



Online Hate Speech

Il peso delle parole d'odio

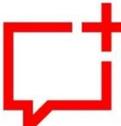
Beatrice Aprile

**Centro
Editoriale**
Accademico
unicollege

**Centro
Editoriale
Accademico
unicollege**

unicollege
Knowledge
and Experience.

 **International
Council®**

 **Accademia
di Italiano®**

 **Adiuva®**

**Centro
Editoriale
Accademico**
unicollege

Via Bolognese 52

50139 Firenze

[https://www.unicollegessml.it/centro-editoriale-accademico/
centro.editoriale.accademico@unicollegessml.it](https://www.unicollegessml.it/centro-editoriale-accademico/centro.editoriale.accademico@unicollegessml.it)

Scientia Unicollege

Collana diretta da Lorenzo Grifone Baglioni

- 1 *Online Hate Speech*, Beatrice Aprile.

How to cite this book / Come citare questo libro:

Aprile B. (2025), *Online Hate Speech. Il peso delle parole d'odio*, Centro Editoriale Accademico, Firenze.

Online Hate Speech

Il peso delle parole d'odio

Beatrice Aprile

Centro Editoriale Accademico - Firenze

ANNO 2025 - ISBN 9791298507029

Indice

- 11 *Introduzione*
- 13 *Hate Speech: dagli atomi ai bit*
 - 1.1 Verso una definizione di Hate Speech
 - 1.2 La regolamentazione dell'odio
 - 1.3 La tutela della libertà di espressione del pensiero
 - 1.4 Gli accordi europei e la Corte EDU
 - 1.5 La Corte costituzionale italiana
 - 1.6 Gli interventi americani
 - 1.7 Conclusioni
- 85 *Rimedi: censura o counterspeech*
 - 2.1 Il Principio di Neutralità
 - 2.2 Il bene Dignità
 - 2.3 Il Principio di Garanzia
 - 2.4 L'autonomia del singolo
 - 2.5 Autogoverno e legittimità politica
 - 2.6 The Marketplace of Ideas
 - 2.7 Counterspeech
 - 2.8 Conclusioni
- 127 *Online Hate Speech e AI*
 - 3.1 Internet e Online Hate Speech
 - 3.2 Il Principio di Neutralità della Rete
 - 3.3 Social Networks e responsabilità degli ISP
 - 3.4 AI nei Social Networks: i possibili rischi
 - 3.5 Conclusioni

189	<i>Il cyberbullismo</i>
	4.1 Inquadramento della fattispecie
	4.2 Dati alla mano
	4.3 Testimonianze
	4.4 Conclusioni
203	<i>Bibliografia</i>
215	<i>Sitografia</i>

A Benedetta e Andrea

*Perché, anche se hanno tentato di spegnere la vostra
Luce, chi è luminoso dentro splenderà sempre*

A nonno Piero

*Perché, se oggi sono diventata la donna che sono,
gran parte del merito è Tuo.
I tuoi insegnamenti rimarranno dentro di me,
per sempre.*

Introduzione

L'obiettivo di questa riflessione è analizzare in profondità l'espressione *Hate Speech* e, in ossequio al principio di salvaguardia dei diritti fondamentali dell'individuo, sottolineare il pericolo concreto e attuale legato alla sua pratica.

In prima istanza, si provvede ad analizzare le fonti che regolano il fenomeno, rispettivamente negli USA e in Europa, concentrando poi prettamente l'attenzione sulle fonti nazionali e sulla giurisprudenza che si è espressa in merito nel corso del tempo.

In seconda battuta, dallo studio delle radici storiche e filosofiche poste alla base del discorso d'odio, si procede ad individuare due fondamentali e opposte correnti di pensiero. Da queste ultime, scaturiscono i rimedi (giuridici e non, si pensi all'impiego degli strumenti di AI) approntati dai singoli Stati per regolamentare, se ritenuto necessario, tale fenomeno. Se da un lato, infatti, domina imperante il principio del *Free Speech* che gode della protezione ad esso garantitagli dal Primo Emendamento della Costituzione Americana, dall'altro lato l'Europa sceglie di adottare la cosiddetta *Censorship*, ossia una politica di censura volta a sopprimere qualsiasi discorso che si prefigga

l'obiettivo di diffondere espressioni discriminatorie, mettendo così in pericolo beni ritenuti meritevoli di tutela da parte degli ordinamenti giuridici dei singoli Stati.

