
Hate speech

Dossier ECPMF



Luglio 2019

Indice generale

Cos'è l'hate speech?.....	3
L'hate speech è in aumento?.....	4
L'hate speech va considerato un crimine?.....	6
Come contrastare l'hate speech su internet?.....	9
Come prevenire l'hate speech?.....	14
Qual è il ruolo dei giornalisti?.....	15

Cos'è l'hate speech?

Sebbene si tratti di un'espressione diffusa, non esiste una definizione univoca di *hate speech* (o discorsi d'odio). Le difficoltà in cui si imbatte la ricerca di una definizione condivisa sono legate al dibattito – [giuridico](#), ma ancor prima politico-filosofico e dunque culturale – su quali siano i confini della libertà di espressione. Come possiamo definire (e quindi contrastare) i discorsi d'odio senza correre il rischio di limitare una libertà fondamentale? Questo [dilemma](#) precede internet, ma la rivoluzione tecnologica lo ripropone con rinnovata forza.

A fornire una base comune alle [diverse definizioni di hate speech](#) in circolazione sono stati i documenti prodotti dalle istituzioni internazionali del secondo dopoguerra. Stando a quanto raccomandato dal [Consiglio d'Europa nel 1997](#), ricadono nei discorsi d'odio quelle “espressioni che diffondono, incitano, promuovono o giustificano l'odio razziale, la xenofobia, l'antisemitismo o altre forme di minaccia basate sull'intolleranza – inclusa l'intolleranza espressa dal nazionalismo aggressivo e dall'etnocentrismo –, sulla discriminazione e sull'ostilità verso i minori, i migranti e le persone di origine straniera”.

Se è vero che l'espressione “hate speech” si è affermata solo negli anni Novanta, l'osservazione del fenomeno e l'impegno a contrastarlo non sono nuovi (in precedenza si preferiva utilizzare l'espressione “[incitamento all'odio](#)”). Per molti decenni l'attenzione si è concentrata soprattutto sull'odio su base razziale, sull'antisemitismo e sul negazionismo. All'alba del nuovo millennio la sensibilità sul tema è cresciuta fino a comprendere le minoranze religiose (anzitutto musulmane, sempre più oggetto di minacce e discriminazioni), mentre altre categorie come le donne, le persone LGBT, i disabili e gli anziani sono considerate potenziali bersagli dei discorsi d'odio relativamente da poco tempo.

In sintesi, indipendentemente dalle forme assunte (scritte o orali, verbali o non verbali, esplicite o implicite) e dalla portata giuridica (eventuali “reati d'odio”), può ricadere all'interno della definizione di hate speech **qualsiasi espressione violenta o discriminatoria nei confronti di altre persone o**

gruppi di persone. Proprio perché l'hate speech colpisce le persone per le loro caratteristiche e/o condizioni personali, le azioni di contrasto al fenomeno hanno bisogno di adattarsi al contesto e ai fenomeni sociali, economici, politici e tecnologici in corso.

Il problema del contrasto ai discorsi d'odio incrocia oggi i dilemmi e le contraddizioni della nascente era digitale. In un [rapporto di recente pubblicazione](#), il Consiglio d'Europa ha inserito i discorsi d'odio all'interno del più vasto problema dell'*information disorder*, un inquinamento dei contenuti su scala globale che vede intrecciarsi le "patologie" dell' hate speech e delle cosiddette *fake news*: la disinformazione nascerebbe dall'incontro tra *mis-information* (diffusione di notizie false ma innocue) e *mal-information* (notizie vere ma diffuse con l'intenzione di colpire).

L'hate speech è in aumento?

Anche se il dibattito sull'incitamento all'odio ha una lunga storia, l'attenzione verso l'hate speech è senza dubbio aumentata negli ultimi anni, in seguito all'acuirsi delle sue manifestazioni nello spazio digitale. Come si osserva in questo [grafico](#) (che indica la frequenza dell'espressione all'interno del corpus di libri in inglese indicizzati su Google Books), il termine è diventato ricorrente solo negli ultimi 30 anni – mentre gran parte delle iniziative adottate dalle autorità e dalla società civile sono ancora più recenti.



Poiché continua a mancare una definizione condivisa sufficientemente precisa, a livello europeo

la rilevazione dei dati sulla ricorrenza dei discorsi d'odio rimane però assai scarsa o, nella migliore delle ipotesi, disomogenea. Solo negli ultimi anni enti come l'[OSCE](#) o l'[Agenzia dell'Unione europea per i diritti fondamentali](#) hanno iniziato a premere per sistematizzare perlomeno i dati nazionali sui crimini d'odio (che si sovrappongono solo in parte ai discorsi d'odio): dal 2016 l'OSCE raccoglie statistiche e compara le modalità di rilevazione dei crimini d'odio da parte delle autorità degli stati membri.

Per quanto riguarda l'*hate speech* in quanto tale, organizzazioni della società civile e università stanno svolgendo un ruolo importante nel [raccogliere dati](#), soprattutto per quanto riguarda i social media. Nell'ambito del progetto europeo [Positive Messengers](#), ad esempio, enti di sette paesi diversi monitorano i discorsi d'odio nei confronti di migranti e rifugiati. Stando a un loro rapporto di fine 2017, una frequenza maggiore di hate speech è correlata a: maggiore accesso a connessioni a banda larga; crescita dei flussi migratori; campagne elettorali; drammi nazionali come gli attacchi terroristici; peggioramento della situazione economica. Analogamente, uno [studio](#) dell'Università di Warwick sulla Germania ha evidenziato una corrispondenza tra l'incidenza dei messaggi razzisti online e la commissione di crimini contro i rifugiati.

