

ITALIA

AMNESTY
INTERNATIONAL



HATE SPEECH

Guida breve per
combattere i discorsi
d'odio online

**CONOSCERLO
E CONTRASTARLO**

Amnesty International è un movimento globale di oltre 10 milioni di persone impegnate in campagne per un mondo dove tutti godano dei diritti umani. La nostra visione è che ogni persona possa godere dei diritti sanciti dalla Dichiarazione universale dei diritti umani e altri standard internazionali sui diritti umani. Siamo indipendenti da qualsiasi governo, ideologia politica, interesse economico o religione e ci finanziamo principalmente grazie alle persone associate e che ci sostengono.

Amnesty International Italia promuove il rispetto della dignità umana e il diritto di ogni persona a esprimere liberamente la propria identità di genere e il proprio orientamento sessuale e si batte contro ogni forma di violenza e discriminazione attraverso le immagini, i comportamenti e le parole.

Nei testi del presente documento abbiamo cercato di adottare terminologie neutre. Laddove non sia stato possibile, abbiamo fatto ricorso al finto neutro, per soli fini di semplificazione, sintesi e leggibilità.

2024 © Amnesty International Sezione Italiana

Immagine di copertina: © Playlab

Grafica: Enrico Calcagno Design

Per maggiori informazioni:
info@amnesty.it
www.amnesty.it

INDICE

INTRODUZIONE 5

PRIMA PARTE

CONOSCERE 6

1	DA DOVE ARRIVA L'HATE SPEECH?	7
	PERCEZIONI ED EMOZIONI	7
	STEREOTIPI E PREGIUDIZI	10
	IL FRAME NARRATIVO	12
	LA PIRAMIDE DELL'ODIO	12
2	RICONOSCERE L'HATE SPEECH	14
	UNA DEFINIZIONE	14
	I MESSAGGI	14
	LA NARRAZIONE DELL'ODIO	16
3	DIS/INFORMAZIONE ONLINE	19
4	EFFETTI OLTRE LA RETE	21
	LE CONSEGUENZE DELL'HATE SPEECH	21
	L'IMPATTO SUI SINGOLI INDIVIDUI	22
	L'IMPATTO SULLA COMUNITÀ	22

SECONDA PARTE

CONTRASTARE 23

1	COMBATTERE L'HATE SPEECH ONLINE	24
	DA DOVE INIZIO?	24
	OSSERVARE	24
	ORGANIZZARE	24
	SICUREZZA TECNOLOGICA	25
	SICUREZZA PSICO-EMOTIVA	26
2	PASSARE ALLA PRATICA	27
	SUGGERIMENTI PRATICI	27
	QUESTIONE DI GRAVITÀ	27
3	REAGIRE ALL'HATE SPEECH	32
	C'È UN MODO GIUSTO?	32
	MODALITÀ POSSIBILI	32
4	RISPONDERE	32
	NARRAZIONE POSITIVA	32
	INTERAGIRE CON L'ALTRO	34
5	ATTIVARSI	35
	LE AZIONI: MONITORAGGIO E CONTRASTO	35
	COME SI ATTIVA LA TFHS	36
	TONI D'ODIO E DISINFORMAZIONE	36
	PERCHÉ INTERVENIRE	37
	IL VALORE DEL GRUPPO	37
	IL MANIFESTO DELLA TASK FORCE HATE SPEECH	38
6	SEGNALARE	40
	MOLTI SOCIAL NETWORK, TANTE DIFFERENZE	40



© Jan van Eck



Introduzione

Appare sempre più evidente il ruolo centrale di ogni singolo individuo e della società civile tutta nel contrasto ai discorsi e ai fenomeni d'odio online. Mentre avanziamo, passo dopo passo, sul lungo percorso che, sotto il profilo legislativo e culturale-educativo, ha come obiettivo quello di giungere a un cambiamento e alla previsione di strumenti e forme di tutela che rendano il web uno spazio dove dove ogni persona possa sentirsi libera e sicura, allo stato attuale legge e tecnologia spesso non sono ancora in grado di rimuovere forme di intolleranza e discriminazione in rete prima che esse abbiano prodotto un danno. In questo contesto la risposta di coloro che fruiscono del web all'*hate speech* può essere immediata: chiunque abbia una tastiera, una connessione e la giusta dose di volontà e controllo può, nella modalità e nella misura più opportuna reagire, dire no all'odio che infesta la rete.

In questa breve guida cerchiamo di fornire nozioni e strumenti essenziali affinché ci si possa impegnare in prima persona in questa impresa: facendo sentire il proprio sostegno a chi è preso di mira dagli *hater*; fornendo a chi sta cercando di costruire la propria opinione una visione alternativa a quella discriminatoria e marginalizzante; segnalando i contenuti che non dovrebbero essere pubblicati perché lesivi dei diritti di coloro che sono più esposti al rischio di essere discriminati. Diritti fondamentali che, come esseri umani, tutti condividiamo.



PRIMA PARTE

CONOSCERE

1

DA DOVE ARRIVA L'HATE SPEECH?

Comprendere i fattori alla base di un fenomeno è indispensabile per poter intraprendere qualsiasi azione volta a contrastarlo in modo efficace. Per poter individuare possibili risposte ai discorsi d'odio online, quindi, è opportuno fare un passo indietro e risalire alle loro origini: dove affondano le radici? Quali meccanismi conducono a essi? Nella prima parte di questa guida proviamo a fornire una risposta, attraverso alcune nozioni chiave.

PERCEZIONI ED EMOZIONI

Per comprendere e contrastare l'*hate speech* (in italiano “discorso d'odio”) è necessario partire da un breve sguardo ai **meccanismi psicologici dai quali deriva la percezione di ciò che avviene intorno a noi**.

Esiste, infatti, un certo numero di fattori che interagiscono – a volte in modo conflittuale – definendo le nostre attitudini e i nostri comportamenti: interrogarsi su come funzionano è cruciale per comprendere l'origine e la persistenza di idee false e per poter elaborare risposte efficaci.

Ciò che chiamiamo “visione del mondo” è data dalla combinazione di diversi elementi: **emozioni, identità, esperienze vissute, valori,**





convinzioni. La visione del mondo determina la percezione che abbiamo delle esperienze così come il significato che attribuiamo loro. Quando ci confrontiamo con un elemento che non corrisponde alla nostra visione del mondo, può prodursi un senso di disagio, la **dissonanza cognitiva**.

Il nostro cervello risponde provando a sopprimere questa sensazione sgradevole e lo fa cambiando il significato dato all'evento, affinché possa tornare a corrispondere alla nostra visione del mondo. Ecco perché, per esempio, le **teorie del complotto** hanno tanto successo: esse consentono di attribuire spiegazioni diverse ai fatti, più conformi alla nostra visione del mondo e, di conseguenza, più confortanti.

• EMOZIONI

Reagiamo in modo differente e complesso agli stimoli interni e esterni, provando diverse tipologie di emozioni che possono guidare i nostri comportamenti. La modalità attraverso la quale interagiscono la parte del nostro cervello relativa alle emozioni e quella del ragionamento può essere comparata all'immagine di una persona a cavallo: la persona (la ragione) avrà controllo limitato su ciò che fa il cavallo (le emozioni)¹. Alcuni temi, come l'immigrazione, suscitano emozioni potenti quali la paura o la rabbia: per parlare in modo costruttivo con chi prova queste sensazioni, è necessario riuscire a **produrre dei contenuti in grado di ridurre ansia e emozioni negative**.

• IDENTITÀ

Pensando a come descriviamo noi stessi, possiamo distinguere due tipologie di identità: quella personale (alla quale riconduciamo aspetti come qualità, ruoli, relazioni, elementi biografici ecc.) e quella sociale, che deriva dall'appartenenza a determinati gruppi (sulla base di nazionalità, etnia, religione, genere, fascia d'età, occupazione, educazione ecc.). Siamo, quindi, il prodotto dell'incrocio di differenti tipologie di identità; l'evoluzione di questa condizione è determinata da circostanze e da scelte coscienti quanto incoscienti².

• ESPERIENZE VISSUTE

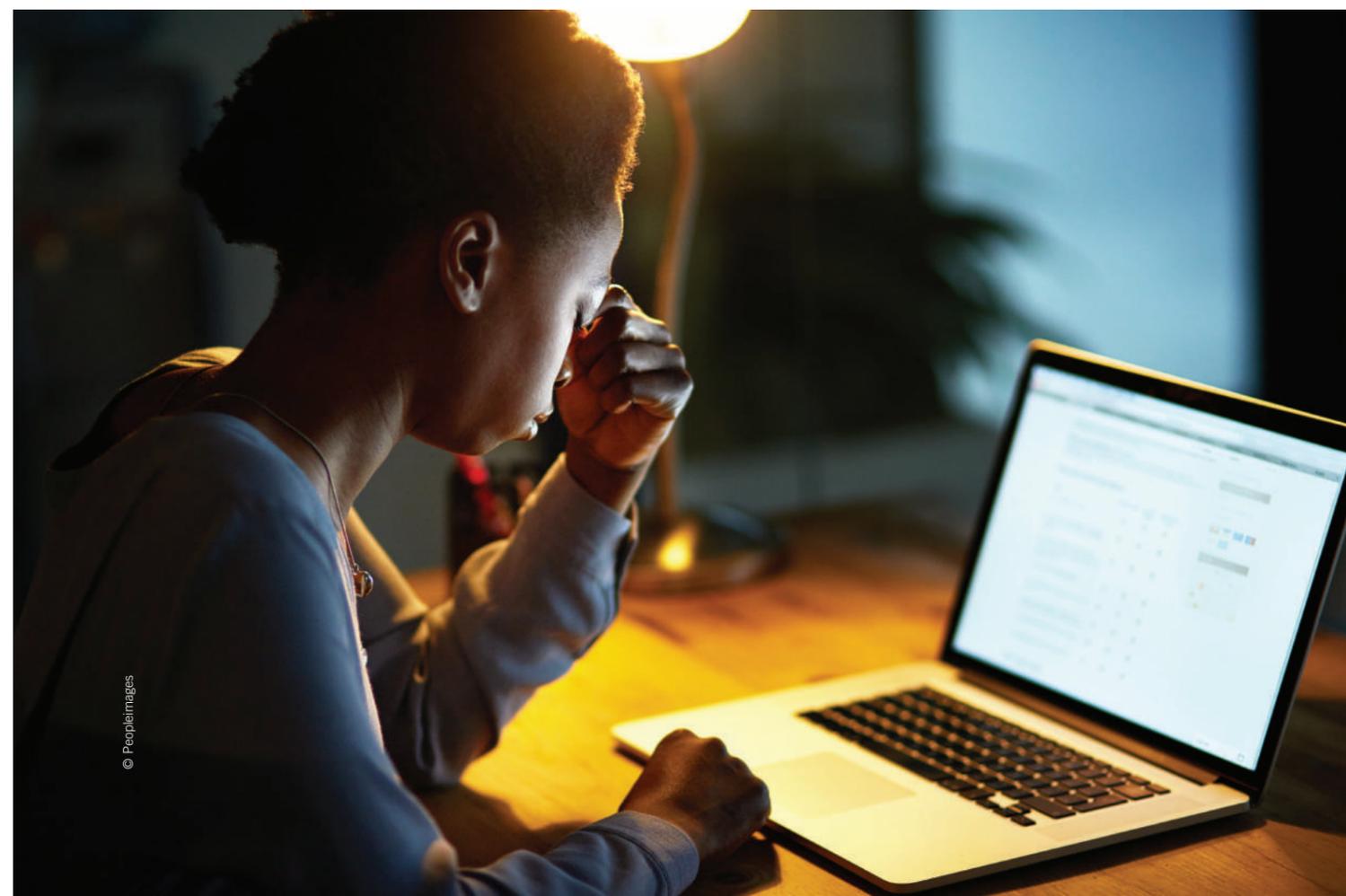
Gli eventi e le relazioni che viviamo nel corso della nostra esistenza, così come il significato che attribuiamo loro, possono modellare il nostro modo di affrontare le sfide. Per questo la maniera in cui ricordiamo e interpretiamo gli avvenimenti – il racconto che ne diamo – è importante quanto ciò che realmente è accaduto. Provare a comprendere e a tenere in considerazione le esperienze delle persone con cui stiamo comunicando può essere essenziale per sviluppare messaggi efficaci.

