

# Online hate speech, diritto penale e libertà di espressione. Utopia od opportunità?\*

Matilde Bellingeri, Federica Delaini

## Abstract

L'avvento delle nuove tecnologie nella governance del discorso pubblico impone di svolgere una riflessione circa l'opportunità di utilizzare il diritto penale quale strumento volto a contrastare (e prevenire) le condotte di online hate speech, le quali hanno assunto potenzialità lesive senza precedenti, portando con sé seri rischi per la sopravvivenza di una società democratica e plurale. A fronte della dimensione transnazionale del fenomeno, si procederà tracciando una mappatura dello stato dell'arte in Italia, mediante il ricorso ad una metodologia comparatistica che abbracci le soluzioni applicate in ambito europeo ed americano. Si rifletterà sull'efficacia delle scelte di politica criminale, valutando, quali rimedi alternativi, gli strumenti di "moderazione algoritmica" in capo agli Internet Service Providers. In conclusione, si indagheranno le interferenze tra i limiti apposti alla libertà di pensiero ed i principi che sorreggono un ecosistema penale costituzionalmente orientato, nel prisma della tutela dei diritti e delle libertà fondamentali degli utenti.

The advent of new technologies in the governance of public speech compels a reflection on the possibility of using criminal law as a means of counteracting (and preventing) online hate speech, which has taken on an unprecedented detrimental potential, leading to serious risks for the survival of a democratic and pluralistic society. In view of the transnational dimension of the phenomenon, it is intended to map the state of the art in Italy, through a comparative methodology that encompasses the solutions applied in the European and American frameworks. The effectiveness of criminal policy measures will be reflected upon, evaluating, by way of alternative remedies, the tools of algorithmic moderation in the hands of the Internet Service Providers. In conclusion, the interactions between the limits posed on freedom of thought and the principles that underpin a constitutional criminal ecosystem will be investigated through the prism of the protection of users' fundamental rights and freedoms.

\* L'articolo è stato sottoposto, in conformità al regolamento della Rivista, a referaggio "a doppio cieco". Il seguente articolo è frutto di riflessioni condivise. Tuttavia, i paragrafi 2, 4 e 7 sono attribuibili a Federica Delaini, mentre i paragrafi 3, 5 e 6 a Matilde Bellingeri. L'introduzione e le conclusioni (paragrafi 1 e 8) sono stati scritti congiuntamente dalle Autrici.

Contributo finanziato dall'Unione europea - Next Generation EU, Missione 4, Componente 1, CUP B31I23000810004

## Sommario

1. Introduzione: il contrasto all'*online hate speech*. – 2. Il panorama giuridico europeo: la mancata inclusione dei discorsi d'odio online tra i c.d. "eurocrimini". – 3. Il quadro normativo e giurisprudenziale italiano. – 4. Esperienze giuridiche di incriminazione dell'*online hate speech* a confronto: Francia, Germania e Spagna. – 5. Lo stato dell'arte oltre Oceano. – 6. Rimedi alternativi: quale ruolo per gli strumenti di "moderazione algoritmica"? – 7. I rapporti tra diritto penale, libertà di manifestazione del pensiero nell'era digitale e democraticità. – 8. Riflessioni conclusive: il ricorso al diritto penale, utopia od opportunità?

## Keywords

diritto penale – discorsi d'odio online – comparazione – libertà di espressione – democrazia plurale

---

*“Il connubio di odio e di tecnologia è il massimo pericolo che sovrasti l'umanità. E non mi riferisco alla sola grande tecnologia della bomba atomica, mi riferisco anche alla piccola tecnologia della vita di ogni giorno: conosco persone che stanno per ore davanti al televisore perché hanno disimparato a comunicare tra di loro”*

Simon Wiesenthal

## 1. Introduzione: il contrasto all'*online hate speech*

La comunicazione al tempo di Internet ha fortemente ridisegnato le coordinate del discorso pubblico: sulle piattaforme digitali ogni individuo diffonde il proprio pensiero a priori dall'appartenenza ad una classe ovvero etnia, in piena rispondenza al paradigma strettamente democratico che permea l'ecosistema digitale. Tuttavia, la fondamentale libertà di manifestazione del pensiero e d'informazione e la tutela del pluralismo e del diritto alla non discriminazione trovano massima frizione nel contesto delle piattaforme online, in quanto luoghi forieri di nuovi e gravi pregiudizi, sia individuali che collettivi<sup>1</sup>.

Per detta ragione, si è ritenuto di cruciale importanza esaminare, quale asse portante del contributo, la *vexata quaestio* dell'opportunità di incriminare i discorsi d'odio online, che costituiscono una *species* del più ampio *genus* degli *hate crimes*<sup>2</sup>, dai quali ereditano le riflessioni in tema di oggettività giuridica tutelata, oltre che i punti di contatto e di

---

<sup>1</sup> P. Stanzione, *Introduzione*, in P. Stanzione (a cura di), *I "poteri privati" delle piattaforme*, Torino, 2022, 9.

<sup>2</sup> L'Organizzazione per la Sicurezza e la Cooperazione in Europa (OCSE) ha definito i crimini d'odio come «*violent manifestations of intolerance and have a deep impact on not only the immediate victim but the group with which that victim identifies him or herself. They affect community cohesion and social stability. A vigorous response is therefore, important both for individual and communal security*», così OCSE, in *Hate Crime Laws. A Practical Guide*, Odih, Varsavia, 2009, 11.

tensione con le fondamentali libertà sopra menzionate<sup>3</sup>.

Nonostante nel panorama internazionale non esista una definizione unitaria di discorso d'odio, gli Stati che si sono interessati al tema ne hanno individuato alcuni tratti comuni: l'incitamento all'odio, la volontà che l'odio non rimanga in uno stato di potenza, bensì venga posto in azione nei confronti del gruppo che ne è destinatario e il fatto che per quest'ultimo sussista anche soltanto il rischio imminente che vengano realizzati atti di violenza e di discriminazione<sup>4</sup>.

Tali elementi sono presenti, altresì, nella loro massima evoluzione, nella recentissima Raccomandazione CM/Rec (2022) del Comitato dei Ministri del Consiglio d'Europa, ove si afferma che con discorso d'odio «si intende qualsiasi forma di espressione mirante a stimolare, promuovere, diffondere o giustificare la violenza, l'odio o la discriminazione nei confronti di una persona o un gruppo di persone, o a denigrare una persona o un gruppo di persone per motivi legati alle loro caratteristiche o situazioni personali, reali o presunte, quali la “razza”, il colore della pelle, la lingua, la religione, la nazionalità o l'origine nazionale o etnica, l'età, la disabilità, il sesso, l'identità di genere e l'orientamento sessuale»<sup>5</sup>.

Da tempo la dottrina penalistica si interroga sull'opportunità di sanzionare penalmente la manifestazione di un pensiero “pericoloso”, qualora esso contrasti con valori di rango superiore alla libertà di espressione<sup>6</sup>. Quest'ultima, infatti, a una prima lettura, sembra incompatibile con qualsiasi forma di limitazione. Appare, dunque, legittimo chiedersi per quale motivo alcuni ordinamenti contemporanei, tra cui il nostro, abbiano attribuito rilevanza penale ai discorsi d'odio<sup>7</sup>.

In quest'ottica, si cercherà di investigare le problematiche insite nell'utilizzo della rete, che a tutt'oggi è in grado di mettere a rischio la sopravvivenza di società eterogenee e multiculturali, data la spiccata attitudine lesiva nei confronti dei valori fondamentali su

<sup>3</sup> Per un maggiore approfondimento in ordine ai diversi modelli di incriminazione dei discorsi d'odio che si sono diffusi a livello europeo, distinguibili in modelli istigatori in contrapposizione a modelli “ampi”, si rinvia a A. Spena, *La parola(-)odio. Sovraesposizione, criminalizzazione e interpretazione dello hate speech*, in *Criminalia*, 2017, 589-593.

<sup>4</sup> G. Ziccardi, *Il contrasto dell'odio online: possibili rimedi*, in *Lessico di etica pubblica*, 9,1, 2018, 39 – 40; A. Portaru, *Freedom of Expression Online: the Code of Conduct on Countering Illegal Hate Speech*, *Revista Romana de Drept European*, 4, 2017, 82-83.

<sup>5</sup> Raccomandazione CM/Rec (2022) 16, del Comitato dei Ministri agli Stati membri sulla lotta contro i discorsi d'odio, adottata dal Comitato dei Ministri il 20 maggio 2022, in occasione della 132esima Sessione del Comitato dei Ministri d'Europa.

<sup>6</sup> M. Pelissero, *Discriminazione, razzismo e diritto penale fragile*, in *Diritto penale e procedura*, 2020, 1017 ss.; R. Bartoli-M. Pelissero-S. Seminara, *Diritto penale. Lineamenti di parte speciale*, Torino, 2022, 25; A. Cadoppi-S. Canestrari-A. Manna-M. Papa, *Diritto penale. Tomo terzo. I delitti contro l'onore e la libertà individuale*, Vicenza, 2022, 6101 ss.; P. Tanzarella, *Discriminare parlando. Il pluralismo democratico messo alla prova dai discorsi d'odio razziale*, Torino, 2020, 19-46; M. Caputo, *La menzogna di Auschwitz, le verità del diritto penale. La criminalizzazione del c.d. negazionismo tra ordine pubblico, dignità e senso di umanità*, in AA.VV. *Verità del precetto e della sanzione penale alla prova del processo*, a cura di G. Forti-G. Varraso-M. Caputo, Napoli, 2014, 309; sul tema si veda anche G. Pavich-A. Bonomi, *Reati in tema di discriminazione: il punto sull'evoluzione normativa recente, sui principi e valori in gioco, sulle prospettive legislative e sulla possibilità di interpretare in senso conforme a Costituzione la norma vigente*, in *Penale Contemporaneo*, 13.10.2014; A. Galluccio, *Punire la parola pericolosa?*, Milano, 2020, 319 ss.

<sup>7</sup> Per un maggior approfondimento in relazione a tale profilo si rimanda all'analisi che si svolgerà nei paragrafi 3 e 4 del presente contributo.

cui si basano le moderne democrazie plurali<sup>8</sup>.

Finalità della disamina è quella di individuare i principali profili di criticità della reintroduzione, all'interno del nostro ordinamento (e non solo), dei c.d. "reati di opinione"<sup>9</sup>, idonei a turbare i valori morali sovra-individuali riconducibili ad un'intera collettività<sup>10</sup>. Al riguardo, la Corte costituzionale ha precisato che il bilanciamento degli interessi è il modello che meglio si adatta alla risoluzione della questione, in quanto consente di considerare i rapporti tra la libera manifestazione del pensiero e il cruciale controvalore che emerge, condizionato dal contesto sociopolitico di riferimento. Tale bilanciamento è chiamato a fronteggiare nuove criticità in considerazione del fatto che, con la pubblicazione online sempre più massiccia di contenuti da parte degli utenti, i fornitori di servizi digitali sono chiamati a svolgere attività di moderazione. L'attuale necessità è quella di trovare il miglior equilibrio possibile tra l'esigenza di proteggere i beni giuridici direttamente colpiti dalla diffusione del contenuto d'odio online e il diritto alla libertà di espressione, su cui si regge altresì l'odierno Stato costituzionale di diritto<sup>11</sup>.

In aggiunta, alcune caratteristiche dei social network impongono la necessità di cooperazione tra gli Stati e le giurisdizioni: basti pensare alla velocità con cui i messaggi si diffondono, alla possibilità di raggiungere un numero indeterminato di destinatari in maniera pressoché simultanea, alla capacità del contenuto offensivo di sopravvivere per un lungo arco di tempo dalla sua immissione (anche in parti del web diverse da quelle in cui era stato originariamente inserito), alla possibilità di un "ritorno imprevedibile", all'anonimato e alla natura transnazionale degli intermediari informatici<sup>12</sup>.

Per fornire una trattazione il più possibile esaustiva, si è deciso di adottare un metodo di tipo comparatistico, al fine di analizzare con maggiore consapevolezza l'opportunità della scelta di politica criminale effettuata dal legislatore italiano, mettendola a confronto con la legislazione radicalmente diversa vigente negli Stati Uniti d'America. Inoltre, si prenderà in considerazione l'esperienza maturata nel continente europeo, dove, passando attraverso l'esperienza giuridica francese, tedesca e spagnola, si cercheranno di analizzare le vicende legislative e giurisprudenziali che hanno segnato il cammino del ricorso alla tecnica incriminatrice in riferimento alle condotte di *online hate speech* e che appaiono, quindi, in grado di spiegarne l'attuale fisionomia.

Infine, ci si è interrogati circa l'efficacia di ricorrere a strumenti diversi dalla tecnica di stampo penalistico per contrastare (e prevenire) il fenomeno in parola, indagando

<sup>8</sup> In relazione ai rapporti con i crimini d'odio, si rimanda a L. D'Amico, *Le forme dell'odio. Un possibile bilanciamento tra irrilevanza penale e repressione*, in *La Legislazione penale*, 17.6.2020, 2 ss.

<sup>9</sup> Sul tema si rinvia a M. Pelissero, *La parola pericolosa. Il confine incerto del controllo penale del dissenso*, in *Questione Giustizia*, 2015, 38 e L. Alesiani, *I reati di opinione. Una rilettura in chiave costituzionale*, Milano, 2006, 141 ss.

<sup>10</sup> A. Spena, *Libertà di espressione e reati di opinione*, in *Rivista italiana di diritto e procedura penale*, 2, 2007, 692 ss.

<sup>11</sup> Quali quelli alla libera manifestazione del pensiero e alla libera informazione (ex artt. 10 e 11 Carta di Nizza). Per un approfondimento in merito, si segnala: B. Sander, *Freedom of Expression in the Age of Online Platforms: The Promise and Pitfalls of a Human Rights-Based Approach to content moderation*, in *Fordham International Law Journal*, 2020, 939 ss.

<sup>12</sup> G. Ziccardi, *Il contrasto dell'odio online: possibili rimedi*, cit., 42.

il ruolo dei meccanismi di moderazione algoritmica nella *governance* del discorso pubblico online.

Attraverso l'approfondimento delle direttive di indagine sopra individuate, l'obiettivo finale della trattazione sarà comprendere se la repressione dell'*online hate speech* possa pregiudicare l'esercizio della libertà di espressione, componente insostituibile di ogni democrazia eterogenea.

## 2. Il panorama giuridico europeo: la mancata inclusione dei discorsi d'odio online tra i c.d. "eurocrimini"

Si è già sottolineato come i discorsi d'odio si pongono in un rapporto di *species* rispetto al più ampio *genus* dei crimini d'odio. Essi si connotano, al pari di questi ultimi, per la sussistenza di un c.d. *bias motive*, ossia il motivo di pregiudizio che li supporta, mentre difettano, rispetto ai primi, della c.d. *criminal offence*, ossia la commissione di un reato<sup>13</sup>. Di talché anche l'*hate speech*, al pari degli *hate crimes*, può essere definito un crimine sorretto e motivato dal pregiudizio<sup>14</sup>, trattandosi, secondo la definizione resa dal Comitato dei Ministri del Consiglio d'Europa nel 1997, di «discorsi suscettibili di produrre l'effetto di legittimare, diffondere o promuovere l'odio razziale, la xenofobia, l'antisemitismo o altre forme di discriminazione o odio basate sull'intolleranza»<sup>15</sup>.

Al fine di comprendere compiutamente la portata lesiva di queste condotte, è, altresì, opportuno evidenziare come, secondo la definizione maggiormente condivisa di crimini d'odio fornita dall'Organizzazione per la Sicurezza e la Cooperazione in Europa nel 2009, essi si sostanziano nell'esternalizzazione di «violente manifestazioni di intolleranza dotate di un profondo impatto non solo sulla vittima diretta bensì anche sul gruppo con cui la vittima si identifica. Essi colpiscono la coesione della comunità e la stabilità sociale. Pertanto, una risposta vigorosa è importante sia per la sicurezza individuale che per quella comune»<sup>16</sup>.

Da tale enunciazione emerge con chiara evidenza l'insita duplicità che connota i cri-

<sup>13</sup> Ocse, *Hate Crime Laws. A practical Guide*, Adhir, Varsavia, 2009, 16, come evidenziato da L. D'Amico, *Le forme dell'odio. Un possibile bilanciamento tra irrilevanza penale e repressione*, in *La Legislazione Penale*, 2020, 6, secondo la quale proprio tale profilo rende impellente e di cruciale importanza il dibattito innestatosi in riferimento all'elemento soggettivo del reato.

<sup>14</sup> L. Goisis, *Crimini d'odio. Discriminazioni e giustizia penale*, Napoli, 2019, 30. Di seguito alcune delle definizioni anche ricordate da L. D'Amico, *Le forme dell'odio*, cit., 2, che si sono succedute in materia nel corso degli anni: B. Perry «il crimine d'odio [...] comporta atti violenti ed intimidatori, generalmente diretti verso gruppi già oggetto di marginalizzazione e stigmatizzazione. Così inteso, è un meccanismo di potere e di oppressione, teso a riaffermare le precarie gerarchie che caratterizzano un dato ordine sociale [...]» in B. Perry, *In the Name of Hate: Understanding Hate Crimes*, Londra 2001, 1 e 10; N. Chakraborti i crimini d'odio vanno individuati in quegli «atti di violenza, intimidazione e ostilità diretti verso persone a causa della loro identità o della loro percepita diversità» in N. Chakraborti-J. Garland, *Hate Crime. Impact, Causes, and Responses*, Los Angeles-Londra 2015, 5; F. Lawrence definisce l'hate crime – o meglio il bias crime – come un «crimine commesso per un motivo di pregiudizio (bias) contro una "caratteristica protetta", propria di un gruppo» in F. M. Lawrence, *Punishing Hate. Bias Crimes under American Law*, Cambridge 1999, 9.

<sup>15</sup> Comitato dei Ministri del Consiglio d'Europa, *Raccomandazione n. R (97)20*.

<sup>16</sup> Ocse, *Hate Crime Laws. A practical Guide*, cit., 11.

mini d'odio e che si riverbera, ineludibilmente, anche sui discorsi d'odio. Innanzitutto, si rileva una duplicità quanto ai destinatari delle condotte, volte a ledere non soltanto il singolo individualmente inteso, bensì concepito alla luce della sua riconducibilità ad una collettività. Tale profilo si riflette sulle ragioni che spingono l'autore all'azione: l'*animus* non si rivolge tanto alla vittima intesa *uti singuli*, quanto piuttosto quale mero *nuncius* del gruppo di appartenenza a cui è realmente indirizzato il contenuto del messaggio d'odio<sup>17</sup>. Ne deriva che tali condotte, definite dalla stessa Ocse "simboliche"<sup>18</sup>, non arrecano pregiudizio soltanto alla sfera personale, bensì sono in grado di danneggiare la coesione e la stabilità di una società democratica, eterogenea e multiculturale, minando la sopravvivenza dell'uguaglianza e della pari dignità dei suoi componenti. Quale corollario, i discorsi d'odio online, rinvenendo la propria matrice negli *bate crimes* e, più nello specifico, nell'*bate speech*, non costituiscono nuovi illeciti, ma configurano unicamente illeciti diversi quanto alle modalità di manifestazione di fattispecie già esistenti. Essi presentano un *quid pluris*, considerato che essi si manifestano nella dimensione della rete, che li ha trasformati e implementati sotto il profilo della permanenza, della viralità, del ritorno imprevedibile (considerato che si tratta di messaggi itineranti), dell'anonimato e del carattere transazionale che assumono le condotte<sup>19</sup>. Per questo motivo, si è posta la necessità di procedere con una normazione *ad hoc* per contrastare fenomeni già noti e inizialmente concepiti a livello normativo nella sola dimensione reale<sup>20</sup>.

Nel corso dell'ultimo ventennio, la domanda di regolamentazione è aumentata in modo significativo<sup>21</sup>, sia nella dimensione offline sia in quella online, e l'incremento dei discorsi d'odio anche sul web ha messo in luce le lacune normative riguardanti la commissione dei crimini in commento anche nel mondo reale. È stato osservato come Internet, da luogo definito, alle origini, intrinsecamente democratico, si è rivelato, invece, in grado di porre in pericolo il fondamentale presidio di garanzia proprio dello Stato inteso come collettività, considerata l'irrefrenabile e quasi incontrollabile capacità di diffondere disinformazione e di incentivare ostilità<sup>22</sup>.

Conseguentemente, al pari dei legislatori nazionali, anche il legislatore europeo si è interessato ai rapporti tra diritto e realtà digitale, altresì, in considerazione della progressiva crescita di potere delle piattaforme digitali e delle criticità legate all'esercizio della libertà di espressione degli utenti nella nuova società dell'informazione. Nell'ecosistema del web, i discorsi d'odio si diffondono tramite i gestori di dati, che ne rafforzano

<sup>17</sup> Per un maggiore approfondimento di tali profili, vedasi L. D'Amico, *Le forme dell'odio*, cit., 3.

<sup>18</sup> Ocse, *Hate Crime Laws. A practical Guide*, cit., 17.

<sup>19</sup> G. Ziccardi, *Il contrasto dell'odio online: possibili rimedi*, cit. 42-44-45.

<sup>20</sup> Si rimanda, ai fini di un maggiore approfondimento della questione, a Oliveri F., *Diritti degli internauti, obblighi degli Stati, responsabilità delle piattaforme digitali: problemi regolativi in materia di odio online*, in *Teoria e Critica della Regolazione Sociale/Theory and Criticism of Social Regulation*, 2, 23, 2021, 116. Per una analisi dettagliata dei caratteri distintivi del discorso d'odio online vedasi UNESCO, *Countering online hate speech*, Parigi, 2015.

<sup>21</sup> In senso difforme, altri rilevano la non necessarietà di una legislazione specifica, atteso che i medesimi interessi possono essere perimenti protetti dalla legislazione di carattere generale: F. Easterbrook, *Cyberspace and the Law of the Horse*, in *University of Chicago Legal Forum*, 1, 1996, 207 e 208.

<sup>22</sup> Oliveri F., *Diritti degli internauti, obblighi degli Stati, responsabilità delle piattaforme digitali*, cit., 106.

notevolmente la portata, tanto dei comportamenti quanto dei gestori stessi. Di tale fenomeno hanno beneficiato le principali aziende tecnologiche, che sono divenute attori di primaria grandezza nelle dinamiche sociali, politiche e culturali e hanno acquisito un potere senza precedenti, la cui regolamentazione assume rilevanza pubblicistica<sup>23</sup>. Il Trattato sul funzionamento dell'Unione europea contempla, agli artt. 9 e 10 - tra gli obiettivi in esso annoverati - la promozione della lotta contro ogni tipo di discriminazione e all'art. 19 assegna chiara rilevanza all'esigenza di contrastare ogni condotta di tal sorta<sup>24</sup>.

L'intervento legislativo europeo in questa materia trova fondamento giuridico in una duplicità di disposizioni. In primo luogo, si segnala la rilevanza dell'art. 83, par. 1, TFUE<sup>25</sup>, che attribuisce alle istituzioni europee una competenza penale "diretta"<sup>26</sup>. Ne deriva che l'Unione europea detiene il potere di richiedere agli Stati membri l'adozione di norme incriminatrici ogniquale volta vengano in rilievo sfere di criminalità particolarmente grave e che presentano una dimensione transnazionale, tra le quali il Trattato annovera anche le forme di "criminalità informatica". Questa nozione, al pari di quella di "cybercrime", lungi dal rivenire una definizione giuridica univoca, abbraccia una molteplicità di comportamenti lesivi di interessi penalmente rilevanti riconducibili alla categoria dei "reati informatici". Nel dettaglio, la criminalità informatica include non soltanto i "reati informatici in senso stretto", ma anche i "reati informatici in senso lato". Se, da un lato, i primi presentano elementi di tipizzazione connessi a procedimenti di automatizzazione dei dati o a una connotazione squisitamente tecnologica, dall'altro, i secondi si sostanziano in figure criminose che, pur non recando elementi di tipizzazione legislativa tecnologicamente caratterizzati, possono essere applicate a fatti commessi tramite la tecnologia. Tali reati potrebbero essere commessi anche offline, al di fuori del cyberspazio; tuttavia, ricevono un contributo rilevante, in termini di capacità lesiva, dall'utilizzo di dati, dispositivi o strutture informatiche<sup>27</sup>.

---

<sup>23</sup> Ivi, 105-106.

<sup>24</sup> P. Falletta, *Analisi normativa in tema di contrasto agli hate speech su Internet e i social media = Regulatory analysis in the field of contrast to hate speech on the Internet and social media*, in *H-ermes. Journal of Communication*, 23, 2023, 29.

<sup>25</sup> In forza del quale «Il Parlamento europeo e il Consiglio, deliberando mediante direttive secondo la procedura legislativa ordinaria, possono stabilire norme minime relative alla definizione dei reati e delle sanzioni in sfere di criminalità particolarmente grave che presentano una dimensione transnazionale derivante dal carattere o dalle implicazioni di tali reati o da una particolare necessità di combatterli su basi comuni. Dette sfere di criminalità sono le seguenti: terrorismo, tratta degli esseri umani e sfruttamento sessuale delle donne e dei minori, traffico illecito di stupefacenti, traffico illecito di armi, riciclaggio di denaro, corruzione, contraffazione di mezzi di pagamento, criminalità informatica e criminalità organizzata».

<sup>26</sup> L. Picotti, *La nozione di "criminalità informatica" e la sua rilevanza per le competenze penali europee*, in *Rivista trimestrale di diritto penale dell'economia*, 4, 2011, 827.

<sup>27</sup> A. Mattarella, *Il cybercrime nell'ordinamento italiano e le nuove prospettive dell'Unione europea e delle Nazioni Unite*, in *Diritto penale e Processo*, 6, 2022, 809-810, che, alla nota 2, in riferimento agli interventi legislativi operati in ambito europeo in relazione a profili di diritto penale sostanziale concernenti la lotta alla criminalità informatica, rinvia alle seguenti direttive europee: direttiva 95/46/CE sulla tutela dei dati personali, direttive in materia di protezione dei diritti d'autore e, in particolare, direttiva 2001/29/CE; la direttiva 2000/31/CE sul commercio elettronico, oltre che alle decisioni quadro contro gli attacchi informatici (2005/222/GAI), contro lo sfruttamento sessuale di minori e la pedopornografia (2004/68/GAI), contro il terrorismo (2002/475/GAI, parzialmente riformata dalla decisione 2008/919/GAI).

