

HATE SPEECH: l'odio nella Rete

Arianna Maceratini

ABSTRACT

Sebbene l'odio e le sue espressioni abbiano accompagnato la storia dell'uomo sin dai tempi più remoti, le peculiarità di Internet introducono alcune caratteristiche di amplificazione temporale e spaziale del fenomeno che portano il contesto virtuale a rappresentare un fertile terreno al dilagare dell'hate speech, fake news e idee estremiste e persecutorie, rendendo arduo il rinvenimento di un'informazione corretta e oggettiva. Alla difficoltà di offrire una definizione univoca e comprensiva dell'hate speech si aggiunge, poi, una diversa sensibilità degli ordinamenti giuridici sul tema, nella comune esigenza di gestire il rapporto, in tensione, tra libertà di espressione e confliggenti diritti costituzionalmente protetti. Una questione critica centrale è data dalla prospettiva di una regolazione del dibattito online, effettuata dai pubblici poteri o da soggetti privati professionalmente coinvolti nella diffusione dei contenuti virtuali, facendo emergere il pericolo di un arbitrario controllo del pensiero ed aprendo ad un intenso confronto sulla responsabilità dei fornitori di servizi digitali. La circostanza che proprio dal Web provengano molteplici segnalazioni e richieste di rimozione dei post incitanti all'odio sembra, in ogni caso, far ripensare il contrasto all'hate speech, al di là delle possibili misure repressive, per introdurre interventi di segno positivo, di supporto delle vittime, iniziative di attivismo digitale e counterspeech. Si porta all'attenzione, inoltre, l'indispensabilità di una complessa opera di educazione all'impiego delle più recenti tecnologie, essenziale a scongiurare ogni rischio di paternalismo informatico e correlata ad una formazione linguistica capace di valorizzare i principi della cittadinanza e della democrazia, favorendo l'apertura del dibattito pubblico con il diffondere gli anticorpi dell'odio virtuale. Le questioni affrontate sollecitano, infine, una profonda riflessione che investe la centralità della persona e dei suoi dati elettronici nonché il ruolo dei social network e degli intermediari virtuali, mettendo in luce il rischio di una collaterale censorship conferita a soggetti che, per vocazione specifica, esercitano attività di natura commerciale, un tema che rimanda alla complessità dei rapporti tra poteri pubblici e privati e ai loro riflessi sul concreto esercizio della libertà di espressione.

Although hatred and its expressions have accompanied the history of man since the most remote times, the peculiarities of the Internet introduce some characteristics of temporal and spatial amplification of the phenomenon which lead the virtual context to represent fertile ground for the spread of hate speech, fake news and extremist and persecutory ideas, making it difficult to find correct and objective information. Adding to the difficulty of offering a univocal and comprehensive definition of hate speech is the different sensitivity of the legal systems on the topic, in the common need to manage the relationship, in tension, between freedom of expression and conflicting constitutionally protected rights. A central critical issue is given by the prospect of regulation of the online debate, carried out by public authorities or by private entities professionally involved in the diffusion of virtual contents, bringing to light the danger of an arbitrary control of the thoughts circulating online and opening up to an intense debate on responsibility of online service providers. The fact that multiple reports and requests for the removal of posts inciting hatred come from the Web seems, in any case, to make us rethink the fight against hate speech, beyond possible repressive measures, to introduce positive, supportive interventions of victims, digital activism initiatives and *counterspeech*. Furthermore, attention is drawn to the indispensability of a complex work of education in the use of the most recent technologies, essential to avoid any risk of IT paternalism and correlated to linguistic training capable of enhancing the principles of citizenship and democracy, favoring the opening of public debate by spreading the *antibodies* of virtual hatred. Finally, the issues addressed call for profound reflection on the centrality of the person and their electronic data as well as the role of social networks and virtual interme-

diaries, highlighting the risk of *collateral censorship* conferred on subjects who, by specific vocation, carry out activities of a commercial nature, a theme that refers to the complexity of the relationships between public and private powers and their effects on the concrete exercise of freedom of expression.

HATE SPEECH: UNA DEFINIZIONE IN EVOLUZIONE

Sebbene l'odio e le sue espressioni abbiano accompagnato la storia dell'uomo sin dai tempi più remoti¹, possedendo una "natura anfibia" in grado di attecchire e di svilupparsi negli ambienti più diversi², le peculiarità di Internet sembrano aver introdotto alcune caratteristiche di amplificazione temporale e spaziale del fenomeno, rappresentando il contesto virtuale un fertile terreno al dilagare dell'hate speech, *fake news* e idee estremiste e persecutorie³, capaci di inquinare le fonti⁴ e di consolidare stereotipi e pregiudizi, rendendo sempre più arduo il rinvenimento di un'informazione corretta e oggettiva⁵.

Senza dubbio non è agevole offrire una definizione univoca e comprensiva dell'hate speech⁶, efficacemente rappresentato come macroconte-

1 Cfr. ad. es. M. N. Campagnoli, *Hate Speech*, in A. C. Amato Mangiameli, G. Saraceni (a cura di), *Cento e una voce di informatica giuridica*, Torino, 2023, p. 253; per un *excursus* delle più significative rappresentazioni dell'odio in riferimento all'evoluzione del pensiero filosofico, si veda F. Cerquozzi, *Dall'odio all'hate speech. Conoscere l'odio e le sue trasformazioni per poi contrastarlo*, in "Tigor: Rivista di scienze della comunicazione e dell'argomentazione giuridica", X (2018), n. 1, pp. 43-45.

2 Cfr. B. G. Bello, *Discorsi d'odio e attivismi digitali: il veleno e i suoi antidoti*, in B. G. Bello, L. Scudieri (a cura di), *L'odio online: forme, prevenzione e contrasto*, Torino, 2022, p. 153.

3 Cfr. G. Ziccardi, *L'odio online. Violenza verbale e ossessioni in rete*, Milano, 2016, p. 15; cfr. F. Di Tano, *I reati informatici e i fenomeni del cyberstalking, del cyber bullismo e del revenge porn*, in T. Casadei, S. Pietropaoli, *Diritto e tecnologie informatiche*, Milano, 2021, p. 168.

4 Cfr. M. N. Campagnoli, *Informazione, social network & diritto*, Milano, 2020, p. 30.

5 Cfr. A. Di Rosa, *Hate speech e discriminazione. Un'analisi performativa tra diritti umani e teorie della libertà*, Modena, 2020, p. 14.

6 Cfr. ad. es. F. Abbondante, *Il ruolo dei social network nella lotta all'hate speech: un'analisi comparata fra l'esperienza statunitense e quella europea*, in "Informatica e diritto", XLIII (2017), n. 1-2, p. 41.

PAROLE-CHIAVE

HATE SPEECH; DIRITTO DI ESPRESSIONE; DIRITTI FONDAMENTALI; WEB; ISP; COUNTERSPEECH; EDUCAZIONE DIGITALE.

KEYWORDS

HATE SPEECH; RIGHT OF EXPRESSION; FUNDAMENTAL RIGHTS; WEB; ISP; COUNTERSPEECH; DIGITAL EDUCATION.

nitore dai contorni imprecisi e mutevoli, atto a raggruppare una varietà di esternazioni aggressive, denigratorie e discriminatorie⁷, una difficoltà di definizione aggravata dalla diversa sensibilità degli ordinamenti giuridici sul tema, nella comune esigenza di gestire il rapporto, spesso in tensione, tra libertà di espressione e confliggenti diritti costituzionalmente protetti⁸. In ogni caso, il termine hate speech è stato introdotto, alla fine degli anni '80, dalla Critical Race Theory per portare ad emergere il razzismo latente nella società americana e nel sistema giuridico di riferimento⁹, facendo rinvio alla distinzione, presente nella giurisprudenza statunitense, tra *speech* e *action*¹⁰. In senso lato, l'hate speech, tenuto conto dei differenti contesti discorsivi, si riferisce all'utilizzo di un linguaggio che "non solo veicola stereotipi negativi associati a gruppi vulnerabili, bensì anche propriamente e direttamente offensivo nei confronti di individui in virtù della loro appartenenza a un gruppo"¹¹, rivelandosi arduo da contrastare poiché obbliga le vittime e l'intera società civile a confrontarsi con termini non solo denigratori, ma, in un certo senso, performativi, "parole in atto" in grado di innescare un crescendo di violenza imprevedibile¹². Va, inoltre, notato come, in

7 Cfr. M. N. Campagnoli, *Hate Speech*, cit., p. 253.

8 Cfr. P. Falletta, *Controlli e responsabilità dei social network sui discorsi d'odio online*, in "MediaLaws – Rivista di Diritto dei Media", (2020), n. 3, p. 147.

9 Cfr. C. Bianchi, *Hate speech. Il lato oscuro del linguaggio*, Roma-Bari, 2021, p. 4.

10 Cfr. A. Di Rosa, *op. cit.*, p. 23.

11 Ivi, p. 11.

12 Il tema discusso non può che richiamare la teoria degli *atti linguistici* di J. L. Austin, secondo la quale, qui necessariamente in breve, il linguaggio non avrebbe l'esclusiva funzione di dichiarare qualcosa, potendo costituire, in alcuni casi, una vera e propria azione. Al riguardo, J. L. Austin, *How to Do*

alcuni casi, l'hate speech non assuma le connotazioni dell'odio esplicito, del disprezzo e dell'insulto insinuandosi in stereotipi culturali e sociali che, a partire dall'impiego di una forma apparentemente neutrale e benevola, si trasformano in pregiudizi e in etichette per gli individui e i gruppi che ne sono soggetti¹³, evidenziando le implicazioni della pragmatica linguistica in termini di oppressione, marginalizzazione ed esclusione¹⁴.