In effetti, l'avvento del web 2.0, dei social network e la diffusione della "cultura dei commenti" hanno portato a una crescita esponenziale dei contenuti digitali generati dagli utenti, così come delle opportunità per le persone di interagire sulla rete. Di conseguenza, i discorsi d'odio su internet hanno potuto propagarsi in modo molto più massiccio che in precedenza.

Come spiega [l'UNESCO](#), si tratta di un caso esemplare delle sfide che pongono tecnologie dall'impatto potenzialmente molto positivo. La nascita di internet era infatti stata accompagnata da una retorica utopistica, che vedeva nella comunicazione virtuale un potenziale enorme di emancipazione e democratizzazione. Ma, come si legge in un [rapporto](#) del Consiglio d'Europa, negli Stati Uniti i gruppi suprematisti sono stati tra i primi a sfruttare la rete per diffondere messaggi razzisti e xenofobi, ad esempio attraverso la creazione di "hate sites". Anche in Europa molti gruppi estremisti sono stati molto rapidi nel sviluppare una sofisticata presenza online.

A partire dagli anni Duemila le autorità hanno dunque dedicato un'attenzione crescente all'hate speech sul web. Nel 2001 il Consiglio d'Europa ha adottato la [Convenzione sul crimine informatico](#), il

primo trattato multilaterale volto a combattere i reati compiuti su internet. Alla Convenzione è seguito, nel 2003, il [Protocollo aggiuntivo](#), un documento che promuove una serie di misure per rendere più efficace il contrasto ai crimini d'odio su internet.

Anche se al momento non esistono dati accurati sulla diffusione dei discorsi d'odio online, sembra esserci consenso sul fatto che l'odio online sia in crescita, sia per diffusione, sia per la varietà delle strategie usate. Ad esempio, il rapporto 2015 della [Commissione europea contro il razzismo e l'intolleranza \(ECRI\)](#), un organismo del Consiglio d'Europa, citava l'aumento dell'hate speech online tra le principali tendenze dell'anno. Un rapporto pubblicato sempre nel 2015 dalla [Relatrice speciale del Consiglio dei diritti umani dell'ONU per le questioni delle minoranze](#) evidenziava una "diffusione senza precedenti" dell'hate speech online, mentre sulla rivista "[EU Internet Law](#)" lo studioso Ioannis Iglezakis definiva internet la "nuova frontiera" dell'hate speech.

L'hate speech va considerato un crimine?

Per quanto riguarda gli stati europei, le misure che le istituzioni prendono (o invitano a prendere) contro i discorsi d'odio sono fortemente influenzate dalle indicazioni e raccomandazioni elaborate a livello internazionale. Ad esempio, le Nazioni Unite hanno predisposto una serie di documenti e raccomandazioni sui discorsi d'odio che gli stati sono invitati a osservare. Già nel 1966 il [Patto internazionale sui diritti civili e politici](#) prevedeva che "qualsiasi appello all'odio nazionale, razziale o religioso che costituisce incitamento alla discriminazione, all'ostilità o alla violenza deve essere vietato dalla legge".

A livello europeo, i discorsi d'odio sono stati un tema di cui il Consiglio d'Europa s'è occupato per decenni, sia dal punto di vista delle azioni e raccomandazioni legislative, sia da quello della giurisprudenza, attraverso la [Corte europea dei diritti dell'uomo](#). L'interesse dell'Unione europea per il tema è invece molto più recente ma parecchio vivace, e [accenna ad aumentare](#). Considerati i poteri di cui dispone l'UE – soprattutto nell'ambito del mercato digitale e in termini di direzione delle politiche degli stati membri – le sue iniziative possono rivelarsi molto incisive. A livello nazionale, pressoché tutti gli stati europei hanno introdotto delle normative per contrastare almeno alcune tipologie di discorsi d'odio.

In linea con le raccomandazioni delle Nazioni Unite e del Consiglio d'Europa, per lungo tempo le autorità nazionali ed europee hanno teso a considerare i discorsi d'odio un vero e proprio reato. In questo modo diventava possibile reprimerli, anche se solamente a precise condizioni, in modo proporzionato e a seguito di una valutazione della magistratura. Il ricorso agli strumenti di ordine penale ha però mostrato sempre più limiti nel corso degli anni Duemila, innanzitutto dal punto di vista quantitativo. Come s'è visto, il concetto di hate speech si è espanso notevolmente, e la stessa mole delle espressioni d'odio in circolazione rende ormai arduo per la magistratura affrontarle una per una.

D'altra parte, l'osservazione del fenomeno e della giurisprudenza ha mostrato i limiti e i pericoli insiti nell'affidarsi alla sola criminalizzazione dei discorsi d'odio: il rischio che i processi finiscano per dare ancora maggiore visibilità a idee e autori ispiratori d'odio; il pericolo che le autorità etichettino come "discorsi d'odio" semplici opinioni in contrasto con le loro, riducendo così lo spazio disponibile per il dissenso; la possibilità che definizioni ampie di hate speech lascino troppo spazio alla discrezionalità dei giudici e finiscano per comprimere la libertà di espressione.