• VALORI

Per Treccani “dal punto di vista dei comportamenti sociali, si tende a considerare come *valore* ogni condizione o stato che l'individuo o più spesso una collettività reputa desiderabile, attribuendogli in genere significato e importanza particolari e assumendolo a criterio di valutazione di azioni e comportamenti”. Sulla base dei valori, quindi, identifichiamo ciò che è bene e ciò che è male, giusto e ingiusto, appropriato e inappropriato, diamo significati a ciò che viviamo e giudichiamo le nostre decisioni e azioni. All'interno di una stessa cultura o comunità possiamo trovare **valori in contrasto tra loro o divisivi**, tali da determinare una polarizzazione.

• CONVINZIONI

Definiamo convinzioni le idee che riteniamo siano vere. Nel corso della nostra esistenza costruiamo un insieme di postulati su come funziona il mondo: più facciamo esperienza nell'ambito di un certo argomento o con un determinato gruppo di persone, più le nostre convinzioni assumono consistenza e si arricchiscono di sfumature. Al contrario, quando la nostra esperienza rispetto a qualcuno o a qualcosa è limitata, potremmo ricorrere a generalizzazioni e a **analogie fondate su esperienze isolate** per colmare le nostre lacune conoscitive.



¹ Haidt J., *Felicità: un'ipotesi. Verità moderne e saggezza antica*, Codice Edizioni, 2008.

² Huber J. e Reynolds C., *Developing intercultural competence through education*, serie Pestalozzi n° 3, Consiglio d'Europa, 2014, p. 13, <https://rm.coe.int/developing-intercultural-enfr/16808ce258>



STEREOTIPI E PREGIUDIZI

I primi possono essere definiti come “**convinzioni relative a caratteristiche, attributi e comportamenti dell’insieme dei membri di determinati gruppi**”³. Uno stereotipo, anche se in apparenza positivo, può produrre un effetto negativo se riduce gli individui di uno stesso gruppo a un’unica dimensione, uniformandoli o *essenzializzandoli*, ossia spiegando le loro condotte e comportamenti attraverso *ciò che sono (essenza)*. Gli stereotipi associati a un gruppo conducono, nella generalità dei casi, allo sviluppo di pregiudizi nei confronti dello stesso.

N.B.

3 COSE DA SAPERE SUGLI STEREOTIPI

- Sono il frutto della commistione di diversi elementi conoscitivi relativi a uno stesso gruppo o a una categoria di persone e si consolidano nella nostra memoria.
- Sono, nella generalità dei casi, socialmente accettati e condivisi.
- Possono attraversare più periodi storici e persistere per lungo tempo.

I pregiudizi, come indica il nome, consistono in un “**un giudizio a priori, un’opinione preconcepita relativa a un gruppo di persone o a una categoria sociale**”⁴ e influenzano la disposizione d’animo, ossia l’inclinazione affettiva e morale, nei confronti di persone e situazioni.

Gli stereotipi e i pregiudizi possono, infine, portare alla discriminazione.

COME NASCE UN PREGIUDIZIO

“*I rom sono*”:
evocando questo stereotipo, per esempio, è possibile **provare rabbia o disgusto, emozioni che influenzano le azioni**. Questa combinazione può dare luogo a un pregiudizio.

³ Legal J.B. E Delouvee S., *Stéréotypes, préjugés et discriminations*, Dunod, 2016.

⁴ Amossy R. e Herschberg Pieerot A., *Stéréotypes et clichés*, Armand Colin, 2014, p. 22.



STEREOTIPI

Conoscenze:
componente cognitiva

PREGIUDIZI

Disposizione d’animo:
componente affettiva

DISCRIMINAZIONE

Azioni: componente
comportamentale



PER RIFLETTERE

- Siamo tutti condizionati da stereotipi e pregiudizi: quali sono i tuoi?
- Come puoi evitare che influenzino eccessivamente le tue attitudini e i tuoi comportamenti?



IL FRAME NARRATIVO

Se facciamo un passo indietro, notiamo che stereotipi e pregiudizi nei confronti di un gruppo si inseriscono in un **insieme più ampio che chiamiamo “narrazione”**.

La narrazione è definita come racconto di fatti, reali o immaginari, scelti e presentati in una sequenza coerente. In altri termini, corrisponde all'azione di “raccontare” una storia. Può avere forme narrative diverse (cinema, letteratura, musica, ma anche pubblicità, giornalismo ecc.). È composta da una serie di caratteristiche (contesto, personaggi, situazioni ecc.) e dai significati che le attribuiamo (politici, economici, sociali o di altra natura). Cambiando e combinando questi elementi (forma narrativa, caratteristiche, significati) possiamo ottenere molte e diverse modalità per raccontare gli stessi fatti.

Essa veicola, inoltre, **significati espliciti o sottintesi**.

È uno **strumento potente**, che può essere utilizzato anche in chiave negativa per mettere un gruppo di persone contro un altro, una maggioranza contro una minoranza, per alimentare leggende e pettegolezzi, per istigare l'odio e la violenza, per stigmatizzare. Per comprenderne la forza, basti pensare che in caso di situazioni di conflitto, per esempio, “l'immagine del nemico” (costruita attraverso una narrazione *ad hoc*) è spesso utilizzata per rafforzare la coesione nazionale, soffocare le controversie e/o permettere il mantenimento dello *status quo*.

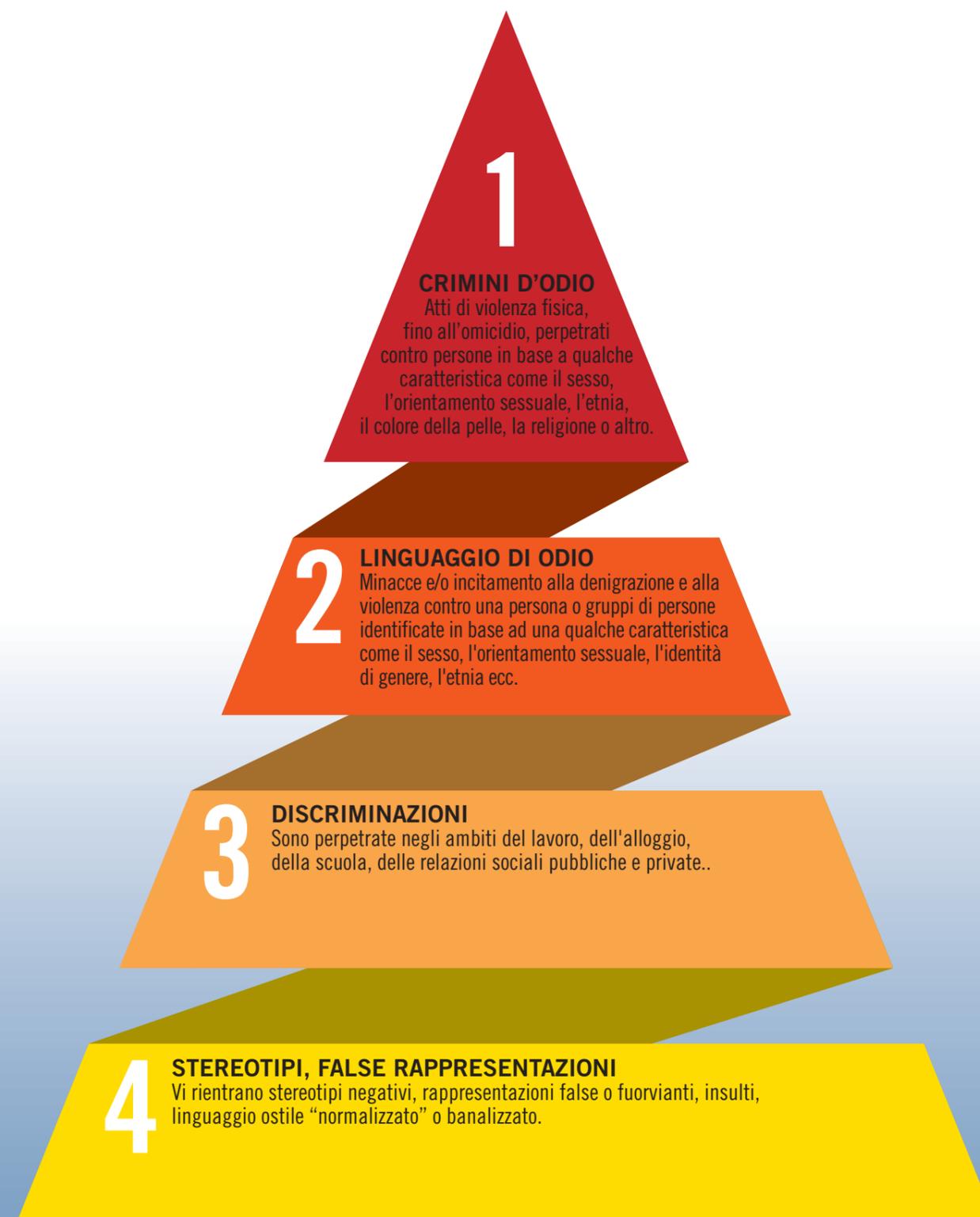
LA PIRAMIDE DELL'ODIO

Nel 2016 presso la Camera dei Deputati è istituita la Commissione sull'intolleranza, la xenofobia, il razzismo e i fenomeni di odio, che sarà intitolata a Jo Cox, deputata presso la Camera dei Comuni del Regno Unito uccisa il 16 giugno 2016.

La Commissione, presieduta dall'allora presidente della Camera Laura Boldrini, era composta da un deputato per ogni gruppo politico, da rappresentanti del Consiglio d'Europa, delle Nazioni Unite, dell'Istat, di centri di ricerca, di associazione ed esperti.