L'hate speech si declina, dunque, in molteplici direzioni accomunate dal suscitare una sempre maggior apprensione degli Stati democratici per le gravissime ricadute di tale fenomeno sul piano personale e sociale¹⁵, potendo contribuire alla delineazione di categorie sostenute da luoghi comuni che pongono le vittime dei discorsi d'odio in una posizione di inferiorità¹⁶, evidenziando come, in ogni situazione di "ingiustizia discorsiva", al di là di chi inciti all'odio, assuma un ruolo di rilievo anche l'astante indifferente e silenzioso, fornendo implicito consenso e legittimazione all'hate speech: discorso d'odio è, dunque, anche quello che riduce l'interlocutore al silenzio togliendogli visibilità, il ruolo di agente e, con essi, ogni diritto di parola e di rappresentanza¹⁷.

Nel tentativo di tracciare alcune linee di delimitazione, va innanzitutto menzionata la definizione del discorso d'odio effettuata dalla Raccomandazione del Consiglio d'Europa n. 20 del 30.10.1997 secondo la quale è considerata hate speech qualsiasi forma espressiva in grado di diffondere, incitare, promuovere o giustificare odio razziale, xenofobia, antisemitismo o altre forme di odio basate sull'intolleranza, compresa quella espressa da nazionalismo aggressivo ed etnocentrismo, la discriminazione e l'ostilità contro le minoranze, i migranti e le persone di

origine straniera. In tempi più recenti, seguendo il dettato della General Policy Recommendation No.15/2015 dell'European Commission against Racism and Intolerance, ECRI, *On combating hate speech*, la definizione del discorso d'odio appare ampliarsi sino ad includere forme apparentemente *politically correct*¹⁸, esplicitandosi nel "fomentare, promuovere o incoraggiare, sotto qualsiasi forma, la denigrazione, l'odio o la diffamazione nei confronti di una persona o di un gruppo, nonché sottoporre a soprusi, insulti, stereotipi negativi, stigmatizzazione o minacce una persona o un gruppo (...) sulla base della razza, del colore della pelle, dell'ascendenza, dell'origine nazionale o etnica, dell'età, della disabilità, della lingua, della religione o delle convinzioni, del sesso, del genere, dell'identità di genere, dell'orientamento sessuale e di altre caratteristiche o stato personale". Tale definizione è stata successivamente ripresa, in maniera pressoché letterale, dalla Commissione sull'intolleranza, la xenofobia, il razzismo e i fenomeni di odio, istituita nel maggio 2016 e intitolata a Jo Cox, deputata presso la Camera dei Comuni del Regno Unito, uccisa il 16 giugno 2016. Dalla relazione finale della suddetta Commissione, *La piramide dell'odio in Italia*, approvata nel luglio 2017¹⁹, si può desumere una definizione dell'hate speech capace di porsi ben oltre l'ambito delle espressioni penalmente perseguibili per estendersi ad ogni diffusione del discorso d'odio, seguendo una piramide comportamentale che, a salire, si concretizza in veri e propri reati. Significativamente poi, nel Documento conclusivo dell'indagine conoscitiva sulla natura, cause e sviluppi recenti del fenomeno dei discorsi d'odio, approvato nel giugno del 2022 dalla Commissione straordinaria per il contrasto dei fenomeni di intolleranza, razzismo, antisemitismo e istigazione all'odio e alla violenza, "Commissione Segre", si può notare come la tutela dai discorsi d'odio rappresenti una garanzia rivolta, innanzitutto, alle categorie tipicamente destinatarie dell'hate speech, focalizzando l'attenzione sulla posizione di svantag-

Things with Words (1962 I ed.), trad. it., *Come fare cose con parole*, Genova, 2019; con specifico riferimento al discorso d'odio, quale tipologia di linguaggio performativo, si rimanda a C. Bianchi, *op. cit.*, pp. 3-14 e 91-135 e a M. N. Campagnoli, *Informazione*, cit., pp. 67 sgg.

13 Cfr. F. Cerquozzi, *op. cit.*, p. 46.

14 Cfr. B. G. Bello, L. Scudieri, *Discorsi d'odio online. Spunti per un dibattito interdisciplinare*, in B. G. Bello, L. Scudieri (a cura di), *L'odio online*, cit., pp. 16-17. Sulla relazione tra il sistema linguistico e il suo uso e sulla rilevanza di un'appropriata educazione linguistica, si veda S. Cavagnoli, *Le parole fanno male. E anche le immagini*, in B. G. Bello, L. Scudieri (a cura di), *L'odio online*, cit., pp. 19-35.

15 Cfr. F. Di Tano, *op. cit.*, p. 168.

16 Cfr. A. Di Rosa, *op. cit.*, p. 15.

17 Cfr. S. Cavagnoli, *op. cit.*, pp. 23-25.

18 Cfr. F. Cerquozzi, *op. cit.*, p. 47.

19 https://www.camera.it/application/xmanager/projects/leg17/attachments/shadow_primapagina/file_pdfs/000/007/099/Jo_Cox_Piramide_odio.pdf

gio discorsivo assunta dalle vittime, sull'effetto di *silencing*²⁰ o effetto *chilling*, letteralmente agghiacciante, derivante dal discorso d'odio tutelando, nel proteggere i soggetti più esposti, l'intero contesto sociale. Nel suddetto Documento della Commissione Segre viene, pertanto, riportata la citata Raccomandazione di politica generale n. 15 dell'ECRI, esponendo idee e concetti ripresi anche dalla recente Raccomandazione del Comitato dei Ministri del Consiglio d'Europa sulla lotta all'hate speech del 20 maggio 2022 e dalle relative Linee Guida²¹.

La più attuale prospettiva sull'hate speech tende, pertanto, a porsi dal punto di vista delle vittime²², consentendo di ancorare la definizione del fenomeno "non al movente in sé del discorso, ma agli effetti che questo può avere, ai rischi che causa per le vittime, reprimendo quei comportamenti che abbiano come obiettivo ed effetto probabile quello di incitare, promuovere, diffondere o giustificare (e, in tal modo, incentivare) violenza, odio e discriminazione"²³. In tale direzione, intendere l'hate speech principalmente come una lesione della libertà di espressione di chi ne è sottoposto permette, altresì, di mettere in luce l'*under-reporting*, ovvero, il tipico effetto paralizzante che impedisce alle vittime di denunciare²⁴, portando ad emergere il lato sommerso del discorso d'odio, ovvero, tutti quei casi non denunciati alle forze dell'ordine per vergogna, mancanza di consapevolezza o paura di ritorsioni²⁵.

20 Cfr. M. N. Campagnoli, *Informazione*, cit., p. 69.

21 È del 25 ottobre 2022 il Report del Congresso dei Poteri Locali e Regionali del Consiglio d'Europa sull'hate speech che, seguendo le Linee Guida adottate dal Comitato dei Ministri il 20 maggio dello stesso anno, evidenzia l'impatto dei discorsi d'odio sul lavoro dei funzionari delle amministrazioni locali e sul corretto funzionamento delle istituzioni democratiche, cfr. F. Abbondante, *op. cit.*, p. 68.

22 M. Martorana, *Il complesso rapporto tra hate speech, libertà e democrazia. Un difficile bilanciamento tra le libertà fondamentali coinvolte e gli sforzi europei per il contrasto ai discorsi d'odio*, in <https://www.altalex.com/documents/news/2022/11/08/complesso-rapporto-hate-speech-liberta-democrazia>

23 Sull'hate speech e sulla discriminazione per motivi razziali si veda, ad esempio, la sentenza della Cassazione Penale, Sez. III, 14.09.2015, n. 36906 secondo la quale ha rilevanza penale la condotta di chi propaga idee fondate sulla superiorità o sull'odio razziale o etnico, ovvero, di chi istiga a commettere o commette atti di discriminazione per motivi razziali, etnici, nazionali o religiosi.

24 Cfr. M. Martorana, *op. cit.*

25 Cfr. G. Ziccardi, *op. cit.*, p. 35.

Va, infine, evidenziato come, nella prevalente qualificazione dell'hate speech, debbano ricorrere contemporaneamente tre elementi, ovvero, la volontà di incitare all'odio, l'istigazione effettiva, cioè idonea ad incentivare fenomeni di violenza e di discriminazione, la traduzione dell'odio in azione o in un rischio reale e imminente per la vittima, potendosi anche solo potenzialmente tramutare in azioni violente o in comportamenti discriminatori²⁶. Si può, dunque, ben comprendere come la vastità delle situazioni raffigurabili mediante i suddetti elementi, faccia propendere per l'adozione di un'elastica *definizione ombrello* del discorso d'odio, sottolineando l'elemento discriminatorio idoneo a ricomprendere comportamenti intenzionali e azioni più generiche nelle quali non è direttamente rinvenibile l'elemento psicologico dell'odio e che, pertanto, non sembrerebbero direttamente sanzionabili²⁷. Per tali ragioni, assumendo una definizione estensiva e in evoluzione dell'hate speech, confermata anche dal *UN Strategy and Plan of Action on Hate Speech*, adottato dalla Comunità internazionale nel 2019²⁸, si nota un progressivo ampliamento dei comportamenti in attenzione e, parallelamente, delle categorie protette dal principio di non-discriminazione, risultando evidente l'elemento comportamentale di collegamento, ovvero, la *negazione dell'umanità dell'altro*²⁹.