Il problema di valutare il rapporto tra il rispetto della libertà d'espressione e il contrasto ai discorsi d'odio è stato al centro di numerose sentenze della Corte europea dei diritti dell'uomo. La Corte si appoggia principalmente sulla [Convenzione europea dei diritti dell'uomo](#), che all'articolo 10 protegge la libertà d'espressione, ma che all'articolo 17 proibisce di abusare delle libertà riconosciute dalla Convenzione per minarne le fondamenta: in sostanza, chi promuove valori in aperto contrasto con quelli della Convenzione non può poi appellarsi alla libertà di espressione che essa garantisce.

Facendo riferimento all'articolo 17, [nel corso dei decenni](#) la Corte europea dei diritti dell'uomo ha varie volte respinto i ricorsi di individui e attori che erano stati perseguiti per le loro affermazioni. Dalla giurisprudenza della Corte, appare chiaro che il diritto alla libertà di espressione non copre ad esempio affermazioni e opinioni negazioniste, che inneggiano al nazionalsocialismo o che invocano la discriminazione razziale.

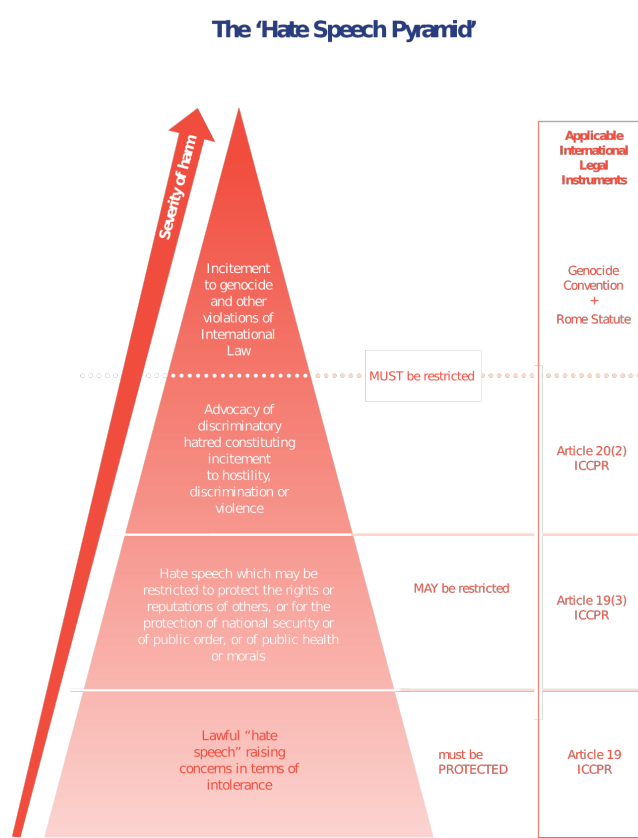
Tuttavia, mentre alcuni casi ricadono chiaramente nell'articolo 17, in altri casi è necessaria una valutazione più sottile dei vari diritti in gioco e delle misure prese dalle autorità nazionali per contrastare determinate affermazioni. In questi casi la Corte valuta con attenzione il profilo del presunto reo, le motivazioni che lo avevano guidato e gli obiettivi che intendeva perseguire, il tipo di affermazioni che aveva fatto e il contesto globale in cui erano inserite, e il grado di circolazione che

avevano raggiunto.

Anche laddove la Corte riconosca come legittimi dei provvedimenti che puniscono e rimuovono determinate espressioni che incitano all'odio, essa richiede comunque che le restrizioni abbiano un preciso fondamento giuridico e che non vi siano alternative a disposizione. Inoltre, le restrizioni devono perseguire degli obiettivi legittimi, come proteggere l'interesse generale o i diritti di altre persone. Dal punto di vista giuridico, la Corte di Strasburgo ha anche riconosciuto l'obbligo per gli stati di proteggere attivamente le vittime dei discorsi d'odio.

A fronte dei numerosi problemi giuridici e quantitativi posti dal considerare i discorsi d'odio un reato, le raccomandazioni [delle Nazioni Unite](#) e [del Consiglio d'Europa](#) invitano gli stati a criminalizzare “solamente le forme gravi ed estreme dei discorsi d'odio”, cioè quelle che incitano concretamente a violenze e discriminazioni contro una persona o un gruppo di persone. In tutti gli altri casi, gli stati sono invece invitati a esplorare strumenti alternativi per prevenire o contrastare i discorsi d'odio.

A fronte dei numerosi problemi giuridici e quantitativi posti dal considerare i discorsi d'odio un reato, le raccomandazioni più recenti [delle Nazioni Unite](#) e [del Consiglio d'Europa](#) invitano gli stati a criminalizzare “solamente le forme gravi ed estreme dei discorsi d'odio”, cioè quelle che incitano concretamente a violenze e discriminazioni contro una persona o un gruppo di persone. In tutti gli altri casi, gli stati sono invece invitati a esplorare strumenti alternativi per prevenire o contrastare i discorsi d'odio.



Fonte: *Article 19, 'Hate speech' explained*

Il caso rumeno

In Romania l'hate speech non è affrontato in maniera molto rigida e sistematica. Diversi studi tra cui il [report dell'ECRI](#), pubblicato nel giugno 2019, evidenziano che i principali obiettivi dei discorsi d'odio nel Paese sono i rom, gli ungheresi, gli ebrei e le persone LGBT, ma negli ultimi anni sono diventati più visibili anche i discorsi d'odio contro i migranti. Nonostante [alcuni passi positivi](#) compiuti negli ultimi cinque anni, i discorsi d'odio risultano estremamente diffusi, mentre la risposta del sistema della giustizia penale ai crimini d'odio è particolarmente carente.