Nel corso dei suoi 14 mesi di attività, la Commissione ascolta 31 soggetti e acquisisce 187 documenti, sulla base dei quali elabora una relazione finale. Tale relazione riconosce “l'esistenza di **una piramide dell'odio** alla cui base si pongono stereotipi, rappresentazioni false o fuorvianti, insulti, linguaggio ostile ‘normalizzato’ o banalizzato e, ai livelli superiori, le discriminazioni e quindi il linguaggio e i crimini di odio”⁵. Questa ricostruzione (che riproduciamo a pagina 13) del processo che conduce alla produzione di discorsi e crimini d'odio, quindi, vede alla base il risultato del processo che unisce gli elementi finora analizzati: pregiudizi e stereotipi si inseriscono in un *frame* narrativo, il quale può dar luogo alle “rappresentazioni false o fuorvianti, insulti,

linguaggio ostile ‘normalizzato’ o banalizzato” appena citati. Questa base rappresenta il terreno ideale per la pratica e la perpetuazione di discriminazioni, a loro volta *humus* indispensabile affinché il linguaggio d'odio si affermi e siano compiuti crimini d'odio.



⁵ Camera dei Deputati, *La piramide dell'odio in Italia*, 2017 https://www.camera.it/applicazioni/xmanager/projects/leg17/attachments/shadow_primapagina/file_pdfs/000/007/099/Jo_Cox_Piramide_odio.pdf



2

RICONOSCERE L'HATE SPEECH

Quando ci troviamo di fronte a un tweet come “O sei maschio o sei femmina, altrimenti sei uno scherzo della natura” o a una foto che mostra una lunga fila di rifugiati, con la scritta “Ci stanno invadendo, per fermarli serve la forza”, ci troviamo di fronte a un discorso d'odio, cosiddetto *hate speech*.

Molto spesso, tuttavia, ci troviamo di fronte a contenuti che sebbene discriminatori, non sono altrettanto espliciti. Come comportarsi allora?

UNA DEFINIZIONE

Un'utile definizione di *hate speech* (la stessa alla quale fa riferimento la Commissione Jo Cox) ci giunge dalla Commissione europea contro il Razzismo e l'Intolleranza (Ecri), ed è contenuta nella Raccomandazione di politica generale n. 15 relativa alla lotta contro il discorso d'odio adottata l'8 dicembre 2015⁶:

“si intende per discorso dell'odio il fatto di fomentare, promuovere o incoraggiare, sotto qualsiasi forma, la denigrazione, l'odio o la diffamazione nei confronti di una persona o di un gruppo, nonché il fatto di sottoporre a soprusi, insulti, stereotipi negativi, stigmatizzazione o minacce una persona o un gruppo e la giustificazione di tutte queste forme o espressioni di odio testé citate, sulla base della “razza”, del colore della pelle, dell'ascendenza, dell'origine nazionale o etnica, dell'età, dell'handicap, della lingua, della religione o delle convinzioni, del sesso, del genere, dell'identità di genere, dell'orientamento sessuale e di altre caratteristiche o stato personale”.

Limitarsi alla ricerca di una corrispondenza tra il contenuto che abbiamo di fronte e questa (o altre) definizioni, tuttavia, può rivelarsi insufficiente quando abbiamo la necessità di stabilire confini precisi entro i quali muoverci, sulla base degli obiettivi della nostra azione.

Se, per esempio, intendiamo segnalare con successo un commento pubblicato su un social network, dobbiamo conoscere le *policy* della piattaforma per capire se il contenuto rientra o meno tra i casi consentiti; se, invece, vogliamo “rispondere” al commento (con una forma di contro-narrazione o di narrazione alternativa, come vedremo più avanti) è bene essere in grado di analizzare al meglio i vari elementi che si celano dietro alle parole che riteniamo essere *hate speech*. Tra questi vi sono i messaggi che il contenuto veicola.

I MESSAGGI

Per capire se ci troviamo di fronte a un caso di *hate speech*, può essere utile individuare i messaggi che, in modo più o meno esplicito, fornisce. Al di là del suo contenuto intrinseco, infatti, l'*hate speech* veicola altri due messaggi⁷.

Il primo è **indirizzato al gruppo attaccato**: ha lo scopo di compromettere il sentimento di sicurezza e libertà delle persone o dei gruppi presi di mira, inducendoli a pensare che non vi sia spazio (ossia che non possano essere accettati e/o integrati) per loro in una determinata società. L'altro messaggio è indirizzato alla **fetta di comunità non identificabile con il gruppo o la categoria sociale attaccati**: l'obiettivo è veicolare l'idea che le opinioni alla base del discorso d'odio siano largamente condivise, anche se non sempre espresse pubblicamente.

In presenza di questo tipo di messaggi, ci troviamo di fronte a un discorso d'odio.

Col tempo questi messaggi diventano componenti del tessuto sociale, contribuendo alla formazione di un terreno fertile per discriminazione, crimini d'odio, violazioni dei diritti umani. L'esistenza di relazioni sociali e di potere squilibrate concorre a causare lo sviluppo di tale condizione, che spesso arriva a definirsi in modo insidioso, poco alla volta, nel corso di un lungo periodo, agendo come un veleno il cui effetto tossico emerge progressivamente⁸. Se facciamo un passo indietro e guardiamo alla piramide dell'odio, ci rendiamo conto che questi messaggi sono quelli che incontriamo nelle false rappresentazioni e in quel linguaggio ostile “normalizzato” o banalizzato, fundamenta dei gradini più alti dei fenomeni d'odio.

⁸ Escande C., *Actualité et prémisses de discours totalitaires. La fabrique des égarés*, Cairn, 2011.

MEMO

È HATE
SPEECH?
RICORDA
CHE...

- Manca una **definizione ampiamente condivisa, precisa ed esaustiva** che distingua tra i discorsi d'odio che, in quanto tale, non devono trovare spazio e la libertà di espressione, che va difesa anche quando il pensiero espresso disturba qualcuno;
- i criteri cambiano fra le leggi nazionali, i documenti internazionali, i codici deontologici (es. dei giornalisti) o i codici di autoregolamentazione dei social network e mutano nel tempo;
- il discrimine principale consiste nell'esistenza di un rischio concreto di promuovere **discriminazione, odio, violenza o segregazione** prodotto dalle espressioni di incitamento all'odio, alla violenza e all'estremismo;
- nel valutare la pericolosità contano vari aspetti, per esempio se **l'odio è rivolto a persone a rischio di discriminazione**.

⁶ <https://rm.coe.int/ecri-general-policy-recommendation-no-15-on-combating-hate-speech-ital/16808b5b04>

⁷ Waldron J., *The harm in hate speech*, Harvard University Press, 2014.



LA NARRAZIONE DELL'ODIO

Sappiamo a questo punto cosa sono **narrazione e hate speech**, osserviamo allora il modo in cui, facendo convergere questi due elementi, è possibile costruire una narrazione dell'odio e come questa possa essere finalizzata al raggiungimento di uno scopo.

La narrazione corrisponde al racconto di fatti, reali o immaginari, scelti e presentati in una sequenza coerente, con uno scopo generale e possibili obiettivi specifici. L'*hate speech* vi trova spazio quando, nello scopo generale oppure tra gli obiettivi specifici del racconto, vi è quello di colpire un gruppo particolare di persone o una persona in particolare. Se, per esempio, vogliamo raccontare il fenomeno dell'immigrazione con lo scopo di rafforzare le politiche di chiusura, potremmo ricorrere a una narrazione che, oltre a enfatizzare le problematiche relative alla gestione dell'immigrazione, attacchi i migranti con affermazioni discriminatorie e disumanizzanti (*hate speech*) con l'obiettivo di renderle invisibili alla popolazione residente (se l'obiettivo specifico sarà raggiunto, sarà più semplice perseguire anche lo scopo generale).

Quando ci troviamo di fronte a casi di *hate speech* che rientrano in una vera e propria narrazione dell'odio, come nel caso di contenuti discriminatori diffusi da personaggi di pubblico rilievo, è molto utile provare a individuare quali sono gli obiettivi specifici o lo scopo generale della campagna condotta da chi li promuove che li promuove: in questo modo potremo cercare di impostare meglio un nostro eventuale intervento nel dibattito sul tema e ci sarà d'aiuto per capire quali messaggi chiave utilizzare.



N.B.

3 CARATTERISTICHE DELL'ODIO ONLINE

L'*hate speech* esiste da secoli, ma **muta nei bersagli, nelle forme, nei modi di circolare**. Quello online pone sfide inedite per come nasce e si diffonde. Le specificità del mezzo che utilizza⁹ fanno sì che idee che un tempo non avrebbero trovato uno spazio significativo in cui essere espresse, oggi attraverso il web possano raggiungere un pubblico ampio e avere una grandissima visibilità.

I social network hanno una struttura "orizzontale", che li rende una cassa di risonanza per i contenuti che vi circolano. La brevità dei contenuti e la velocità nella loro redazione fa sì che i messaggi siano spesso semplificati e estremizzati. Inoltre, l'esistenza delle cosiddette *echo chamber*, spazi virtuali (per esempio i gruppi Facebook) in cui sono diffuse informazioni volte a rafforzare un unico punto di vista rispetto a un certo tema, facilita la polarizzazione delle opinioni degli utenti¹⁰.

⁹ Gagliardone I., Danit G., Alves T. e Martnez G., *Combattere les discours de haine sur internet*, Organisation des Nations Unies pour l'éducation, la science et la culture, 2015.

¹⁰ Per approfondire è utile il paper "Echo Chambers on Facebook" di Walter Quattrociocchi, Antonio Scala e Antonio Scala e Cass R. Sunstein, 2016 <https://poseidon01.ssrn.com/delivery.php?ID=412117110095126112115067067102004011022024001018005001005071083094009016071073089025052011123123001038027091102124107090112074041023045020021083065077086078003065008011009009071120013125114009031068030125016120089071098109126105064081019009011117090&EX T=pdf>

Ecco le 3 principali caratteristiche che rendono l'*hate speech* online un fenomeno da combattere tanto difficile quanto urgente:

- 1 La permanenza nel tempo:** il discorso d'odio può restare online per molto tempo, in differenti formati e su più piattaforme; più a lungo resta accessibile, maggiore è il suo potenziale in termini di danni.
- 2 Itinerante e ricorrente:** l'architettura delle piattaforme dove l'*hate speech* dimora influenza molto la dinamica della sua diffusione, che può essere itinerante e ricorrente. Un contenuto rimosso, infatti, può apparire sotto un altro nome e/o titolo sulla stessa piattaforma o altrove.
- 3 Dietro a un monitor:** le idee di anonimato e di impunità associate all'utilizzo di internet e alimentate dalle modalità di interazione sui social network favoriscono l'espressione di opinioni di odio. Chi fa *hate speech* spesso non associa conseguenze dirette ai propri atti e non percepiscono il potenziale impatto dei messaggi d'odio sulla vita reale delle persone; se fossero offline si comporterebbero allo stesso modo? Probabilmente non sempre.



N.B.

CONDIVIDERE O NON CONDIVIDERE?

Quando un contenuto incontrato online ci indigna, **una reazione istintiva può essere quella di dividerlo**, per prendere una posizione critica o di condanna nei confronti del contenuto stesso. Ma è sempre opportuno?

L'Ethical Journalism Network ha elaborato un test in 5 punti¹¹ al quale ogni giornalista dovrebbe sottoporsi prima di riportare un caso di *hate speech*, riflettendo sull'impatto del messaggio che veicola. Un test che si rivela utile per chiunque utilizzi in modo costante i social network, dal momento che è sufficiente avere un discreto numero di persone che ci seguono per far giungere ciò che possiamo lontano, sia per numeri che per distanza.