26 Cfr. *ivi*, p. 21; cfr. M. N. Campagnoli, *Hate Speech*, cit., p. 255.

27 Cfr. A. Di Rosa, *op. cit.*, p. 71 e p. 81. Nella direzione di rilevare il fattore discriminatorio, come elemento caratteristico dell'hate speech, si pone la Direttiva del Parlamento europeo e del Consiglio sui servizi di media audiovisivi, Direttiva 2010/13/ (Direttiva AVMS), rinviante all'art. 21 della Carta dei Diritti fondamentali dell'Unione Europea centrato sul principio di non-discriminazione <https://eur-lex.europa.eu/IT/legal-content/summary/audiovisual-media-services-directive-avmsd.html>, sull'argomento cfr. anche M. N. Campagnoli, *Hate Speech*, cit., p. 255.

28 Cfr. A. Di Rosa, *op. cit.*, p. 67. L'ICCPR all'art. 19 comma 3, in riferimento alla libertà di espressione, esprime specifiche condizioni di punibilità concernenti comportamenti che, genericamente, non sarebbero sanzionati, indicati nel 2° comma, mentre l'art. 20 della stessa si riferisce al divieto di incitamento all'odio.

29 Cfr. A. Di Rosa, *op. cit.*, p. 71. L'hate speech costituisce, quindi, una categoria più vasta rispetto a quella degli *hate crimes* essendovi ricompresi comportamenti non sempre punibili penalmente, cfr. *ivi*, p. 73.

LIBERTÀ DI ESPRESSIONE E HATE SPEECH

La libertà di espressione rappresenta, un aspetto cruciale delle democrazie contemporanee³⁰, essendo radicata nel tessuto costituzionale dei moderni ordinamenti giuridici quale indispensabile condizione per l'esercizio di ogni altra libertà fondamentale, nonché punto di intersezione di una pluralità di diritti più specifici³¹. Per quanto concerne il dibattito sulla libertà di espressione si può rilevare un dissimile orientamento tra le Corti europee e la Corte Suprema degli Stati Uniti che corrisponde a differenti interpretazioni e modalità nella tutela della libertà di manifestazione del pensiero e del diritto ad informare e ad essere informati. È noto, infatti, come gli Stati Uniti siano estremamente attenti nel garantire la competizione e il confronto tra idee mentre l'ottica europea sia più centrata nel bilanciamento di interessi pubblici e privati a protezione dei fondamentali diritti dei paesi democratici³². Naturalmente, "queste diverse sensibilità oggi sono sottoposte sempre più a confronto, alla luce dell'avvento di un mezzo come Internet che ha permesso di collegare utenti e con loro idee, opinioni e manifestazioni di pensiero"³³.

Nel contesto europeo, specificamente, la libertà di espressione del pensiero si relaziona ed è in rapporto di bilanciamento con eterogenei diritti costituzionalmente garantiti, venendo sottoposta a specifiche e precise limitazioni riconducibili, in primo luogo, all'esigenza di prevenirne eventuali abusi. In relazione a ciò, appare esemplare la tutela apprestata dall'art. 10 della CEDU il quale, dopo aver affermato il fondamentale rilievo della libertà di parola, stabilisce che le limitazioni alla stessa debbano essere legislativamente previste e proporzionali agli obiettivi esplicitati dal secondo comma della suddetta disposizione³⁴. A tal proposito, va rilevato come tale Convenzione sanzioni, in ogni

30 Sulle sfide poste dallo spazio virtuale alla libertà di espressione si rinvia a F. J. Ansuátegui Roig, *Libertà d'espressione: ragione e storia*, Torino, 2018, pp. 201 sgg.

31 Cfr. C. M. Reale, M. Tomasi, *Libertà d'espressione, nuovi media e intelligenza artificiale: la ricerca di un nuovo equilibrio nell'ecosistema costituzionale*, in "DPCE online", (2022), n. 1, pp. 325-326.

32 Cfr. G. Ziccardi, *op. cit.*

33 Cfr. G. Pitruzzella, O. Pollicino, S. Quintarelli, *Parole e potere. Libertà d'espressione, hate speech e fake news*, Milano, 2017, p. 3.

34 Cfr. *ivi*, p. 5.

caso, l'incitamento alla discriminazione e all'odio senza che ciò venga inteso come un'indebita limitazione della libertà di espressione³⁵ e come, nel corso del tempo, la stessa Corte Europea abbia offerto importanti indicazioni in materia di negazionismo, arricchendo quanto originariamente stabilito dalla CEDU con l'art. 17, disposizione che punisce l'abuso di un diritto, sebbene riconosciuto dalla Convenzione, laddove esso appaia strumentale al diniego o alla limitazione di altri diritti o libertà³⁶. In questo modo, la Corte Europea, come anche il Comitato per i diritti umani delle Nazioni Unite che ha qualificato il discorso d'odio quale *punishable abuse* della libertà di espressione³⁷, è ricorsa più volte proprio all'art. 17 nei casi di hate speech negando *ab origine* la possibilità di un bilanciamento con altri diritti, riservando una particolare attenzione "al dibattito pubblico e all'esistenza di un interesse generale, che legittimano e giustificano un grado di protezione più intenso rispetto a quelle manifestazioni di pensiero che rispondono esclusivamente all'etichetta di *commercial speech*"³⁸. Tali indicazioni - confermate anche dalla Decisione Quadro 2008/913/GAI in tema di hate speech³⁹, dalla Convenzione per la salvaguardia dei diritti dell'uomo e delle libertà fondamentali⁴⁰, nonché da diverse pronunce della Corte di giustizia UE sui sistemi di *screening* e filtraggio delle informazioni⁴¹ - appaiono dirette a tutelare la centralità della persona, facendo emergere la *funzione sociale* della libertà

35 Cfr. F. Cerquozzi, *op. cit.*, p. 47.

36 Cfr. G. Pitruzzella, O. Pollicino, S. Quintarelli, *op. cit.*, p. 7; cfr. A. De Rosa, *op. cit.*, pp. 45 sgg.

37 Sull'argomento, cfr. *ivi*, pp. 40-41.

38 G. Pitruzzella, O. Pollicino, S. Quintarelli, *op. cit.*, p. 8.

39 La Decisione Quadro 2008/913/GAI del Consiglio del 28.11.2008 sulla lotta contro talune forme ed espressioni di razzismo e xenofobia mediante il diritto penale costituisce il primo provvedimento centrato sul contrasto a specifiche forme espressive di razzismo e xenofobia con gli strumenti del diritto penale, sebbene la sua trasposizione, effettuata dagli Stati dell'Unione, abbia condotto a modalità differenziate e non sempre in reciproca coordinazione. In Italia, in seguito a tale decisione, è stato approvato il d.lgs. 21/2018, art. 7 che ha introdotto gli artt. 604-bis e 604-ter c.p. cfr. *ivi*, p. 22.

40 Cfr. M. Orofino, *La libertà di espressione tra Costituzione e Carte europee dei diritti. Il dinamismo dei diritti di una società in trasformazione*, Torino, 2014, p. 39 sgg.

41 Cfr. G. Pitruzzella, O. Pollicino, S. Quintarelli, *op. cit.*, p. 20; cfr. M. Bassini, *Libertà di espressione e social network, tra nuovi "spazi pubblici" e "poteri privati". Spunti di comparazione*, in "Rivista Italiana di Informatica e Diritto", (2021), n. 2, p. 47.

di espressione in un contesto democratico e genuinamente pluralista⁴².

L'approccio statunitense alla libertà di manifestazione del pensiero e, conseguentemente, in materia di hate speech, sembra caratterizzato da una maggiore sensibilità libertaria⁴³ rispetto al modello continentale: al tal proposito, l'ideale *marketplace of ideas* può essere considerato come il punto di arrivo di un percorso interpretativo compiuto dalla giurisprudenza costituzionale statunitense attraverso il quale si sono a lungo confrontate una concezione più tradizionalista nell'individuare alcuni limiti al *free speech* in circostanze comunicative ritenute idonee a determinare la commissione diretta o indiretta di atti illeciti, ed una visione più duttile, ritenente necessaria una pericolosità concretamente riferibile al discorso d'odio⁴⁴. Nel contesto statunitense, inoltre, la prevalente interpretazione dottrinale del Primo Emendamento, mediante i principi del *clear and present danger*⁴⁵, consente una tutela pressoché assoluta del diritto di parola, da cui deriva il divieto, per i poteri pubblici, di interferire con la *freedom of speech*⁴⁶, sebbene una parte della dottrina abbia messo in luce l'effetto *chilling* del discorso d'odio sulle vittime facendo emergere, nella difficoltà di queste ultime di partecipare efficacemente al dibattito pubblico, una tensione tra il 1 Emendamento e il 14 Emendamento, sull'uguaglianza⁴⁷. In tale confronto, le più significative pronunce sul Primo Emendamento sono

42 Cfr. G. Ziccardi, *op. cit.*, pp. 47-48.

43 Cfr. *ivi*, p. 48. "Le corti che si sono pronunciate nelle vicende in cui era in questione una possibile equiparazione di Internet (e dei social network in particolare) a un *public forum*, hanno opportunamente delimitato il campo di applicazione di questa analogia, evitando di trarne conseguenze che avrebbero verosimilmente presentato riflessi 'di sistema' ben più invasivi", *ibidem*.

