



*Dipartimento di
Scienze Politiche*

*Cattedra di Diritto di
Internet: Social Media e
Discriminazione*

L'evoluzione del contrasto all'*hate speech* tra online ed offline

RELATORE

Prof. Pietro Santo
Leopoldo Falletta

CANDIDATA

Assia Leoni
Matr. 075132

ANNO ACCADEMICO 2015/2016

Indice

INTRODUZIONE	3
CAPITOLO 1: L' <i>HATE SPEECH</i>	
1.1. La definizione di <i>hate speech</i>	4
1.2. Le componenti dell' <i>hate speech</i>	6
1.3. L'avvento del cyberspazio: odio online ed offline	8
CAPITOLO 2: IL CONTRASTO ALL' <i>HATE SPEECH</i>	
2.1 Il contrasto all' <i>hate speech</i> nell'Ordinamento Italiano	12
2.2 La normativa dell'Unione Europea in materia di <i>hate speech</i>	16
2.3 Il consiglio d'Europa	18
2.4 La giurisprudenza della Corte Europea dei Diritti dell'Uomo (CEDU)	19
2.5 Un progetto europeo per contrastare l'odio: PRISM	20
CAPITOLO 3: LA RISPOSTA DI INTERNET ALL' <i>HATE SPEECH</i>	
3.1 Le policy e la responsabilità dei social network	22
3.2 L' <i>hate speech policy</i> di YouTube	24
3.3 Facebook e l' <i>hate speech</i>	24
3.4 L'approccio di Twitter all'odio online	26
CONCLUSIONI	28
BIBLIOGRAFIA E SITOGRAFIA	31
ABSTRACT	32

INTRODUZIONE

Nel presente studio andremo ad analizzare il crescente fenomeno dell'*hate speech*, categoria della giurisprudenza americana che da qualche anno è entrata a far parte anche di quella europea, identificativa di un genere di discorsi che hanno il fine di esprimere odio ed intolleranza nei confronti di una persona od un gruppo sia questo razziale, etnico, religioso, di genere o orientamento sessuale. In nostro obbiettivo è quello di inquadrare il fenomeno dal punto di vista normativo, facendo riferimento all'ambito nazionale e sovranazionale ed esaminare i provvedimenti messi in atto per arginare questo fenomeno all'interno dei *social network*.

All'interno del primo capitolo, dopo aver effettuato una panoramica sulle diverse definizioni che si sono date nel corso degli anni per definire questo fenomeno, abbiamo proceduto ad un'analisi delle sue componenti; effettuando uno studio su come tale fenomeno si sia evoluto attraverso l'avvento di internet, il quale ha infatti dato un'accelerazione non indifferente alla comunicazione facendo trovare i media tradizionali in una situazione di profonda arretratezza rispetto ai nuovi strumenti digitali, ciò soprattutto data la peculiarità di questi mezzi. Si è proceduto quindi ad un'analisi di confronto tra l'*hate speech* online e quello offline, definito <<tradizionale>>, per vedere in cosa differiscono e per comprendere meglio quanto l'avvento di internet abbia realmente influito su questo fenomeno.

Nel secondo capitolo invece ci siamo soffermati su come tale fenomeno possa essere contrastato, senza però andare a ledere il diritto alla libertà d'espressione: lavoro che necessita di una continua opera di bilanciamento tra la garanzia alla libertà d'espressione e il rispetto della dignità umana. Andremo quindi ad analizzare quale sia stato l'approccio a questo fenomeno a vari livelli: partendo dal nostro ordinamento interno e quindi dall'articolo 21 della Costituzione in cui viene sancito il diritto alla libertà di parola, attraverso un veloce excursus su come la giurisprudenza italiana in merito si sia evoluta; passando per il ruolo che l'Unione Europea ha avuto nel definire e contrastare l'*hate speech*, fino ad analizzare il ruolo svolto dagli organismi internazionali quali il Consiglio d'Europa e la Corte Europea dei Diritti dell'Uomo, che si sono impegnati ad emanare un vasto corpo di leggi a riguardo. Il tutto sarà avvalorato dall'analisi di sentenze a riguardo, per meglio aiutarci a comprendere il punto di vista della giurisprudenza.

Con il terzo capitolo infine andremo ad analizzare come internet si sia approcciato all'*hate speech* e come abbia cercato di tenere sotto controllo questo fenomeno, analizzando il primis il ruolo degli *Internet Service Provider* e la possibilità di un loro intervento preventivo sul fenomeno, per poi giungere ad un'analisi comparativa tra le varie *policy* dei *social network*; anche in quest'ultima fase con l'aiuto di alcune sentenze in merito.

Capitolo 1: L'hate speech

1.1 La definizione di hate speech

L'*hate speech*, solitamente tradotto in italiano con l'espressione <<discorsi d'odio>>, va collocato all'interno di quella categoria di illeciti definita <<*Hate crimes*>> dentro la quale si trovano le violenze commesse nei confronti di persone in ragione della loro appartenenza ad un gruppo sociale, individuate sulla base della razza, dell'etnia, della religione, dell'orientamento sessuale, dell'identità di genere o di particolari condizioni fisiche o mentali¹. L'*hate speech* è una categoria della giurisprudenza americana che da qualche anno è entrata a far parte anche di quella europea e ricomprende al suo interno tutti i discorsi d'odio e le espressioni che contengono elementi discriminatori, indica cioè un genere di parole e discorsi che hanno lo scopo di esprimere odio e intolleranza verso una persona o un gruppo. Questo fenomeno è in crescita in tutta Europa e sfrutta il web, in particolar modo i *social media*, facendo appello alla libertà d'espressione, sancita dall'articolo 10 della Convenzione Europea dei Diritti dell'Uomo. La CEDU è però allo stesso tempo uno dei primissimi documenti internazionali che contiene un'esplicita affermazione del principio di non discriminazione.² L'articolo 14 della CEDU prevede che << il godimento dei diritti e delle libertà riconosciuti nella presente convenzione deve essere assicurato senza nessuna discriminazione, in particolare quelle fondate sul sesso, la razza, il colore, la lingua, la religione, le opinioni politiche o quelle di altro genere l'origine nazionale, la ricchezza la nascita o ogni altra condizione.>> Tale articolo è stato successivamente revisionato, approfondito e inserito nel Trattato sul Funzionamento dell'Unione Europea dove i contenuti della tutela fino ad allora incentrati su cause di discriminazione per motivi di sesso e nazionalità, sono stati ampliati fino a ricomprendere tutti i diritti civili, politici e sociali. Molti studiosi, come analizzeremo in seguito riportando le definizioni che ognuno di loro dà dell'*hate speech*, ritengono che non esista una nozione condivisa di cosa realmente vada inteso come *hate speech*; le differenze possono essere sensibili spostandosi da un ordinamento giuridico all'altro.³ Per fare qualche esempio, nell'ambito del contrasto al razzismo è considerato *hate speech* qualsiasi discorso vada a pregiudicare le minoranze etniche; in modo differente, per gli intellettuali di stampo femminista, l'*hate speech* è individuato nella pornografia, o per finire i membri della comunità ebraica lo identificano in forme di negazionismo. Una nozione che può essere presa a riferimento la troviamo nella Convenzione o Patto Internazionale sui Diritti Civili o Politici, dove l'Articolo 20 enuncia: << 1. Qualsiasi propaganda a favore della guerra deve essere vietata dalla legge. 2. Qualsiasi appello all'odio nazionale, razziale o

¹ M. MENSÌ – P. FALLETTA, *Il diritto del web. Casi e materiali*, Padova, 2015, p.180.

² M. MENSÌ – P. FALLETTA, *op. cit.*, p. 174.

³ G. ZICCARDI, *L'odio online. Violenza verbale e ossessione in rete*, Milano, 2016, p. 19.

religioso che costituisca incitamento alla discriminazione, all'ostilità o alla violenza deve essere vietato dalla legge>>. Possiamo quindi tracciare una prima definizione di hate speech dove troviamo sei elementi centrali divisibili in due gruppi: tre <<ambiti>> (i nazionalismi, il razzismo e la religione, quali strumenti di discriminazione e di odio) e l'incitamento a tre comportamenti (discriminazione, ostilità e violenza).⁴ Un'altra fonte molto interessante per poter giungere ad una più completa definizione di *hate speech* ci è fornita poi dalla Raccomandazione n. 20 del 30 ottobre 1997 dell'Consiglio d'Europa sull'*hate speech*. Il termine viene qui interpretato come atto a comprendere ogni qual forma di espressione che diffonda, inciti, promuova o giustifichi odio razziale, xenofobia, antisemitismo o altre forme di odio basate sull'intolleranza, compresa quella espressa da nazionalismo aggressivo ed etnocentrismo. Anche l'ICERD, Convenzione internazionale sull'eliminazione di ogni forma di discriminazione razziale del dicembre 1965, obbliga i governi a punire atti di discriminazione razziale, sia nelle istituzioni pubbliche che in ambito privato.

Nonostante le molte sfaccettature ed interpretazioni che riguardano la definizione di *hate speech* possiamo giungere alla conclusione che siano necessari almeno tre requisiti affinché un'espressione possa considerarsi *hate speech*: 1) la volontà e l'intenzione di incitare l'odio con ogni mezzo di comunicazione; 2) l'esortazione vera e propria, idonea a causare atti di odio e violenza sui soggetti presi di mira; 3) la concretezza del rischio che gli atti di violenza o discriminazione si verifichino. Ciò comporta che se le affermazioni non hanno l'intento di incitare terzi all'odio, non rientrano nell'*hate speech*.

Affinché possa ritenersi *hate speech* è quindi è fondamentale che l'incitamento all'odio o alla violenza sia connesso strettamente all'idea di danno, discriminazione o violenza. Nel 1954 Gordon Allport⁵, elaborò una scala del pregiudizio e della discriminazione nel tentativo di "misurare" l'odio nella società (la quale venne definita come Allport's scale of Prejudice and Discrimination).⁶ La scala consta di cinque gradi di gravità dei comportamenti.

- 1) *Antilocution*. È il gradino più basso della scala, nel quale i membri del gruppo di maggioranza deridono, spesso con stereotipi negativi o immagini non veritiere, un gruppo di minoranza.
- 2) *Avoidance*. Qui i componenti del gruppo di maggioranza evitano volontariamente i membri del gruppo di minoranza, con un ovvia conseguenza, l'isolamento.

⁴ G. ZICCARDI, *L'odio online. Violenza verbale e ossessione in rete*, Milano, 2016.

⁵Gordon Allport figura di spicco della psicologia statunitense- è stato un noto psicologo della cosiddetta psicologia sociale ed esponente dell'approccio cognitivista, fondamentali soprattutto per quanto riguarda la nostra analisi i suoi studi sul pregiudizio .

⁶ G.W. ALLPORT, *The Nature of Prejudice*, Cambridge, 1949; trad. it. , *La natura del pregiudizio*, Firenze, 1973.

- 3) *Discrimination*. In questo terzo livello viene impedito alla minoranza di avere accesso in condizioni di eguaglianza alle opportunità, ai beni e ai servizi della società in cui i gruppi convivono.
- 4) *Physical attack*. Il quarto stadio rappresenta l'aggressione fisica. Prendono qui forma i cosiddetti *hate crimes*. I componenti dei gruppi di minoranza sono presi di mira con atti vandalici o attacchi fisici violenti.
- 5) *Extermination*. Si può infine arrivare alla volontà del gruppo di maggioranza di sterminare i componenti del gruppo di minoranza, eliminando l'esistenza stessa degli avversari.

Quindi nonostante le difficoltà che abbiamo incontrato nel delineare una chiara e univoca definizione di *hate speech*, dopo una prima analisi inizia ora a prendere forma ciò che caratterizzerà la base di questo vasto fenomeno. Andiamo ora ad analizzarla più nello specifico.

1.2. Le componenti dell'*hate speech*

Un altro metodo attraverso il quale si è tentato di definire *l'hate speech*, si caratterizza per l'individuazione di cinque componenti⁷: il mezzo attraverso il quale avviene la comunicazione; il contenuto del messaggio, il destinatario, il contesto all'interno del quale avviene e gli effetti che questo provoca.

