



## #Silence Hate: Study on hate Speech Online in Belgium, Czech Republic, Germany and Italy

Il **report europeo #Silence Hate** è la sintesi di **quattro studi** condotti in **Belgio, Repubblica Ceca, Germania e Italia sull'hate speech online**, nell'ambito del progetto BRICKS - Building Respect on the Internet by Combating Hate Speech.

Secondo la **definizione** data dal **Consiglio d'Europa nel 1997**, per **“discorsi d'odio”** si intendono: **“tutte le forme di espressione che diffondono, incitano, promuovono o giustificano l'odio razziale, la xenofobia, l'antisemitismo o altre forme di odio basate sull'intolleranza, inclusa l'intolleranza espressa da nazionalismo ed etnocentrismo aggressivi, la discriminazione e l'ostilità nei confronti delle minoranze, dei migranti e delle persone di origine straniera.”**

I **discorsi d'odio** sono **divenuti** una **questione rilevante** sia a livello politico che nel dibattito pubblico, assumendo particolare importanza **nel 2015**. In **Italia nel 2014 l'UNAR** (Ufficio Nazionale Antidiscriminazioni Razziali) ha registrato ben **347 casi di espressioni razziste su social network**, di cui 185 casi su Facebook, Twitter e Youtube. Questa tendenza è aumentata nel 2015 ed in particolare dall'estate 2015 si è assistito ad un incremento dei **discorsi di odio diretti ai rifugiati**, soprattutto rifugiati **musulmani, e a persone che operano in loro favore** come volontari, politici e attivisti. Questi **fenomeni** sono chiaramente **collegati alla crisi Europea dei rifugiati**, che ha portato non solo all'aumento delle forme di supporto verso i migranti da parte della società civile ma anche ad un incremento dei discorsi d'odio online.

I **forum dei media online**, i commenti agli articoli e le **pagine Facebook dei giornali nazionali e locali** sono **gli spazi virtuali in cui hanno origine i discorsi di odio** rivolti contro rifugiati e cittadini di origine straniera.

Nei quattro paesi presi in esame dalla ricerca emergono **quadri normativi differenti**. Mentre in **Germania** è molto difficile definire i discorsi d'odio in senso giuridico, essendoci una differenziazione molto generale tra espressioni di opinioni accettabili e non accettabili, in **Repubblica Ceca** l'ordinamento giuridico persegue chi esprime discorsi d'odio anche online. In **Italia e Belgio** non vi è una legislazione specifica riguardante l'*hate speech* anche se vi sono una serie di norme sull'incitamento all'odio razziale e altre forme di discriminazione.

In generale, per quanto riguarda le **pratiche giudiziarie**, uno degli **aspetti più critici** dei **discorsi d'odio online** è che gli **autori** di tali commenti **raramente vengono riconosciuti come responsabili** dal momento che spesso è molto difficile determinare la loro identità. Inoltre, un'altra questione cruciale riguarda la **difficile identificazione** delle **persone responsabili della rimozione dal web di contenuti che incitano all'odio** e infrangono la legge.



Nonostante alcuni paesi, come Italia e Repubblica Ceca, abbiano creato enti pubblici appositi per prevenire e monitorare il diffondersi sul web dei discorsi di incitamento all'odio, dall'analisi dei casi studio emerge un quadro complesso e problematico. Un monitoraggio, durato 3 mesi, delle notizie e degli articoli pubblicati su diversi siti web nei 4 paesi presi in esame ha evidenziato che:

- I **topic** che innescano discorsi d'odio sono **diversi e dipendono dai contesti politici e culturali dei singoli stati**. Nessun topic è esente dal rischio di tali commenti, anche se quelli che presentano un maggior rischio di attirare commenti d'odio sono quelli che fanno leva sul **coinvolgimento emozionale** delle persone, in particolare gli **eventi violenti**.
- La **moderazione dei contenuti on line** fa la differenza. La pre-moderazione è sempre meno utilizzata sia perchè richiede maggiori risorse umane sia per la diffusione sempre maggiore di Facebook come canale di comunicazione, che non prevede la pre-moderazione dei post. La post-moderazione è la più comune forma di moderazione usata in Belgio, in una certa misura in Repubblica Ceca e Germania e solo a volte in Italia. La moderazione è un importante strumento di prevenzione dei discorsi d'odio. L'evidenza mostra che laddove la moderazione è assente o insufficiente i discorsi d'odio si propagano senza limiti.
- Il **profilo degli hate speakers** non sembra avere particolari caratteristiche: sono uomini e donne, con differenti idee politiche, residenti in differenti regioni e di età diverse. Nello studio italiano i ricercatori hanno identificato diversi "profili" di *hate speakers*: Rassegnati, Arrabbiati e Aggressivi.
- L'**atteggiamento dei giornalisti nei confronti della partecipazione online**. Le redazioni sono in genere favorevoli alla partecipazione online degli utenti attraverso i commenti. Per i giornali il "traffico" generato dai commenti negativi può essere **economicamente conveniente**. Tuttavia, la diffusione dei discorsi d'odio online e il deleterio impatto di questo fenomeno sulla reputazione dei giornali sta portando alcune redazioni a riconsiderare la partecipazione online come un elemento di disturbo.
- **Le differenti strategie di regolazione delle discussioni**. Quasi tutti gli intervistati concordano sull'idea che le sezioni dei commenti debbano essere moderate, con poche eccezioni. Gli strumenti menzionati dagli intervistati sono: norme e politiche pubbliche, pre-moderazione, moderazione attiva, rimozione dei commenti, termini vietati, tempi limitati di discussione, "classificazione" degli utenti, recompense e programmi fedeltà, non apertura o chiusura delle discussioni. Tutte queste strategie seguono un "approccio del No", ovvero sono orientate soltanto a limitare la discussione e punire il comportamento non desiderato. In Belgio si sta sperimentando un approccio proattivo e alternativo che consiste nel pubblicare online solo i commenti più interessanti.
- **Approcci innovativi alla regolazione delle discussioni su internet** sono stati sperimentati soprattutto in Belgio. In generale alcuni giornalisti credono che gli stessi lettori possano assumere un ruolo attivo nella riduzione dei discorsi d'odio online. Vengono quindi incentivati ad essere i filtri iniziali contro i commenti d'odio, moderando e denunciando gli abusi e vengono premiati i commenti positivi. Altre strategie utilizzate sono la diversificazione delle forme di partecipazione, ad esempio attraverso format specifici entro cui elaborare commenti più strutturati. La costruzione di **community active** è l'obiettivo principale alla quale la maggioranza delle testate ambiscono.