Politici, mass media e utenti dei media online risultano ugualmente attivi nel generare hate speech. Ma le statistiche sui crimini e i discorsi d'odio sono scarse, anche perché manca un organismo incaricato di monitorarli. Il [Consiglio nazionale per combattere la discriminazione \(CNCD\)](#) è l'organismo per la promozione della parità più importante della Romania; si occupa soprattutto di rispondere a ricorsi. Il Consiglio nazionale audiovisivo (CNA) monitora invece i media audio-visivi, avviando autonomamente pratiche e sanzionando all'occorrenza comportamenti discriminatori.

Dal 2002 il CNCD ha analizzato oltre 6.300 casi di discriminazione, che sono risultati in sanzioni amministrative o raccomandazioni pubbliche. Nel 2006, il Consiglio ha condannato il gruppo Destra Nuova per aver pubblicato sul proprio sito articoli xenofobi e razzisti, ordinandone la rimozione. L'hate speech è usato spesso durante le [campagne elettorali](#): il [CNCD ha sanzionato](#) anche politici di alto rango, incluso l'ex presidente rumeno [Traian Basescu](#), per dichiarazioni discriminatorie contro le minoranze.

Per quanto riguarda [la cornice legale e gli strumenti regolatori riguardanti l'hate speech](#), il codice penale rumeno del 2014 contiene alcune misure contro i discorsi d'odio. Il Paese ha trasposto le disposizioni della [direttiva europea sui servizi media audiovisivi](#), che proibisce contenuti che incitano all'odio. I media sono considerati responsabili dei contenuti delle pubblicazioni e le trasmissioni, oltre che della [moderazione dei commenti](#) riferiti agli articoli pubblicati online. Allo stato presente il quadro rumeno offre però ben poche regolamentazioni che possano essere applicate ai nuovi media, alle aziende informatiche e ai fornitori di servizi internet.

Come contrastare l'hate speech su internet?

Quando parliamo di hate speech su internet gli attori principali sono le testate online e i blog, le piattaforme tecnologiche e gli altri intermediari di internet, come i motori di ricerca, i fornitori di servizi, i social network, e così via. In particolare, negli ultimi anni è cresciuta l'attenzione sul ruolo delle grandi aziende digitali nel diffondere l'hate speech, ed è aumentata la pressione affinché questi soggetti si impegnassero di più nel contrastarlo sulle proprie piattaforme, smettendo di considerarsi come dei meri intermediari. Anche la Commissione europea [ha chiesto](#) "maggiore responsabilità" e ha richiamato le piattaforme tecnologiche ad agire in modo più deciso e tempestivo nella prevenzione, individuazione e rimozione dei contenuti illegali pubblicati dai loro utenti, minacciando di ricorrere a [interventi legislativi](#).

Una questione importante è proprio quella della responsabilità giuridica delle piattaforme per i

contenuti prodotti autonomamente dai loro utenti. A fare giurisprudenza in questo campo è stata la sentenza della Corte europea per i diritti dell'uomo nel caso "[Delfi vs. Estonia](#)" (2015), che riguardava la responsabilità per i commenti postati dagli utenti sul portale di notizie estone "Delfi AS". La Corte ha stabilito che il portale aveva il "controllo editoriale" sulla sezione commenti, quindi avrebbe dovuto impedire la pubblicazione dei commenti illegali, procedendo a una limitazione "proporzionata" e "giustificata" della libertà di espressione. Ponendo la responsabilità a carico degli intermediari, la sentenza, che si discosta da quanto si evince dalla [Direttiva europea sul commercio elettronico](#) (che però non tratta direttamente il tema), è stata oggetto di un acceso dibattito.

Contrastare l'hate speech online è per certi versi più complicato che combatterlo altrove. Come indica [l'UNESCO](#), i discorsi d'odio su internet sono infatti contraddistinti da alcune caratteristiche che li rendono più sfuggenti: l'anonimato dei loro autori; la permanenza, ovvero la capacità dei messaggi d'odio di restare online nel tempo; *l'itineracy*, cioè la capacità di propagarsi in piattaforme ed ambienti diversi da quelli dove i messaggi sono stati creati; il carattere intergiurisdizionale e transnazionale dei contenuti e delle piattaforme che li ospitano.

Mentre è ormai [pacifico](#) che le aziende private operanti sul web sono soggette alle stesse norme e agli stessi standard sui diritti umani che si applicano nello spazio offline, l'individuazione del regime giuridico applicabile nei vari casi non è sempre semplice. Se infatti i gestori di servizi internet sono soggetti alle leggi nazionali del paese in cui operano e i motori di ricerca sono vincolati ad un regime multiplo – quello del paese in cui sono registrati e quelli dei diversi paesi in cui offrono i propri servizi – per i social network con una diffusione globale la questione si fa più complessa. In assenza di un'autorità giurisdizionale sovranazionale, e in virtù delle limitate capacità degli singole giurisdizioni nazionali, di fatto i social network tendono ad operare principalmente in accordo con i propri termini di servizio.

