

La solidarietà: primo strumento di contrasto all'incitamento all'odio (hate speech)

Lucia Tria

SOMMARIO: 1. - *L'hate speech, fenomeno connaturato alla storia dell'uomo.* 2. - *Incitamento all'odio online e offline.* 3. - *Fake news e trolling.* 4. - *La particolare "efficacia" dell'incitamento all'odio online.* 5. - *Le fonti giuridiche.* 6. - *Il difficile bilanciamento tra il diritto alla libertà di espressione e contrasto ai discorsi di incitamento all'odio nella giurisprudenza della Corte di Strasburgo.* 7. - *La Corte di giustizia UE.* 8. - *Il piano d'azione di Rabat.* 9. - *La difficile individuazione di strategie di contrasto.* 10. - *Le bugie sono più potenti della verità.* 11. - *La reazione dell'Europa.* 12. - *La risposta giurisdizionale.* 13. - *Combattere la solitudine e la marginalizzazione.* 14. - *Conclusioni.*

1. L'hate speech, fenomeno connaturato alla storia dell'uomo.

L'incitamento all'odio (hate speech, in inglese) è un fenomeno che pone moltissimi problemi sociali e giuridici – nazionali, europei e globali – soprattutto nella sua versione on-line.

Si tratta di un fenomeno che si collega alle discriminazioni in quanto è intenzionalmente diretto a spingere all'intolleranza verso persone o gruppi vulnerabili alle discriminazioni e con modalità tali da propagarsi in modo efficace.

Purtroppo, l'hate speech – al pari delle discriminazioni – è un fenomeno connaturato alla storia dell'uomo, la misoginia così come il razzismo sono radicati nella cultura occidentale, pertanto gli unici strumenti di contrasto non possono essere la rimozione dei singoli contenuti e/o l'irrogazione di sanzioni penali.

Infatti, tali strumenti – il cui utilizzo deve comportare la possibilità per le vittime di difendersi adeguatamente, non solo in sede giudiziaria – servono a punire comportamenti "patologici", che dovrebbero rappresentare delle eccezioni, mentre si deve puntare alla diffusione di prassi e comportamenti "fisiologicamente" corretti e rispettosi del principio di eguaglianza, andando ad analizzare le cause del fenomeno.

La differenza dell'era tecnologica è rappresentata dal mezzo a disposizione. Se un tempo era solo l'élite a poter diffondere il proprio pensiero (tramite la televisione e i giornali), oggi con internet chiunque può farlo.

Tuttavia, benché spesso l'élite continui ad utilizzare discorsi d'odio per contrastare la dissidenza, tale costume non viene collegato – neppure dal punto di vista emulativo – con la generale diffusione dei discorsi d'odio da parte dei comuni cittadini.

Si deve anche considerare che – come risulta emblematicamente dalla giurisprudenza della Corte europea dei diritti dell'uomo – nel contrasto al fenomeno dell'incitamento all'odio si deve effettuare un adeguato bilanciamento tra la tutela delle vittime delle espressioni offensive (perché manifestazioni di razzismo, xenofobia, antisemitismo, nazionalismo aggressivo, misoginia, discriminazione nei confronti delle minoranze e degli immigrati) e il diritto dei singoli individui (compresi i giornalisti e i politici) di esprimersi liberamente e di “offendere, scioccare o turbare” gli altri.

2. Incitamento all'odio online e offline.

Tale bilanciamento, già di per sé complesso, diviene ancora più difficile per il fatto che vanno presi in considerazione anche altri fenomeni come i troll – persone che intralciano il normale svolgimento di una discussione inviando messaggi provocatori, irritanti, falsi o fuori tema con lo scopo di disturbare, provocare reazioni forti negli altri agendo sia sui profili di comuni utenti, che di personaggi pubblici o sulle pagine di aziende e brand conosciuti – e dei fake account, cioè dei falsi profili creati su Facebook, Twitter così come su altre piattaforme, che possono essere prodotti in serie da software dedicati oppure creati a mano.

Nel primo caso si tratta di quelli che sono chiamati “bot”¹, cioè di profili automatizzati creati per un certo scopo (ri-tuittare solo alcuni tipi di

¹ Il termine BOT corrisponde ad una abbreviazione di ROBOT, nei paesi anglosassoni, indica un programma autonomo che nei [social network](#) fa credere all'utente di comunicare con un'altra persona umana. I BOT sono in continua evoluzione, anno dopo anno, sicché è sempre più difficile distinguere un BOT da una persona umana. L'inizio della storia dei BOT viene di solito individuato nella teorizzazione da parte dell'informatico britannico Alan Turing di un test per capire se una macchina fosse effettivamente in grado di imitare il comportamento umano, basato sull'analisi di una conversazione tra un individuo e un computer effettuata da una persona esterna, onde stabilire se gli scambi della conversazione consentissero o meno di distinguere le risposte della macchina da quelle della persona. Se tale distinzione risultava impossibile, voleva dire che la macchina aveva superato il test di Turing (erano gli anni Cinquanta).

Il primo BOT ad avvicinarsi al superamento della prova teorizzata da Turing fu [ELIZA](#) nel 1966, un programma che fingeva di essere uno psicoterapista e che regolava le sue risposte in base alle cose scritte dal suo interlocutore. In pratica iniziava la conversazione chiedendo quale fosse il problema della persona, analizzava la risposta alla ricerca di una serie di parole chiave e, sulla base di queste, dava una risposta. ELIZA era quindi capace di creare conversazioni verosimili in un ambito piuttosto ristretto, cosa che comunque fu sufficiente per superare il test. BOT del medesimo tipo sono poi stati sperimentati alla

contenuti, come nel caso dell'ISIS, oppure adescare incauti utenti per ricattarli anche sessualmente) che finiscono per vivere di vita propria senza più alcun intervento umano. In linea generale, si tratta di un qualsiasi account in cui qualcuno si spacci per qualcun altro – magari rubandogli l'identità – oppure di qualcosa che produce online attività assimilabili a quelle di un essere umano, sempre più simili a quelle di un utente autentico.

Nel secondo caso, invece, all'utenza “creatrice” può corrispondere una persona in carne e ossa, che quindi può compiere azioni precise spesso con obiettivi chiari, quali la diffamazione, lo stalking fino all'insulto e alla denigrazione generica o al “trolling” politico – di cui spesso si parla – ma anche per promuovere prodotti o indirizzare le discussioni.

Il fenomeno della creazione di falsi profili sulla piattaforme web non nuovo. Già nel 2013 acquistare un pacchetto da mille follower o “amici” taroccati costava meno di 10 euro. Basta fare qualche ricerca su Google con richieste quali “buy twitter followers” o “buy facebook friends” o “likes” per rendersi conto che poco è cambiato. Società specializzate di social media marketing – in realtà vere e proprie “bot farm” come la misteriosa Rantic che sforna fino a 50mila fake al giorno, anche se la maggior parte ha sede in India, Bangladesh e Filippine – utilizzando numerosissimi intermediari fanno da anni ricchi affari, abbeverandosi a un comparto che solo nel 2014 ha mosso su Facebook 200 milioni di dollari. Ad essere cambiati sono semmai gli scopi per i quali queste sterminate guarnigioni di false utenze vengono fabbricate. Infatti, mentre prima si parlava di mero “social doping”, cioè di un aumento dei follower e del seguito su Facebook etc., anche in ottica commerciale o di popolarità, magari per scalare nel

fine degli anni Novanta da diversi sistemi di chat come Instant Messenger di AOL negli Stati Uniti e Windows Messenger di Microsoft. Davano informazioni sulle notizie, sugli orari del cinema e sul meteo, cose che oggi si possono ottenere dai “personal digital assistant” (assistenti personali digitali) come Siri su iPhone o a Google, ricevendo direttamente una risposta nella pagina dei risultati prima della classica lista di link. I programmi per le chat dei primi tempi di Internet sparirono con l'emergere dei social network e in seguito delle APP, mettendo fuori circolazione per quasi un decennio i BOT. I progressi nei sistemi di intelligenza artificiale e soprattutto la grandissima diffusione delle applicazioni per scambiarsi messaggi, come Messenger e WhatsApp, hanno dato nuova vita ai BOT, per ragioni commerciali e cioè nell'idea che i BOT possano cambiare almeno in parte l'attuale mercato delle applicazioni (oggi governato da Facebook e Google), offrendo nuove e insolite opportunità ai produttori di contenuti e a chi vende cose online. Inoltre, per la creazione di nuovi modi di fare pubblicità, e soprattutto per una pubblicità personalizzata e calibrata sugli interessi del singolo cliente, l'utilizzazione dei BOT – e di quanto hanno imparato nel tempo dai loro interlocutori – appare strategica (per ulteriori notizie vedi: “Cosa sono i bot”, www.ilpost.it 21 aprile 2016.). Nel tempo si sono molto diffusi ed ora si stanno affermando i CHATBOT, che rappresentano l'evoluzione dei c.d. “personal digital assistant” che da vent'anni aiutano gli utenti con interazioni “ricche” gestite da BOT. Così come per i BOT si tratta di “strumenti”, quel che conta è l'uso che se ne fa, tenendo presente che sono molto invasivi.

ranking dei motori di ricerca, oggi si spazia su diversi fronti: dal dating al business delle recensioni finte, dalla propaganda interna alla criminalità organizzata passando per la lotta politica².