Per quanto riguarda il mezzo, una delle domande che si pone la giurisprudenza è se sia possibile far coincidere *l'hate speech* online con quello trasmesso tramite i tradizionali mezzi di comunicazione. Analizzando la sentenza del tribunale di St Paul, Minnesota: <<Whoever places on public or private property a symbol, object, appellation, characterization or graffiti, including, but not limited to, a burning cross or Nazi swastika, which one knows or has reasonable grounds to know arouses anger, alarm or resentment in others on the basis of race, color, creed, religion or gender commits disorderly conduct and shall be guilty of misdemeanor>>⁸ troviamo quindi quali elementi che caratterizzano *l'hate speech*; simboli, oggetti, graffiti, appellativi. Ed il mezzo attraverso il quale il

⁷<http://sosracismo.eu/wp-content/uploads/2016/07/Hate-Crime-and-Hate-Speech-in-Europe.-Comprehensive-Analysis-of-International-Law-Principles-EU-wide-Study-and-National-Assessments.pdf>

⁸ R.A.V. v City of Saint Paul 505 U.S. 377 (1992). È un caso della Corte Suprema degli Stati Uniti d'America che coinvolge *l'hate speech* e la clausola di libertà di parola del primo Emendamento della Costituzione americana. Un gruppo di ragazzi, tra cui R.A.V. che al momento della sentenza era un adolescente, ha posizionato una croce –fatta con pezzi di legno presi dalle gambe di una sedia– nel giardino di una casa di una famiglia afroamericana, per poi dargli fuoco. Il ragazzo è stato accusato tra l'altro dalla *St. Paul, Minnesota, Bias Motivated Crime Ordinance* che vieta l'esposizione di un simbolo che si ha motivo di sapere suscitare rabbia o risentimento sulla base di razza colore credo religione o sesso. La Corte Suprema ha invece respinto questa accusa in quanto inammissibile per insufficiente chiarezza nella formulazione della norma. La corte suprema afferma che l'ordinanza della città di Saint Paul è incompatibile con il primo emendamento; è vero che la condotta incriminata ricadeva nell'ambito delle *fighting words*, ma ciò non vuol dire che risulti per forza fuori dalla protezione del primo emendamento. Per limitare una particolare manifestazione di pensiero deve sussistere un motivo diverso dal mero contenuto.

messaggio viene veicolato, ovvero il modo di comunicare, è identificato come elemento chiave per decidere se si è o meno dinanzi ad una formula offensiva e quindi se c'è o meno la punibilità. Da questa sentenza dunque emerge l'importanza attribuita al mezzo di comunicazione utilizzato, infatti la Corte Suprema specifica che per limitare una particolare manifestazione di opinione, non ci si può affidare al mero contenuto del discorso, ma deve sussistere qualcosa di più. Per poter essere regolamentata l'*hate speech* abbisogna poi dell'analisi del contenuto dei messaggi e, come abbiamo già precedentemente analizzato, sono inclusi nella categoria i messaggi che hanno ad oggetto genere, razza ed etnia, disabilità e religione. Per quanto riguarda la terza componente, ovvero il destinatario del messaggio, è opportuno verificare se sia presente un destinatario diretto o meno. Affinché si tratti di *hate speech* deve sussistere la presenza di uno specifico destinatario, non possono infatti considerarsi appartenenti a questa categoria quei discorsi che hanno avuto luogo in pubblico: essi raggiungono un pubblico numericamente più elevato ma, esso, non essendo direttamente interessato a recepire il messaggio, può prenderne atto o meno. Questo è un po' ciò che accade sul web, infatti una mail contenente messaggi d'odio deve essere considerata diversamente rispetto allo stesso messaggio d'odio postato su un sito web. Altro ruolo fondamentale per la comprensione dell'*hate speech* è da attribuire allo scopo e al contesto, infatti prima di poter affermare o meno se siamo dinanzi ad un caso di *hate speech* dobbiamo contestualizzare la situazione. Per questa ragione solitamente vengono esclusi da suddetta fattispecie i discorsi rientranti nella categoria della satira, così come commedie e parodie nonostante a volte contengano contenuti discriminatori. Infatti lo stesso messaggio veicolato da un fumettista satirico può facilmente divenire perseguibile come *hate speech* se pronunciato in un qualsiasi altro luogo. Tuttavia anche per quanto riguarda la satira, che troviamo protetta dall'ONU, dalla carta di Nizza ed anche dalla CEDU, esistono dei limiti non sempre facilmente interpretabili. Negli Stati Uniti ad esempio molti gesti che da noi potrebbero risultare offensivi o razzisti sono accettati in quanto tutelati dal primo Emendamento della Costituzione, a meno che non si identifichi nell'atto un intento intimidatorio. Infine per quanto concerne gli effetti, il bisogno di limitare la libertà d'espressione e condannare l'*hate speech* si rileva sulla considerazione dei suoi effetti negativi che si riversano sia sui destinatari diretti che sulla società.

Le cinque componenti esposte sopra, rappresentano un sostegno nella definizione di *hate speech*. Non bisogna considerarle indipendenti tra loro, posso infatti intrecciarsi e sovrapporsi, ma non saranno sempre onnipresenti nei casi di *hate speech*, di conseguenza ogniquale volta verrà applicata una legislazione differente a seconda di quale prevalga.⁹ A tale riguardo infatti, si pensa che ci sia il

⁹<http://sosracismo.eu/wp-content/uploads/2016/07/Hate-Crime-and-Hate-Speech-in-Europe.-Comprehensive-Analysis-of-International-Law-Principles-EU-wide-Study-and-National-Assessments.pdf>

bisogno di intervenire sulle leggi dell'odio online per quanto riguarda il mezzo internet, per ampliarle e renderle più complete. Bisognerebbe ovvero, migliorare la legislazione su tale tema.

1.3. L'avvento del cyberspazio: odio online ed offline

Non si può parlare di odio online senza analizzare come si sono evolute negli anni le modalità di produzione di informazione nella società tecnologica, in particolar modo in relazione ai social network. Lo studioso Yochai Benkler¹⁰ suddivide l'ambiente elettronico in due categorie; la prima dove troviamo i produttori di informazione professionali, la seconda invece, più recente, riguarda la produzione di contenuti più informale, e comprende gli utenti, gli individui, che grazie alle nuove tecnologie hanno potuto ideare informazione a basso costo e hanno avuto la possibilità di diffonderla su larga scala ottenendo un'audience non indifferente. Il nodo della questione sta nel fatto che la prima *sfera pubblica*¹¹ era controllata dai media tradizionali, e perciò era più semplice supervisionare ciò che era veicolato; adesso invece la sfera pubblica vede sempre come protagonisti i mass media, ma il controllo non è più di loro esclusiva competenza, è infatti internet, con l'economia dell'informazione condivisa, ad essersi inserito nella democrazia stessa. Nella modernità inoltre sono i grandi fornitori quali Facebook, Twitter e Google a diventare parte in causa diretta nella decisione di quali contenuti rimuovere, ma a questo dedicheremo un capitolo più avanti. Quello che è bene tenere a mente quando si parla di odio online è che l'odio che si è creato non è virtuale, ma reale, ovvero gli effetti sulle persone sono uguali a quelli conseguenti ad aggressioni fisiche e reali. Infatti se un messaggio contenente incitazione all'odio riesce a risultare ben fatto e diventare virale, non è importante quanto un ipotetico soggetto sia distante dalla diffamazione, il pericolo sarà ugualmente molto elevato, nonché reale e concreto, data la globalità di internet.

Sul finire degli anni Novanta iniziarono ad essere pubblicati, in America, i primi studi sul problema *dell'hate speech online*, sui diritti civili nel cyberspazio e sulle nuove sfide poste alla società da internet. La novità principale consisteva nella possibilità di comunicazione e nell'accesso all'informazione in tempo reale e ad un basso costo. Queste innovazioni però avevano un rovescio della medaglia, ovvero facilitare il compito a chi intendeva utilizzare internet con il fine di far circolare *hate speech* e di incitare alla violenza. Allo stesso tempo si palesava il fatto che il tradizionale concetto di giurisdizione, con i limiti fisici che ne conseguivano, era destinato a decadere, vista la fluidità con la quale i contenuti viaggiavano attraverso la rete. Ci si domandava

¹⁰ Y. BENKLER, *La ricchezza della rete*, Milano, 2007.

¹¹ *Ibidem*.

come la tutela garantita dal primo Emendamento si potesse estendere anche al web e alla presenza di espressioni d'odio online. D'altronde l'odio, nella sua versione online, non cambiava forma rispetto alla sua tradizionale nozione, ma internet offriva per la prima volta la possibilità di creare messaggi (d'odio) e trasmetterli senza alcuno sforzo, riuscendo a portarli su larga scala. In questi stessi anni alcune organizzazioni cercarono di offrire contromisure per tracciare l'odio online e segnalare eventuali siti web contenenti espressioni violente, ma la difficoltà nel riuscirci erano abbastanza evidenti. Nonostante gli interventi dei provider¹², che tentarono di trovare un espediente per il contenimento dell'odio online. Agli inizi del nuovo millennio gli studiosi iniziarono a controllare con maggiore attenzione il contenuto dei siti web estremisti, cosa che ai tempi era più semplice data la quantità decisamente inferiore rispetto ad oggi.

Secondo un report dell'UNESCO del 2015 intitolato *Countering Online Hate Speech*¹³, sarebbero principalmente quattro le differenze tra l'hate speech online e l'hate speech per così dire offline¹⁴, innanzitutto, internet e le varie piattaforme in esso presenti consentono la permanenza dell'odio online facendo sì che il messaggio veicolato rimanga in circolo per molto tempo; più esso rimane in circolo e maggiore sarà il danno che può infliggere sulle potenziali vittime. L'architettura che caratterizza le diverse piattaforme può permettere ai messaggi di rimanere in circolazione per periodi più o meno lunghi; in social quali Twitter è molto facile dare maggiore visibilità ai messaggi d'odio, grazie ai cosiddetti *trending topics*. La seconda differenza è il *ritorno imprevedibile* dell'odio: il messaggio d'odio può infatti facilmente “tornare in vita” anche se cancellato, sfruttando un'altra piattaforma, mediante la nuova diffusione da parte di un altro utente. La terza differenza enfatizza l'importanza che può avere nel mondo online *l'anonimato*, garantito dall'utilizzo di pseudonimi o nomi fittizi. Infatti il credersi anonimi fa sentire gli utenti più a proprio agio nell'esprimere odio; tale meccanismo riduce l'esposizione al rischio di essere scoperti, e di conseguenza di incorrere in cause legali. La quarta differenza è la *trans nazionalità*, che incrementa l'effetto dell'*hate speech* e rende più complicato individuare degli espedienti legali per combatterlo. Infatti, quando si utilizzano i nuovi media per trasmettere un messaggio d'odio, è molto probabile che vengano coinvolti oltre i destinatari del messaggio anche altri attori in modo consapevole o meno; inoltre per via della *trans nazionalità* della rete è possibile aggirare le regole di giurisdizioni che i Paesi si sono prefissati o, ancor di più, è possibile riavviare l'attività in Stati accuratamente scelti affinché la libertà d'espressione abbia maggiori garanzie.

¹² “Internet Service Provider (ISP), soggetti che forniscono servizi di connessione, trasmissione, memorizzazione dati, anche attraverso la messa a disposizione delle proprie apparecchiature per ospitare siti. È essenzialmente un intermediario, che stabilisce un collegamento tra chi intende comunicare un'informazione e i destinatari della stessa.”

M. MENSI – P. FALLETTA, *Il diritto del web. Casi e materiali*, Padova, 2015

¹³ <http://unesdoc.unesco.org/images/0023/002332/233231e.pdf>

¹⁴ G. ZICCARDI, *op. cit.*, pag. 78.

