, al limitare dell'immunità di cui le piattaforme godono, di richiedere invece a queste di dimostrare di avere tutti i sistemi che permettano di identificare e rimuovere i contenuti illeciti.

Sarebbe infatti ben poco pratico per piattaforme, dove all'interno vengono caricati milioni di contenuti ogni giorno, di essere ritenute responsabili se uno di questi in particolare sfugge al loro controllo.

Secondo Zuckerberg questo sistema dovrebbe essere proporzionato alla dimensioni della piattaforma, riconoscendo così che le aziende più piccole non possano essere in grado di sviluppare la stessa infrastruttura di Facebook e Google.

L'aggiornamento più recente in ambito "Sezione 230" è avvenuto ad ottobre 2021 con l'introduzione del "Justice Against Malicious Algorithms Act", un disegno di legge che eliminerebbe le protezioni derivanti dalla Sezione per i fornitori di servizi relativi agli

---

<sup>40</sup> NBCNews: Zuckerberg calls for changes to tech's Section 230 protections. 2021

algoritmi di raccomandazione personalizzati, nel caso in cui presentano contenuti che forniscono consapevolmente o incautamente informazioni che contribuiscono a lesioni fisiche o emotive gravi.

Rimane comunque estremamente aperto e controverso il dibattito attorno a questa normativa, sono molte infatti le preoccupazioni secondo cui, qualsiasi tentativo di indebolimento della Sezione 230, potrebbe causare un effettivo aumento della censura, derivato dalla perdita dell'esenzione di responsabilità da parte delle piattaforme.

## Conclusioni

Come visto finora, le soluzioni elaborate al fine di arginare fenomeni di disinformazione al momento sono essenzialmente per lo più in codici di autoregolamentazione, “regole d’uso” delle varie piattaforme, meccanismi di segnalazione di messaggi lesivi da parte degli utenti, individuazione e rimozione di contenuti e account falsi da parte dei gestori delle piattaforme stesse.

Questa soluzione, nella quale vi è una distorsione nella visione del mondo e della realtà, dettata da algoritmi che premiano il volume delle interazioni, e non la qualità delle informazioni, non può reputarsi ottimale.

Vi è sostanzialmente una delega di funzioni e responsabilità spettanti alle istituzioni competenti affidate alle piattaforme tecnologiche.

È presente inoltre una mancanza di definizioni certe delle categorie di contenuti da cancellare e di adeguati strumenti di impugnazione delle decisioni da parte degli utenti che ne sono oggetto. Importante infine considerare come il potere di rimozione di determinati contenuti determina il rischio della cancellazione di elementi di indagine.

Le limitazioni alla libertà di espressione, giustificate da esigenze di contrasto alla disinformazione o di repressione dell’odio online, potrebbero portare a soffocare espressioni di dissenso, attraverso accuse verso le opposizioni politiche di fomentare l’odio sociale.

Soprattutto, l’eliminazione di messaggi di hate speech dai social rappresenta solo un rimedio palliativo, che non previene né ostacola la formazione nell’opinione pubblica di idee basate su intolleranza e discriminazione, ma semplicemente ne impedisce o tenta di impedirne la diffusione attraverso internet, spostandone la manifestazione altrove, ove forse essa è meno visibile.

Questi fenomeni che si stanno manifestando in rete vanno, pertanto, affrontati non solo in funzione delle particolarità del mezzo – il web – mediante cui vengono esternati, ma soprattutto attraverso l’analisi delle cause sottostanti.

Così come il contrasto allo hate speech, anche quello alle fake news e, più in generale, alla disinformazione richiede la promozione di politiche volte all’educazione e responsabilizzazione dei cittadini, nonché un’azione tesa ad alimentare il pluralismo informativo.

Su queste basi, la presenza online di notizie controverse unitamente ad altre in grado di confutarle può avviare una discussione pubblica in grado di portare a una crescita sociale che renda più critici rispetto a ciò che si legge, più esigenti della prova e della provenienza delle notizie, più autosufficienti nel pensiero.

La soluzione, pertanto, non può trovarsi nella limitazione dei contenuti sul web, bensì nel suo opposto: cioè nel favorire diverse e credibili fonti di informazione, nel promuovere l'alfabetizzazione all'uso dei media e del digitale e diffondere – anche a livello governativo – informazioni affidabili sulle materie di pubblico interesse. Servirebbero, altresì, regole più chiare e maggiore trasparenza dei social media sul marketing politico realizzato per il loro tramite, specie sulla sua fonte, applicando ad esso gli stessi standard delle altre pubblicità, incluso un trattamento grafico che lo renda più riconoscibile.

In conclusione, trasparenza da parte dei gestori delle piattaforme, cultura della legalità, pluralismo, alfabetizzazione all'informazione, educazione alla conoscenza e percorsi di sviluppo delle capacità cognitive, non censure tecniche e normative, possono rappresentare gli strumenti più adeguati per contrastare la propagazione sistematica della disinformazione online.



## Bibliografia

R, MCNAMEE (2018) Zucked. Come aprire gli occhi sulla catastrofe di Facebook. Nutrimenti.