Ecco i cinque elementi sui quali riflettere (tutti strettamente connessi tra loro) per capire quale potrebbe essere l'impatto del discorso d'odio e se, di conseguenza, può essere più utile – rispetto all'obiettivo di contrastare la discriminazione - dividerlo per prenderne le distanze oppure no.

1 Lo status dell'*hater*

Chi diffonde il discorso d'odio è in una posizione tale da influenzare le ragioni per cui promuove un certo messaggio d'odio? Ha un ruolo istituzionale?

2 La portata del discorso

Quanto lontano arriverà il discorso? L'*hater* è una persona che, attraverso il mezzo scelto, in questo caso il web, realisticamente raggiungerà molte persone oppure no?

3 Obiettivi del discorso

Il discorso intende in modo deliberato causare un danno a una persona o a un gruppo di persone? L'*hater* in che modo trae beneficio da tale discorso, qual è il suo interesse?

4 Il contenuto stesso

Incita alla violenza e, in tal senso, costituisce un pericolo? Quali sono le sue caratteristiche intrinseche?

5 Contesto sociale-economico-politico

Chi potrebbe essere colpito negativamente dal messaggio? Il contesto sociale-economico-politico in cui il contenuto è diffuso, rende il bersaglio del discorso vulnerabile? Tale contesto espone il bersaglio al rischio (presunto o concreto) di subire forme di discriminazione/violenza/esclusione?

Più avanti nella guida approfondiremo meglio questo tipo di analisi anche ai fini dell'attività pratica di contrasto all'*hate speech* (vedi pag. 27, "Questione di gravità").

¹¹ <https://ethicaljournalismnetwork.org/resources/infographics/5-point-test-for-hate-speech-english>

3

DIS/INFORMAZIONE ONLINE

Nell'ambito della comunicazione, un'informazione è definita come un fatto riferito da fonti identificabili, verificabili e quindi controllabile. Alle informazioni possono essere associate spiegazioni e/o interpretazioni che pongono i fatti in questione all'interno di una cornice sociale, culturale, economica, politica, geopolitica ecc. L'inevitabile scelta dell'angolazione dalla quale analizzare i fatti, fa sì che l'informazione non possa essere totalmente oggettiva e neutrale¹².

Con la crescita di internet e l'avvento dei social network sono cambiate le modalità attraverso le quali l'informazione è prodotta e diffusa:

- grazie a una tecnologia ampiamente accessibile **ogni utente può creare contenuti** e farli circolare in modo semplice;
- il **consumo** dell'informazione non è più privato, bensì pubblico;
- la **velocità** attraverso la quale è diffusa l'informazione è altissima e aumentata dall'utilizzo di smartphone e tablet;
- il fatto che l'informazione circoli con estrema rapidità tra i fruitori **riduce le possibilità che essa venga messa in discussione**¹³.

Un fenomeno del quale si è sentito parlare molto negli ultimi anni, è quello relativo alla circolazione di **false notizie**, le cosiddette *fake news*. L'attenzione mediatica è divenuta tale da far sì che nel 2017 questa definizione fosse scelta da Collins Dictionary come parola dell'anno. Il fenomeno della diffusione di false notizie messe in circolazione con determinate finalità tuttavia non è nuovo, ma anch'esso ha senz'altro subito un'evoluzione con l'ascesa del web e dei social media.

Si tratta di un termine da utilizzare con prudenza. In primo luogo perché è

¹² Le Centre pour l'Education aux Médias et à l'Information (CLEMI), « Qu'est ce qu'une info ? » 2018.

¹³ Wardle C. e Derakhshan H., *Information disorder: Toward an interdisciplinary framework for research and policy making*, Consiglio d'Europa, 2017.

¹⁴ Wardle C. e Derakhshan H., *Information disorder: Toward an interdisciplinary framework for research and policy making*, Consiglio d'Europa, 2017, <https://rm.coe.int/information-disorder-toward-an-interdisciplinary-framework-for-research/168076277c>

COME SI DIFFONDE UN CONTENUTO¹⁴





una definizione inadatta a descrivere **un fenomeno tanto complesso come quello della disinformazione**, mosaico composto da moltissimi tasselli, online e offline. Il termine *fake news*, inoltre, è stato ripreso da numerosi politici per screditare le informazioni che non perseguono i loro interessi, divenendo una forma di attacco nei confronti della libertà di stampa. È quindi necessario esaminare con attenzione i casi di disinformazione, volta per volta, tenendo in considerazione questo contesto.

Vi è, indubbiamente, un legame tra disinformazione e *hate speech* online: **la disinformazione, nelle sue varie forme, può alimentare gli assunti alla base di affermazioni discriminatorie e d'odio.**

SIAMO FACT CHECKER ANCHE NOI

Condividendo false informazioni, inganniamo coloro che si fidano di noi. Contribuiamo così a nutrire la disinformazione.

Per questo di fronte a informazioni o notizie che ci sembrano ambigue o poco attendibili, è bene provare a svolgere una verifica. Esistono numerosi portali e siti che ci vengono in aiuto (una lista di fonti utili è a pag. 40).

FOCUS

¹⁵ Wardle C. e Derakhshan H., *Information disorder: Toward an interdisciplinary framework for research and policy making*, Consiglio d'Europa, 2017, <https://rm.coe.int/information-disorder-toward-an-interdisciplinary-framework-for-research/168076277c>

N.B.

DISINFORMAZIONE ONLINE

Le tipologie di disinformazione nelle quali ci possiamo imbattere nel web sono molte. Di seguito le più comuni e rilevanti.

- 1 Contenuti “confezionati”:** quando un contenuto è fabbricato e confezionato da zero, completamente falso, con l'obiettivo di ingannare il pubblico e creare pregiudizi.
- 2 False connessioni:** quando il titolo non corrisponde al contenuto.
- 3 Contesti e descrizioni falsi:** quando informazioni vere e illustrazioni autentiche che non hanno legami tra loro sono

associati per perseguire un certo obiettivo; quando una foto o un video sono pubblicati con una descrizione erronea.

- 4 Fonti false:** quando i portali d'informazione sono contraffatti¹⁵.



PER RIFLETTERE

- In che modo ti informi? Come condividi le informazioni?
- I contenuti condivisi sui social network dei tuoi contatti sono una delle tue fonti? Quanto ti fidi di questi contenuti?
- Consulti mai media che non riflettono la tua visione politica e sociale?

4

EFFETTI OLTRE LA RETE

LE CONSEGUENZE DELL'HATE SPEECH

Quando si parla di discorsi d'odio online è importante considerare che il loro impatto non si limita all'influenza che possono avere sul dibattito in rete relativo a certi temi sensibili. L'*hate speech* è in grado di produrre **effetti e ripercussioni sulle singole persone attaccate, così come sull'intera comunità.**

UGUALI IN DIGNITÀ E DIRITTI: DALLA DICHIARAZIONE UNIVERSALE DEI DIRITTI DELL'UOMO

La Dichiarazione universale dei diritti dell'uomo¹⁶ sancisce in modo chiaro il diritto alla non discriminazione di ogni individuo. Di seguito gli articoli della Dichiarazione a cui possono essere ricondotte le violazioni perpetrate per mezzo dell'*hate speech* nei confronti di un singolo individuo o di un gruppo di individui.

ART. 1

*“Tutti gli esseri umani nascono **liberi ed eguali in dignità e diritti**. Essi sono dotati di ragione e di coscienza e devono agire gli uni verso gli altri in spirito di fratellanza”.*

ART. 2

*“Ad ogni individuo spettano tutti i diritti e tutte le libertà enunciate nella presente Dichiarazione, **senza distinzione alcuna, per ragioni di razza, di colore, di sesso, di lingua, di religione, di opinione politica o di altro genere, di origine nazionale o sociale, di ricchezza, di nascita o di altra condizione**.”*

Nessuna distinzione sarà inoltre stabilita sulla base dello statuto politico, giuridico o internazionale del paese o del territorio cui una persona appartiene, sia indipendente, o sottoposto ad amministrazione fiduciaria o non autonomo, o soggetto a qualsiasi altra limitazione di sovranità”.

ART. 5

*“Nessun individuo potrà essere sottoposto a tortura o a **trattamento o a punizioni crudeli, inumani o degradanti**”.*

ART. 7

*“Tutti sono eguali dinanzi alla legge e hanno diritto, senza alcuna discriminazione, ad una eguale tutela da parte della legge. Tutti hanno diritto ad una **eguale tutela contro ogni discriminazione che violi la presente Dichiarazione** come contro qualsiasi incitamento a tale discriminazione”.*

FOCUS



L'IMPATTO SULLE SINGOLE PERSONE

L'*hate speech* è fondato su elementi di discriminazione, per questa ragione colpisce nella dignità le persone attaccate e mina il loro senso di sicurezza, in quanto individui già considerati socialmente vulnerabili e potenzialmente vittime di discriminazioni multiple.

Quando l'attacco perpetrato attraverso l'*hate speech* è personale assume spesso la forma di trattamenti disumanizzanti o degradanti.

I discorsi d'odio si trasformano in **etichette con le quali è assegnata un'identità alle persone**; un marchio che non viene apposto senza conseguenze: "L'attribuzione da parte di altri di identità inadatte e il divario tra queste e la percezione che ognuno ha di sé nuocciono al benessere psicologico e all'adattamento sociale dell'individuo"¹⁷.

L'IMPATTO SULLA COMUNITÀ

L'*hate speech* contribuisce a una narrazione che relega all'inferiorità chi viene ricondotto (dall'*hater*) ad alcune categorie sociali¹⁸.

In questo modo i discorsi d'odio creano e alimentano **"condizioni che indeboliscono il valore dei diritti e delle libertà delle persone stigmatizzate"**, la loro possibilità di vivere senza ostacoli e di difendere i loro interessi civili. L'*hate speech* attacca i diritti che garantiscono una eguale partecipazione alla vita civile e un "dibattito pubblico aperto dove gli interessi di tutti sono presi in considerazione e ciascuno ha una voce che sarà ascoltata" (l'eguaglianza nella dignità e nei diritti, il diritto di prendere parte alla direzione degli affari pubblici del paese, la partecipazione alla vita culturale della comunità ecc.).

I discorsi d'odio possono avere, quindi, un **impatto negativo sulla qualità della vita sociale**.

¹⁷ Huber J., Reynolds C., *Développer la compétence interculturelle par l'éducation*, Serie Pestalozzi n° 3, Consiglio d'Europa, 2014, <https://rm.coe.int/developing-intercultural-enfr/16808ce258>

¹⁸ Girard C., « *Pourquoi punir les discours de haine?* », Esprit, 2015.



© Amnesty International / Richard Burton

SECONDA PARTE

CONTRASTARE





1

COMBATTERE L'HATE SPEECH ONLINE

DA DOVE INIZIO?

Dopo aver passato in rassegna alcune nozioni e elementi teorici, per avvicinarci alla conoscenza di un fenomeno complesso come quello dell'*hate speech*, nella seconda parte di questa guida passiamo a esaminare aspetti pratici e a fornire consigli su come contrastare concretamente l'odio online.