L'intossicazione è estesissima anche sul lato degli affari: basti considerare che, stando ai numeri del Ponemon Institute, il 10% dei database delle aziende interpellate è costituito da falsi profili potenzialmente dannosi per la sicurezza delle informazioni.

Fatti come la petizione per il referendum sulla Brexit, gonfiato a dismisura da utenze false prodotte in serie da bot i cui codici sono stati diffusi sul forum per smanettoni 4chan, ci raccontano infine che il peggio deve ancora arrivare.

A titolo di esempio si può considerare 4chan – forum e piattaforma dove si ritrovavano hacker e smanettoni, facile e immediato da usare e garanzia di anonimato assoluto – con il passare del tempo è stato messo sotto accusa per essere diventato il luogo preferito da pedofili, malintenzionati e haters, ma questo non è servito ad abolirne la capacità offensiva. Infatti, ha lasciato spazio ad 8chan (o Infinitechan), nato nel 2013 da una scissione, creato da Fredrick Bennan, sviluppatore statunitense disabile, già creatore di una community di «uomini vergini».

Qui hanno trovato la loro casa gli estremisti dell'Alt Right, diffondendo per lo più post su teorie complottiste, nei quali si ritrova tutta la simbologia del neonazismo, dai *meme* di Pepe the Frog fino ai *dingoes* australiani. La Cina di Xi Jinping o la Germania del multiculturalismo di Angela Merkel sono dipinte come il nemico mentre gli Stati Uniti di Donald Trump vengono esaltati come la terra promessa di un ritorno della supremazia bianca. Con la vittoria di Donald Trump, la diffusione di siti come Breitbart, creatura di Steve Bannon, i messaggi di odio e razzismo propinati dai partiti populistici, si è creato l'hotbed, il terreno fertile, per l'innesco della miccia. Infatti, in questi anni numerosi sono state gli allarmi degli analisti di intelligence sul pericolo di una nuova ondata di radicalizzazione negli ambienti di estrema destra con il rischio di pericolosi attentati.³

3. Fake news e trolling.

² S. Cosimi, *Fenomeno fake: l'invasione degli utenti falsi, dalla politica social al dating online*, www.repubblica.it, 29 luglio 2016.

³ M. Serafini, *Che cos'è 8chan, il forum preferito dai suprematisti bianchi*, www.corriere.it 28 aprile 2019

Alle precedenti osservazioni deve essere aggiunto che spesso il regno incontrastato dei troll e degli account fake sono le c.d. fake news o bufale, anch'esse prive di una definizione generalmente acquisita.

Letteralmente il termine equivale a notizia falsa o bufala ed è usato per indicare articoli redatti con informazioni inventate, ingannevoli o distorte, resi pubblici nel deliberato intento di disinformare o diffondere bufale attraverso i mezzi di informazione tradizionali o via Internet, per mezzo dei social media.

Secondo la definizione coniata da Melissa Zimdars, docente di comunicazione al Merrimack College di North Andover (Massachusetts), si tratta di notizie che “inventano del tutto le informazioni, disseminano contenuti ingannevoli, distorcono in maniera esagerata le notizie vere”.

Deve essere, però, considerato che: la finalità e le modalità con le quali si utilizzano le fake news e il trolling non sono sempre illecite.

Può, infatti, accadere che:

a) l'utilizzo delle fake news avvenga – in modo chiaro e trasparente – per scopi umoristici o per fare satira, come accade in Italia per il sito “Lercio”, che dal 2012 fa circolare per il web critiche più o meno mordaci su qualsiasi argomento di attualità, tempo libero o sport, con l'unica prerogativa di raccontare fatti e accadimenti mai avvenuti, con l'idea – secondo gli ideatori – di cercare di inculcare senso critico, col paradosso, nella convinzione gli utenti debbano venire comunque pungolati per risvegliarsi dalla passività intellettuale in cui si trovano, anche se capita di ottenere l'effetto contrario;-

b) si faccia trolling con toni provocatori e dissacranti, talvolta al limite dell'offensivo, ma solo per provocare utili approfondimenti su certi argomenti o comunque per manifestare “opinioni estreme” senza intenti di odio, ma esclusivamente per stimolare discussioni ragionate e interventi lunghi e contestualizzati, contestando l'impostazione dei social media, che sono stati ideati per enfatizzare l'aggiornamento rapido e l'intervento breve e secco. Con un effetto simile a quello ottenuto dal passaggio dai media cartacei alla televisione, laddove in quest'ultima la degenerazione dei talk show è sempre più all'ordine del giorno, talvolta appositamente cercata perché garantisce ascolti maggiori;

c) all'origine di un messaggio rabbioso, provocatorio oppure condito di turpiloquio non vi sia un troll, un cliente insoddisfatto. Sicché per il gestore di una pagina social non è semplice distinguere tra un comune

disturbatore e un cliente insoddisfatto, sebbene le motivazioni che li spinge all'azione siano assai differenti sebbene mentre sia necessario ignorare o bloccare il troll, il cliente invece vada seguito e ascoltato, per provvedere repentinamente alla risoluzione dei suoi problemi. Ne deriva che, per scoprire la natura della lamentela, lo strumento più efficace è quello invitare l'utilizzatore in una conversazione privata, togliendolo quindi dalle attenzioni pubbliche: si tratti della mail, dei servizi di messaggistica, del telefono o di qualsiasi altra forma d'interazione. Il cliente deluso sarà ben disponibile a un confronto, nella speranza di risolvere le problematiche incontrate, mentre il troll tenderà a dileguarsi, non avendo ovviamente alcuna questione reale da riferire all'assistenza;

d) alcuni clienti usano legittimamente per il loro profilo nomi di fantasia o fotografie generiche per ragioni di privacy: quindi l'utilizzo di tali strategie non deve portare a compiere il facile errore di considerare tutti coloro che se ne avvalgono dei troll.

Poiché la democrazia vive di un dialogo continuo tra le parti, costituito anche da contestazioni e disaccordi, scioperi e proteste, una risposta legale ed esclusivamente repressiva dei troll come dell'hate speech potrebbe compromettere il corretto equilibrio che è la base della democrazia.

Si rileva che il dissenso, la contestazione, la protesta di piazza, sono spesso attuate tramite discorsi di hate speech ed espressioni offensive, ma sono generalmente anche l'espressione di un malcontento, di un qualcosa che non funziona nella società. E tutto questo non può essere relegato nella definizione di trolling o addirittura di hate speech. Una democrazia non può vivere nascondendo i disagi sociali di parte della popolazione.

Negli USA, pur non dubitandosi della dannosità per le persone dell'hate speech tuttavia si teme maggiormente la censura, nella consapevolezza che occorra una particolare attenzione nel regolamentare i discorsi d'odio.

Infatti, una semplificazione del problema scade facilmente nella censura, che ovviamente colpisce i discorsi dei cittadini, non certo quelli delle élite. Per questo è necessario che la normalità sia il libero fluire delle parole online e solo in casi gravi si ricorra alla soppressione dei discorsi d'odio, preferendo delle risposte alternative, come politiche volte alla riduzione delle discriminazioni sociali, politiche educative per la responsabilizzazione dei cittadini, preparandoli alla comprensione dei propri diritti e all'identificazione dell'hate speech nella vita reale e online.

Va, però, considerato che si vanno diffondendo invenzioni di “bufale” effettuate con scopi tutt’altro che nobili e molto inquietanti:

1) movente economico, consistente nel consentire ai gestori dei siti web (grazie a un sistema di redirect da un sito all’altro) di incassare lauti guadagni con i banner posizionati o che vengono aperti, spesso inconsapevolmente, dai lettori, sul presupposto che le fake news, per come sono strutturate, attirano clic;

2) fine politico, consistente nel veicolare facili e immotivati allarmismi riguardanti temi come la salute, le crisi economiche e i migranti onde far crescere, in termini di popolarità, quei partiti che si fanno propugnatori di misure drastiche per farvi fronte. Come si è verificato, tempo fa, per la notizia dell’abolizione da parte della Chiesa svedese luterana del genere maschile nei testi biblici in riferimento al Creatore.