S. FRENKEL, C. KANG(2021) Facebook: L'inchiesta Finale. Torino, Einaudi.

B. KAISER (2019) La dittatura dei dati. Milano, HarperCollins

C. CEPERNICH (2017). Le campagne elettorali al tempo della networked politics. Urbino, Edizioni Laterza.

C. BIANCHI (2021). Hate Speech. Il lato oscuro del linguaggio. Bari, Edizioni Laterza.

## Sitografia

AVAAZ, Health Professionals Sound Alarm Over Social Media Infodemic: [avaaz.org](https://avaaz.org)

AVAAZ, Avaaz: <https://secure.avaaz.org/page/it/>

AVAAZ, Sondaggio: l'87% delle persone in Germania, Francia, Spagna e Italia vuole che Facebook "corregga il record" per combattere le notizie false: [avaaz.org](https://avaaz.org)

AVAAZ, Libro bianco: Correggere il record - Le correzioni come antidoto alla disinformazione: [avaaz.org](https://avaaz.org)

CHANZUCKERBERG, Chanzuckerberg: [chanzuckerberg.com/about/our-approach/](https://chanzuckerberg.com/about/our-approach/)

C-SPAN, C-Span: <https://www.c-span.org>

FACEBOOK, Fact-Checking on Facebook: What Publishers Should Know: <https://www.facebook.com/journalismproject/programs/third-party-fact-checking/publisher-faqs>

FACEBOOK, How News Feed Works: <https://www.facebook.com/help/1155510281178725>

FACEBOOK, Zuckerberg M, A Blueprint for Content Governance and Enforcement: <https://www.facebook.com/notes/751449002072082/>

FACEBOOK, 2018 Preparing for Elections ( <http://archive.vn/CyIzG> )

FACEBOOK, An Update on Our Work to Keep People Informed and Limit Misinformation About COVID-19 (<https://about.fb.com/news/2020/04/covid-19-misinfo-update/>)

FACEBOOK, Community Standards: <https://transparency.fb.com/it-it/policies/community-standards/>

FACTCHECK, Hospital Payments and the COVID-19 Death Count: <https://www.factcheck.org/2020/04/hospital-payments-and-the-covid-19-death-count/>

INTERNETASSOCIATION, Economic Value of Internet Intermediaries and the Role of Liability Protections: <https://internetassociation.org/wp-content/uploads/2017/06/Economic-Value-of-Internet-Intermediaries-the-Role-of-Liability-Protections.pdf>

LEADSTORIES, Doctors Are NOT 'Encouraged' By Hospitals And AMA To Overcount COVID-19 Deaths: <https://leadstories.com/hoax-alert/2020/05/fact-check-doctors-are-not-encouraged-to-overcount-COVID19-deaths.html>

NBC, News At South Carolina Rally, Donald Trump Defiant on Muslim Ban: <https://www.nbcnews.com/politics/2016-election/south-carolina-rally-trump-defiant-steadfast-muslim-ban-n475951>

NBCNEWS, Zuckerberg calls for changes to tech's Section 230 protections: <https://www.nbcnews.com/tech/tech-news/zuckerberg-calls-changes-techs-section-230-protections-rcna486>

NEWSGUARDTECH, NewsGuard Nutrition Label for globalresearch.ca: <https://www.newsguardtech.com/special-reports/coronavirus-misinformation-tracking-center/>

NY TIMES, Twitter Permanently Bans Trump, Capping Online Revolt: <https://www.nytimes.com/2021/01/08/technology/twitter-trump-suspended.html>

NY TIMES, Ellick A, Westbrook A, Kessel J, The Seven Commandments of Fake News: <https://www.nytimes.com/video/Adam-B-Ellick>

TWITTER, Permanent suspension of @realDonaldTrump: [blog.twitter.com/en\\_us/topics/company/2020/suspension](https://blog.twitter.com/en_us/topics/company/2020/suspension)

IL POST, Facebook ha sospeso l'account di Donald Trump per almeno due settimane: <https://www.ilpost.it/2021/01/07/facebook-sospeso-account-donald-trump-due-settimane/>

JUSTICE, Attorney General William P. Barr Delivers Remarks at the National Association of Attorneys General 2019 Capital Forum: <https://www.justice.gov/opa/speech/attorney-general-william-p-barr-delivers-remarks-national-association-attorneys-general>

KNIGHT FOUNDATION, Why Does Junk News Spread So Quickly Across Social Media: [https://kf-site-production.s3.amazonaws.com/media\\_elements/files/000/000/142/original/Topos\\_KF\\_White-Paper\\_Howard\\_V1\\_ado.pdf](https://kf-site-production.s3.amazonaws.com/media_elements/files/000/000/142/original/Topos_KF_White-Paper_Howard_V1_ado.pdf)

SNOPEs, How Fake Accounts Constantly Manipulate What You See on Social Media: <https://www.snopes.com/news/2020/06/24/how-fake-accounts-constantly-manipulate-what-you-see-on-social-media/>

TECH CRUNCH, Facebook and YouTube remove Trump video calling extremists ‘special’: <https://techcrunch.com/2021/01/06/facebook-and-youtube-remove-trump-video-calling-extremists-special/>

TECHNOLOGY REVIEW, Hao K, How Facebook uses machine learning to detect fake accounts: <https://www.technologyreview.com/2020/03/04/905551/how-facebook-uses-machine-learning-to-detect-fake-accounts/>

UN, tackles ‘infodemic’ of misinformation and cybercrime in COVID-19 crisis: <https://www.un.org/en/un-coronavirus-communications-team/un-tackling-‘infodemic’-misinformation-and-cybercrime-covid-19>

USCODE, Protection for private blocking and screening of offensive material: [https://uscode.house.gov/view.xhtml?req=\(title:47%20section:230%20edition:prelim\)\)](https://uscode.house.gov/view.xhtml?req=(title:47%20section:230%20edition:prelim)))

WALL STREET JOURNAL, Facebook Executives Shut Down Efforts to Make the Site Less Divisive: <https://www.wsj.com/articles/facebook-knows-it-encourages-division-top-executives-nixed-solutions-11590507499>

WIKIPEDIA, Propaganda: <https://it.wikipedia.org/wiki/Propaganda>