Malgrado gli elementi cui ci si riferisce con espressioni d'odio siano pressoché gli stessi (etnia, nazionalità con una crescita di offese contro religione e classi sociali), l'odio online e l'odio offline si distinguono per tre rilevanti aspetti: le misure ideate per controbattere l'odio tradizionale si dimostrano spesso inefficaci per combattere l'odio nella rete; c'è inoltre una notevole differenza tra un messaggio trasmesso online ed uno offline, hanno infatti un diverso impatto sociale; c'è poi un'ulteriore problematica per quanto concerne la riuscita nell'applicare la legge ad un fenomeno reso così fluido dalla capacità di internet di <<esistere>> contemporaneamente in più Paesi del mondo. L'UNESCO individua gli aspetti più problematici dell'odio online, concentreremo la nostra analisi su due in particolare, l'idea di *permanence* e quella di *itinerancy*. Nel primo caso come accennato in precedenza, più il materiale "odioso" resta in rete e più è facile che la vittima si senta minacciata: questo può dare molto potere a chi ha inserito l'hate speech online. Allo stesso tempo si potrebbero limitare i danni se si riuscisse ad eliminare il contenuto anche in una fase avanzata. Inoltre le differenze che caratterizzano le varie piattaforme online possono far sì che l'argomento dell'espressione d'odio rimanga vivo più o meno a lungo. Ad Esempio su Twitter è più semplice e veloce la circolazione di messaggi d'odio grazie ai *trending topics* ma, allo stesso tempo questo meccanismo permette a tali messaggi di essere più evidenti e di conseguenza eliminabili. Al contrario Facebook dà vita a conversazioni multiple parallelamente, con miriadi di commenti che non sono facilmente individuabili (a meno che non siano segnalati), facilitando la creazione di spazi dove gli aggressori possono offendere, aggredire e discriminare i soggetti presi di mira. Il secondo aspetto, l'*itinerancy*, sottolinea il carattere itinerante dell'hate speech online, infatti anche nel caso in cui il messaggio sia stato rimosso è sempre possibile che questo riemerga, attraverso un'altra piattaforma, in spazi online differenti o semplicemente con una nuova denominazione. Se un sito web viene oscurato, ad esempio, può sempre essere reso di nuovo attivo grazie ad un servizio di web hosting¹⁵ con politiche meno serrate, o "appoggiandosi" ad un Paese con meno rigorose leggi sull'hate speech. In definitiva siamo dinanzi ad un fenomeno in crescita, il cui contrasto richiede un impegno collettivo. Attuare delle misure repressive generiche risulterebbe inefficace considerando la natura di internet, la quale richiede invece degli interventi mirati; bisognerebbe permettere ai fruitori di tali servizi di intervenire attraverso delle segnalazioni, che richiederanno però una ponderazione con il diritto alla libertà d'espressione.

Per reagire a tale fenomeno sono state consigliate alcune strategie; prima fra tutte l'educazione, con il conseguente aumento di consapevolezza riguardo le conversazioni online. Bisognerebbe poi

¹⁵ Hosting: prestatori di servizi di memorizzazione d'informazione.

delineare con precisione la responsabilità degli opinion makers¹⁶, di figure politiche e istituzionali e dei mass media professionali, che di sovente sono i primi a veicolare espressioni d'odio a fini elettorali o di incremento di audience (Basti pensare per esempio ad un personaggio quale Matteo Salvini, che ha un fortissimo seguito sui social, e li utilizza spesso con fini razzisti e di propaganda dell' hate speech). Sarebbe poi opportuno attivarsi attraverso la legge per considerare l' hate speech un vero e proprio crimine in tutti gli Stati, così da coordinare nel migliore dei modi le azioni d'intervento; discutere riguardo la responsabilità dei contenuti d'odio, e sulla necessità di una maggiore trasparenza per quanto riguarda la gestione delle informazioni da parte delle community online e dei social media in genere. Potrebbe infine essere interessante sviluppare << misure tecniche utili a gestire il traffico di odio in rete: si pensi a sistemi di filtraggio di quei contenuti individuati come estremi o, addirittura, alla possibilità di influenzare i discorsi in rete, e la loro visibilità o meno, tramite algoritmi >>¹⁷. Cercheremo per l'appunto adesso di analizzare come la giurisprudenza nazionale ed europea si muova all'interno di questo delicato argomento e quali soluzioni ipotizzi per contrastare l' hate speech.

¹⁶ È noto che all'interno della rete, le opinioni si formano in pochi social media, i quali, più importanti sono gestiti dai cosiddetti *opinion makers*, soggetti in grado di influenzare l'opinione pubblica, consolidando l'ostilità tra opposti schieramenti. Rischiando di giungere alla cosiddetta *polarizzazione di gruppo*.

¹⁷ Come afferma G. ZICCARDI, in *L'odio online. Violenza verbale e ossessione in rete*, Milano, 2016

Capitolo 2: Contrasto all'*hate speech*

2.1. Il contrasto all'*hate speech* nell'Ordinamento Italiano

Prima di affrontare la questione del contrasto all'*hate speech* all'interno del nostro ordinamento è opportuno considerare l'importanza che è attribuita in Italia così come in ogni stato con base democratica alla libertà d'espressione, diritto che viene posto allo stello livello delle libertà costituzionali. Troveremo infatti non poche difficoltà nel contrastare l'odio online e ciò probabilmente deriva dall'esperienza totalitaria che ha vissuto l'Italia nel periodo precedente alla nascita della Costituzione, dove ogni forma di libertà era assente nella sua quasi totalità. Per questo motivo si è cercato di lì in poi di tutelare in modo stringente la libertà di espressione¹⁸, per evitare il ripetersi di un simile accaduto¹⁹. Questa quindi è la ragione per la quale risulta così complicato regolare un mezzo quale è internet; soprattutto dato che in un primo momento si cercherà di regolamentarizzarlo ricorrendo alle norme utilizzate per il resto dei mezzi di diffusione; ma ciò non sarà possibile, infatti come affermato nella sentenza del 6 giugno 2000 n. 137, internet rappresenta un mezzo di informazione <<del tutto peculiare, al quale, vertendo in materia penale, non può essere estesa in via analogica la disciplina dettata per la stampa o per la radio o la televisione>>²⁰.

Dopo questo breve premissa, arriviamo adesso ad analizzare la normativa interna al nostro Paese in materia di *hate speech*. La prima normativa alla quale si fa riferimento è la legge del 13 ottobre 1975, n.654²¹, la quale ha immesso <<il reato di istigazione alla discriminazione o alla violenza nei confronti di determinati gruppi nazionali, etnici o razziali>>²². Una prima modifica²³ a suddetto articolo si ebbe in seguito ad un periodo di manifestazioni di forte intolleranza, con il decreto legge n. 122/1993, convertito poi nella cosiddetta legge Mancino²⁴ con la quale ci fu un estensione delle norma, ovvero, venivano puniti oltre a coloro che incitavano alla discriminazione, anche chi incitava a commettere o commetteva atti di discriminazione, così da poter punire ogni comportamento che aveva quale fine quello di portare a fenomeni discriminatori. In secondo luogo

¹⁸ Art.21 della costituzione: << Tutti hanno diritto di manifestare liberamente il proprio pensiero con la parola, lo scritto e ogni altro mezzo di diffusione.>>.

¹⁹ G. GARDINI, *Le regole dell'informazione. Principi giuridici, strumenti, casi*. Torino 2009.

²⁰ Giudice dell'Udienza Preliminare presso il Tribunale di Oristano, Sentenza 25 maggio 2000 (dep. 6 giugno 2000), n. 137, Est. Tuveri.

²¹ Legge n. 654 del 1975, con la quale si ratifica la Convenzione internazionale dell'ONU sull'eliminazione di ogni forma di discriminazione razziale conclusa a New York il 21 dicembre 1965.

²² M. MENSI – P. FALLETTA, *Il diritto del web. Casi e materiali*, Padova, 2015, p.192

²³ L'Articolo 3 della Legge n. 654/1975, emendata dalla Legge n. 205/1993 articolo 1: Salvo che il fatto costituisca più grave reato, anche ai fini dell'attuazione della disposizione dell'articolo 4 della convenzione, è punito: a) con la reclusione sino a tre anni chi diffonde in qualsiasi modo idee fondate sulla superiorità o sull'odio razziale o etnico, ovvero incita a commettere o commette atti di discriminazione per motivi razziali, etnici, nazionali o religiosi; b) con la reclusione da sei mesi a quattro anni chi, in qualsiasi modo, incita a commettere o commette violenza o atti di provocazione alla violenza per motivi razziali, etnici, nazionali o religiosi.

²⁴ Legge n. 205 del 25 giugno del 1993.

venne poi aggiunto come fenomeno sanzionatorio anche la discriminazione per causa religiosa, da aggiungere a quella per razza, etnia e nazionalità già esistenti. In poche parole, mentre con la legge del 13 ottobre 1975 n.654 si tutelava genericamente l'ordine pubblico, ora con la legge Mancino viene messa sotto tutela la dignità umana. Il problema che però si venne profilando riguardava il fatto che questa legge rischiava di ledere il principio di libertà d'espressione, tant'è che si fece un'ultima modifica da apportare all'articolo 13 con la legge 24 febbraio 2006 n.85, dove l'azione di <<diffondere>> venne sostituita con quella di <<propagandare>> termine che ha sicuramente un senso più specifico, così da rendere alla legge un significato con una più stretta interpretazione e meno difficilmente equivocabile. Così facendo il legislatore cercò di determinare in modo più specifico l'area dei comportamenti perseguibili, restringendola ai pensieri fattivi e non più alla sola manifestazione del pensiero, riducendo le pene per i reati di propaganda di idee o teorie di superiorità di una razza o gruppo di persone, e di istigazione alla discriminazione razziale o religiosa. Con sanzioni detentive meno punitive si è compiuto un passo in avanti, consentendo una maggiore libertà d'espressione che ha portato ad una maggiore circolazione delle idee a vantaggio della democrazia.

In tempi recenti si era tornati sulla legge Mancino per discutere²⁵ l'eventualità di aggiungere un'altra fattispecie di *hate speech*, in base all'orientamento sessuale. Il nuovo decreto legge all'articolo 1, quindi estende la perseguibilità degli atti discriminatori compiuti a causa dell'orientamento sessuale o dell'identità di genere, sancendo così l'equivalenza tra le discriminazioni causate da motivi razziali e quelle causate dall'identità di genere o dall'orientamento sessuale delle persone. Inoltre il disegno legge di Scalfarotto²⁶, tuttora in corso di esame alla commissione, estenderebbe l'articolo 3 della legge Mancino inserendo altre due fattispecie tra le condotte di istigazione e violenza, ovvero quelle per omofobia e trans fobia; questa modifica qualora venisse approvato, rischierebbe di compromettere la libertà d'espressione ma fortunatamente la legge Mancino è molto precisa difatti punisce l'istigazione a commettere una discriminazione o una violenza, e non delle semplici manifestazioni di opinioni, anche quando esprimano un pregiudizio, per cui non dovrebbe andare ad intaccare troppo suddetta libertà.

Un altro degli strumenti utilizzati dalla giurisprudenza italiana per contrastare il fenomeno dell'*hate speech* è il codice penale; difatti agli articoli 594 e 595 vengono riconosciuti il reato di <<ingiuria>> e il reato di <<diffamazione>>. Per quanto concerne l'articolo 594 al primo comma è punito <<Chiunque offende l'onore o il decoro di una persona presente>> con un aggravante nel caso in cui << l'offesa consiste nell'attribuzione di un fatto determinato>> o << qualora l'offesa sia commessa in presenza

²⁵ Disegno di legge n. 842, Norme in materia di sensibilizzazione e repressione della discriminazione razziale, per l'orientamento sessuale e l'identità di genere. Modifiche alla legge 13 ottobre 1975, n. 654.