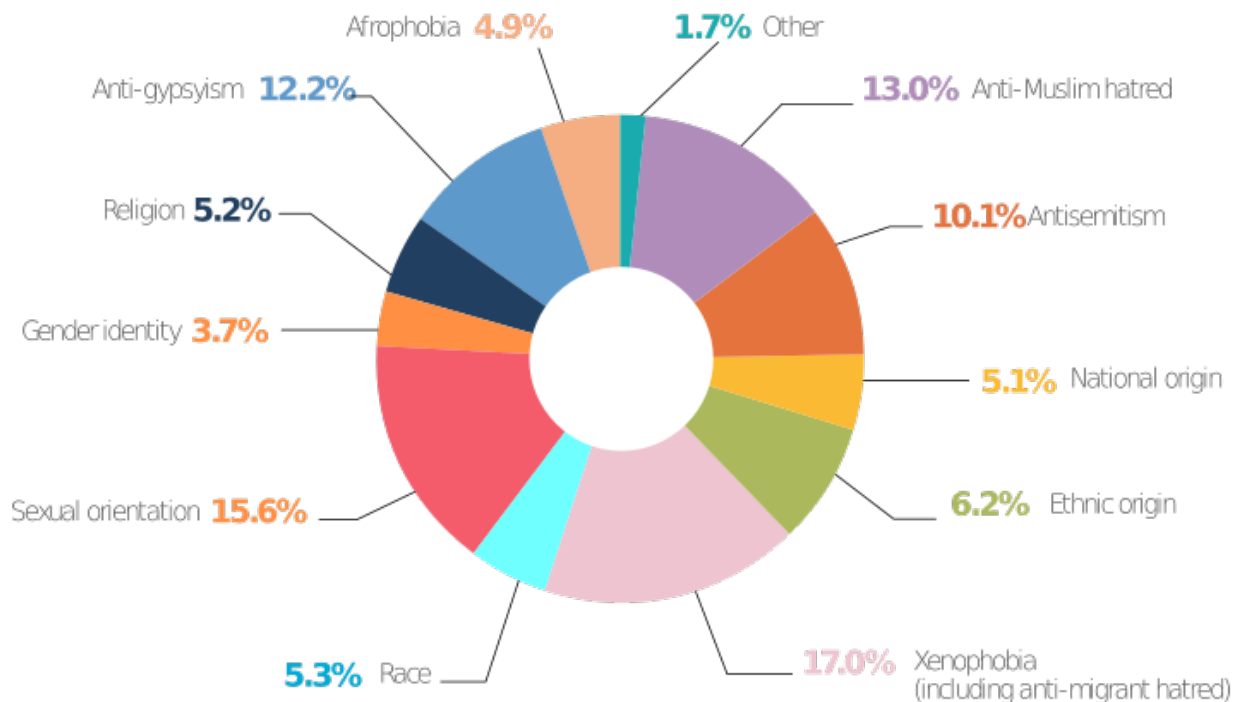
Dunque, se l'hate speech online non è intrinsecamente diverso dall'hate speech espresso negli spazi offline, la natura della sfera online pone sfide uniche e peculiari che rendono difficile l'individuazione delle responsabilità, lo sviluppo di risposte legali adeguate e l'applicazione delle norme vigenti. Per questa ragione sono necessari approcci nuovi, che prendano in considerazione le peculiarità delle tecnologie digitali.

Nel maggio 2016 la Commissione europea e quattro grandi piattaforme di social media

(Facebook, Twitter, YouTube e Microsoft) hanno adottato un [codice di condotta](#) per contrastare i discorsi d'odio illegali online. Anche se non si tratta di misure giuridicamente stringenti, in base all'accordo le aziende si impegnano a rimuovere i contenuti d'odio illegali entro 24 ore dalla segnalazione, agendo in linea con le leggi nazionali, in particolare con quelle che traspongono la [decisione quadro](#) dell'Unione europea sul razzismo e la xenofobia e la [direttiva europea sul commercio elettronico](#).

L'applicazione del codice di condotta viene monitorata e descritta in rapporti periodici. Secondo l'[ultimo rapporto](#) (febbraio 2019), le aziende coinvolte hanno rimosso in media il 71,7% dei contenuti segnalati dagli utenti e dalle organizzazioni della società civile. I casi di rimozione sono [cresciuti costantemente](#) dal maggio 2017 al dicembre 2018 (data dell'ultima elaborazione del dato). Sussistono tuttavia delle differenze tra le piattaforme: Twitter ad esempio è piuttosto lento o restio a rimuovere contenuti, dando seguito a solo il 43,5% delle segnalazioni. Sono emerse anche differenze nazionali rilevanti: le percentuali di rimozione dei contenuti vanno dal 38% del Portogallo al 100% di Cipro.

Grounds of hatred



Tipologie dei contenuti d'odio rimossi dalle piattaforme che hanno sottoscritto il Codice di condotta dell'UE ([gennaio 2018](#))

La scelta della Commissione europea di affidare un ruolo tanto preminente alle aziende private [è stata criticata](#) sotto diversi aspetti. Questo approccio sembra infatti andare verso una [privatizzazione](#) del sistema di tutela dei diritti fondamentali, per cui grandi piattaforme tecnologiche si trovano di fatto ad agire come legislatori, giudici ed esecutori, [aggirando la supervisione da parte della magistratura](#). Come sostiene il [Centre for Democracy and Technology](#), un *think tank* che si occupa di tutela dei diritti civili sulla rete, l'assenza di supervisione giudiziaria [è molto problematica](#), soprattutto perché le questioni legate all'hate speech avrebbero bisogno di un dibattito pubblico aperto e trasparente.