Questa impostazione teorica trova riscontro, tra gli altri, nell'art. 14 della "Convenzione sul Cybercrime", le cui disposizioni processuali si applicano a tutti gli illeciti commessi attraverso sistemi informatici e a quelli per i quali è necessaria la raccolta della prova elettronica<sup>28</sup>, nonché negli artt. 171, lett. a)-bis e 171-ter, comma 2, lett. a-bis), l. 633/1941 (c.d. "Legge sul diritto d'autore"), che sanzionano la diffusione abusiva realizzata mediante l'immissione in un sistema di reti telematiche di un'opera dell'ingegno protetta<sup>29</sup>.

Alla luce di quanto detto e in virtù dell'ampiezza della nozione di "criminalità informatica", la stessa è in grado di abbracciare tutte le fattispecie commesse mediante l'apporto della tecnologia, che non deve costituirne il mezzo essenziale, essendo sufficiente che venga utilizzata incidentalmente per la commissione del reato. Di conseguenza, anche i discorsi d'odio, ogniqualvolta realizzati nell'ambiente digitale, ricadono nella categoria dei c.d. "eurocrimini"<sup>30</sup>, in relazione ai quali, l'Unione europea detiene una potestà legislativa "diretta" ai sensi dell'art. 83, par. 1, TFUE<sup>31</sup>.

Secondariamente, molte volte la Commissione europea ha proposto di ampliare l'elenco dei c.d. "reati europei" ai crimini d'odio e all'incitamento all'odio, anche online<sup>32</sup>, mediante il ricorso all'art. 83, par. 2, TFUE<sup>33</sup>, che prevede gli ambiti di competenza penale "indiretta" dell'Unione europea<sup>34</sup>, trattandosi di forme di criminalità particolarmente grave, la cui commissione dà luogo ad un forte impatto sulla persona e sulla

---

Sul piano processuale, rinvia, invece, alle direttive sul mandato d'arresto europeo (2002/584/GAI) e sull'applicazione del principio del reciproco riconoscimento delle decisioni di confisca (2006/783/GAI), che includono la "criminalità informatica" nelle liste di reati per cui si prescinde, in conformità con il principio del mutuo riconoscimento, dal requisito della doppia incriminazione per l'esecuzione diretta dei provvedimenti emessi dall'autorità giudiziaria dello Stato richiedente.

<sup>28</sup> Art. 14 della Convenzione sulla criminalità informatica del Consiglio d'Europa, c.d. "Convenzione sul Cybercrime", Budapest, 23 novembre 2001, per un cui maggiore approfondimento a livello internazionale si rinvia a K. Ghimire, *International Perspective of Cyber Law: Specially Focused on Cybercrime Convention*, in *NJA Law Journal*, 2021, 307 ss.

<sup>29</sup> R. Flor, *Tutela penale e autotutela tecnologica dei diritti d'autore nell'epoca di internet. Un'indagine comparata in prospettiva europea ed internazionale*, Cedam, 2010.

<sup>30</sup> Per tali intendendosi tra le «sfere di criminalità particolarmente grave che presentano una dimensione transnazionale» ai sensi dell'art. 83, §1, co. 2 del Trattato sul Funzionamento dell'Unione europea, come precisato da N. Cardinale, *Il Parlamento europeo chiede l'inserimento della violenza di genere tra i c.d. eurocrimini ai sensi dell'art. 83, § 1, co. 2 del TFUE* in [sistemapenale.it](http://sistemapenale.it), 07 dicembre 2021.

<sup>31</sup> A conferma delle considerazioni sopra svolte, si segnala la rilevanza della direttiva (UE) 2024/1385 del Parlamento europeo e del Consiglio, del 14 maggio 2024, sulla lotta alla violenza contro le donne e alla violenza domestica, adottata ai sensi dell'art. 83, par. 1, TFUE.

<sup>32</sup> P. Falletta, *Analisi normativa in tema di contrasto agli hate speech su Internet e i social media*, cit., 34 - 35; Per un maggior approfondimento vedasi anche *Study to support the preparation of the European Commission's initiative to extend the list of EU crimes in Article 83 of the Treaty on the Functioning of the EU to hate speech and hate crime*, Commissione europea, novembre 2021, 32 ss.

<sup>33</sup> In forza del quale «Allorché il ravvicinamento delle disposizioni legislative e regolamentari degli Stati membri in materia penale si rivela indispensabile per garantire l'attuazione efficace di una politica dell'Unione in un settore che è stato oggetto di misure di armonizzazione, norme minime relative alla definizione dei reati e delle sanzioni nel settore in questione possono essere stabilite tramite direttive. Tali direttive sono adottate secondo la stessa procedura legislativa ordinaria o speciale utilizzata per l'adozione delle misure di armonizzazione in questione, fatto salvo l'articolo 76».

<sup>34</sup> L. Picotti, *La nozione di "criminalità informatica" e la sua rilevanza per le competenze penali europee*, cit., 829.

comunità di appartenenza<sup>35</sup>, da cui deriva una seria minaccia per i valori democratici. Tuttavia, nessuna di tali basi giuridiche è stata finora impiegata per criminalizzare le condotte di *online hate speech*. L'approccio europeo è, difatti, storicamente risultato timido, sostanziosamente nell'adozione, a alternanza, di strumenti di *hard law*, i cui destinatari sono apparsi (con esito deludente, anche se, per alcuni profili, comprensibilmente, per le ragioni sopraccitate) i prestatori di servizi informativi (d'ora in poi, PSI), insinuando il dubbio che si trattasse di provvedimenti principalmente funzionali alla tutela del mercato interno anziché dei diritti fondamentali e, nella specifica materia penalistica, di strumenti di *soft law*, ossia atti di cooperazione giudiziaria<sup>36</sup>.

Ne è risultata l'adozione di una metodologia minimalista<sup>37</sup>, innestata sulla regolamentazione del mondo digitale esclusivamente nei suoi tratti essenziali (in particolare, la responsabilità dei PSI) e non a mezzo di "scelte forti"<sup>38</sup> di incriminazione delle condotte in parola, la cui diffusione viene contrastata a livello europeo in un'ottica preventiva e non repressiva<sup>39</sup>, cui corrisponde, quale contraltare, anche la rilevanza strategica dei soggetti appena menzionati nella circolazione di contenuti nel cyberspazio<sup>40</sup>.

Al contrario, è stato rilevato che è la potenza stessa della tecnologia ad esigere la messa a punto di misure di *governance* di Internet mediante l'adozione di strumenti di *hard law*, atteso che quelli di *soft law*, seppure utili e recanti propositi assai meritevoli di considerazione, non risultano sufficienti ad assicurare una adeguata salvaguardia dei diritti degli utenti<sup>41</sup>.

I primi strumenti di *hard law* maggiormente risalenti nel tempo erano stati concepiti a livello di diritto derivato: il riferimento è alla direttiva 2000/31/CE del Parlamento europeo e del Consiglio dell'8 giugno 2000 volta all'incentivazione del commercio elettronico<sup>42</sup>, che è stata definita «l'unico atto di respiro generale specificamente dedicato ai servizi digitali»<sup>43</sup>. Essa, infatti, ha costituito lo strumento principale che ha

<sup>35</sup> Commissione europea, *Un'Europa più inclusiva e protettiva: estendere l'elenco dei reati riconosciuti dall'UE all'incitamento all'odio e ai reati generati dall'odio*, COM (2021)777 final.

<sup>36</sup> I. Anrò, *Online hate speech: la prospettiva dell'Unione europea tra regolamentazione della condotta dei prestatori di servizi intermediari e ricorso al diritto penale*, in *Osservatorio sulle fonti*, 16, 1, 2023, 14-15.

<sup>37</sup> A riguardo vedasi O. Lynskey, *Regulating Platforms Power*, in *LSE Law, Society and Economy Working Papers*, 1, 2017.

<sup>38</sup> Alle quali ha fatto riferimento, con specifico riguardo alla regolazione di Internet, M. Bassini, *Libertà di espressione e social network, tra nuovi "spazi pubblici" e "poteri privati" Spunti di comparazione*, cit., 73.

<sup>39</sup> Vedasi, sul punto, quanto si dirà più compiutamente infra in relazione agli strumenti di moderazione algoritmica nel par. 7.

<sup>40</sup> M. Bassini, *Libertà di espressione e social network, tra nuovi "spazi pubblici" e "poteri privati" Spunti di comparazione*, cit., 73.

<sup>41</sup> C. Confortini, *Diffamazione e discorso d'odio in internet Note a margine di Cass.*, [ord.] 16.5.2023, n. 13411, in *Persona e mercato*, 4, 2023, 703, che richiama sul punto L. Floridi, *The End of an Era: from Self-Regulation to Hard Law for the Digital Industry*, in *Philosophy & Technology*, 34, 2021, 621.

<sup>42</sup> Direttiva 2000/31/CE del Parlamento europeo e del Consiglio dell'8 giugno 2000 *relativa a taluni aspetti giuridici dei servizi della società dell'informazione, in particolare il commercio elettronico, nel mercato interno* (Direttiva sul commercio elettronico), in GUUE L 178 del 17 luglio 2000, p. 1 ss. Per un maggior approfondimento in relazione all'evoluzione della normativa europea vedasi I. Anrò, *Online hate speech*, cit., 16 ss. e G. Morgese, *Moderazione e rimozione dei contenuti illegali online nel diritto dell'Unione europea*, in *federalismi.it*, 12 gennaio 2022, 80 ss.

<sup>43</sup> M. Bassini, *Libertà di espressione e social network, tra nuovi "spazi pubblici" e "poteri privati" Spunti di*

legiferato in merito agli spazi di autonomia e ai profili di responsabilità dei gestori delle piattaforme digitali. È d'obbligo rilevare come, poiché è trascorso oltre un ventennio dalla pubblicazione, detta direttiva non può considerarsi (e non si poteva considerarla nemmeno allora) capace di disciplinare efficacemente le diverse condotte d'odio e discriminatorie che si realizzano a mezzo di Internet<sup>44</sup>.

Inoltre, la scarna e datata disciplina ivi contenuta si connota per essere stata «mobilitata da più di un decennio di giurisprudenza quantomeno creativa»<sup>45</sup> della Corte di Giustizia dell'Unione europea, la quale ha rivoluzionato il proprio orientamento, spingendosi a riconoscere potere sempre maggiore ai PSI, tenuti, dietro ordine dello Stato membro, ad eliminare i contenuti oggetto di segnalazione. Da tale circostanza è anche derivata la configurabilità, in capo ai medesimi, di un dovere di sorveglianza sul materiale circolante sul *web*, oltre che un certo margine di discrezionalità quanto alla qualificazione del medesimo<sup>46</sup>.

Il ricorso alla tecnica incriminatrice propria del diritto penale è stato oggetto di un tentativo di utilizzo da parte dell'Unione europea mediante la Decisione quadro 2008/913/GAI del Consiglio del 28 novembre 2008<sup>47</sup>, la quale obbliga gli Stati membri a punire, mediante l'adozione di sanzioni effettive, proporzionate e dissuasive, le condotte di istigazione pubblica alla violenza o all'odio, indipendentemente dal fatto che esse si configurino online oppure offline, in ragione della razza, della religione e dell'ascendenza ovvero dell'origine nazionale o etnica<sup>48</sup>. La Decisione quadro mira a realizzare un'opera di armonizzazione minima delle legislazioni nazionali degli Stati membri, i quali possono decidere di ampliare la sfera dei motivi di odio che giustificano una repressione penale. Le difficoltà riscontrate nella sua lunga gestazione, sono riconducibili, principalmente, all'esigenza di realizzare un bilanciamento tra la libertà di espressione e la lotta contro il razzismo e la xenofobia nel complesso quadro delle diverse cornici giuridiche nazionali<sup>49</sup>, che presentano tratti differenziali anche in

---

*comparazione*, cit., 71.

<sup>44</sup> Per un approfondimento circa i punti salienti della legislazione europea in commento, vedasi P. Falletta, *Controlli e responsabilità dei social network sui discorsi d'odio online*, cit., 150-151. A ciò aggiungasi che, la detta direttiva si connota per una contraddizione in termini: se, da una parte, l'art. 15 statuisce l'assenza di un obbligo di sorveglianza in capo ai PSI, dall'altra parte si prevedono, nelle *Disposizioni finali*, obblighi di rimozione dei contenuti a fronte di ordine in tal senso delle autorità pubbliche ed oneri di notifica (c.d. «*notice and take down*») che consentono alle dette autorità di individuare e prevenire attività illecite ai sensi dell'art. 21. Si è, così, consentita una certa espansione della libertà di espressione, creando significative esenzioni di responsabilità in capo ai PSI e riconoscendo in capo ai medesimi un obbligo di rimozione a posteriori, una volta appurata l'illiceità dei contenuti ospitati.

<sup>45</sup> M. Bassini, *Libertà di espressione e social network, tra nuovi "spazi pubblici" e "poteri privati"* *Spunti di comparazione*, in questa *Rivista*, 2, 2021, 71.

<sup>46</sup> P. Falletta, *Analisi normativa in tema di contrasto agli hate speech su Internet e social media*, cit., 30-31.

<sup>47</sup> Decisione quadro 2008/913/GAI del Consiglio del 28 novembre 2008 sulla lotta contro talune forme ed espressioni di razzismo e xenofobia mediante il diritto penale, in GUUE, L. 328 del 6 dicembre 2008.

<sup>48</sup> P. Falletta, *Analisi normativa in tema di contrasto agli hate speech su Internet e social media*, cit., 29, che evidenzia che, successivamente, il Parlamento europeo, con la risoluzione del 14 marzo 2013 ha sottolineato la necessità di includere anche le manifestazioni di antisemitismo, intolleranza religiosa, antiziganismo, omofobia e transfobia nell'ambito applicativo della Decisione quadro in esame.

<sup>49</sup> Difatti, numerose e serie sono state le difficoltà che si sono frapposte all'adozione della Decisione; a

termini sociali e culturali. Difatti, è stato rilevato che uno dei principali profili di problematicità e che meriterebbe maggiore attenzione, riguarda la dimensione culturale della questione. Si ravvisa la necessità di sensibilizzare sia le popolazioni sia le stesse piattaforme digitali, al fine di aumentare la consapevolezza delle potenzialità lesive insite nelle conversazioni d'odio online<sup>50</sup>.

Più di recente, le istituzioni europee hanno adottato un approccio di autoregolamentazione e co-regolamentazione, in un'ottica di proficua collaborazione con le piattaforme online, a cui è stato attribuito il potere di censurare i contenuti illeciti<sup>51</sup>, nonostante trattasi di comunità virtuali, che perseguono finalità commerciali<sup>52</sup>, aliene dall'interferenza dei poteri statali<sup>53</sup> e a qualsiasi circuito democratico<sup>54</sup>.

Come anticipato, una delle principali criticità emerse ha riguardato l'assimilazione dei PSI ai caratteri della pubblicità, in quanto diventati centri di straordinaria concentrazione di potere, conoscenza e ricchezza<sup>55</sup>.

Su questa scia si è posto anche il Regolamento 2022/2065/UE del Parlamento europeo e del Consiglio del 19 ottobre 2022<sup>56</sup>, noto come "Digital Services Act" o "DSA", con efficacia dal 17 febbraio 2024. L'intervento legislativo europeo si è radicato nell'esigenza improrogabile di garantire uniformità tra le legislazioni: la presenza di quadri giuridici differenti è apparsa non più sostenibile per regolamentare un fenomeno tipicamente transfrontaliero e che incide negativamente anche sul mercato interno. Sono così state ideate innovative forme di collaborazione tra le istituzioni e i PSI, tenendo conto del loro ambito di azione e della possibilità che si verificano condotte d'odio, ovvero di discriminazione, all'interno delle piattaforme<sup>57</sup>.

Il Regolamento ha disegnato una commistione tra *private* e *public governance* del discorso pubblico online, mediante il consolidamento della cooptazione tra poteri pubblici e

---

ciò aggiungasi il consistente ritardo riscontrato da numerosi Stati Membri in sede di trasposizione: basti pensare che l'Italia vi ha provveduto soltanto nel 2017 e che altri, quali la Romania e l'Estonia, sono stati destinatari di una lettera di messa in mora ex art. 258 TFUE dal parte della Commissione europea in ragione del perdurante mancato recepimento, a conferma della delicatezza del bilanciamento da operarsi alla luce di differenti tradizioni culturali, sociali e giuridiche, così, I. Anrò, *Online hate speech*, cit., 31-32.

<sup>50</sup> G. Ziccardi, *Il contrasto dell'odio online: possibili rimedi*, cit. 42-45.

<sup>51</sup> P. Falletta, *Analisi normativa in tema di contrasto agli hate speech su Internet e social media*, cit., 30.

<sup>52</sup> Nell'era digitale, si parla, infatti, di compenetrazione con effetti orizzontali della libertà di espressione, così M. Bassini, *Libertà di espressione e social network, tra nuovi "spazi pubblici" e "poteri privati" Spunti di comparazione*, cit., 68 e 70 che richiama S. Gardbaum, *The "Horizontal Effect" of Constitutional Rights*, in *Michigan Law Review*, 102, 2003, 387 ss.; nel panorama italiano, vedasi, sul punto, anche gli scritti di A. Gentili, *Diritti fondamentali e rapporti contrattuali. Sulla efficacia orizzontale della Convenzione europea dei diritti dell'uomo*, in *Nuova giurisprudenza civile commentata*, 1, 2016, 183 ss.; A. Zoppini, *Il diritto privato e le "libertà fondamentali" (Principi e problemi della Drittwirkung nel mercato unico)*, in *Rivista di diritto civile*, 3, 2016, 712 ss.

<sup>53</sup> M. Bassini, *Libertà di espressione e social network, tra nuovi "spazi pubblici" e "poteri privati" Spunti di comparazione*, cit., 71.

<sup>54</sup> C. Confortini, *Diffamazione e discorso d'odio in internet*, cit., 699.

<sup>55</sup> *Ibid.*

<sup>56</sup> Regolamento (UE) 2022/2065 del Parlamento europeo e del Consiglio del 19 ottobre 2022 relativo a un mercato unico dei servizi digitali e che modifica la direttiva 2000/31/CE (regolamento sui servizi digitali), in GUUE L 277 del 27 ottobre 2022, 1 ss.

<sup>57</sup> P. Falletta, *Controlli e responsabilità dei social network sui discorsi d'odio online*, in questa *Rivista*, 1, 2020, 147.

privati<sup>58</sup> e ciò in ragione dell'ampiezza acquisita dal mercato digitale, essenziale strumento di informazione e di partecipazione al dibattito politico e culturale. L'approccio adottato è stato definito "risk-based", in quanto vengono fissate regole di condotta e meccanismi di *enforcement* diversificati in funzione dell'attività svolta dai *providers*, della loro dimensione e della loro capacità di incidere sui valori fondamentali dell'Unione. Il legislatore europeo ha, quindi, optato per il superamento di una disciplina unitaria e per l'adozione di paradigmi regolatori differenziati, allo scopo di assicurare la legalità delle comunicazioni che hanno luogo sul web<sup>59</sup>.

Nel complesso, gli Stati membri, attraverso l'introduzione delle normative descritte, hanno cercato di contrastare «l'opacità dello spazio digitale e garantire la costruzione di una solida democrazia digitale»<sup>60</sup>, tentando di colmare il rapido processo di erosione delle di loro sovranità che le società private stavano realizzando. Nondimeno, l'azione legislativa si è concentrata eccessivamente sui profili strettamente economici legati al corretto funzionamento del mercato interno e del mercato digitale, anziché sulla repressione delle condotte in grado di porre in pericolo la sopravvivenza delle democrazie europee. Ne è seguita una stretta connessione tra i provvedimenti legislativi riguardanti l'ecosistema digitale e il concetto di sovranità<sup>61</sup>.

Quanto al formante giurisprudenziale europeo, si evidenzia come la Corte di Giustizia abbia tratto ispirazione dai risultati della Corte Europea dei diritti dell'Uomo, che storicamente hanno manifestato diffidenza nei confronti del *cyberspace*<sup>62</sup>. Nel solco dei propri precedenti, i giudici di Strasburgo hanno confermato che l'adozione di misure di contrasto ai discorsi d'odio mediante il ricorso allo strumento del diritto penale deve essere considerata priva di profili di illiceità, apponendo, tuttavia, quale limite, la circostanza che detti discorsi presentino un maggior grado di offensività tale da poterli inquadrare nelle forme di *hate speech*, le cui manifestazioni più gravi costituiscono fattispecie di abuso dell'art. 10<sup>63</sup>.

Pur essendo evidente l'impegno profuso dalle istituzioni europee e, in particolare, dalla Commissione, nel contrasto ai discorsi d'odio online, tale azione presenta tuttora delle criticità, alla luce della mancata adozione di direttive idonee a estendere l'ambito dei c.d. "eurocrimini" e ricondurre tali reati informatici alle materie di legislazione penale europea in forza dell'art. 83, par. 1, TFUE, ovvero, in subordine, dell'art. 83, par. 2, TFUE<sup>64</sup>.

Nel settembre 2020, è stato annunciato, ad opera della Presidente della Commissione

<sup>58</sup> C. Confortini, *Diffamazione e discorso d'odio in internet*, cit., 702-703.

<sup>59</sup> S. Braschi, *Il nuovo Regolamento sui servizi digitali: quale futuro per la responsabilità degli Internet Service Provider?*, in *Diritto Penale E Processo*, 3, 2023, 368-369.

<sup>60</sup> P. Falletta, *Analisi normativa in tema di contrasto agli hate speech su Internet e social media*, cit., 36.

<sup>61</sup> Per un maggior approfondimento della questione, vedasi S. Torregiani, *Il Data Act: una versione europea del Data Nationalism?*, in *Rivista italiana di informatica e diritto*, 5, 2, 2023, 135.

<sup>62</sup> CEDU, *Sanchez c. Francia*, ric. 45581/15 (2021).

<sup>63</sup> Come nel caso del contesto politico, così P. Dunn, *Carattere eccezionale dell'"hate speech" e nuove forme di responsabilità per contenuti di terzi nella giurisprudenza EDU. Nota a C.edu, Sanchez c. Francia, 15 maggio 2023*, in *Osservatorio costituzionale*, 6, 2023, 243-244.

<sup>64</sup> Per un maggior approfondimento a riguardo vedasi I. Anrò, *Online hate speech*, cit., 32 ss.

Ursula Von Der Leyen, l'obiettivo di inserire le condotte di *hate speech* e *hate crime*<sup>65</sup> all'interno di quest'ultimo novero, proposta cui è seguita, in data 18 gennaio 2024, l'approvazione da parte del Parlamento europeo di una Risoluzione volta a «Estendere l'elenco dei reati riconosciuti dall'UE all'incitamento all'odio e ai reati generati dall'odio»<sup>66</sup>.

La Risoluzione evidenzia che tutte le manifestazioni di odio e di intolleranza, inclusi l'incitamento e i reati generati dall'odio, sono incompatibili con i valori dell'Unione, quali la dignità, la libertà, la democrazia, l'uguaglianza, lo Stato di diritto e il rispetto dei diritti umani, compresi i diritti delle persone appartenenti a minoranze, sanciti dall'art. 2 TUE<sup>67</sup>. Oltre a ciò, è stato osservato come la libertà di espressione costituisca un valore fondamentale delle società democratiche e non debba essere limitata in modo ingiustificato; qualsiasi legislazione a riguardo dovrebbe, pertanto, basarsi sui principi di necessità e proporzionalità, così da garantirne l'esercizio in conformità all'art. 11 della Carta dei diritti fondamentali dell'Unione europea, che non potrebbe mai essere utilizzato come scudo per la commissione di reati<sup>68</sup>.

La Risoluzione è stata trasmessa alla Commissione e al Consiglio, esortando quest'ultimo a adottare una Decisione che ampli l'elenco dei c.d. "reati europei", affinché la Commissione possa avviare la seconda fase della procedura di modifica del Trattato sul funzionamento dell'Unione europea. La progettazione di un'iniziativa comune a livello europeo e la realizzazione di un solido quadro unionale, mediante l'adozione di misure globali, rappresenterebbe la risposta più efficace alle sfide di cui sopra<sup>69</sup>, cui, però, osta la necessità che la Decisione venga adottata dal Consiglio all'unanimità, soglia assai difficile da raggiungere, attesa la titolarità, in capo a ciascuno Stato, del potere di veto e in considerazione della diversità dei contesti sociali e culturali che fanno da sfondo ai quadri giuridici degli Stati membri. All'interno di un quadro democraticamente ordinato, plurale e eterogeneo quale quello europeo, un epilogo da ritenersi necessario e imprescindibile viene precluso dall'agire democratico stesso del sistema di funzionamento, che assegna cruciale rilevanza alla voce di ciascuno Stato membro. Per tale ragione, il Parlamento ha raccomandato al Consiglio la modifica dell'art. 83 TFUE, al fine di renderlo soggetto ad una maggioranza qualificata rafforzata, anziché all'attuale unanimità, attraverso l'attivazione della c.d. "clausola passerella"<sup>70</sup> di cui

<sup>65</sup> Comunicazione della Commissione del 9 dicembre 2021 dal titolo «Un'Europa più inclusiva e protettiva: estendere l'elenco dei reati riconosciuti dall'UE all'incitamento all'odio e ai reati generati dall'odio» (COM (2021)0777), citata in I. Anrò, *Online hate speech*, cit., 32 ss.

<sup>66</sup> Risoluzione del Parlamento europeo del 18 gennaio 2024 (P9\_TA (2024)0044), «Estendere l'elenco dei reati riconosciuti dall'UE all'incitamento all'odio e ai reati generati dall'odio» [2023/2068(INI)] (A9-0377/2023).

<sup>67</sup> Ivi, 3.

<sup>68</sup> Ivi, 8.

<sup>69</sup> Come confermato, oltre che dalla Risoluzione in commento, anche dalla Comunicazione della Commissione al Parlamento Europeo e al Consiglio «Un'Europa più inclusiva e protettiva: estendere l'elenco dei reati riconosciuti dall'UE all'incitamento all'odio e ai reati generati dall'odio» COM/2021/777 final.