²⁶ Disegno di legge S. 1052.

di più persone>>; con l'articolo 595 poi è considerato perseguibile << Chiunque, fuori dei casi indicati nell'articolo precedente, comunicando con più persone offende l'altrui reputazione>> anche in questo caso con un inasprimento della pena nel caso in cui << l'offesa consiste nell'attribuzione di un fatto determinato>> oppure << se l'offesa è recata col mezzo della stampa o con qualsiasi altro mezzo di pubblicità, ovvero in atto pubblico [...] Se l'offesa è recata a un Corpo politico, amministrativo o giudiziario, o ad una sua rappresentanza, o ad una Autorità costituita in collegio.>>²⁷. C'è quindi una mancanza per quanto riguarda la diffamazione consumata attraverso il mezzo internet, nonostante sia ormai il più potente strumento di diffusione dell'informazione²⁸. A tal proposito la Suprema Corte aveva specificato che il legislatore, pur tenendo presente l'esistenza di nuovi strumenti di comunicazione, non ha sentito il bisogno di modificare o integrare la legge in merito. Evidenziando però la profonda differenza presente tra la diffusività e la pervasività di internet rispetto a quelle delle trasmissioni radio televisive la Corte afferma che sarà presente un più severo trattamento penale; difatti precisa la <<diffamazione tramite internet costituisce un'ipotesi di diffamazione aggravata ai sensi dell'articolo 595, comma 3, c.p., in quanto commessa con altro mezzo di pubblicità [rispetto alla stampa]>> dato che <<essendo internet un potente mezzo di diffusione di notizie, immagini ed idee, anche attraverso tale strumento di comunicazione si estrinseca il diritto di esprimere le proprie opinioni, tutelato dall'articolo 21 della costituzione che, per essere legittimo, deve essere esercitato rispettando le condizioni e i limiti dei diritti di cronaca e di critica>>²⁹. In breve, per la Suprema Corte, internet è considerato il mezzo di comunicazione più democratico, essendo accessibile quasi a tutti con un costo non troppo oneroso e permettendo un'ampia fruibilità dei contenuti immessi in qualsivoglia parte del mondo. Proprio grazie alla sua distintiva diffusività e al suo carattere transnazionale, su internet diviene sempre più difficile riuscire a bloccare qualsiasi tipo di *hate speech*, se non facendo uso di appositi mezzi coercitivi, legalmente riservati alla pubblica autorità, che oltretutto spesso non soddisfano per via di strumenti tecnici inadeguati; come infatti sottolinea la Corte, le procedure legali hanno spesso una lunga tempistica, rispetto alla velocità alla quale il <<messaggio odioso>> si propaga.

La Corte Suprema tende poi a sottolineare che in caso di diffamazione si è in presenza di un evento non fisico ma psicologico, che consiste nella percezione da parte di terzi dell'eventuale *hate speech*. Per la Corte quindi il reato si ritiene consumato non al momento dell'effettiva diffusione, ma dal momento che tale messaggio viene percepito da soggetti terzi rispetto all'offeso e all'offensore. Proprio per questo la Corte distingue in modo netto il momento in cui il messaggio d'odio viene inserito online, da quello che si verifica successivamente, in cui utenti terzi connettendosi prendono

²⁷ <http://www.diritto.it/codici/2>

²⁸ M. MENSI – P. FALLETTA, *op. cit.*, p.157

²⁹ Cass. Pen., sez. V, 1 luglio 2008 n. 31392.

visione dell'accaduto, verificandone l'atto. Proprio per questo motivo la Suprema Corte ritiene che il tempo ed il luogo in cui viene consumato l'*hate speech* non corrispondono con il momento e la sede del server dal quale vengono immessi, ma come si deduce da quanto detto sopra la diffamazione deve ritenersi compiuta nel tempo e nel luogo in cui le offese sono percepiti da terzi nella rete, anche nel caso in cui il sito web sia registrato in uno Stato estero.

Negli ultimi anni si è cercato di contrastare in modo sempre più stringente il fenomeno dell'*hate speech* sul web, e si è raggiunto un importante risultato con la nascita nel luglio del 2015 della Carta Italiana dei Diritti in Internet³⁰, voluta dal Presidente della Camera Laura Boldrini. Questa Carta è stata istituita con lo scopo di immettere una serie di valori e principi nella migliore fruibilità di internet; nell'articolo 13 è trattata la sicurezza in rete dove si afferma che <<non sono ammesse limitazioni della libertà di manifestazione del pensiero>>, ma allo stesso tempo <<deve essere garantita la tutela della dignità delle persone da abusi connessi a comportamenti quali l'incitamento all'odio, alla discriminazione e alla violenza>>. Dunque come abbiamo avuto modo di notare, la disciplina italiana in merito è particolarmente complessa ed articolata e di conseguenza non sempre di facile interpretazione e applicazione.

Per quanto concerne le denunce in caso di atti discriminatori, in Italia ci si riferisce di norma a due organismi, l'Osservatorio per la Sicurezza contro gli Atti Discriminatori (OSCAD) e l'Ufficio Nazionale Antidiscriminazioni Razziali (UNAR)³¹. Il primo nasce da una collaborazione tra Polizia di Stato e Carabinieri nel 2010, con il fine di incrementare la protezione contro le discriminazioni per tutti coloro che sono vittime di *hate speech*, e per promuovere il diritto all'uguaglianza davanti alla legge. L'OSCAD riceve segnalazioni sia da singoli privati, che da istituzioni ed associazioni, e svolge un ruolo di vigilanza per quanto concerne le discriminazioni per razza, nazionalità, credo religioso, genere, età, lingua, disabilità fisica o mentale, orientamento sessuale ed identità di genere. Un'altra delle sue mansioni è quella di attivare, in seguito alle segnalazioni, delle procedure d'intervento, mobilitando le forze di polizia, e seguendo l'evoluzione degli atti discriminatori denunciati, restando in contatto con gli organismi di tutela per elaborare misure di contrasto e prevenzione ai crimini d'odio. Altri strumenti per la segnalazione di eventuali atti discriminatori sono lasciati a disposizione dell'UNAR, istituto nato in Italia con il decreto legislativo 9 luglio 2003 n. 215, con lo scopo di promuovere la parità di trattamento e di vigilare sull'effettiva applicazione degli strumenti di tutela esistenti contro le discriminazioni; al fine di rimuovere le

³⁰ La Mozione *Quintarelli ed altri*, n. 1-01031, mozione concernente iniziative per la promozione di una carta dei diritti in internet e per la governance della rete, passata alla camera il 3 novembre 2015 all'unanimità.

³¹ http://www.cospe.org/wp-content/uploads/2016/03/ricerca_odionon%C3%A8opinione.pdf

discriminazioni fondate sulla razza o sull'origine etnica. L'UNAR attraverso un *contact centre*, riceve segnalazioni, denunce e testimonianze con l'impegno di fornire assistenza alle vittime durante i procedimenti in sede sia amministrativa che giudiziaria. È stato stipulato un protocollo d'intesa tra il Ministero dell'Interno ed il Dipartimento delle pari opportunità della Presidenza del Consiglio dei ministri, dove si stabilisce che la Polizia debba segnalare al *contact centre* dell'UNAR i casi di discriminazione che sono sotto la tutela civile e non costituiscono reati mentre l'UNAR è obbligato a informare la Polizia dei crimini d'odio e degli atti di razzismo penalmente rilevanti. L'UNAR e l'OSCAD riportano inoltre alla Polizia Postale tutte le discriminazioni che hanno a che fare con atti discriminatori o razzisti compiuti sul web. È inoltre presente una quantità non indifferente di Organizzazioni Non Governative (ONG) che insieme ad associazioni svolgono un'importante ruolo nella prevenzione e nel contrasto all'*'hate speech* e che sono inserite nell'elenco degli enti legittimati ad agire in giudizio a tutela delle vittime di discriminazioni razziali.

2.2. La normativa dell'Unione Europea in materia di *'hate speech*

L'Unione Europea si è molto prodigata per cercare di arginare il fenomeno dell'*'hate speech*, attuando diversi interventi e risoluzioni approvate dal Parlamento Europeo, che incitano gli stati membri a compiere delle azioni concrete per contrastare il fenomeno dell'odio online. I principali elementi in materia di non discriminazione sono l'articolo 13 del Trattato di Amsterdam, l'articolo 21 della Carta di Nizza ed infine le Direttive 2000/43/CE e 2000/78/CE.

Il principio di non discriminazione fa parte dei principi fondamentali dell'Unione Europea ed ha ottenuto il riconoscimento formale nell'Unione Europea con il Trattato di Amsterdam, il quale è andato a modificare l'articolo 13 del trattato sulla Comunità Europea, oggi articolo 19 del TFUE. Dopo tale revisione la tutela si è ampliata arrivando a ricomprendere l'insieme di tutti i diritti civili politici e sociali, e non più per i soli motivi di discriminazione per sesso o nazionalità. Difatti l'articolo 19 del TFUE esordisce con: << Fatte salve le altre disposizioni dei trattati e nell'ambito delle competenze da essi conferite all'Unione, il Consiglio, deliberando all'unanimità secondo una procedura legislativa speciale e previa approvazione del Parlamento europeo, può prendere i provvedimenti opportuni per combattere le discriminazioni fondate sul sesso, la razza o l'origine etnica, la religione o le convinzioni personali, la disabilità, l'età o l'orientamento sessuale.>>.

Troviamo poi, sempre a tutela di tali finalità, la Carta dei Diritti fondamentali dell'Unione Europea³² (Carta di Nizza) che nell'articolo 21 afferma che: << E' vietata qualsiasi forma di

³² La Carta dei Diritti fondamentali dell'Unione Europea, anche rinominata Carta di Nizza, è stata proclamata una prima volta a Nizza il 7 dicembre 2000 ed una seconda volta, in una versione adattata, a Strasburgo il 12 dicembre 2007. La Carta è stata ideata per raccogliere in un testo organico i diritti così come risultano dalle tradizioni

discriminazione fondata, in particolare, sul sesso, la razza, il colore della pelle o l'origine etnica o sociale, le caratteristiche genetiche, la lingua, la religione o le convinzioni personali, le opinioni politiche o di qualsiasi altra natura, l'appartenenza ad una minoranza nazionale, il patrimonio, la nascita, gli handicap, l'età o le tendenze sessuali. Nell'ambito d'applicazione del trattato che istituisce la Comunità europea e del trattato sull'Unione europea è vietata qualsiasi discriminazione fondata sulla cittadinanza, fatte salve le disposizioni particolari contenute nei trattati stessi.>>.

Trattando l'ambito della normativa dell'Unione Europea non si possono non citare le Direttive 2000/43/CE e 2000/78/CE emanate dal Consiglio dell'Unione Europea. La direttiva 200/43/CE, la quale impedisce ogni qualsivoglia forma di discriminazione per motivi di razza, origine etnica, sia in ambito privato che pubblico; e la direttiva 200/78/CE; la quale amplia il divieto di discriminazione a fattori legati alla religione, all'età, all'orientamento sessuale, e per quanto concerne l'occupazione e le condizioni di lavoro. In conclusione, queste due direttive lette in combinato danno una definizione del principio di non discriminazione, che può essere sovrapposta a quello di eguaglianza. Tali Direttive mirano a dar vita ad un quadro normativo generale, dove gli stati, rimanendo nel proprio ordinamento, possono partecipare attivamente alla lotta contro la discriminazione su base razziale ed etnica, allo scopo finale di dare piena effettività al principio di parità di trattamento ed assicurare la crescita di società democratiche.³³

Infine è bene soffermarsi un momento sull'impegno della Commissione Europea che oltre a promuovere dal punto di vista finanziario politiche per contrastare l'*hate speech*, si occupa di monitorare attraverso dei report annuali lo sviluppo della Carta dei Diritti fondamentali dell'Unione; ha inoltre realizzato una strategia per un'attuazione effettiva della Carta con il fine di garantire la correttezza delle azioni dell'Unione a sostegno dei diritti fondamentali. Si è inoltre servita di questi report per raccomandare agli Stati membri l'inserimento del divieto dell'*hate speech* nella legislazione penale interna, con il fine di permettere ai cittadini di trarre beneficio di specifiche norme di contrasto del fenomeno, che potranno essere applicate dai tribunali.

Tutto ciò ci dimostra come l'Unione Europea si sia dotata con il tempo di una consolidata base giuridica per quanto riguarda la lotta contro la discriminazione in ogni sua sfaccettatura; andremo adesso ad analizzare come le organizzazioni internazionali si sono mosse nell'azione di disciplina e contrasto all'*hate speech*.

costituzionali, dagli obblighi internazionali comuni dei paesi dell'UE, dalla Convenzione europea per la salvaguardia dei diritti dell'uomo e delle libertà fondamentali, dalle Carte sociali adottate dall'UE e dal Consiglio d'Europa e dalla giurisprudenza della Corte di giustizia dell'Unione europea e della Corte europea dei diritti dell'uomo. Così da diventare <<certezza del diritto>> dell'Unione Europea, grazie alla sua chiarezza.

