

Hate speech e social media

A che punto siamo e come difendersi

Dr. Eleonora Benecchi
Istituto Media e Giornalismo
Università della Svizzera italiana

Hate Speech online

Cosa intendiamo davvero



Linguaggio
abusivo

ESPLICITO

Hate Speech non nasce online

Hate Speech è:

un linguaggio che *attacca* o *sminuisce*, che incita alla violenza o all'odio nei confronti di *gruppi*, sulla base di *caratteristiche specifiche* come l'aspetto fisico, la religione, discendenza, origine nazionale o etnica, orientamento sessuale, identità di genere o altro, e può verificarsi con *stili linguistici diversi*.

(Fortuna e Nunes 2018, p. 5)

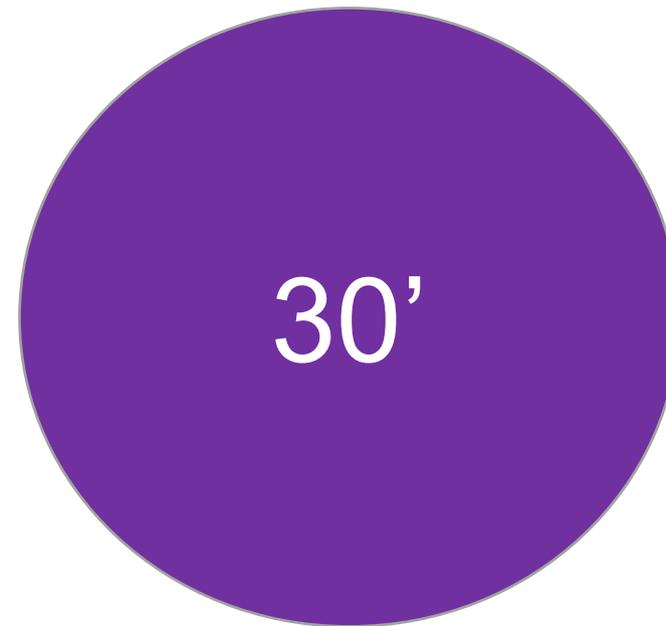
Hate Speech online

Questioni di gender

Questioni di gender



**Gruppo target:
donne**



Dati Stop Hate Speech

Ogni 30 secondi viene pubblicato un tweet aggressivo o offensivo nei confronti di una donna.

Questioni di gender



**Gruppo target:
donne**

62.35%

Dati Stop Hate Speech

Il 62,35% delle donne che usano Twitter afferma di essere stata aggredita sulla piattaforma.

Questioni di gender



**Gruppo target:
donne**



Dati Stop Hate Speech

Il 26% delle donne tra i 18 e i 24 anni ha subito molestie online.



Gruppo target:
donne

Questioni di gender

Ti è già successo...

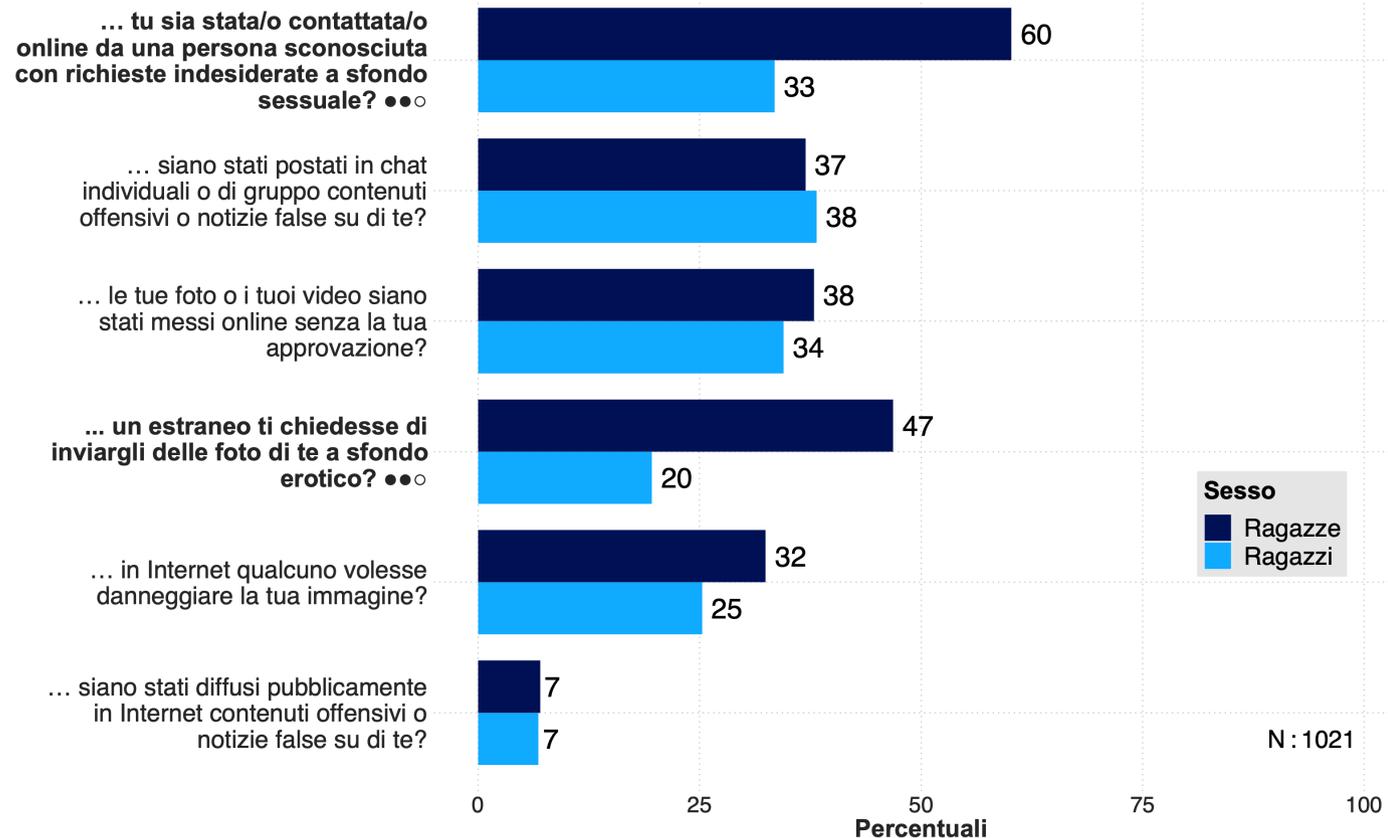


Figura 51: Cybermobbing e molestie sessuali in base al sesso



Gruppo target:
donne

Questioni di gender

Già prima dei social media gran parte delle molestie informatiche erano, in realtà, forme di **molestie di genere**.

Barak (2005)

Sui social le donne hanno maggiori probabilità di essere prese di mira da un **vocabolario fortemente connotato in termini di genere** (invece che da insulti di genere neutro).

(Megarry 2014, 50).



Gruppo target: donne

Questioni di gender

«Nell'ultimo anno aumentano percentualmente i messaggi di **odio, di disprezzo e di insulto** nei confronti delle **donne**.

Si colpisce la donna, perché **la donna che lavora e ha un ruolo è un concorrente**. Lo è sempre stata dal punto di vista del maschilismo e lo è a maggior ragione oggi, in un momento in cui le ben note vicende della pandemia hanno messo in difficoltà l'economia e le famiglie.

La paura della concorrenza e del ruolo si esprime nel **negare la professionalità e la parità** e lo si fa anche con un certo linguaggio».

(Progetto SPID, 2022).

Questioni di gender



Gruppo target:
donne

L'**imbarbarimento verbale** del linguaggio sui social media può avere un effetto di **normalizzazione** della **misoginia** e del razzismo, **spostando i confini normativi** di ciò che è consentito dire pubblicamente e ciò che non lo è.

Hate speech online

Le caratteristiche

Hate Speech online: le caratteristiche

L'anonimato spinge le persone a essere **più violente** a parole rispetto alle situazioni di compresenza fisica

(Poland, 2016: 22-24).

L'anonimato percepito può **eliminare la paura** di essere ritenuti **responsabili** di hate speech e può anche dare la sensazione che le **normali regole** di comportamento **non si applichino**.

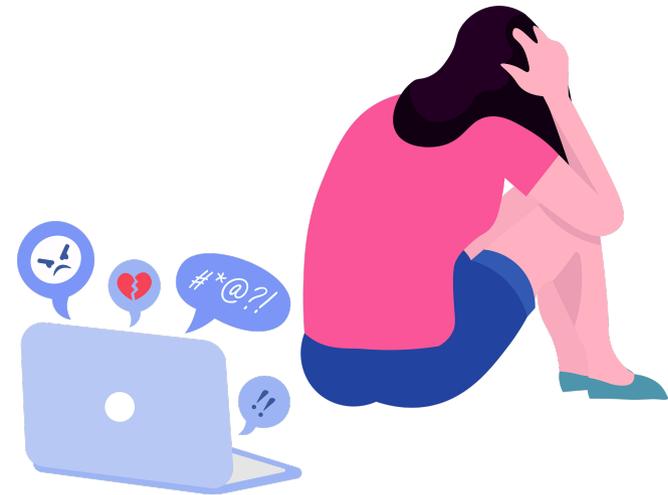
Poiché chi incita all'odio online **non è fisicamente presente**, non deve preoccuparsi di una reazione fisica.

(Delgado e Stefancic, 2014: 322)



ANONIMATO

Hate Speech online: le caratteristiche



La comunicazione online spesso significa che l'impatto immediato degli atti di parola **non è visibile all'autore**.

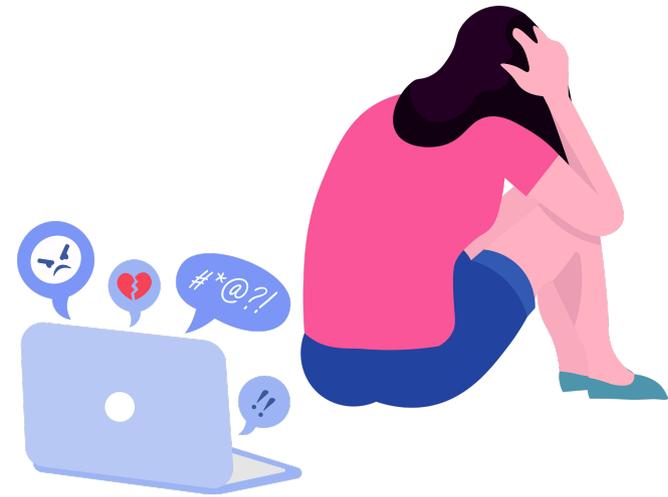
Se non si può vedere la **ferita emotiva** provocata dal proprio discorso d'odio online, è più probabile che se ne **minimizzi l'importanza**.

Quando è online, chi ricorre a hate speech **non può vedere** i volti di **altre persone** che potrebbero **disapprovare** ciò che dice.

(Citron, 2014: 59)

INVISIBILITÀ

Hate Speech online: le caratteristiche



Internet **aiuta a riunire le persone**, ogni tipo di persona. Permette a persone che altrimenti si sentirebbero isolate, o che si dedicherebbero ad altro, di **riunirsi e diffondere le loro voci**, molte delle quali paranoiche e **cariche di odio**

(Sunstein 2007: 57-58)

I gruppi d'odio usano Internet per **attrarre nuovi membri** e garantire che i membri esistenti **si sentano connessi**

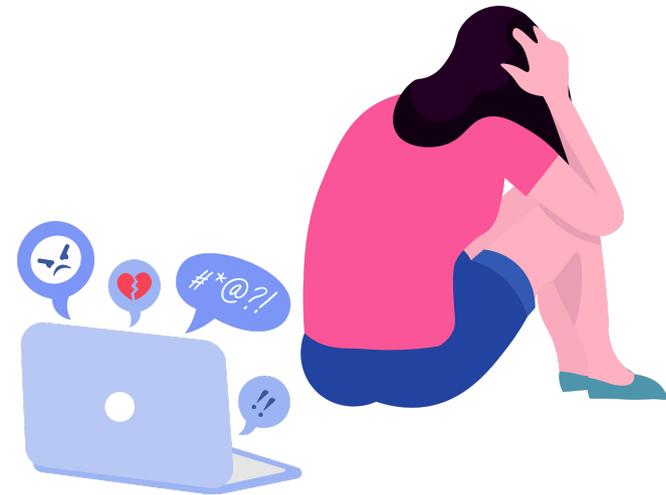
(Citron, 2014: 61-62)

In questo caso, i discorsi d'odio non sono rivolti alle vittime, ma **a persone che la pensano come loro**.