44 Cfr. F. Abbondante, *op. cit.*, p. 44.

45 Cfr. *ibidem*. Sul principio del *clear and present danger* si veda anche G. Ziccardi, *op. cit.*, pp. 53-56.

46 Cfr. F. Abbondante, *op. cit.* p. 46. Ogni eventuale limitazione della libertà di manifestazione del pensiero sembra, dunque, ammissibile solo nei casi in cui venga dimostrato un *compelling interest* la cui valutazione di compatibilità sia soggetta ad uno *strict scrutiny*, cfr. *ibidem*.

47 Cfr. *ivi*, p. 47. Parimenti l'industria pornografica ridurrebbe le vittime, in questo caso donne, ad una condizione di sottomissione e di riduzione al silenzio avvicicabile, sotto molti aspetti, al discorso d'odio. Su questo tema e per una critica alla pornografia centrata sulla difesa dei diritti civili delle donne, a partire dall'impiego di un linguaggio adeguato, si rimanda a C. Bianchi, *op. cit.*, pp. 43-90.

riferibili alla Corte Suprema la quale, evidenziando la necessità di un nesso tra parola e azione⁴⁸, sembra prevedere un margine di tutela particolarmente ampio per le forme di espressione del dissenso politico richiedendo, come condizione di punibilità, gli elementi della condotta illecita imminente, elemento temporale, dell'elevata probabilità del verificarsi, elemento probabilistico, infine, dell'intenzione di provocare la condotta, elemento finalistico-psicologico⁴⁹.

Le criticità emergenti dai differenti orientamenti, europeo e statunitense, qui solo tratteggiate, vanno, poi, ricondotte alla Rete, ambito transnazionale per definizione, capace di porre in costante confronto eterogenei sistemi giuridici favorendo i contesti territoriali disponibili ad una maggior protezione delle espressioni d'odio, talvolta aggirando le regole sulla giurisdizione⁵⁰. "Non è casuale che si discuta, a proposito, di *forum-shopping*, ossia della possibilità di pubblicare determinati contenuti nel paese il cui ordinamento garantisca il livello di tutela più ampio. Internet rischia di alimentare un corto circuito in cui sarebbe impossibile decifrare lo standard di tutela applicabile ai vari fenomeni che prendono corpo in rete, determinando uno scenario in cui la discrezionalità dei giudici costituirebbe l'ago della bilancia per risolvere i conflitti costituzionali destinati ad emergere"⁵¹. In definitiva, in assenza di uno standard comune sullo *speech on line*, la determinazione di ciò che è lecito o meno, verrebbe rimessa alla discrezionalità delle Corti di giustizia, in riferimento alle quali si è già rilevata tutta la distanza sussistente tra l'orientamento europeo e statunitense.

L'ODIO ONLINE

Pur nell'evidente indistinzione *onlife* del mondo fisico da quello analogico⁵², le condotte riconducibili all'hate speech, appoggiandosi alle

48 Cfr. F. Abbondante, *op. cit.*, pp. 47-48.

49 Cfr. G. Pitruzzella, O. Pollicino, S. Quintarelli, *op. cit.*, pp. 24-27.

50 Cfr. G. Ziccardi, *op. cit.*, p. 220.

51 G. Pitruzzella, O. Pollicino, S. Quintarelli, *op. cit.*, p. 35.

52 Il rimando è, naturalmente, a L. Floridi, *La quarta rivoluzione. Come l'infosfera sta trasformando il mondo*, Milano, 2017.

peculiarità di Internet, si differenziano da simili comportamenti adottati offline per alcune caratteristiche della Rete, quali la diffusione virale delle notizie, capace di generare un inedito *effetto doppler* informativo⁵³, la persistente accessibilità di ciò che viene pubblicato⁵⁴, la sostanziale impossibilità di effettuare una verifica delle fonti ed un effettivo controllo della propagazione delle notizie⁵⁵, potendo i contenuti incriminati trovare diffusione nella medesima piattaforma o altrove con un nome diverso, seguendo meccanismi *itineranti* di difficile ricostruzione. A tali incognite si aggiungono la presunzione dell'anonimato combinata alla possibilità di delineare false identità virtuali⁵⁶ e alla sovraesposizione delle informazioni personali⁵⁷ che espongono l'utente ai rischi di una falsa rappresentazione di sé e di sottomissione ad una pervasiva profilazione elettronica, un *social sorting* capace di incrementare e radicalizzare l'odio online⁵⁸.

Alle tradizionali caratteristiche forme d'odio, riconducibili a gruppi o minoranze portatrici di una qualche "diversità" è da aggiungere un ulteriore aspetto tipico dell'odio online, ovvero, la sua *canalizzazione interpersonale*, cioè l'impiego di espressioni dirette a colpire non tanto alcune categorie *target*, bensì soggetti determinati, spesso individuati per caratteristiche fisiche o di genere⁵⁹. L'attenzione si rivolge, infine, agli impieghi sempre più estesi degli algoritmi sia in discorsi generati automaticamente - principalmente nel contesto politico-elettorale con *bots* capaci di distorcere la realtà e di modulare la pubblica opinione⁶⁰ - sia come filtro della comunicazione virtuale, in grado di alimentare la *clusterizzazione* e la *polarizzazione* informativa⁶¹, realizzando un

53 Cfr. M. N. Campagnoli, *Informazione*, cit., p. 71.

54 Cfr. F. Di Tano, *op. cit.*, p. 169; cfr. S. Cavagnoli, *op. cit.*, p. 31.

55 Cfr. M. N. Campagnoli, *Informazione*, cit., pp. 24 sgg.

56 Per un esame delle principali differenze tra le manifestazioni di odio online e offline, cfr. G. Ziccardi, *op. cit.*, pp. 78-81.

57 Cfr. B. G. Bello, L. Scudieri, *op. cit.*, p. 3.

58 Cfr. G. Ziccardi, P. Perri, *L'odio online tra profilazione, big data e protezione dei dati personali*, in B. G. Bello, L. Scudieri (a cura di), *L'odio online*, cit., pp. 92-93.

59 Cfr. G. Ziccardi, *op. cit.*, p. 232 sgg; cfr. B. G. Bello, L. Scudieri, *op. cit.*, pp. 4-6.

60 Cfr. C. M. Reale, M. Tomasi, *op. cit.*, p. 328.

61 Per un primo approccio al concetto della *filter bubble* e all'inquietante idea delle *echo chambre* si veda C. R. Sunstein, *#republic*, Bologna, 2017, Z. Bauman, *Consumo, dunque sono*,

pluralismo ovattato⁶² che radicalizza i conflitti. Su questo tema, possiamo ricordare il meccanismo di selezione e di oscuramento dei contenuti nelle piattaforme digitali attuato tramite lo *Shadow ban* - termine riferito ad un'azione di moderazione che consente di escludere e celare un preciso utente da una comunità online o di rendere invisibili i contenuti da questi pubblicati ad altri utenti - impiegato nel limitare la visibilità di determinati materiali sulla base di criteri non esplicitati, impedendone la diffusione⁶³. Di questi aspetti problematici occorre, dunque, tener conto nel tratteggiare l'ipotesi di un "libero mercato delle idee" nella Rete che, in modo ben diverso, attribuisce al fruitore finale dell'informazione virtuale un onere di vaglio e di approfondimento conoscitivo assai arduo da adempiere, considerata la vastità delle fonti e delle informazioni⁶⁴. Di conseguenza, "il rischio è che il mondo della comunicazione e dell'informazione si automatizzi, alternativamente si standardizzi o si frammenti in una serie di bolle non interoperabili, privandosi di quel dinamismo e di quella forza trasformativa necessari a che la libertà di espressione e informazione continui a rappresentare un motore di crescita individuale e collettiva"⁶⁵.

D'altro canto, se la tecnologia costituisce una minaccia essa, al contempo, offre un utile contributo alla protezione dei diritti fondamentali della persona con strumenti di protezione dei dati personali e delle linee di comunicazione, con meccanismi di elusione di controlli algoritmici sempre più sofisticati⁶⁶ o con l'impiego di sistemi di moderazione automatica come selectori di *fake news* e nell'individuazione dell'hate speech, consentendo un'analisi semantica delle espressioni d'odio⁶⁷, in grado di verificare rapi-

Roma-Bari, 2010; B-C Han, *Nello sciame. Visioni del digitale*, Roma, 2015.

62 Cfr. C. M. Reale, M. Tomasi, *op. cit.*, p. 331.

63 A tal proposito, il *think-tank* Australian Strategic Policy Institute (ASPI) ha mostrato come molti degli hashtag connessi alla comunità LGBTQIA+ siano stati oscurati da TikTok in Bosnia, in Giordania e in Russia, delineando un peculiare intreccio tra informazione e questioni geopolitiche sottese ai diritti umani, cfr. *ivi*, p. 333.