³³ M. MENSI – P. FALLETTA, *op. cit.*, p.176

2.3. Il consiglio d'Europa

Il consiglio d'Europa ha dimostrato una particolare attenzione per la discriminazione sul web, ha infatti affermato che << considers racism not as an opinion but as a crime³⁴>>; oltretutto l'azione del consiglio è di più ampia portata infatti al suo interno troviamo disciplinata non solo la lotta al discorso razzista, ma ad ogni forma di *hate speech*, che per il Consiglio si estende ad ogni forma di discorso d'odio, su base razziale, di origine etnica, religiosa, ed anche sociale. A tal proposito troviamo la Carta Sociale Europea³⁵ e la Convenzione Quadro per le Protezione delle Minoranze Nazionali³⁶, le quali hanno il fine di proteggere da ogni forma di discriminazione. Per quanto riguarda la prima essa vieta la discriminazione in tutte le sue manifestazioni per questioni di razza, colore, religione o provenienza nazionale; mentre la seconda proibisce ogni qualsiasi forma di discriminazione che si fondi sull'appartenenza di un individuo ad una minoranza nazionale, facendo impegnare gli Stati firmatari ad adottare di conseguenza le misure adeguate per promuovere la piena uguaglianza tra le persone.

In un momento successivo rispetto ai citati interventi sono poi stati adottati due accordi internazionali che potrebbero risultare molto utili nel risolvere la lotta ai crimini informatici, nello specifico ai comportamenti discriminatori sul web. In primis la Convenzione del Consiglio d'Europa sulla criminalità informatica, entrata in vigore il 1 luglio 2004, la quale ha come obbiettivo quello di dar vita ad efficaci strumenti di lotta al *cybercrime* e di rendere più chiara la normativa dei Paesi aderenti, anche non facenti parte del consiglio d'Europa. La convenzione delinea una guida per gli Stati che intendono sviluppare una legislazione nazionale esaustiva contro la criminalità informatica e fornisce anche un quadro per cooperare a livello internazionale in questo campo. In secondo luogo, l'approvazione del Protocollo addizionale alla suddetta convenzione per quanto riguarda l'incriminazione di atti aventi natura razzista e xenofobica attuati attraverso sistemi informatici. Il documento si prefigge due obbiettivi principali; quello di accordare il diritto penale degli Stati aderenti e quello di migliorare la cooperazione internazionale tra gli Stati al fine di una più efficace lotta contro il razzismo e la xenofobia in internet.

Il Consiglio d'Europa ha predisposto una Commissione Europea contro il razzismo e l'intolleranza (ECRI) operante a partire dagli anni novanta, il cui obbiettivo fondamentale è combattere il razzismo e la discriminazione razziale in Europa, difendendo i diritti fondamentali degli individui; ha inoltre il compito di monitorare attraverso la pubblicazione periodica di report le situazioni

³⁴ <http://www.assembly.coe.int/nw/xml/XRef/X2H-Xref-ViewHTML.asp?FileID=9903&lang=en>

³⁵ La Carta sociale europea è un trattato del Consiglio d'Europa, adottato a Torino nel 1961 e rivisto a Strasburgo nel 1996. La Carta riconosce i diritti umani e le libertà e stabilisce un meccanismo di controllo per garantirne il rispetto da parte degli Stati.

³⁶ Convenzione internazionale stabilita nel quadro del Consiglio d'Europa al fine di garantire la tutela delle minoranze. Entrata in vigore nel 1998.

all'interno dei singoli Stati. Un esempio tangibile del lavoro di questa Commissione lo troviamo nel rapporto sull'Italia pubblicato da quest'ultimo il 21 febbraio 2012³⁷, all'interno del quale si evidenziava il bisogno di adottare misure allo scopo di rafforzare il ruolo dell'UNAR e, di adottare le misure necessarie per proteggere i migranti, con un particolare riguardo al principio di non respingimento, ed infine si raccomandava alle autorità italiane di rendere più intensi gli sforzi per contrastare la diffusione di materiale di propaganda razzista, xenofoba e antisemita via Internet.

2.5. La giurisprudenza della Corte Europea dei Diritti dell'Uomo (CEDU)

La Corte Europea dei Diritti dell'Uomo si è occupata per più volte di trovare una risposta alla questione dell'*hate speech*, e lo ha fatto applicando il principio di non discriminazione in modo non assoluto, bensì bilanciandolo con gli altri diritti presenti nella stessa CEDU. Françoise Tulkens, giudice e vicepresidente della Corte Europea dei Diritti dell'Uomo nel 2015, ha tentato di fare ordine in un campo pieno di difficoltà interpretative, delineando l'approccio dei giudici europei in materia di *hate speech*³⁸. La prima cosa sulla quale si è soffermata è la profonda differenza tra l'azione della Corte Europea, segnata dal ricordo della seconda guerra mondiale, con tutto ciò che ne consegue; e la Corte Suprema degli Stati Uniti d'America, dove è molto forte la libertà d'espressione come si deduce dal primo emendamento della loro costituzione.

Quando la Corte si trova dinanzi al tema dell'incitamento all'odio e deve pesarlo in relazione alla libertà d'espressione, solitamente utilizza due approcci³⁹: uno di tipo più ampio, con il quale vengono escluse determinate espressioni dalla protezione dell'articolo 17, articolo per il divieto all'abuso di diritto⁴⁰, grazie al quale è tutelata la libertà di manifestazione del pensiero ma con un'attenzione particolare a proibirne l'abuso; uno più restrittivo, dove a volte la protezione è relativa, in quanto ci si sofferma sui limiti forniti dall'articolo 10 sulla libertà d'espressione comma 2 della Convenzione⁴¹. Utilizzare l'uno o l'altro all'interno di un procedimento modifica totalmente

³⁷ <https://www.coe.int/t/dghl/monitoring/ecri/Country-by-country/Italy/ITA-CbC-IV-2012-002-ITA.pdf>

³⁸ G. ZICCARDI, *op. cit.*, p. Pag.43.

³⁹ Ibidem, p.45.

⁴⁰ << Nessuna disposizione della presente Convenzione può essere interpretata nel senso di comportare il diritto di uno Stato, un gruppo o un individuo di esercitare un'attività o compiere un atto che miri alla distruzione dei diritti o delle libertà riconosciuti nella presente Convenzione o di imporre a tali diritti e libertà limitazioni più ampie di quelle previste dalla stessa Convenzione >>.

⁴¹ << 1. Ogni persona ha diritto alla libertà d'espressione. Tale diritto include la libertà d'opinione e la libertà di ricevere o di comunicare informazioni o idee senza che vi possa essere ingerenza da parte delle autorità pubbliche e senza limiti di frontiera. Il presente articolo non impedisce agli Stati di sottoporre a un regime di autorizzazione le imprese di radiodiffusione, cinematografiche o televisive. 2. L'esercizio di queste libertà, poiché comporta doveri e responsabilità, può essere sottoposto alle formalità, condizioni, restrizioni o sanzioni che sono previste dalla legge e che costituiscono misure necessarie, in una società democratica, alla sicurezza nazionale, all'integrità territoriale o alla pubblica sicurezza, alla difesa dell'ordine e alla prevenzione dei reati, alla protezione della salute o della morale, alla

il senso dell'azione, dove la differenza principale è la mancanza di un processo di bilanciamento. I casi che sono stati analizzati nella giurisprudenza facendo riferimento all'articolo 10 tengono presente il diritto di libertà di manifestazione del pensiero e lo pesano in rapporto alle tutele dell'articolo 10 comma 2. Con l'articolo 17 diversamente non c'è un bilanciamento, le *hate speech* vengono limitate solo in base ad una ponderazione del loro contenuto. Proprio per questo motivo, Tulkens, ha la convinzione che l'articolo 17 vada usato con moderazione, in quanto se ne potrebbe abusare per commettere degli illeciti. Per esempio basandosi sull'articolo 17, la Corte Costituzionale turca ha legittimato la procedura di sciogliere i partiti politici⁴².

La CEDU ha dichiarato che la protezione dell'articolo 10 non ha valenza nei casi in cui l'attacco perpetrato nei confronti di un gruppo etnico sia di portata troppo generica, soprattutto nei casi in cui si vadano ad intaccare i valori della convenzione stessa quali la tolleranza, la pace sociale e la non discriminazione.⁴³ Un caso sul quale vorrei soffermarmi è *Delfi AS v. Estonia* del 16 giugno 2015, che ha rappresentato il primo nel suo genere, in cui la Corte è stata chiamata ad esaminare un ricorso sulla responsabilità per commenti creati dagli utenti su un portale internet di news. La società che gestisce il portale commerciale di news è stata ritenuta responsabile dalle Corti Nazionali per i commenti d'odio postati dagli utenti su un articolo da loro pubblicato. La società era intervenuta eliminando i commenti dopo sei settimane e secondo la Corte non sussisteva la violazione dell'articolo 10; i commenti erano ritenuti illeciti in quanto incitavano all'odio contro la compagnia di navigazione, sussisteva quindi la violazione dei diritti altrui ed era stata alimentata la violenza contro tale compagnia. Per la Corte in casi come questi dove i commenti sono ritenuti una vera e propria minaccia all'integrità degli individui, è ammissibile imporre una responsabilità da parte degli Stati in capo alle società senza incorrere nella violazione dell'articolo 10. Nel caso in cui le società falliscono nel trovare le giuste misure per rimuovere i commenti illeciti, se viene valutata la natura estrema dei commenti e l'insufficienza delle misure per rimuoverli senza ritardo, una sanzione moderata è accettabile e ritenuta una sanzione proporzionata.

2.6. Un progetto europeo per contrastare l'odio: PRISM

Il progetto PRISM, *Preventing, Redressing and Inhibiting Hate Speech in New Media*, nato in occasione della Giornata internazionale per l'eliminazione della discriminazione razziale, è un

protezione della reputazione o dei diritti altrui, per impedire la divulgazione di informazioni riservate o per garantire l'autorità e l'imparzialità del potere giudiziario>>.

⁴² G. ZICCARDI, *op. cit.*, p.46

⁴³ Caso *Ivanov v. Russia* 20 febbraio 2007

progetto co-finanziato dal programma Fundamental Right e Citizenship dell'Unione Europea che vede coinvolti cinque Paesi: l'Italia,⁴⁴ la Francia, la Spagna, la Romania, ed il Regno Unito.

Questa campagna vuole mostrare attraverso dati raccolti in anni di ricerca, che il fenomeno dell'*hate speech* non è per nulla trascurabile. Si pensi infatti che solamente nell'anno 2014 in Italia ci sono stati 596 casi di crimini d'odio⁴⁵, molti dei quali riguardanti episodi di razzismo o xenofobia. Il lavoro che è stato svolto nell'ambito di questo progetto ha preso via tramite una serie di studi ed inchieste, che hanno tentato di mappare il fenomeno dell'*hate speech* sui blog, sui siti web e sui social network. Si è inoltre cercato di coinvolgere e mobilitare il <<grande pubblico>> a riguardo, attraverso interventi nelle scuole e campagne di sensibilizzazione. Il progetto mira a designare delle linee guida per riconoscere e contrastare l'*hate speech* a tutti i suoi livelli, partendo da un dialogo con i vari soggetti coinvolti nel fenomeno; partendo dalle potenziali vittime, fino a raggiungere le istituzioni locali e nazionali. Ha inoltre il fine di coinvolgere anche le direzioni europee dei principali social network così da modificare alla radice le policy, nel tentativo di poter meglio intervenire a circoscrivere l'*hate speech*.

È a tal fine stato ideato un sito ufficiale PRISM, con la struttura di un blog, contenente articoli e contributi di esperti ed attivisti sul tema dell'*hate speech*, con il fine ultimo di diventare un punto di incontro e di riferimento per il dibattito Europeo sui discorsi d'odio. Lo slogan del progetto è <<No all'odio, no all'intolleranza sul web>>, e rientra in una strategia precisa che mira ad utilizzare un linguaggio alternativo; come ha spiegato Filippo Miraglia, vicepresidente dell'Arci, si è difatti creato un breve <<spot cartoon>>⁴⁶ con lo scopo di coinvolgere anche i più giovani in questa lotta contro l'*hate speech* con l'obiettivo di un netto contrasto a questo fenomeno.

⁴⁴ Per quanto concerne l'Italia le organizzazioni che si sono impegnate in questo progetto sono l'Arci, l'Ufficio Nazionale Antidiscriminazioni Razziali, l'Istituto di studi Giuridici Internazionali del CNR, la fondazione di ricerche ANCI-Cittalia e l'associazione carta di Roma.