(Cohen-Almagor, 2015).

COMUNITÀ

Hate Speech online: le caratteristiche



Su Internet, il **tempo** che intercorre **tra l'averne un pensiero o un sentimento e l'esprimerlo** a un particolare individuo che si trova a grande distanza, o a un gruppo di persone o a un pubblico di massa, può essere una **questione di secondi**.

Internet incoraggia forme di hate speech che sono **spontanee**, nel senso di reazioni **di pancia**, giudizi **non ponderati**, osservazioni **fuori dagli schemi**, commenti **non filtrati**.

"Twitta in fretta, pentiti con calma". Potrebbe essere la morale di un'alta percentuale di discorsi d'odio online.

ISTANTANEITÀ

Hate Speech online: le caratteristiche

INVISIBILITÀ

ISTANTANEITÀ

ANONIMATO

COMUNITÀ

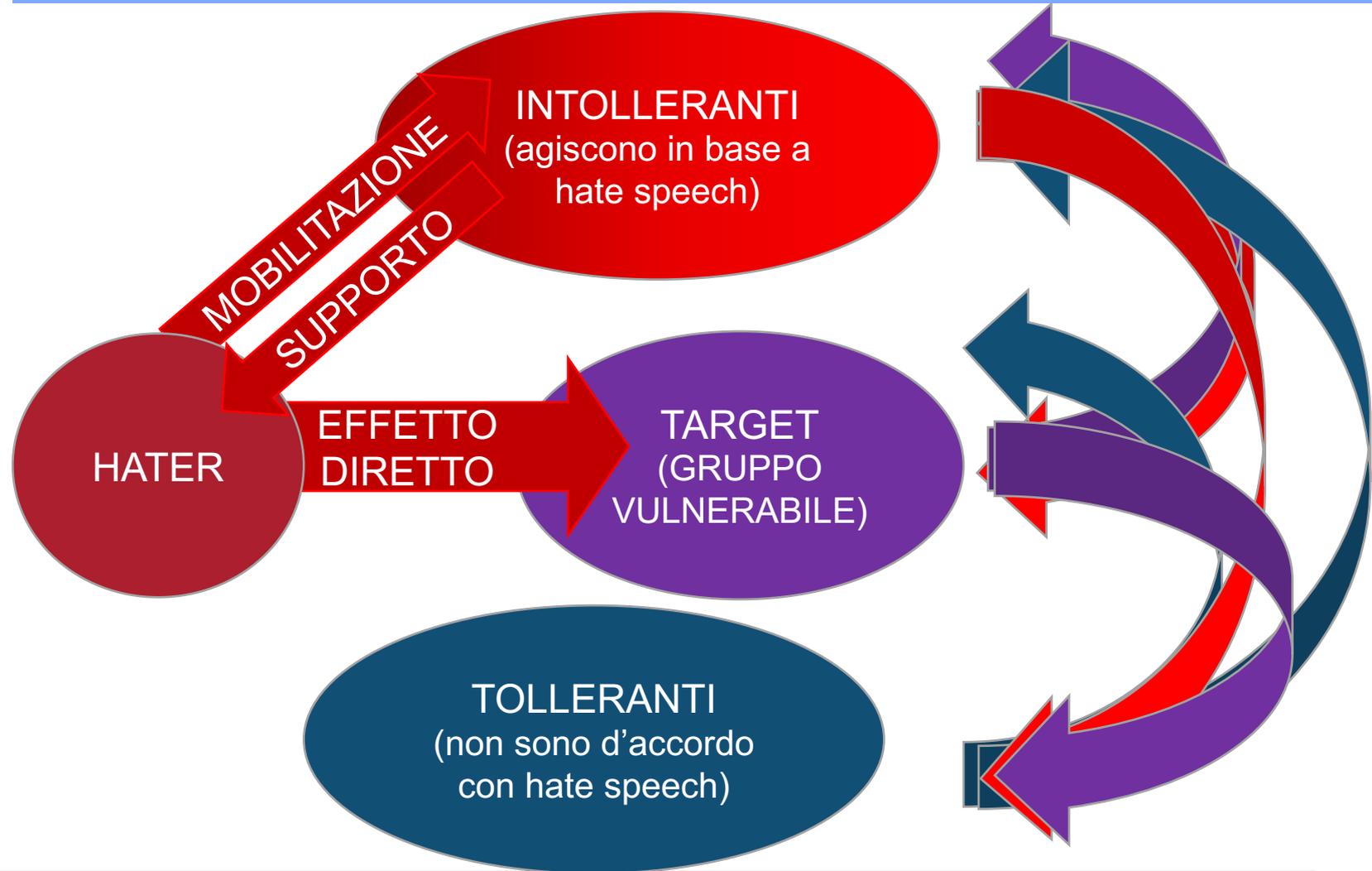
Hate Speech online

Lo schema comunicativo

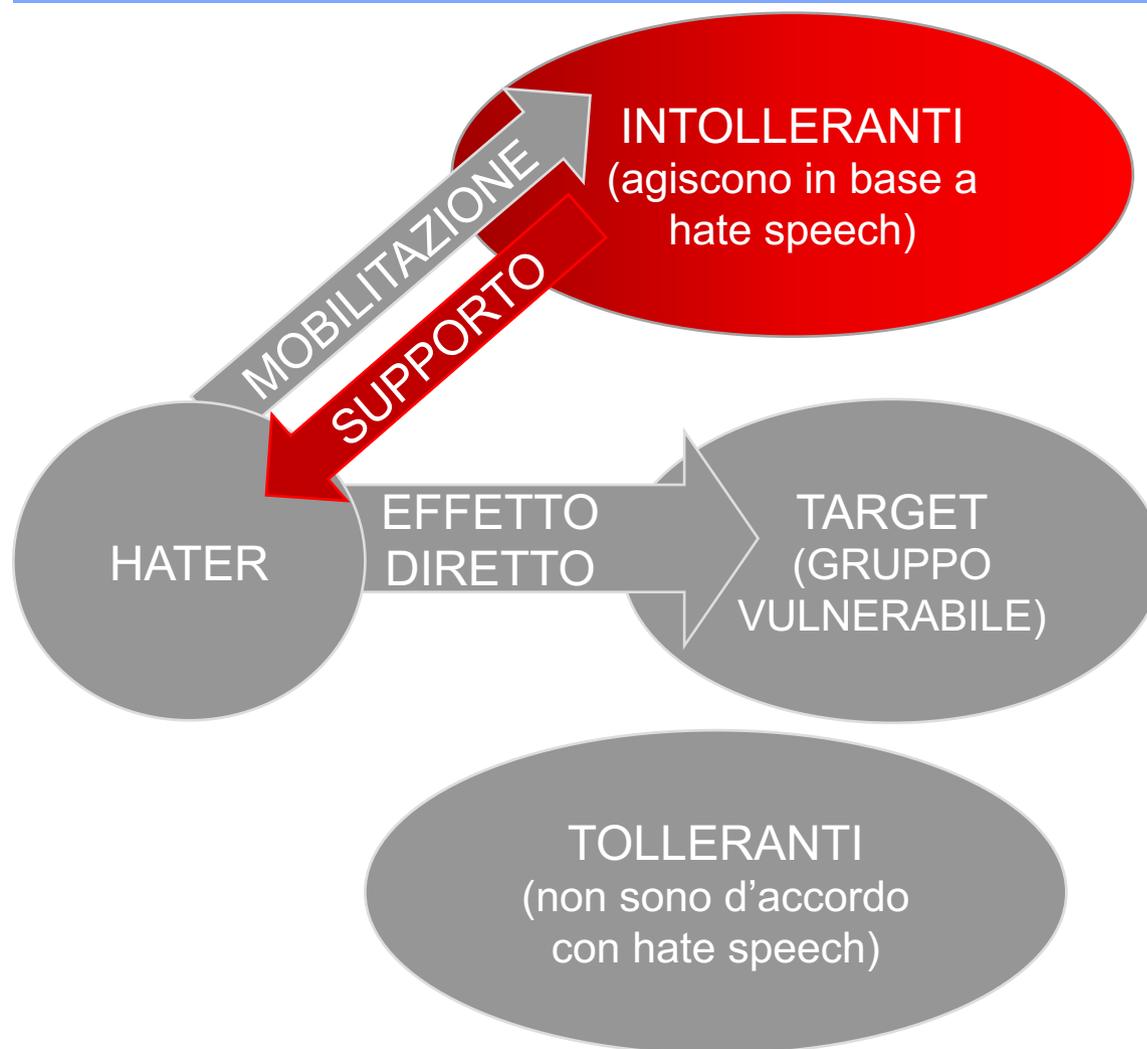
Lo schema comunicativo dell'hate speech



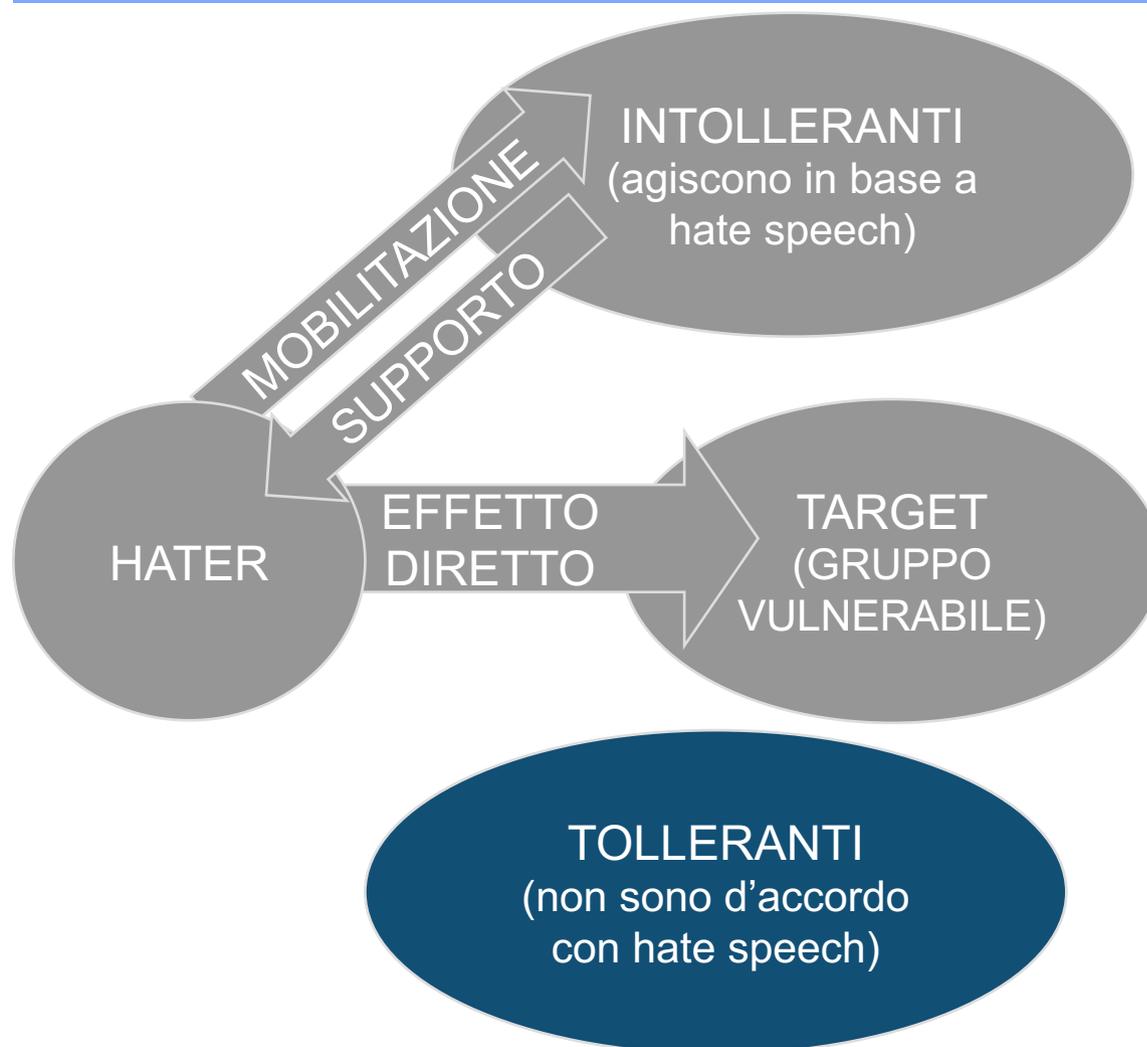
Lo schema comunicativo dell'hate speech



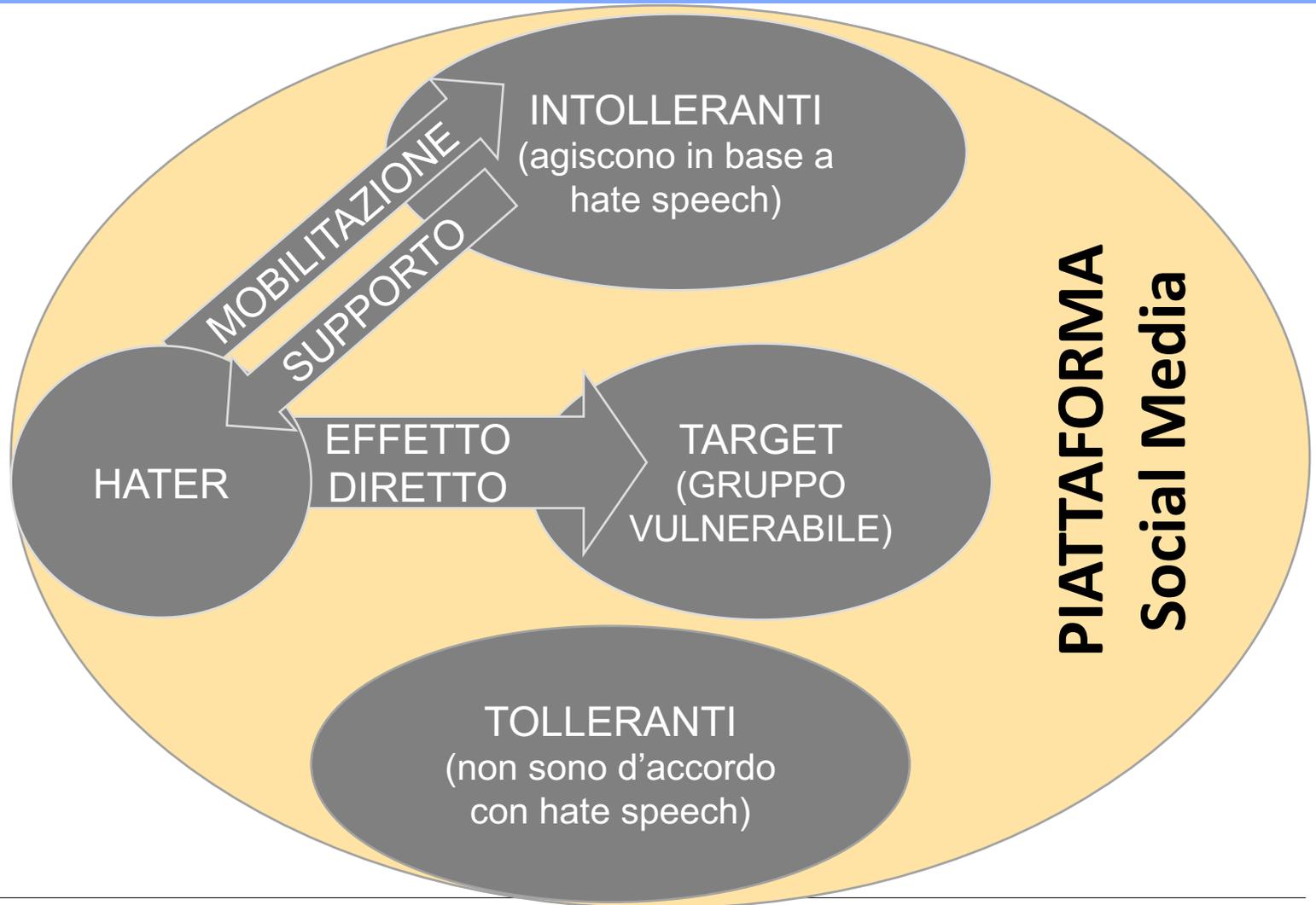
Lo schema comunicativo dell'hate speech