<sup>70</sup> Risoluzione del Parlamento europeo del 18 gennaio 2024 (P9\_TA (2024)0044), 8.

all'art. 48, par. 2, TUE<sup>71</sup>.

In conclusione, il Parlamento europeo deplora con forza il fatto che sono trascorsi quasi due anni dalla pubblicazione della Comunicazione e che il Consiglio non ha compiuto alcun progresso in merito, sebbene sia stato in grado di estendere rapidamente l'elenco dei c.d. "eurocrimini" per altri fini. Tale inazione risulta ancor più riprovevole alla luce dell'aumento statistico dell'incitamento all'odio e dei reati generati dall'odio<sup>72</sup>.

Alla luce del quadro sopra delineato, emerge con chiarezza che il problema sottostante al mancato avvio di un intervento legislativo mediante lo strumento della direttiva, sulla base dell'art. 83, par. 1 (e, eventualmente, par. 2), TFUE, sia da ricondursi a una volontà di stampo politico. A tal proposito, è stata rilevata la mancanza di una preparazione socioculturale adeguata riguardo il grado di lesività insito nelle condotte in esame. Nel proseguo si approfondirà come diversi Stati europei abbiano criminalizzato il discorso d'odio online, ciascuno secondo la propria sensibilità nazionale. Tuttavia, per un'efficace neutralizzazione non può dirsi sufficiente la sola incriminazione, altresì, considerato che il ricorso alla "clava penale"<sup>73</sup> segna il fallimento delle democrazie contemporanee nel loro impegno a educare seriamente e in modo effettivo i propri cittadini al rispetto e all'uguaglianza<sup>74</sup>.

La natura globale di Internet rende impossibile una regolamentazione giuridica onnicomprensiva del cyberspazio e il coinvolgimento di tutti gli attori coinvolti nella catena di propagazione dei contenuti si rivela indefettibile. Nella sua duplice veste di trasmettitore e destinatario, è necessario configurare in capo all'utente sia una specifica responsabilità, sia una specifica forma di tutela<sup>75</sup>. Dunque, è indifferibile cercare delle alternative per limitare la pubblicazione dei discorsi d'odio e minimizzare i pregiudizi arrecati dagli stessi, mediante uno sforzo coordinato che veda la cooperazione delle componenti legislativa, tecnologica e socioculturale. Questa soluzione permetterebbe di mitigare la diffusione senza influire negativamente sul libero flusso di conoscenze, idee e informazioni all'interno della rete<sup>76</sup>.

Per risolvere in maniera definitiva le problematiche connesse alle conversazioni d'odio online, che nascono altresì da fenomeni di emarginazione sociale, diventa cruciale

---

<sup>71</sup> La relativa attivazione dà luogo ad un procedimento di revisione dei Trattati in forma semplificata tramite il quale è possibile disporre il passaggio dalla procedura legislativa speciale a quella ordinaria, ovvero disporre la sostituzione della regola dell'unanimità con quella della maggioranza qualificata per l'assunzione delle delibere da parte del Consiglio. Difatti, il potere deliberativo è attribuito al Consiglio europeo, che lo esercita mediante l'adozione di una decisione assunta all'unanimità, a condizione della mancata opposizione dei Parlamenti nazionali e previa approvazione del Parlamento europeo, che si pronuncia a maggioranza dei suoi membri.

<sup>72</sup> Risoluzione del Parlamento europeo del 18 gennaio 2024 (P9\_TA (2024)0044), 7.

<sup>73</sup> A. Pugiotto, *Le parole sono pietre? I discorsi di odio e la libertà di espressione nel diritto costituzionale*, in *Diritto penale contemporaneo Rivista trimestrale*, 3, 2013, 5.

<sup>74</sup> G. Giorgini Pignatiello, *Profili Comparati e problemi costituzionali della legislazione contro l'omotransfobia. Il caso spagnolo e quello italiano*, in *Diritto pubblico comparato ed europeo*, 4, 2020, 1022.

<sup>75</sup> *Avis relatif à la lutte contre la haine en ligne (A - 2021 - 9)*, NOR: CDHX2122366V, JORF n°0170 du 24 juillet 2021, Texte n° 79, in [legifrance.gouv.fr](http://legifrance.gouv.fr).

<sup>76</sup> J. Banks, *Regulating hate speech online*, in *International Review of Law, Computers & Technology*, 24, 3, 2010, 238-239.

ricorrere a strumenti quali l'educazione digitale, anche già in ambito scolastico, la formazione di tutta la popolazione e, infine, il ruolo della politica legislativa, in grado di sensibilizzare l'opinione pubblica<sup>77</sup>.

Soltanto promuovendo, a livello sia nazionale sia europeo, l'integrazione sociale e le sue potenzialità, nel lungo termine si potrà garantire coesione e sicurezza, impedire la marginalizzazione delle singole appartenenze politiche<sup>78</sup> e radicare la cultura del contrasto ai discorsi d'odio online.

Il primo baluardo contro la fitta trama di ostilità reciproche che pervade il mondo digitale<sup>79</sup> è l'educazione e la consapevolezza del digitale stesso. Tale assunto può aiutare a spiegare per quale ragione lo strumento forte del diritto penale, seppure importante, non è sufficiente per contrastare efficacemente le condotte d'odio, obiettivo essenziale di ogni democrazia plurale.

### 3. Il quadro normativo e giurisprudenziale italiano

Volendo ora indagare quale sia il ruolo legittimamente ascrivibile al diritto penale nel contrasto alle condotte d'odio online, ci si concentrerà sull'analisi normativa e giurisprudenziale interna.

In Italia non esiste uno strumento normativo che riconosce e sanziona in modo specifico l'*hate speech* (online); tale condotta è perseguita (seppur con difficoltà) quando ricade sotto il cappello di fattispecie penali esistenti, come quelle previste dagli artt. 604-bis c.p. (propaganda e istigazione a delinquere per motivi di discriminazione razziale etnica e religiosa) ovvero art. 595 c.p. (diffamazione), aggravata ai sensi dell'art. 604-ter c.p.. Di tali disposizioni parleremo più diffusamente infra, ritenendosi ora preliminarmente opportuno contestualizzare l'evoluzione normativa con riferimento alla criminalizzazione delle condotte d'odio genericamente intese, la quale ha condotto all'inserimento delle citate fattispecie entro il perimetro del Codice penale.

Nel 1975 l'Italia con la l. 654/1975<sup>80</sup>, c.d. legge Reale, ha dato esecuzione alla Convenzione ONU «sulla eliminazione di tutte le forme di discriminazione razziale», firmata a New York nel 1966 (G.U. n. 337/1975)<sup>81</sup>. Tale, ha introdotto nel nostro ordinamento

---

<sup>77</sup> *Avis relatif à la lutte contre la haine en ligne* (A - 2021 - 9), cit.

<sup>78</sup> D. Piccione, *L'espressione del pensiero ostile alla democrazia, tra diritto penale dell'emotività e psicologia collettiva*, in questa *Rivista*, 3, 2018, 87-88.

<sup>79</sup> A. Spena, *La parola (-) odio*, cit., 577-578.

<sup>80</sup> Prima dell'entrata in vigore della c.d. L. Reale, con la l. 645/1952 (c.d. L. Scelba) l'Italia ha dato attuazione alla XII disposizione transitoria e finale della Costituzione imponendo il divieto di riorganizzazione del partito fascista. Successivamente, con la L. 962 del 9 ottobre 1967, nota come legge «per la prevenzione e repressione del delitto di genocidio», l'Italia ha punito (art. 8) «chiunque pubblicamente istiga a commettere alcuno dei delitti preveduti negli articoli da 1 a 5, è punito per il solo fatto dell'istigazione, con la reclusione da tre a dodici anni. La stessa pena si applica a chiunque pubblicamente fa l'apologia di alcuno dei delitti preveduti nel comma precedente».

<sup>81</sup> L'art. 3 della L. in esame incriminava chiunque diffondesse, in qualsiasi modo, idee fondate sulla superiorità o sull'odio razziale o incitasse in qualsiasi maniera alla discriminazione o alla commissione di atti di violenza o provocazione alla violenza nei confronti di certe persone in quanto appartenenti ad un gruppo nazionale, etnico o razziale.

autonome fattispecie di reato caratterizzate dalla matrice razzista: la propaganda razzista, l'incitamento alla discriminazione razziale e agli atti di violenza nei confronti dei medesimi soggetti e, infine, la costituzione di associazioni e di organizzazioni con scopo di incitamento all'odio o alla discriminazione razziale. Condotte tutte riconducibili alla nozione di *bate speech* e di *bate crime*, in rapporto di specie a genere<sup>82</sup>.

Successivamente, la l. 205/1993, c.d. legge Mancino, ha introdotto tra i motivi di incitamento alla commissione di atti discriminatori (art. 3 della l. 654/1975<sup>83</sup>) sia il fattore etnico, idoneo a connotare la propaganda, che quello religioso. Tale normativa è altresì nota per aver inserito, sempre nell'ambito dell'art. 3 della legge Reale, l'aggravante dell'odio razziale (etnico, nazionale, religioso)<sup>84</sup>.

In questo complesso impianto antidiscriminatorio si è innestata, nel 2006, la l. 85/2006, recante «modifiche al Codice penale in materia di reati di opinione», la quale ha nuovamente modificato il citato art. 3, sostituendo il termine «diffusione» con quello di «propaganda» ed il termine «incitamento» con quello di «istigazione»<sup>85</sup>. La modifica non è di poco conto se si considera che la qualificazione del reato deve oggi corrispondere a condotte di maggiore gravità (propaganda e istigazione in luogo della diffusione e dell'incitamento).

Nel 2008, poi, la Decisione Quadro 2008/913/GA<sup>86</sup> (attuata in Italia tramite la l. 115/2016) ha modificato nuovamente l'art. 3, introducendo un successivo comma 3-bis, attraverso il quale è stata prevista la misura della reclusione da due a sei anni nei casi in cui la propaganda, l'istigazione e l'incitamento, comportino il concreto pericolo di diffusione e si fondino «in tutto o in parte sulla negazione della Shoah o dei crimini di genocidio, dei crimini contro l'umanità e dei crimini di guerra, come definiti dagli articoli 6, 7 e 8 dello statuto della Corte penale internazionale, ratificato ai sensi della legge 12 luglio 1999, n. 232».

Solo un anno dopo, nel 2017 (l. 167/2017), si è proceduto ad un'ultima modifica del comma 3-bis attraverso l'introduzione, dopo il riferimento alla negazione, della «minimizzazione in modo grave o l'apologia», con il chiaro fine di omologarsi ulteriormente

<sup>82</sup> Il testo originario dell'art. 3 della l. in esame puniva: «chi incita in qualsiasi modo alla discriminazione, o incita a commettere o commette atti di violenza o di provocazione alla violenza, nei confronti di persone perché appartenenti a un gruppo nazionale, etnico o razziale». Cfr. paragrafo 2 del presente scritto.

<sup>83</sup> Il nuovo art. 3 della l. 654/1975, come novellato dalla c.d. legge Mancino puniva: «a) con la reclusione fino a tre anni chi diffonde in qualsiasi modo idee fondate sulla superiorità o sull'odio razziale o etnico, ovvero incita a commettere o commette atti di discriminazione per motivi razziali, etnici, nazionali o religiosi; b) con la reclusione da sei mesi a quattro anni che, in qualsiasi modo, incita a commettere o commette violenza o atti di provocazione alla violenza per motivi razziali, etnici, nazionali o religiosi».

<sup>84</sup> L. Goisis, *Crimini d'odio*, cit.13.

<sup>85</sup> Viene punito non più chi «diffonde in qualsiasi modo» ma chi «propaganda idee fondate sulla superiorità o sull'odio razziale o etnico»; non più chi «incita», ma chi «istiga a commettere o commette atti di discriminazione per motivi razziali, etnici, nazionali o religiosi»; nonché chi «istiga» anziché chi «incita» a commettere o commette atti di violenza o atti di provocazione alla violenza per motivi razziali, etnici o religiosi. Si segnala che la Corte di cassazione con la pronuncia n. 34713 del 2016 ha sostenuto che vi è continuità normativa tra il testo anteriore alle modifiche ex l. 85/2006 e quello successivo (per un'analisi completa sul punto si segnala il commento alla citata pronuncia, in *Rivista Penale*, 10, 2016, 895 ss.).

<sup>86</sup> Decisione Quadro del Consiglio del 28 novembre 2008.

alle indicazioni contenute nella citata Decisione Quadro.

L'approdo di tale evoluzione normativa è rappresentato dal d.lgs. 21/2018, il quale ha trasferito la disciplina entro il perimetro del Codice penale (titolo XII «dei delitti contro la persona», capo III, «dei delitti contro la libertà individuale», sezione I-bis, dedicata ai «delitti contro l'uguaglianza»<sup>87</sup>), introducendo gli artt. 604-bis c.p., rubricato «propaganda e istigazione a delinquere per motivi di discriminazione razziale etnica e religiosa» e 604-ter c.p., il quale disciplina un'autonoma circostanza aggravante, per i reati commessi per finalità di discriminazione o di odio etnico, nazionale, razziale o religioso, ovvero al fine di agevolare l'attività di organizzazioni, associazioni, movimenti o gruppi che hanno tra i loro scopi le medesime finalità.

Le citate figure delittuose sono comunemente definite crimini d'odio, estrinsecandosi non solo nella loro componente fattiva, intesa come compimento materiale di un atto discriminatorio, bensì anche in quella dialettica ed epistemica, nota come *hate speech*.

L'inserimento nel Codice penale di tali disposizioni<sup>88</sup> deve essere apprezzato in forza del maggior rilievo attribuito ai relativi precetti, i quali mirano a tutelare l'essere umano come individuo, in quanto tale non discriminabile per motivi razziali, etnici, nazionali o religiosi<sup>89</sup>. Appare dunque inevitabile il contrasto che potrebbe verificarsi tra il diritto alla libera manifestazione del pensiero<sup>90</sup> (disciplinato a livello sovranazionale e interno, dagli artt. 10 CEDU e 21 Cost.), il divieto di abuso del diritto (art. 17 CEDU) e la pari dignità dei consociati (art. 3 Cost.). E proprio nella tutela della pari dignità sociale deve essere individuato il perno essenziale della normativa in esame.

L'individuazione del bene giuridico tutelato dalle disposizioni sopra richiamate ha assunto, e continua ad assumere, una rilevanza decisiva. Fino a quando queste disposizioni sono state interpretate come finalizzate esclusivamente alla tutela dell'ordine pubblico<sup>91</sup>, la giurisprudenza ha incontrato difficoltà nel riconoscere i presupposti per una loro concreta applicazione<sup>92</sup>. Tuttavia, il quadro interpretativo è mutato con

<sup>87</sup> Sui delitti contro l'uguaglianza, nella manualistica, cfr.: F. Bacco, *Norme antidiscriminatorie*, in D. Pulitanò, (a cura di), *Diritto penale. Parte speciale, I, Tutela penale della persona*, Torino, 2019, 403 ss.; a livello monografico, L. Goisis, *Crimini d'odio*, cit., 263 ss.

<sup>88</sup> Si segnala che le disposizioni in oggetto hanno abrogato l'art. 3 della l. 654 del 1975 e l'art. 3 del d.l. 122 del 1993, convertito nella c.d. legge Mancino, pur mantenendone inalterato il testo.

<sup>89</sup> Cfr. M. Pelissero, *Discriminazione, razzismo e diritto penale fragile*, in *Diritto penale e procedura*, 2020, 1017 ss.

<sup>90</sup> Cfr. G. de Vero, *La giurisprudenza della Corte di Strasburgo*, in *Delitti e pene nella giurisprudenza delle Corti europee*, (a cura di) G. de vero, Torino, 2007, 46 ss.

<sup>91</sup> La sentenza spartiacque, da questo punto di vista, è rappresentata dalla n. 341/2001, secondo la quale: «il diritto alla libera manifestazione del pensiero non può dilatarsi sino a comprendere la diffusione di idee ed altre condotte che neghino la personalità e la dignità dell'uomo, valori questi che sono affermati dalla Costituzione come principi fondamentali e non tollerano alcuna forma di gerarchia fondata sull'appartenenza ad un gruppo etnico, nazionalità e razza»; «il diritto alla libera manifestazione del pensiero, tutelato dall'art. 21 della Costituzione, non può essere esteso fino alla giustificazione di atti o comportamenti che, pur estrinsecandosi in una esternazione delle proprie convinzioni, ledano tuttavia altri principi di rilevanza costituzionale ed i valori tutelati dall'ordinamento giuridico interno ed internazionale». In altre parole, per la Corte di cassazione l'esternazione delle convinzioni equivale ad atti e comportamenti che ledono la pari dignità sociale, a prescindere dagli effetti esplicati in concreto (cfr.: i motivi della decisione).

<sup>92</sup> Per una disamina dei beni giuridici in gioco, si veda, P. Tanzarella, *Discriminare parlando. Il pluralismo democratico messo alla prova dai discorsi d'odio razziale*, Torino, 2020, 12; G. Puglisi, *La parola acuminata*.

l'affermarsi della dottrina maggioritaria, che ha visto in esse una tutela plurioffensiva, adottando una lettura personalista e riconoscendo, altresì, la protezione dei beni giuridici dell'uguaglianza e della pari dignità dei cittadini<sup>93</sup>.

L'art. 604-bis c.p.<sup>94</sup> comma 1, lettera a), limita la sanzionabilità alla propaganda, all'istigazione a commettere e alla commissione di atti discriminatori. Ai fini della configurazione della condotta di propaganda, secondo il prevalente orientamento dottrinale, non basterebbe una generica diffusione di idee, rendendosi piuttosto necessario che la dichiarazione esternata sia idonea a raccogliere un considerevole numero di consensi in ordine al pensiero manifestato<sup>95</sup>.

Le condotte di istigazione a commettere e di commissione di atti discriminatori non richiederebbero, invece, la pubblicità del fatto. L'istigazione può verificarsi nei confronti di una sola persona, in deroga alla disposizione generale di cui all'art. 115 c.p.; d'altro canto, gli atti discriminatori possono essere realizzati anche in privato. Inoltre, non è neppure richiesto che l'atto discriminatorio configuri un illecito penale, essendo sufficiente il suo carattere discriminatorio per motivi razziali o etnici (come per la condotta di propaganda) o nazionali o religiosi. Da questo punto di vista le condotte di propaganda e di istigazione devono necessariamente incidere sull'altrui volontà, ovvero devono essere concretamente idonee a «provocare l'immediata esecuzione di delitti o, quantomeno, la probabilità che essi vengano commessi in un futuro più o meno prossimo»<sup>96</sup>. Il momento consumativo si individua nell'istante in cui la propaganda e l'istigazione vengono esternate, diventando percepibili e conoscibili, senza che sia necessario che siano anche accolte dai destinatari né che si verifichi un ulteriore evento materiale ed esterno.

Ai fini della configurazione è fondamentale accertare in concreto la pericolosità delle stesse, tenendo in considerazione il contesto in cui si è svolto il fatto, il ruolo dell'agente e le attitudini dei destinatari<sup>97</sup>.

Il richiamo alla concreta idoneità della condotta, combinato con l'anticipazione san-

---

*Contributo allo studio dei delitti contro l'uguaglianza, tra aporie strutturali ed alternative alla pena detentiva*, in Rivista Italiana di Diritto e Procedura penale, 2018, 1331 ss.

<sup>93</sup> La dignità è stata ben definita come «insieme delle condizioni necessarie a uno sviluppo della persona che le consenta di vivere un'esistenza piena», in M. Caputo, *La menzogna di Auschwitz, le verità del diritto penale. La criminalizzazione del c.d. negazionismo tra ordine pubblico, dignità e senso di umanità*, in AA.VV. *Verità del precetto e della sanzione penale alla prova del processo*, a cura di G. Forti-G. Varraso-M. Caputo, Napoli, 2014, 309; sul tema si veda anche G. Pavich-A. Bonomi, *Reati in tema di discriminazione: il punto sull'evoluzione normativa recente, sui principi e valori in gioco, sulle prospettive legislative e sulla possibilità di interpretare in senso conforme a Costituzione la norma vigente*, in *Penale Contemporaneo*, 13.10.2014, 12 ss.; L. Scaffardi, *Oltre i confini della libertà di espressione. L'istigazione all'odio razziale*, Padova, 2009, 239.

<sup>94</sup> Art. 604-bis c.p., c. 1: «Salvo che il fatto costituisca più grave reato, è punito: a) con la reclusione fino ad un anno e sei mesi o con la multa fino a 6.000 euro chi propaga idee fondate sulla superiorità o sull'odio razziale o etnico, ovvero istiga a commettere o commette atti di discriminazione per motivi razziali, etnici, nazionali o religiosi».

<sup>95</sup> F. Bacco, *Norme antidiscriminatorie* cit., 406.

<sup>96</sup> Cass. pen., sez. I, 7 ottobre 2009, n. 40552, in *Dejure*.

<sup>97</sup> In tal senso, giurisprudenza consolidata afferma che l'odio razziale o etnico è integrato da un sentimento idoneo a determinare il concreto pericolo di comportamenti discriminatori, e non da qualsiasi sentimento di generica antipatia, insofferenza o rifiuto riconducibile a motivazioni attinenti alla razza, alla nazionalità o alla religione, cfr: Cass. pen., sez. V, 22 luglio 2019, n. 32862.

zionatoria caratteristica delle fattispecie analizzate, le quali non richiedono l'effettiva attuazione dell'istigazione o della propaganda, consente di inquadrare il reato tra quelli a pericolo concreto<sup>98</sup>. Tale orientamento, del resto, consente di evitare il rischio di un possibile conflitto con il principio di offensività, il quale potrebbe essere compromesso da una sanzione penale nei confronti di un soggetto autore di condotta propagandistica concretamente inidonea, per i mezzi, le modalità e i destinatari, a ledere il bene giuridico tutelato. In un campo, come quello che ci occupa, in cui il bilanciamento dei beni giuridici in gioco è così delicato, inquadrare il reato nel novero di quelli a pericolo concreto permette di garantire il rispetto della Costituzione<sup>99</sup>. È dunque necessario svolgere un accertamento in concreto sulla pericolosità della propaganda e dell'istigazione, tenendo conto del contesto, del ruolo dell'agente e delle attitudini dei destinatari<sup>100</sup>.

La ricostruzione in termini di pericolo concreto è stata accolta anche nella risalente (seppur estremamente significativa) pronuncia della Corte costituzionale in materia di reati di opinione<sup>101</sup> oltretutto, in modo analogo, anche dalla Corte EDU nelle sue numerose sentenze in materia di negazionismo<sup>102</sup>.

Con riferimento alla previsione di cui alla lett. b)<sup>103</sup> dell'art. 604-bis c.p., comma 1<sup>104</sup>, essa concerne l'istigazione a commettere o la commissione di atti di violenza o di pro-

<sup>98</sup> In tal senso A. Cadoppi-S. Canestrari-A. Manna-M. Papa, *Diritto penale. Tomo terzo. I delitti contro l'onore e la libertà individuale*, Vicenza, 2022, 6105 ss.; L. Goisis, *Crimini d'odio*, cit., 283, dove parla di «reati di pericolo concreto (implicito)»; contra A. Galluccio, *Punire la parola pericolosa?*, Milano, 2020, 416.

<sup>99</sup> R. Bartoli-M. Pelissero-S. Seminara, *Diritto penale*, cit., 26.

<sup>100</sup> Cfr.: Cass. pen., sez. V, 29 gennaio 2020, n. 3722, in *Dejure*; Cass. pen., sez. I, 16 gennaio 2020, n. 1602, in *Dejure*; Cass. pen., sez. V, 30 luglio 2019, n. 34815, in *Dejure*; Cass. pen., sez. V, 22 luglio 2019, n. 32862, in *Dejure*.

<sup>101</sup> Corte cost., 4 maggio, 1970, nella quale si afferma che l'apologia punibile ai sensi dell'art. 414, ultimo comma c.p., non sarebbe la manifestazione di pensiero pura e semplice, bensì quella che per le sue modalità integra un comportamento concretamente idoneo a provocare la commissione di delitti. Invero, la libertà di manifestazione del pensiero, garantita dall'art. 21, primo comma della Costituzione troverebbe i suoi limiti non soltanto nella tutela del buon costume, ma anche nella necessità di proteggere altri beni di rilievo costituzionale e nell'esigenza di prevenire e far cessare turbamenti della sicurezza pubblica, la cui tutela costituisce una finalità immanente del sistema.

<sup>102</sup> Si segnalano le sentenze della Corte EDU nelle quali si è fatta applicazione di questi due criteri: *Perincekin* (2015), un caso di negazionismo del genocidio armeno, giudicato non concretamente pericoloso. La Corte ha, in tale occasione, ribadito che le limitazioni della libertà di espressione sono giustificate se le dichiarazioni di odio si innestano in un clima di tensione politica e sociale e se le dichiarazioni incitano in modo diretto o indiretto alla violenza o giustificano la violenza, l'odio o l'intolleranza. Si segnala altresì, con esiti opposti, *Diendonné* (2015), dove uno spettacolo pubblico, camuffato da manifestazione artistica, ha integrato una presa di posizione d'odio antisemita, capace di integrare l'abuso del diritto di manifestazione del pensiero ex art. 17 CEDU. Per un ulteriore approfondimento di tali pronunce si veda anche G. Forti, *Le tinte forti del dissenso nel tempo dell'ipercomunicazione pulviscolare. Quale compito per il diritto penale?* in *Rivista Italiana di Diritto e Procedura Penale*, 4, 2018.

<sup>103</sup> Art. 604 bis, comma 1, lett. b): «con la reclusione da sei mesi a quattro anni chi, in qualsiasi modo, istiga a commettere o commette violenza o atti di provocazione alla violenza per motivi razziali, etnici, nazionali o religiosi».

<sup>104</sup> Per ordine di completezza si segnala che il comma 2 dell'articolo in esame vieta le organizzazioni, le associazioni, i movimenti o i gruppi aventi tra i propri scopi l'incitamento alla discriminazione o alla violenza per motivi razziali, etnici, nazionali o religiosi. In seno a questa condotta associativa, parte della dottrina ravvisa un residuo di tutela del bene giuridico dell'ordine pubblico (L. Goisis, *Crimini d'odio*, cit. 285).

vocazione alla violenza per scopi discriminatori. La violenza può essere esercitata sia direttamente sulle persone, che anche sulle cose dalle stesse possedute.