⁴⁵ <http://www.cilditalia.org/blog/le-parole-sono-armi-fermare-lhate-speech-con-prism/>

⁴⁶ <http://www.cartadiroma.org/editoriale/contrareodiolinguaggisocial/>

Capitolo 3: La risposta di Internet all'*hate speech*

3.1. Le Policy e la responsabilità dei social network

Si è cercato di seguito di analizzare la responsabilità dei social network per ciò che viene immesso nei loro server; partendo dalla figura degli ISP, che sono degli intermediari che hanno lo scopo di stabilire un contatto tra chi vuole comunicare qualcosa e chi deve riceverla, si è arrivati alla conclusione che la loro attività non è passibile di responsabilità se si limita ad un'attività di ordine meramente tecnico, automatico e passivo, il che sottende che l'intermediario non conosca né controlli le informazioni trasmesse o memorizzate. Secondo la Corte infatti se il prestatore ha svolto un ruolo neutro, meramente tecnico comportante la mancanza di conoscenza o controllo dei dati memorizzati su richiesta di un inserzionista non può ritenersi responsabile.⁴⁷ Materiale riguardo le responsabilità degli ISP lo troviamo nell'articolo 17 del decreto legislativo n.70 del 2003, il quale riporta la clausola secondo cui non esiste un obbligo di sorveglianza a carico dei *provider* sulle informazioni che trasmette, o un obbligo di ricercare in modo attivo fatti o circostanze che indichino la presenza di attività illecite. Tuttavia può diventare civilmente responsabile dal momento in cui avendo ricevuto una richiesta dall'autorità giudiziaria, non impedisce prontamente l'accesso ai contenuti illeciti o non abbia provveduto ad informare l'autorità competente. Perché sussista un obbligo di informazione e collaborazione con l'autorità competente, si necessita di una segnalazione con gli estremi identificativi dei singoli contenuti da eliminare in quanto pubblicati in violazione degli altri diritti.

I provider dei social media, nota Jeffrey Rosen⁴⁸, utilizzano solitamente definizioni di *hate speech* in stile europeo applicate però nel rispetto della tradizione americana. Il problema più lampante è che lo stanno facendo senza trasparenza o supervisione, affidandosi a membri dello staff delle *policy* che riuniscono e mettono insieme le funzioni del giudice, del legislatore e dell'esecutore. Non è un lavoro sbagliato quello che svolgono ma sarebbe auspicabile avere maggiore trasparenza in queste *policy*, anche nel rispetto dei valori costituzionali. Alcune migliorie sono state apportate sul finire del 2015 probabilmente anche in seguito ai vari attentati terroristici di Bruxelles e Parigi - che hanno aperto gli occhi su come al giorno d'oggi i social siano diventati anche strumento per reclutare combattenti o più semplicemente dove è facile far scaturire espressioni d'odio: coerentemente con quanto appena affermato è stato infatti istituito un codice di condotta sulle espressioni illegali e di odio online tra la Commissione Europea e le aziende italiane delle maggiori

⁴⁷ M. MENSI – P. FALLETTA, , *op. cit.*, p.145

⁴⁸ Jeffrey Rosen, giurista, pubblicò nel 2013 un articolo su *Insights on Law and Society*, dal titolo <<Who decides? Civility v. hate speech on the internet.>> (Chi decide? Civiltà contro espressioni d'odio in internet) http://www.americanbar.org/publications/insights_on_law_and_society/13/winter_2013/who_decides_civility_v_hate_speech_on_the_internet.html

piattaforme: Facebook, Twitter, YouTube e Microsoft. Per garantire che nonostante ciò internet rimanga un luogo sicuro, queste aziende si sono impegnate affinché la lotta contro l'odio online sia combattuta in modo più efficace; hanno infatti rivisto le loro *policy* di funzionamento così da rimuovere o disabilitare l'accesso, laddove si presentino espressioni d'odio, in meno di 24 ore. In base a questo codice di condotta tali aziende avranno il compito di riuscire a coniugare la libertà d'espressione con la tutela degli utenti; dovranno inoltre dotarsi di un team appositamente formato, e preoccuparsi di educare i propri utenti a cosa è possibile pubblicare e cosa no, incoraggiandoli a segnalare laddove si scontrino con contenuti che incitano all'odio o alla violenza. Attraverso questo Codice di condotta si attuerà una privatizzazione delle funzioni statali⁴⁹, infatti dal testo si deduce che non saranno più le autorità di Polizia ad intervenire ma le aziende private che avranno l'onere di istruire i loro utenti su cosa è lecito e cosa no, in base ai termini di servizio. Saranno inoltre sempre i privati a incoraggiare i propri utenti a segnalare laddove ritengano ci sia un comportamento non conforme; e sarà ugualmente compito delle aziende private rimuovere, senza rispettare il principio di innocenza fino a condanna definitiva e senza attuare un giusto processo, i contenuti che riterranno illeciti. Attuando tale codice di condotta << le leggi statali e comunitarie vengono degradate a mera base per i termini di servizio delle aziende del web, che invece diventano vere e proprie leggi.^{50>>; comportando una palese illegittimità dell'Atto Europeo, infatti la Carta dei Diritti Fondamentali dell'Unione Europea prevede all'articolo 52 comma 1 che << Eventuali limitazioni all'esercizio dei diritti e delle libertà riconosciuti dalla presente Carta devono essere previste dalla legge e rispettare il contenuto essenziale di detti diritti e libertà. Nel rispetto del principio di proporzionalità, possono essere apportate limitazioni solo laddove siano necessarie e rispondano effettivamente a finalità di interesse generale riconosciute dall'Unione o all'esigenza di proteggere i diritti e le libertà altrui.>>⁵¹. Oltretutto l'approccio utilizzato nel codice sembra risultare poco efficiente, difatti non va a risolvere il problema alla radice, ma si limita ad una <<pulizia>> di internet effettuata dalle aziende in un modo poco controllato; ottenendo così una lesione ai danni degli utenti i quali si trovano limitati dei loro diritti quale (quale il diritto alla libertà d'espressione), sulla base delle policy delle aziende, che si trovano ad affrontare un compito più grande di loro, quale effettuare un bilanciamento tra i diritti, per decidere cosa sia opportuno rimuovere e cosa no, compito che oltretutto non prevede nessuna tutela per i casi di abuso o di limitazioni della libertà d'espressione. Andiamo ad analizzare nel dettaglio come le maggiori aziende si sono mosse a riguardo.}

⁴⁹ <http://www.valigiablu.it/accordo-ue-social-hatespeech/>

⁵⁰ Ibidem.

⁵¹ <http://eur-lex.europa.eu/legal-content/IT/TXT/?uri=CELEX%3A12012P%2FTXT>

3.2. L'*hate speech* policy di YouTube

Per quanto riguarda ad esempio l'*hate speech* policy di YouTube, la regola è quella di rimuovere i contenuti se si è in presenza di *hate speech*, se vengono violati i termini di servizio o a seguito di ordini di tribunali o inchieste governative. All'interno di YouTube quando si parla di incitamento all'odio << ci si riferisce a contenuti che incitano alla violenza o all'odio nei confronti di individui o gruppi sulla base di determinati attributi, come ad esempio: razza o etnia, religione, disabilità, sesso, età, stato di veterano di guerra e orientamento/identità sessuale. C'è una labile distinzione tra ciò che è considerato incitamento all'odio e ciò che non lo è. Per esempio, in linea generale è consentito criticare uno Stato o una Nazione, ma non è consentito inviare commenti rancorosi o cattivi nei confronti di un gruppo di persone solo sulla base della loro etnia.>>⁵². Quello di YouTube sembra un approccio europeo, ma al momento della decisione se lasciare o meno aperto l'accesso ad un video, si fa riferimento anche agli standard americani. Gli utenti possono segnalare i video in caso di nudità, abuso di animali, hate speech, incitazione alla violenza, o promozione del terrorismo sia attraverso una semplice segnalazione, sia inviando una <<segnalazione di abuso>>, con la quale si può attuare una comunicazione più completa; le richieste saranno poi valutate dallo staff di Google che deciderà il da farsi. Viene poi sottolineato che a volte determinati contenuti anche non violando le normative della community potrebbero ugualmente essere sottoposti a limitazioni, per esempio geografiche o d'età.

Ci fu, ad esempio, un caso nel 2007 in cui un giudice turco ordinò ai provider del paese di bloccare l'accesso a Google in risposta ad un video che insultava il fondatore della Turchia moderna, Atatürk, cosa considerata un crimine per la legge turca alla fine dopo un'analisi per verificare quali video fossero illegali in Turchia, quali fossero contro le policy di YouTube e quali invece fossero da considerare di *protected speech* si decise di eliminare il video in violazione della legge turca ma solamente in Turchia. Il giudice turco non fu però soddisfatto e chiese di eliminare il video in tutto il mondo, YouTube si rifiutò di farlo e così fu il governo turco a bloccare l'accesso alla piattaforma in Turchia; accesso che fu riabilitato nel marzo del 2007 dopo che i legali di YouTube inviarono la documentazione della rimozione del video al procuratore turco.⁵³

3.3. Facebook e l'*hate speech*

Facebook, così come un po' tutti i social media, ha fornito all'*hate speech* una cassa di risonanza non indifferente, grazie alla sua immediatezza, la pervasività, la diffusività dei contenuti erga

⁵² <https://support.google.com/youtube/answer/2801939?hl=it>

⁵³ G.ZICCARDI, *op. cit.*, p. 76

omnes, l'amplificazione e la replicabilità del messaggio tra più utenti e soprattutto per via alla sua validazione sociale che avviene attraverso i *like* e le condivisioni. Difatti secondo l'ODIHR, *l'Office for Democratic Institutions and Human Rights*, nel 2013 sono stati registrati 472 casi di crimini d'odio e, stando ai dati dell'UNAR, per la prima volta in quell'anno le discriminazioni online hanno superato quelle nell'ambito della vita pubblica e del luogo di lavoro, infatti ben 354 casi sono avvenuti attraverso i *media* in particolar modo nei *social media*.⁵⁴ L'anno successivo lo stesso ente ha registrato 347 casi di espressioni d'odio sui social, di cui oltre il 50% sono stati riscontrati su Facebook.

Proprio nell'agosto del 2014 i dirigenti di Facebook furono accusati di essere più veloci nella rimozione di nudi o immagini/post con riferimenti sessuali, rispetto a post con contenuti omofobi o razzisti⁵⁵. I vertici del social network si erano subito impegnati, garantendo una maggiore attenzione nell'identificare nel modo più rapido determinati tipi di contenuti con le successive conseguenze, nonostante ciò nel novembre i procuratori di Amburgo avevano già avviato un fascicolo contro il *managing director* di Facebook per l'Europa del nord, centrale e dell'est, Martin Ott, incolpandolo della lentezza nelle risposte alle istanze di rimozione. Le Policy di Facebook hanno comunque avuto una notevole riforma verso la fine del 2015, difatti ora troviamo nella sua policy una definizione ben specifica di cosa si debba intendere per *hate speech*. Innanzitutto viene fatta una premessa nella quale si dichiara che ci troviamo in un <<luogo>> ideato per condividere le nostre esperienze e i temi che ci sono a cuore, e che per questo potremmo incorrere in pensieri contrastanti dal nostro che potrebbero dar vita a conversazioni difficilmente gestibili; e che quindi per tutelare l'intera *community* potrebbe essere necessario rimuovere o limitare alcuni tipi di contenuti. Per la sezione riguardante l'incitamento all'odio troviamo che <<Facebook rimuove i contenuti che incitano all'odio, compresi quelli che attaccano direttamente una persona o un gruppo di persone in base a: razza; etnia; nazionalità di origine; affiliazione religiosa; orientamento sessuale; sesso; disabilità o malattia.>> si afferma poi che tuttavia << sono consentiti messaggi umoristici, satirici o commenti relativi a questi argomenti>>;⁵⁶ per concludere Facebook si impegna a garantire la rimozione di messaggi che incitano all'odio sottolineando allo stesso tempo l'importanza di controbattere questo tipo di discorsi con un dibattito costruttivo nel quale esporre il proprio punto di vista fornendo precise informazione al fine di creare un ambiente più sicuro e rispettoso.