In Italia, ad esempio, ha fatto molto discutere il caso di un sito web che diffondeva informazioni false in chiave politica contro Paolo Gentiloni, all’epoca appena divenuto Presidente del Consiglio dei Ministri, che si è risolto grazie a due diverse inchieste che hanno chiarito che quel portale faceva parte di un più vasto insieme di realtà simili riconducibile a una società con sede in Bulgaria, gestita da italiani.

Spesso si discute del rischio di esposizione del nostro Paese a campagne di fake news finalizzate al condizionamento delle opinioni degli elettori in vista delle elezioni politiche. E si ipotizza una strategia di destabilizzazione analoga a quella che avrebbe già influenzato gli elettori olandesi, francesi e, come si è scoperto, anche inglesi durante le ultime tornate elettorali e la Brexit. In un articolo di qualche tempo fa Jason Horowitz corrispondente per l’Italia del New York Times ha sostenuto che tale situazione sia potenzialmente in grado di mettere a rischio l’integrità dell’Unione già seriamente minata negli ultimi anni.

Inoltre, un’analisi pubblicata dal Socialpolitico.it ha dimostrato che, ad esempio, la pubblicazione di bufale riguardanti presunte connessioni tra immigrati e aumento della criminalità, ad opera di portali riconducibili alla destra radicale, riscuote notevole successo in rete, facendo crescere la paura e generando un clima d’intolleranza nei confronti degli stranieri.

Di recente, in alcuni documenti inviati al Congresso USA, Facebook, Twitter e Google hanno svelato, numeri alla mano, che dagli account dall’azienda russa Internet Reserach Agency – che altro non è se non la famigerata “fabbrica dei troll” di San Pietroburgo, epicentro della

propaganda pro Cremlino – la rete è stata inondata di decine di migliaia di post su argomenti “divisivi”, dall’omosessualità alla guerra in Siria, che hanno raggiunto milioni di utenti tra il gennaio 2015 e l’agosto 2017, centrando in pieno anche la campagna per il voto che ha visto trionfare Donald Trump su Hillary Clinton.

In particolare, nei documenti inviati al congresso Usa e svelati dal New York Times, Facebook attribuisce alla “fabbrica” 80mila contenuti di quel tipo, che hanno raggiunto 29 milioni di persone; la stessa Internet Research Agency ha investito nella piattaforma social 100mila dollari per rendere più visibili i propri post. Quanto a Twitter, gli account messi in relazione con l’azienda russa e poi sospesi sono stati 2.700, scoperti fra settembre e novembre 2016 (in piena corsa per le presidenziali USA), con all’attivo un totale di 131mila tweet. Anche Google imputa a Internet Research Agency di aver fatto incetta di servizi pubblicitari e canali Youtube per diffondere messaggi divisivi, anche sulla guerra in Siria, che vede la Russia di Putin fra i protagonisti in campo. Il motore di ricerca spiega anche che l’azienda russa ha speso 4.700 dollari in pubblicità.

Tali rivelazioni hanno fatto aumentare le preoccupazioni anche negli USA, specialmente dopo la pubblicazione da parte di Fq Millennium dell’inchiesta sul c.d. Russiagate e sulla “fabbrica dei troll”, alias Internet Reserach Agency, a firma Anna Lesnevskaia, ex dipendente dell’Agenzia.

4. La particolare “efficacia” dell’incitamento all’odio online.

Pur non essendo l’hate speech online intrinsecamente differente dalle analoghe espressioni offline, tuttavia è caratterizzato da alcune peculiarità che dipendono dal mezzo (internet): l’estrema diffusione, la permanenza, la facilità di condivisione.

Si comprende, pertanto, che in conseguenza della diffusione di internet si siano avuti studi specifici sull’incitamento all’odio online.

Uno degli studi più interessanti sull’incitamento all’odio online è il Progetto UMATI effettuato in Kenya quasi contemporaneamente al piano RABAT, cioè tra il settembre del 2012 e il marzo del 2013, al fine di analizzare l’hate speech online durante il periodo elettorale, caratterizzato da accuse che anticipavano possibili violenze. Il progetto utilizzava la definizione di hate speech di Susan Benesch (Dangerous Speech: A Proposal to Prevent Group Violence), che si focalizza sulle espressioni in grado di catalizzare violenza. I risultati dello studio evidenziarono la

complessità dei collegamenti tra l'hate speech online e la violenza nella vita reale. Il progetto mostrò, infatti, che l'hate speech online non può essere il solo precursore della violenza nella vita reale, ma una finestra di comprensione delle conversazioni offline (e quindi può essere utile per scoprire problemi sociali). Inoltre, l'hate speech può essere una reazione agli eventi che accadono nella vita reale.

Inoltre, tempo fa su iniziativa dell'allora Presidente della Camera Laura Boldrini, è stata istituita presso la Camera la Commissione "Jo Cox" su fenomeni di odio, intolleranza, xenofobia, e razzismo, dalla cui relazione finale (approvata il 6 luglio 2017) – "La piramide dell'odio in Italia" – sono emersi dati molto interessanti, anche se sconfortanti, essendosi accertato che i soggetti più colpiti dai discorsi d'odio nel nostro Paese sono gli immigrati, le donne e gli omosessuali e, in particolare, che: 1) le donne sono di gran lunga le maggiori destinatarie del discorso d'odio online; 2) l'Italia è il Paese con il più alto tasso del mondo di ignoranza sull'immigrazione; 3) ed è anche quello maggiormente omofobo in ambito UE.

Ne deriva che sia dal progetto UMATI sia dalla relazione finale della Commissione Jo Cox emerge che i discorsi di odio online vanno posti in relazione con quanto si verifica offline sono spesso poco esaminato. Infatti, diversamente da quanto comunemente si fa nei recenti dibattiti non ci si dovrebbe concentrare sulla criminalizzazione del mezzo (internet) – come se la sua velocità di diffusione fosse la causa scatenante dell'hate speech – ma si dovrebbero analizzare le vere cause del problema. Da tale impostazione deriva che la soluzione – l'unica prevista – sia la soppressione dei discorsi d'odio, tanto che la risposta richiesta ai social network è univoca in tal senso.

Ma la psicologia sociale ci insegna che la cancellazione di una espressione d'odio non risolve il problema se alla base vi è un malcontento, una discriminazione sociale. L'autore semplicemente trasferirà il suo malcontento altrove, nel mondo reale, casomai in famiglia.

In altri termini, come per tutte le discriminazioni, si deve puntare alla prevenzione oltre ad una repressione efficace.

Non basta "oscurare" senza una reale comprensione non solo delle peculiarità del mezzo (internet), ma anche delle cause sottostanti il problema stesso.

Inoltre, è bene diffondere una adeguata comprensione del funzionamento di internet, onde evitare che qualcuno sia portato a

considerare più difficile, se non impossibile, essere rintracciato e di essere protetto da una sorta di anonimato.

In questa ottica alcuni Governi hanno chiesto nuove politiche di eliminazione dell'anonimato in rete (real names policies), ma iniziative del genere (che anche Facebook assecondò, per un certo periodo, cercando di costringere gli utenti a non usare i nomi reali, anche se probabilmente soprattutto per motivi commerciali) hanno sempre trovato una forte opposizione da parte delle associazioni per i diritti civili, in quanto in contrasto con il diritto alla privacy e con la libertà di espressione.

In realtà, l'anonimato o lo pseudo-anonimato consentono a molte persone di poter denunciare non solo reati, ma anche atti di discriminazione contro di loro, senza il timore di doverne subire delle ritorsioni (quali licenziamenti, minacce, aggressioni).

Deve anche essere rilevato che l'hate speech, inteso come linguaggio provocatorio, tagliente, che tende a ridicolizzare, è spesso l'unico mezzo utilizzabile per raccogliere l'attenzione della platea online.

Inoltre, nei momenti critici (es. elezioni), l'hate speech può essere facilmente soggetto a manipolazioni, con accuse verso le opposizioni politiche di fomentare l'odio sociale e quindi può diventare strumentale alla soppressione della dissidenza. Si tratta, infatti, di uno strumento che spesso è utilizzato dal gruppo dominante per reprimere la dissidenza e imporre la propria visione del mondo (durante l'apartheid in Sud Africa, le leggi sull'hate speech erano usate per criminalizzare le critiche verso la supremazia bianca).