I motivi razziali, etnici, nazionali o religiosi devono emergere in maniera univoca dalle idee, mentre risulta essere privo di rilievo l'eventuale, diverso, intento del soggetto attivo. A differenza di quanto previsto per le condotte di cui alla lett. a), il riferimento agli atti violenti sembra integrare, di per sé, un comportamento intrinsecamente illecito.

Una questione particolarmente dibattuta tra dottrina e giurisprudenza sembra riguardare l'elemento soggettivo. La giurisprudenza prevalente ritiene che la propaganda di idee discriminatorie e l'istigazione alla commissione di atti discriminatori (prima parte, lett. a) integrerebbero un'ipotesi di dolo generico; diversamente, la commissione di atti di discriminazione (seconda parte, lett. a) e la commissione di violenza o di atti di provocazione alla violenza sarebbero sorrette dal dolo specifico, in quanto «in tali ultime ipotesi il motivo ispiratore eccede la condotta discriminatoria o violenta, mentre nel caso della propaganda e dell'istigazione tale motivo è incluso nelle idee propagandate o negli atti discriminatori istigati»<sup>105</sup>. La dottrina, d'altro canto, ritiene che entrambe le condotte (lettere a e b) debbano essere punite a titolo di dolo generico<sup>106</sup>.

Un ultimo aspetto che merita di essere esaminato è quello concernente il trattamento sanzionatorio<sup>107</sup>.

Rispetto alle condotte disciplinate alla lettera a) dell'art. 604-bis c.p., appare incongruo omologare condotte dotate di un differente disvalore, ponendo sul medesimo piano la propaganda, l'istigazione e la materiale realizzazione di atti discriminatori (reclusione fino a un anno e sei mesi o con la multa). Peraltro, l'alternatività della pena indebolisce l'efficacia del precetto, ammettendo una rilevanza, per così dire, "bagatellare", di fatti pur offensivi di beni di rango così elevato. Con riferimento alla graduazione della sanzione prevista alle lettere a) e b) dell'art. 604-bis c.p., essa si considera proporzionata: la propaganda e l'istigazione (lettera a) sono punite meno gravemente (fino a un anno e sei mesi di reclusione o con la multa fino a 6.000 euro) rispetto all'istigazione e agli atti violenti (reclusione da sei mesi a quattro anni) disciplinati alla lettera b).

Un discorso diverso deve essere fatto per la fattispecie di cui all'art. 604-ter c.p., ossia l'aggravante di natura teleologica<sup>108</sup>. Nello specifico, essa prevede un aumento di pena sino alla metà per i reati puniti con la pena diversa da quella dell'ergastolo, quando sono commessi per finalità di discriminazione o di odio etnico, nazionale, razziale o religioso, ovvero al fine di agevolare organizzazioni aventi i medesimi scopi.

<sup>105</sup> Cass. pen., sez. III, 7 maggio 2008, n. 37581, in *Dejure*.

<sup>106</sup> Cfr.: G. Fiandaca – E. Musco, *Diritto Penale. Parte speciale*, vol. II, tomo I, Bologna, 2024, 243; L. Picotti, cit., 145; G. Forti-S. Riondato-S. Seminara, *Commentario breve al Codice Penale*, VII ed., Milano, 2024, 2381.

<sup>107</sup> La delicatezza della materia oggetto d'esame e la tensione che questa può comportare con il diritto alla libera manifestazione del pensiero impongono una puntuale attenzione al momento sanzionatorio. Tale necessità è stata peraltro avvertita dalla Corte europea dei diritti dell'uomo, la quale ha avuto sin da subito la necessità di specificare che ogni sanzione imposta in questo ambito deve essere proporzionata allo scopo legittimo perseguito (cfr.: Corte EDU, *Handyside c. UK*, par. 49,18 del 7 dicembre 1976).

<sup>108</sup> L'aspetto peculiare risiede nella sottrazione di siffatta circostanza aggravante al sistema di bilanciamento di circostanze di cui all'art. 69 c.p., eccetto che nel caso di attenuante per la minore età di cui all'art. 98 c.p. Questa impostazione è eloquente del notevole disvalore che connota tali condotte, tanto da imporre un trattamento punitivo differenziato e meno favorevole per il reo.

La finalità di discriminazione o di odio non deve essere rigidamente intesa come movente psicologico: essa sussiste sia quando l'azione concreta è finalizzata a rendere percepibile all'esterno la sua causale, sia quando evidenzia un pregiudizio di tipo etnico, nazionale, razziale o religioso. Non sembra inoltre essere necessario uno specifico scopo di incitamento<sup>109</sup>.

Il fondamento di tale disposizione non dà luogo a problemi interpretativi quando il reato base presenta una specifica e autonoma direzione lesiva, alla quale si aggiungono le finalità tipizzate dalla norma.

Problemi di diversa natura sorgono con particolare riferimento al reato di diffamazione, in quanto il confine tra gli artt. 604-bis e 595 c.p. nella sua forma aggravata è assai labile: le condotte di propaganda e istigazione recano solitamente in sé una nota diffamatoria nei confronti dei soggetti appartenenti alle razze, etnie, nazioni o religioni. Da diversi anni pendono in Parlamento proposte di legge tese a estendere le incriminazioni contenute negli artt. 604-bis e ter c.p. alle discriminazioni fondate su motivi sessuali. Il contrasto vede, da un lato, coloro i quali si oppongono al sacrificio della libertà di manifestazione del pensiero; da altro lato, quelli che ravvisano nella negazione di un riconoscimento all'individuo diverso, un'offesa intollerabile di un diritto inviolabile proiettato sull'eguaglianza. Rispetto a questa contrapposizione, è difficile discostarsi dal contenuto della pronuncia della Corte EDU *Gunduz v. Turchia*, con la quale si è affermato che: «la tolleranza e il rispetto per l'eguale dignità di tutti gli esseri umani costituiscono il fondamento di una società democratica e pluralistica. Ne deriva che nelle società democratiche può ritenersi necessario sanzionare, e così prevenire, ogni forma di espressione che diffonda, istighi, promuova o giustifichi l'odio basato sull'intolleranza»<sup>110</sup>. Si segnala che l'Italia risulta essere uno dei pochi paesi in Europa che non dispone di specifiche incriminazioni dei reati d'odio in ambiti legati alla sfera sessuale, in palese contrasto con gli atti sovranazionali<sup>111</sup>.

<sup>109</sup> Cass. pen., sez. V, 22 luglio 2019, n. 32862, in *Dejure*, per quanto concerne i riferimenti alla giurisprudenza della Corte EDU. Si esclude, ai fini della configurazione, l'esigenza che l'azione risulti intenzionalmente diretta a dare luogo, in futuro o nell'immediato, al concreto pericolo di comportamento discriminatorio.

<sup>110</sup> Corte EDU, sez. I, 4 dicembre 2003, n. 35071/97.

<sup>111</sup> Si segnala, per completezza, che il disegno di legge noto come d.d.l Zan, recante «Misure di prevenzione e contrasto della discriminazione e della violenza per motivi fondati sul sesso, sul genere, sull'orientamento sessuale, sull'identità di genere e sulla disabilità», naufragato in Senato, perseguiva l'obiettivo di modificare l'attuale assetto dei motivi discriminatori delle norme esaminate aggiungendo dopo le parole «per motivi razziali, etnici, nazionali o religiosi» l'inciso «oppure fondati sul sesso, sul genere, sull'orientamento sessuale, sull'identità di genere o sulla disabilità». Tale estensione avrebbe dovuto essere riferita esclusivamente alle condotte di istigazione al compimento o al compimento di atti discriminatori (art. 604-bis c.p. comma 1 lett. b), nonché alla condotta associativa di cui al secondo comma dell'art. 604-bis c.p. ed alla circostanza aggravante di cui all'art. 604-ter c.p.

Con riferimento agli atti sovranazionali, si segnalano: il Trattato sul funzionamento dell'Unione Europea che stabilisce che l'Unione mira in tutte le sue attività a combattere la discriminazione sulla base di diversi fattori, incluso l'orientamento sessuale; l'art. 21 della Carta dei Diritti Fondamentali; la direttiva 2000/78/CE, che stabilisce un quadro generale per la parità di trattamento in materia di occupazione e di condizioni di lavoro, ribadendo il diritto universale di tutti all'uguaglianza davanti alla legge e alla protezione contro le discriminazioni quale diritto universale; la Raccomandazione Rec CM2010(05) del Comitato dei Ministri del Consiglio d'Europa agli Stati membri sulle misure dirette a combattere la discriminazione fondata sull'orientamento sessuale o l'identità di genere (adottata il 31 marzo 2010); le Linee guida per la promozione e la tutela dell'esercizio di tutti i diritti umani da parte di lesbiche, gay,

Le problematiche relative ai discorsi d'odio online vengono affrontate principalmente nelle aule di giustizia, a causa della difficoltà, del legislatore italiano e di quello europeo, di rispondere in maniera adeguata e repentina alle sfide poste dall'incessante evoluzione tecnologica<sup>112</sup>. Di conseguenza, la risposta di un singolo tribunale a un caso specifico tende a configurarsi come un riferimento generale per l'intera materia, contribuendo alla sua definizione normativa. Secondo i dati diffusi dall'allora ministro della Giustizia Marta Cartabia nel 2022<sup>113</sup>, l'introduzione di queste fattispecie penali non avrebbe assolto all'auspicata funzione di deterrenza.

Nonostante un esponenziale aumento dei discorsi d'odio<sup>114</sup>, si è riscontrato un numero esiguo di procedimenti giudiziari. Tra il 2016 e il primo semestre del 2021, i procedimenti iscritti (prevalentemente nel nord Italia, con le percentuali maggiori nelle grandi città di Roma con il 12,62%, e Milano con il 4,85%) non hanno superato le 300 unità, sia con riferimento alla forma di manifestazione della propaganda e dell'istigazione, quanto alla configurazione dell'aggravante, la cui applicazione risulta più diffusa, a discapito del delitto di cui all'art. 604-bis c.p.<sup>115</sup>.

Rispetto ai flussi dei procedimenti definiti dalle sezioni G.i.p./ G.u.p. e dibattimentali, si è potuto osservare che nell'80 % dei casi l'iscrizione delle ipotesi di reato viene poi archiviata; nei casi di rinvio a giudizio, relativamente scarsi, prevale la condanna (circa il 40%). Solo la metà dei procedimenti penali riguardanti ipotesi delittuose aggravate ai sensi dell'art. 604-ter c.p. si conclude con una condanna aggravata.

Si segnala, inoltre, che in alcuni dei casi in cui il rinvio a giudizio avviene ai sensi dell'art. 595 c.p., aggravato ex art. 604-ter c.p., la configurazione del reato viene successivamente esclusa a causa dell'impossibilità di individuare la persona destinataria dell'offesa<sup>116</sup>,

---

bisessuali, transgender e intersessuali, adottati dal Consiglio europeo il 24 giugno 2013; la Risoluzione del 14 febbraio 2019 adottata dal Parlamento europeo, con il quale si invita la Commissione a garantire priorità ai diritti delle persone LGBT+; la Risoluzione del 1 marzo 2021 del Parlamento europeo.

<sup>112</sup> Si ritiene opportuno segnalare che negli anni Novanta le sentenze penali aventi ad oggetto il reato di *hate speech* sono state piuttosto rare. Nello specifico, la lett. a), art. 3 legge Reale, così come modificato dalla c.d. legge Mancino, non ha mai trovato applicazione ai fini della valutazione del reato di diffusione di idee fondate sulla superiorità o sull'odio razziale o etnico. In alcune (rare) circostanze ha trovato invece applicazione il reato di incitamento a commettere atti di discriminazione, sia con riferimento alla violazione dell'art. 3 lett. a), sia in relazione all'aggravante ex art. 3 della legge Mancino. Interessante, al riguardo, è osservare come inizialmente i Giudici abbiano inteso interpretare tali fattispecie delittuose alla luce del criterio del pericolo concreto, sulla scorta di un precedente che aveva dimostrato riluttanza a offrire una tutela anticipata per la mera diffusione di frasi solo ipoteticamente foriere di atti violenti (cfr.: Cass. pen., sez. III, 24 novembre 1998, n. 434).

<sup>113</sup> Dati forniti dall'allora Ministra Marta Cartabia, in seno alla Commissione straordinaria per il contrasto dei fenomeni di intolleranza, razzismo, antisemitismo e istigazione all'odio e alla violenza, presieduta dalla Sen. Segre, l'8 febbraio 2022.

<sup>114</sup> Il Barometro dell'odio 2023-2024 di Amnesty International, rivela un aumento significativo dei contenuti problematici o di *hate speech*: dal 2019 a luglio 2024 il tasso di discorsi offensivi, discriminatori o che incitano all'odio è passato dal 10 per cento al 15 per cento. In particolare, i contenuti che incitano alla discriminazione e alla violenza sono triplicati, superando il tre per cento del corpus analizzato. Nell'ambito della ricerca è stato inoltre condotto un sondaggio, in collaborazione con Ipsos, per rilevare quale fosse la posizione dell'opinione pubblica rispetto all'attivismo e alle varie forme di protesta. È emerso che il 48 per cento delle persone intervistate vede le manifestazioni come un passatempo o una moda, mentre il 17 per cento non crede che tutti, in Italia, dovrebbero avere il diritto di manifestare.

<sup>115</sup> Dati forniti dall'allora Ministra Marta Cartabia nel 2022.

<sup>116</sup> Cfr.: Cass. pen., sez. V, 22 luglio 2019, n. 32862, in *Dejure*.

con la conseguente disapplicazione della circostanza aggravante in esame. Un'ulteriore ipotesi problematica riguarda il caso di rinvio a giudizio per l'ipotesi di reato di cui all'art. 604-bis c.p., successivamente riqualificato in diffamazione aggravata. Il rischio, in questo caso, è rappresentato dal proscioglimento per difetto di querela<sup>117</sup>.

L'analisi svolta fino a questo momento evidenzia come lo strumento del diritto penale possa rivelarsi utile, in quanto idoneo a stigmatizzare comportamenti che ledono beni giuridici fondamentali per il funzionamento della democrazia. Tuttavia, stante le criticità emerse nel corso della trattazione, non sembra essere sufficiente a contrastare in modo efficace la diffusione dell'odio online.

#### **4. Esperienze giuridiche di incriminazione dell'online hate speech a confronto: Francia, Germania e Spagna**

Volendo ora rivolgere l'attenzione ad altri sistemi giuridici che hanno operato scelte di politica criminale analoghe a quelle effettuate dal legislatore italiano, merita di essere menzionato l'ordinamento francese, che si contraddistingue per la previsione di un complesso sistema di contrasto all'odio online articolato su tre livelli: i contenuti vietati, i contenitori - ossia i soggetti che ospitano tali contenuti - e i controllori esterni (magistratura, autorità amministrativa, autorità indipendente, ecc.). In Francia, Internet non è un'area esente dalla legge<sup>118</sup>.

Più nel dettaglio, l'odio online viene punito ricorrendo alla sussunzione sia sotto i reati previsti dal Codice penale, come l'apologia di terrorismo o di minaccia di morte, sia sotto i reati previsti dalla legge sulla libertà di stampa del 29 luglio 1881<sup>119</sup>, che sanziona gli attacchi alla presunzione di innocenza, alla dignità delle vittime, il negazionismo, l'insulto, la diffamazione e la provocazione, in modo graduale a seconda della loro portata più o meno pubblica o del loro carattere discriminatorio, nonché l'apologia di alcuni reati<sup>120</sup>.

La suddetta legge punisce l'espressione di qualsivoglia commento che possa assumere una dimensione pubblicistica, vale a dire "penetrare in un modo o nell'altro nel pubblico dominio", e di conseguenza si applica anche alle manifestazioni pubblicate su Internet<sup>121</sup>.

Tuttavia, è importante ricordare che la legge sulla stampa fu adottata per proteggere la

---

<sup>117</sup> Dati forniti dall'allora Ministra Marta Cartabia nel 2022 *Norme antidiscriminatorie*.

<sup>118</sup> L. Marguet, *La complexité du dispositif juridique de lutte contre la haine en ligne*, in *Les sciences sociales et les défis éthiques de la recherche*, 16, 2023, 1-5.

<sup>119</sup> *Loi du 29 juillet 1881 sur la liberté de la presse*.

<sup>120</sup> Vedasi riguardo a detti reati gli artt. 35 ter, 35 quater, 24 bis, 33, 29, 23 e 24 della l. sulla stampa del 1881, citati da L. Marguet, *La complexité du dispositif juridique de lutte contre la haine en ligne*, cit., 10, la quale rimanda, sul punto, per un maggiore approfondimento a L. Marguet, *La répression de la provocation, de la diffamation et des injures non publiques représentant un caractère raciste ou discriminatoire en France*, in *Revue des droits et libertés fondamentaux*, 20, 2019.

<sup>121</sup> Art. 23 della legge sulla libertà di stampa del 1881, modificata dalla l. 2004-575 del 21 giugno 2004 per la fiducia nell'economia digitale (c.d. "LCEN"), citata da L. Marguet, *La complexité du dispositif juridique de lutte contre la haine en ligne*, cit., 16.

libertà di espressione. Ne consegue che, quando un discorso d'odio rientra nel suo ambito di applicazione, all'autore vengono garantite una serie di tutele procedurali, quali un termine di prescrizione molto breve (di tre mesi), l'inammissibilità della custodia cautelare, della comparizione con rito immediato e del pronunciamento anticipato di colpevolezza. Inoltre, il giudice istruttore e il tribunale di primo grado sono vincolati alla qualificazione dei fatti come descritti negli atti che hanno dato origine all'azione penale<sup>122</sup>.

Questo quadro giuridico è attualmente in evoluzione. Alcune iniziative legislative mirano a escludere talune condotte dal campo di operatività della legge sulla stampa. Tra queste, la l. 1109 del 24 agosto 2021, la cosiddetta “legge di rafforzamento dei principi della Repubblica”, ha introdotto una serie di disposizioni relative alla lotta contro i discorsi d'odio e i contenuti illeciti nella sfera digitale<sup>123</sup>.

La principale novità di interesse è costituita dall'introduzione, ad opera dell'art. 36, di una nuova fattispecie criminosa all'art. 233-1-1 del *Code pénal*, rubricata “messa in pericolo della vita altrui tramite la diffusione di informazioni relative alla vita privata, familiare o professionale”<sup>124</sup>. Tale disposizione criminalizza il *doxing*<sup>125</sup>, ossia la rivelazione di informazioni relative alla vita privata, familiare o personale di una persona che consentano di identificarla o di localizzarla, allo scopo di esporla a un rischio diretto di danno, nella misura in cui il responsabile della rivelazione non avrebbe potuto ignorare il verificarsi di un pregiudizio alla persona medesima o a una cosa<sup>126</sup>.

La figura incriminatrice è stata concepita facendo ricorso alla forma di imputazione

<sup>122</sup> Artt. 50 e 51 della l. sulla libertà di stampa del 1881. In riferimento ai profili procedurali, si rinvia a R. Evan, *La procédure pénale en droit de la presse*, Gazette du Palais, 2019.

<sup>123</sup> *Loi n° 2021-1109 du 24 août 2021 confortant le respect des principes de la République*, entrata in vigore a seguito dell'attentato a Samuel Paty, decapitato all'uscita della scuola dove insegnava, per aver mostrato immagini di Maometto, intesa principalmente a contrastare i c.d. “separatismi religiosi”, in particolare quelli di matrice islamista, che secondo il legislatore francese – e alla luce degli attentati che la Francia ha dovuto sopportare nell'ultimo decennio – rappresentano una minaccia per i valori della *République*. È stata così introdotta una normativa restrittiva che, secondo alcuni commentatori e gli stessi vescovi francesi, pone forti limiti all'esercizio della libertà religiosa, così E. Tawil, *La recente legge francese sulla laicità dello Stato contro i separatismi religiosi*, intervento reso durante una lezione di Diritto pubblico presso l'Université Paris II Panthéon-Assas, lumsa.it, 21 novembre 2022.

<sup>124</sup> La formulazione dell'art. 233-3-1 del Codice penale francese è la seguente: «L'atto di rivelare, diffondere o trasmettere, con qualsiasi mezzo, informazioni relative alla vita privata, familiare o professionale di una persona che ne permettano l'identificazione o la localizzazione al fine di esporre tale persona o i membri della sua famiglia a un rischio diretto di danno alla loro persona o ai loro beni di cui l'autore non poteva non essere a conoscenza, è punito con tre anni di reclusione e una multa di 45.000 euro. Quando i fatti sono commessi a danno di un pubblico ufficiale, di una persona incaricata di un pubblico servizio o che ricopre una carica elettiva pubblica o di un giornalista, ai sensi c. 2 dell'art. 2 della l. del 29 luglio 1881 sulla libertà di stampa, le pene sono aumentate a cinque anni di reclusione e a 75.000 euro di multa. Quando i fatti sono commessi a danno di un minore, le pene sono aumentate a cinque anni di reclusione e 75.000 euro di multa. Quando i fatti sono commessi a danno di una persona la cui particolare vulnerabilità, dovuta all'età, alla malattia, all'infermità, alla disabilità fisica o mentale o alla gravidanza, è evidente o nota all'autore del reato, le pene sono aumentate a cinque anni di reclusione e 75.000 euro di multa. Nel caso in cui i reati siano commessi attraverso la stampa scritta o audiovisiva o attraverso la comunicazione pubblica online, si applicano le disposizioni specifiche delle leggi che regolano queste materie per quanto riguarda la determinazione della responsabilità».

<sup>125</sup> L. Marguet, *La complexité du dispositif juridique de lutte contre la haine en ligne*, cit., 17.

<sup>126</sup> P. Januel, *Séparatisme: les principales dispositions de la loi*, in dalloz-actualite.fr, 07 settembre 2021.

soggettiva della responsabilità introdotta dal Codice penale del 1994 all'art. 121, cc. 2 e 3, che si colloca in una posizione intermedia tra i concetti che, nella scienza penalistica italiana, assumono la qualifica di “dolo eventuale” e “colpa cosciente”. Ai fini della configurazione della *mise en danger*, è necessario che l'autore abbia coscienza del pericolo per l'interesse tutelato, presunto dalle modalità di aggressione<sup>127</sup>. È evidente che il legislatore francese ha attribuito rilevanza al solo stato di pericolo dell'oggettività giuridica tutelata oltre che alle caratteristiche con cui si esplica il comportamento, anticipando la soglia della rilevanza penale della condotta.

Questa soluzione si inserisce nel solco della giurisprudenza costituzionale che ha posto l'attenzione non soltanto sull'importanza del mantenimento dell'ordine pubblico e dei diritti dei destinatari delle manifestazioni d'odio, ma anche sui canoni di proporzionalità e necessità che devono guidare l'accertamento concreto dell'illiceità del fatto. A tale proposito, numerose pronunce del Consiglio costituzionale francese hanno affrontato il tema dell'impatto delle norme penali sulla libertà di espressione, definita come «una condizione della democrazia e una delle garanzie del rispetto di altri diritti e libertà»<sup>128</sup>. In linea generale, il Consiglio non ha sollevato obiezioni nei confronti delle scelte del legislatore di incriminare condotte di abuso della libertà di cui si tratta, a condizione che le stesse, in ossequio all'art. 34 Cost., compromettano «l'ordine pubblico e i diritti di terzi»<sup>129</sup> e superino positivamente un triplice vaglio, che ruota intorno ai canoni di necessità, adeguatezza e proporzionalità<sup>130</sup>. Tale valutazione viene operata mediante un approccio di tipo casistico e avendo particolare riguardo alla certezza dell'illiceità del messaggio o, al contrario, all'incertezza di tale illiceità. Ne deriva che il rischio che la violazione possa essere qualificata come proporzionata è tanto maggiore quanto più la qualificazione giuridica dei comportamenti è suscettibile di dare luogo a interpretazioni differenti o contrastanti<sup>131</sup>.

Le successive disposizioni della l. 1109 del 24 agosto 2021 (sino all'art. 48), attraverso la combinazione di nuove previsioni normative e l'estensione ai gestori di siti web delle responsabilità già sancite dalle leggi sulla stampa, tentano di disciplinare il tema spinoso della diffusione dei messaggi d'odio, d'incitamento alla violenza o d'induzione alla commissione di reati, in tutti gli aspetti, talora sfuggenti, che concorrono a delineare il quadro sociologico dell'odierno radicalismo politico di ispirazione islamica affermatosi in Francia<sup>132</sup>.

In quest'ottica, l'ordinamento francese punisce anche le molestie informatiche: l'art. 222-33-2-2 del Codice penale precisa che il reato di molestie si configura anche quando

<sup>127</sup> G. Fornasari - A. Menghini, *Percorsi europei di diritto penale*, Cedam, 2005, 84.

<sup>128</sup> Consiglio cost. francese, 10 giugno 2009, n. 580 DC.

<sup>129</sup> Art. 34 Cost. francese. Per un commento sul punto, vedasi Consiglio cost. francese, d28 febbraio 2012, n. 2012-647 DC, par. 5, citata in *Commento alla decisione 2020-801 del 18 giugno 2020*, cit., 18.

<sup>130</sup> Consiglio cost. francese, 20 maggio 2011, n. 2011-131 QPC, Consiglio cost. francese, 15 dicembre 2017, n. 2017-682 QPC, par. 4; Consiglio cost. francese, 8 gennaio 2016, n. 2015-512 QPC par. 5-8, citate in *Commento alla decisione 2020-801 del 18 giugno 2020*, cit.

<sup>131</sup> *Ibid.*, come risulta dal *Commento alla decisione 2020-801 del 18 giugno 2020*, cit., 10.

<sup>132</sup> Così, A. Tira, *La legge francese n. 1109 del 24 agosto 2021 sul “rafforzamento del rispetto dei principi della Repubblica”*, in *Stato, Chiese e Pluralismo confessionale*, 14, 2021, 25. Si veda *ex multis*, per questo delicato profilo, il saggio di F. Benslama, *Un furieux désir de sacrifice. Le surmusulman*, in *Media Diffusion*, 2016.

gli atti sono stati commessi attraverso un servizio di comunicazione online (*scapel*<sup>133</sup>). Dal 2018 il reato si configura anche «quando tali parole o comportamenti vengono imposti alla stessa vittima da più persone, in modo concertato o su istigazione di una di esse, anche se ciascuna di queste persone ha agito ripetutamente»<sup>134</sup>. Tale precisazione mira a sanzionare le c.d. “molestie da branco”.