Di recente la Presidente della Camera Laura Boldrini, in occasione della giornata contro la violenza sulle donne, ha pubblicato una serie di *hate speech* a lei rivolti su Facebook, denunciando come tale

⁵⁴ *Discorsi d'odio e social media; Criticità, strategie e pratiche di intervento*, pdf a cura di CARLA SCARAMELLA, che per l'Arci ha coordinato il Progetto PRISM.

⁵⁵ G. ZICCARDI, *op. cit.*, p.239

⁵⁶ <https://www.facebook.com/communitystandards#hate-speech>

piattaforma, che afferma di battersi contro questo tipo di messaggi d'odio, non riesca in realtà molto spesso ad arginare questo fenomeno. Richard Allan, responsabile per l'Europa delle policy di Facebook, dispiaciuto dell'accaduto ha ben accettato un invito a Roma per incontrarsi con la Presidente Boldrini, la quale ha suggerito in prima persona alcuni accorgimenti che secondo lei potrebbero aiutare nell'eliminare l'*hate speech*; da parte sua Facebook si è impegnato nell'ascoltare la Presidente, e nel riconfermare l'impegno di Facebook per bannare l'*hate speech*, affermando che <<Abbiamo avuto un incontro molto positivo e produttivo con la Presidente Boldrini e abbiamo discusso su come lavorare insieme per assicurarci che tutti siano sicuri nel momento in cui sono online. Non c'è posto per l'*hate speech* su Facebook e lavoreremo per sviluppare progetti insieme>>⁵⁷

3.4. L'approccio di Twitter all'odio online

Per quanto concerne Twitter questo risulta il social con l'approccio di stampo più statunitense, difatti fino a pochi anni fa non condannava esplicitamente l'*hate speech* ma si limitava a eliminare i *tweet* con contenuti vietati in alcuni Paesi. Un primo passo in questa direzione ci fu sul finire del 2015, come affermato in precedenza, quando avvenne un accordo in Germania tra Facebook, Google e Twitter e le autorità del Governo tedesco. Quest'ultimo richiedeva alle suddette società di stabilire delle possibili azioni per rimuovere i contenuti online con chiaro riferimento all'odio razziale; con l'accordo i social network si impegnavano alla rimozione dei contenuti da tutte le piattaforme dei social in Germania, entro le ventiquattro ore. Questo impegno da parte della Germania fu scaturito dall'arrivo di moltissimi emigranti e rifugiati politici che causarono un aumento non indifferente delle espressioni razziste online, soprattutto da parte di gruppi nazionalisti e di estrema destra; in quell'occasione Heiko Mass, Ministro della Giustizia tedesco, intervenne affermando che se i limiti del *free speech* vengono oltrepassati arrivando ad incitare l'odio con offese e minacce è giusto che quel tipo di commenti vengano dunque cancellati dalla rete, in non più di una giornata di tempo⁵⁸. Inoltre dall'accordo scaturisce che quando si tratta di debellare l'*hate speech*, tali società facciano riferimento alla legge tedesca fuorché alle proprie policy, essendo infatti la Germania molto dura per quanto riguarda i commenti pubblici che hanno come fine quello di incitare all'odio o alla violenza. Un ulteriore aumento delle misure di controllo per quanto riguarda il social network Twitter è avvenuto un paio di mesi fa a seguito delle elezioni presidenziali americane, che hanno visto come vincitore il candidato repubblicano Donald Trump. La pagina ha riscontrato un aumento degli utenti che <<si approfittano dell'apertura e usano

⁵⁷ <http://www.webnews.it/2016/11/30/boldrini-facebook-hate-speech/>

⁵⁸ <http://www.theverge.com/2015/12/16/10287498/facebook-twitter-google-hate-speech-germany>

Twitter per essere offensivi verso gli altri>>⁵⁹, ha quindi deciso di intervenire su due punti in particolare: rendere più semplici gli strumenti utili agli utenti per gestire i contenuti che si ritengono minacciosi ed attuare una formazione dello staff che avrà poi il compito di analizzarli. Per quanto concerne il primo punto è stata immessa nel sistema la funzione << toglì voce>> che è un espediente meno drastico rispetto alla rimozione di un *followers* che ci consente di non vedere temporaneamente i *tweet* di questa persona che possono apparirci indesiderati. Inoltre si cercherà di rendere più semplice la funzione di segnalazione così da motivare gli utenti a farne uso laddove ritengono ce ne sia il bisogno. L'altra azione riguarda invece la preparazione dello staff; si dibatteva infatti sulla difficoltà, per esempio, di una segnalazione esaminata in Asia, << chi in Asia esamina le segnalazioni potrebbe non riconoscere qualcosa che accade in Europa o negli Stati Uniti come un contenuto d'odio >>⁶⁰, bisogna quindi far sì che si crei un'universale familiarità su cosa riconosciamo essere *hate speech* nonostante possa a prima vista non sembrare tale. Dunque le sessioni in cui verranno formati i membri dello staff avranno cura di includere un approfondimento culturale seguito da una contestualizzazione storica dei contenuti d'odio.

Comunque andando ad analizzare l'ultima policy aggiornata di Twitter il social network afferma: <<Crediamo nella libertà di espressione e di parola, ma questa filosofia di base non ha senso se le persone non dicono la propria opinione per paura che qualcuno le faccia tacere. Al fine di assicurarci che le persone si sentano sicure nell'esprimere opinioni e convinzioni diverse, non tolleriamo atteggiamenti offensivi che sconfinano nell'abuso, incluse molestie, minacce o intimidazioni tese a mettere a tacere altri utenti>> viene poi specificato sotto la voce condotta detestabile che <<Non puoi promuovere la violenza contro una persona o attaccarla direttamente o minacciare altre persone in base a razza, etnia, origine nazionale, orientamento sessuale, sesso, identità sessuale, religione, età, disabilità o malattia. Non accettiamo nemmeno account il cui scopo primario sia incitare altri utenti a recare danno in base a queste categorie.⁶¹>>.

Notiamo quindi come anche Twitter si sia aggiornato nel corso degli anni nel modo più propenso per stare al passo con ciò che accade sulla sua piattaforma e limitarne gli inconvenienti, ha anche annunciato che sarà difficile eliminare di punto in bianco tutte le condotte offensive da Twitter ma che si impegnerà nel miglior modo possibile per far sì che i diritti di tutti siano rispettati, andando ad intervenire man a mano ogni qualvolta che lo riterrà opportuno sulla propria policy.

⁵⁹<http://www.cartadiroma.org/news/twitter-introduce-nuove-misure-lhate-speech-strumenti-piu-semplifici-usare-gli-utenti-formazione-allo-staff-social-network-ammette-nessuna-azione-condurre-all/>

⁶⁰ Ibidem.

⁶¹ <https://support.twitter.com/articles/93870>

Conclusione

Alla luce di quanto fin qui analizzato la prima cosa che ci si palesa è la difficoltà che si è avuta negli anni per contrastare il fenomeno crescente dell'*hate speech*, senza andare ad intaccare troppo la libertà d'espressione che rappresenta uno dei diritti cardine delle società a stampo democratico. Il lavoro svolto si è incentrato sul come contrastare il fenomeno dell'*hate speech* e quindi come disciplinare la libertà d'espressione attraverso il mezzo internet. Nell'ordinamento Italiano come abbiamo riscontrato durante il nostro studio è stato ancora più difficile contrastare questo fenomeno, infatti memore dell'esperienza totalitaria, l'Italia si è preoccupata di difendere in modo ancora più stringente la libertà d'espressione sottolineando all'articolo 21 della nostra Costituzione che la libertà di pensiero è permessa <<con la parola, lo scritto e ogni altro mezzo di diffusione.>> . All'interno dell'ordinamento Europeo abbiamo constatato la nascita di un vasto apparato di leggi dovuto principalmente al fatto che, innanzitutto, ogni singolo Stato ha una propria disciplina che viene poi integrata con quella comunitaria e secondo poi avendo l'Europa una maggior intenzione di tutelare il cittadino membro, e di conseguenza limitarne maggiormente la libertà d'espressione, ciascuna delle istituzioni europee e sovranazionali a lei affiliate hanno disciplinato la materia attraverso specifici provvedimenti. Il tutto è diventato ancora più complesso con la nascita dei social network, si è infatti riscontrata una non indifferente difficoltà nel tutelare i diritti dei singoli all'interno di queste piattaforme, soprattutto in quanto la maggioranza di esse essendo di stampo statunitense sono, come abbiamo analizzato, meno inclini a limitare la libertà d'espressione per contrastare i discorsi d'odio. Negli Stati Uniti vige difatti il principio della libertà sacra e inviolabile dell'individuo, l'unico strumento sul quale ci si basa per disciplinare i casi di *hate speech* è il primo Emendamento, con il quale vengono stabilite delle linee guida per capire in quali casi l'opinione dell'individuo vada tutelata e in quali invece si sia in presenza di un discorso d'odio perseguibile. Di conseguenza si è cercato inizialmente di limitare l'*hate speech* responsabilizzando i c.d. *provider*, rimediando ai casi di odio che si verificavano sulle piattaforme o cercando di prevenire tali fatti attraverso una possibilità d'intervento diretto dei *provider*. Ma oltre al fatto di aver analizzato che non è possibile considerare un provider responsabile di ciò che circola sulle proprie piattaforme, se non in specifici, delimitati casi, si venne manifestando la preoccupazione dei *provider*, che nonostante non fossero tenuti a rispettare i limiti imposti dal primo Emendamento, non volevano ledere la propria immagine pubblica apparendo come dei servizi che soffocassero la libertà d'espressione. Si palesò oltretutto la difficoltà che avrebbero riscontrato i *provider* dal punto di vista di gestione dei costi nel caso in cui fosse stato immesso un principio di obbligatorietà di controllo preventivo dei contenuti. Come abbiamo avuto modo di verificare attraverso il nostro studio, la mossa successiva fu quella di tentare di gestire il problema dall'interno, attraverso una modifica delle *policy*. Negli ultimi 10 anni si è difatti iniziato a

vedere un maggior impegno da parte dei social network come Facebook, YouTube e Twitter i quali si sono impegnati nel modificare le proprie policy affinché ci sia la libertà d'espressione ma nella misura in cui questa non vada ad intaccare i diritti e le libertà altrui. Si è inoltre andata a migliorare la funzione di <<segnalazione>> che permette agli utenti dei social di segnalare ciò che ritengono lesivo dei propri diritti o non in linea con gli standard della pagina, così da facilitarne il lavoro; non è infatti possibile utilizzare esclusivamente, come si è pensato di fare, degli algoritmi per eliminare i contenuti non in linea con gli standard della pagina, e questo perché all'interno dei social subentrano troppe varianti che non renderebbero possibile individuare i discorsi d'odio o elementi di diffamazione solamente attraverso l'identificazione di parole o frasi chiave, che potrebbero essere utilizzate da un eventuale algoritmo.

La modifica delle Policy è avvenuta in modo graduale e non parallelamente ad ogni social; per esempio Twitter è il social che in modo più tardivo ha dato una definizione di *hate speech*. Un aiuto per rendere più chiaro ed omogeneo il concetto di *hate speech* e quindi di conseguenza per comprendere cosa fosse legalmente perseguibile come tale è arrivato dall'Unione Europea, la quale ha redatto un codice di condotta per tentare di uniformare la definizione di *hate speech* e le conseguenti pratiche legali da intraprendere.

Come abbiamo avuto modo di riscontrare, anche dallo svolgimento di cause poste all'attenzione della Corte, si è oggi abbastanza in grado di limitare l'*hate speech* e di punirlo in modo adeguato così da difendere i diritti, l'onore e la reputazione dei soggetti lesi; nonostante ciò non è sempre possibile identificare i discorsi d'odio presenti sulle varie piattaforme, e anche quando questi vengono segnalati da utenti, in quanto percepiti come lesivi dei propri diritti, non è detto che risultino in contrasto con le policy della pagina e quindi da eliminare.