64 Cfr. G. Pitruzzella, O. Pollicino, S. Quintarelli, *op. cit.*, pp. 78-79.

65 C. M. Reale, M. Tomasi, *op. cit.*, p. 336.

66 Cfr. G. Ziccardi, P. Perri, *op. cit.*, p. 95.

67 Cfr. G. Ziccardi, *op. cit.*, pp. 97-100.

damente una vastità di dati difficilmente gestibili senza l'ausilio della tecnologia⁶⁸. In questa direzione sono, tuttavia, da rilevare le criticità nel determinare criteri matematici e procedure informatiche di controllo dei contenuti, dal momento che il contesto svolge un ruolo determinante nel rinvenimento delle fattispecie riconducibili all'hate speech, mediante un'attività di interpretazione e di valutazione che rinvierebbe, in ogni caso, ad uno spazio di arbitrarietà decisionale⁶⁹: lo stesso contenuto in paesi diversi, e anche su pubblici diversi dello stesso paese, potrebbe essere valutato in modi diametralmente opposti. L'affidare, infine, all'opacità delle procedure algoritmiche di valutazione dei contenuti la tutela di diritti fondamentali e di interessi giuridicamente rilevanti potrebbe comportare un'inquietante compressione della libertà di espressione o di informazione degli interessati senza la possibilità di comprendere a pieno le basi decisionali⁷⁰. "Il rischio, chiaramente, è che gli strumenti algoritmici per contrastare il dilagare di forme di odio online, in assenza di un *online due process*, vengano utilizzati per silenziare forme di espressione che rientrano invece nell'ombrello dei diritti nell'ambito dell'informazione, o che gli stessi meccanismi di moderazione della comunità online producano e riproducano diseguaglianze e discriminazioni"⁷¹. Il rischio che si prospetta, in altri termini, è che tali sistemi informatici possano restringere se non eliminare la libertà di scelta e di valutazione soggettiva⁷², per instaurare una sorta di "dittatura dell'algoritmo"⁷³ rispecchiante una società sempre più spersonalizzata⁷⁴. Più opportuno e fattibile sembra, pertanto, l'utilizzo di sistemi informatici in grado di assistere l'operato-

re umano nella selezione di insiemi discorsivi degni di attenzione, laddove il sistema rilevi la necessità di un approfondimento e di una decisione⁷⁵, in attività preparatorie o di controllo della moderazione, oltre che nel caso di moderazione reattiva, sollecitata dagli utenti. Ciò, oltre ad evidenziare come il discernimento umano appaia, in ogni caso, indispensabile nel valutare l'inadeguatezza o meno di un contenuto⁷⁶, porta all'attenzione l'indispensabilità di un'efficace alfabetizzazione informatica, nonché di adeguati percorsi educativi e formativi tesi ad agevolare gli utenti nell'individuazione di *fake news* e discorsi d'odio, espliciti e impliciti, per procedere con la segnalazione e la rimozione degli stessi⁷⁷.

IL DIRITTO E L'HATE SPEECH: SOGGETTI E RESPONSABILITÀ

I riflessi delle innovazioni informatiche nelle dinamiche comunicative accendono, come si è visto, un intenso dibattito sul bilanciamento della libertà di informazione con ulteriori diritti fondamentali⁷⁸ rendendo controverso stabilire se e in quale misura il diritto debba intervenire⁷⁹. In altri termini, una questione centrale è rappresentata dalla prospettiva di una regolazione del dibattito online effettuata dai poteri pubblici o privati, un argomento assai controverso in linea di principio e, a maggior ragione, nel Web quale contesto

68 Cfr. C. M. Reale, M. Tomasi, *op. cit.*, p. 329.

69 Cfr. G. Pitruzzella, O. Pollicino, S. Quintarelli, *op. cit.*, p. 99-101.

70 Cfr. C. M. Reale, M. Tomasi, *op. cit.*, p. 329. Il problema della mancanza di trasparenza nei sistemi di moderazione delle piattaforme online è stato evidenziato anche nel citato Documento della Commissione Segre sottolineando l'importanza di offrire un'adeguata regolamentazione nell'utilizzo degli algoritmi, M. Martorana, *op. cit.*

71 C. M. Reale, M. Tomasi, *op. cit.*, p. 335.

72 Cfr. G. Ziccardi, *op. cit.*, p. 243.

73 Cfr. S. Rodotà, *Il mondo nella rete. Quali i diritti quali i vincoli*, Roma-Bari, 2019, p. 37.

74 Cfr. *ibidem*.

75 Cfr. G. Pitruzzella, O. Pollicino, S. Quintarelli, *op. cit.*, p. 111. Sono tipici i casi di *falsi positivi*, cioè valutati dal sistema informatico come degni di nota, ma che, in realtà, non costituiscono un problema in riferimento ad un determinato ordinamento giuridico, e di *falsi negativi*, ovvero, di contenuti potenzialmente lesivi di diritti ma non segnalati dal sistema informatico, cfr. *ivi*, pp. 111-112.

76 Cfr. *ibidem*. I sistemi cognitivi di supporto alla moderazione potrebbero risultare funzionali ad escludere la coincidenza della figura del gestore della piattaforma digitale e dell'eventuale censore dei contenuti, cfr. *ivi*, p. 115.

77 Cfr. F. Cerquozzi, *op. cit.*, p. 51. A tal proposito si può qui ricordare la *fake news challenge* e l'esperienza degli "eroi di YouTube" ovvero, il riconoscimento, denominato *trusted flagger*, che YouTube conferiva agli utenti che con maggior costanza contribuivano alla moderazione della comunità virtuale, anche mediante la sottotitolazione e il supporto tecnico, segnalando materiali potenzialmente offensivi e rimettendo alla decisione aziendale la rimozione o meno dei contenuti in esame, cfr. G. Pitruzzella, O. Pollicino, S. Quintarelli, *op. cit.*, pp. 125-125.

78 Cfr. *ivi*, p. 94

79 Cfr. C. M. Reale, M. Tomasi, *op. cit.*, p. 334.

contraddistinto da una peculiare cifra libertaria⁸⁰, facendo emergere il pericolo di un arbitrario controllo del pensiero circolante online⁸¹.

In secondo luogo, la rimozione di materiali dalla Rete richiede un intervento da parte dei soggetti professionalmente coinvolti nella diffusione dei contenuti virtuali, aprendo ad un dibattito sulla responsabilità dei fornitori di servizi online (ISP) – intesi come “*essential facilities*” democratiche⁸² o, seguendo una recente definizione, come *common carriers* di interesse pubblico⁸³ - delle piattaforme digitali, sottoposte ad una notevole responsabilizzazione in seguito alla sentenza Google-Spain del 2014, e dei *social network*, ad oggi considerabili come il mezzo privilegiato per la trasmissione del discorso di odio⁸⁴, tratteggiando un quadro sfaccettato che mette in luce tutta la difficoltà di rinvenire univocamente, e mediante le tradizionali categorie giuridiche, responsabilità e rimedi giuridici. Anche su queste tematiche si può notare, poi, un differente orientamento tra gli USA e l’Europa. Nello specifico, il dibattito statunitense sembra concentrarsi nel valutare se le piattaforme digitali private possano attuare scelte di moderazione dei contenuti non consentite ai pubblici poteri, in quanto direttamente vincolati dal Primo Emendamento⁸⁵. Su questo punto, la Section 230 del Communication Decency Act⁸⁶, pur prevedendo la possibilità, in casi particolari, di effettuare una moderazione, in buona fede, dei contenuti di terzi, esclude qualsiasi responsabilità dei fornitori di servizi per i materiali circolanti in Rete, esonerandoli da

80 Le pronunce della Corte Suprema, coerenti con il principio del *free marketplace of ideas*, si pongono in direzione della massima garanzia, anche nel Web, della libertà di espressione, cfr. G. Pitruzzella, O. Pollicino, S. Quintarelli, *op. cit.*, pp. 31-34.

81 Cfr. G. Ziccardi, *op. cit.*, p. 30.

82 Cfr. P. Falletta, *op. cit.*, p. 148.

83 Cfr. M. Bassini, *Libertà di espressione e social network*, cit., p. 46.

84 Cfr. M. N. Campagnoli, *Informazione*, cit., p. 18; cfr. F. Abbondante, *op. cit.*, 43.

85 Cfr. M. Bassini, *Libertà di espressione e social network*, cit., p. 49. La giurisprudenza statunitense è parsa opporre una chiusura alla possibilità di estensione del Primo Emendamento, proponendo un’interpretazione restrittiva della *state action doctrine*, ivi, pp. 50-52.

86 La Section 230 del CDA, introdotta dal Congresso nel 1996, regola la figura degli ISP e mette al riparo i *social network* dai contenuti pubblicati dai loro utenti, in quanto non considerati come “*media company*”.

ogni indagine preventiva⁸⁷: qualunque controllo rappresenterebbe, infatti, un freno alla libera circolazione delle idee. Nella prospettiva statunitense, “il mezzo è un elemento neutrale della diffusione del pensiero e, dunque, più esso ha capacità diffusiva più merita tutela perché consente di realizzare quella diversità di idee che arricchiscono il dibattito pubblico consentendo l’emergere, nell’ambito della competizione fra idee, di quella portatrice di verità”⁸⁸. Tale ottica sembra aver subito, tuttavia, una parziale revisione proprio in seguito alla diffusione progressiva dei discorsi d’odio che ha indotto le principali piattaforme digitali a prevedere meccanismi di controllo e di rimozione dell’hate speech tramite algoritmi, gruppi di esperti e spesso mediante gli stessi *social network* che, in maniera spesso discordante e con un occhio rivolto al marketing, hanno introdotto nei propri contratti regole volte a prevenire la propagazione di messaggi di tenore discriminatorio e offensivo⁸⁹.

Da questa impostazione appare parzialmente distaccarsi l’ambito europeo il quale, come si è visto, è caratterizzato da previsioni costituzionali dirette a tutelare ampiamente, sebbene in modo non assoluto, la libertà di manifestazione del pensiero⁹⁰, in bilanciamento con altri principi e diritti di pari rilievo, quali la dignità umana, il diritto all’immagine, alla privacy e alla protezione dei dati personali. In secondo luogo, a partire dall’ormai datata Direttiva 2000/31/CE, all’ISP non spetta alcun obbligo di sorveglianza sui contenuti immessi in Rete dagli utenti, né di ricercare attivamente fatti o circostanze che indichino lo svolgimento di attività illecite, eccettuato il caso nel quale il provider assuma un ruolo attivo nel veicolare l’informazione.