È evidente, quindi, che i provvedimenti di soppressione dei discorsi d'odio possono non essere il mezzo migliore per contrastarli, e possono portare anche ad abusi, senza in realtà risolvere nulla nel caso in cui vi sia un problema sociale.

Paradossalmente le limitazioni alla libertà di espressione, giustificate da esigenze di repressione dell'odio online, possono portare addirittura alla recrudescenza del fenomeno.

Inoltre alcune organizzazioni per la tutela dei diritti umani hanno più volte sostenuto che la libertà di espressione e l'uguaglianza si sostengono l'una l'altra, sicché le limitazioni alla libertà di espressione, la rimozione delle espressioni d'odio, spesso possono avere gravi conseguenze sull'uguaglianza.

In tale prospettiva è pacifico che si deve puntare sull'esistenza di media indipendenti e pluralistici, strumento essenziale per una concreta lotta alla discriminazione e la promozione di intese interculturali.

5. Le fonti giuridiche.

Sono molteplici le fonti che disciplinano il contrasto all'incitamento all'odio, adottando definizioni non coincidenti⁴:

a) la Dichiarazione universale dei diritti umani del 1948 che all'art. 7 prevede il diritto alla tutela da discriminazioni;

b) la Convenzione internazionale sui diritti civili e politici (ICCPR) sancisce il diritto alla libertà di espressione (art. 19) e la proibizione degli appelli all'odio nazionale, razziale e religioso che costituiscano incitamento alla discriminazione e alla violenza (art. 20), stabilendo che la libertà di espressione non è un diritto assoluto, ma le restrizioni devono essere previste dalla legge e devono essere necessarie. Inoltre, la libertà di espressione può essere limitata solo per la tutela del diritto alla reputazione, della pubblica sicurezza o dell'ordine pubblico, della salute pubblica o della pubblica morale. Quindi non sono ammissibili divieti generalizzati, ma occorre che siano relativi a specifici contenuti;

c) la Convenzione per la prevenzione e punizione del crimine di genocidio (1951) mira a tutelare i gruppi definiti per razza, etnia e nazionalità, comprendendo anche quelli religiosi e quindi si limita a vietare le azioni commesse con l'intento di distruggere, in tutto o in parte, gruppi razziali, etnici, nazionali o religiosi;

d) la Convenzione internazionale sull'eliminazione di ogni forma di discriminazione razziale (ICERD, 1969) che si limita a condannare le discriminazioni relative alla razza e all'etnia, ma prevede la punizione anche per la mera diffusione di messaggi d'odio o superiorità razziale, laddove invece la "Convenzione internazionale sui diritti civili e politici" richiede, per la punibilità, che ci sia l'intento di seminare odio, intento che, quindi, deve essere provato. Questo è rilevante per gli spazi online, all'interno dei quali può essere sufficiente la sola diffusione dei messaggi d'odio per un intervento;

e) la Convenzione per l'eliminazione di discriminazione contro le donne (CEDAW, 1981) si riferisce agli atti di violenza e discriminazione

⁴ Vedi, per tutti: B. SAETTA, *Contrastare l'hate speech online: questioni aperte e alcune proposte* in www.valigiablu.it 18 febbraio 2017

basati sul genere e sull'identità sessuale e, in questo ambito, chiede agli Stati di prendere misure contro la diffusione di materiale pornografico degradante che dipinge le donne come oggetti;

f) la Carta dei diritti fondamentali dell'Unione europea prevede la libertà di espressione all'articolo 11, il quale proibisce l'abuso dei diritti e di conseguenza devono ritenersi inammissibili quelle limitazioni alla libertà di espressione che non siano previste dagli scopi della Carta.

Peraltro, pur non esistendo una definizione comune di "incitamento all'odio", in genere nelle legislazioni ci si riferisce a discorsi di discriminazione, ostilità e violenza, quindi invocazioni contro persone identificate con gruppi sociali o demografici. Talvolta sono inclusi anche discorsi di incitamento alla violenza, e altri che favoriscono un clima di pregiudizio e intolleranza, che poi possono eventualmente portare alla discriminazione mirata e ad attacchi violenti.

Nel linguaggio comune la definizione è più ampia, finendo per ricomprendere insulti contro le autorità o espressioni dispregiative per individui particolarmente esposti al pubblico.

Quindi, dal punto di vista giuridico si parla di "incitamento all'odio" con riferimento a gruppi o categorie di persone e non a singoli. In quanto nel caso in cui le vittime siano dei singoli individui, si fa ricorso ad altri strumenti per contrastare eventuali espressioni d'odio, dalle leggi penali (diffamazione, ingiurie, stalking) ad azioni civili di risarcimento del danno oppure ad esposti alle Autorità di Polizia (es. art. 1 TULPS).

Tra quelli indicati, il documento internazionale che tuttora ha maggior rilievo per il contrasto al fenomeno, in ambito mondiale, è il "Patto internazionale sui diritti civili e politici" (adottato dall'Assemblea Generale ONU il 16 dicembre 1966 ed entrato in vigore il 23 marzo 1976), il cui art. 20 vieta "Qualsiasi appello all'odio nazionale, razziale o religioso che costituisca incitamento alla discriminazione, all'ostilità o alla violenza deve essere vietato dalla legge", con una definizione ampia che però può anche portare ad abusi nel contrasto, con riguardo alle limitazioni della libertà di espressione.

6. Il difficile bilanciamento tra il diritto alla libertà di espressione e contrasto ai discorsi di incitamento all'odio nella giurisprudenza della Corte di Strasburgo.

A tale ultimo proposito, dalla nutrita giurisprudenza della Corte europea dei diritti dell'uomo in materia risulta in modo emblematico come

sia difficile effettuare in sede giudiziaria il contrasto al fenomeno dell'incitamento all'odio.

La Corte EDU, muovendo dalla premessa dall'inesistenza di una definizione universalmente riconosciuta dell'espressione "discorsi dell'odio", con consolidato indirizzo afferma che nel contrasto al fenomeno dell'incitamento all'odio si deve effettuare un adeguato bilanciamento tra la tutela delle vittime delle espressioni offensive dell'incitamento all'odio reale e serio all'estremismo (come tali, manifestazioni di razzismo, xenofobia, antisemitismo, nazionalismo aggressivo, misoginia, discriminazione nei confronti delle minoranze e degli immigrati) e i diritti dei singoli, specialmente giornalisti e uomini politici, di esprimersi liberamente e di "offendere, scioccare o turbare" gli altri (si veda *Handyside c. Regno Unito*, 7 dicembre 1976, § 49).

Tale bilanciamento può, ovviamente, portare a soluzioni differenti, caso per caso.

Comunque la Corte ribadisce che:

a) "la tolleranza e il rispetto della dignità di tutti gli esseri umani costituiscono il fondamento di una società democratica e pluralista", sicché, "in via di principio, si può considerare necessario, nelle società democratiche, sanzionare e cercare di prevedere tutte le forme di espressione che diffondono, incitano, promuovono o giustificano l'odio basato sull'intolleranza ..." (*Erbakan c. Turchia*, sentenza del 6 luglio 2006, §56);

b) la Convenzione europea dei diritti dell'uomo si fonda su valori democratici e su un sistema istituzionale che ripudia l'estremismo.

La Corte, inoltre, ha identificato un certo numero di forme di espressione che devono essere considerate offensive e contrarie alla Convenzione (specialmente il razzismo, la xenofobia, l'antisemitismo, il nazionalismo aggressivo, la discriminazione nei confronti delle minoranze e degli immigrati). In questa ottica il Comitato dei Ministri del Consiglio d'Europa, ha adottato la Raccomandazione n. R 97 (20) sui "discorsi dell'odio".

In estrema sintesi, la Corte EDU ha stabilito alcuni parametri che permettono di qualificare un "discorso dell'odio" al fine di escluderlo da ogni tutela dalla libertà di espressione (art. 10 della Convenzione) o ancora dalla libertà di riunione e associazione (art. 11 della Convenzione).

Tale esclusione viene effettuata secondo due modalità previste dalla Convenzione:

a) applicando l'art. 17 (divieto dell'abuso di diritto) della Convenzione quando lo scopo è l'odio e costituisce una negazione dei valori fondamentali sanciti dalla Convenzione;

b) applicando le limitazioni previste al paragrafo 2 degli artt. 10 e 11 della Convenzione, strada prescelta quando il discorso, anche se d'odio, non è distruttivo dei valori fondamentali su cui si basa la Convenzione.

L'art. 17 della Convenzione ha lo scopo di evitare che strumentalizzando i principi convenzionali si possano svolgere attività o compiere atti che mirino alla distruzione dei diritti e delle libertà sanciti dalla Convenzione stessa.