OSSERVARE

Dopo essersi informati per conoscere in linea teorica il fenomeno, è importante osservarlo: in questo modo è possibile imparare a comprenderlo, a riconoscerlo, a individuare le modalità di intervento più adatte, a sviluppare risposte pertinenti e efficaci. **Osservare attivamente** ciò che accade sui social network, quindi, è il secondo passo da compiere. L'attività di osservazione consente, infatti, di "misurare la temperatura" della società, di identificare e seguire le tendenze nazionali e internazionali e di analizzare come il web reagisce ai fatti di attualità.

ORGANIZZARE

Nella fase di osservazione attiva, possiamo iniziare a individuare alcuni elementi per **circoscrivere** la nostra attività di contrasto all'*hate speech*.

• DOVE?

Su quali piattaforme online vogliamo concentrare la nostra attività? Social network (Facebook, Instagram, X, YouTube, TikTok ecc.), forum, portali d'informazione? Per iniziare può essere una buona idea scegliere una sola piattaforma, ampliando il proprio raggio d'intervento solo se e quando si è pronti per farlo.

• COSA?

Quali sono i temi sui quali vorreste focalizzarvi? Anche in questo caso partire da un solo argomento (discriminazione di genere, islamofobia, antiziganismo, omobitransfobia ecc.) per poi allargare lo spettro delle tematiche monitorate è consigliabile.

• CON QUALI STRUMENTI?

Per aiutarvi a individuare i contenuti di vostro interesse, potreste ricorrere a strumenti quali l'impostazione di notifiche su Google (Google Alerts, che consente di essere avvisati ogni volta che vengono pubblicati online contenuti con le parole chiave e le caratteristiche da voi indicati).

In alternativa, alcuni social network consentono di cercare contenuti pubblicati attraverso l'indicazione di parole-chiave.

In caso aveste la necessità di verificare delle notizie e/o informazioni, potete ricorrere ad alcuni portali che offrono dati e informazioni attendibili: in appendice a questa guida trovate una lista utile.

SICUREZZA TECNOLOGICA

Prima di iniziare a contrastare l'*hate speech* online, occorre prendere in considerazione dei **fattori di rischio**. Se vogliamo, per esempio, provare ad arginare la degenerazione del discorso tossico inserendovi in dibattiti in cui il livello di odio è elevato (su social network, portali d'informazione, forum ecc.), dovremmo chiederci per prima cosa se siamo sicuri di voler intraprendere questa battaglia utilizzando un account che riconduce direttamente alla nostra identità. Si tratta di una scelta personale, tuttavia, se si vuole svolgere questa attività in modo continuativo, è **consigliabile evitare di esporsi personalmente**.

MEMO

TUTELA LA TUA IDENTITÀ

- Evita l'utilizzo di **indirizzi email personali** al momento della creazione di un account sui social network.
- Scegli **password sicure** e tra loro diverse per ogni account e indirizzo email.
- Crea un **nuovo profilo** per portare avanti l'attività di contrasto all'*hate speech* sul social network scelto.
- Non rispondere a commenti e messaggi utilizzando account e indirizzi personali.
- Se gestisci un sito o una pagina web (un blog per esempio), assicurati che non vi siano informazioni tali da poter ricondurre alla tua identità.
- Quando ti colleghi, ricorri alla **VPN – Virtual Private Network** (online sono disponibili applicazioni e programmi scaricabili gratuitamente) per mascherare il tuo indirizzo IP.
- Quando pubblichi un contenuto, alcune applicazioni inseriscono anche la **localizzazione**: assicurati preventivamente di aver disattivato questa funzione.
- Non indicare tra le informazioni sui social network (o su altre piattaforme che prevedono questa possibilità) il tuo indirizzo privato, né quello della tua occupazione professionale.



SICUREZZA PSICO-EMOTIVA

Contrastare il discorso d'odio può essere impegnativo sotto il profilo psicologico. È importante, dunque, prendere alcuni accorgimenti, a partire dalla **gestione del tempo da dedicare all'attività**, che dovrebbe essere limitato: imponetevi di non trascorrere più di alcune ore al giorno contrastando l'*hate speech online*; evitate di farlo durante la notte, soprattutto se vivete soli.

Considerate poi che i contenuti di fronte ai quali potreste trovarvi possono turbarvi e, a lungo andare, esporvi a livelli di stress non gestibili. Ponetevi, quindi, dei limiti anche rispetto alla **quantità di contenuti** d'odio ai quali reagire: è davvero necessario leggerne di supplementari o visionare un ulteriore video? **Stabilite quando ne avete abbastanza e non andate oltre**, ma spegnete il computer (o mettete da parte smartphone e tablet) e dedicatevi ad altre attività capaci di distrarvi. Passeggiate, meditate, praticate sport, leggete, dipingete: svolgete qualsiasi attività che riesca a farvi davvero "staccare", dimenticando l'*hate speech* e tutto ciò che vi ruota intorno.

È essenziale, inoltre, non essere soli: **agire in gruppo** può essere una soluzione in grado di offrire sicurezza e sostegno sotto il profilo emotivo e emozionale. Poter condividere le esperienze con chi contrasta i discorsi d'odio, supportandosi in modo reciproco, consente di evitare il senso di solitudine e lo scoraggiamento che possono accompagnare attività stressanti dal punto di vista psico-emotivo.

Quando siete testimoni di discorsi d'odio che si configurano come attacchi personali, intervenite valutando caso per caso qual è la modalità migliore. Potete decidere di segnalare il caso al solo social network o moderatore oppure potreste riportarlo – se sussistono le condizioni – alla Polizia postale. Fate sentire alla persona presa di mira che non è sola, suggerendo anche, quando lo riterrete adeguato, dove poter trovare sostegno di tipo psicologico e/o legale.

ATTENZIONE AL CROLLO EMOTIVO

Non adottare misure adeguate a controllare il livello di stress che può essere generato da un'attività come quella del contrasto all'odio online può condurre al cosiddetto *burn-out*. Questo fenomeno è definito da Treccani come "sindrome da stato patologico che si verifica in individui che svolgono professioni di aiuto. Ne sono interessati medici, poliziotti, infermieri, psicologi ecc. Compare in figure professionali che devono sostenere in modo adeguato il proprio stress psicoemotivo e quello della persona assistita. Se la fase di logoramento psicologico non è gestita o non risulta controllata, si osserva una progressione del danno psichico e fisico [...]".

FOCUS



2

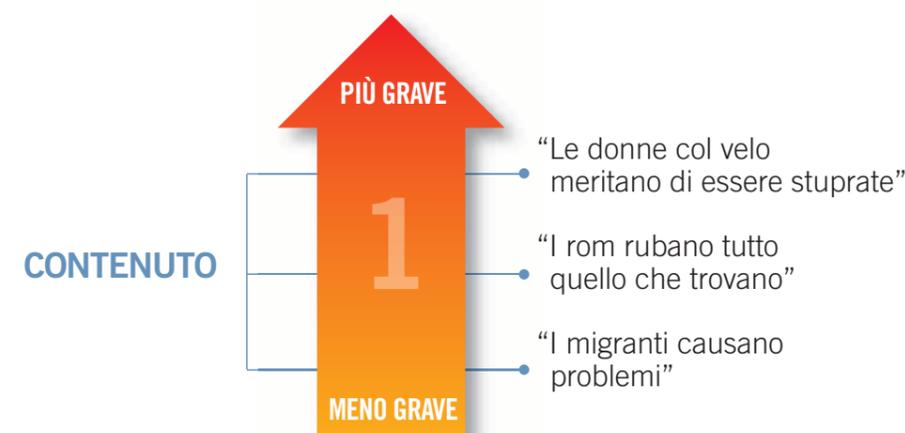
PASSARE ALLA PRATICA

SUGGERIMENTI PRATICI

Vi siete informati, avete osservato questo fenomeno, considerato elementi e fattori di rischio e organizzato la vostra attività. Prima di iniziare, però, è bene fare un ultimo passaggio: individuare un metodo pratico per riconoscere più facilmente l'*hate speech*.

QUESTIONE DI GRAVITÀ

Per comprendere il **grado di tossicità** delle parole e dei messaggi incontrati, potete far riferimento alla seguente griglia di lettura, basata su 5 aspetti la cui analisi ci aiuta nella valutazione.



Alcune frasi, per il loro **contenuto intrinseco**, possono essere considerate più gravi di altre. Un insulto moderato è meno grave dell'incitamento ad agire contro una comunità.





Una stessa frase può essere più o meno grave a seconda del contesto in cui è diffusa. Possiamo individuare **due tipologie di contestualizzazione**. La prima rinvia a differenti aspetti legati alla disseminazione del messaggio: la **visibilità** dell'*hater* e la **tipologia di mezzo** usato per la diffusione del messaggio, che concorrono a determinarne la **portata** (il pubblico che può raggiungere).

Altrettanto importante è la contestualizzazione relativa allo specifico *background* storico, politico, sociale, economico nel quale si inquadra il contenuto e all'**attualità**.



La stessa frase può essere più o meno grave in funzione del **tono** utilizzato.

È possibile che l'autore di un'affermazione discriminatoria possa non averlo considerato o che, una volta compresa la violenza del tono, se ne penta.

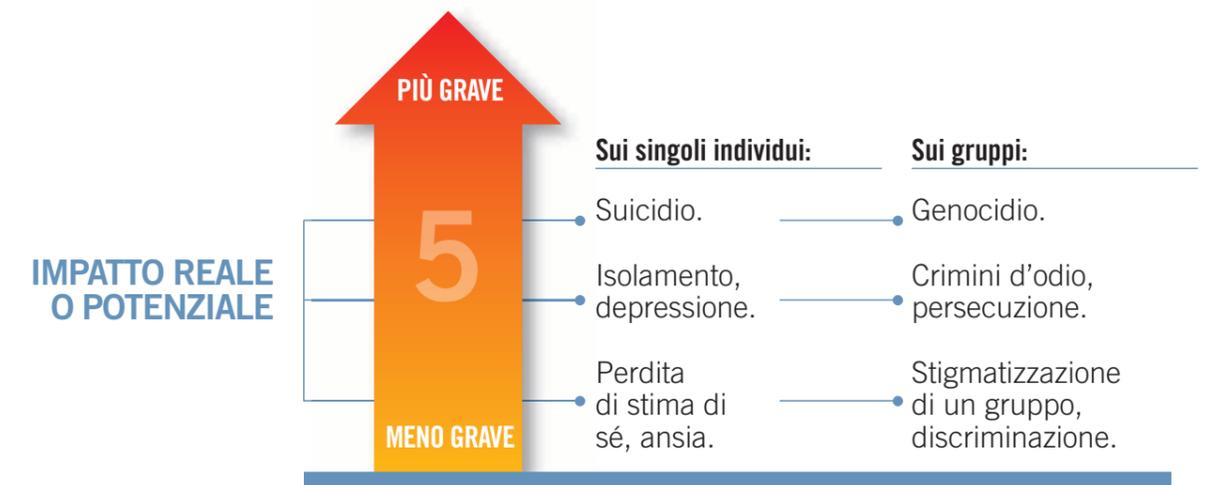
Quando, al contrario, una persona agisce con l'**intenzione di promuovere contenuti discriminatori e di nuocere** a qualcuno, la stessa frase acquisisce maggiore gravità. A parità di tono l'intenzione fa la differenza, come nell'esempio sopra.