Sempre secondo i critici, rimuovere i discorsi d'odio dal web non significa rimuoverli dalla società, ma piuttosto eliminare la possibilità di contrastarli con argomentazioni equilibrate ed informate (il cosiddetto "counter-speech"). Altre contestazioni riguardano la scarsa trasparenza degli algoritmi che operano un primo filtraggio dei contenuti segnalati alle piattaforme, la qualità del lavoro dei moderatori, l'eccessiva fiducia riposta dalla Commissione europea negli strumenti tecnici e nei cosiddetti "trusted flaggers", utenti certificati che si impegnano a segnalare alle piattaforme i contenuti illegali.

Un'altra critica ricorrente riguarda la scarsa trasparenza prevista del Codice di condotta, la cui stessa elaborazione è stata accusata di opacità. Non sono infatti previste banche dati pubbliche sull'operato delle piattaforme e sulle loro decisioni, né sono chiari i criteri che guidano le loro valutazioni: si preoccupano di applicare i termini di servizio oppure le normative europee e nazionali? Rispetto al ruolo delle aziende è intervenuto anche il Consiglio d'Europa, che in [uno studio](#) del 2011 metteva in guardia: "le aziende non sono immuni da interferenze ingiustificate. Le loro decisioni a volte nascono da pressioni politiche dirette o da obblighi economici motivati politicamente, giustificati dal rispetto dei loro termini di servizio".

Il caso tedesco

La Germania è uno degli stati europei più intransigenti sulla rimozione di dell'hate speech online. Nel 2015, in parallelo alla decisione del governo tedesco di accogliere fino a un milione di rifugiati e all'aumento di episodi di crimini d'odio, inclusi diversi [attacchi ai giornalisti](#) come riportato in questo [video](#), il Paese [ha chiesto](#) a Facebook, Twitter e Youtube di sottoscrivere un accordo per impegnarsi a cancellare contenuti d'odio entro 24 ore dalla segnalazione, nel rispetto delle [leggi nazionali](#) e con la collaborazione di organizzazioni come [Correctiv](#).

Il governo tedesco ha però ritenuto insufficiente la rimozione dei contenuti d'odio effettuata dalle grandi piattaforme, e ha dunque introdotto la [Network Enforcement Law \(NetzDG\)](#). Il ministro della Giustizia [ha spiegato](#) che la legge era necessaria poiché la "radicalizzazione verbale è spesso il primo passo verso la violenza fisica". La nuova normativa è entrata in vigore il 1° gennaio

2018, e ha dunque introdotto la [Network Enforcement Law \(NetzDG\)](#). Il ministro della Giustizia [ha spiegato](#) che la legge era necessaria poiché la “radicalizzazione verbale è spesso il primo passo verso la violenza fisica”. La nuova normativa è entrata in vigore il 1° gennaio 2018, e prevede per i social network che non cancellano contenuti considerati illegali una sanzione che può arrivare a 50 milioni di euro.

La legge è stata criticata da deputati del Partito dei verdi come Renate Kuenast, secondo la quale la paura di incorrere in una sanzione potrebbe portare a cancellare contenuti offensivi senza essere certi della loro illegalità. [La Federazione dei giornalisti tedeschi](#) ha attirato l'attenzione sulle “responsabilità giornalistiche dei contenuti che non possono essere delegate agli operatori di piattaforme come Facebook perché potrebbero cancellare contenuti per motivi commerciali anziché preoccupazioni editoriali”. Affermando a sua volta che la libertà di espressione va protetta a tutti i costi, Andrus Ansip, Commissario europeo per il mercato unico digitale, [ha sottolineato](#) la necessità di migliorare la [media literacy](#) e il pensiero critico per porre rimedio al fenomeno dell'hate speech.

Il caso francese

La [legge sulla stampa del 1881](#) inquadra la libertà di espressione e punisce anche i discorsi d'odio illegali. Dopo gli attacchi terroristici degli ultimi anni, il Paese ha assistito a un massiccio incremento dell'hate speech, [utilizzato anche da alcuni politici](#).

Nel 2015 il governo ha avviato un [piano nazionale biennale contro razzismo e antisemitismo](#), al quale sono stati destinati 100 milioni di euro. La misura, oltre ad inasprire le leggi sul crimine d'odio, ha comportato l'ideazione di un piano di [lotta contro l'hate speech online](#), che prevedeva la creazione di un'[unità specifica dedicata alla protezione degli utenti internet dalla diffusione dell'odio](#), a partire dalle piattaforme social. A tale [unità](#) sono stati attribuiti diversi compiti tra cui la vigilanza sulla rete, la sorveglianza delle piattaforme social e la facilitazione delle segnalazioni di hate speech per gli internauti.