Infine, si intende fornire uno spunto di riflessione sull'odio online in quanto tema interessante ai fini dei suoi effetti soprattutto sui nostri giovani ragazzi. In merito, si è provveduto a raccogliere le testimonianze di giovani di oggi e giovani di ieri per vedere come quello che un tempo prendeva il nome di bullismo, oggi, a causa della Rete, abbia raggiunto una portata tale da essere diventato incontenibile assumendo nuovi e pericolosi caratteri.

Capitolo Primo

Hate speech: reale o virtuale?

1.1 Verso una definizione di Hate Speech

Per capire come risolvere l'annoso (e ormai evidente) problema dei cosiddetti discorsi d'odio, occorre preliminarmente tracciarne i confini. Un primo intervento da segnalare in tal senso è quello della Commissione Jo Cox istituita nel 2017 in Italia presso la Camera dei deputati. Qui la Commissione ha definito l'hate speech come

l'istigazione, la promozione o l'incitamento alla denigrazione, all'odio o alla diffamazione nei confronti di una persona o di un gruppo di persone, o il fatto di sottoporre a soprusi, molestie, insulti, stereotipi negativi, stigmatizzazione o minacce tale persona o gruppo, e comprende la giustificazione di queste varie forme di espressione, fondata su una serie di motivi, quali la razza, il colore, la lingua, la religione o le convinzioni, la nazionalità o l'origine nazionale o etnica, nonché l'ascendenza, l'età, la disabilità, il sesso, l'identità di genere, l'orientamento sessuale e ogni altra caratteristica o situazione personale.

Lo scopo di tale intervento è quello di sottolineare la gravità del fenomeno del discorso d'odio veicolato attraverso la rete Internet. Si tratta di una

definizione particolarmente ampia, volta a ricomprendere non solo fenomeni di stampo meramente diffamatorio ma altresì tutte quelle espressioni d'odio che vengono usate e diffuse nel mondo virtuale. Si può infatti parlare di una vera e propria piramide d'odio che, in quanto tale, si sviluppa su diversi livelli¹. Alla base vi sono gli stereotipi, il linguaggio ostile “banalizzato”, le rappresentazioni non veritiere e gli insulti. Ai livelli superiori invece, troviamo le discriminazioni e quindi il linguaggio d'odio costituito da minacce e incitamenti alla violenza contro le minoranze. Infine, i crimini d'odio (veri e propri atti di violenza fisica, costituenti reato, compiuti contro persone in base a caratteristiche come il sesso, l'orientamento sessuale, l'etnia, la religione o altro).

In realtà, già nel 1954, Gordon Allport, psicologo sociale statunitense, elabora una scala del pregiudizio e della discriminazione, al fine di classificare i livelli d'odio presenti nella società. Egli suddivide i discorsi d'odio in cinque categorie a seconda della loro intensità. Il primo gruppo è definito Antilocution e si riferisce a quel linguaggio stereotipato utilizzato dai componenti di un gruppo di maggioranza al fine di dileggiare gruppi

¹<https://www.medialaws.eu/wp-content/uploads/2019/05/1.-Pitruzzella.pdf> (8.2.2025)

di minoranza. Il secondo livello di gravità viene definito Avoidance e viene raggiunto da quelle persone che volontariamente ne evitano altre costringendo così all'isolamento coloro che vengono considerati diversi. Nel terzo livello, la Discrimination, trovano spazio le discriminazioni volte a creare disparità di trattamento tra il gruppo di maggioranza e le minoranze, impedendo a quest'ultime di poter avere le stesse opportunità dei soggetti non discriminati. L'aggressione fisica, il cosiddetto Physical Attack, costituisce la quarta fase. In questo caso, si parla di veri e propri hate crimes. Ne costituiscono esempio il linciaggio delle persone di colore o le violenze contro la comunità LGBTQIA+. L'apice si raggiunge mediante la cosiddetta Extermination, che ha luogo quando il gruppo di maggioranza vuole eliminare l'esistenza stessa degli avversari. Gli esempi più noti sono il genocidio nazista nei confronti degli ebrei, i crimini di guerra commessi contro i nativi americani, il genocidio in Ruanda e gli episodi di pulizia etnica in Bosnia e durante le guerre nell'ex-Jugoslavia (Allport 1954).