87 Cfr. F. Abbondante, *op. cit.*, 53.

88 *Ibidem*.

89 Cfr. *ibidem*.

90 Ad eccezione dell’ordinamento tedesco che, per ragioni storiche, comprende già nella Costituzione la possibilità di attuare significative deroghe alla libertà di espressione, i Paesi dell’Unione si sono dotati, nel tempo, di sanzioni dell’incitamento all’odio razziale, etnico e religioso, temperato dal bilanciamento dei giudici nazionali nel considerare la rilevanza delle circostanze concrete in modo da evitare l’anticipazione della soglia di punibilità alla mera condotta espressiva, cfr. ivi, pp. 56-59.

ne⁹¹. Viene così descritta una posizione di presunta neutralità del fornitore di servizi online alla quale corrisponde un obbligo di pronta rimozione o di disabilitare l'accesso alle informazioni su ordine della pubblica autorità, giudiziaria o amministrativa di vigilanza, unito ad oneri di comunicazione alla stessa in modo da agevolare attività investigative e di prevenzione di illeciti⁹². Dagli artt. 16 e 17 del d. lgs. n. 70/2003 si può, anche qui, notare l'assenza di un obbligo generale, a carico dell'ISP, di sorveglianza in ordine ai contenuti veicolati, prevedendo che il provider sia comunque responsabile nel caso in cui abbia avuto consapevolezza dell'illiceità degli stessi e distinguendo - come mostrano le pronunce della Cassazione Civile sez. I 19 marzo 2019 n. 7708 e della Cassazione Civile sez. I 19 marzo 2019 n. 7709 - il caso in cui la consapevolezza del contenuto illecito del documento sia stata autonomamente conseguita, ovvero derivi da una richiesta dell'autorità giudiziaria o amministrativa, poiché, nella prima ipotesi, la responsabilità si avrà solo laddove l'illiceità sia manifesta, mentre, nel secondo caso è da escludersi ogni valutazione in merito da parte dell'ISP, affermandosene la responsabilità civile per la mera inadempienza di quanto disposto dall'autorità⁹³. Seguendo tali linee interpretative e il principio del *notice and take down*⁹⁴, si è pronunciato

La Corte EDU ha escluso l'applicazione dell'art. 10 della Convenzione laddove i discorsi incitanti all'odio si traducano in attacchi veementi contro gruppi per motivi religiosi o etnici e incompatibili con i valori proclamati e garantiti dalla Convenzione, in particolare con la tolleranza, pace sociale e non-discriminazione, cfr. *ivi*, p. 59.

91 Sul *content provider* e sui più recenti indirizzi giurisprudenziali riferibili a tale figura professionale, con speciale riguardo al tema dell'hate speech, si veda G. Ziccardi, P. Perri, *op. cit.*, pp. 102-104.

92 Cfr. ad es. S. D'Alterio, *ISP: la responsabilità per le pubblicazioni degli utenti. I criteri di imputazione delle responsabilità civile e penale a carico degli Internet Service Provider in base al d.lgs. n. 70/2003* in <https://www.altalex.com/documents/news/2021/02/05/responsabilita-civile-e-penale-internet-service-provider-per-pubblicazioni-di-utenti>

93 Tali principi trovano una significativa applicazione nelle pronunce della Cassazione Civile sez. I 19 marzo 2019 n. 7708 e Cassazione Civile sez. I 19 marzo 2019 n. 7709 le quali, previa esclusione dell'attribuzione a Yahoo inc. e a Yahoo Italia s.r.l. della qualità di ISP attivo, ne affermano la responsabilità civile per omessa rimozione di video in violazione del diritto di autore di Reti televisive s.p.a., a seguito della comunicazione del carattere illecito della pubblicazione effettuata dalla parte interessata, RT s.p.a. La Suprema Corte afferma dunque la non necessità, ai fini della conoscenza dell'illecito da parte dell'ISP, della richiesta dell'autorità competente, valutando l'informazione della parte interessata quale fonte di conoscenza comunque acquisita, ai sensi dell'art 16 d.lgs 70/2003.

94 Un approfondimento dell'argomento in B. Panattoni, *Il sistema di controllo successivo: obbligo di rimozione dell'ISP e meccanismi*

il Tribunale di Napoli Nord sez. II civ. con la nota ordinanza del 3 novembre 2016, nella controversia che ha visto la madre di Tiziana Cantone - giovane donna suicidatasi a seguito della pubblicazione in Internet di video attinenti alla propria sfera intima - contrapporsi a Facebook Ireland l.t.d. In particolare, il Tribunale di Napoli Nord ha condannato Facebook ad eliminare le pubblicazioni riferite agli specifici link o Url indicati dall'istante, respingendo, tuttavia, la domanda di condanna della convenuta ad impedire ulteriori analoghi inserimenti, attesa l'espressa esclusione, nel citato decreto legislativo, di un obbligo di autonoma vigilanza in merito ai futuri materiali veicolati in Rete.

Il progressivo ampliamento delle funzioni dei fornitori di servizi online, tale, in alcuni casi, da ricalcare atteggiamenti caratteristici dei pubblici poteri⁹⁵, parrebbe, tuttavia, sollecitare un'interpretazione sempre più estesa del ruolo e delle responsabilità riferibili a tali soggetti: in questa direzione sembra, infatti, porsi il d.l. 18.02.2015 n. 7⁹⁶, provvedimento che stabilisce la formulazione di un elenco di siti web impiegati in attività terroristiche, soggetto ad aggiornamento periodico da parte delle forze di polizia postale e delle comunicazioni e in cui confluiscono le segnalazioni della polizia giudiziaria (art. 2, co. 2) prevedente un correlato obbligo per i "fornitori di connettività" di inibire l'accesso ai siti inseriti nella lista, quando ne faccia richiesta la magistratura (art. 2, co. 3)⁹⁷. Nella medesima prospettiva di espansione degli obblighi e *di notice and take down*, in "Diritto Penale Contemporaneo", (2018), n. 5, pp. 249-263.

95 Cfr. M. Bassini, *Libertà di espressione e social network*, cit., p. 55.

96 Si fa qui riferimento al Decreto-Legge 18 febbraio 2015, n. 7 "Misure urgenti per il contrasto del terrorismo, anche di matrice internazionale, nonché proroga delle missioni internazionali delle Forze armate e di polizia, iniziative di cooperazione allo sviluppo e sostegno ai processi di ricostruzione e partecipazione alle iniziative delle Organizzazioni internazionali per il consolidamento dei processi di pace e di stabilizzazione. (15G00019) (GU Serie Generale n.41 del 19-02-2015) <https://www.gazzettaufficiale.it/eli/id/2015/02/19/15G00019/sg>

97 Nel citato Decreto-Legge viene, inoltre, stabilito che, laddove vi siano concreti elementi, tali da ritenere che specifici delitti con finalità di terrorismo vengano agevolati dalla Rete, si possa ordinare ai fornitori di servizi di memorizzazione o di immissione e gestione di informazioni la rimozione del materiale illecito accessibile al pubblico, garantendo comunque, ove tecnicamente possibile, la fruizione dei contenuti estranei alle condotte illegali. È infine da rilevare come, nella predisposizione dell'elenco, sia previsto l'intervento del giudice sebbene quest'ultimo, in assenza di ulteriori chiarimenti legislativi sul punto, sembri assumere un ruolo per lo più formale, cfr. F. Abbondante, *op. cit.*, p. 64.

delle responsabilità degli ISP, sin quasi a delineare i tratti di una *culpa in vigilando*⁹⁸, si inserisce altresì la già menzionata sentenza della Corte di Cassazione del marzo 2019 n. 7708⁹⁹: in tale pronuncia, la Suprema Corte ha, infatti, chiarito come la responsabilità dell'host provider sussista qualora questi non abbia provveduto all'immediata rimozione dei contenuti illeciti, nonché quando abbia continuato a pubblicarli, nonostante ricorrano congiuntamente le seguenti condizioni: a) sia a conoscenza dell'illecito perpetrato dal destinatario del servizio, per averne avuto notizia dal titolare del diritto leso oppure *aliunde*; b) l'illiceità dell'altrui condotta sia ragionevolmente constatabile, onde egli sia in colpa grave per non averla positivamente riscontrata, alla stregua del grado di diligenza che è ragionevole attendersi da un operatore professionale della rete in un determinato momento storico; c) abbia avuto la possibilità di attivarsi utilmente. All'host provider appartiene, in ogni caso, la disponibilità di adottare norme di principio e codici etici conformi ai valori della community, il potere disciplinare di imporre sanzioni verso chi contravvenisse alle regole stabilite, nonché un più generico potere di controllo sulla comunità virtuale¹⁰⁰. Significativa, in tal senso, è stata la decisione di Facebook e di Instagram, nel 2019, di oscurare le pagine di due movimenti neofascisti, CasaPound e Forza Nuova, e di numerosi esponenti delle suddette associazioni inneggianti ripetutamente all'odio e alla violenza, una circostanza che ha condotto, ancora una volta, ad interrogarsi sul ruolo e sui compiti delle piattaforme virtuali, direttamente chiamate in causa nella decisione sulla rimozione o meno dei contenuti¹⁰¹, quesiti aggravati dalla possibilità di censurare materiali non espressamente vietati dalla legge statutale¹⁰².