Le aziende come Facebook, Twitter e Youtube rientrano nello statuto di “[ospiti](#)” secondo la legge francese. Ciò comporta che non sono responsabili di quanto viene pubblicato sulle loro piattaforme, a meno che non ne siano stati allertati. Dopo la segnalazione [la legge](#) li obbliga a riferire il contenuto manifestamente illegale alla magistratura. Nel 2016 [due organizzazioni della società civile](#) hanno rilevato che le tre piattaforme social avevano rimosso solo una percentuale minima (4% per Twitter, 7% per Youtube e 34% per Facebook) dei 586 contenuti segnalati. La scarsa reattività delle piattaforme ha portato nel 2017 a un ricorso da parte di [sei organizzazioni anti-razziste e anti-omofobe](#). Nel luglio 2019, l'Assemblea nazionale francese [ha approvato](#) un disegno di legge che vorrebbe imporre ai giganti dell'informatica come Facebook e Twitter [la rimozione dei contenuti d'odio entro 24 ore](#). In caso di mancato riscontro nei tempi stabiliti o di mancata offerta agli utenti dei mezzi necessari per denunciare simili contenuti i network potrebbero andare incontro a sanzioni che prevedono una pena carceraria e un'ammenda tra i 250.000 e i 1,25 milioni di euro a seconda che il ‘propagatore’ di odio sia una persona fisica o giuridica. Il nuovo disegno di legge, paragonato alla Network Enforcement Law (NetzDG) tedesca, è stato considerato ben più severo dal punto di vista delle sanzioni previste e ha attirato critiche e timori analoghi di “censura abusiva” da parte della società civile, ma non dei politici d'opposizione. Per essere approvato il disegno deve passare il vaglio del Senato.

Per quanto riguarda la questione [dei confini tra la libertà di espressione e l'hate speech](#), il [dibattito](#) francese si è sviluppato attorno a due filoni principali: in primo luogo, la definizione del confine che separa l'incitamento all'odio da un crimine d'odio; in secondo luogo, la ricerca di un percorso educativo o ri-educativo che accompagni la rimozione dei contenuti ritenuti pericolosi per superare la ‘sensazione di censura’. Inoltre, dopo l'attentato a *Charlie Hebdo*, in Francia si è dibattuto su come i confini della libertà d'espressione varino a seconda di chi è oggetto di hate speech (il cosiddetto ‘[doppio standard](#)’).

Come prevenire l'hate speech?

Molti osservatori, inclusi l'[UNESCO](#) e la [Relatrice speciale delle Nazioni unite sulle questioni delle minoranze](#), ritengono che le misure di ordine legale non siano sufficienti per contrastare l'hate speech. Servono anche delle risposte sociali, culturali ed educative, che vadano ad affrontare le cause recondite del fenomeno, considerato sintomo di un problema più profondo.

Secondo l'[UNESCO](#) l'educazione e la formazione alla cittadinanza digitale consapevole sono una chiave fondamentale per prevenire i discorsi d'odio online. Esse includono l'educazione ai diritti umani e all'utilizzo sicuro di internet, la promozione dell'*information* e [media literacy](#) e lo sviluppo di capacità critiche. Nella moltitudine di [iniziative](#) caratterizzate da questo approccio emergono essenzialmente tre obiettivi comuni: informare, analizzare e agire; sono obiettivi complementari tra loro.

Sul fronte delle attività di informazione possono essere citate le [campagne](#) avviate in questi anni per creare consapevolezza sul fenomeno e sulle sue conseguenze, [la raccolta e la disseminazione](#) di informazioni sull'hate speech e la divulgazione delle norme di riferimento. Per quanto concerne l'analisi del fenomeno, risulta innanzitutto essenziale identificare [i discorsi d'odio](#), le loro radici e le loro diverse [forme](#), in modo da poterli riconoscere e mettere a nudo. Anche a livello delle azioni da intraprendere ci sono diverse alternative. Scrivere contro l'hate speech, monitorare i media, comprenderne le dinamiche e ribattere alterandone il contenuto sono tra le strade più utilizzate.

L'iniziativa [No Hate Speech](#), promossa dal Consiglio d'Europa e rivolta ai giovani, è un esempio della complementarità dei diversi obiettivi dell'approccio educativo. Da una parte il movimento promuove una [campagna](#) contro l'hate speech in oltre 40 paesi; dall'altra mira ad accrescere l'alfabetizzazione sui discorsi d'odio, attraverso [varie risorse](#) e un [manuale](#) che affronta la questione dalla prospettiva dei diritti umani. Il movimento opera inoltre come una [piattaforma](#) di monitoraggio, affinché gli utenti possano segnalare e discutere contenuti d'odio online.

I giovani non sono l'unico gruppo a cui si rivolgono le [iniziative educative](#) sui discorsi d'odio:

sono numerosi anche gli strumenti di formazione mirati alle forze dell'ordine, alla magistratura, agli insegnanti e ad altri membri della società civile. Ad esempio [il progetto europeo Prism](#) sviluppa [attività trasversali](#) con l'obiettivo di elaborare strategie e pratiche efficaci per stimolare un migliore utilizzo del linguaggio, contrastare i discorsi d'odio e promuovere la cultura del rispetto.

Per tutti questi progetti di prevenzione e contrasto all'hate speech, lo sviluppo delle competenze digitali è considerato come un elemento indispensabile per prevenire, mettere a nudo e combattere i discorsi d'odio online. Ciò è visibile anche dalla vasta gamma di formati digitali utilizzati ([video](#), [infografiche](#), siti web, blog, social media), che rendono possibile raggiungere più persone in maniera trasversale.

Tuttavia, sebbene negli ultimi anni siano state condotte ricerche trasversali e comparate di ampio respiro come quella realizzata nell'ambito del progetto [Positive Messenger](#), non ci sono valutazioni esaustive sui risultati delle risposte educative e culturali ai discorsi d'odio e non risulta chiaro quale sia il loro impatto effettivo. Sembra ad ogni modo assodato che siano necessarie simili iniziative di contrasto ai discorsi d'odio, alternative alle misure di ordine legale.