Attenzione: **l'impatto dell'*hate speech* non è necessariamente proporzionale alla violenza del tono e del contenuto**. Per esempio, un'argomentazione ingannevole inserita in un discorso sui problemi sanitari e sociali all'interno degli insediamenti rom, espressa in una forma educata, può fare più danni rispetto a insulti e commenti percepiti come estremi in modo immediato.



Sulla base del modo in cui gli individui o i gruppi sociali sono **percepiti all'interno della società**, alcuni possono essere considerati più vulnerabili di altri, come nel caso delle **minoranze**.

Per esempio, un insulto rivolto alle persone cristiane sarà più pregiudizievole in un paese in cui la maggioranza appartiene a un'altra confessione religiosa. La stessa espressione applicata a persone o a gruppi diversi può avere effetti differenti.



Per valutare al meglio l'impatto reale o potenziale di un'affermazione tossica, è necessario provare a **guardarla dalla prospettiva della persona o del gruppo attaccato**. Un individuo esterno può tendere a minimizzare l'impatto; dire, per esempio, che una frase contenente *hate speech* è stata scritta “per gioco” o “per ridere” può essere profondamente mortificante per chi è stato preso di mira.





MEMO

COSA TENERE A MENTE NEL VALUTARE L'IMPATTO

Nella prima parte di questa guida, dedicata alla conoscenza del fenomeno, ci siamo soffermati sull'importanza della valutazione dell'impatto che il discorso d'odio può avere, per capire se e quando può essere opportuno condividere e condannare o meno un contenuto. In questa seconda parte, invece, abbiamo ripreso il tema della valutazione dell'impatto in un'altra prospettiva: quella di comprendere quale può essere la migliore strategia di reazione, sulla base del rischio presunto o concreto generato dall'*hate speech*.

Concentriamoci su:

- il contenuto e il tono, poiché vi è grande differenza tra generalizzazioni negative ("Gli immigrati non lavorano") e insulti moderati e espressioni estreme ("Troia, domani ti stupro");
- l'intenzione dell'autore, perché, per esempio, "I gay sono contronatura" è grave se scritto in un'email indirizzata a un amico, ma è ancora più grave e può rappresentare un pericolo se postato sulla pagina Facebook di una persona gay;

- i bersagli, che possono essere gruppi o individui vulnerabili per come sono considerati nella società, per come sono raccontati e rappresentati nei media e per la loro situazione personale (l'esistenza dell'antisemitismo rendere la frase "Gli ebrei sono avidi" più pericolosa de "I politici sono avidi");
- il contesto che è alla base di tutte le considerazioni che possiamo fare, poiché le circostanze storiche e culturali, il mezzo impiegato, il pubblico raggiunto, l'autorità di chi pubblica, le tensioni verso il gruppo bersaglio costituiscono, insieme, ciò che rende più o meno pericoloso un messaggio ("Gli immigrati rovinano il paese" è male sul blog di un 16enne, ma è molto peggio come tweet di una persona che rappresenta il governo);
- la valutazione dell'impatto reale o potenziale su individui, gruppi, società deve tenere in considerazione la percezione di chi è colpito dal discorso d'odio, che può contare più delle impressioni dell'autore o di terzi rispetto al contenuto stesso.



N.B.

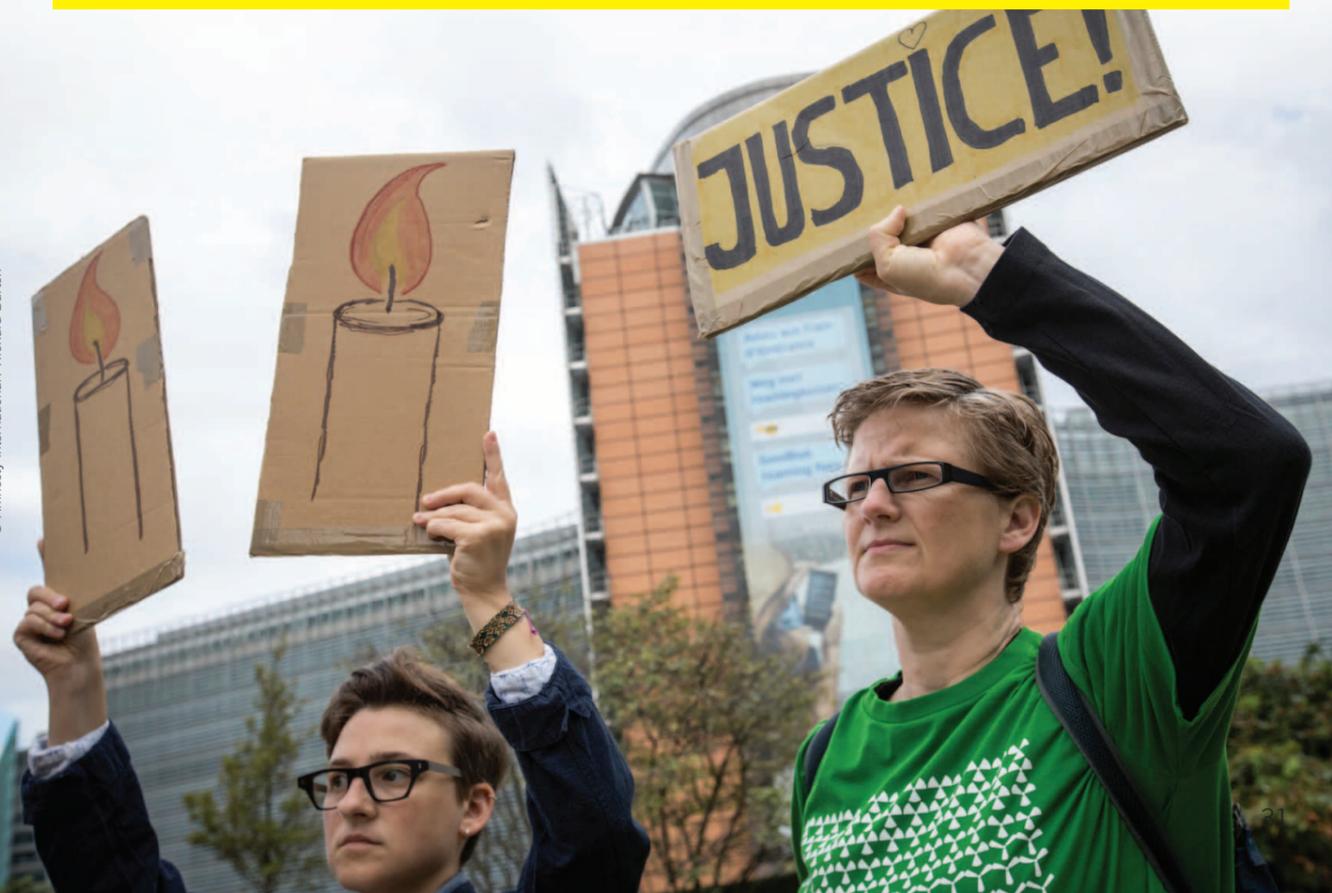
7 INDICATORI D'ODIO NEL LINGUAGGIO

Spia di possibile *hate speech* è anche il linguaggio usato (che però va sempre valutato nel contesto). Oltre alle chiare espressioni offensive è possibile incontrare:

- 1 Parole neutre o positive rese spregiative** da certi usi o alterazioni. Esempi: coso, giovinastra, intellettualoide, pseudoattore.
- 2 Il richiamo a false credenze:** il migrante contagioso, la zingara rapitrice di bimbi.
- 3 Il raggruppare i soggetti** in «noi» (vittime) e «loro» (minaccia),

presentati come gruppi omogenei e alternativi, con uso enfatico di «noi, voi, loro, questi, 'sti, quelli».

- 4 Il far sentire supportato chi aggredisce** e lontane le vittime, minimizzando gli atti d'odio "muoiono in continuazione".
- 5 Ridurre il bersaglio a una sua caratteristica:** una condizione momentanea "disperato", uno stato giuridico "clandestina", la nazionalità.
- 6 Numeri distorti** o presentati con falsa precisione che dà un illusorio senso di oggettività.
- 7 L'accostamento di due affermazioni** per farle sembrare collegate quando non lo sono: «Gli immigrati fanno lo sciopero della fame? In Italia 5 milioni di poveri faticano a mangiare».





3

REAGIRE ALL' *HATE SPEECH*

C'È UN MODO GIUSTO?

Reagire all' *hate speech* aiuta a **non far sentire sole le persone o le categorie prese di mira** e a **segnalare che l'odio non è una norma sociale accettata**. Il consenso percepito, infatti, influenza il rischio che si sviluppino nuove e più estreme espressioni d'odio e che dalle parole si passi ai fatti. Cerchiamo di capire insieme quali sono le varie possibilità d'intervento.

MODALITÀ POSSIBILI

Studi e indagini disponibili non hanno ancora individuato quali tra le varie tipologie di intervento siano più efficaci nel prevenire e ridurre l'odio online, tuttavia vi sono diverse strade percorribili. **In risposta all'odio online** si può replicare all'autore in privato o in pubblico; si può segnalare l'episodio al gestore del social network o del sito, o all'amministratore del gruppo; in casi estremi si può tentare una denuncia civile o penale.

Se si avverte un pericolo immediato, per sé o per altri, conviene **contattare le forze dell'ordine**. Inoltre si può **segnalare alle vittime** che l'Ufficio nazionale anti razzismo (Unar) può assisterle, con un **contact center gratuito e multilingue attivo 24 ore su 24**.

Parte dei suggerimenti di questo capitolo, sono frutto dell'esperienza della Task Force Hate Speech di Amnesty International Italia (vedi pagina 35).

4

RISPONDERE

NARRAZIONE POSITIVA

Una possibile forma di contrasto all'odio online consiste nel rispondere alla narrazione negativa dominante proponendo una **visione alternativa**: per sradicare l' *hate speech*, non basta sanzionarlo o dimostrarlo falso, bisogna anche offrire una narrazione opposta a cui aderire.

La narrazione è un racconto che offre una visione coerente di eventi e personaggi, per esempio: "gli immigrati hanno portato la criminalità, l'Europa non ci difende, ma un nostro eroe li fermerà". **L' *hate speech* racchiude sempre una narrazione d'odio**, a cui si possono opporre due tipi di narrazioni positive:

- la **contro-narrazione** è una risposta diretta a uno specifico



messaggio d'odio: si rivolge a chi lo conosce, per evidenziarne le incoerenze e i sottintesi (es. la logica del capro espiatorio), indebolirlo e minarne la credibilità, con strumenti vari che vanno dal *fact checking* al rovesciamento ironico della prospettiva;

- la **narrazione alternativa** invece mira a cambiamenti di lungo termine, con campagne che promuovono punti di vista alternativi. Non reagisce a un messaggio, ma a una prospettiva generale (es. "i migranti ci invadono"), per smontarla e proporre un'altra, rivolgendosi a un pubblico generale.

N.B.

CONTRONARRAZIONE:
COSA FARE E COSA
NON FARE

Di seguito alcune indicazioni che può essere utile seguire per replicare ai contenuti d'odio, frutto dell'esperienza della Task Force Hate Speech della sezione italiana di Amnesty International (focus a pag. 35).