Ma se si pensa di poter dare una definizione compiuta ed esaustiva di hate speech, si cade in errore. Infatti, ancora oggi, nonostante i numerosi interventi in merito da parte della dottrina, della

giurisprudenza e del legislatore, non esiste un concetto di discorso d'odio unanimemente condiviso. A titolo esemplificativo, se guardiamo al report *Hate Crime and Hate Speech in Europe: Comprehensive Analysis of International Law Principles, EU-wide Study and National Assessments*, prodotto nell'ambito del progetto PRISM (Preventing, Repressing & Inhibiting Hate Speech in New Media² finanziato dalla Commissione Europea, si parla di hate speech in termini di

denigration of the reputation of a social group, stereotyped by some particular national, racial or religious characteristics, accompanied by incitement to hostility, violence and discrimination against that group.

In questo caso, vediamo come l'elemento che rileva ai fini della definizione di hate speech sia la lesione alla reputazione di un gruppo sociale. Al contrario, Giovanni Ziccardi (Ziccardi 2016) afferma che per aversi hate speech occorre che l'espressione utilizzata manifesti un incitamento volontario e idoneo a causare atti d'odio, di discriminazione o violenza. Rileva qui l'elemento della pericolosità concreta del linguaggio stesso. Altri

²https://migrant-integration.ec.europa.eu/library-document/hate-crime-and-hate-speech-europe-comprehensive-analysis-international-law_en (12/08/2024).

(Pino 2008, 293-294) invece ritengono di poter definire hate speech un discorso finalizzato a promuovere odio nei confronti di certi individui o gruppi, impiegando epiteti che denotano disprezzo nei loro confronti a causa della loro connotazione razziale, etnica, religiosa, culturale o di genere. Ritengono quindi necessaria la presenza del dolo intenzionale.

Anche in territorio statunitense si fatica a trovare un concetto compiuto e definito di discorso d'odio. Diversamente dal pensiero europeo però, molti autori privilegiano l'elemento del danno al fine di delineare i confini di tale concetto. Dal canto suo, Susan Benesch focalizza la propria attenzione sull'analisi di una sottocategoria di hate speech, il cosiddetto *dangerous speech*, individuando così le variabili da ritenere rilevanti per stabilire la danosità del discorso d'odio³. Occorre dunque un locutore potente in grado di esercitare un'influenza ad alto livello, un pubblico recettivo che sia animato da rimostranze e paure che il locutore può coltivare, un contesto sociale o storico che si predisponga alla violenza ed infine mezzi di diffusione influenti. Anche Jeremy Waldron (2012), nel tentativo di stabilire i confini del discorso d'odio, insiste

³ <https://www.dangerousspeech.org/libraries/guide> (12/08/2024).

sul danno insito nell'hate speech e provocato – suggerisce quest'ultimo – non tanto o, meglio, non solo dall'odio contenuto nel messaggio ma dagli strumenti di pubblicazione del messaggio stesso. Egli infatti parla di

danno causato agli individui e ai gruppi mediante la trasfigurazione del nostro ambiente sociale per mezzo di dichiarazioni visibili, pubbliche e semipermanenti, finalizzate a non considerare come equi cittadini i membri dei gruppi di minoranza.