Con riguardo all'applicazione dell'art. 17 cit. – considerato di carattere eccezionale – va rilevato che, dal quadro complessivo della giurisprudenza della Corte relativa al rapporto fra gli artt. 10 e 17 CEDU è possibile, ad esempio, ipotizzare che non è tanto il carattere diretto o indiretto, esplicito o implicito del fenomeno negazionista di un genocidio a determinare l'applicabilità dell'art. 17 CEDU, quanto il fatto che sia «immediatamente chiaro» ed evidente il fine non consentito cui l'espressione esaminata è diretta.

Così, con la decisione Dieudonné M'Bala M'Bala c. Francia del 20 ottobre 2015, la Corte EDU ha dichiarato irricevibile il ricorso presentato Dieudonné M'Bala M'Bala – un comico con attività politiche, che era stato condannato in sede penale per pubblici insulti diretti ad una persona o gruppo di persone sulla base della loro origine o appartenenza ad una certa comunità etnica, nazione, razza o religione, nella specie persone di origine o fede ebraica – che era diretto a ricomprendere la condotta posta a base della condanna penale nell'ambito applicativo dell'art. 10 CEDU, che tutela la libertà di espressione. In particolare, in data 26 dicembre 2008, alla fine di uno spettacolo allo “Zenith” di Parigi, Dieudonné M'Bala M'Bala invitò Robert Faurisson, un accademico che ha ricevuto diverse condanne in Francia per le sue opinioni negazioniste e revisioniste, principalmente per la sua negazione dell'esistenza delle camere a gas nei campi di concentramento, a raggiungerlo sul palco per ricevere un “premio per infrequentabilità e insolenza”. Il premio, che aveva la forma di un candeliere a tre bracci con una mela su ogni braccio, gli fu conferito da un attore con indosso ciò che fu descritto come un “indumento di luce” – un

pigiama a righe con cucita una stella gialla recante la scritta “ebreo” – che quindi recitava la parte di un deportato ebreo in un campo di concentramento.

Invece la Grande Camera nel caso *Perinçek c. Svizzera* – relativo ad un uomo politico turco, laureato in giurisprudenza, che, partecipando ad alcune conferenze in Svizzera, aveva negato pubblicamente che il massacro e le deportazioni del popolo armeno, compiute dall’Impero ottomano a partire dal 1915, costituiscano un genocidio – ha dato prevalenza alla libertà di espressione (sentenza 15 ottobre 2015).

Le pronunce più significative rese sui “discorsi dell’odio” si riferiscono ai temi dell’odio razziale, religioso, d’orientamento sessuale e negazionista e, relativamente ai regimi totalitari, alla politica e all’odio nazionale e incostituzionale e seguono il medesimo andamento estremamente “casistico”⁵.

7. La Corte di giustizia UE.

La Corte di giustizia UE, dal canto suo, si è occupata della questione principalmente con riguardo all’esatta applicazione dell’art. 22-bis della direttiva 3 ottobre 1989, 89/552/CEE, relativa al coordinamento di determinate disposizioni legislative, regolamentari e amministrative degli Stati membri concernenti l’esercizio delle attività televisive, come modificata dalla direttiva 30 giugno 1997, 97/36/CE nonché dell’art. 3 della Direttiva 2000/31 relativa a taluni aspetti giuridici dei servizi della società dell’informazione, in particolare il commercio elettronico, nel mercato interno («Direttiva sul commercio elettronico»). Articolo quest’ultimo che consente ad uno Stato membro di limitare la libera circolazione di un determinato servizio della società dell’informazione proveniente da un altro Stato membro, con specifico e proporzionato provvedimento, se lo considera necessario per una delle ragioni elencate fra cui: ordine pubblico, in particolare per l’opera di prevenzione, investigazione, individuazione e perseguimento in materie penali, quali la tutela dei minori e la lotta contro l’incitamento all’odio razziale, sessuale, religioso o etnico, nonché contro le violazioni della dignità umana della persona; tutela della sanità pubblica; pubblica sicurezza, compresa la salvaguardia della sicurezza, e della difesa

⁵ A. Mascia, *Discorsi dell’odio secondo la giurisprudenza della Corte europea dei diritti dell’uomo*, www.antonellamascia.wordpress.com 4 maggio 2014.

nazionale (vedi: sentenza 20 dicembre 2017, causa C-434/15, Asociación Profesional Elite Taxi c. Uber Systems SpainSL).

Nelle relative decisioni la CGUE ha, fra l'altro, stabilito che:

- il controllo sull'applicazione del diritto dello Stato membro di origine che si applica alle trasmissioni televisive e sull'osservanza delle disposizioni della direttiva compete solo allo Stato membro dal quale partono le trasmissioni e che lo Stato membro di ricezione non è autorizzato ad esercitare un proprio controllo per ragioni attinenti ai settori coordinati dalla direttiva (vedi, in tal senso, sentenze 10 settembre 1996, causa C-11/95, Commissione/Belgio, punti 34 e 86, nonché De Agostini e TV-Shop, punto 27);

- la direttiva non contiene alcuna definizione dei termini di cui all'art. 22-bis della stessa;

- i lavori preparatori delle direttive 89/552 e 97/36 non contengono indicazioni pertinenti circa il senso e la portata della nozione di «incitamento all'odio» e confermano che all'art. 22-bis della direttiva il legislatore dell'Unione ha voluto prevedere un motivo di divieto basato sull'ordine pubblico diverso da quelli intesi, in particolar modo, alla tutela dei minori;

- ne consegue che la portata dell'art. 22-bis della direttiva deve essere determinata conformemente al senso abituale nel linguaggio corrente dei termini impiegati in tale articolo, tenendo conto del contesto in cui essi sono utilizzati e degli scopi perseguiti dalla direttiva (vedi: sentenze 22 settembre 2011, cause riunite C-244/10 e C-245/10, punti 39 e ss. nonché 10 marzo 2005, causa C 336/03, punto 21 e giurisprudenza ivi citata);

- per quanto concerne le parole «incitamento» e «odio», occorre rilevare che esse prendono in considerazione, da una parte, un'azione destinata ad orientare un determinato comportamento e, dall'altra, un sentimento di animosità o di ripulsa contro un insieme di persone;

- così la direttiva, attraverso l'impiego della nozione di «incitamento all'odio», ha per scopo di prevenire qualsiasi ideologia irrispettosa dei valori umani, segnatamente iniziative che praticano l'apologia della violenza con atti terroristici contro una comunità determinata di persone;

- pertanto il motivo di divieto, desunto dal “pregiudizio alla comprensione fra i popoli”, si può considerare incluso nella nozione di «incitamento all'odio basato su differenze di razza, sesso, religione o

nazionalità» ai sensi della direttiva e, conseguentemente, rientra nei settori coordinati da quest'ultima;

- la Mesopotamia Broadcast e la Roj TV che contribuiscono, secondo il giudice del rinvio tedesco, ad attizzare gli scontri violenti tra le persone di etnia turca e curda in Turchia e ad esacerbare le tensioni tra i Turchi e i Curdi che vivono in Germania ed arrecano quindi pregiudizio alla comprensione fra i popoli (sentenza 22 settembre 2011, cause riunite C 244/10 e C 245/10, punto 43);

- quindi, il citato art. 22-bis non osta a che uno Stato membro prenda, in applicazione di una normativa generale in materia di associazioni, misure nei confronti di un ente di radiodiffusione televisiva stabilito in un altro Stato membro, per il motivo che le attività e gli obiettivi del medesimo ente violano il divieto di arrecare “pregiudizio alla comprensione fra i popoli”, purché le suddette misure non impediscano – il che deve essere verificato dal giudice nazionale – la ritrasmissione propriamente detta sul territorio dello Stato membro di ricezione delle trasmissioni televisive realizzate dal suddetto ente a partire dall'altro Stato membro;

- la direttiva non osta a che uno Stato membro, applicando una normativa generale relativa alla tutela dei consumatori contro la pubblicità ingannevole, adotti provvedimenti nei confronti di un inserzionista a motivo di una pubblicità televisiva trasmessa da un altro Stato membro, purché tali provvedimenti non impediscano la ritrasmissione in sé e per sé sul suo territorio delle trasmissioni televisive provenienti dall'altro Stato membro suddetto (sentenza 9 luglio 1997, cause riunite C-34/95, C-35/95 e C-36/95).