Ne abbiamo avuto la prova in epoca recente, in cui è emersa una realtà davvero spiacevole all'interno della piattaforma di Facebook: è stata infatti riscontrata la presenza di <<gruppi chiusi>> all'interno dei quali si incita allo stupro virtuale.⁶² Vengono ovvero postate foto sottratte da profili o <<rubate>> per strada a ragazze intente nelle loro normalissime attività quotidiane, per essere poi condivise in questi gruppi nei quali ci si sbizzarrisce nel commentarle in modo a dir poco disdicevole, sminuendo la donna a mero oggetto sessuale, ledendone ogni qualsivoglia diritto al decoro, alla decenza e alla reputazione; la cosa che ha più sbalordito è che dopo la segnalazione di una ragazza, che su Facebook ha pubblicato un post nel quale denuncia l'accaduto, nessuno sia intervenuto, o meglio, moltissime persone hanno una volta venute a conoscenza di questo fatto segnalato a Facebook alcune di queste pagine ma per la maggior parte dei casi la risposta di

⁶²<http://www.ilfattoquotidiano.it/2017/01/18/facebook-e-la-cultura-dello-stupro-sul-web-a-nessuno-interessa-delle-vittime/3323664/>

Facebook è stata che questi non violino le norme della pagina. Ci si chiede quindi come sia possibile che pagine in cui si dia vita ad un simile svilimento della figura e del ruolo della donna e in cui non manchino commenti che sicuramente non rispettano i miei diritti di donna non risulti violare le normative di Facebook.

Quindi sembra evidente che siano stati fatti degli enormi passi in avanti per quanto riguarda la tutela dei diritti online e la lotta all'*hate speech*, ma che c'è ancora molto su cui lavorare per far sì che queste piattaforme possano realmente divenire dei luoghi sicuri in cui scambiare le proprie opinioni o manifestare liberamente i propri pensieri senza incorrere in violazione di alcun tipo, né dei diritti altrui e tanto meno dei propri.

Il problema è, come disse Umberto Eco che <<I social media danno diritto di parola a legioni di imbecilli che prima parlavano solo al bar dopo un bicchiere di vino, senza danneggiare la collettività. Venivano subito messi a tacere, mentre ora hanno lo stesso diritto di parola di un Premio Nobel. È l'invasione degli imbecilli>>⁶³.

⁶³ <http://www.lastampa.it/2015/06/10/cultura/eco-con-i-parola-a-legioni-di-imbecilli-XJrvezBN4XOoyo0h98EfiJ/pagina.html>

BIBLIOGRAFIA

- M. MENSI – P. FALLETTA, *Il diritto del web. Casi e materiali*, Padova, 2015.
- G. ZICCARDI, *L'odio online. Violenza verbale e ossessione in rete*, Milano, 2016.
- G.W. ALLPORT, *The Nature of Prejudice*, Cambridge, 1949; trad. it. , *La natura del pregiudizio*, Firenze, 1973.
- Y. BENKLER, *La ricchezza della rete*, Milano, 2007.
- G. GARDINI, *Le regole dell'informazione. Principi giuridici, strumenti, casi*. Torino 2009.
- *Discorsi d'odio e social media; Criticità, strategie e pratiche di intervento*, pdf a cura di CARLA
- SCARAMELLA, che per l'Arci ha coordinato il Progetto PRISM.

SITOGRAFIA

- Federaciòn sos racismo: <http://sosracismo.eu>
- UNESCO: <http://unesdoc.unesco.org>
- COSPE onlus: <http://www.cospe.org>
- Parliamentary Assembly: <http://www.assembly.coe.int>
- Consiglio d'Europa: <https://www.coe.int>
- Coalizione Italiana: <http://www.cilditalia.org>
- Carta di Roma: <http://www.cartadiroma.org>
- American bar association: <http://www.americanbar.org>
- Valigia blu: <http://www.valigiablu.it>
- EUR-lex: <http://eur-lex.europa.eu>
- Facebook: <https://www.facebook.com>
- WebNews: <http://www.webnews.it>
- The verge: <http://www.theverge.com>
- Twitter: <https://support.twitter.com>
- Il fatto quotidiano: <http://www.elfattoquotidiano.it>

Abstract

In this study we analyze the growing phenomenon of hate speech, usually translated in Italian with the expression << discorsi d'odio >>, could be placed within that category of crimes defined << Hate crimes >> within, which it is possible to find also crimes related to discrimination. In particular, that category includes violent acts committed against persons belonging to a specific social group, identified on the basis of race, ethnicity, religion, sexual orientation, gender identity to any special physical or mental state. The hate speech is a category of American jurisprudence that Just few years ago has become part of the European law. It includes not only hate speeches but also expressions that contain discriminatory elements. Indeed, it indicates words and speeches which are meant to express hatred and intolerance toward individuals and groups of people. Our goal is to test the effectiveness of the solution methods that were applied in the case law on the matter, by examining the measure put in place to stem this phenomenon within social networks.

This phenomenon is growing throughout Europe and is taking advantage of the web, especially social media, appealing to freedom of expression, guaranteed by article 10 of the European Convention on human rights. As we will discuss later on this paper, by bringing the definitions of hate speech provided by many scholars, we will see how they believe that there is a shared notion of what must be understood as hate speech. The differences can be sensitive when moving from one legal system to another. After analyzing the various definitions exist, we came to the conclusion that it takes at least three requirements that an expression can be considered hate speech: 1) the will and intent to incite hatred by any means of communication; 2) the exhortation itself, such as to cause acts of hatred and violence against individuals targeted; 3) the concreteness of risk that acts of violence or discrimination occur. This means that if the statements do not have the intent incite hatred, do not fall within the third hate speech. Moreover, in order to be considered hate speech it is crucial to understand if the incitement to hatred or violence is strictly connected to the idea of damage, discrimination or violence.

To do so, five components are delineated: the means by which communication occurs; the content of the message, the recipient, the context within which it takes place and the effects that this causes.

Once analyzed the phenomenon in every facet and definition, We will proceed through a comparative analysis between hate speech committed online and offline.

This comparison will help us to examine the differences between those two forms of hate speech, and to better understand how the advent of internet - and particularly social media - has impacted

this phenomenon. In fact in the late 1990s, the first studies on the problem of hate speech online, on civil rights in cyberspace and on new challenges posed to society by the internet, was published in America . The main novelty consisted in the possibility of communication and access to information in real time and at a low cost. Such innovations had a downside, which facilitated the task to those who wanted to use the internet to circulate hate speech and incite violence. At the same time showed that the traditional concept of jurisdiction, with the physical limits that has achieved, is destined to decay, given the fluidity with which the contents were traveling across the network.

We find indeed many difficulties in combating hatred online, especially in our system. The main reason why those difficulties occur is probably related to the totalitarian experience faced by Italy in the period leading up to the birth of the Constitution, where every form of freedom was absent in its almost totality, and this reason has attempted thereafter to protect so tightening freedom of expression. Another reason could be the fact that initially we tried to regulate our system using the standards used for the rest of the means of dissemination; nevertheless, as we saw in this analysis, this is not possible. In fact, as the ruling June 6, 2000 n. 137 states, the internet is a means of information << quite peculiar, to whom, he lectured in criminal matters, cannot be extended, mutatis mutandis the discipline dictated by the press or on radio or television >>.

Currently, the largest legal instrument within our legal system to counter hate speech is the law left-handed flanked by the Penal Code. Indeed articles 594 and 595 recognizes the crime of << injury>> and criminal << defamation >>. As we will examine, there is a deficiency regarding the defamation consumed through the internet, although it is now the most powerful tool of dissemination of information. But in that regard, the Court has not seen fit to make changes, noting that the internet due to its peculiarities will be a hypothesis of aggravated libel under article 595, paragraph 3, Penal Code, as committed by other advertising medium than printing. Lately a milestone has been reached with the birth in July 2015 Italian Internet Rights Charter, promoted by the speaker of the House Laura Boldrini, aiming to enter a series of values and principles in the use of the internet.

Furthermore, the European and International Order has been analyzed, and has emerged that the EU is very concerned to try to stem the phenomenon of hate speech, implementing various

interventions and resolutions adopted by the European Parliament, to encourage Member States to take concrete action to counter the phenomenon of hate online. The main elements that have been used are the article 13 of the Amsterdam Treaty, article 21 of the Charter of Nice and finally directives 2000/43/EC and 2000/78/EC. The Council of Europe demonstrated a focus on discrimination on the web, has asserted that << considers racism not as an opinion but as a crime >>. Moreover the impact of Council's action is even wider. In fact inside is not only regulated the fight against the hate speech, but to all forms of hate speech, which for the Council extends to all forms of hate speech, racially based ethnic, religious, and social. As to the the European Court of human rights, is has dealt several times to find an answer to the question of hate speech, and did so by applying the principle of non-discrimination so not absolute, but balancing it with other rights on the same Convention. In fact, when the Court is confronted to the subject of hate speech and must weigh in relation to freedom of expression, usually using two approaches: one larger type, which excluded certain expression from the protection of article 17, article for the prohibition on the abuse of right which is protected freedom of expression but with particular attention to prohibit its abuse; then a more restrictive one, where sometimes the protection is relative, since it focuses on the limits provided in article 10 paragraph 2 of the Convention on freedom of expression. The use of either the first or the second within a process review totally the sense of action, where the main difference is the lack of a balancing process.

In the rest of our study, we tried to analyze the social responsibility to what is entered on their server, starting with the figure of ISP. It is possible to define ISP as intermediaries that are intended to establish contact between those who want to communicate something and who should receive. It was concluded that their activity is not subject to liability if limited to a purely technical order, automatic and passive, which implies that the broker does not know check the information which is transmitted or stored. According to the Court, if the lender has been neutral, purely technical involving the lack of knowledge or control of data stored at the request of an advertiser, in that case he cannot be considered a responsible. Social media providers usually definitions of hate speech in the European style applied though respecting tradition. The most glaring problem is that they act without transparency or oversight, relying on staff members of the policy gathering and bring together the functions of judge, legislator and of the performer. It is not wrong that they play but it would be desirable to have greater transparency in these policies, even in respect of constitutional values. Some improvements has been evidenced at the end of 2015 when it was set up a code of conduct on illegal and expressions of hate online between the European Commission and

companies of the major platforms: Facebook, Twitter, YouTube and Microsoft. According to this code of conduct these companies will be able to combine freedom of expression with the protection of users; they must also adopt a team specially trained, and bothering to educate their users on what to publish and what not, encouraging them to report where clashing with contents that incite hatred or violence.

Finally, The problem from the inside has been studied through an analysis of social policy, in order to see how they try to counteract this phenomenon and how they have evolved over the years. Over the past 10 years has indeed started seeing greater use of social networks such as Facebook, YouTube and Twitter, which have pledged to change their policy so that there is freedom of expression but to the extent that this does not affect the rights and freedoms of others.

For Instance, within YouTube when we talk about hate speech << It refers to content inciting violence or hatred against individuals or groups based on certain attributes, such as race or ethnicity, religion, disability, gender, age, veteran status, and sexual orientation/gender identity. There is a faint distinction between what is considered hate speech and what is not. >>. Regarding the policy of Facebook, have had a significant reform by the end of 2015, in fact now we find in his policy a very specific definition of what constitutes hate speech, emphasizing the importance of counter this phenomenon not only through censorship but through constructive debate in which expose their views by providing precise information in order to create a safer environment and respectful. Twitter has been the social media, that most recently has aligned its policy by inserting a definition of hate speech, is this because we have seen through our analysis which happens to be the platform that uses more American approach. The first step is to sign the agreement in Germany with German Government authorities in which they would be busy with Facebook and YouTube to react more promptly to eliminate the hate speech; In addition a further increase of control measures with regard to the social network Twitter is advent a few months ago as a result of the U.S. presidential election, which saw the winner in the Republican candidate Donald Trump. The page has experienced an increase in users who << take advantage of opening and using Twitter for being abusive toward others >>, and has therefore decided to intervene in order to improve the protection against hate speech.

To conclude, it was very difficult to counter the growing phenomenon of hate speech, especially being careful to not go to affect too much the freedom of expression that represents one of the key rights of democratic mold. It is crucial, therefore, to say that today we are quite capable of restricting hate speech and punish appropriately so as to defend the rights, honor and reputation of

the injured parties. Although this is not always possible to identify the hate speeches on various platforms, and even when they are reported by users because perceived as detrimental to their rights, does not mean that they are contrary to the policy of the page, and then to delete. Therefore, it seems clear that enormous progress has been made as regards the protection of human rights and the fight against online hate speech, but that there is still a lot to work to make these platforms To let them become safe places where exchange their opinions or to freely express their thoughts without any violation of any kind nor the rights of others, much less of their own.