Nel classificare il discorso d'odio come dannoso, Andrew Sellars⁴ sostiene che è altresì necessario guardare al contesto in cui questo si inserisce e alla percezione che il destinatario del discorso ha in relazione all'espressione utilizzata nei suoi confronti. Su quest'ultimo punto troviamo una corrispondenza con quanto sostenuto da Gianmarco Gometz⁵, il quale ritiene che alcune espressioni vengano fatte rientrare nel discorso d'odio in quanto recepite come tali dai destinatari del discorso, indipendentemente dall'intenzione del locutore. Egli, quindi, parla di discorso non intenzionalmente diretto a provocare l'odio ma idoneo a

⁴https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2882244 (7.2.2025)

⁵https://d1vbhhq6ow083.cloudfront.net/contributi/Go-metz.M_Lodio.pdf (6.2.2025)

produrlo. È questo il caso dell'art. 8, secondo comma, della legge italiana n. 962 del 1967, che punisce con la reclusione da tre a dodici anni chiunque pubblicamente istighi o faccia l'apologia di: atti diretti a commettere genocidio, deportazione a fini di genocidio, morte per genocidio o deportazione, atti diretti a commettere genocidio mediante limitazione delle nascite e atti diretti a commettere genocidio mediante sottrazione di minori.

Ecco, quindi, che in tutte le definizioni che, a titolo esemplificativo, abbiamo fornito in relazione al discorso d'odio, possiamo notare come emerge imperante un requisito ritenuto essenziale perché si possa parlare di hate speech: l'ambiente all'interno del quale il messaggio d'odio si diffonde. Lo stesso Ziccardi (2016) distingue tra espressione d'odio utilizzata all'interno di un contesto informale e colloquiale quale può essere un bar rispetto ad un ambiente virtuale quale può essere un social network. Ma prima di capire come Internet abbia inciso sul linguaggio d'odio, cambiane per sempre la portata e gli effetti, è necessario analizzare l'approccio che il legislatore ha avuto nel corso del tempo rispetto al fenomeno odio.

1.2 La regolamentazione dell'odio

Nel 1951 entra in vigore la Convenzione sulla prevenzione e la punizione del crimine di genocidio, con lo scopo di proteggere i gruppi definiti in base alla razza, all'etnia o alla nazionalità. Con particolare riferimento al discorso d'odio, essa punisce solo chi realizza atti che pubblicamente incitano all'odio, intesi come "atti commessi con l'intenzione di distruggere, in tutto o in parte, un gruppo nazionale, etnico, razziale o religioso".

Nel 1969 viene stipulata la cosiddetta ICERD (The International Convention on the Elimination of All Forms of Racial Discrimination), la quale prevede che gli Stati parte si impegnino per adottare misure finalizzate ad eliminare ogni forma di discriminazione razziale e a promuovere il rispetto e la comprensione tra tutte le razze. Ciò significa, ai sensi dell'art. 2, che:

- (a) Ciascuno Stato contraente si impegna a non commettere alcun atto o pratica di discriminazione razziale nei confronti di persone, gruppi di persone o istituzioni e di garantire che tutti i cittadini, le autorità e le istituzioni pubbliche, nazionali e locali, agiscono conformemente a tale obbligo;
- (b) Ciascuno Stato contraente si impegna a non sponsorizzare, difendere o sostenere la discriminazione razziale da parte di persone o organizzazioni;
- (c) Ciascuno Stato parte adotta

misure efficaci per riesaminare le politiche governative, nazionali e locali, e modificare, revocare o annullare le leggi e i regolamenti che hanno l'effetto di creare o perpetuare ovunque la discriminazione razziale; (d) Ciascuno Stato parte vieta e pone fine, con tutti i mezzi opportuni, compresa la legislazione come richiesto dalle circostanze, la discriminazione razziale da qualsiasi persona, gruppo o organizzazione.

Successivamente, nel 1976, viene emanato il cosiddetto ICCPR (International Covenant on Civil and Political Rights) il quale, all'art. 20, prevede che:

Qualsiasi difesa dell'odio nazionale, razziale o religioso che costituisce incitamento a discriminazione, ostilità o violenza, deve essere proibito dalla legge.