- **Non legittimare i modi e i linguaggi che avversiamo**: non attaccare con un "contro-odio", né usare gli stessi stereotipi invertiti (il "buon migrante" contrapposto al cattivo).
- La narrazione d'odio è forte perché l'odio è un'emozione; quindi, per controbilanciarla, **bisogna far leva anche sulle emozioni**, potenziando quelle positive e inclusive, non conflittuali.
- Bisogna inoltre tener conto di **emozioni e significati soggettivi** attribuiti dal nostro interlocutore ai fatti e ai dati.
- Le **differenze di gravità** delle espressioni d'odio vanno considerate nel decidere come rispondere.
- Di solito conviene **non puntare a far cambiare idea agli *hater* bensì a chi li segue**, che spesso è mosso da emozioni forti ma estemporanee, più suscettibili di modifica.
- Contrastare l' *hate speech* può avere grossi costi emotivi, perché ci espone all'odio contro altri e possiamo diventarne bersagli noi stessi. Se però gli contrappiamo narrazioni positive, possiamo vedere l' *hate speech* non come una spaccatura sociale irrimediabile, ma come **un fallimento comunicativo temporaneo**, riparabile.



INTERAGIRE CON L'ALTRO

Quando rispondiamo a un post o a un *tweet*, oppure a un commento, dobbiamo ricordare che dietro a quel testo o a quella immagine c'è una persona. Come relazionarci con questa persona?

- **Non essere aggressivo ma assertivo:** affermare le proprie tesi in modo chiaro e diretto ma rispettoso, esigendo a nostra volta rispetto.
- Consideriamo l'interlocutore **non come incarnazione di una categoria** ma come persona ed esigiamo che faccia lo stesso verso di noi e coloro di cui si parla.
- Possiamo **dargli del lei** e chiamarlo per nome: "Buongiorno Mario, può spiegarmi cosa intende?".
- Chiediamogli se **intendeva davvero quanto affermava**, se lo direbbe nel mondo reale, ripetendoglielo con parole diverse.
- **Non usare termini che creino tifoserie** contrapposte quali "razzista", "ignorante" o "buonista".
- Se ci si confronta su un testo, **invitiamo a leggerlo** per intero per essere certi che lo conosca.
- Invitiamo a **restare in tema** (evitando il benaltrismo) e a non generalizzare: il crimine di un singolo non colpevolizza l'intero gruppo.
- **Chiediamo le fonti** di un'affermazione, e chiarimenti se è confusa.
- Mostriamo **comprensione dello stato d'animo** che ha portato a certe affermazioni.
- **Usiamo l'ironia solo se sappiamo padroneggiarla**, per esempio per sdrammatizzare e non per aggredire.



5 ATTIVARSI

La **Task Force Hate Speech** è un gruppo di lavoro di Amnesty International Italia il cui mandato è contrastare i discorsi d'odio online. Il progetto nasce nel 2016, dopo l'osservazione di un nuovo fenomeno sui social network: la già diffusa abitudine degli utenti a pubblicare commenti violenti e intolleranti nei confronti delle minoranze moltiplica la propria forza, sospinta dalle tragedie dei migranti nel Mediterraneo.

Amnesty International Italia inoltre osserva che in rete linguaggi e toni aggressivi veicolano contenuti spesso infondati, amplificandoli a dismisura. Viene quindi ideata per la prima volta una forma di attivismo specializzata ed organizzata nella difesa dei diritti umani nel mondo online. La TFHS tutela i gruppi che sono più a rischio di discriminazione per motivi legati all'etnia, alla religione, al genere, all'identità di genere, all'orientamento sessuale, all'età, alla disabilità, alle condizioni socio-economiche e alla missione sociale. Le categorie di intervento si sono ampliate nel tempo non solo distinguendo tra le forme di razzismo legate alla migrazione e quelle legate alla cittadinanza, ma osservando nuove forme di intolleranza nei confronti delle Ong che si attivano per la tutela dei diritti umani e dell'ambiente.

LE AZIONI: MONITORAGGIO E CONTRASTO

Per contrastare discriminazione e intolleranza online, la TFHS vuole offrire una risposta concreta, collettiva e incisiva che vede protagonista la società civile. La Task Force **interviene sulle pagine Facebook dei quotidiani e degli organi di informazione in generale**, principali fonti di notizie per la maggior parte delle persone. Attraverso la loro azione, i membri della Task Force partecipano al processo di cambiamento che l'organizzazione intende produrre nella società e nella vita delle persone. Molti quotidiani, nell'ottica di aumentare le interazioni, danno a titoli e ad articoli un taglio che mira a generare commenti di odio, proponendo narrazioni stereotipate rispetto alle minoranze i cui diritti sono meno tutelati.

La TFHS in questi casi interviene perché **dove non c'è contraddittorio le posizioni tendono ad estremizzarsi**. Le attivazioni hanno un impatto diverso a seconda che si tratti di grandi testate oppure di quotidiani locali. Sui grandi quotidiani nazionali si incontrano persone con posizioni molto diverse ed il dibattito è spesso più equilibrato: esiste la possibilità che il commento sia letto da molte persone, se si riesce ad assicurargli la visibilità necessaria. Sui quotidiani locali invece i commenti sono pochi, per cui gli interventi della TFHS possono essere più visibili, ma il pubblico è più ristretto.



COME SI ATTIVA LA TFHS

Il monitoraggio delle pagine viene fatto individualmente, scegliendo di volta in volta quali testate controllare e su quali post concentrarsi. Quando rilevano un post che tocca specifiche tematiche e sotto il quale sono già presenti o potrebbero comparire commenti d'odio, i membri della TFHS lo comunicano al gruppo e passano all'azione:

1. segnalando i commenti che violano le policy di Facebook, primariamente quelle relative all'incitamento all'odio;
2. rispondendo ai commenti;
3. scrivendo commenti al post, per precisare alcuni punti e fornire informazioni che il titolo non dà (e a volte neanche l'articolo), introducendo una narrazione alternativa;
4. supportando i membri della TFHS e le altre persone che esprimono posizioni equilibrate, mettendo *like*, integrando e ribadendo.

TONI D'ODIO E DISINFORMAZIONE

I membri della TFHS usano **sempre toni civili, evitando di categorizzare le persone** con cui interloquiscono: un obiettivo fondamentale è quello di abbassare i toni del confronto. Nel corso dell'attività evitano di schierarsi a livello partitico, tenendo come unico punto di riferimento il rispetto dei diritti umani, in base alle posizioni di Amnesty International.

Spesso **la disinformazione rappresenta il primo anello della catena che genera discorsi di odio**. Un buon inizio, nelle attivazioni, prevede quindi di presentare i fatti corretti spostando l'attenzione dalle persone e dai giudizi che vengono espressi su di esse. Nei commenti la TFHS fa riferimento a dati completi, citando fonti attendibili, contrastando le generalizzazioni che spesso colpiscono i gruppi discriminati.

Le generalizzazioni possono infatti produrre una forte polarizzazione del discorso, che sui social tende a semplificare questioni complesse, creando schieramenti contrapposti. Tali questioni richiederebbero invece analisi approfondite e soprattutto consapevolezza delle conseguenze che i discorsi d'odio hanno sulle vite delle persone. In particolare, la TFHS interviene cercando di innescare una prospettiva legata più all'umano e alle esperienze delle singole persone, che ai gruppi di appartenenza. Nel farlo si impegna in un lavoro di messa in luce e invito alla riflessione su stereotipi e pregiudizi, mostrandone l'inconsistenza.



PERCHÉ INTERVENIRE

Attraverso questi interventi, che ribaltano completamente il discorso d'odio, **si tutelano persone che rischiano di rimanere bersagli di chi ne svilisce l'esperienza e l'esistenza** stessa, mettendone in dubbio sofferenze, bisogni, diritti, rivendicazioni e l'oppressione che subiscono.

Intervenire con una contro narrativa efficace, basata su tecniche di comunicazione non-violenta che garantiscono assertività e rispetto di interlocutori e interlocutrici, permette di ampliare punti di vista, prospettive, piani di discorso e di avere effetti su quella "maggioranza silenziosa" che legge e non commenta. Queste persone rischiano infatti di trovarsi completamente scoraggiate dalla lettura di commenti di odio privi di contraddittorio e reiterati, formandosi una percezione del senso comune negativa. Agire su questo tipo di rischio è molto importante per Amnesty International, al fine di incoraggiare le persone ad attivarsi, guidate dall'esempio di chi già lo fa, e di garantire visibilità all'azione in difesa dei diritti umani.

IL VALORE DEL GRUPPO

Contrastare i discorsi d'odio può essere impegnativo dal punto di vista psicologico. Viene però in aiuto **il ruolo del gruppo che sostiene la contro-narrazione** rispetto ai commenti d'odio con i *like* e con i commenti di rinforzo. Ogni attivista sente dunque il supporto del gruppo e meno vulnerabilità. Oltre a consentire di confrontarsi e di condividere le modalità, il gruppo gioca un ruolo fondamentale quando entrano nuove forze in campo.

Infatti, sebbene la formazione fornisca le competenze necessarie, alla prima attivazione **la reazione può essere di panico e di insicurezza**. Da qui la figura *buddy*: una persona che aiuta nell'inserimento attraverso call dedicate e confronto continuo. Attivandosi su Facebook, l'attivista non deve necessariamente usare il profilo personale ma può anche avere un profilo dedicato, ben costruito in termini di credibilità. Capita a volte, durante le attivazioni di gruppo, che qualcuno degli autori dei commenti d'odio riconosca l'azione della Task Force e provi a sminuirlo. Anche in questo caso il gruppo fa la differenza nel gestire al meglio la situazione, senza togliere credibilità all'azione.



IL MANIFESTO DELLA TASK FORCE HATE SPEECH

Per hate speech (in italiano “discorsi d’odio”) s’intende “qualsiasi forma di espressione mirante a stimolare, promuovere, diffondere o giustificare la violenza, l’odio o la discriminazione nei confronti di una persona o un gruppo di persone, o a denigrare una persona o un gruppo di persone per motivi legati alle loro caratteristiche o situazioni personali, reali o presunte, quali il colore della pelle, la lingua, la religione, la nazionalità o l’origine nazionale o etnica, l’età, la disabilità, il sesso, l’identità di genere e l’orientamento sessuale”¹⁹, le condizioni socio-economiche, la missione sociale o qualsiasi altra caratteristica o stato personale.

La Task Force Hate Speech (TFHS) della sezione italiana di Amnesty International è formata da un gruppo di persone - diverse per età, genere, formazione, provenienza - che, a titolo volontario e dopo un’adeguata preparazione, agiscono ogni giorno sui social networks, principalmente Facebook, attraverso attività di osservazione della rete e di attivazione laddove si ravvisano potenziali discorsi d’odio, per estendere la difesa e la promozione dei diritti umani anche al mondo online. Nella consapevolezza della pericolosità e pervasività del fenomeno dell’hate speech, la TFHS monitora notizie e fatti di cronaca pubblicati da testate giornalistiche online o altri portali di informazione sui social networks, segnalando alle piattaforme eventuali violazioni e abusi, e intervenendo attivamente nelle conversazioni tra utenti, in presenza di commenti che esprimono odio e intolleranza.