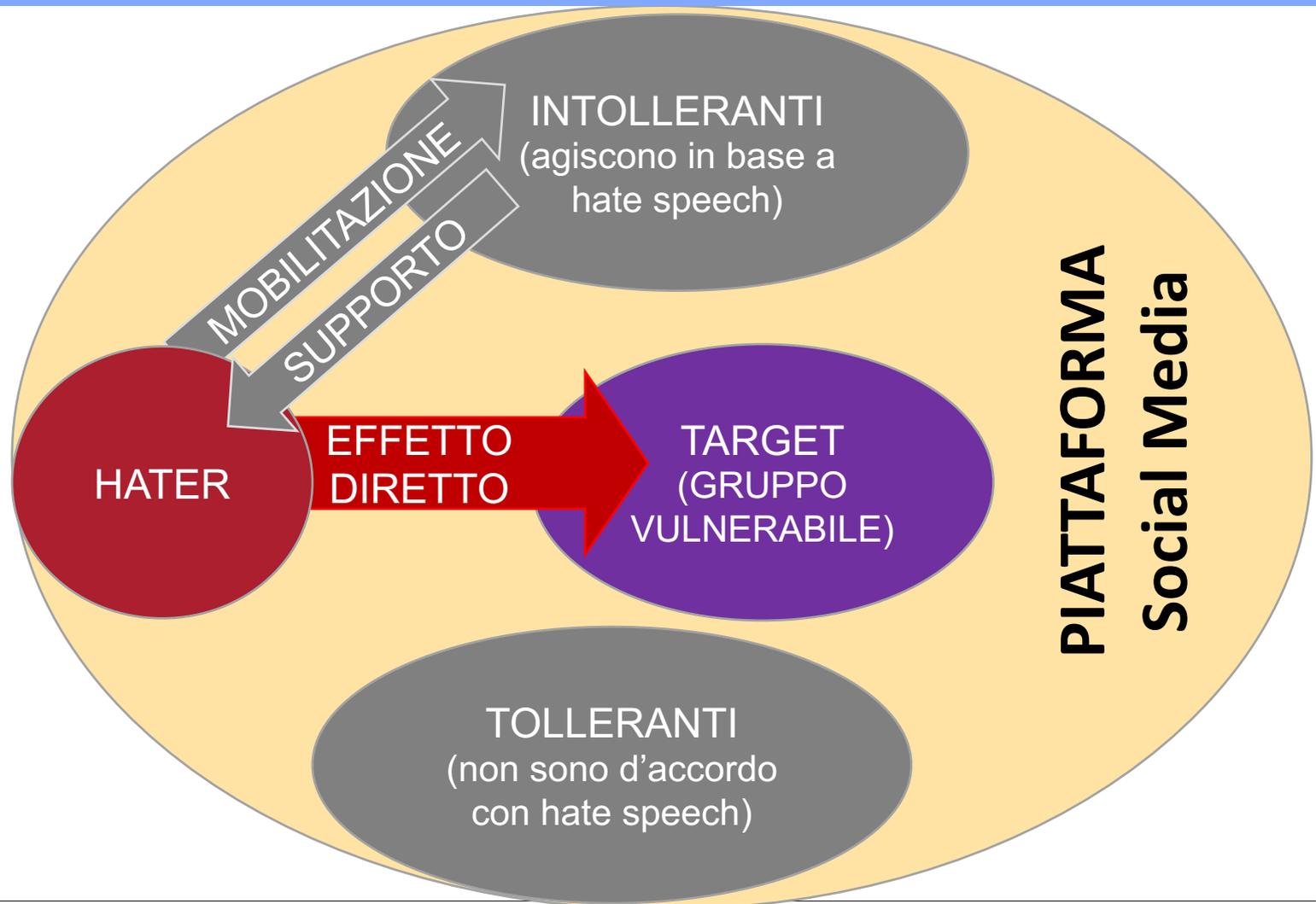
Lo schema comunicativo dell'hate speech



Lo schema comunicativo dell'hate speech



Lo schema comunicativo dell'hate speech



Hate Speech Online

Cosa fare?

STRATEGIA: prevenire
METODO: igiene digitale e filtri

Strategia prevenire

INVISIBILITÀ

DIMINUISCE IMPATTO



ISTANTANEITÀ

DIMINUISCE IMPATTO



ANONIMATO

COMUNITÀ

Strategia: prevenire



TIP #1

RIMUOVETE I
VOSTRI DATI



TIP #2

HTTPS://



TIP #3

PROTEGGETE
LOCATION



TIP #4

2-STEP
VERIFICATION



TIP #5

ESPANDETE
LA PRIVACY
ONLINE



TIP #6

THIRD-PARTY
PERMISSION



TIP #7

PIXEL
DI
TRACCIAMENTO



TIP #8

INSTALLA
GLI UPDATE



TIP #9

USA
PASSWORD
SICURE



TIP #10

PROTEGGETE
I DISPOSITIVI
FISICI



FILTRARE

Strategia: prevenire

Posso evitare di ricevere commenti o vedere post contenenti Hate Speech?

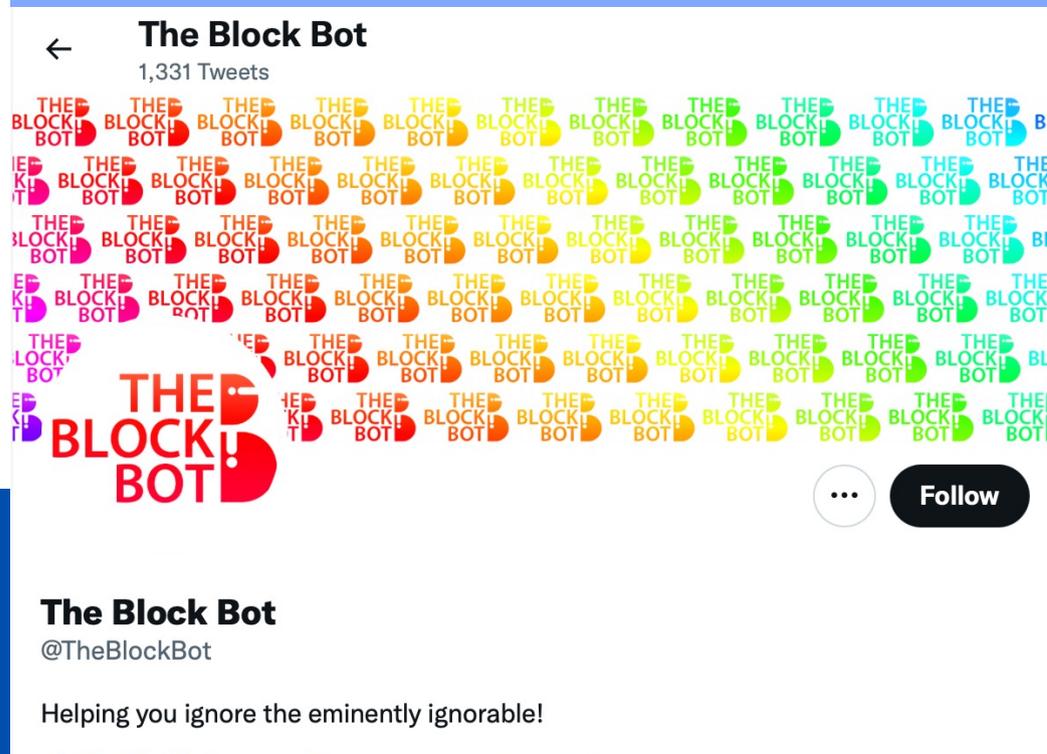
Una strategia possibile è quella **del filtraggio** che si fa forte degli strumenti tecnologici offerti dentro e fuori dai social.

Si agisce **prima che l'Hate Speech si verifichi**.