La densità del sistema disegnato dal legislatore francese, pone, tuttavia, numerosi interrogativi: la quantità di norme applicabili non è sempre garanzia di qualità o di efficacia, bensì, al contrario, può essere sinonimo di complessità, che ostacola l'intelligibilità del sistema<sup>135</sup>.

Analogamente a quanto avvenuto in Francia, anche l'ordinamento giuridico tedesco ha introdotto una legislazione finalizzata a contrastare, in maniera quanto più complessiva e sistemica, l'incitamento all'odio online e la commissione di reati sui social network, con la *Network Enforcement Law* (c.d. NetzDG)<sup>136</sup> (c.d. NetzDG), entrata in vigore il 1° ottobre 2017 e successivamente riformata dalla l. del 3 giugno 2021.

È bene premettere che la Costituzione tedesca<sup>137</sup>, all'art. 5, par. 1, tutela il diritto individuale di «esprimere e diffondere [...] opinioni con parole, scritti e immagini». Il concetto di opinione deve essere inteso in senso ampio e non può essere limitato, a meno che l'espressione non si configuri come una “critica abusiva”, concetto che segna la sottile linea di demarcazione tra l'opinione critica (che merita protezione) e la condotta illecita<sup>138</sup>.

Il libero esercizio del diritto di opinione non è soggetto a riserve di natura contenutistica: il suo ambito di applicazione è esteso a qualunque contenuto<sup>139</sup> indipendentemente dalla sua forma di manifestazione<sup>140</sup>.

A tutt'oggi, l'ordinamento tedesco non stabilisce una definizione di discorso d'odio<sup>141</sup>: neppure la *NetzDG* ha introdotto una definizione giuridica univoca. Ciononostante, il § 1, c. 3, della *NetzDG* fornisce un'importante direttiva ermeneutica per attribuire rilevanza penale a queste condotte: il legislatore ha operato un rinvio al reato di cui al § 130 del Codice penale tedesco (*Strafgesetzbuch*, in forma abbreviata *StGB*), il cui c. 1 punisce l'“istigazione all'odio” nei confronti di un gruppo nazionale, razziale, religioso

<sup>133</sup> Cass. francese, 10 marzo 2020, n. 19-81026; Cass. francese, 17 settembre 2019, n. 18-834.

<sup>134</sup> L. 703 del 3 agosto 2018 sul rafforzamento della lotta alla violenza sessuale e di genere.

<sup>135</sup> L. Marguet, *La complexité du dispositif juridique de lutte contre la haine en ligne*, cit., 5.

<sup>136</sup> L. del 03 giugno 2021 *Gesetz zur Verbesserung der Rechtsdurchsetzung in sozialen Netzwerken* (Legge per migliorare l'applicazione della legge nei social network) che prevede alcuni obblighi per i social network, tra i quali la rimozione entro 24 ore dei contenuti illegali diffusi, come quelli che incitano all'odio o alla diffamazione. Per un maggiore approfondimento, si rinvia a O. Pollicino – G. De Gregorio, *Hate speech: una prospettiva di diritto comparato*, in *Giornale di diritto amministrativo*, 4, 01 luglio 2019, 433.

<sup>137</sup> Più propriamente la Legge fondamentale della Repubblica federale di Germania, c.d. *Grundgesetz*.

<sup>138</sup> V. Claussen, *Fighting hate speech and fake news. The Network Enforcement Act (NetzDG) in Germany in the context of European legislation*, in *Media Laws*, 3, 3, 2018, 121.

<sup>139</sup> Per un maggior approfondimento vedasi M. Lothar, *Die wehrhafte Demokratie als verfassungsimmanente Schranke der Meinungsfreiheit*, in *Zeitschrift für das Juristische Studium*, 2, 2010, 155.

<sup>140</sup> A. Espinoza – J. Rivas Alberty, *Limitaciones Inmanentes de la libertad de expresión. El discurso de odio en Alemania, España y el TEDH*, in *Revista de Derecho Público Contemporáneo*, 2, 2020, 224-225.

<sup>141</sup> J. Rinceanu, *Verso una forma di polizia privata nello spazio digitale? L'inedito ruolo dei provider nella disciplina tedesca dei social network*, in *Sistema penale, Studi in onore di Antonio Fiorella: Volume 1. Tre press*, 2021, 4.

o di un gruppo determinato dall'origine etnica, contro parti della popolazione o contro un individuo, a causa della sua appartenenza a uno dei suddetti gruppi o a una parte della popolazione<sup>142</sup>. Dunque, il discorso d'odio diventa illecito e non costituzionalmente tutelato, ogni qualvolta esso raggiunga l'intensità dell'istigazione<sup>143</sup>. La soglia della punibilità individuata è paragonabile a quella prevista dall'art. 604-bis del Codice penale italiano<sup>144</sup>.

Ferma restando l'applicazione della disposizione sopra esaminata, anche altre fattispecie vengono attualmente impiegate per perseguire i discorsi d'odio. È questo il caso dell'ingiuria (§ 185 *StGB*), della diffamazione (§ 186 *StGB*), della menzogna diffamatoria (§ 187 *StGB*), della diffamazione e menzogna diffamatoria contro personaggi della vita politica (§ 188 *StGB*) e dell'incitamento pubblico a commettere reati (§ 111 *StGB*)<sup>145</sup>.

Al fine di garantire il principio della certezza del diritto e della legalità in materia penale, il legislatore tedesco ha tassativamente elencato i contenuti da ritenersi illeciti al § 1, par. 3, della NetzDG, positivizzando un vero e proprio catalogo dei discorsi d'odio penalmente rilevanti<sup>146</sup>: § 86 *StGB* (Diffusione di mezzi di propaganda di organizzazioni incostituzionali), § 86a *StGB* (Uso di contrassegni di organizzazioni incostituzionali), § 89 *StGB* (Preparazione di un grave atto di violenza che mette in pericolo lo Stato), § 91 *StGB* (Istruzioni per commettere un grave atto di violenza che mette in pericolo lo Stato), § 100a *StGB* (Falsità a scopo di tradimento della Patria), § 111 *StGB* (Pubblico incitamento a commettere reati), § 126 *StGB* (Turbativa della pace pubblica mediante la minaccia di reati), §§ 129-129b *StGB* (Formazione di associazioni criminali, Formazione di associazioni terroristiche, Associazioni criminali e terroristiche all'estero), § 130 *StGB* (Istigazione all'odio), § 131 *StGB* (Apologia della violenza), § 140 *StGB* (Remunerazione ed apologia di reati), § 166 *StGB* (Vilipendio di culti, comunità religiose e associazioni ideologiche), § 184b *StGB* i.V.m. § 184d *StGB* (Distribuzione, acquisizione e possesso di scritti pedopornografici in relazione alla messa a disposizione di contenuti pornografici per mezzo di trasmissioni radiotelevisive o telematiche), §§ 185-187 *StGB* (Ingiuria, Diffamazione, Menzogna diffamatoria), § 201a *StGB* (Violazione della sfera intima e dei diritti della personalità attraverso ripresa di immagini), § 241 *StGB* (Minaccia) o § 269 *StGB* (Falsificazione di dati probatori)<sup>147</sup>.

Quanto al formante giurisprudenziale, la Corte costituzionale federale tedesca (*Bundesverfassungsgericht*, per brevità BVerfG) ha precisato che in tutte le ipotesi riconducibili al § 130 si ha una violazione della dignità umana, che costituisce il nucleo essenziale di ciascun diritto fondamentale e non può essere oggetto di bilanciamento con nessuno di essi. Pertanto, ogni qualvolta la dignità umana venga lesa, non potranno essere presi

<sup>142</sup> § 130, c. 1, *StGB*.

<sup>143</sup> J. Rinceanu, *Verso una forma di polizia privata nello spazio digitale?*, cit., 4.

<sup>144</sup> Rubricata "Propaganda e istigazione a delinquere per motivi di discriminazione razziale etnica o religiosa" per il cui approfondimento si rimanda al par. 3 del presente contributo.

<sup>145</sup> J. Rinceanu, *Verso una forma di polizia privata nello spazio digitale?*, cit., 5.

<sup>146</sup> V. Claussen, *Fighting hate speech and fake news*, cit., 118.

<sup>147</sup> O. Pollicino – G. De Gregorio, *Hate speech: una prospettiva di diritto comparato*, cit., nota 59.

in considerazione gli interessi della libertà di espressione<sup>148</sup>. Poiché, secondo la Corte costituzionale federale tedesca, tutti i diritti umani costituiscono una concretizzazione del principio di dignità, è necessaria una motivazione particolarmente accurata per affermare che l'esercizio di un diritto fondamentale, quale la libertà di espressione, possa giustificare una violazione della dignità umana, per sua natura stessa inviolabile<sup>149</sup>. Ne deriva che anche questo ordinamento giuridico concepisce la proporzionalità come strumento di controllo della conformità delle fattispecie incriminatrici in commento ai diritti fondamentali<sup>150</sup>. Più nel dettaglio, la giurisprudenza costituzionale tedesca richiede che esse perseguano «uno scopo legittimo e [siano] idonee, necessarie e, in senso stretto, proporzionate al raggiungimento di tale scopo»<sup>151</sup>.

Nel modello tedesco, il concetto di dignità umana vincola il diritto al rispetto della persona. Per questo motivo, tale nozione viene spesso concepita in termini di violazione, che può consistere nella denigrazione, nella stigmatizzazione, nella persecuzione ovvero in qualsiasi altra condotta che metta in dubbio la qualità della persona che la subisce<sup>152</sup>.

Infine, si precisa che, secondo la Corte Suprema federale di giustizia (*Bundesgerichtshof*, in forma abbreviata *BGH*), l'odio rilevante ai sensi del § 130 *StGB*, si sostanzia in un comportamento che colpisce i sentimenti, ovvero l'intelletto, di un'altra persona, oggettivamente idoneo oppure soggettivamente determinato a creare, ovvero rafforzare, un atteggiamento emotivamente ostile nei confronti di parte della popolazione, che vada oltre il semplice rifiuto o disprezzo<sup>153</sup>. È evidente che il carattere doppiamente lesivo dei crimini in questione<sup>154</sup> venga sottolineato anche dalla giurisprudenza tedesca. Volgendo, infine, lo sguardo all'ordinamento spagnolo, la l. organica n. 1 del 30 marzo 2015<sup>155</sup>, in linea con l'evoluzione del diritto internazionale e, in particolare, della giurisprudenza della Corte europea dei diritti dell'uomo, ha rafforzato le politiche criminali antidiscriminatorie<sup>156</sup>, introducendo l'art. 510 del Codice penale, rubricato “reato di incitamento all'odio e alla violenza”. Successivamente, il novero dei crimini d'odio è stato ulteriormente ampliato dalla riforma introdotta dalla l. organica n. 6 del 2022<sup>157</sup>.

<sup>148</sup> BVerfG, 06 settembre 2000, n. 1056, par. 43.

<sup>149</sup> A. Espinoza – J. Rivas Alberty, *Limitaciones Inmanentes de la libertad de expresión*, cit., 228.

<sup>150</sup> Ivi, 223.

<sup>151</sup> BVerfGE, 24 febbraio 1961, n. 12, 205.

<sup>152</sup> BVerfG, 04 febbraio 2010, n. 369, par. 32.

<sup>153</sup> Concetti per i quali J. Rinceanu, *Verso una forma di polizia privata nello spazio digitale?*, cit., 4 rimanda, quanto a formante giurisprudenziale, a BGH, Urt. v. 27.7.2017, 3 StR 172/17, BGH, Urt. v. 3.4.2008, 3 StR 394/07 e, quanto a dottrina, a Sternberg-Lieben, Schittenhelm, § 130, *Rd. 5a*, in Schönke, Schröder, *StGB – Strafgesetzbuch Kommentar*, Aufl. 2019, 30.

<sup>154</sup> Per un cui maggiore approfondimento si rinvia al par. 2 del presente contributo.

<sup>155</sup> Che modifica la l. organica 10/1995 del 23 novembre 1995 sul Codice penale.

<sup>156</sup> G. Giorgini Pignatiello, *Profili Comparati e problemi costituzionali della legislazione contro l'omotransfobia*, cit., 1006.

<sup>157</sup> L. complementare alla l. 15/2022 per la parità di trattamento e la non discriminazione, che modifica la l. organica n. 23/1995, e ha introdotto altre due fattispecie, l'aporofobia e l'antiziganismo. Si segna che l'introduzione di questi reati è stata rafforzata con l'elaborazione del *Protocollo d'azione per le forze di sicurezza e i corpi d'armata per i crimini d'odio e i comportamenti che violano le norme giuridiche sulla discriminazione* e la creazione di 53 uffici di Procura specializzati, un Procuratore di Sezione presso la Corte Suprema e 52

Merita una menzione il fatto che, secondo un approccio da molti definito “olistico”, il *Código penal* disciplini in modo dettagliato gli atti di discriminazione basati sull’orientamento sessuale e l’identità di genere dall’art. 510 all’art. 521<sup>158</sup>. Nello specifico, l’art. 510<sup>159</sup> sanziona la condotta di chiunque incoraggi, promuova ovvero inciti pubblicamente discriminazione, odio, ostilità, oppure violenza, direttamente o indirettamente, o leda la dignità delle persone attraverso atti che comportano umiliazione, disprezzo o discredito di un gruppo, ovvero di un’associazione, ovvero di una parte di esso o contro una persona determinata, a causa della sua appartenenza a esso, per motivi razzisti, antisemiti o per altri motivi riferiti all’ideologia, alla religione o alle convinzioni, alla situazione familiare, all’appartenenza dei suoi membri ad un’etnia, una razza o una nazione, la loro origine nazionale, il loro sesso, orientamento o identità sessuale, per ragioni di genere, malattia o disabilità<sup>160</sup>. Il c. 3 costituisce il punto nevralgico per la criminalizzazione del *ciberodio* oppure *odio online*: esso stabilisce che la disposizione in commento si applica anche quando «i fatti si sono svolti attraverso un social media, via Internet o attraverso l’uso di tecnologie informatiche, in modo da renderli accessibili a un gran numero di persone»<sup>161</sup>.

In sintesi, anche la Spagna, al pari del legislatore tedesco, ha anticipato la soglia della rilevanza penale delle condotte, incriminando quelle che siano da sole idonee a incitare la commissione di *hate speech*. Inoltre, come il legislatore francese, ha previsto un apposito comma per punire la realizzazione del reato sul web.

È bene precisare che a differenza dell’art. 510, che criminalizza le condotte di *hate speech*, l’art. 22, al c. 4, del Codice penale prevede un’aggravante generica per motivi discriminatori, la cui sussistenza darebbe luogo a un crimine d’odio in senso stretto<sup>162</sup>. Considerato che il combinato disposto di tale aggravante con il reato di provocazione e di apologia alla commissione di crimini d’odio, punito all’art. 18<sup>163</sup>, potrebbe supplire

---

procuratori provinciali per il servizio di crimini d’odio e discriminazione, così R.V. Candaliya, *Tratamiento jurídico del discurso y los delitos de odio en Irlanda y España: una visión penal comparativa*, in *Revista General de Derecho Penal*, 39, 2023, 22. Per un maggiore approfondimento circa la norma citata vedasi E. Nuñez Castaño, *Libertad de Expresión y Derecho Penal: la criminalización de los discursos extremos*, Aranzadi/Civitas, 2022, 238-241.

<sup>158</sup> L. Goisis, *La violenza di genere in ottica comparata. La recente novella spagnola. Verso la progressiva affermazione di un modello consensualistico*, in *Genius Rivista di studi giuridici sull’orientamento sessuale e l’identità di genere*, del 16 maggio 2023, 11.

<sup>159</sup> Per un approfondimento dei delitti d’odio nel quadro giuridico spagnolo sussiste un ampio numero di contributi ed opere di riferimento, come R. Alcacer Guirao, *La libertad de Odio. Discurso intolerante y protección penal de las minorías*, Marcial Pons, Madrid, 2020; M. Díaz Y García Conlledó, *El discurso del odio y el delito de odio de los artículos 510 y 510 bis del Código Penal: Necesidad de limitar*, in *Boletín Límites a la Libertad de Expresión, Juezas y Jueces para la Democracia*, 5, 2018; A. Galán Muñoz, *¿Juntos o revueltos? Algunas consideraciones y propuestas sobre la cuestionable fundamentación y distinción de los delitos de odio y el discurso del odio*, in J. León Alapont, *Temas Claves de Derecho Penal. Presente y futuro de la política criminal en España*, Bosch Editor, Barcelona, 2021.

<sup>160</sup> Art. 510, c. 1, Codice penale spagnolo.

<sup>161</sup> I. G. Benito, *Ciberodio: Un estudio de derecho penal comparado*, in *Cuadernos de RES PUBLICA en derecho y criminología*, 3, 2024, 16.

<sup>162</sup> D. P. Herradón, *La responsabilidad de las entidades deportivas como consecuencia de los delitos de odio por razones religiosas que se producen en sus instalaciones*, in S. Meseguer Velasco-E. García Antón Palacios, *Deporte, diversidad religiosa y Derecho*, Aranzadi, Cizur Menor, 2020, 221.

<sup>163</sup> J. Bernal Del Castillo, *La justificación y enaltecimiento del genocidio en la Reforma del Código Penal de 2015*, in

alla punizione anche dei discorsi d'odio<sup>164</sup>, ciò ha indotto a dubitare circa la necessità della fattispecie autonoma di cui all'art. 510<sup>165</sup>. Difatti, tra le criticità riscontrate nella formulazione, si menziona la portata apparentemente priva di limiti, che secondo alcuni condurrebbe all'estensione del suo ambito applicativo sino al "parossismo"<sup>166</sup>. A causa della mancata specificazione degli elementi tipici, l'art. 510 si riferirebbe a ogni condotta "in procinto" di raggiungere un livello di eccezionale gravità. Secondo un orientamento, la via dell'azione penale intermedia, che predilige la tecnica dell'aggravamento dei reati di espressione attraverso l'art. 22, c. 4, per colmare l'indeterminatezza concettuale dell'art. 510, sarebbe più equilibrata in termini sanzionatori. Tale approccio predilige la creazione di un'ampia rete dell'azione penale<sup>167</sup>.

Ciò nonostante, l'applicazione dell'art. 22, c. 4, porta con sé il rischio di snaturare l'essenza stessa dei discorsi d'odio, quali manifestazioni del pensiero proiettate verso l'esterno e capaci di valicare i confini dell'individuo, raggiungendo il gruppo di riferimento. Il filone interpretativo sopra menzionato non terrebbe debitamente in considerazione che, così ragionando, l'attenzione si concentrerebbe sul "microconflitto"<sup>168</sup> anziché sulla sua rilevanza super-individuale. In altre parole, si perderebbe di vista l'impatto sul gruppo che il discorso d'odio è in grado di generare e, quindi, la sua dimensione strutturale. L'unica via percorribile sembra quella di distinguere con nettezza gli ambiti applicativi dell'art. 510, strumento marcatamente eccezionale di tutela avanzata<sup>169</sup>, sia nell'ambito dell'istigazione che in quello del discorso diffamatorio, e dell'art. 22, c. 4, quale mera circostanza aggravante. Di conseguenza, una corretta interpretazione dell'art. 510 richiede una lettura particolarmente restrittiva; di qui, l'innalzamento della soglia di gravità richiesta ai fini della rilevanza penale della condotta attraverso la necessaria connotazione collettiva del discorso d'odio, a prescindere dal mezzo utilizzato<sup>170</sup>. L'applicazione della circostanza aggravante esaminata dovrebbe limitarsi ai commenti offensivi indirizzati a un individuo in quanto tale e non alla sua collettività di appartenenza (come invece accade nei discorsi d'odio)<sup>171</sup>.

In linea generale, l'intervento legislativo spagnolo si è caratterizzato per un approccio sistematico mediante l'adozione di un ricco impianto normativo, che assomma diverse tecniche incriminatrici<sup>172</sup> e che, al pari di quello francese, ha suscitato profonde critiche

---

*In Dret Penal. Revista para el análisis del Derecho*, 4, 2015, 9-12.

<sup>164</sup> I. G. Benito, *Ciberodio: Un estudio de derecho penal comparado*, cit., 29.

<sup>165</sup> A. Galán Muñoz, "¿Juntos o revueltos? Algunas consideraciones y propuestas sobre la cuestionable fundamentación y distinción de los delitos de odio y el discurso del odio", in J. León Alapont, *Temas Claves de Derecho Penal. Presente y futuro de la política criminal en España*, Bosch Editor, Barcelona, 2021, 327-328.

<sup>166</sup> Così, R. A. Guirao, *Dimensiones del discurso de odio*, in *Derecho penal y orden constitucional: límites de los derechos políticos y reformas pendientes*, Tirant lo Blanch, 2022, 41.

<sup>167</sup> Di tale orientamento dà atto I. G. Benito, *Ciberodio: Un estudio de derecho penal comparado*, cit., 30.

<sup>168</sup> *Ibid.*

<sup>169</sup> I. G. Benito, *Ciberodio: Un estudio de derecho penal comparado*, cit., 25.

<sup>170</sup> *Ivi*, 30.

<sup>171</sup> *Ivi*, 31.

<sup>172</sup> G. Giorgini Pignatiello, *Profili Comparati e problemi costituzionali della legislazione contro l'omotransfobia*, cit., 1010.

in ragione della sua notevole forza espansiva<sup>173</sup>. Già prima della riforma del 2022, si riteneva che le condotte tipizzate dagli interventi legislativi del 1995<sup>174</sup> e del 2015<sup>175</sup> potessero difficilmente conciliarsi con i principi fondamentali che governano un diritto penale democratico, in quanto ritenuti capaci di oltrepassare i limiti di un intervento punitivo minimo<sup>176</sup>.

In riferimento alla criminalizzazione dei discorsi d'odio anche nelle manifestazioni indirette, il formante giurisprudenziale ha mantenuto un orientamento ondivago<sup>177</sup>: se negli 90<sup>178</sup> ha riconosciuto la punibilità delle stesse, nei primi anni 2000<sup>179</sup> si è orientato in senso restrittivo, tornando ad abbracciare soltanto nel 2016 l'orientamento estensivo che ne aveva caratterizzato l'operato alle origini, altresì, allineandosi alla giurisprudenza della Corte EDU<sup>180</sup>.

Il Tribunale costituzionale spagnolo ha precisato che un discorso d'odio che incitava alla violenza elogiando l'autore di attività terroristiche non può rientrare nel contenuto costituzionalmente protetto del diritto alla libertà di espressione<sup>181</sup>. La giurisprudenza costituzionale è giunta a conclusioni simili in relazione a espressioni o campagne di carattere razzista o xenofobo, in quanto la Costituzione non garantisce il diritto di diffondere un determinato significato della storia o una concezione del mondo con l'intento deliberato di sminuire e discriminare persone o gruppi per ragioni legate a qualsiasi condizione personale, etnica o sociale<sup>182</sup>.

Per determinare se una condotta costituisce una manifestazione punibile, la giurisprudenza costituzionale osserva il seguente giudizio sommario: in primo luogo, valuta se si tratta di un'espressione d'odio basato sull'intolleranza, formulata in modo aggressivo e caratterizzata da ostilità inequivoca nei confronti di altri individui. Secondariamente, verifica se la diffusione mediante i mezzi di comunicazione produce un effetto equiparabile a quello dei mezzi di diffusione tradizionali, come i giornali e i notiziari televisivi. Tuttavia, tale giudizio non può ritenersi esaustivo. L'impatto potenziale, la lesione della dignità e il pericolo per il tessuto sociale sono presi in esame come elementi di carattere assoluto, senza considerare altri criteri, come l'interesse generale o il contributo

<sup>173</sup> M. Revenga Sánchez, *El discurso del odio: entre la trivialización y la hiper-penalización*, in *Liber Amicorum per Pasquale Costanzo*, 13-2-2019, 1-14, giurcost.org

<sup>174</sup> L. organica n. 10/1995 sul Codice penale.

<sup>175</sup> L. organica n. 1/2015 sul Codice penale.

<sup>176</sup> G. Giorgini Pignatiello, *Profili Comparati e problemi costituzionali della legislazione contro l'omotransfobia*, cit., 1008-1009.

<sup>177</sup> Cfr. M. Iacometti, *Il Tribunale costituzionale spagnolo verso l'ipertrofia del concetto di "discorso del odio" e la eccessiva compressione della libertà di espressione?*, in *Rivista Associazione Italiana dei Costituzionalisti*, 1, 2017.

<sup>178</sup> Ricostruzione effettuata da G. Giorgini Pignatiello, *Profili Comparati e problemi costituzionali della legislazione contro l'omotransfobia*, cit., nota 60, mediante il riferimento a Tribunale cost. spagnolo, n. 214, 11 novembre 1991.

<sup>179</sup> Tribunale cost. spagnolo, n. 235, 07 novembre 2007.

<sup>180</sup> Tribunale cost. spagnolo, n. 112, 20 giugno 2016.

<sup>181</sup> *Ibid.*, richiamato da A. Espinoza – J. Rivas Alberty, *Limitaciones Inmanentes de la libertad de expresión*, cit., 246-247.

<sup>182</sup> Tribunale cost. spagnolo, n. 214, 11 novembre 1991.

alla formazione di un'opinione pubblica libera e pluralista<sup>183</sup>. Seguendo questa logica, gli organi giudiziari arriverebbero a sostenere che determinate condotte, pur non godendo di tutela costituzionale, non potrebbero essere punite. Tale risultato verrebbe raggiunto mediante l'oggettivizzazione delle garanzie proprie del diritto penale che derivano dall'effetto di irradiazione dei diritti fondamentali operanti nel caso specifico. L'influenza di tali diritti sul piano penale comporta il rafforzamento del principio di legalità. Il diritto fondamentale può, pertanto, operare come causa di esclusione dell'antigiuridicità o come causa di giustificazione dell'eccesso nell'esercizio di un diritto, a condizione che non ne venga snaturato il contenuto<sup>184</sup>.