8. Il piano d'azione di Rabat.

Consapevole delle difficoltà di effettuare l'indicato bilanciamento – tra tutela delle vittime e tutela della libertà di espressione – l'ONU ha avviato negli anni una serie di consultazioni che sono sfociate nel piano d'azione di Rabat (Rabat è la città del Marocco dove fu approvato il documento), lanciato dall'Alto Commissario per i diritti umani dell'ONU nel 2013. L'iniziativa, ha cercato di spostare il dibattito su basi legali solide, chiarendo la portata degli obblighi di uno Stato ai sensi dell'art. 20 della Convenzione e proponendo un test per l'identificazione dei messaggi d'odio.

Nel piano Rabat si è stabilito che gli Stati dovrebbero: 1) includere nelle loro legislazioni nazionali precise definizioni dei termini-chiave come odio, discriminazione, violenza, ostilità, di cui all'art. 20 cit.; 2) abrogare la blasfemia e la diffamazione delle religioni, trattandosi di norme incompatibili con la libertà di espressione e di religione; 3) dotarsi di un sistema giudiziario diretto a garantire un'interpretazione uniforme delle norme in materia di incitamento all'odio e ad applicare le sanzioni criminali soltanto nei casi più gravi.

Il piano si concentra dunque sull'esigenza di una serie di risposte differenti ai discorsi d'odio, che non siano limitate alla sanzione criminale o alla semplice soppressione del discorso.

Tali risposte dovrebbero promuovere un dialogo interculturale, il pluralismo e la diversità e misure concrete a tutela delle minoranze e dei gruppi vulnerabili, sul presupposto secondo cui l'hate speech quasi sempre nasce a seguito di una precedente stigmatizzazione ed umanizzazione di un gruppo specifico ed è la conseguenza di conflitti tra gruppi o categorie differenti.

Il documento differenzia 3 tipi di hate speech: 1) espressioni costituenti reato penale; 2) espressioni non costituenti reato penale ma illecito civile; 3) espressioni che non sono illecite ma sollevano problemi in termini di tolleranza, civiltà e rispetto per gli altri.

Secondo il Piano Rabat la definizione di "incitamento all'odio", quindi, dovrebbe essere il più possibile ristretta e precisa, limitandosi ai discorsi d'odio pericolosi, cioè quelli che hanno elevata probabilità di catalizzare la violenza contro un gruppo.

Il rapporto Rabat propone anche di tenere in conto la popolarità dell'oratore (maggiore popolarità comporta maggiore diffusione del messaggio), lo stato emotivo degli ascoltatori, il contenuto (in particolare se è una vera e propria chiamata all'azione), il tipo di termini utilizzati.

9. *La difficile individuazione di strategie di contrasto.*

Dal piano Rabat risulta confermato che l'incitamento all'odio – come le discriminazioni – sono fenomeni connaturati alla storia dell'uomo, essendo la misoginia, il razzismo, l'omofobia etc. sentimenti radicati nella cultura mondiale.

Per questo gli unici strumenti di contrasto non possono essere la rimozione dei singoli contenuti e/o l'irrogazione di sanzioni penali, in

quanto è la mancanza di una sanzione sociale dell'aggressività che rende più facile diffondere i discorsi d'odio in televisione o sui giornali o anche online.

Infatti, la rimozione e le sanzioni penali – il cui utilizzo deve comportare la possibilità per le vittime di difendersi adeguatamente, non solo in sede giudiziaria – servono a punire comportamenti “patologici”, che dovrebbero rappresentare delle eccezioni, mentre si deve puntare alla diffusione di prassi e comportamenti “fisiologicamente” corretti e rispettosi del principio di eguaglianza, andando ad analizzare le cause del fenomeno, muovendo dalla premessa che la differenza dell'attuale era tecnologica è rappresentata dal mezzo a disposizione.

Anche perché resta l'esigenza di non limitare la libertà di espressione, anche se in concreto conciliare le opposte esigenze di contrasto delle condotte patologiche e di tutela della libertà di espressione non è facile soprattutto oggi che, con l'utilizzazione di internet, la potenzialità della diffusione dei messaggi di odio è enormemente aumentata e, di conseguenza, è enormemente aumentata la diffusione delle c.d. fake news, bufale o false notizie (le quali che si differenziano dalle notizie vere, in quanto “non superano la verifica dei fatti”⁶), che spesso ne rappresentano il substrato e che, come risulta da un recente e documentato studio svolto dal Massachusetts Institute of Technology (MIT) – sulla base di un'osservazione effettuata nel periodo tra il 2006 e il 2017 – sono maggiormente “gradite” delle notizie vere.

La scelta di efficaci strategie di contrasto ai discorsi d'odio è ulteriormente complicata dal fatto che la finalità e le modalità con le quali si utilizzano le notizie false ed altri stratagemmi in rete non sono sempre illecite.

Infatti, il dissenso, la contestazione, la protesta di piazza, sono spesso attuate tramite discorsi di contestazione ed espressioni offensive, che sono generalmente anche l'espressione di un malcontento, di un qualcosa che non funziona nella società. Tutto questo non può, di per sé, essere relegato nella definizione di hate speech in quanto la democrazia non può vivere nascondendo i disagi sociali di parte della popolazione⁷.

⁶ Lo studio – dal titolo *The spread of true and false news online* – è stato pubblicato su Science il 9 marzo 2018. Per un commento vedi: AM TESTA, *L'amara verità sulle notizie false*, in www.internazionale.it 12 marzo 2018

⁷ G. LOVINK, *L'abisso dei social media: Nuove reti oltre l'economia dei like*, la Feltrinelli, 2016

Per questo negli USA, pur non dubitandosi della dannosità per le persone dell'hate speech, tuttavia si teme maggiormente la censura, nella consapevolezza che occorra una particolare attenzione nel regolamentare i discorsi d'odio.

Resta da considerare che negli ultimi tempi si vanno diffondendo invenzioni di “bufale” effettuate con scopi tutt'altro che nobili e molto inquietanti: 1) con movente economico, come, ad esempio, può accadere con atti di concorrenza sleale che si sostanziano nella diffusione di notizie errate sui prodotti di imprese concorrenti; 2) con finalità di tipo politico, consistenti nel veicolare facili e immotivati allarmismi riguardanti temi come la salute, le crisi economiche e i migranti onde far crescere, in termini di popolarità, quei partiti che si fanno propugnatori di misure drastiche per farvi fronte.

10. Le bugie sono più potenti della verità.

In questa complessa situazione il dato che può considerarsi maggiormente inquietante è che, come risulta dal citato studio del MIT, si desume che le notizie false si diffondono sempre molto più rapidamente, più ampiamente, più profondamente (cioè, con catene di retweet lunghe il doppio, e dieci volte più veloci) di quelle vere e che questo accade per tutte le categorie di informazione, ma è particolarmente accentuato per l'informazione politica

In altri termini le bugie sono più potenti della verità e questo non dipende dagli algoritmi perché sono gli esseri umani a incrementarne la diffusione.

Ciò accade perché le notizie false attivano emozioni più potenti – prime tra tutte, paura e disgusto – e anche perché sono più “nuove” e, quindi, non solo suscitano maggior curiosità e sorpresa, ma appaiono anche più utili a capire il mondo e a prendere decisioni, più prestigiose da diffondere e dunque dotate di maggior valore.

Di questo si ha conferma anche nei già citati gli studi specifici sull'incitamento all'odio online, come il Progetto UMATI effettuato in Kenya, il piano RABAT, nonché nella relazione finale (approvata il 6 luglio 2017) – “La piramide dell'odio in Italia” – della Commissione “Jo Cox”⁸

⁸ Jo Cox era una parlamentare laburista britannica schierata contro BREXIT, che è stata uccisa giovedì 16 giugno 2016 a Birstall, nel nord dell'Inghilterra, da un uomo l'ha aggredita in strada sparandole e ferendola ripetutamente con un coltello, il quale secondo alcuni testimoni prima e durante l'attacco alla

dalla quale sono emersi dati molto interessanti e concordanti con quelli del progetto UMATI nel senso che i discorsi di odio online vanno posti in relazione con quanto si verifica offline.

In sintesi in tutti gli studi più accreditati – a partire da quello del MIT – si sostiene che l’approccio di tipo sociologico, psicologico – spesso poco praticato – è l’unico che, nel medio-lungo periodo, potrebbe portare a dei risultati positivi. Del resto, questo è confermato anche dalla “Lettera aperta al Presidente della Repubblica” Sergio Mattarella scritta il 6 febbraio 2019 dagli psicoanalisti appartenenti alla storica Società Psicoanalitica Italiana (SPI) e pure dal coevo Documento dell’Associazione Italiana degli Psicologi (AIP), entrambi aventi ad oggetto l’analisi delle implicazioni di tipo psichiatrico, psicologico e psicosociale sull’intera società delle disposizioni in materia di immigrazione contenute nel c.d. decreto sicurezza.