FILTRARE

Strategia: prevenire



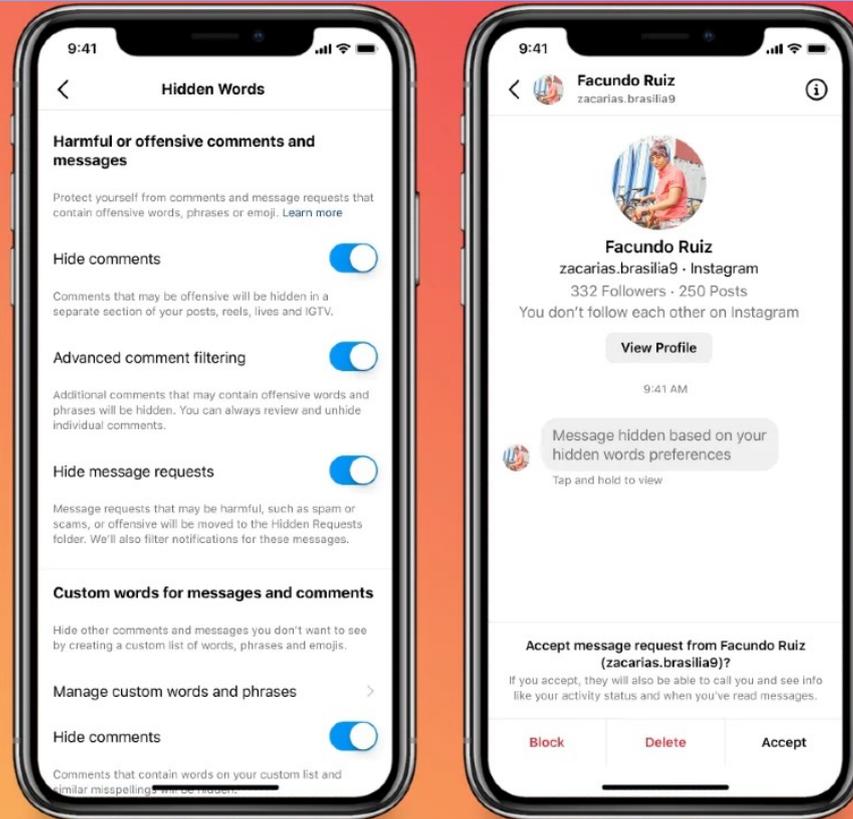
The Block Bot consente di aggiungere persone a una **lista di blocco condivisa** per abusi

<https://www.theblockbot.com>



FILTRARE

Strategia: prevenire

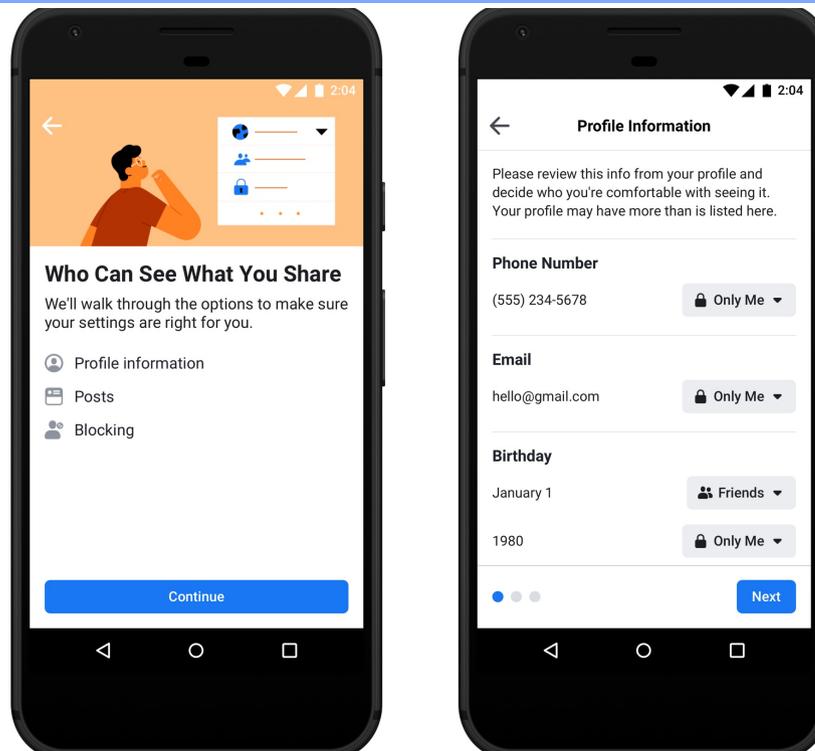


Grazie al filtro **Parole nascoste** su IG è possibile **filtrare automaticamente** i contenuti offensivi nei commenti e nei messaggi.



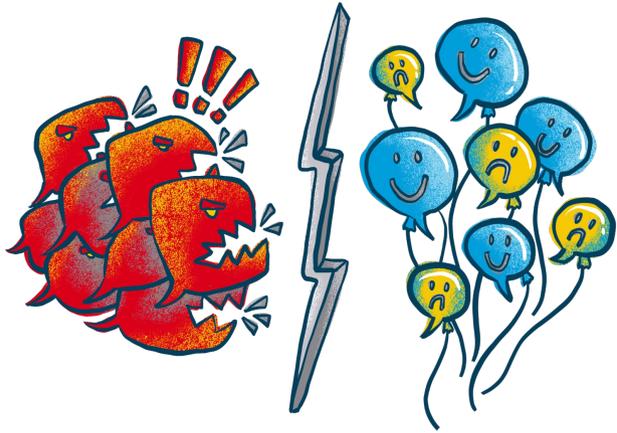
FILTRARE

Strategia: prevenire



Ogni due mesi ricevo una notifica sul calendario che mi ricorda di prendermi mezz'ora per verificare che **non siano state aggiunte nuove funzioni** che abbiano **intaccato la mia privacy**. Fate controlli regolari!

STRATEGIA: Silenzio



Strategia: Silenzio

Posso fare finta di niente di fronte a un post che contiene Hate Speech?

La **strategia del silenzio** è uno dei modi per combattere l'hate speech

DEFINIZIONE

**MA ATTENZIONE A QUANDO
E COME USARLA**



QUANDO USARLA

Strategia: Silenzio

- A pubblicare il post è stato un **troll o un account fake**
 - I troll e gli account fake cercano attenzione/interazioni, non rispondere evita di cadere nel loro gioco
- Si tratta di un **insulto personale**
 - Non è possibile intavolare una discussione

Strategia: Silenzio



DOCUMENTARE

Fare sempre **screen shot datati** dei post contenenti hate speech rivolti a voi o a altri prima di intraprendere qualunque tipo di azione

È importante che siano documentati:

- Numero delle minacce
- Dettagli (date, ore, immagine della minaccia)
- Numero delle persone/account coinvolti
- Gravità dell'attacco (implicito/esplicito)

**DA FARE
PRIMA**

Strategia: Silenzio



**COME
USARLA**

Il post viene **ignorato solo apparentemente**.

Dato che dovrebbe essere applicata quando ci si trova di fronte a troll o «odiatori seriali» questa strategia va **SEMPRE** associata alla **segnalazione** dei post offensivi sulla piattaforma di riferimento

Strategia: silenzio / Metodi: Segnalare o bloccare

INVISIBILITÀ

DIMINUISCE IMPATTO

ISTANTANEITÀ

ANONIMATO

DIMINUISCE IMPATTO

COMUNITÀ

STRATEGIA: Silenzio
METODO: Segnalare

Strategia: Silenzio / Metodo: Segnalare



VANTAGGI

Non vi obbliga a impiegare tempo per gestire gli hater.

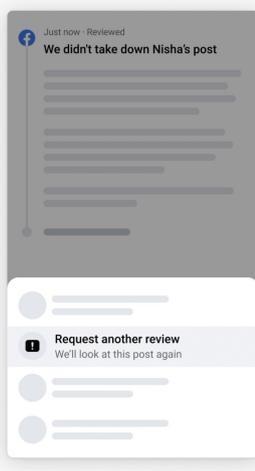
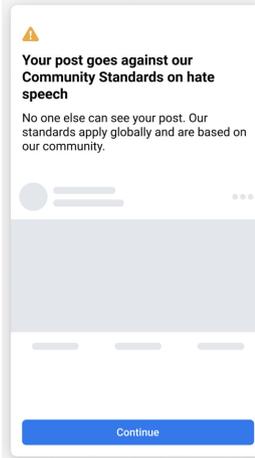
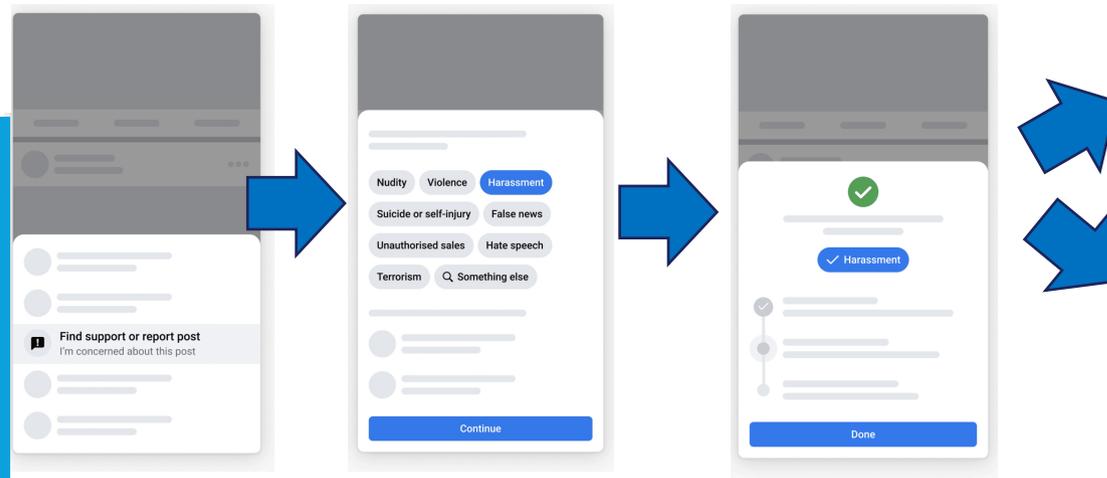
Strategia: Silenzio / Metodo: Segnalare



VANTAGGI

Le persone che sono state segnalate **non sapranno da chi arriva la segnalazione**, mentre l'utente che l'ha inviata potrà controllare direttamente lo stato della segnalazione.

Strategia: Silenzio / Metodo: Segnalare



VANTAGGI

SEGNALAZIONE DI UN POST

Può farla chiunque non solo l'utente interessato

Strategia: Silenzio / Metodo: Segnalare



SEGNALAZIONE DI UN POST SU TWITTER

Le segnalazioni di abusi possono provenire solo dall'account della vittima

← Segnala un problema

Aiutaci a capire il problema. Qual è il problema con questo Tweet?

Questo Tweet non mi interessa

È un account sospetto o di spam

Contiene materiale offensivo o dannoso

Manifesta pensieri di autolesionismo o suicidio

[Scopri di più](#) su come segnalare le violazioni delle nostre regole.



METODO

Strategia: Silenzio / Metodo: Segnalare



ATTENZIONE!

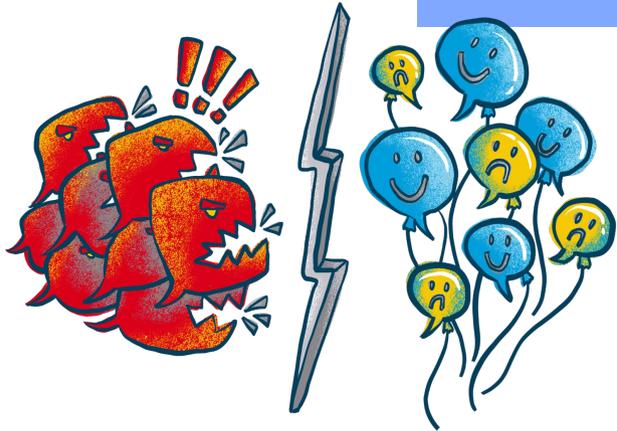
METODO

Per Twitter la minaccia deve essere "**diretta e specifica**".

Non si tratta solo di dire che ti deve succedere qualcosa, ma che ti succederà qualcosa.
Dove accadrà, a opera **di chi**, **con cosa**.

STRATEGIA: Silenzio
METODO: Bloccare

L'autoregolamentazione: le Policy interne



DEFINIZIONE

Uno dei principali strumenti utilizzati per combattere l'hate speech online è il **blocco o la sospensione del messaggio o dell'account** dell'utente «incriminato».

Strategia: Silenzio / Metodo: Bloccare



**QUANDO
USARLA**

Sui profili privati social, **bloccate, bloccate, bloccate.**

Eliminate la possibilità per odiatori seriali e troll di continuare «facilmente» a molestarvi.

Non dobbiamo alle persone violente il nostro tempo o la nostra attenzione.