Il quadro comparativo sopra tracciato mostra come gli ordinamenti penalistici esaminati abbiano previsto l'incriminazione delle condotte d'odio, anche online, mediante forme di anticipazione della soglia della rilevanza penale, al fine di incriminare quelle condotte che possano anche soltanto mettere a repentaglio l'esistenza, o il godimento, dei beni giuridici tutelati. Con la precisazione che tali comportamenti devono essere rivolti all'individuo in quanto membro del suo gruppo di appartenenza. A tal fine, le scelte di politica criminale analizzate si sono caratterizzate per il ricorso alla tecnica dell'istigazione o della messa in pericolo.

In Francia, è stato sottolineato come la libertà di manifestazione del pensiero costituisca una condizione essenziale per la sopravvivenza stessa della democrazia e l'incriminazione è stata operata mediante la creazione di una figura criminosa basata sulla forma di imputazione soggettiva della "*mise en danger*".

Per quanto riguarda il quadro tedesco, è importante sottolineare l'impegno legislativo nel tentativo di definire con rigore le condotte d'odio, mediante l'introduzione di un elenco tassativo di contenuti da qualificare come illeciti. Inoltre, la fattispecie di discorso d'odio è individuata tramite il rinvio al reato di istigazione all'odio, soluzione che trova fondamento nella prevalenza del principio della dignità umana su ogni altro diritto fondamentale dell'individuo, inclusa la libertà di espressione.

In Spagna, si è posta la questione del mancato utilizzo della specifica figura incriminatrice appositamente introdotta, costruita sul modello istigatorio, al pari di quanto occorso anche nell'ordinamento tedesco. Conseguentemente, nella prassi talvolta si ricorre al combinato disposto degli artt. 18 e 22, c. 4, e, quindi, alla fattispecie generica di istigazione alla commissione di delitti aggravata dai motivi discriminatori, a causa della ritenuta indeterminatezza concettuale dell'art. 510 del Codice penale. Tale corollario costituisce uno dei principali motivi delle critiche rivolte a un così ricco impianto normativo, ritenuto in grado di superare i limiti di un intervento punitivo minimo.

In sede di accertamento, principalmente secondo un giudizio scrupoloso presidiato dai principi di proporzionalità, adeguatezza e necessità, è demandato al formante giurisprudenziale il compito di verificare lo stato di pericolo della libertà di espressione dell'autore del discorso d'odio e la capacità di sopravvivenza del nucleo di individui cui le parole pericolose si rivolgono.

Dall'analisi delle diverse esperienze giuridiche emerge un'identità di vedute anche in merito all'oggetto della tutela giuridica: il discorso d'odio deve porre in pericolo l'indi-

<sup>183</sup> A. Espinoza – J. Rivas Alberty, *Limitaciones Inmanentes de la libertad de expresión*, cit., 248.

<sup>184</sup> Ivi, 249.

viduo non come singolo, bensì come componente di una collettività e, in ultima analisi, dell'agire democratico plurale, multiculturale e eterogeneo.

In conclusione, si ritiene di concordare circa la preferibilità dell'utilizzo di strumenti di tutela penale avanzata, atteso che l'elevata soglia di gravità varcata da questi crimini giustifica la specifica incriminazione dei discorsi d'odio, anche online, nell'ottica di una salvaguardia quanto più effettiva del bene giuridico super-individuale tutelato, pilastro dello Stato di diritto democratico.

### 5. Lo stato dell'arte oltre Oceano

L'orientamento nordamericano, basato sul Primo emendamento della Costituzione<sup>185</sup>, non tollera alcuna interferenza da parte dei poteri pubblici nell'esercizio della libertà di espressione e non prevede limitazioni con riguardo ai contenuti espressi.

Questa ideologia liberale ha per molto tempo giustificato l'astensionismo dalla regolamentazione dell'*hate speech*; in tale contesto, la libertà di parola del singolo può subire limitazioni solo in presenza di altri diritti costituzionalmente tutelati, e sempre entro margini molto ristretti. Per distinguere le condotte coperte dal Primo emendamento la giurisprudenza della Corte Suprema ha fatto ricorso ai criteri del *clear and present danger*<sup>186</sup> e delle *fighting words*, il cui comune assunto è quello di ritenere che la manifestazione del pensiero possa essere limitata non tanto in relazione al contenuto del messaggio, bensì in quanto idonea a tradursi in atto violento. In altre parole, secondo tali teorie, la limitazione del pensiero può avvenire tutte le volte in cui i discorsi d'odio producono conseguenze tangibili e dannose per l'ordine pubblico.

Una prima formulazione del concetto di *clear and present danger* è stata fornita dal giudice Holmes<sup>187</sup> nel caso *Schenk v. United States* del 1919 e in *Abrams v. United States*<sup>188</sup>. In tale ultima pronuncia egli scrisse: «the United States constitutionally may punish speech that produces or is intended to produce a clear and imminent danger». Con queste parole la Corte Suprema scelse di non indicare i criteri in base ai quali graduare la pericolosità del discorso, al fine di permettere allo Stato di scegliere di vietare le opinioni apparentemente minacciose, analizzando caso per caso e tenendo in considerazione il contesto e la natura delle parole utilizzate.

In una pronuncia successiva degli anni Cinquanta, *Dennis v. United States*<sup>189</sup>, la Corte

<sup>185</sup> «Congress shall make no law respecting an establishment of religion or prohibiting the free exercise thereof; or abridging the freedom of speech, or of the press; or the right of the people peaceably to assemble, and to petition the Government for a redress of grievances».

<sup>186</sup> Il primo a teorizzare questa dottrina fu il Giudice della Corte Suprema Holmes, nel caso *Schenk v. United States*, nel 1919; successivamente sono risultate rilevanti al riguardo anche le pronunce *Chaplinsky v. New Hampshire*, del 1942 e *Brandenburg v. Ohio*, del 1969. Si segnala altresì la lettura di A. Fricano, *Prove tecniche di resilienza costituzionale: l'assalto a Capitol Hill e la censura mediatica di Donald Trump*, in *Rivista del Gruppo Pisa*, fascicolo speciale monografico, 3, 2021, 742-743.

<sup>187</sup> Si segnala A. Lewis, *Freedom for thought that we hate*, New York, 2007, 11-20, per una rassegna delle opinioni del Giudice Holmes nelle sentenze della Corte Suprema.

<sup>188</sup> *Abram v. United States*, 250 U.S. (1919).

<sup>189</sup> *Dennis v. United States*, 345 U.S. 494 (1951), inerente ad un caso di presunta incostituzionalità dello *Smith Act* del 1940, il quale stabiliva l'illegalità di costituire organizzazioni atte a sovvertire qualsiasi

Suprema, conscia dell'impossibilità di determinare la probabilità che una minaccia sovversiva si traducesse in azione, attribuì agli aggettivi *clear and present danger* portata preventiva: «the words cannot mean that, before the Government may act, it must wait until the putsch is about to be executed». Con tali parole si inaugurò un nuovo filone destinato a far scuola nel futuro diritto penale c.d. «prepressivo»<sup>190</sup>. In particolare, la Corte rimodulò il criterio della punibilità alla stregua di un pericolo astratto ed elaborò una nuova teoria del rischio presunto: se il pregiudizio è davvero grave, ai fini della punibilità dell'autore della condotta non è necessario che l'istigazione sia idonea a indurre qualcuno, in un intervallo temporale ristretto, alla sua commissione.

Successivamente<sup>191</sup>, a partire dal caso *Brandenburg v. Ohio*<sup>192</sup>, il test del *clear and present danger* iniziò ad essere applicato in modo più rigoroso. In particolare, la Corte Suprema dichiarò incostituzionale l'*Ohio criminal Syndicalism Statute*, approvato nel 1919, che prevedeva la punizione dell'istigazione alla violenza finalizzata alla promozione di riforme industriali o politiche. In tale occasione, recuperando la necessaria carica di offensività della fattispecie di istigazione, la Corte rovesciò la condanna e dichiarò protetto dal Primo emendamento qualunque forma di incitamento a usare la forza o a violare la legge (ad eccezione dell'ipotesi di produzione di un'azione illegale imminente e idonea al raggiungimento dello scopo). Nel passaggio chiave della sentenza si legge che un discorso può essere punito se si verificano due condizioni: i) intenzionalità di adottare un comportamento violento e ii) l'idoneità delle parole utilizzate a produrre un pericolo. L'idoneità e l'imminenza, quali requisiti necessari per violare i margini del Primo emendamento, sono stati successivamente affinati in *Hess v. Indiana*<sup>193</sup> e in *N.A.A.P. v. Claiborne Hardware Co*<sup>194</sup>, pur rimanendo ambigui.

Analizzando ora il criterio delle cosiddette *fighting words*, è opportuno precisare che tale concetto comprende tutte quelle espressioni verbali e non verbali che manifestano un

---

governo negli Stati Uniti mediante la forza e la violenza. Il ricorrente, Eugene Dennis, fu condannato per aver tentato di costituire un partito di matrice comunista.

<sup>190</sup> Per il significato di questo neologismo e per un'analisi della recente diffusione globale della giustizia penale «prepressiva», si rinvia a E. Grande, *Il dispositivo penale della paura. Diffusione di un'ideologia, in Democrazia e Diritto*, 3, 2014,17 e ss.

<sup>191</sup> In costanza del periodo della guerra del Vietnam.

<sup>192</sup> Un leader del Ku Klux Klan venne condannato per il reato di istigazione a un illecito penale o alla violenza (previsto dal legislatore dell'Ohio) per aver organizzato una manifestazione razzista in cui aveva paventato il ricorso alla vendetta nei confronti delle istituzioni politico-giuridiche degli Stati Uniti (Presidente, Congresso e Corte Suprema), che a suo giudizio non avrebbero tutelato abbastanza i bianchi nei confronti dei neri e degli ebrei. Alla manifestazione, cui molti parteciparono armati, il leader dell'organizzazione razzista aveva invitato un cameraman della televisione per farsi filmare. Una volta trasmesso il servizio egli fu processato e successivamente condannato.

<sup>193</sup> 414 U.S. 105 (1973). Il caso riguardava un manifestante che dopo essere stato disperso dalla polizia aveva gridato «we'll take the fucking streets», senza indicare quando. Tale istigazione fu ritenuta non capace di condurre a un'azione violenta imminente. Non c'era prova, sostenne la Corte, che le «parole fossero volte a produrre e fossero idonee a produrre un disordine imminente».

<sup>194</sup> 458 U.S. 886 (1982). In questa vicenda, inerente il boicottaggio di alcuni negozi i cui titolari erano stati accusati di discriminare persone nere, la Corte ritenne che la seguente affermazione rivolta alla gente: «se becchiamo uno di voi entrare in questi negozi razzisti vi rompiano il collo», pronunciata da uno dei responsabili dell'associazione nazionale per i diritti della gente di colore, fosse protetta dal Primo emendamento, in quanto «mere advocacy of the use of force or violence does not remove speech from the protection of the first Amendment», ivi, 928.

chiaro intento di incitare alla violenza da parte di chi le pronuncia. Si tratta, dunque, di dichiarazioni che sono immediatamente in grado di generare disordine nel contesto pubblico, in quanto finalizzate a tradursi in comportamenti idonei a suscitare reazioni violente<sup>195</sup>.

I capisaldi giurisprudenziali di questo criterio sono enucleati dalla seguente casistica. Nel caso *Chaplinsky v. New Hampshire*<sup>196</sup>, un testimone di Geova, che si era issato su un banchetto collocato sul marciapiede nel centro di Rochester, New Hampshire, inveiva contro le organizzazioni religiose. Invitato dallo sceriffo a interrompere la sua condotta, egli reagiva insultandolo con frasi come «you are a god damned racketeer» e a «damned Fascist». Arrestato, fu condannato in base a una legge nazionale che puniva la condotta di chi offendeva intenzionalmente una persona in un luogo pubblico. In merito, la Corte Suprema, investita della questione per verificarne l'incostituzionalità, all'unanimità la dichiarò infondata, ritenendo che le espressioni usate avrebbero dovuto essere ricomprese nella categoria delle *fighting words*, in quanto non idonee ad appor- tare alcun valore aggiunto al dibattito pubblico<sup>197</sup>. Nel caso specifico, la Suprema Corte individuò nella carica aggressiva delle espressioni utilizzate il discrimine per distinguere le parole violente da quelle non violente, concludendo di non poterle considerare, in quella circostanza, mera espressione di un punto di vista e, quindi, libera manifestazione del pensiero<sup>198</sup>.

Questo primo approdo venne successivamente testato nel famosissimo caso di *hate speech* americano, ovvero *Village of Skokie v. National Socialist Party of America*<sup>199</sup>. La vicenda trae origine dal divieto imposto a un partito filonazista di sfilare in corteo in una cittadina dell'Illinois, prevalentemente abitata da appartenenti alla comunità ebraica. I partecipanti alla manifestazione avrebbero, come in altre occasioni, voluto sfilare indossando uniformi naziste e mostrando cartelli con scritte del seguente tenore: «free speech for the White man» e «Free speech for White America». La Corte Suprema, chiamata a valutare la legittimità del divieto imposto, affermò che non era possibile escludere dal dibattito pubblico le espressioni ovvero l'esposizione di simboli che, solo potenzialmente, avrebbero potuto causare rabbia o risentimento nel pubblico<sup>200</sup>. A fronte di tale posizione, anche la Corte dell'Illinois stabilì che la marcia avrebbe semplicemente rappresentato una forma simbolica libertà di espressione.

Successivamente con *R.A.V. v. City of St. Paul*<sup>201</sup>, un gruppo di ragazzi adolescenti venne condannato per aver collocato una croce fiammeggiante davanti al giardino di proprietà di una famiglia afroamericana. La condanna si fondava su un'ordinanza della città di St. Paul, nel Minnesota, la quale vietava di occupare il suolo pubblico o privato con simboli, oggetti, graffiti o altro «including, but not limited to, a burning cross or Nazi swastika, which one knows or has reasonable grounds to know arouses anger, alarm or resentment in

<sup>195</sup> P. Tanzarella, *Discriminare parlando*, cit., p. 57 e ss.

<sup>196</sup> *Chaplinsky v. New Hampshire*, 568, U.S.315 (1942).

<sup>197</sup> *Chaplinsky v. New Hampshire*, 568, U.S.315 (1942).

<sup>198</sup> *Ibid.*: «are not essential part of any exposition of ideas».

<sup>199</sup> 432 U.S. 43 (1977).

<sup>200</sup> *Ibid.*

<sup>201</sup> 505, U.S. 377 (1992).

*others on the basis of race, color, creed, religion or genders*». La Suprema Corte del Minnesota, interessata della questione, sottopose alla Corte Suprema federale un dubbio di costituzionalità con riferimento al Primo emendamento. La sentenza fu redatta principalmente dal giudice Scalia, secondo il quale sarebbe stato un errore includere nella categoria delle *fighting words* le idee, facendo esclusivo riferimento al loro contenuto e non, invece, alla loro effettiva pericolosità. In effetti, ai fini dell'applicazione del criterio in esame, sarebbe rilevante solo il modo in cui l'idea odiosa viene espressa<sup>202</sup>. Sulla base di tale ragionamento si spiega, quindi, il motivo della decisione di incostituzionalità dell'ordinanza della città di St. Paul: essa non era diretta a punire un modo – pericoloso e offensivo – di espressione<sup>203</sup>, bensì esclusivamente, a selezionare i messaggi intolleranti da condannare, ovvero quelli riguardanti la razza, il genere e la religione, trascurando altri tipi di discorsi che ben avrebbero potuto urtare la sensibilità di altre categorie di persone (come, per esempio, quelli riguardanti l'orientamento sessuale, non inclusi nel divieto)<sup>204</sup>.

I principi sanciti nella sentenza *R.A.V.*<sup>205</sup> sono stati ulteriormente ribaditi in *Black v. Virginia*<sup>206</sup>, dove la croce fiammeggiante era stata innalzata nel giardino di una famiglia afroamericana. In questo caso venne sollevato un dubbio di costituzionalità sulla legge della Virginia volta a condannare «*any person [...], with the intent of intimidating any person or group [...], to burn a cross on the property of another, a highway or other public place*», specificando inoltre che «*[a]ny such burning [...] shall be prima facie evidence of an intent to intimidate a person or group*»<sup>207</sup>. Sulla scia della giurisprudenza *R.A.V.*, la Corte Suprema dichiarò l'incostituzionalità di tale disposizione, affermando che le leggi volte a punire le condotte descritte, negando la possibilità di indagare sui motivi che hanno condotto a compierle, devono necessariamente ritenersi incostituzionali. La valutazione del caso concreto deve essere sempre essere ammessa, in quanto solo attraverso di essa si potrebbe stabilire la vera natura dell'atto, se intimidatorio o meramente espressivo (come ritenuto nel caso specifico) di un dissenso, privo dell'intenzione di minacciare qualcuno<sup>208</sup>.

<sup>202</sup> *Ibid.*, così il Giudice Scalia: «*the reason why fighting words are categorically excluded from the protection of the First Amendment is not that their content communicates any particular idea, but that their content embodies a particularly intolerable (and socially unnecessary) mode of expressing whatever idea the speaker wishes to convey*».

<sup>203</sup> *Ibid.*, «*has not singled out an especially offensive mode of expression – it has not, for example, selected for prohibition only those fighting words that communicate ideas in a threatening (as opposed to a merely obnoxious) way*».

<sup>204</sup> *Ibid.*, «*the ordinance applies only to 'fighting words' that insult, or provoke violence, 'on the basis of race, color, creed, religion, or gender'. Displays containing abuse invective, no matter how vicious or severe, are permissible unless they are addressed to one of the specified disfavored topics. Those who wish to use 'fighting words' in connection with other ideas – to express hostility, for example, on the basis of political affiliation, union membership, or homosexuality – are not covered*».

<sup>205</sup> Si segnala che i giudici statali hanno riscontrato difficoltà applicative dei principi sanciti nella citata sentenza, sul punto si richiama l'analisi di V. M. Manetti, *L'incitamento all'odio razziale tra realizzazione dell'eguaglianza e difesa dello Stato*, Torino, 2005, 103 – 137, la quale ricorda come il Giudice Stevens avrebbe ribadito, nella sua opinione dissenziente, che il caso *RAV* avrebbe assicurato una tutela maggiore al pensiero razzista rispetto a quella conferita al *commercial speech*.

<sup>206</sup> 538, U.S. 343 (2003).

<sup>207</sup> Va. Code Ann. § 18.2-423 (1996).

<sup>208</sup> Questa argomentazione è stata criticata dai sostenitori della teoria critica della razza c.d. *critical race theory*. Per questi ultimi, i discorsi d'odio non costituiscono semplici opinioni fuori dal coro ma, essendo diretti contro minoranze vulnerabili (come la comunità afroamericana negli Stati Uniti), rappresentano atti di violenza verbale volti ad avvilire le vittime, a ridurle al silenzio e a emarginarle sul piano sociale e

Posta in questi termini la teoria delle *fighting words*, così come quella del *clear and present danger*, risulta relegata ai casi limite, facendo apparire gli Stati Uniti come i massimi garanti della libertà di manifestazione del pensiero<sup>209</sup>. Entrambe, tuttavia, mostrano chiaramente che la Corte Suprema non sia mai riuscita (o non abbia mai voluto) classificare in maniera tassativa le ipotesi di punibilità dei discorsi d'odio in relazione al pericolo concreto prodotto, lasciando ai giudici ampio spazio di apprezzamento a seconda delle singole circostanze. Ciò, diversamente rispetto a quanto avviene nelle Corti Europee ove, come visto<sup>210</sup>, sulla base dell'art. 10 CEDU e dell'art. 11 della Carta dei diritti fondamentali dell'Unione europea, pur proclamando la libertà di espressione come diritto fondamentale, si delineano le situazioni giuridiche nelle quali essa trova espressione<sup>211</sup>. Nel contesto attuale dell'esperienza giuridica oltreoceano, sono le modalità espressive a giocare un ruolo centrale, costituendo l'unico criterio per valutare l'applicazione del Primo Emendamento. In particolare, per quanto riguarda i mezzi di informazione, le differenze nelle loro caratteristiche giustificerebbero una differenziazione negli standard applicativi dello stesso Emendamento<sup>212</sup>.

Con specifico riferimento alla libertà di espressione online, a partire dalla nota sentenza *Reno v. ACLU*, del 1997, relativa all'incostituzionalità del *Communication Decency Act*

---

politico. Ci si può anche chiedere se intimidazioni come la pratica delle croci infiammate – tipica del *Ku Klux Klan* – siano forme di espressione ammesse in un dibattito di idee degno di questo nome. Infine, in un'ottica europea, la sentenza colpisce anche per le sue considerazioni sbrigative sui diritti delle minoranze: su questo tema, la Corte si limitò infatti ad affermare che, per proteggerli, il legislatore non deve per forza ricorrere a misure che restringono la libertà di espressione ma dispone di alternative.

<sup>209</sup> Per un esame ulteriore dei casi di *hate speech* rispetto ai quali hanno trovato applicazione le dottrine esaminate si rimanda a E. Grande, *Il dispositivo penale della paura*, cit. 47 ss.

<sup>210</sup> Cfr., paragrafo 2.

<sup>211</sup> O. Pollicino, *La prospettiva costituzionale sulla libertà di espressione nell'era di Internet*, in *Rivista di Diritto di Media*, 1/2018, cit. 3 ss.; M. Bassini, *Internet e libertà di espressione. Prospettive costituzionali e sovranazionali*, Roma, 2019, 163 ss.

<sup>212</sup> Il *leading case* è rappresentato dalla decisione *Red Lion Broad. Co. v. FCC*, del 1969, nell'ambito della quale la Corte Suprema ha valutato la legittimità della c.d. *fairness doctrine*, una policy della *Federal Communications Commission (FCC)* (fondata sul «Radio Act» del 1927 e sulla sez. 307 (licenze) del Titolo 47 (radiodiffusione) dello *US Code*), introdotta nel 1949, in base alla quale la Commissione, in considerazione della scarsità delle risorse nella diffusione radiofonica e televisiva, subordinava la concessione della licenza a trasmettere all'impegno da parte dei titolari delle relative licenze a fornire un servizio di informazione pubblica orientato al pluralismo e alla correttezza. Il tema della scarsità delle risorse d'accesso al mezzo televisivo viene poi ripreso anche in altre decisioni successive ove la Corte sancisce la legittimità costituzionale di alcune disposizioni c.d. *must carry* finalizzate a garantire il pluralismo delle emittenti radiotelevisive imponendo, ad esempio, che una percentuale delle reti via cavo fosse riservata all'emittenza locale. La giurisprudenza della Corte Suprema in relazione alla limitazione della libertà di espressione in ambito radiotelevisivo si spinge sino a pronunciarsi anche sugli stretti contenuti veicolati tramite tali mezzi. In particolare, nella decisione *FCC v. Pacifica Foundation* del 1978, dove conferma il potere della *FCC* di regolare, in circostanze limitate (stabilendo orari e condizioni di trasmissione), programmi che pur non osceni fossero qualificabili come indecenti. A fianco della motivazione concernente la scarsità del bene, nella decisione in esame uno dei perni in relazione al quale il relatore per la Corte ritiene di differenziare il regime di applicabilità del Primo emendamento al mezzo radiotelevisivo è legato alla passività dell'utenza ed alla sua estrema diffusione presso le abitazioni private. Questa giurisprudenza restrittiva, inaugurata per i mass media radiotelevisivi, non si è poi estesa alla carta stampata la quale, anche dopo tale decisione, ha continuato a godere nelle sentenze della Corte Suprema di un'ampia libertà di espressione.

del 1996<sup>213</sup>, non hanno trovato applicazione i precedenti giurisprudenziali aventi ad oggetto la diffusione del discorso tramite l'emittenza radiotelevisiva. La Corte, nella persona del Giudice Stevens, operò un netto distinguo, affermando che, Internet non dovesse essere considerato altrettanto invasivo dei *mass media* tradizionali, in quanto l'accesso alle informazioni online richiederebbe un'attività di ricerca da parte dell'utente, in ragione della quale, esso, non potrebbe qualificarsi passivo<sup>214</sup>.

Circa vent'anni dopo il caso *Reno*, la Corte suprema con la sentenza *Packingham v. North Carolina* del 19 giugno 2017<sup>215</sup> ebbe l'opportunità di tornare a ragionare sul diritto di Internet<sup>216</sup>, seppur in un caso non strettamente attinente ai discorsi di odio online<sup>217</sup>. L'esordio della parte motivazionale della sentenza è significativo nella misura in cui evidenzia che il Primo emendamento trova piena attuazione se chiunque ha «accesso a luoghi in cui possa parlare ed ascoltare, e poi, dopo riflessione, parlare ed ascoltare ancora»<sup>218</sup>.

Un altro passaggio rilevante della sentenza è quello in cui specifica che, se in passato potevano nutrirsi perplessità riguardo al luogo più importante per liberamente esprimere la propria opinione, oggi (nel 2017) non vi sono più dubbi: tali sono il *cyber spazio* e i *social media*<sup>219</sup>. Se, quindi, anche in rete possono essere commessi alcuni reati, secondo la Corte non sono giustificati divieti volti a colpire in maniera integrale l'esercizio dei diritti di cui al Primo emendamento sul web «tessuto connettivo della nostra moderna società e cultura»<sup>220</sup>.

L'attuale modello statunitense non ha raggiunto un soddisfacente equilibrio tra la libertà di espressione e gli altri interessi di pari livello, non riuscendo a concretizzare in maniera efficace la tutela dalle condotte di *online hate speech* a causa della difficoltà di anticipare la tutela ad un momento in cui altri diritti fondamentali non hanno ancora subito una lesione apprezzabile<sup>221</sup>. Un simile approccio sembra costituire un terreno fertile per i siti Internet che promuovono contenuti discriminatori

---

<sup>213</sup> Atto normativo adottato dal Congresso nel 1996 a protezione dei minori dalla diffusione di materiale osceno o pornografico.

<sup>214</sup> 521 U.S. (1997). Una traduzione italiana della sentenza citata può essere consultata in *Foro Italiano*, 198, VI, 23.

<sup>215</sup> 582 U.S. (2017), No. 15-1194.

<sup>216</sup> V. Neri, *Il lato oscuro della Rete e l'esigenza di una legislazione responsabile. Due recenti casi emblematici*, in *Federalismi.it*, 22 novembre 2017, 22, 2017.