Qual è il ruolo dei giornalisti?

Se il ruolo dei social network nel propagare l'odio è ampiamente dibattuto, [meno attenzione](#) viene dedicata agli organi di informazione tradizionali. Già nel 1978 l'[Unesco richiamava](#) le responsabilità dei mass media nella lotta al razzismo, sottolineandone il ruolo nel "dare espressione ai popoli oppressi [...] diffondendo informazioni sugli obiettivi, le aspirazioni, le culture e i bisogni di tutti i popoli [...] senza distinzione di razza, sesso, lingua, religione o nazionalità." Un ventennio più tardi, con riferimento alle guerre di dissoluzione jugoslava, il Consiglio d'Europa richiamava [la responsabilità dei media](#) nell'opporci alla violenza e al linguaggio dell'odio.

Il [monitoraggio dei media](#) in diversi contesti nazionali ha messo in evidenza come questi veicolino spesso discorsi d'odio, anche se non sempre in modo consapevole. Questo è particolarmente vero nei periodi elettorali, quando emerge il problema della "[notiziabilità dell'odio](#)", ovvero della circolazione virale di polemiche e provocazioni: i media che decidono di darvi spazio contribuiscono a diffondere i discorsi d'odio. Emerge da ciò l'importanza delle [occasioni di formazione](#)

rivolte ai giornalisti per farli [riflettere](#) su come riportare [notizie](#) senza ricadere nei discorsi d'odio, e per diffondere [buone pratiche](#).

Che questo aspetto sia centrale nell'esercizio della professione giornalistica è confermato anche dal fatto che il contrasto all'hate speech trova spazio in numerosi codici di autoregolamentazione professionale. Secondo il database "[Accountable journalism](#)" – la più grande collezione al mondo di codici etici in vigore nelle organizzazioni della stampa – ammontano a 120 i documenti ufficiali in cui viene affrontato il tema.

Dal canto suo, l'[Ethical Journalism Network](#) – una rete che ha tra i suoi obiettivi quello di aiutare i giornalisti di tutto il mondo ad affrontare [i dilemmi etici posti dalla cosiddetta "era della post-verità"](#) – ha di recente [divulgato](#) un [test](#) per mettere in guardia i giornalisti dall'utilizzo e dalla divulgazione inconsapevole dei discorsi d'odio. Prima di decidersi a riportare una dichiarazione, il test suggerisce di chiedersi chi siano il mittente e i destinatari del messaggio, quali obiettivi abbia chi lo lancia, quale sia il contenuto del messaggio e in quale contesto verrà diffuso.

Queste domande aprirebbero la strada a ulteriori discussioni su cosa sia o non sia una notizia: anche animata dalle migliori intenzioni, la scelta di non offrire visibilità a un contenuto d'odio si espone al rischio di "censurare" un fenomeno che può essere combattuto solamente nella misura in cui se ne è al corrente. Un giornalista che voglia contrastare efficacemente il discorso d'odio deve impegnarsi anzitutto a non diffonderlo, o a rendere noti e magari criticare coloro che ne fanno uso?

Ancora una volta, i dubbi e le domande sono molteplici, e risposte univoche sembrano non esistere. Certamente esiste un problema

HATE SPEECH
TURNING THE PAGE OF HATE:
A MEDIA CAMPAIGN FOR
TOLERANCE IN JOURNALISM

When it comes to hate speech, journalists and editors must pause and take the time to judge the **potential impact** of offensive, inflammatory content.
The following test, developed by the EJN and based on international standards, highlights questions in the **gathering, preparation and dissemination** of news and helps place what is said and who is saying it in an **ethical context**.

2 REACH OF THE SPEECH
How far is the speech traveling?
Is there a **pattern** of behaviour?

3 GOALS OF THE SPEECH
How does it benefit the **speaker** and their **interests**?
Is it **deliberately intended** to cause harm to others?

A 5 POINT TEST FOR JOURNALISTS

4 THE CONTENT ITSELF
Is the speech **dangerous**?
Could it incite **violence** towards others?

5 SURROUNDING CLIMATE
SOCIAL / ECONOMIC / POLITICAL
Who might be **negatively affected**?
Is there a history of **conflict or discrimination**?

1 STATUS OF THE SPEAKER
How might their **position** influence their **motives**?
Should they even be **listened to** or just **ignored**?

DONT **SENSATIONALISE!**
AVOID THE **RUSH** TO PUBLISH
TAKE A **MOMENT OF REFLECTION**

EthicalJournalismNetwork.org
CC BY NC SA SHARE IT! EJN

antico (l'odio espresso tra esseri umani) e certamente esistono nuove tecnologie in grado di veicolarlo (ad esempio i social network). Al contempo però esistono anche cittadini, istituzioni e organizzazioni transnazionali che danno vita, con metodo democratico, a un dibattito cruciale del nostro tempo, e che si impegnano nella ricerca di soluzioni possibili.

*Dossier realizzato nell'ambito del progetto European Centre
for Press and Media Freedom (ECPMF)*

Osservatorio Balcani e Caucaso Transeuropa

Unità operativa del Centro per la Cooperazione Internazionale

www.balcanicaucaso.org

redazione@balcanicaucaso.org

Enti finanziatori: Commissione Europea
Provincia autonoma di Trento
Comune di Rovereto