Strategia: Silenzio / Metodo: Bloccare



MUTARE account che segui



Alice

Segui e muta

Bob



Alice vuole seguire Bob ma non vuole vedere i suoi tweet



Vuole vedere solo contenuti postati da Bob che la riguardano

Bob non vede che Alice l'ha mutato

MUTARE account che NON segui



Alice

muta

Bob



Alice NON vuole seguire Bob



Bob continua a menzionare e taggare Alice

Bob non vede che Alice lo ha mutato

BLOCCARE account



Alice

blocca

Bob



Alice NON vuole essere seguita e seguire Bob



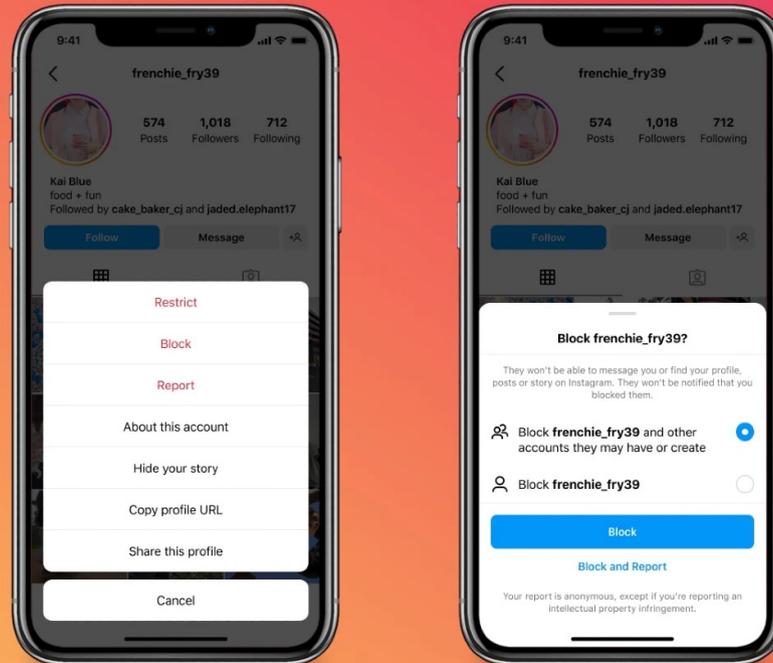
Alice non ricevere notifiche da Bob

Bob vede che è bloccato se va sul profilo di Alice

METODO



Strategia: Silenzio / Metodo: Bloccare



METODO

Quando si blocca qualcuno su Instagram si ha la possibilità di **bloccare anche gli altri account** che questa persona potrebbe avere o creare, limitando in questo modo il più possibile le interazioni con utenti indesiderati

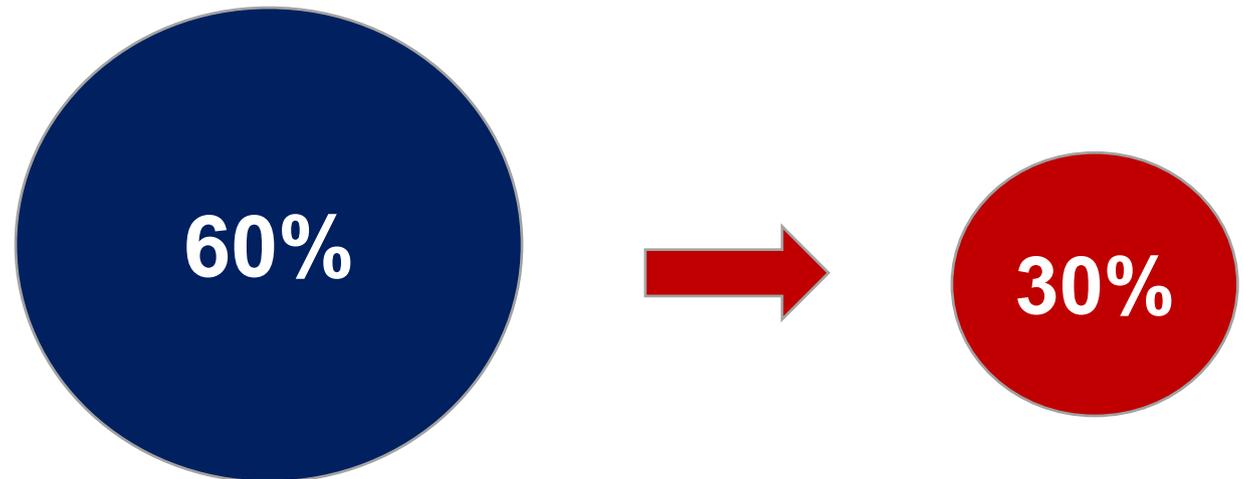
STRATEGIA: Silenzio

Gli svantaggi di segnalare e bloccare



**SCARSA
EFFICACIA**

Limiti tecnologici

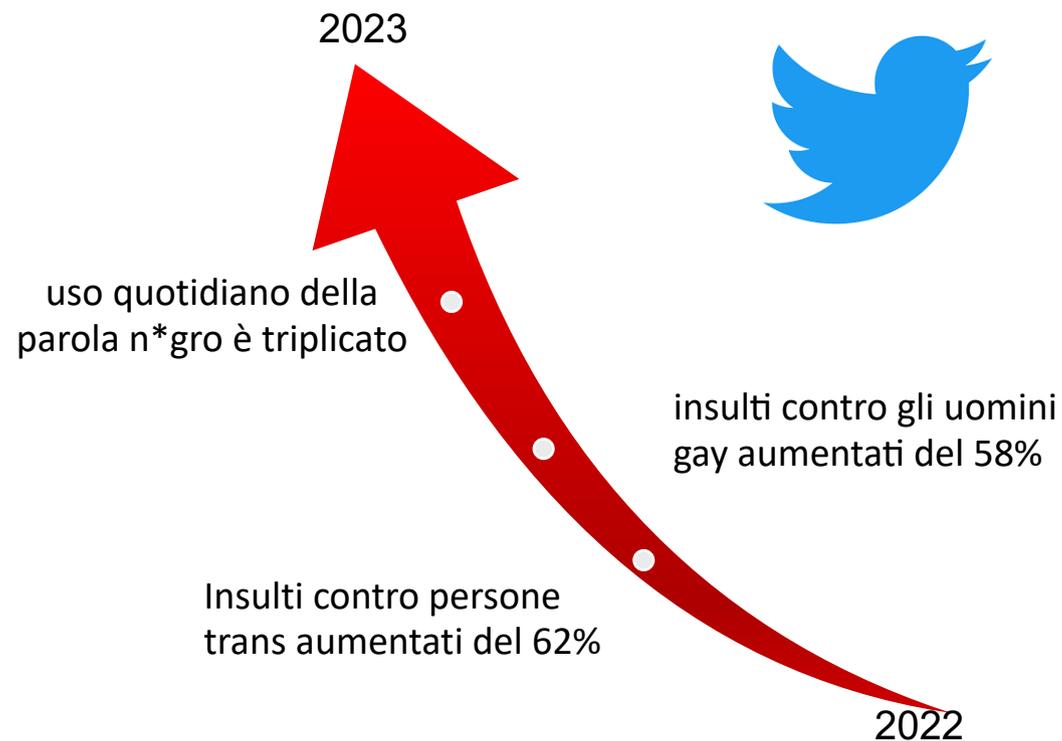


Nel 2023 Twitter è passata dall'intervenire sul 60% dei tweet segnalati a solo il 30% in casi di contenuti antisemiti (Anti-Defamation League)

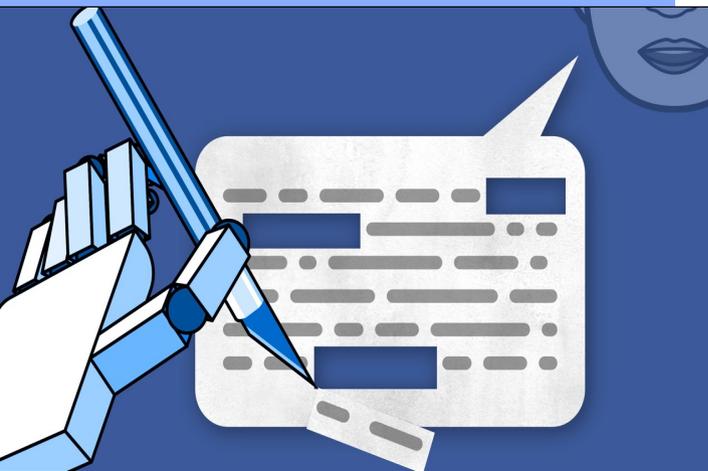


**SCARSA
EFFICACIA**

Limiti tecnologici



(Center for Countering Digital Hate)



**TREND
NEGATIVO**

Limiti tecnologici

Alcuni social hanno firmato un **Codice di Condotta** con la Commissione europea (2019). Si sottopongono a una **valutazione annuale** per determinare se rispettano il codice. C'è un **trend negativo**



 **YouTube** migliorati



rimossi meno post



più tempo a valutare i post segnalati
rimossi meno post





PROBLEMA FREE SPEECH

Limiti tecnologici

SVANTAGGI Hate speech/libertà di parola

Il blocco di un discorso d'odio può ridurre l'impatto sulla società, ma si corre sempre il **rischio di violare la libertà di parola.**

Pertanto, il rimedio preferenziale all'hate speech è quello di **aggiungere parola, non di toglierla**

(Richards e Calvert 2000)

Contrastare l'odio con la parola

parola

/pa·rò·la/

sostantivo femminile

Contrastare l'odio con la parola

"Se c'è tempo per smascherare **attraverso la discussione** le falsità e le fallacie, per scongiurare il male attraverso i **processi di educazione**, il **rimedio** da applicare è **più parola**, non il silenzio forzato".

Giudice Brandeis, 1927

parola

/pa·rò·la/

sostantivo femminile

Contrastare l'odio con la parola

Le strategie esistenti per combattere l'hate speech sui social media si limitano per lo più alla **cancellazione dei contenuti**, al **shadow banning** (il post esiste ancora ma non è più visibile) e alla **sospensione degli account**.

Alcuni lavori recenti hanno proposto un **approccio alternativo**:

rispondere all'incitamento all'odio utilizzando il **contro-discorso** (una specifica forma di contro-narrazione) che gli utenti possono creare manualmente seguendo alcune **linee guida e buone pratiche**.

Contrastare l'odio con la parola



CONTRO-
NARRAZIONE

CONTRO-
DISCORSO

Contro-narrazione: espressioni che **contrastano** un'altra **narrazione o visione in generale**, non necessariamente in risposta a una precisa dichiarazione. Categoria che potrebbe includere forme di **educazione, propaganda** positiva e **informazione** pubblica

Contro-discorso: contenuto che contrasta o contraddice contenuti d'odio o in generale discorsi lesivi, **in risposta** a una **particolare dichiarazione**.

Strategia: contro-discorso

Contro-discorso



Un contro-discorso è una risposta che fornisce **un feedback non negativo** attraverso **diversi** tipi di **argomentazioni** e **stili** ed è considerata l'approccio più efficace per contrastare l'hate speech (Benesch, 2014; Schieb e Preuss, 2016).

DEFINIZIONE

Contro-discorso

(1) Odio i musulmani. Non dovrebbero esistere.

(2) Anche i musulmani sono esseri umani. Le persone possono scegliere la propria religione.

(3) Sei davvero uno stupido idiota che pensa al contrario e crede alle negatività sull'Islam.

Un contro-discorso come quello della frase 2 è una risposta appropriata e non negativa alla frase 1, mentre quella della frase 3 non lo è, poiché inasprisce la conversazione.

ESEMPIO

Contro-discorso

Ci si rivolge agli spettatori con un contro-discorso per 4 ragioni principali:

- 1) Si spera di **cambiare il punto di vista** degli spettatori.
- 2) Si spera di **mobilitare** coloro che **non sono d'accordo** con il post, ma che non interverrebbero se non sollecitati.
- 3) Si vogliono **amplificare i contenuti negativi** per rafforzare l'idea che ci sono delle **norme contrarie** comunemente accettate, ma silenti.
- 4) Si vogliono **sostenere** coloro che sono **bersaglio del discorso d'odio**. In questo modo, si cerca di **mitigare** l'impatto negativo del discorso.



OBIETTIVI

Contro-discorso



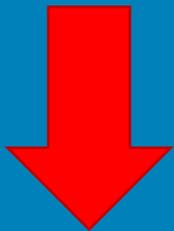
Un contro-discorso ha successo se:

- ha un **impatto favorevole** sull'utente originale (hater), spostando il suo discorso se non anche le sue convinzioni. Ciò è solitamente indicato da **scuse o ritrattazioni**, oppure dalla **cancellazione del post**.
- influenza positivamente le **norme discorsive** del "**pubblico**" di una conversazione.

Contro-discorso



EFFICACIA



È più probabile che il contro-discorso abbia successo in risposta a coloro che **non sono profondamente impegnati nell'odio o nell'estremismo.**

I troll o gli «odiatori seriali» sono **più difficili da influenzare** utilizzando il contro-discorso.

Tuttavia, anche gli scambi con i troll possono **modificare le norme del discorso e le convinzioni del pubblico.**

Contrastare l'odio con la parola



SVANTAGGI



Il contesto oltre che il tempo è importante.