<sup>217</sup> Nel 2008 il legislatore del North Carolina introdusse una nuova fattispecie di reato consistente nel prevedere una pena per coloro i quali, iscritti nel registro dei criminali sessuali, accedevano ad un *commercial social networking web site* pur consapevoli del fatto che il sito avrebbe permesso a soggetti minorenni di diventare membri della community o di creare e mantenere pagine personali sul sito: «Offense. - It is unlawful for a sex offender who is registered in accordance with Article 27° of Chapter 14 of the General Statutes of access a commercial social networking Web site where the sex offender knows that the site permits minor children to become members or to create or maintain personal Web pages or have commercial social networking Web sites» (North Carolina General Statutes, 14-202.5(a)).

<sup>218</sup> Corte Suprema federale, *Packingham v. North Carolina*, 582 U.S. (2017), slip opinion, 4.

<sup>219</sup> *Ibid.*, 5.

<sup>220</sup> *Ibid.*, 10.

<sup>221</sup> M. Lamanuzzi, *Il "lato oscuro della rete": odio e pornografia non consensuale. Ruolo e responsabilità dei gestori delle piattaforme social oltre la net neutrality*, in *La Legislazione Penale*, 24 maggio 2021, 12 e ss.

Negli anni più recenti, con l'obiettivo di trovare un equilibrio tra la salvaguardia del libero mercato delle idee e il principio di uguaglianza, i legislatori statali<sup>222</sup> hanno cercato di dotarsi di atti normativi anti *hate speech* per tutelare la dignità dei propri cittadini appartenenti alle cosiddette categorie deboli, quali gli immigrati di nuova generazione, gli afro-americani, gli omosessuali e le minoranze religiose, che reclamano una maggior protezione dalle istituzioni, proprio in ragione dell'aumento della frequenza di fenomeni di intolleranza nei loro confronti.

### **6. Rimedi alternativi: quale ruolo per gli strumenti di “moderazione algoritmica”?**

In questo panorama complesso, in cui le pubblicazioni online sono sempre più numerose e provengono da categorie anche molto diverse tra loro, individuare la soluzione atta a contrastare efficacemente il fenomeno in esame significa trovare il giusto coordinamento tra la legge, la sanzione penale e altre possibili misure tecnico-giuridiche. In tale ottica, a partire dalla prima decade degli anni 2000, le piattaforme digitali hanno iniziato a utilizzare gli strumenti di moderazione e a redigere policy dei contenuti diffondibili in rete<sup>223</sup>, con l'obiettivo di impedire la pubblicazione, ovvero consentire la rimozione di messaggi, immagini o espressioni ritenuti deplorabili o inappropriati, sulla base di criteri stabiliti unilateralmente dalle stesse piattaforme<sup>224</sup> e, quindi, altamente discrezionali.

A partire dal 2010 l'attenzione delle scelte legislative e politiche con riferimento alle

---

<sup>222</sup> Al riguardo si può far riferimento al report diffuso dall'Anti-Defamation League del 2012, con il quale si è documentato che 43 Stati e il distretto di Washington D.C. hanno approvato leggi sui crimini d'odio. Si segnala inoltre che sono parecchie le vie attraverso le quali gli Stati hanno cercato di disciplinare tale fenomeno, soprattutto a seguito dell'attentato alle Torri Gemelle di New York. In particolare, non sono mancate le prese di posizione di alcuni professori universitari licenziati a seguito delle rispettive posizioni antipatriottiche e di condanna alle politiche estere degli Stati Uniti. Per una panoramica completa sul punto si richiama E. Grande, *Il dispositivo penale della paura*, cit., 60-61.

<sup>223</sup> M. Monti, *Privatizzazione della censura e Internet Platforms: la libertà di espressione e i nuovi censori dell'agorà digitale*, in *Rivista Italiana di informatica e diritto*, 1, 2019, 37.

<sup>224</sup> Può essere di interesse, a conferma dell'elevata discrezionalità della classificazione dei contenuti e della conseguente individuazione di quelli vietati, richiamare alcuni recenti interventi dell'*Independent Oversight Board di Facebook*, relativi a contenuti rimossi dalla piattaforma perché ritenuti di incitamento all'odio e riabilitati in quanto riconducibili alla critica politica. È il caso del post contraddistinto da un'immagine di una serie tv turca raffigurante un combattente, con la didascalia «se la lingua del Kafir (non musulmano) si scaglia contro il Profeta, allora la spada deve essere tirata fuori dal fodero» e le connesse definizioni del presidente francese Emmanuel Macron come «il diavolo». Il Comitato, nel caso in esame aveva revocato (decisione 2020-007-FB-FBR) la cancellazione del post originariamente disposta da Facebook che lo aveva ritenuto di incitamento all'odio, qualificandolo come critica politica alla risposta di Macron nei confronti della violenza di matrice religiosa. Un altro caso è quello della pubblicazione di una (presunta) citazione di Goebbels, ministro della propaganda del regime nazista, secondo cui «invece di appellarsi agli intellettuali, le discussioni dovrebbero fare leva sulle emozioni e sugli istinti». Questa frase, in un primo momento rimossa da Facebook in quanto ritenuta una forma di elogio e di supporto al regime nazista, era poi stata nuovamente pubblicata a seguito della decisione del citato Comitato (decisione 2020-005-FB-UA), in quanto volta a creare un parallelo tra il concetto espresso dalla citazione e la presidenza di Donald Trump e quindi riconducibile alla critica politica nei confronti del Presidente.

tecniche di regolazione della libertà di espressione è costantemente aumentata, scegliendosi di intervenire nella regolazione delle infrastrutture digitali attraverso l'elaborazione di forme di responsabilità sussidiaria a carico dell'intermediario<sup>225</sup>, evitando forme di restrizione e di sanzione che possano ledere la libertà di espressione dei singoli individui. La circostanza che fossero le piattaforme aderenti a dover effettuare una valutazione sostanziale dei contenuti da moderare (e rimuovere), facendosi quindi carico di un delicato bilanciamento fra i valori di rilievo costituzionale in gioco, prospettava una pericolosa equiparazione dei PSI agli organi giudicanti, chiamati a sindacare la legalità delle condotte espressive<sup>226</sup>.

In Europa, la normativa relativa alla moderazione dei PSI si compone di misure legislative, non legislative e volontarie, sia di carattere generale (applicabili a tutti i contenuti) che di carattere speciale (con riferimento agli specifici obiettivi da raggiungere)<sup>227</sup>. Il punto di riferimento è rappresentato dal *Code of conduct on countering illegal hate speech online*, atto a carattere non vincolante, adottato il 30 maggio 2016 e immediatamente sottoscritto dai principali attori del mondo digitale. Questo strumento nasce dall'esigenza di costruire un'alleanza tra gli attori privati (PSI) e le istituzioni europee<sup>228</sup>. La disciplina di tale codice affida alle piattaforme il compito di eliminare contenuti qualificati come atti di incitamento all'odio.

Seppur pregevole in quanto rappresenta un tentativo concreto di coniugare i diversi strumenti utilizzati per impedire il dilagare del fenomeno di odio online, tale codice

<sup>225</sup> In una prima fase, la Commissione europea faceva per lo più ricorso a strumenti di auto-regolazione. Tra questi, il rimando è soprattutto al Codice di condotta dell'UE per contrastare l'illecito incitamento all'odio (2016), nonché al Codice di buone pratiche sulla disinformazione (2019). Progressivamente, l'approccio della Commissione è mutato nella direzione di una maggiore regolazione dall'alto, attraverso il ricorso sempre più diffuso a strumenti di *hard law*. Si vedano, in particolare: la direttiva (UE) 2018/1808 del Parlamento europeo e del Consiglio, del 14 novembre 2018, recante modifica della direttiva 2010/13/UE, relativa al coordinamento di determinate disposizioni legislative, regolamentari e amministrative degli Stati membri concernenti la fornitura di servizi di media audiovisivi (direttiva sui servizi di media audiovisivi), in considerazione dell'evoluzione delle realtà del mercato (2018) OJ L61/69; la direttiva (UE) 2019/790 del Parlamento europeo e del Consiglio, del 17 aprile 2019, sul diritto d'autore e sui diritti connessi nel mercato unico digitale e che modifica le direttive 96/9/CE e 2001/29/CE (2019) OJ L130/92; il Regolamento (UE) 2021/784 del Parlamento europeo e del Consiglio, del 29 aprile 2021, relativo al contrasto della diffusione di contenuti terroristici online (2021) OJ L172/79. Si veda, da ultimo il c.d. Digital Services Act: Proposta di Regolamento del Parlamento europeo e del Consiglio relativo a un mercato unico dei servizi digitali (legge sui servizi digitali) e che modifica la direttiva 2000/31/CE, COM (2020) 825.

<sup>226</sup> G. Vasino, *Censura privata e contrasto all'hate speech nell'era delle Internet Platforms*, in *Federalismi*, 8 febbraio 2023, 130-159.

<sup>227</sup> Direttiva 2000/31/CE sul commercio elettronico.

<sup>228</sup> Commission staff working document *Countering racism and xenophobia in the EU: fostering a society where pluralism, tolerance and non-discrimination prevail*, in Brussels, 15.3.2019, SWD (2019) 110 final. Si segnala che l'elaborazione dell'atto è stata preceduta dal Colloquium *Tolerance and respect: preventing and combating anti-Semitic and anti-Muslim hatred in Europe* (internet Forum, dicembre 2015) in cui la Commissione manifestava un interesse all'individuazione di strumenti adeguati a prevenire la diffusione di discorsi discriminatori con specifico riferimento alle condizioni di alcune minoranze (promuovendo un approccio inizialmente fondato, appunto, sul *soft law*). La pubblicazione del Codice è stata seguita dalla emanazione di una comunicazione contenente importanti linee guida relative all'implementazione del Codice (si veda: *Communication from the Commission to the European Parliament, the Council, the European economic and social committee and the Committee of the regions Tackling Illegal Content Online - Towards an enhanced responsibility of online platforms*, Bruxelles, 28 settembre 2017, COM (2017) 555).

presenta profili controversi che meritano di essere analizzati per comprendere se (e come) la moderazione possa rappresentare uno strumento di contrasto efficace.

In primo luogo, è necessario evidenziare la difficoltà di trovare, all'interno degli Stati membri, una definizione condivisa di odio online. Il Codice rimanda alla formulazione contenuta nella Decisione Quadro 2008/913/GAI<sup>229</sup>, la quale identifica l'*hate speech* come «*incitement to hatred*» e non come «*incitement to violence, discrimination e hostility*»<sup>230</sup>. A seconda della giurisdizione di riferimento, le condotte ascrivibili ai discorsi d'odio penalmente rilevanti presentano delle differenze considerevoli. A loro volta le piattaforme e gli intermediari digitali tendono a definire autonomamente il concetto di *hate speech* sanzionabile ai sensi dei loro termini e condizioni d'utilizzo, adottando nozioni<sup>231</sup> notevolmente più ampie e aperte rispetto alle fattispecie considerate dai sistemi giuridici nazionali. Il campo di applicazione di tali standard privati rischia di risultare estremamente ampio e, per certi versi, indefinito, con il pericolo di compromettere la lotta al contrasto al fenomeno dell'*online hate speech*<sup>232</sup>.

Secondariamente, si deve affrontare il problema della fallibilità dello strumento algoritmico come ausilio alla moderazione e il suo conseguente impatto sui diritti fondamentali degli utenti e delle categorie vulnerabili.

Le piattaforme devono necessariamente fare ricorso a strumenti di moderazione algoritmica complessi e raffinati<sup>233</sup>, basati sull'utilizzo di sistemi di intelligenza artificiale e *machine learning*<sup>234</sup>. Tale soluzione, sebbene sia l'unica effettivamente praticabile, cela evidenti rischi connessi al loro utilizzo<sup>235</sup>. I *software* di moderazione automatica dei contenuti sono inevitabilmente soggetti a errore in quanto funzionano sulla base di modelli probabilistici. Ciò significa che, anziché operare con certezza assoluta, questi sistemi effettuano delle previsioni su quale contenuto sia appropriato o meno, con un margine di incertezza che può portare a decisioni imprecise<sup>236</sup>. Il problema è rappresentato dal fatto che l'errore ha un impatto sulle comunità tradizionalmente emarginate

<sup>229</sup> Per una lettura critica della Decisione Quadro del 2008, si rimanda a T. Moschetta, *La decisione quadro 2008/913/GAI contro il razzismo e la xenofobia: una «occasione persa» per l'Italia?* in *Rivista di Diritto dell'Economia, dei Trasporti e dell'Ambiente*, vol. XII, 2014, 31, ss.

<sup>230</sup> B. Bukovska, *The European Commission's Code of Conduct for Countering Illegal Hate Speech Online. An analysis of freedom of expression implications*, in *Transatlantic working group*, 4, ove afferma: «*incitement to hatred makes the proscribed outcome an emotional state or opinion, rather than the imminent and likely risk of a manifested action (discrimination, hostility or violence)*».

<sup>231</sup> Si veda, in tal senso, R. Wilson-M. Land, *Hate Speech on social media: Content Moderation in Context*, in *Connecticut Law Review*, 52(3), 2021, 1029-1076.

<sup>232</sup> P. Dunn, *Moderazione automatizzata e discriminazione algoritmica: il caso dell'hate speech*, in *Rivista Italiana di Informatica e Diritto*, fasc. 1-2022.

<sup>233</sup> J. Grimmelmann, *The Virtues of Moderation*, in *Yale Journal of Law and Technology*, 17, 2015, 42-109.

<sup>234</sup> Vi sono tre modelli di moderazione: a) umana (o manuale); b) di intelligenza artificiale (moderazione algoritmica o automatica); c) ibrida (combinazione di a) e b)), in questo caso la funzione dei sistemi di intelligenza artificiale è quella di operare una scrematura preventiva dei contenuti pubblicati dagli utenti e rimettere al moderatore umano solo i casi più ambigui). I sistemi ibridi sono quelli che hanno assunto maggior rilievo.

<sup>235</sup> O. Pollicino, *Judicial protection of fundamental rights on the Internet*, in *Oxford*, 2021, 193-194.

<sup>236</sup> G. De Gregorio-O. Pollicino-P. Dunn, *Digitisation and the central role of intermediaries in a post-pandemic world*, in *MediaLaws.eu*, 2021; G. Sartor-A. Loreggia, *The impact of algorithms for online content filtering or moderation*, 2020.

e discriminate<sup>237</sup>, data la ricorrente presenza di *bias* involontari<sup>238</sup>. Nonostante i tentativi di perfezionamento dei dispositivi, permane un'incapacità della macchina di comprendere il metatesto dell'espressione esaminata. Gli elementi intrinseci del contenuto pubblicato quali, ad esempio, sfumature semantiche<sup>239</sup>, o l'uso ironico o satirico di una determinata terminologia, sfuggono allo screening del sistema; in altre parole, con la moderazione online, diversamente rispetto ai discorsi d'odio offline, si rinunciarebbe alla (irrinunciabile) contestualizzazione della condotta.

In sintesi, le numerose disfunzionalità algoritmiche idonee a generare un ampio numero di falsi positivi restano numerose. Se riletto nella cornice della tutela dei diritti fondamentali dell'utente, questo aspetto delinerebbe uno scenario particolarmente grave, anche in considerazione dei plurimi studi<sup>240</sup> multidisciplinari che hanno dimostrato che l'errore del sistema di moderazione automatizzato impatterebbe maggiormente sulle comunità tradizionalmente marginalizzate e discriminate<sup>241</sup>. Gli studi citati sembrano mostrare un approccio dell'algoritmo orientato ad un criterio di uguaglianza formale, teso al mantenimento (e perseguimento) degli equilibri sociali preesistenti, per evitare di creare ulteriori disuguaglianze<sup>242</sup>. Ciò tuttavia, non può essere accettato,

---

<sup>237</sup> Un interessante approfondimento sul tema dei bias discriminatori negli algoritmi delle piattaforme digitali è stato fornito da S.U. Noble, nel suo libro *Algorithms of Oppression: How Search Engines Reinforce Racisms*, New York, 2018). Come ormai ampiamente documentato dalla letteratura, il margine di errore dei sistemi automatizzati utilizzati per rilevare l'*hate speech* ha un impatto particolarmente significativo sulle comunità tradizionalmente marginalizzate e discriminate. In effetti, sono sempre più numerosi gli studi che si concentrano sullo sviluppo di tecniche di *debiasing* per i moderatori automatici, ma il problema rimane lontano dalla soluzione. Per esempio, è stato osservato che i contenuti pubblicati da membri delle comunità afroamericana o LGBTQIA+ sono spesso ingiustamente penalizzati per presunta violazione delle normative contro l'*hate speech* o il *toxic speech*. Le cause di questo fenomeno sono molteplici. Ad esempio, i dataset utilizzati per addestrare gli algoritmi non sono sempre di alta qualità, soprattutto perché non riflettono adeguatamente il linguaggio e le modalità comunicative proprie di questi gruppi minoritari. Inoltre, molte volte i gruppi marginalizzati utilizzano termini ed espressioni che, seppur potenzialmente offensivi in altri contesti, sono intrinsecamente parte della loro comunicazione. L'incapacità degli algoritmi di cogliere queste sfumature e il loro significato implica un alto rischio di errori, con la possibilità di falsi positivi.

<sup>238</sup> G. Ziccardi, *Odio online. Violenza verbale e ossessioni in rete*, Milano, 2016, 97 ss.

<sup>239</sup> La non sensibilità dell'algoritmo è stata oggetto di numerosi studi, in tempi in cui ancora non si faceva largo uso dei sistemi di IA, finalizzati ad analizzare la capacità del sistema di rilevare correttamente tweet razzisti nei confronti della comunità afroamericana, cfr., ad esempio, I. Kwok-Y. Wang, *Locate the Hate: Detecting Tweets against Blacks*, in *Proceedings of the AAAI Conference on Artificial Intelligence*, Washington, DC, USA, vol. 27, 2013.

<sup>240</sup> P. Dunn, *Moderazione automatizzata*, cit., 137, il quale osserva come tale effetto distorsivo si rileva anche a livello di *content curation* e non solo nell'*hard moderation* in senso stretto. Più specificamente, è stato sottolineato in dottrina come il meccanismo algoritmico tenda a perseguire una generale massimizzazione dell'*engagement* dei gruppi di maggioranza con l'effetto di ridurre sempre di più gli spazi riservati a gruppi marginalizzati.

<sup>241</sup> Questa tendenza, registrata in relazione a diversi gruppi minoritari quali, ad esempio, la comunità LGBTQ+85 e la comunità afroamericana si scontrerebbe con il meccanismo di funzionamento dell'algoritmo. Le principali disfunzionalità che ancora si rilevano si sostanziano in *lexical bias* (meccanica associazione di una parola a un contenuto tossico) e *dialectal bias* (i quali derivano dall'incapacità del sistema di cogliere l'uso peculiare che si fa di un determinato lessico all'interno di una subcategoria linguistica o dialetto).

<sup>242</sup> Si segnalano al riguardo per completezza le interessanti riflessioni e le proposte correttive, basate sui meccanismi di *bias transforming*, di S. Watcher-B. Mittelstadt-C. Russel, *Bias Preservation in Machine Learning: The Legality of Fairness Metrics Under EU Non-Discrimination Law*, in *Virginia Law Review*, 123(3),

comportando altrimenti un secondario effetto discriminatorio molto superiore per alcune minoranze a base razziale o fondato sull'orientamento sessuale<sup>243</sup>. In definitiva, l'effetto paradossale dell'utilizzo della moderazione algoritmica, seppur indispensabile, consisterebbe nel rovesciamento del principio di uguaglianza sostanziale sotteso al percorso normativo euro-unitario, modellato sull'accrescimento delle tutele e finalizzato ad eliminare ogni forma di disparità di trattamento<sup>244</sup>.

Gli aspetti analizzati, seppur brevemente, hanno evidenziato come la disciplina di moderazione del Codice possa avere un impatto sui diritti fondamentali degli utenti. La Commissione europea ha mostrato una crescente consapevolezza in tal senso, come documentato dal Regolamento (UE) 2021/784<sup>245</sup>. Anche il *Digital Services Act* contiene alcune disposizioni in tal senso, come l'articolo 12, il quale richiede agli intermediari di applicare condizioni generali dei loro servizi in modo «equo, trasparente, coerente, diligente, tempestivo, non arbitrario, non discriminatorio e proporzionato», nel pieno rispetto dei diritti e degli interessi legittimi di ogni parte coinvolta. Inoltre, l'articolo 17 impone alle piattaforme online di predisporre sistemi interni di gestione dei reclami da attuarsi «in modo tempestivo non discriminatorio, diligente e non arbitrario». Sebbene tali disposizioni rappresentino un sicuro passo in avanti, da più parti è stato rilevato come in realtà raffigurino petizioni di principio, non essendo corredate di apparati applicativi definiti e sviluppati. Il sistema introdotto dal DSA incentiverebbe un incremento dell'utilizzo su vasta scala di sistemi di moderazione automatizzati senza, tuttavia, prevedere adeguati rimedi a tutela dell'individuo<sup>246</sup> né, tantomeno, sufficiente attenzione con riferimento al rischio di un'iniqua rimozione dei contenuti.

Secondo l'opinione di chi scrive, il nodo cruciale della moderazione risiede nell'individuazione della soglia oltre la quale il margine di errore degli strumenti algoritmici di moderazione automatica (falso positivo) non può più essere considerato accettabile. In conclusione, sebbene tali strumenti siano validi e, anzi, indispensabili, essi da soli non sono sufficienti. La soluzione, provvisoria e necessariamente legata al contesto storico, non può che consistere nell'integrazione tra diritto e tecnologia, alla luce dei risultati emersi dalla discussione pubblica.

---

2021, 735 ss.

<sup>243</sup> Si segnalano alcune analisi statistiche, le quali sembrano dimostrare come l'incapacità dello strumento di moderazione automatizzata di cogliere il contesto porti ad un maggior numero di contenuti eliminati di utenti appartenenti alla comunità delle *drag queen* rispetto a contenuti discriminatori di nazionalisti bianchi, cfr. O. Dias Thiago-D. M. Antonioli-A. Gomes, *Fighting Hate Speech, Silencing Drag Queens? Artificial Intelligence in Content Moderation and Risks to LGBTQ Voices Online*, in *Sexuality and Culture*, 25, 2021, 700 ss.

<sup>244</sup> G. Vasino, *Censura privata*, cit. 146-147.

<sup>245</sup> Il Regolamento all'art. 5 prevede che un fornitore di servizi riconosciuto come esposto a contenuti terroristici debba predisporre misure specifiche volte a contrastarne la diffusione: nell'applicare tali misure, tuttavia, dovrà tenere pienamente conto dei diritti e degli interessi legittimi degli utilizzatori (ivi inclusa la libertà di espressione e di informazione) e, allo stesso tempo, agire in maniera diligente e, soprattutto, non discriminatoria. È inoltre disposta, all'art. 10, la predisposizione di meccanismi di reclamo a tutela degli utenti i cui contenuti siano stati rimossi, con l'obbligo per il fornitore di rendere decisioni motivate e fatto salvo l'eventuale ricorso all'autorità amministrativa o giudiziaria dello Stato.

<sup>246</sup> J. Barata, *The Digital Services Act and Its Impact on the Right to Freedom of Expression: Special Focus on Risk Mitigation Obligations*, in *DSA Observatory*, 27 July 2021.

## 7. I rapporti tra diritto penale, libertà di manifestazione del pensiero nell'era digitale e democraticità

Dal quadro sin qui delineato emerge chiaramente l'assunzione di una dimensione e di una rilevanza pubblicistica da parte dei PSI, che si evince chiaramente dalla definizione recentemente data ai social network di «un pezzo di infrastruttura democratica immateriale». In un Paese democratico, non è accettabile che i crimini d'odio online vengano classificati e trattati come mere vicende private, gestite a livello contrattuale tra PSI e utenti, dal momento che le dette piattaforme costituiscono strumenti di cittadinanza essenziali per la democrazia e spazi globali di confronto<sup>247</sup>. Limitazioni alla libertà di espressione sono ammissibili esclusivamente se imposte da un ente statale e non per ottemperare a logiche aziendali<sup>248</sup>.

La migrazione da un "Internet delle reti" ad un "Internet delle piattaforme"<sup>249</sup>, ha concentrato nelle mani di pochi attori privati (si badi che le più grandi piattaforme digitali occupano posizioni oligopolistiche o di monopolio sul *social marketplace of ideas*<sup>250</sup>) una quantità di potere ordinativo (contrattuale ed economico) che si configura come auto-sufficiente e autonomo. Ne deriva che il cyberspazio aspirerebbe a rendersi immune da interventi pubblicistici. Di qui, la precarietà e fragilità delle libertà degli utenti<sup>251</sup>.

Si pone, di conseguenza, il problema di trasferire "la Costituzione nella rete"<sup>252</sup>. Ciò richiede di procedere a una costituzionalizzazione delle interazioni che hanno luogo sul web per garantire non solo una società reale democratica, ma anche una società digitale democratica. La rivoluzione informatica sta modificando l'idea stessa di realtà e la rete è reale prima ancora che virtuale; essa non si può concepire come semplice estensione immateriale delle dinamiche della società, bensì come luogo elettivo per la costruzione della società civile<sup>253</sup>.

L'esigenza di strutturare e presidiare i fenomeni digitali in maniera costituzionalmente

<sup>247</sup> P. Falletta, *Controlli e responsabilità dei social network sui discorsi d'odio online*, cit., 148.

<sup>248</sup> Per un maggiore approfondimento sul tema vedasi V. Claussen, *Fighting hate speech and fake news*, cit., 135 ss.

<sup>249</sup> C. Confortini, *Diffamazione e discorso d'odio in internet*, cit., 694, che pone in risalto la progressiva crescita di un'economia basata sui dati c.d. «data driven economy» con richiamo, altresì, alla Comunicazione della Commissione al Parlamento europeo, al Consiglio, al Comitato Economico e Sociale europeo e al Comitato delle Regioni «Costruire un'economia dei dati europea», COM/2017/09.

<sup>250</sup> Cfr. K. Klonick, *The New Governors: The People, Rules, and Processes Governing Online Speech*, in *Harvard Law Review*, 131, 2018, 1630. Vedasi anche F. Pasquale, *Privacy, Antitrust, and Power*, in *George Mason Law Review*, 4, 2013, 20, 1015 ss.; L.M. Khan, *Amazon's Antitrust Paradox*, in *Yale Law Journal*, 126, 2017, 710 ss. e L. Califano, *La libertà di manifestazione del pensiero... in rete; nuove frontiere di esercizio di un diritto antico. Fake news, hate speech e profili di responsabilità dei social network*, in *Federalismi.it*, 26, 2021, 9.