In riferimento a tali questioni, anche l'Unione Europea, come testimoniato dalle Linee guida della Commissione del 2017 concernenti gli obblighi di sorveglianza sulle informazioni trasmesse o memorizzate dai fornitori di servizi online, ha sviluppato

98 Cfr. P. Falletta, *op. cit.*, p. 151.

99 Cass. civ., sez. I, 19 marzo 2019 n. 7708. Una critica del suddetto orientamento giurisprudenziale in M. Bassini, *La Cassazione e il simulacro del provider attivo: Mala tempora currunt*, in "MediaLaws – Rivista di Diritto dei Media", (2019), n. 2, pp. 248 ss.

100 Cfr. P. Falletta, *op. cit.*, p. 156.

101 Cfr. G. Ziccardi, *op. cit.*, p. 77.

102 Cfr. P. Falletta, *op. cit.*, pp. 156-158; cfr. M. Bassini, *Libertà di espressione e social network*, cit., pp. 53-55.

un indirizzo volto ad una maggiore responsabilizzazione degli ISP, al di là di quanto previsto dalla Direttiva 2000/31/CE, aggravandone gli oneri con il sancire i principi del *take-down*, esprimente la necessità di un'espressa previsione statutale del procedimento di tempestiva rimozione dei contenuti illeciti, e dello *stay-down* inteso come prevenzione della ricomparsa di contenuti illeciti analoghi a quelli già oggetto di *take down*¹⁰³. Alla base della più recente regolamentazione si colloca la recrudescenza di episodi delittuosi connessi alle comunicazioni digitali, – basti qui ricordare la "blue whale challenge" o la "blackout challenge" – nonché il notevole progresso delle tecnologie informatiche, tale da consentire, attraverso meccanismi di ricerca basati su sofisticati algoritmi, di gestire tempestivamente ed efficacemente le iniziative di *take down* e *stay down* attivabili dai providers. In definitiva, l'assenza di un generale obbligo di sorveglianza da parte degli ISP sembra arrestarsi di fronte a casi specifici che impongano non solo obblighi di rimozione dei contenuti illeciti, a prescindere da specifici provvedimenti della pubblica autorità, ma anche controlli *sine die* affinché non si ripetano condotte dello stesso tipo o anche solo equivalenti¹⁰⁴.

È da sottolineare, in aggiunta, come il collegare la rimozione dei contenuti illeciti alla richiesta del soggetto leso e alla conseguente tempistica¹⁰⁵ si contrapponga all'immediatezza di diffusione della notizia, eventualmente aggravata dalla presenza di account anonimi¹⁰⁶ o riferibili ad utenti difficilmente individuabili essendo sottoposta, solo in seconda istanza, al sindacato giurisdizio-

103 Tale onere veniva escluso dalla citata Ordinanza del Tribunale di Napoli Nord sez. II civ. 3 novembre 2016 nella quale si affermava, infatti, che "non ritenendosi sussistente un dovere di Facebook di verificare in via anticipata il contenuto dei post e dei commenti immessi dagli utenti, non appare in conseguenza configurabile a suo carico il dovere di inibire, in via generale, un caricamento sulla sua piattaforma "di ogni video, immagini, notizie o articoli riferiti alla persona della ricorrente" essendo, invece, possibile impedire, nell'ottica del dovere di controllo "successivo" (perché preceduto da una denuncia in cui sono stati individuati gli URL o i LINK dal contenuto lesivo) e "mirato" (perché diretto a impedire nuovi caricamenti degli URL o dei LINK già segnalati), il solo nuovo caricamento degli stessi specifici links comunicati"

104 P. Falletta, *op. cit.*, p. 153.

105 Secondo un Report della Commissione europea, citato dalla Rete Nazionale Per Il Contrasto Ai Discorsi E Ai Fenomeni D'Odio, riferito al 2022, solo il 64, 4 % delle segnalazioni sarebbero state gestite nelle 24 ore previste, <https://www.retecontrolodio.org>

106 Sulle incognite e le criticità riferibili all'anonimato degli utenti digitali, si veda G. Ziccardi, *op. cit.*, pp. 95-97.

nale la verifica dell'inosservanza delle regole di convivenza della comunità virtuale, nonché l'accertamento delle lesioni di situazioni giuridicamente rilevanti¹⁰⁷. Vi è, infine, l'evenienza che "la richiesta di rimozione sia infondata, strumentale e peggio ancora tenda alla eliminazione di contenuti che sono espressione della libertà di informazione e alimentano il dibattito pubblico"¹⁰⁸, ovvero, la possibilità di limitazioni private alla libertà di espressione¹⁰⁹. In tale contesto, estremamente sfaccettato, si inserisce, dunque, il ruolo di "vigile disimpegno" degli ISP nonché il dibattito sui profili di responsabilità riferibili a tali soggetti che rimanda a questioni di assoluto rilievo quali la libertà di manifestazione del pensiero, di informazione e di associazione e, in secondo luogo, la libertà di impresa e gli sviluppi del commercio elettronico¹¹⁰. Al riguardo, sul presupposto che una tempestiva ed efficace rimozione dei contenuti illeciti online non possa essere efficacemente garantita dalla pubblica autorità quanto dalle piattaforme digitali, in prima linea nell'assicurare una rapida e ramificata reazione alle segnalazioni degli utenti¹¹¹, la Commissione Europea, il 31 maggio 2016, in accordo con Facebook, Twitter, YouTube e Microsoft ha varato un Codice di condotta - cui è seguita, nel marzo 2018, una Raccomandazione - concernente misure di contrasto ai contenuti illegali online la quale, pur essendo non vincolante, rappresenta una "morbida pressione" nell'opposizione all'odio online. Secondo tale Codice, le piattaforme si impegnano ad autoregolarsi e a sorvegliare le proprie community, predisponendo rapide ed efficaci procedure nel vaglio delle segnalazioni effettuate dagli utenti dei contenuti incitanti all'odio, per rimuoverli o renderli inaccessibili¹¹² in un termine massimo di 24 ore dalla segnalazione del contenuto illegale, pur notando criticamente come la ristrettezza dei termini previsti per l'azione degli intermediari escluda la possibilità di un approfondimento nel processo di verifica dei contenuti, derivandone un aggravio del rischio di errore e di rimozione

di speech leciti¹¹³. Di recente, anche l'AGCOM, il 29 luglio 2022, ha pubblicato uno schema di Regolamento recante disposizioni di rilievo sul principio di non-discriminazione e di lotta all'hate speech¹¹⁴, un provvedimento che traccia significative indicazioni nel bilanciamento dei diritti fondamentali inevitabilmente compromessi nel discorso d'odio con la tutela della libertà di espressione e di associazione e, sempre recentemente, il Consiglio d'Europa ha adottato una nuova Raccomandazione di contrasto all'odio¹¹⁵, online e offline, accompagnata da un'Appendice con principi e Linee-guida i quali, oltre ad invitare gli Stati membri ad imporre legislativamente agli intermediari digitali idonee misure di contrasto all'hate speech, esprimono un approccio globale e *multistakeholder* al fenomeno¹¹⁶. Da ultimo, non certo per importanza, va ricordato come il *Digital Services Act* (DSA)¹¹⁷ sia inteso ad assicurare un ambiente online più sicuro, migliorando i procedimenti di rimozione dei contenuti illeciti ed imponendo stringenti standard di responsabilità per le piattaforme digitali con la previsione di ulteriori obblighi a carico di tali soggetti¹¹⁸, tra cui quello di offrire informazioni esplicite sulle modalità di moderazione dei contenuti e, per gli intermediari di maggiori dimensioni, mediante la previsione dell'adozione di codici di condotta e di meccanismi di prevenzione dei rischi sistemici concernenti i diritti e le libertà fondamentali degli individui.

107 Cfr. P. Falletta, *op. cit.*, p. 156.

108 G. Pitruzzella, O. Pollicino, S. Quintarelli, *op. cit.*, p. 82.

109 Cfr. M. Bassini, *Libertà di espressione e social network*, cit., p. 46.

110 Cfr. G. Ziccardi, *op. cit.*, pp. 93-94

111 Cfr. P. Falletta, *op. cit.*, p. 154.

112 Cfr. C. M. Reale, M. Tomasi, *op. cit.*, p. 334.

113 Cfr. F. Abbondante, *op. cit.*, p. 65. Nel suo primo anno di applicazione l'impegno preso dalle grandi piattaforme digitali non ha, tuttavia, sortito l'effetto immaginato, tant'è che queste ultime, in seguito alla minaccia di una regolazione più rigorosa da parte dell'Ue, hanno sottoscritto un reciproco accordo per la creazione di un database condiviso - diretto, tuttavia, alla rimozione di immagini e video a sfondo terroristico, cfr. *ibidem*.

114 Il Regolamento in questione, adottato in seguito alla delibera 292/22/CONS, sostituisce il precedente Regolamento recante disposizioni in materia di rispetto della dignità umana e del principio di non discriminazione e di contrasto all'hate speech approvato con delibera 157/19/CONS, art. 11, "Abrogazioni".

115 COMMITTEE OF MINISTERS TO MEMBER STATES, Recommendation on Combating Hate Speech, 20 maggio 2022 (CM/Rec(2022)16[1]).