Il contesto **deve consentire la discussione** e i **processi di educazione reciproca**.

Questi standard elevati **non sono sempre disponibili** per la vittima di un discorso tossico e misogino.

Contrastare l'odio con la parola



SVANTAGGI

Ogni situazione è **diversa!**

Ci sono situazioni in cui l'unico obiettivo dell'hater è quello di ferire.

In questi casi la strategia del silenzio combinata al blocco è migliore.

Se decidete di attuare la strategia del contro-discorso ci sono alcune cose che sarebbe meglio fare **preventivamente.**

Contrastare l'odio con la parola



**DA FARE
PRIMA**

Protegetevi.

Pensate a come la vostra identità online e/o profilo potrebbe essere usata contro di voi.

(vedete i **tips per l'igiene digitale** alla fine della lezione)

Contrastare l'odio con la parola



**DA FARE
PRIMA**

Ricordate che avete davanti una persona.

Dietro ogni commento, per quanto violento sia, c'è sempre una persona.

Attenzione a non scadere in un contro-discorso violento quanto l'hate speech

Nota bene: per bot e fake account il discorso è a parte

Contrastare l'odio con la parola



**DA FARE
PRIMA**

Non etichettate le persone, ma i comportamenti.

Non usate etichette come «bigotto», «razzista», o insulti come «idiota» ecc

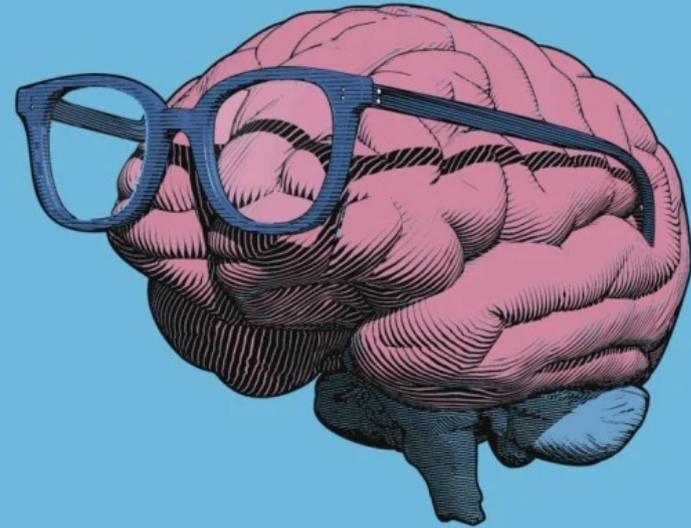
Non presupponete le cattive intenzioni della persona

Non siate aggressive/i

Strategia: contro-discorso

METODO: usare i fatti

Metodo: Usare i fatti



Presentare fatti per correggere inesattezze o percezioni errate: si cerca di persuadere correggendo le inesattezze.

DEFINIZIONE

Metodo: Usare i fatti

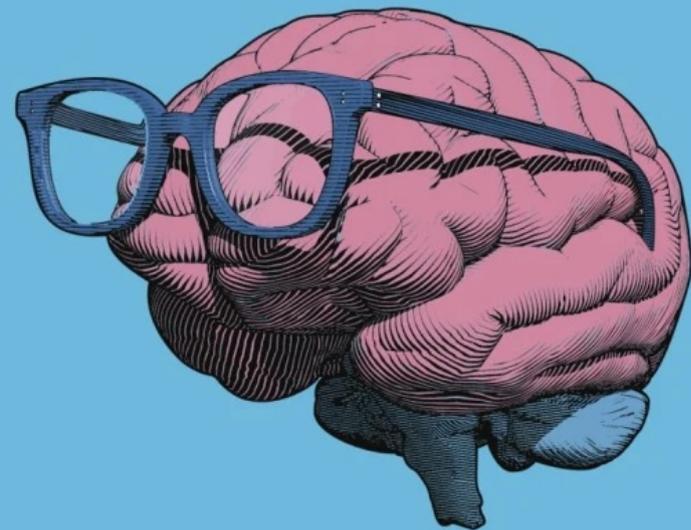
Es.

Commento in risposta a un video di un'intervista in cui l'intervistato afferma che l'omosessualità è innaturale, dannosa e distruttiva per la società.

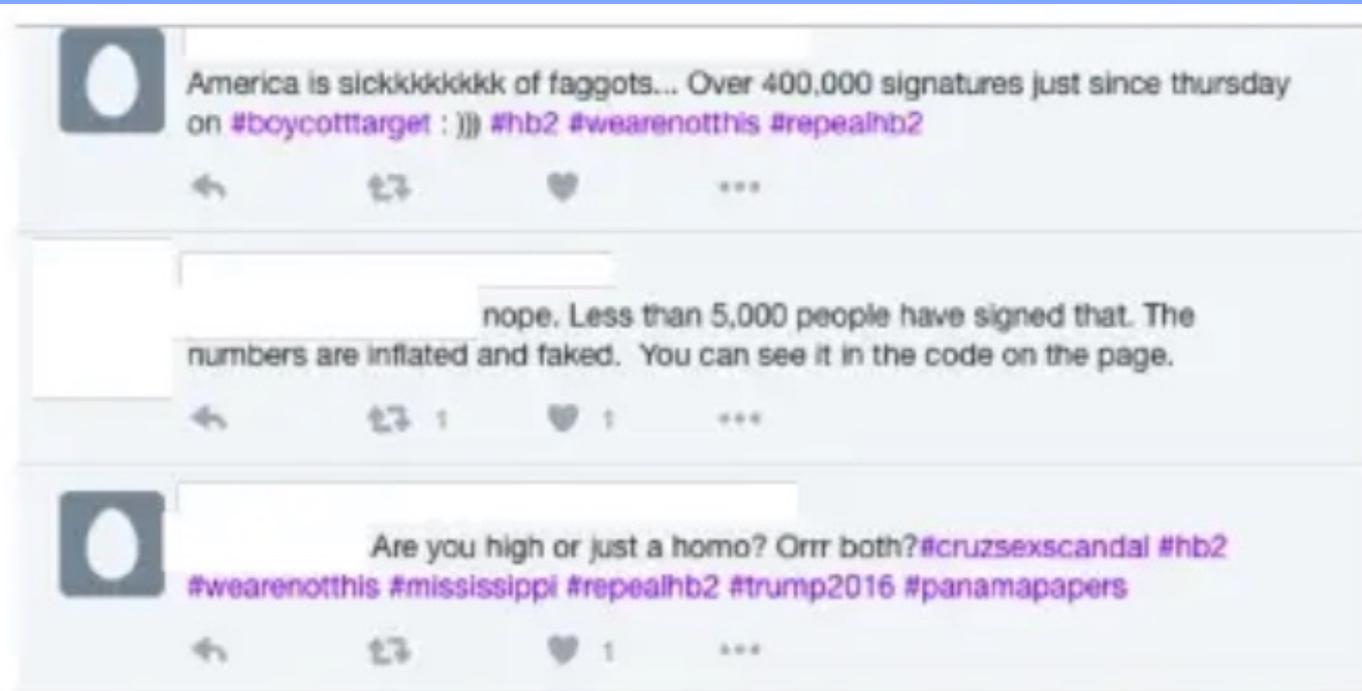
"In realtà l'omosessualità è naturale. Quasi tutte le specie animali conosciute hanno le loro comunità gay. Che si tratti di un leone o di una balena, hanno o avevano (se sono in via di estinzione) una comunità gay. Anche il matrimonio è un atto innaturale. Anche se ci sono alcune specie che hanno relazioni più lunghe con un partner, la maggior parte di quelle conosciute non lo fa".

ESEMPIO

Metodo: Usare i fatti



EFFICACIA



Quando l'oratore originale è **radicato** nelle sue opinioni, tende a trovare **un modo per adattare i nuovi fatti presentati alla conclusione a cui è già giunto** ("ragionamento motivato") o a **trovare prove diverse** a sostegno della sua posizione.

Strategia contro-discorso / Metodo: Usare i fatti

INVISIBILITÀ

ISTANTANEITÀ

ANONIMATO

COMUNITÀ

STRATEGIA: Contro-discorso

METODO: Piegare

Strategia: Piegare



«Piegare» indica una strategia che consiste nel: dare **deliberatamente** una risposta comportamentale distorta alla mossa dannosa di un oratore e precisamente, una risposta **riparatrice**, che può trasformare la mossa **in un contributo diverso e meno dannoso**.

Comportarsi

COME SE

Strategia: Piegare

Es. Promozione via social di un convegno scientifico sul cambiamento climatico. Viene pubblicata una foto dei relatori. Tutti uomini.

Utente 1: «**Wow**, un panel di soli uomini!»

Utente 2: "**Hai ragione, è terribile**. Pensare che la studiosa X ha appena pubblicato un articolo straordinario sulla responsabilità ambientale".

Seguendo il commento dell'utente 2, altri utenti aggiungono altri nomi di studiose di spicco.



Strategia: Piegare

L'utente 2 risponde **come se** l'utente 1 stesse richiamando l'attenzione su un problema **piuttosto che** gongolare per il suo tweet politicamente scorretto.

L'utente 2 **piega la forza dell'enunciato dell'utente 1**, lasciando intatto il suo contenuto letterale.

La mossa dell'utente 2 **disinnesca** così il **potenziale sessista** del commento dell'utente 1.

In particolare, il presupposto che sia giusto - forse, persino "brillantemente politicamente scorretto" - escludere le donne dagli eventi.



Strategia: Piegare



La strategia di piegare funziona anche se tutti gli altri partecipanti **hanno capito le reali intenzioni** dell'oratore «scorretto» e sanno che **anche chi «piega» le ha capite?**

Comportarsi

COME SE

Strategia: Piegare



Nell'esempio, gli altri utenti potrebbero aver capito che **l'utente 2 sa** che la risposta dell'utente 1 era volutamente politicamente scorretta e quindi sta dando **deliberatamente una risposta distorta**.

Assecondando il gioco, i conversanti **resistono** (più o meno consapevolmente) alla **direzione problematica** che la conversazione avrebbe altrimenti preso:

i convegni di soli uomini **non sono mai considerati un'opzione** (tanto meno da appoggiare con orgoglio)



Strategia: Piegare



ciò che avviene
ufficialmente in una
conversazione



ciò che i partecipanti
pensano sia avvenuto
nella conversazione

**Comportarsi
COME SE**

Il **divario** tra **verbale della conversazione** e **appunti mentali** dei partecipanti crea lo spazio per **piegare**

Strategia: Piegare



SVANTAGGI: la piegatura è praticabile **solo se l'hate speech rimane implicito**

Dichiarazioni apertamente e palesemente foriere di hate speech possono lasciare poco spazio a risposte migliorative.

Può quindi essere particolarmente difficile usare la piegatura.

SVANTAGGI

Strategia: Piegare



SVANTAGGI: si protegge la «faccia» dell'oratore colpevole di hate speech

In alcune circostanze potremmo voler (e forse dovremmo) denunciare esplicitamente l'hate speech.

Potremmo volere che chi pratica hate speech si impegni pubblicamente a dichiarare ciò che stava cercando di fare implicitamente, per attirare l'attenzione di tutti sulla dannosità che certe mosse conversazionali implicano.

SVANTAGGI

Strategia: Piegare



SVANTAGGI: piegare è impegnativo dal punto di vista cognitivo.

- È necessario afferrare la mossa prevista dall'interlocutore (cioè darle un'interpretazione corretta);
- percepire il suo potenziale dannoso;
- trovare un'interpretazione migliorativa di quella mossa;
- e comportarsi in modo tale da rendere manifesta questa interpretazione (cioè fornire una risposta distorta).

Tutto questo richiede una risposta rapida, abile e ingegnosa per controbattere.