<sup>251</sup> C. Confortini, *Diffamazione e discorso d'odio in internet*, cit., 702 che evidenzia come, al fine di contemperare i beni in discussione, il *Digital Services Act* abbia fortemente inciso sul contratto stipulato tra la piattaforma digitale e l'utente, il quale, pur rimanendo espressione di un potere di tipo privatistico, viene integrato nel suo contenuto, allo scopo di tutelare un bene di rango superiore, l'ordine giuridico del mercato, nell'ambito del quale la tutela della concorrenza si coniuga (incidentalmente) con la protezione dei diritti fondamentali della persona.

<sup>252</sup> F. Oliveri, *Diritti degli internauti, obblighi degli Stati, responsabilità delle piattaforme digitali*, cit., 116.

<sup>253</sup> M. Ferraris, *Metafisica del Web*, lezione tenuta presso il Centro Nexa, Politecnico di Torino, in *nexa.polito.it*, 08.01.2020,

conforme è proporzionale al potere acquisito dalle grandi società informatiche nel plasmare l'ecosistema web, potere che richiede parimenti di essere democratizzato. Di riflesso, il diritto a essere protetti dai discorsi d'odio online deve trovare fondamento all'interno della nuova democrazia digitale e deve costituire oggetto di bilanciamento con la libertà di informazione, di espressione e di comunicazione, nonché con i diritti posti a presidio dell'ordinamento democratico plurale e multiculturale che caratterizzano la fisionomia di ogni Stato di diritto.

La stessa Corte di Strasburgo per oltre vent'anni ha seguito un approccio definito "militante"<sup>254</sup> che, mediante l'adozione della tecnica decisionale tipica dei Paesi di *Common Law*, ossia il metodo *case by case*<sup>255</sup>, prevede un bilanciamento quanto più adeguato e effettivo tra la libertà di espressione e la tutela del diritto alla dignità e a non subire discriminazioni. In particolare, la Corte EDU ha tenuto conto in maniera omnicomprensiva del materiale illecito commesso, in piena ottemperanza ai principi di offensività e di materialità del diritto penale, cristallizzati nel nostro quadro positivo all'art. 25, c. 2, Cost. In tale ottica, il ricorso allo strumento repressivo forte appare giustificato al fine di presidiare i fondamentali valori democratici dell'uguaglianza e del pluralismo e deve essere finalizzato a punire qualsiasi condotta che - fermi restando gli essenziali principi sopra citati - sia in grado di arrecare pregiudizio a tali valori. I giudici di Strasburgo sono giunti, così, ad attribuire "carattere eccezionale" ai discorsi d'odio, in forza del quale il vaglio di legittimità delle misure restrittive della libertà di espressione assume maglie più larghe<sup>256</sup>. Di qui, seguendo una linea di severo contrasto del fenomeno, la Corte EDU ammette il ricorso al diritto penale per contrastare l'*online hate speech*.

Appare fuori discussione come, a tutt'oggi, spazi di condivisione quali i social network si concretizzino in luoghi di esercizio del diritto di manifestare liberamente il proprio pensiero<sup>257</sup>, sancito dall'art. 10 CEDU (che, al par. 2, ne disciplina le eventuali restrizioni), dagli artt. 11 e 21 della Carta dei diritti fondamentali dell'UE (che la sancisce, nello specifico, nella accezione relativa al principio di non discriminazione)<sup>258</sup> e, non da ultimo, dall'art. 21 della Costituzione italiana.

Anche la nostra Corte costituzionale ha affrontato il tema della repressione penale delle c.d. "parole pericolose"<sup>259</sup>, che sono state storicamente inquadrate nell'ambito dei reati posti a tutela dell'"ordine pubblico", come, a titolo esemplificativo, le norme volte al contrasto delle condotte discriminatorie<sup>260</sup>. Il nostro legislatore ha tentato di

<sup>254</sup> C. Caruso, *L'hate speech a Strasburgo: il pluralismo militante del sistema convenzionale*, in *Quaderni costituzionali*, 37, 4, 2017, 963-984.

<sup>255</sup> Su questo aspetto, cfr. V. Cinà, *Libertà di espressione e importanza del contesto: la Corte europea dei diritti dell'uomo ridefinisce il perimetro della protesta politica*, in *La Nuova giurisprudenza civile commentata*, 6, 2021, 1379 ss.

<sup>256</sup> P. Dunn, *Carattere eccezionale dell'"hate speech" e nuove forme di responsabilità per contenuti di terzi nella giurisprudenza EDU*, cit., 249.

<sup>257</sup> C. Confortini, *Diffamazione e discorso d'odio in internet*, cit., 698.

<sup>258</sup> P. Falletta, *Analisi normativa in tema di contrasto agli hate speech su Internet e social media*, cit., 29.

<sup>259</sup> Fin dalla sentenza della Corte cost., 26 gennaio 1957, n. 1 avente ad oggetto la previsione della l. 645 del 1952, *Norme di attuazione della XII disposizione transitoria e finale della Costituzione*, che incrimina le condotte di apologia di fascismo.

<sup>260</sup> Vedasi, per un maggiore approfondimento, G. Puglisi, *La parola acuminata. Contributo allo studio dei*

condizionare l'esercizio della libertà di manifestazione del pensiero sotto lo spettro del bene giuridico "ordine pubblico", a cui è stata inizialmente ispirata l'introduzione di fattispecie che incriminano forme di incitamento e di istigazione a commettere reati. È questo il caso della diffusione di ideologie fondate sull'incitamento all'odio razziale che hanno posto la questione dell'idoneità della libertà di pensiero a impedire la repressione di manifestazioni aggressive di dissenso ideologico idonee a negare i valori di fondo della convivenza civile<sup>261</sup>.

Merita di essere menzionata la sopravvivenza dei c.d. "reati di opinione", contenuti nel Codice Rocco e concepiti dalla Consulta sotto forma di limiti esterni all'art. 21 Cost., che hanno reso necessaria l'enucleazione di una serie di principi per verificare la conformità a Costituzione degli strumenti posti a tutela di una democrazia minacciata da linguaggi discriminatori<sup>262</sup>. L'opera della Corte costituzionale, mossa verso la salvaguardia dell'impianto repressivo del dissenso, ha individuato gli argini irrinunciabili della libertà di espressione, sia in termini di contenuti del messaggio informativo, sia di diffusione dello stesso attraverso qualsiasi mezzo<sup>263</sup>.

Il passaggio a un "Internet delle piattaforme" ha fatto sì che, se, in un primo momento, i social media sono stati concepiti come spazio di agevolazione del dibattito pubblico<sup>264</sup>, essi sono, a tutt'oggi, assurti al rango di "pietra angolare dell'ordine democratico"<sup>265</sup>.

Pertanto, sebbene, originariamente, la repressione dei discorsi d'odio sia stata giustificata in forza della nozione di "ordine pubblico", attualmente appare maggiormente corretto inquadrarli nel novero delle componenti atte a garantire la solidità della democrazia nella dimensione digitale (forse, ormai, si potrebbe dire, nell'era "post" digitale). A conferma di ciò, il parametro utilizzato dalla Consulta è divenuto il più elastico principio del pluralismo, inteso quale distinta e ulteriore libertà che trova specifica cristallizzazione nell'art. 21 Cost. e che garantisce, altresì, il diritto ad «esteriorizzare liberamente il proprio pensiero con la parola, lo scritto e ogni altro mezzo di diffusione»<sup>266</sup>. Peraltro, tale diritto assolve non soltanto un'esigenza di carattere individuale, ma anche una vera e propria "funzione sociale", assicurando alla collettività il contributo del pensiero di tutti i consociati, oltre che la libera discussione e il confronto tra idee, perseguendo interessi pubblici e privati diversi da quelli del solo titolare<sup>267</sup>.

---

*delitti contro l'eguaglianza, tra aporie strutturali ed alternative alla pena detentiva*, in *Rivista italiana di diritto e procedura penale*, 2018, 1329 ss.

<sup>261</sup> L. Califano, *La libertà di manifestazione del pensiero...in rete; nuove frontiere di esercizio di un diritto antico*, cit., 10-11.

<sup>262</sup> *Ibid.*

<sup>263</sup> *Ivi*, 3.

<sup>264</sup> P. Falletta, *Analisi normativa in tema di contrasto agli hate speech su Internet e i social media*, cit., 23.

<sup>265</sup> Così Corte cost., 17 aprile 1969, n. 84 e 04 febbraio 1965, n. 9. Per una definizione analoga della libertà di espressione, inserita «fra i diritti primari e fondamentali» nella giurisprudenza costituzionale italiana vedasi Corte cost. 28 gennaio 1981, n. 1 che la definisce «cardine del regime democratico».

<sup>266</sup> Così, L. Califano, *La libertà di manifestazione del pensiero...in rete; nuove frontiere di esercizio di un diritto antico*, cit., 5, richiamando i seguenti precedenti giurisprudenziali: Corte cost., 14 luglio 1988, n. 826, 07 dicembre 1994, n. 420 e 18 ottobre 2002, n. 422.

<sup>267</sup> L. Califano, *La libertà di manifestazione del pensiero...in rete; nuove frontiere di esercizio di un diritto antico*, cit., 8.

Di talché, i sostenitori della necessità del diritto di cittadinanza dei “reati di opinione” nel nostro ordinamento valorizzano la protezione fornita a beni costituzionalmente rilevanti da forme di tutela anticipata, vista la pericolosità insita nell’*hate speech* di compromettere i processi di integrazione dei valori fondanti la convivenza civile<sup>268</sup>.

In senso parzialmente difforme, la Corte EDU ha per molto tempo richiamato la “funzione didattico-terapeutica”<sup>269</sup> che può essere dispiegata dalla pubblica manifestazione del pensiero, seguendo l’approccio della dottrina statunitense che ammette la legittimità di ogni manifestazione, anche dichiaratamente espressiva di odio razziale o etnico, a condizione che non sia tale da produrre un pericolo chiaro ed imminente che giustifichi l’intervento repressivo dello Stato<sup>270</sup>. Tale logica non è molto diversa da quella seguita dalla Corte costituzionale italiana e dalle successive pronunce dei tribunali di merito nell’opera di adeguamento della portata applicativa delle diverse ipotesi di istigazione, che sono state limitate laddove, per eccessiva ampiezza e genericità, si potesse dire consentito incriminare condotte tutelate dall’art. 21 Cost., muovendo dalla distinzione fra ciò che è “pensiero puro” e ciò che, invece, è “principio in azione”<sup>271</sup>. Così, ad esempio, la Corte ha rilevato la contrarietà all’art. 21 dell’istigazione alla disobbedienza delle leggi di ordine pubblico e dell’odio fra le classi sociali di cui all’art. 415 c.p., nella parte in cui quest’ultima non specificava che l’istigazione, per essere punita, oltre che una pubblica esternazione, deve comportare un concreto pericolo per la pubblica incolumità<sup>272</sup>. Merita una menzione il fatto che, in relazione al primo reato citato, la Consulta abbia ravvisato anche la violazione dell’art. 25, c. 2, Cost. Ciò in quanto la fattispecie è indeterminata e prescinde da un qualsiasi esame circa l’effettivo verificarsi di un pericolo per l’interesse tutelato, non precisando le modalità con cui l’istigazione si potrebbe distinguere dalla diffusione della persuasione di ideologie e dottrine politiche, sociali, filosofiche o economiche, e, quindi, penalmente perseguibile, senza violare il precetto costituzionale dell’art. 21 Cost. Ne deriva che la Consulta, nel solco del c.d. “diritto penale materiale”<sup>273</sup>, ha attribuito rilievo alle concrete e specifiche modalità con le quali è stata compiuta l’istigazione, allo scopo di indagare laddove la stessa rivesta carattere di effettiva pericolosità per l’esistenza di beni costituzionalmente protetti e integri un comportamento concretamente idoneo a promuovere discorsi d’odio. Sol-

<sup>268</sup> M. Manetti, *L’incitamento all’odio razziale tra realizzazione dell’eguaglianza e difesa dello Stato*, in *Scritti in onore di Gianni Ferrara*, Padova, 2006, 103 ss.

<sup>269</sup> L. Califano, *La libertà di manifestazione del pensiero...in rete; nuove frontiere di esercizio di un diritto antico*, cit., 11.

<sup>270</sup> Vedi, sul punto, considerazioni svolte nel paragrafo 5.

<sup>271</sup> L. Califano, *La libertà di manifestazione del pensiero...in rete; nuove frontiere di esercizio di un diritto antico*, cit., 11 e 12 che richiama Corte cost. 16 marzo 1962, n. 19 nella parte in cui afferma che «se per turbamento dell’ordine pubblico bisogna intendere l’insorgere di un concreto ed effettivo stato di minaccia per l’ordine legale, mediante mezzi illegali idonei a scuoterlo (...) è chiaro che non possono essere considerate in contrasto con la Costituzione le disposizioni legislative che effettivamente ed in modo proporzionato siano volte a prevenire e reprimere siffatti turbamenti».

<sup>272</sup> Corte cost., 5 aprile 1964, n. 108 e 5 giugno 1978, n. 71.

<sup>273</sup> Del quale si può parlare, all’interno del nostro ordinamento penalistico, in forza del divieto di punire l’intenzione (*cogitationis poenam nemo patitur*), cristallizzato all’art. 25, c. 2, Cost.; sul punto, si riportano di seguito gli orientamenti della giurisprudenza costituzionale maggiormente rilevanti: Corte cost., 02 novembre 1996, n. 370 sull’art. 708 c.p. e 05 luglio 2010, n. 249 sulla qualità di immigrato «irregolare».

tanto quando l'istigazione è diretta a commettere gli atti che costituiscono – secondo una valutazione legislativa definita immune da irragionevolezza – un pericolo per il bene costituzionalmente protetto e non si limita a una mera critica di fatti specifici, per i quali si esercita democraticamente il controllo dell'opinione pubblica, tali condotte assumeranno rilevanza penale<sup>274</sup>.

In sintesi, un diritto penale costituzionalmente orientato impone l'incriminazione dei discorsi d'odio nella misura in cui questi ultimi possano configurarsi come la premessa di un'azione delittuosa<sup>275</sup> volta a ledere la collettività di appartenenza del destinatario, compromettendone la coesione e pregiudicandone l'ordine democratico eterogeneo.

## **8. Riflessioni conclusive: il ricorso al diritto penale, utopia od opportunità?**

Alla luce delle considerazioni di cui sopra, pare irrinunciabile svolgere alcune riflessioni conclusive circa il quadro tracciato in punto di rilevanza penale dei discorsi d'odio online.

Le condotte d'odio assumono nell'ecosistema digitale, una dimensione transfrontaliera e transnazionale che amplifica enormemente gli effetti pregiudizievoli del messaggio illecito e rende estremamente difficoltosa l'adozione di efficaci strumenti per contrastarlo. Di conseguenza, appare ancora più impellente la necessità che l'Unione europea adotti direttive in questa materia, atteso che tali condotte, in ragione della loro portata, sono in grado di assumere un elevatissimo grado di lesività per i beni giuridici coinvolti. Ne discende che i discorsi d'odio online possono essere ricondotti alla categoria della "criminalità informatica" di cui all'art. 83, par. 1, TFUE, in riferimento alla quale le istituzioni europee detengono una competenza legislativa "diretta". A ciò si aggiunga che la Commissione ha a lungo esaminato la possibilità di punire tali condotte ricorrendo alla competenza "indiretta" in materia penale dell'Unione europea, ai sensi dell'art. 83, par. 2, TFUE.

Tuttavia, nessuna di queste soluzioni di intervento legislativo è stata intrapresa e, allo stato attuale, l'inerzia del legislatore europeo nell'incriminare a livello sovranazionale questi comportamenti non è più tollerabile. L'esigenza che si impone è quella di realizzare il tanto auspicato quadro unionale di tutela giuridica, anche ai fini del ravvicinamento dei quadri degli Stati membri. L'adozione di un efficace sistema di repressione non può prescindere da un'opera di preparazione socioculturale che miri a far comprendere il grado di pericolosità insito nelle condotte in esame e che consenta di coniugare l'utilizzo della tecnica penalistica con il formante tecnologico e con quello sociale, al fine di radicare la cultura di contrasto ai discorsi d'odio online. È, quindi, necessario uno sforzo coordinato tra queste componenti per contenerne la circolazione e mitigare i fenomeni pregiudizievoli che ne derivano. Diventa, quindi, essenziale promuovere la consapevolezza e l'educazione digitale di tutte le fasce della popolazione. A tal fine, la politica legislativa, sia a livello nazionale che europeo, riveste un ruolo di primaria

---

<sup>274</sup> Corte cost., 5 giugno 1978, n. 71.

<sup>275</sup> L. Califano, *La libertà di manifestazione del pensiero*, cit., 12 e 13.

importanza.

Avendo riguardo al quadro normativo italiano, le fattispecie analizzate non sono in grado di stigmatizzare in modo efficace il diffondersi dei discorsi d'odio online. Tuttavia, lo strumento del diritto penale, impiegato nel doveroso rispetto dei principi di tassatività, precisione e determinatezza, corollari ineludibili del principio di legalità, continua a rappresentare un baluardo irrinunciabile per punire quei comportamenti che ledono beni giuridici fondamentali. In questo contesto, la scelta di prevedere l'incriminazione dei discorsi d'odio (anche online) mediante la tecnica incriminatrice dei reati a pericolo concreto, consente di evitare il rischio di un possibile conflitto con il principio di offensività, il quale potrebbe essere compromesso qualora si applicasse una sanzione penale a una condotta concretamente inoffensiva, per i mezzi, le modalità e i destinatari. In un ambito così sensibile, in cui è necessario bilanciare accuratamente i beni giuridici coinvolti, inquadrare il reato tra quelli a pericolo concreto contribuisce a salvaguardare il rispetto della Costituzione.

Nello specifico, l'opera della Consulta si è incentrata sulla salvaguardia dell'impianto repressivo del dissenso, al fine di preservare i beni giuridici del pluralismo, dell'uguaglianza, della dignità umana e della democrazia. Analogamente a quanto riscontrato negli ulteriori quadri giuridici analizzati, anche la nostra Corte costituzionale si è occupata di indagare la linea di demarcazione tra l'espressione di pensiero puro e una manifestazione che, al contrario, racchiude in sé non una mera divulgazione, bensì una condotta in azione, idonea a porre in pericolo l'oggettività giuridica tutelata. Soltanto quest'ultima condotta si connota per una duplicità di motivazioni e di destinatari: non è rivolta all'individuo singolarmente inteso, ma piuttosto alla collettività di appartenenza che viene colpita tramite il singolo.

Altresì dallo studio dei sistemi europei analizzati è emerso il favore verso l'utilizzo dello strumento del diritto penale, a discapito di altri possibili, allo scopo precipuo di punire specificamente le condotte di *hate speech*, anche nella dimensione online. L'elevata gravità di queste condotte, potenzialmente idonee a compromettere il valore super-individuale su cui si regge la collettività, ha indotto a far ritenere preferibile la via dell'utilizzo di strumenti eccezionali di tutela avanzata.

Tuttavia, nei Paesi che hanno effettuato scelte di incriminazione, estremamente insidiosa è apparsa la strutturazione dell'ecosistema penale all'interno della complessa realtà digitale. Ciò in quanto si è ravvisata la contestuale necessità di garantire il rispetto dei principi di matrice costituzionale che sorreggono la tecnica penalistica. Primo tra tutti viene in discussione il principio di offensività, atteso che, nella selezione dei fatti penalmente rilevanti, sia il legislatore che il giudice, devono punire le sole condotte idonee a ledere, ovvero porre in pericolo, i beni giuridici tutelati. Viene poi in rilievo il principio di legalità, di cui all'art. 25, comma 2, Cost., che si declina (per quanto più di interesse) nel principio di riserva di legge e di determinatezza, tassatività e precisione. Non da ultimo, occorre tenere in considerazione anche il principio di materialità del diritto penale, il cui contenuto massimo si sostanzia nel brocardo latino "cogitationis poenam nemo patitur", in virtù del quale nessun individuo può essere punito per aver concepito nella sola forma del pensiero la commissione di un fatto criminoso che non si sia tradotto in azione, ovvero non sia stato messo in atto.

Nel quadro giuridico comparato è emerso che la complessità della disciplina dettata e l'indeterminatezza delle fattispecie incriminatrici introdotte, concepite principalmente secondo le coordinate del modello istigatorio, hanno condotto, nella prassi, alla mancata applicazione di queste ultime a vantaggio delle fattispecie comuni. È quanto è accaduto in Germania e Spagna. Si può, quindi, rilevare che non sempre a uno sforzo legislativo sistematico o, finanche, definito di carattere “olistico”, corrisponda l'efficacia della disciplina, la cui eccessiva forza espansiva compromette l'intelligibilità del sistema, al pari di quanto verificatosi in Francia.

Da ultimo, ai fini dell'accertamento della responsabilità degli autori, la giurisprudenza francese, tedesca e spagnola rimanda a un vaglio di proporzionalità, necessità e adeguatezza da eseguirsi nel caso specifico, in cui si prendano in considerazione sia i diritti fondamentali che vengono in rilievo, sia il rispetto dei principi che permeano la tecnica penalistica e la cui operatività deriva anche dall'effetto di irradiazione dei primi. Anche il cammino della Corte di Giustizia dell'Unione europea e della Corte di Strasburgo, nella faticosa opera di individuazione degli argini irrinunciabili della libertà di espressione, e la disamina offerta dal quadro normativo americano hanno condotto a esiti analoghi a quelli della Consulta italiana. Il diritto penale deve, necessariamente, continuare a essere utilizzato nel contrasto ai discorsi d'odio online come strumento principale, seppur integrato nell'ambito di una strategia complessa che passi anche attraverso la prevenzione.

Da questo punto di vista, la moderazione algoritmica realizzata dai PSI si presenta come uno strumento volto a realizzare un evidente vantaggio sociale, determinato dalla riduzione dell'inquinamento dell'ecosistema informazionale digitale. Come visto, ciò che deve essere necessariamente tenuto in debita considerazione è la fallibilità dei meccanismi di moderazione: un certo margine di errore è purtroppo ineludibile. Il rischio è rappresentato dal fatto che la moderazione attuata dalle piattaforme si riveli un silenziamento delle categorie già discriminate. Per evitare di tradire la *ratio* che ne giustifica l'utilizzo, è necessario pretendere una soglia di accettabilità dell'errore più esigente. Elaborare una strategia efficace di contrasto al fenomeno dell'odio significa necessariamente mettere in relazione le coordinate ermeneutiche del diritto e della tecnologia, alla luce del contesto sociale in cui ci si trova. Al riguardo, una soluzione valida oggi potrebbe non esserlo tra alcuni anni. Sarà compito primariamente del formante politico-legislativo e di quello giurisprudenziale individuare un equilibrio nell'ideazione e nell'interpretazione delle norme penali, al fine di preservare la flessibilità necessaria a fronteggiare temi come quello in esame, soggetto a mutamenti sempre più repentini. In conclusione, nell'era delle democrazie digitali, emerge con chiara evidenza una c.d. “eterogenesi dei fini”: se da un lato è indubbio che le piattaforme costituiscono una componente dell'infrastruttura democratica e favoriscono la libertà di manifestazione del pensiero, quest'ultima, nata come libertà da possibili ingerenze del potere statale, è assurta al rango di «diritto coesistente al regime di libertà garantito dalla Costituzione»<sup>276</sup> e subisce i rischi connessi alle mutate modalità di diffusione delle idee tipiche della modernità tecnologica.

A fronte di ciò, seppure non si possa negare la “funzione didattico-terapeutica” svolta

---

<sup>276</sup> Corte cost., 23 marzo 1968, n. 11.

dalla censurabilità etica del web, occorre considerare che la Corte di Strasburgo, al pari della nostra Consulta, ha rilevato come il ricorso allo strumento pesante del diritto penale possa dirsi giustificato per la tutela di fondamentali valori democratici. Ciò in quanto i discorsi d'odio, in generale e ancor di più quelli che si dispiegano online, presentano un *quid pluris* rispetto alle altre condotte criminose, idoneo a compromettere i valori fondamentali dell'agire democratico plurale nell'era digitale.

La rete è divenuta il «centro di un grande motore di democrazia»<sup>277</sup> e, in tale contesto, l'*online hate speech* riveste uno *status* peculiare, in quanto la gravità e la pericolosità insite in esso per i valori fondamentali della convivenza civile costituiscono di per sé un giustificativo per il ricorso alla tecnica penalistica. Questa è l'unica in grado di difendere il fulcro immutabile di norme fondamentali garantiste su cui si regge il nostro sistema costituzionale, secondo lo schema del «diritto penale minimo», che persegue il fine di razionalizzare e minimizzare la violenza dell'intervento punitivo, vincolandolo a limiti rigidi imposti a tutela dei diritti della persona<sup>278</sup>.

A ciò dovrebbe senz'altro aggiungersi l'adozione di un approccio chirurgico<sup>279</sup> nella progettazione delle fattispecie incriminatrici e nella loro interpretazione, finalizzato a garantire il rispetto dei canoni di legalità, offensività e materialità che permeano l'ordinamento penalistico.

Soltanto un diritto penale costituzionalmente orientato può garantire la repressione di condotte effettivamente e gravemente pregiudizievoli, non soltanto della dignità della vittima, bensì, anche della sopravvivenza della democrazia eterogenea e multiculturale cui essa appartiene. Del resto, «sostituire il diritto penale con qualcosa di meglio potrà avvenire solo quando avremo sostituito la nostra società con una società migliore»<sup>280</sup>.

---

<sup>277</sup> L. Califano, *La libertà di manifestazione del pensiero*, cit., 6.

<sup>278</sup> L. Ferrajoli, *Cos'è il garantismo*, in *Criminalia*, 2014, 130.

<sup>279</sup> G. Ziccardi, *Il contrasto dell'odio online: possibili rimedi*, cit., 38.

<sup>280</sup> A. Baratta, *Criminologia critica e critica del diritto penale: introduzione alla sociologia giuridico-penale*, Milano, 2019, 1982.