Questo tipo di approccio equivale ad un “cambiamento di rotta” e, diversamente da quanto comunemente si fa nei dibattiti in argomento, a non concentrarsi sulla criminalizzazione del mezzo (internet) – come se la sua velocità di diffusione fosse la causa scatenante dell’hate speech – ma sull’analisi delle vere cause del problema.

Infatti in tale prospettiva è pacifico che si deve puntare sull’educazione delle nuove generazioni e, al tempo stesso, sulla lotta alla solitudine degli adulti e sulla diffusione di relazioni interpersonali – e interculturali – solidali.

11. La reazione dell’Europa.

In Europa – dove, a causa delle tristi esperienze storiche di nazismo, fascismo e comunismo, la soppressione dei discorsi d’odio è sentita come un modo per promuovere la democrazia – il dibattito sul tema dell’hate speech online si è molto intensificato nell’ultimo decennio e sono state adottate molteplici iniziative.

Così, da un lato, il Consiglio d’Europa ha portato avanti la lotta a ogni forma di discriminazione e di espressione d’odio, nel pieno rispetto delle libertà fondamentali di opinione e di espressione, anche attraverso specifiche risoluzioni dell’Assemblea parlamentare, relative alle sfide e alle

Cox, ha gridato “Britain first!”, il nome di un partito di destra e anti-europeista e, più in generale, uno slogan usato per esprimere posizioni nazionaliste.

responsabilità dei media e del giornalismo online, alla cyber-discrimination e all'odio online.

Mentre per quanto attiene all'Unione europea, si è avuta l'adozione da parte della Commissione, a maggio 2016, del Codice di condotta sull'illecito incitamento all'odio online, con il quale le aziende informatiche si sono impegnate ad affiancare la Commissione e gli Stati membri nell'affrontare la sfida di garantire che le piattaforme online non offrano opportunità di diffusione "virale" di forme illegali di incitamento all'odio.

L'idea di fondo è che l'Unione europea si basa su una serie di valori: rispetto della dignità umana, libertà, democrazia, uguaglianza, Stato di diritto, diritti fondamentali. L'UE e i suoi Stati membri, insieme alle società di social media e altre piattaforme, devono adoperarsi affinché internet non diventi un rifugio sicuro per l'illecito incitamento all'odio e la violenza.

Ma alcuni rilevano che la delega da parte della politica alle società tecnologiche comporta un problema, poiché in molti casi viene dato per scontato che il pregiudizio e l'odio siano dei comportamenti devianti, mentre gli studi in materia contestano questa ipotesi, come si è detto.

Altri sottolineano che il Codice di condotta è stato elaborato dall'UE assieme alle grandi compagnie digitali, ma non assieme alle testate giornalistiche e televisive, esistendo notevoli divergenze fra gli Stati membri con riguardo alle modalità di repressione dei discorsi d'odio online nonché su quelli promossi dai mezzi di informazione tradizionali.

12. La risposta giurisdizionale.

Come si è detto, la risposta giurisdizionale per contrastare l'hate speech online è resa particolarmente difficile per il fatto che un'ulteriore caratteristica dell'hate speech online e degli altri fenomeni ad esso collegati (a partire dalle notizie false) è quella della trans nazionalità.

Al riguardo va, in primo luogo, sottolineato che nell'ambito delle espressioni d'odio i confini sono sempre molto labili e, oltretutto, variano a seconda dello Stato nonché, in molti casi, anche a seconda del Continente.

Un esempio classico è la diversa disciplina vigente nei diversi Stati europei nonché in Europa, nel suo complesso, rispetto a quella degli Stati Uniti d'America.

In particolare, benché le grandi aziende del web siano americane, tuttavia la maggior parte degli utilizzatori di social network e portali web non sono americani, pertanto assumono una importante valenza sia

l'assenza di una univoca definizione di hate speech – categoria, peraltro, elaborata negli anni dalla giurisprudenza americana – che, a volte, ha riflessi anche all'interno degli USA, sia il fatto che negli Stati Uniti la condanna dello hate speech – sia sul piano giuridico che nelle conversazioni al bar – sta in un equilibrio elastico ma spesso problematico con la libertà di parola, principio tutelato dal Primo emendamento della Costituzione degli Stati Uniti (e fondante, con le sue regole, di ogni democrazia).

Ne deriva che il principale strumento di contrasto è rappresentato dalla diffusione dei metodi più efficaci per riconoscere l'informazione spazzatura e starne alla larga, visto che questo elemento è basilare.

Al riguardo si propone:

a) come metodo d'individuazione più semplice – perché richiede soltanto di fare attenzione all'URL del sito che si va ad aprire – quello relativo alla fonte delle false informazioni e/o dei discorsi d'odio, muovendo dal presupposto che spesso questi portali scelgono nomi simili a quelli di giornali autorevoli per indurre il lettore in inganno (*es. Rebubblica, Il Fatto Quotidiano, La Gazzetta della Sera, Il Messaggio ecc*);

b) di controllare la sezione “Chi siamo” del sito. Infatti i più onesti di questi blog dicono chiaramente che si tratta di satira;

c) in caso di notizia che riporta dichiarazioni sospette di un personaggio noto, di fare sempre fare un check, associando su Google la stessa frase alla persona cui è riferita per verificare se è riportata da altre testate, allo stesso modo si deve fare una analoga ricerca inversa per immagini su Google per verificare se una fotografia che appare in un sito sia stata pubblicata da altri portali e se riguardi veramente la notizia che ci interessa

d) come ulteriore metodo quello consistente nel controllare se il link postato conduce effettivamente al sito indicato. Molte di queste realtà fanno rimbalzare l'utente, infatti, da un portale all'altro per accalappiare più clic possibili.

Comunque, mantenere alta la soglia di attenzione – ed educare a farlo – non è sufficiente per combattere un fenomeno dilagante.

Per questo social network, colossi del web e Governi nel corso degli anni stanno mettendo a punto misure sempre più sofisticate per contrastare le fake news. Facebook si è rivolto agli utenti, chiedendo di segnalare tempestivamente le bufale e chi le produce. Google, invece, ha deciso di colpire le realtà che guadagnano alle sue spalle (con il metodo dei clic)

bloccando i propri banner posizionati all'interno dei blog sospetti. Tutte queste strategie si basano sul prevalente, se non esclusivo, uso degli algoritmi per selezionare i messaggi da eliminare.

Si tratta però di un approccio che va integrato con il coinvolgimento di uomini e donne esperti in informatica, onde evitare che si verifichino limitazioni della libertà di espressione quando non sono opportune.

Alcuni Paesi, come la Germania, hanno introdotto multe salatissime (fino a 50 milioni) per combattere i crimini legati ai messaggi di odio.

Eppure, benché la legge tedesca – nella quale si prevede che uno staff di 50 dipendenti del ministero della Giustizia vigili sull'applicazione delle nuove norme – sia la prima legge al mondo di questo tipo ed anche la più severa, va tuttavia osservato che poco dopo la sua entrata in vigore (ottobre 2017) e l'irrogazione delle prime sanzioni (a partire da novembre 2018) ai primi di gennaio 2019 la Germania ha subito il più grave cyber-attacco della sua storia, con diffusione dei dati sensibili ai danni di un migliaio di personalità tedesche di spicco, fra cui anche molti membri di partiti compresa la Cancelliera Angela Merkel.

L'autore di questo "attacco" pare sia stato un hater ("odiatore") di 20 anni che, nel confessare, ha detto di aver commesso il fatto perché alcune dichiarazioni delle personalità colpite lo avevano "fatto arrabbiare".

13. Combattere la solitudine e la marginalizzazione.

Anche questo caso tedesco dimostra come per contrastare discorsi d'odio e/o cyber-attacchi le sanzioni "esemplari" sono inutili o comunque insufficienti mentre la strategia più efficiente è composita, nel senso che comporta l'adozione di differenti misure⁹, tra loro coordinate e condivise almeno in ambito UE, come: 1) obblighi di trasparenza per le piattaforme online con riguardo ai criteri di valutazione (non basati solo sulle intelligenze artificiali e gli algoritmi) e ai provvedimenti adottati; 2) definizioni precise dei termini chiave fissate per legge; 3) previsioni normative specifiche e circostanziate delle limitazioni alla libertà di espressione; 4) adozione di una pluralità di risposte che non siano limitate alla semplice repressione e che siano proporzionate al tipo di contenuto; 5) sistemi giudiziari coordinati che garantiscano l'interpretazione uniforme delle norme, sia in ambito nazionale sia in ambito europeo.