SVANTAGGI

Strategia contro-discorso

INVISIBILITÀ

ISTANTANEITÀ

ANONIMATO

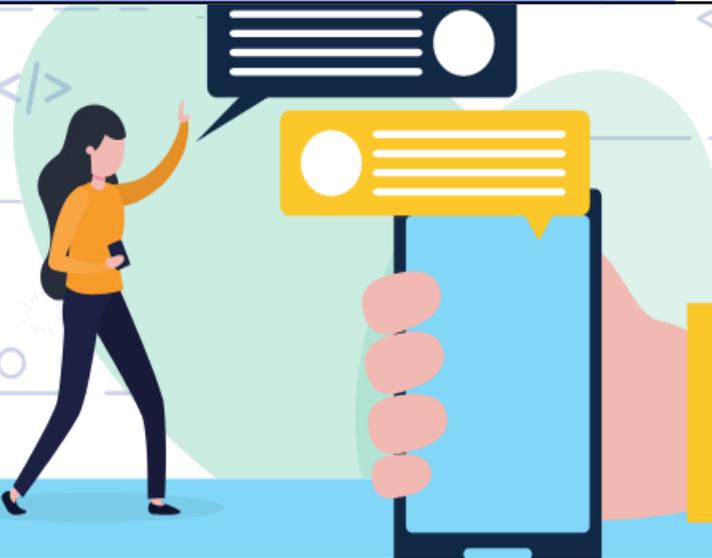
DIMINUISCE IMPATTO

COMUNITÀ



STRATEGIA: contro-discorso

METODO: esplicitare



DEFINIZIONE

Strategia: esplicitare

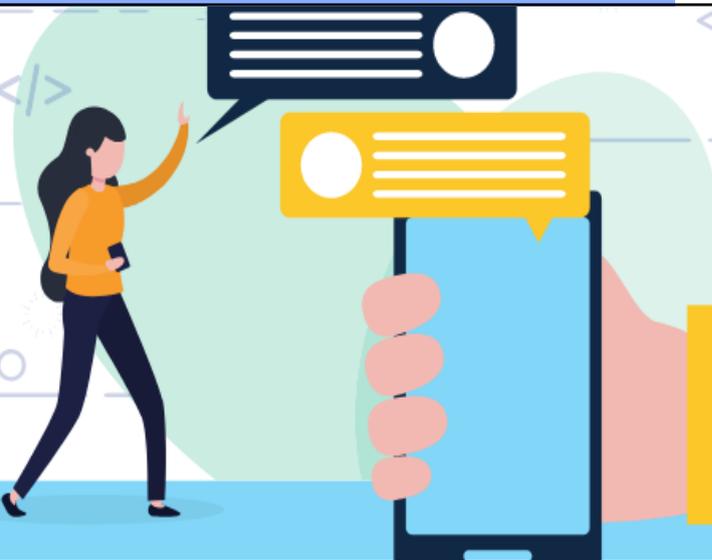
Sottolineare l'ipocrisia o le **contraddizioni** nelle affermazioni dell'utente (odiatore).

In questo caso si possono anche **etichettare** i tweet come hate speech, razzisti, bigotti, misogini, ecc.

Attenzione!

È importante **stigmatizzare le parole**, NON le persone.

Se le parole sono **stigmatizzate**, gli oratori che **non si percepiscono come odiatori**, sono spesso pronti a modificare questi post.



OPZIONI

Strategia: esplicitare

Esplicitare può assumere diverse forme.

Chi «esplicita» può:

1. **esplicitare** l'hate speech contenuto implicitamente nel messaggio **senza rifiutarlo** apertamente;
2. **chiedere** all'oratore **di esplicitare lui stesso** il contenuto discutibile.
3. **esplicitare E rifiutare** l'hate speech contenuto nel messaggio

(Tirrell 2018)



ESEMPIO 1

Metodo: esplicitare

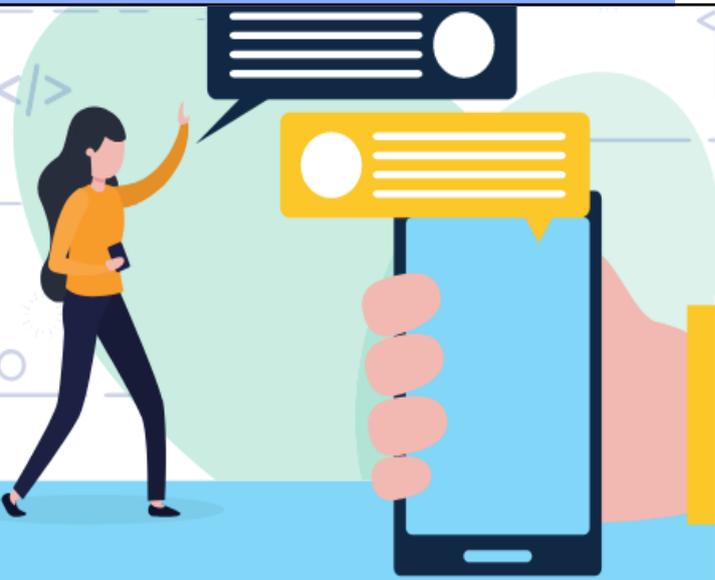
Es.

Nel 2013, un utente di Twitter ha twittato il suo sdegno per il fatto che Nina Davuluri (da lui erroneamente identificata come araba) fosse stata incoronata Miss America.

Altri utenti hanno «etichettato» il suo tweet come «razzista» oltre a correggere la supposizione che Davuluri fosse araba

Due giorni dopo l'utente si è scusato con un riferimento specifico all'etichetta razzista, dicendo:

"@MissAmerica mi dispiace di essere stato scortese e 'razzista', per favore rispondi con un tweet così tutti sapranno che è vero".



EFFICACIA



Metodo: esplicitare

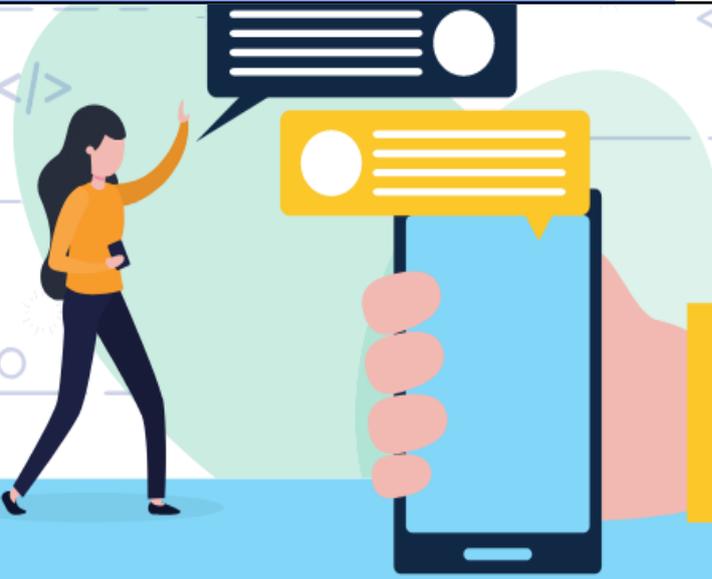
Cosa si può ottenere con questa strategia:

Etichettare il discorso come «discorso d'odio»

→ può aiutare gli **spettatori** a **identificare, interpretare e rispondere**.

Spiegare all'oratore originale **perché** la sua dichiarazione è odiosa o pericolosa.

→ Può anche aiutare a **educare l'interlocutore** in modo che non ripeta l'errore.



Strategia: esplicitare

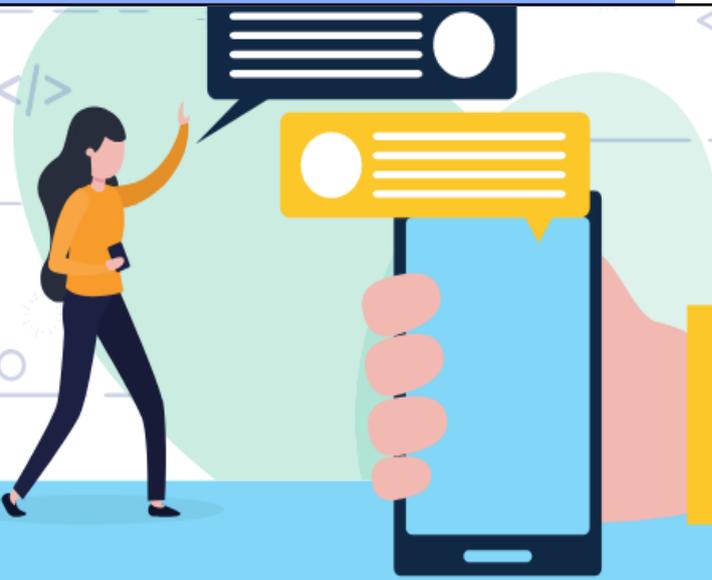
Chi esplicita può essere tenuto a fornire **motivazioni** contro il punto di vista che contesta.

Soprattutto se si appartiene a **gruppi svantaggiati** ci si mette a rischio di **sfruttamento epistemico** cioè il fenomeno per cui

"le **persone privilegiate** costringono le persone emarginate a **istruirle sulla natura della loro oppressione**".

EFFICACIA





EFFICACIA



Strategia: esplicitare

Esplicitare rischia di **rendere evidenti** ipotesi dannose perché domande come

«I cittadini che appartengono a minoranze valgono quanto gli altri?»

sono state sollevate e (dolorosamente) affrontate più e più volte.

Sfidando esplicitamente i pregiudizi, chi ferma li rende disponibili cognitivamente ai partecipanti.

Questo può influenzarli verso **scelte e comportamenti bigotti**.

Strategia contro-discorso

DIMINUISCE IMPATTO
INVISIBILITÀ



ISTANTANEITÀ

ANONIMATO

DIMINUISCE IMPATTO
COMUNITÀ



STRATEGIA: contro-discorso

METODO: avvertire



DEFINIZIONE

Metodo: avvertire

Avvertire delle **conseguenze offline o online**:

- ricordare all'oratore il **danno** che un discorso d'odio può arrecare al **gruppo di riferimento**, poiché le parole possono **catalizzare l'azione**.
- ricordare all'interlocutore che **molte persone nel suo mondo offline** (compresi datori di lavoro, amici, familiari e futuri datori di lavoro) possono **vedere ciò che viene pubblicato online**.
- ricordare all'interlocutore la **permanenza** della comunicazione online.
- ricordare all'interlocutore le possibili **conseguenze online** di un discorso odioso o pericoloso, come il **blocco**, la **segnalazione** e la **sospensione** dell'account.

Metodo: avvertire



DEFINIZIONE

Es.

Gli ex allievi dell'Università dell'Illinois a Champaign-Urbana hanno **avvertito** gli studenti che i datori di lavoro attuali e futuri avrebbero potuto vedere i loro tweet con l'hashtag #FuckPhyllis, creato per attaccare Phyllis Wise, l'allora rettore della loro università, con contenuti misogini e razzisti.

(Erano arrabbiati perché aveva deciso di non concedere un giorno di sospensione in un lunedì molto freddo).

Molti studenti hanno **risposto cancellando i tweet di hate speech**.

Metodo: avvertire



EFFICACIA



Questa strategia di contro-discorso porta in diversi casi alla **ritrattazione** o alla **cancellazione** dei tweet offensivi.

Ciò può essere dovuto al fatto che le persone **dimenticano**, finché non glielo si ricorda, **che i loro discorsi online possono avere gravi conseguenze offline.**

Online possono diventare disinibiti, agendo come se

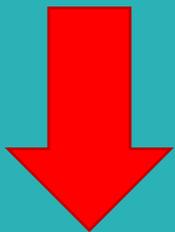
"vivessero in una dimensione di finzione, separata e distante dalle richieste e dalle responsabilità del mondo reale"

(Suler, 2004, p. 323).

Metodo: avvertire



SVANTAGGI



Avvertire di **conseguenze offline** può facilmente **degenerare** in una serie di **minacce e in molestie** nei confronti degli avversari online, a volte anche da parte di un gran numero di persone (dogpiling).

Si può essere tentati di vedere queste strategie come utili o addirittura costruttive, se usate contro chi pratica hate speech.

Attività come il doxing o le minacce non devono essere confuse con un contro-discorso costruttivo.

Strategia contro-discorso

DIMINUISCE IMPATTO

INVISIBILITÀ



DIMINUISCE IMPATTO

ANONIMATO



ISTANTANEITÀ

DIMINUISCE IMPATTO

COMUNITÀ



STRATEGIA: contro-discorso

METODO: umorismo e sarcasmo



DEFINIZIONE

Metodo: umorismo e sarcasmo

L'umorismo è uno degli strumenti più potenti utilizzati per combattere i discorsi d'odio.

Può **attenuare** i conflitti e essere usato per **attirare l'attenzione sull'argomento**.

L'umorismo in contesti online **attenua l'ostilità, offre sostegno** agli altri interlocutori e incoraggia la **coesione sociale**

(Marone 2015).



ESEMPIO

Metodo: umorismo e sarcasmo



Iyad El-Baghdadi
@iyad_elbaghdadi

26 Dec

ISIS leaders: We urgently call upon every Muslim to join the fight, especially those in the land of the two shrines (Saudi Arabia), rise.