116 Cfr. M. Martorana, *op. cit.*

117 Regolamento (UE) 2022/2065 Del Parlamento Europeo e Del Consiglio, 19 ottobre 2022, relativo a un mercato unico dei servizi digitali in <https://eur-lex.europa.eu/legal-content/IT/TXT/?uri=CELEX:32022R2065>

118 Cfr. B. G. Bello, L. Scudieri, *op. cit.*, p. 12.

CONCLUSIONI

La circostanza che proprio dal Web provengano molteplici segnalazioni e richieste di rimozione dei *post* incitanti all'odio sembra far ripensare il contrasto all'hate speech al di là delle possibili misure repressive¹¹⁹, per introdurre interventi di segno positivo, di supporto delle vittime, iniziative di "attivismo digitale"¹²⁰ e *counterspeech* ovvero narrazioni alternative indipendenti ed affidate agli stessi utenti che intervengono nel discorso in modo da dissuadere l'aggressore, senza sottovalutarne il potenziale dannoso e, al contempo, garantendo il confronto democratico¹²¹. Si porta all'attenzione, inoltre, l'indispensabilità di una complessa opera di educazione all'impiego delle più recenti tecnologie¹²², un percorso valorizzato nella guida UNESCO *Addressing Hate Speech Through Education. A guide for policy-makers* del 27 marzo 2023¹²³, essenziale a scongiurare ogni rischio di paternalismo informatico e correlato ad una formazione linguistica capace di valorizzare i principi della cittadinanza e della democrazia¹²⁴, favorendo l'apertura del dibattito pubblico¹²⁵ in modo da diffondere gli *anticorpi* dell'odio virtuale¹²⁶, nella direzione indicata dal No Hate Speech Movement del Dipartimento della Gioventù del Consiglio d'Europa e testimoniata dalle molteplici iniziative della Comunità internazionale¹²⁷.

119 Cfr. G. Ziccardi, *op. cit.*, p. 81; cfr. A. De Rosa, *op. cit.*, p. 209.

120 Per una definizione ed un approfondimento del fenomeno degli *attivismi digitali* e dei *soggetti attivisti* nel contrasto al discorso d'odio, si rimanda a B. G. Bello, *op. cit.*, pp. 155-175.

121 Sulla complessità del *counter speech*, fenomeno che a volte assume connotazioni implicite, in altri casi caratteristiche più esplicite e di diretto contrasto del discorso d'odio, nell'intento di non alimentare la spirale della violenza verbale, si veda F. Cerquozzi, *op. cit.*, pp. 51-52; cfr. G. Ziccardi, P. Perri, *op. cit.*, p. 96.

122 Cfr. G. Ziccardi, *op. cit.*, p. 208; cfr. G. Ziccardi, P. Perri, *op. cit.*, p. 95.

123 <https://unesdoc.unesco.org/ark:/48223/pf0000384872/PDF/384872eng.pdf.multi>

124 Cfr. S. Cavagnoli, *op. cit.*, p. 34.

125 Cfr. C. M. Reale, M. Tomasi, *op. cit.*, p. 332. Sulla riformulazione dello spazio pubblico in Internet e sui nessi tra discorso d'odio e modelli democratici, si rimanda a F. J. Ansuátegui Roig, *op. cit.*, pp. 204-215

126 Cfr. G. Ziccardi, *op. cit.*, p. 31.

127 Al riguardo, a titolo esemplificativo, possiamo ricordare la promozione da parte delle Nazioni Unite di una giornata internazionale dedicata al contrasto dell'hate speech che, a partire dal 18 giugno del 2022, si celebrerà ogni anno nella medesima data.

Le questioni affrontate sollecitano, infine, una profonda riflessione che investe la centralità della persona e dei suoi dati elettronici - essendo evidente come l'incrementare il livello di protezione degli stessi, garantendone parallelamente la libera circolazione, possa costituire un disincentivo a discorsi d'odio mirati sul singolo individuo e, pertanto, estremamente lesivi¹²⁸ - nonché il ruolo dei *social network* e degli intermediari virtuali, mettendo in luce il rischio di una *collateral censorship* conferita a soggetti che, per vocazione specifica, esercitano attività di natura commerciale¹²⁹, un tema che rimanda alla complessità dei rapporti tra poteri pubblici e privati e ai loro riflessi sul concreto esercizio della libertà di espressione¹³⁰.

Arianna Maceratini è ricercatrice in Filosofia del Diritto e professore aggregato di Informatica Giuridica presso il Dipartimento di Giurisprudenza dell'Università degli Studi di Macerata.

Arianna Maceratini is researcher in Philosophy of Law and adjunct professor of Legal Informatics at the Department of Law of the University of Macerata.

arianna.maceratini@unimc.it

OPERE CONSULTATE

F. Abbondante, *Il ruolo dei social network nella lotta all'hate speech: un'analisi comparata fra l'esperienza statunitense e quella europea*, in "Informatica e diritto", XLIII (2017), n. 1-2, pp. 41-68.

A. C. Amato Mangiameli, G. Saraceni (a cura di), *Cento e una voce di informatica giuridica*, Torino, 2023.

F. J. Ansuátegui Roig, *Libertà d'espressione: ragione e storia*, Torino, 2018.

J. L. Austin, *Come fare cose con parole*, Genova, 2019.

128 Cfr. G. Ziccardi, P. Perri, *op. cit.*, pp. 98-101.

129 Cfr. G. Pitruzzella, O. Pollicino, S. Quintarelli, *op. cit.*, p. 45.

130 Cfr. M. Bassini, *Libertà di espressione e social network*, cit., p. 44; cfr. C. M. Reale, M. Tomasi, *op. cit.*, p. 336.

M. Bassini, *La Cassazione e il simulacro del provider attivo: Mala tempora currunt*, in “MediaLaws – Rivista di Diritto dei Media”, (2019), n. 2, pp. 248-257.

M. Bassini, *Libertà di espressione e social network, tra nuovi “spazi pubblici” e “poteri privati”. Spunti di comparazione*, in “Rivista Italiana di Informatica e Diritto”, (2021), n. 2, pp. 43-64.

Z. Bauman, *Consumo, dunque sono*, Roma-Bari, 2010.

B. G. Bello, L. Scudieri (a cura di), *L’odio online: forme, prevenzione e contrasto*, Torino, 2022.

C. Bianchi, *Hate speech. Il lato oscuro del linguaggio*, Roma-Bari, 2021.

M. N. Campagnoli, *Hate Speech*, in A. C. Amato Mangiameli, G. Saraceni (a cura di), *Cento e una voce di informatica giuridica*, Torino, 2023, pp. 253-257.

M. N. Campagnoli, *Informazione, social network & diritto*, Milano, 2020.

F. Cerquozzi, *Dall’odio all’hate speech. Conoscere l’odio e le sue trasformazioni per poi contrastarlo*, in “Tigor: Rivista di scienze della comunicazione e dell’argomentazione giuridica”, X (2018), n. 1, pp. 42-53.

S. D’Alterio, *ISP: la responsabilità per le pubblicazioni degli utenti. I criteri di imputazione delle responsabilità civile e penale a carico degli Internet Service Provider in base al d.lgs. n. 70/2003*, <https://www.altalex.com/documents/news/2021/02/05/responsabilita-civile-e-penale-internet-service-provider-per-pubblicazioni-di-utenti>; Sito consultato il 26/09/2023

A. Di Rosa, *Hate speech e discriminazione. Un’analisi performativa tra diritti umani e teorie della libertà*, Modena, 2020.

F. Di Tano, *I reati informatici e i fenomeni del cyberstalking, del cyber bullismo e del revenge porn*, in T. Casadei, S. Pietropaoli, *Diritto e tecnologie informatiche*, Milano, 2021, pp. 165-178.

P. Falletta, *Controlli e responsabilità dei social network sui discorsi d’odio online*, in “MediaLaws – Rivista di Diritto dei Media”, (2020), n. 3, pp. 146-158.

L. Floridi, *La quarta rivoluzione. Come l’infosfera sta trasformando il mondo*, Milano, 2017.

B-C Han, *Nello sciame. Visioni del digitale*, Roma, 2015.

M. Martorana, *Il complesso rapporto tra hate speech, libertà e democrazia. Un difficile bilanciamento tra le libertà fondamentali coinvolte e gli sforzi europei per il contrasto ai discorsi d’odio*, <https://www.altalex.com/documents/news/2022/11/08/complesso-rapporto-hate-speech-liberta-democrazia>; Sito consultato il 26/09/2023

M. Orofino, *La libertà di espressione tra Costituzione e Carte europee dei diritti. Il dinamismo dei diritti di una società in trasformazione*, Torino, 2014.

B. Panattoni, *Il sistema di controllo successivo: obbligo di rimozione dell’ISP e meccanismi di notice and take down*, in “Diritto Penale Contemporaneo”, (2018), n. 5, pp. 249-263.

G. Pitruzzella, O. Pollicino, S. Quintarelli, *Parole e potere. Libertà d’espressione, hate speech e fake news*, Milano, 2017.

C. M. Reale, M. Tomasi, *Libertà d’espressione, nuovi media e intelligenza artificiale: la ricerca di un nuovo equilibrio nell’ecosistema costituzionale*, in “DPCE online”, (2022), n. 1, pp. 325-336.

S. Rodotà, *Il mondo nella rete. Quali i diritti quali i vincoli*, Roma-Bari, 2019.

C. R. Sunstein, *#republic*, Bologna, 2017.

G. Ziccardi, *L’odio online. Violenza verbale e ossessioni in rete*, Milano, 2016.