⁹ Vedi, per tutti: B. SAETTA, *Contrastare l'hate speech online: questioni aperte e alcune proposte* in www.valigiablu.it 18 febbraio 2017, cit.

Comunque, se davvero si vuole arrivare a dei risultati soddisfacenti, almeno nel medio-lungo periodo, bisogna andare alla causa dei fenomeni e quindi promuovere il dialogo interculturale, il pluralismo, la tutela delle minoranze e delle diversità.

Questo vuol dire diffondere la solidarietà e rapporti sociali aperti e basati sulla fiducia reciproca, valori che sono alla base delle moderne democrazie.

Non va, infatti, dimenticato che spesso gli haters – cioè gli odiatori di professione che presidiano la rete – sono persone abbastanza istruite, ma sole, in quanto prive di una vita sociale, per questo sfogano la loro rabbia sulle piattaforme social, proprio come sembra sia accaduto nel cyber-attacco tedesco. Infatti, solitamente si diventa haters per noia o perché si vuole attenzione. In altre parole si cerca online quello che non si ha nella vita reale, ovvero delle relazioni con gli altri, per questo si prova felicità quando gli utenti reagiscono alle provocazioni, perché così si può continuare a conversare e/o ad essere al centro dell'attenzione¹⁰.

D'altra parte, anche i “proseliti” spesso sono persone sole e marginalizzate e, comunque, prive delle capacità critiche idonee a comprendere la reale finalità del messaggio d'odio e ad accorgersi che spesso si basa su notizie false.

Quindi la prima cosa che tutti noi siamo chiamati a fare è cercare di combattere la solitudine e la marginalizzazione nella quale possono venire a trovarsi le persone.

Questo significa che per la “crescita” della UE e dei singoli Stati membri non si deve puntare solo sul PIL – che non misura l'aumento delle diseguaglianze e delle marginalizzazioni nei diversi territori – ma, secondo quel che si dice da anni in sede ONU, si deve puntare al reale benessere (welfare) della società, tenendo conto di indicatori come salute, conciliazione fra lavoro e tempi di vita, cultura, istruzione, sostenibilità ambientale, qualità dei servizi, del paesaggio e delle relazioni sociali: essendo questi i fattori che rendono la vita soddisfacente e degna di essere vissuta.

Siamo cioè chiamati a passare dalla attuale società basata sulla competizione esasperata ad una società basata sulla cooperazione, sulla fiducia reciproca e su relazioni solidali.

¹⁰ Vedi: *Come difendersi dai troll*, www.fastwebnet.it 2 Maggio 2017.

È la stessa “ricetta” che serve per combattere le discriminazioni e i messaggi d’odio (che ad esse si collegano), per la quale le scienze sociali propongono tre diversi appositi sistemi di riduzione dei pregiudizi, il primo dei quali è favorire gli scambi sociali cui si aggiungono agire attraverso l’educazione ed elevare il grado di consapevolezza sul problema del pregiudizio.

L’applicazione di tali strumenti ai troll e agli haters vale, ovviamente, quando si tratta di singoli individui e non nei confronti di realtà più complesse ed inquietanti che cercano di condizionare la vita politica di interi Paesi, perché in quest’ultimo caso l’azione di contrasto deve essere incisiva e possibilmente coordinata tra tutti gli Stati, salva la tutela delle vittime.

Comunque, già se si riuscisse a diffondere una maggiore solidarietà fra le persone si potrebbero avere dei buoni risultati e si verrebbe a realizzare un modello di società conforme a quello avuto di mira dai nostri Costituenti e anche dagli iniziatori del “progetto europeo”, oggi così fragile proprio su questo fronte.

In questo, il lavoro svolto dai giornalisti è di grande importanza, ma anche il ruolo dell’Avvocatura può essere relevantissimo perché, come ha detto il Presidente del CNF Andrea Mascherin del dare inizio al primo G7 dell’avvocatura nel settembre 2017 (iniziativa di cui il CNF è stato promotore) l’avvocato “può essere soggetto garante della pacifica convivenza per la nuova cittadinanza globale”.

Da tempo, del resto, l’Avvocatura italiana dimostra di avere piena consapevolezza della necessità di impegnarsi in prima linea per la salvaguardia dei diritti umani e fondamentali e attraverso una svolta “culturale” ma anche “etica” che riguarda il nostro Paese ma pure tutta l’Europa in attuazione non solo di quanto prescrive l’art. 4, comma secondo, della Costituzione italiana – secondo cui: “ogni cittadino ha il dovere di svolgere, secondo le proprie possibilità e la propria scelta, un’attività o una funzione che concorra al progresso materiale o spirituale della società” – ma anche in attuazione del Codice deontologico degli Avvocati europei e del Nuovo Codice deontologico degli Avvocati italiani, che ha fra i suoi obiettivi la valorizzazione della “funzione sociale della difesa”.

14. Conclusioni.

Se, come si è detto, è necessario diffondere i valori della solidarietà e dei rapporti sociali aperti e basati sulla fiducia reciproca vuol dire che bisogna riscoprire l'attualità e la centralità dei valori che sono alla base delle moderne democrazie.

La disgregazione degli interessi – contingenti e particolari – rispetto ai valori¹¹, alla quale abbiamo assistito negli ultimi decenni ha “offuscato” agli occhi della maggioranza dei cittadini italiani ed europei, scatenando nelle popolazioni egoismi, aggressività etc. e una diffusa sfiducia negli altri e nel futuro nonché il dramma della solitudine: tutti elementi che direttamente o indirettamente alimentano le discriminazioni, i discorsi d'odio e favoriscono il “gradimento” delle notizie false, grazie alla grande diffusione garantita da internet e al sostanziale abbandono di programmi educativi volti ad incoraggiare il pensiero critico, non solo nei bambini ma anche negli adulti.

In sintesi, siamo tutti chiamati a porre al centro delle condotte degli Stati e delle persone la regola dell'uguaglianza che implica il rispetto della pluralità dei valori che rendono ricca la vita in una società.

Questo vuol dire che ognuno riconosce agli altri consociati pari dignità e quindi che alla base principio democratico vi è la solidarietà ed anche il divieto di discriminazioni (e, quello, dell'incitamento all'odio, che ne deriva), perché le discriminazioni si basano sulla negazione della suddetta pari dignità e sulla umiliazione dell'altro.

In questa direzione si pone anche la recente istituzione, da parte del Senato, di una Commissione straordinaria contro odio, razzismo e antisemitismo e istigazione all'odio e alla violenza proposta dalla Senatrice a vita Liliana Segre.

Questa Commissione speciale, infatti, ha il compito di “osservare, vigilare, studiare e proporre iniziative atte a contrastare eventi e manifestazioni di razzismo, antisemitismo, intolleranza, istigazione all'odio e alla violenza nei confronti di persone o gruppi sociali sulla base dell'etnia, la religione, la provenienza, l'orientamento sessuale, l'identità di genere o di altre particolari condizioni fisiche o psichiche”.

E nella mozione istitutiva si rileva, fra l'altro, l'assenza di una normativa ad hoc contro l'hate speech, e si riconosce l'esistenza di “una capillare diffusione attraverso vari mezzi di comunicazione e in particolare

¹¹ Secondo la felice opinione dello storico Giovanni De Luna.

sul web di fenomeni di odio, intolleranza, razzismo, antisemitismo e neofascismo, che pervadono la scena pubblica accompagnandosi sia con atti e manifestazioni di esplicito odio e persecuzione contro singoli e intere comunità”.

La Commissione, quindi, ha proprio l’obiettivo di dare un segnale da parte del Senato di consapevolezza della diffusione del fenomeno e anche di analizzarne le cause e contribuire alla relativa prevenzione.

Si tratta, quindi, di un approccio che è del tutto in linea con gli studi in materia.

Inoltre l’istituzione della Commissione ha un valore “simbolico” di grande rilievo visto che proviene dal Senato e questo valore come tale può contribuire a contrastare un fenomeno che, giorno dopo giorno, miete vittime tra gli “odiatori”, i seguaci e i soggetti colpiti.

Del resto, la nostra Costituzione è esplicita nel senso di avere di mira una società che vuole essere pluralista, sull’assunto secondo cui nella cultura altrui troviamo anche qualcosa di nostro¹².

L’istituzione della suddetta Commissione, come si desume anche dalla autorevolezza Senatrice a vita che l’ha proposta, si pone in linea con la volontà espressa dai nostri Padri Costituenti.

E rappresenta un’occasione per ricordarla.

¹² Vedi M. Graziadei cit.