Per difendere i diritti umani Amnesty International invita tutt’altro che all’uso di uno stile di comunicazione assertivo, al fine di affermare, in modo chiaro e diretto, i propri diritti ed esprimere le proprie opinioni in modo

¹⁹ Raccomandazione CM/Rec(2022)161 del comitato dei Ministri del Consiglio d’Europa agli stati membri sulla lotta contro i discorsi d’odio.

adeguato al contesto comunicativo, senza prevaricare l’altro, nel rispetto delle altre persone e nella convinzione che tutt’altro che abbiano lo stesso valore e debbano quindi godere dei medesimi diritti. Così facendo auspica che questo stile possa diffondersi nel mondo social e diventare prassi comune condivisa nella società civile.

I NOSTRI PRINCIPI

- **No Space for Hate**
Nessuno spazio per l’odio. La TFHS si impegna a costruire spazi di confronto online non ostili e a sensibilizzare coloro che non partecipano abitualmente al dibattito
- **L’odio ci riguarda**
Siamo consapevoli dell’impatto reale dell’odio in rete e di come il mondo online sia connesso a quello offline. Crediamo nell’importanza del contrasto alla normalizzazione e banalizzazione di parole violente e di odio, il più delle volte alimentate da rappresentazioni false e fuorvianti della realtà, nella consapevolezza che il modo in cui ci esprimiamo contribuisce a costruire la realtà che ci circonda e plasmare la nostra percezione dei fenomeni.
- **Tutela delle libertà di espressione**
La TFHS contrasta i discorsi di odio e le discriminazioni nel rispetto delle libertà di opinione e di espressione previste dal nostro ordinamento, promuovendo narrazioni alternative (articolo 21, Costituzione).
- **Consapevolezza della nostra missione**
Attraverso una riflessione costante sulle pratiche applicate, le persone attiviste della TFHS si pongono l’obiettivo di attivare spazi di dialogo costruttivo e sensibilizzazione, senza mai avere la pretesa di parlare “al posto di”.
- **Utilizzo di linguaggio ampio**
La TFHS si impegna a utilizzare un linguaggio ampio, non sessista, rappresentativo di tutte le soggettività, nonché uno stile di comunicazione accessibile e assertivo, perché crediamo fortemente che le parole abbiano un peso.



- **Intersezionalità**
Uno dei nostri pilastri è il principio di intersezionalità, che consiste nella considerazione della molteplicità degli aspetti che compongono l’identità.
Siamo infatti consapevoli della complessità dei fenomeni e dei diversi livelli di oppressione sistemica a cui sono sottoposte le categorie solitamente oggetto dei discorsi di odio.
- **Ricorso a fonti autorevoli e ufficiali**
La TFHS si impegna ad essere responsabile e attendibile, intervenendo con commenti imparziali, basati su fonti autorevoli e ufficiali, contrastando le fake news e contribuendo ad una proficua e accessibile narrazione alternativa.

LE STRATEGIE

Al fine di prenderci cura anche degli spazi online, abbiamo elaborato delle strategie di attivazione:

- Invitiamo a considerare ciascuna persona ha una storia – unica – alle spalle. Tuttavia evitiamo di suggerire compassione. Questo atteggiamento rischia di ridurre i soggetti a ‘casi’. Rigorosa e condivisa dalla TFHS è la necessità di diffondere l’idea del rispetto dei diritti umani e della vita come fondamento della società civile (Dichiarazione Universale dei Diritti Umani).
- Dialoghiamo tra persone, non tra categorie o ‘fazioni’. Quindi evitiamo di generalizzare, e non utilizziamo mai termini che possano creare contrapposizioni tra tifoserie, quali per esempio “fascista”, “razzista”, “ignorante”, “populista”.
- Cerchiamo di instaurare un dialogo con l’utente e se possibile un confronto, comunicando in modo pacato e non giudicante, e chiedendo di riformulare e argomentare soprattutto laddove ci sembra manchi chiarezza o sia basato il messaggio su notizie confuse, non verificate, o palesemente false.

- Richiamiamo le testate giornalistiche a produrre un’informazione corretta, verificata, e adeguata al loro ruolo e all’utenza, in linea con il codice deontologico dei giornalisti, con il Manifesto di Venezia, con la Carta di Roma; inoltre, segnaliamo loro se un titolo è fuorviante rispetto al contenuto di un articolo o è in contrasto con esso.
- Spesso i siti delle testate sono a pagamento. Perciò, i commenti si limitano al titolo dell’articolo pubblicato. In questi casi richiamiamo l’attenzione dell’utenza e di chi commenta sulla notizia, invitando alla ricerca di ulteriori informazioni e fornendo indicazioni su dove sia possibile trovare notizie e dati attendibili e verificati per evitare il propagarsi di fake news capaci di alimentare i discorsi d’odio.
- Interveniamo su casi urgenti soprattutto sui fatti di cronaca che suscitano grande clamore. Quando un soggetto è particolarmente sotto tiro, attivarsi immediatamente con commenti o segnalazioni anche alle stesse testate può essere importante per far sentire meno sola la persona presa di mira e per tutelare la sua dignità.
- Facciamo presente che alcuni commenti possono sfociare nel penale e che quindi possono configurarsi come reati perseguibili.
- Inoltriamo segnalazioni e richieste di rimozione di contenuti d’odio alle piattaforme: fermo restando il diritto di tutt’altro che alla libertà di espressione, in presenza di commenti profondamente offensivi, reiterati, e dei quali sia chiaro ed esplicito l’intento di incitare all’odio verso una persona o un gruppo di persone, è infatti necessario segnalare il commento alla piattaforma per ottenerne la cancellazione.



6

SEGNALARE

Quando vi imbattete in un contenuto che ritenete essere *hate speech* su Facebook, Instagram, X (ex Twitter), YouTube, TikTok ecc. potete segnalarlo al social network chiedendone la rimozione. È importante leggere gli standard della comunità, per capire quali sono quei contenuti che violano le regole autoimposte dalle piattaforme. Questo consente di evitare di segnalare quei contenuti che non hanno alcuna possibilità di essere rimossi, ci aiuta a riflettere su quali sono gli attuali limiti posti dalle piattaforme ed evita la frustrazione che potrebbe manifestarsi – a lungo andare – nell’eseguire segnalazioni che non vanno a buon fine.

MOLTI SOCIAL NETWORK.
TANTE DIFFERENZE

Nel 2016 alcune aziende IT (il gruppo iniziale, poi ampliato, includeva Facebook, Twitter, YouTube, Microsoft) hanno convenuto con la Commissione europea un codice di condotta col quale, riconoscendo il proprio ruolo cruciale nel contrasto all’*hate speech* online, si impegnano pubblicamente ad adottare alcune misure atte a realizzare tale obiettivo. Tra queste: la **predisposizione “di procedure chiare ed efficaci per esaminare le segnalazioni** riguardanti forme illegali di incitamento all’odio”; l’esame di tali segnalazioni “alla luce delle regole e degli orientamenti da esse predisposti per la comunità degli utenti e, ove necessario, delle leggi nazionali di recepimento della decisione quadro 2008/913/GAI”; l’esame in meno di 24 ore “della maggior parte delle segnalazioni valide” miranti alla rimozione di forme illegali di incitamento all’odio e, se necessario, rimuovono tali contenuti o ne disabilitano l’accesso”²⁰.

Tra il 2016 e il 2022 la Commissione europea, attraverso una rete di organizzazioni della società civile che si occupano di contrastare l’*hate speech* della quale fa parte anche Amnesty International Italia, ha condotto un esercizio di monitoraggio per valutare lo stato di applicazione del codice di condotta, segnalando contenuti ritenuti illegali e monitorando l’esito delle segnalazioni. Quello che abbiamo potuto osservare, partecipando direttamente a questa attività, è una grande varietà in termini di tempi e modalità con cui le varie piattaforme intervengono. Questo ci consente di affermare, che sia importante conoscere bene le regole e gli strumenti messi a disposizione dalle singole piattaforme quando si decide di attivarsi su una di esse.

UN REGOLAMENTO PER UNA RETE PIÙ SICURA

A febbraio 2024 entra in vigore il Digital Services Act, regolamento europeo volto a rendere il web uno spazio sicuro in cui ogni persona veda tutelati i propri diritti, che richiama in modo esplicito quanto precedentemente stabilito dal codice di condotta in materia di incitamento all’odio²¹. In Italia l’ente incaricato di coordinare i servizi digitali nell’ambito del DSA è l’Autorità garante delle Comunicazioni.

²⁰ https://ec.europa.eu/commission/presscorner/detail/it/IP_16_1937

²¹ <https://eur-lex.europa.eu/legal-content/IT/TXT/PDF/?uri=CELEX:32022R2065>

Considerando l’aggiornamento frequente delle policy da parte dei social network, così come delle modalità di segnalazione, appare poco funzionale fornire qui una descrizione. Ci limitiamo, per tanto, a fornire i link ai quali è possibile trovare le indicazioni e le informazioni necessarie, per alcuni dei principali social network.



FACEBOOK

Policy

<https://transparency.meta.com/it-it/policies/community-standards/?source=https%3A%2F%2Fit-it.facebook.com%2Fcommunitystandards%2F>

Come segnalare

<https://it-it.facebook.com/help/263149623790594>



INSTAGRAM

Policy

<https://help.instagram.com/477434105621119>

Come segnalare

https://help.instagram.com/165828726894770?helpref=page_content



X (EX TWITTER)

Policy

<https://help.twitter.com/it/rules-and-policies/hateful-conduct-policy>

Come segnalare

<https://help.twitter.com/it/using-x/communities#report-tweet>



YOUTUBE

Policy

https://support.google.com/youtube/answer/2801939?hl=it&ref_topic=9282436

Come segnalare

<https://support.google.com/youtube/answer/2802027?hl=it&co=GENIE.Platform%3DAndroid>



TIKTOK

Policy

https://www.tiktok.com/community-guidelines/it-it/safety-civility/?enter_method=left_navigation

Come segnalare

<https://support.tiktok.com/it/safety-hc/report-a-problem>



MEMO

SEGNALARE A OSCAD E UNAR

Sono diverse le risorse istituzionali a disposizione di chi subisce o assiste a fenomeni d'odio.

- Una possibilità, per chi è vittima di reati a sfondo discriminatorio, i cosiddetti *hate crime*, è quella di rivolgersi all'**Osservatorio per la sicurezza contro gli atti discriminatori**

del ministero dell'Interno. Chi "subisce un **evento penalmente rilevante** in relazione alla razza/etnia, credo religioso, orientamento sessuale/identità di genere e disabilità" può **scrivere a oscad@dcpc.interno.it**

- L'**Ufficio nazionale antidiscriminazioni razziali** offre un numero verde gratuito per le vittime o i testimoni di discriminazione al quale possono **riportare segnalazioni** su "comportamenti e

fatti discriminatori" o rivolgersi per **chiedere informazioni, orientamento e supporto** al fine di prevenire o contrastare tali episodi o per comprendere quale percorso possa essere il più adatto al caso vissuto. Il servizio è **gratuito e multilingue** ed è **attivo dal lunedì al venerdì al numero 800 90.10.10.**

Il contact center può anche essere contattato compilando un **modulo online.**





Amnesty International Italia

Via Goito 39 - 00185 Roma
Tel: (+39) 06 44.90210
Fax: (+39) 06 44.90.243
info@amnesty.it
www.amnesty.it