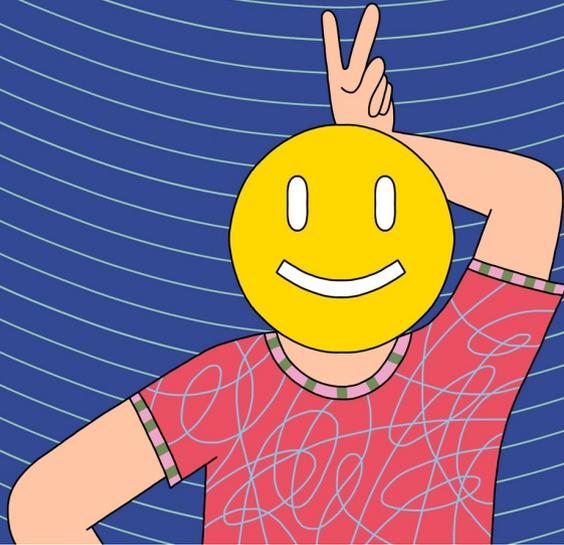
 Follow

@iyad_elbaghdadi Too busy being part of a civilised and functioning society. Also, Sherlock S04 in 4 days. I can't miss the first episode.

9:15 AM - 28 Dec 2015

  71  385

2015 l'attivista per i diritti umani Iyad El-Baghdadi ha twittato parte di un appello all'azione dell'ISIS. Utenti musulmani di Twitter hanno risposto motivando umoristicamente le ragioni per cui non aderivano.



EFFICACIA



Metodo: umorismo e sarcasmo

È efficace per:

- **Neutralizzare** discorsi d'odio e pericolosi che vengono visti come potenti o intimidatori.
- **Attirare** l'attenzione di un pubblico più vasto sul contro-discorso.
- **Ammorbidire** un messaggio che altrimenti sarebbe duro o aggressivo.



Strategia contro-discorso

DIMINUISCE IMPATTO

INVISIBILITÀ



DIMINUISCE IMPATTO

ANONIMATO



ISTANTANEITÀ

DIMINUISCE IMPATTO

COMUNITÀ



CREA CONTRO-COMUNITÀ



STRATEGIA: contro-discorso

METODO: affiliazione ed empatia

Metodo: affiliazione ed empatia



DEFINIZIONE

Cambiare il tono di una conversazione d'odio è un modo efficace per **prevenire** l'escalation e **porre fine allo scambio**. I controrelatori possono:

- Usare un tono **amichevole, empatico** o pacificatore nelle risposte ai messaggi che usano un tono ostile, odioso o violento.
- **Affiliarsi all'interlocutore** originale per **stabilire un legame** (ad esempio, anch'io sono un conservatore, ma...).
- **Affiliarsi al gruppo bersaglio** del discorso d'odio per **generare empatia** (ad esempio, quello che hai detto mi ha ferito in quanto asiatico...).

Metodo: affiliazione ed empatia



ESEMPIO

Es.

In un lungo scambio di battute con un utente intenzionato a usare insulti razziali in occasione del Martin Luther King Day negli Stati Uniti, un altro utente di Twitter ha ripetutamente risposto con empatia:

"ti auguro ancora amore»

"per quanto tu possa odiare, non ti augurerò mai altro che amore»

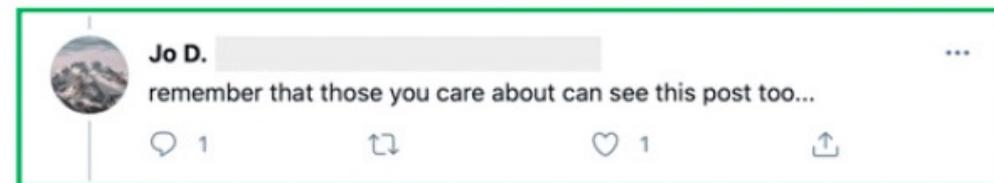
"non riderei mai del tuo omicidio. Non deriderei mai il tuo lutto. Non deriderei mai la tua lotta per l'uguaglianza".

Il primo utente ha risposto con un tweet di razzista dopo l'altro, ma **alla fine si è scusato.**

Conclusioni

IRONICO

BASATO SU EMPATIA



AVVERTE DELLE CONSEGUENZE

EFFICACIA



Conclusioni



EFFICACIA



Solo la risposta empatica ha prodotto effetti consistenti, anche se relativamente piccoli, rispetto alla riduzione di hate speech xenofobo.

Questo risultato indica **un ruolo centrale dell'empatia** nel combattere i discorsi di odio online e fa eco a ricerche precedenti che documentano come le conversazioni di persona che incoraggiano l'empatia possono ridurre l'ostilità verso i gruppi emarginati

Metodo: affiliazione ed empatia



EFFICACIA



Quando chi controbatte è in grado di **trasformare gli scambi d'odio in conversazioni civili**, l'ipotesi del contatto suggerisce che **le persone coinvolte possono diventare meno odiose o estremiste** nel tempo.

Questo è un **risultato positivo** anche quando l'interlocutore originale non ritratta, non si scusa e non cambia opinione.

Strategia contro-discorso

DIMINUISCE IMPATTO

INVISIBILITÀ

SPINGE A SENTIRE



DIMINUISCE IMPATTO

ANONIMATO

SPINGE A USCIRE



ISTANTANEITÀ

DIMINUISCE IMPATTO

COMUNITÀ

CREA CONTRO-COMUNITÀ



Conclusioni

Conclusioni

L'abuso online non è un problema tecnologico; è un **problema sociale alimentato dalla tecnologia**.

Le soluzioni migliori sono quelle che non solo **disinnescano** il potere di Internet di amplificare l'abuso, ma incoraggiano anche **cambiamenti nelle norme sociali**, ponendo il comportamento scorretto al di là di un **limite** immaginario.

ACCENTO
SU
SOCIAL

Conclusioni



ACCENTO SU SOCIAL

Quando le persone parlano di molestie online, una delle risposte più comuni è "Beh, cosa ti aspettavi da Internet?".

Se vogliamo davvero cambiare i nostri spazi online, la risposta di tutti noi deve essere:
di più.

Da fare sempre:

TIPS per un'igiene digitale



TIP #1

RIMUOVETE I VOSTRI DATI

Prima di ogni strategia

Controllate tutte le **impostazioni sulla privacy** su ogni App che usate.

Poche App hanno bisogno di accedere ai vostri contatti, foto o posizione, ad esempio, ma molte lo richiedono come se fosse standard.

Potete anche limitare l'accesso a [solo] quando state utilizzando l'App



TIP #1bis

**RIMUOVETE I
VOSTRI DATI
DAVVERO**

Prima di ogni strategia

Le nostre **informazioni personali** circolano liberamente su Internet. E i data broker le stanno vendendo ai migliori offerenti.

Dovete andare su ciascun sito e richiedere che le vostre informazioni vengano rimosse, cancellate o sopprese. Il processo è diverso per ogni broker di dati. Alcuni onoreranno la vostra rimozione e non vi aggiungeranno mai più. Altri vi aggiungeranno di nuovo.

Esistono diversi servizi online che faranno la procedura per conto vostro. Es:

<https://incogni.com>



TIP #2

HTTPS://

Prima di ogni strategia

Navigate sempre usando HTTPS://

Significa che tutto quello che viene scritto su un sito **non può essere intercettato** e letto da estranei e nemmeno dal sito stesso.

In tutti i browser (Chrome, Firefox, Safari, ecc.) è possibile **attivare la modalità di navigazione in https**

Qui trovate **tutorial** su come attivare la navigazione https nei diversi browser

<https://www.eff.org/https-everywhere/set-https-default-your-browser>



TIP #3

PROTEGGETE LOCATION

Prima di ogni strategia

Protegete la vostra location mentre siete online

Accedere a siti web dal lavoro o da casa consente ai proprietari dei siti di tracciare il vostro indirizzo IP, che a volte può aiutarli a localizzarvi. Queste informazioni possono finire nelle mani di troll o hater.

Potete usare Tor che mantiene la vostra posizione anonima

<https://www.torproject.org/>

Può essere lento e dovrete spesso inserire dei captcha dunque potete anche usarlo solo al bisogno (ad esempio se siete sotto attacco)

Oppure potete usare un VPN commerciale, non vi accorgete di nulla a livello di navigazione ma la vostra location sarà protetta

NON USATE VPN GRATUITI: vendono a terzi i vostri dati



TIP #4

2-STEP VERIFICATION

Prima di ogni strategia

La verifica in due passaggi e l'autenticazione a due fattori aggiungono un ulteriore livello di sicurezza agli account a cui si accede online, compresi gli indirizzi e-mail e i conti bancari.

Si inseriscono il nome utente e la password, ma prima di accedere è necessario inserire un secondo fattore: due ostacoli da superare.

Se qualcuno tenta di reimpostare la vostra password e avete attivato la 2FA, sarà molto più difficile per lui accedere al vostro account.

Controllate tutte le app, i servizi web e i siti in cui siete registrate/i e scegliete la «2-step verification»



TIP #5

ESPANDETE LA PRIVACY ONLINE

Prima di ogni strategia

Prendete in considerazione l'installazione di **estensioni o plugin del browser per ridurre al minimo la vostra impronta online.**

Questi strumenti vi proteggono principalmente dalla pubblicità o dalle aziende che vi seguono su più siti attraverso i cookie e altri metodi che tracciano le vostre abitudini di navigazione.

Tuttavia, questo tipo di blocco aiuta anche a ridurre al minimo le informazioni personali che i siti web possono scoprire su di voi, evitando così che vengano compilate o finiscano online.

Alcuni esempi di plug-in per il browser sono Adblock Plus o Ghostery

Per utilizzare alcuni siti potreste doverli disabilitare momentaneamente.

Prima di ogni strategia



TIP #6

THIRD-PARTY PERMISSION

Un'applicazione di terze parti è un prodotto **che non fa parte del servizio principale che state utilizzando**, ma che ha accesso al vostro account principale o alle sue credenziali.

Per le applicazioni che non utilizzate regolarmente, prendete in considerazione la possibilità di **revocare l'accesso** alle autorizzazioni delle applicazioni di terze parti da Twitter, Facebook, Instagram e dal vostro telefono.



TIP #7

PIXEL DI TRACCIAMENTO

Prima di ogni strategia

I pixel di tracciamento, insieme ai cookie, sono un altro metodo per tracciare la posizione dell'utente ai fini del targeting degli annunci. I pixel di tracciamento sono incorporati nelle immagini.

Quando aprite un'e-mail, scaricate automaticamente l'immagine e la richiesta di immagine fa sapere immediatamente al server che l'avete aperta, insieme ad altre informazioni come il dispositivo, il browser e potenzialmente la posizione.

Assicuratevi che il **caricamento delle immagini** sia **disattivato** quando controllate la posta elettronica sul portatile e sul telefono.

Anche l'utilizzo di una VPN o di Tor vi aiuterà a mantenere l'anonimato dai pixel di tracciamento.



TIP #8

INSTALLA GLI UPDATE

Prima di ogni strategia

Che si tratti di un telefono, di un portatile, di un tablet o di un computer desktop, non c'è mai un momento opportuno per installare patch e aggiornamenti.

È facile continuare a ignorare questo aspetto.

Se non volete impostare l'aggiornamento automatico di tutto, cercate di **non rimandarlo all'infinito**.

Assicurarsi di disporre degli aggiornamenti software e delle correzioni di bug più recenti è un modo efficace per ridurre i rischi per la sicurezza e, idealmente, per correggere le vulnerabilità prima che qualcuno possa **sfruttarle** sui vostri dispositivi.

Prima di ogni strategia



TIP #9

USA PASSWORD SICURE

Il modo migliore per mantenere sicuro il vostro account online è **utilizzare password sicure e uniche.**

Ciò significa che la password deve essere abbastanza complessa da essere difficile da decifrare.

Dovreste inoltre utilizzare una password diversa per ogni account, nel caso in cui uno di essi venga compromesso.



TIP #10

PROTEGGETE I DISPOSITIVI FISICI

Prima di ogni strategia

Oltre a proteggere i vostri dati online, dovrete **proteggere anche i vostri dispositivi fisici.**

Alcuni di questi accorgimenti sono di buon senso, come ad esempio assicurarsi di non lasciare incustoditi il laptop e il telefono.

È anche una buona idea prestare attenzione quando si collegano dispositivi USB al computer, poiché potrebbero contenere malware che potrebbero danneggiare il sistema o esporre i vostri dati personali, quindi assicuratevi di collegare un dispositivo USB al vostro computer personale solo se vi fidate della persona che ve lo ha fornito.

Assicuratevi anche che il vostro computer sia dotato di crittografia completa del disco. In questo modo, se qualcuno ruba il vostro portatile e copia i file, vedrà solo stringhe di codice.

Grazie per l'attenzione!