Gli autori di *Countering Hate Speech Online* (Gagliardone, Gal, Alves, Martinez 2015) rilevano come sussistano alcune differenze tra i due documenti legislativi appena citati. Infatti, essi in primo luogo ritengono che il concetto di "difesa dell'odio" introdotto nell'ICCPR sia più specifico del discorso discriminatorio descritto nell'ICERD, poiché si ritiene necessario prendere in considerazione l'intento dell'autore e non l'espressione d'odio da sola considerata – questo perché il termine *advocacy* viene interpretato nell'ICCPR come la necessità

che sussista come requisito l'intento di seminare odio. La mera diffusione di messaggi di superiorità razziale o di odio, o addirittura di incitamento alla discriminazione razziale o alla violenza, sarà punibile ai sensi dell'ICERD. Ma nel ICCPR, affinché il reato sia vietato è necessario dimostrare l'intento di incitare all'odio ai sensi dell'articolo 20. Infine, nel 1981, la Convenzione sull'eliminazione di tutte le forme di discriminazione contro le donne impegna gli Stati parte a condannare la discriminazione contro le donne e a prevenire, indagare, perseguire e punire gli atti di violenza di genere.

Dal piano internazionale, ora ci caliamo in quelli che sono stati i tentativi definitivi messi in atto dall'Unione Europea al fine di delimitare i confini del discorso d'odio. Un primo tentativo definitivo del discorso d'odio può rinvenirsi nella Raccomandazione del Comitato dei Ministri del Consiglio d'Europa del 30 ottobre 1997. Tuttavia, tale disposizione risulta non solo priva di vincolatività in quanto contenuta in un atto di soft law, ma soprattutto incompleta, dal momento che non include tra le manifestazioni d'odio alcuni fenomeni quali l'omofobia e la misoginia, nonché altre forme di espressioni d'odio ad personam, quali il cyberstalking e il cyberbullismo (Allegri 2012). Si sopperi-

sce quindi a tale lacuna mediante il Protocollo addizionale alla Convenzione di Budapest sulla criminalità informatica firmata a Strasburgo il 28 gennaio 2003⁶. Questo mira alla criminalizzazione dei comportamenti di natura razzista e xenofoba diffusi tramite l'utilizzo di sistemi informatici. Il Protocollo prevede che gli Stati parte definiscano come reato la diffusione o altre forme di messa a disposizione del pubblico per il tramite di un sistema informatico: di materiale razzista e xenofobico (articolo 3); di materiale che neghi, minimizzi in modo palese, approvi o giustifichi degli atti che costituiscano la fattispecie di genocidio o crimine contro l'umanità, come definiti dal diritto internazionale e riconosciuti come tali da una decisione definitiva del Tribunale militare internazionale o ogni altra corte internazionale (articolo 6).

Dal canto suo, l'articolo 21⁷ della Carta di Nizza sancisce il divieto di discriminazioni quale

⁶<http://documenti.camera.it/leg17/dossier/Pdf/ES0420.pdf> (12/08/2024).

⁷ 1. È vietata qualsiasi forma di discriminazione fondata, in particolare, sul sesso, la razza, il colore della pelle o l'origine etnica o sociale, le caratteristiche genetiche, la lingua, la religione o le convinzioni personali, le opinioni politiche o di qualsiasi altra natura, l'appartenenza ad una minoranza nazionale, il patrimonio, la nascita, la disabilità, l'età o l'orientamento sessuale.

2. Nell'ambito d'applicazione dei trattati e fatte salve disposizioni specifiche in essi contenute, è vietata qualsiasi discriminazione in base alla nazionalità.

principio giuridicamente vincolante. Su tale base, sono state adottate due importanti direttive: la direttiva 2000/43/Ce del Consiglio, del 29 giugno 2000, che attua il principio della parità di trattamento fra le persone indipendentemente dalla razza e dall'origine etnica, e la direttiva 2000/78/Ce del Consiglio, del 27 novembre 2000, che stabilisce un quadro generale per la parità di trattamento in materia di occupazione e di condizioni di lavoro, indipendentemente da religione o convinzione personale, disabilità, età o orientamento sessuale. Nello stesso anno, il Consiglio d'Europa adotta la Raccomandazione n. 6 di politica generale allo scopo di perseguire la lotta contro la diffusione di materiale razzista, xenofobo e antisemita via Internet. Successivamente, nel 2008, il Consiglio d'Europa adotta la decisione-quadro 2008/913/Gai sulla lotta contro talune forme ed espressioni di razzismo e xenofobia mediante lo strumento del diritto penale. Tale decisione, ai sensi dell'articolo 1, prevede che:

ciascuno Stato membro adotta le misure necessarie affinché i seguenti comportamenti intenzionali siano resi punibili: a) l'istigazione pubblica alla violenza o all'odio nei confronti di un gruppo di persone, o di un suo membro, definito in riferimento alla razza, al colore, alla religione, all'ascendenza o all'origine nazionale o etnica; b) la perpetrazione di uno degli atti di cui

alla lettera a) mediante la diffusione e la distribuzione pubblica di scritti, immagini o altro materiale; c) l'apologia, la negazione o la minimizzazione grossolana dei crimini di genocidio, dei crimini contro l'umanità e dei crimini di guerra, quali definiti agli articoli 6, 7 e 8 dello statuto della Corte penale internazionale, dirette pubblicamente contro un gruppo di persone, o un membro di tale gruppo, definito in riferimento alla razza, al colore, alla religione, all'ascendenza o all'origine nazionale o etnica, quando i comportamenti siano posti in essere in modo atto a istigare alla violenza o all'odio nei confronti di tale gruppo o di un suo membro.

Il Parlamento europeo, con una Risoluzione approvata il 14 marzo 2013, ha evidenziato l'esigenza di una revisione della decisione-quadro 2008/913/Gai, in modo da includervi anche le manifestazioni di antisemitismo, intolleranza religiosa, antiziganismo, omofobia e transfobia.

Nel 2022, infine, il Comitato dei Ministri del Consiglio d'Europa ha adottato la Raccomandazione CM/Rec (2022) sulla lotta contro l'incitamento all'odio definendo tale fenomeno come:

qualsiasi tipo di espressione che incita, promuove, diffonde o giustifica violenza, odio o discriminazione contro un individuo o un gruppo di persone a causa di caratteristiche come la razza, il colore della pelle, la lingua, la religione, la nazionalità, l'origine nazionale o etnica, l'età, la disabilità, il sesso, l'identità di genere e l'orientamento sessuale.

La raccomandazione compie un passo in avanti rispetto al passato, perché identifica in modo più specifico diverse tipologie di discorsi di odio distinguendo tra quelli che hanno una rilevanza sul piano penale, quindi condotte che si configurano come reati, quelli che sono prevedono (solo) sanzioni sul piano civile e amministrativo, ed infine le espressioni ritenute non abbastanza gravi da essere limitate dalla Convenzione Europea dei Diritti Umani ma che richiedono l'adozione di risposte alternative (produzione di contro-narrazioni, attività di sensibilizzazione, educazione interculturale).

Infine, prima di capire come l'Italia si sia regolata nel corso del tempo in relazione al fenomeno dell'hate speech, di seguito osserviamo i principali interventi legislativi posti in essere da altri Stati. Nel Regno Unito, per esempio, secondo quanto riportato dall'Osservatorio Italiano sui diritti, nel 1965 è stato emanato il Race Relations Act, nel 1986 il Public Order Act ed infine nel 2006 il Racial and Religious Hatred Act.

In Germania invece, il discorso d'odio è sanzionato penalmente dall'articolo 130 del Codice penale tedesco, il quale dispone che:

chi, in maniera tale da disturbare la pace pubblica, incita all'odio o alla violenza contro elementi della popolazione o lede la dignità di altre persone attraverso insulti o offese è punito con una pena detentiva da tre mesi a cinque anni.

Il comma 2 dell'art. 130 prevede una pena detentiva fino a tre anni o una pena pecuniaria per chi commette gli stessi illeciti attraverso la diffusione di opere scritte. Inoltre, in ambiente tedesco, il 1° gennaio 2018 è entrata in vigore una nuova legge contro l'hate speech online. Lo scopo principale della legge è obbligare i principali social networks, come Facebook e Twitter, a rispettare le severissime leggi tedesche in materia di diffamazione, incitamento all'odio e minacce, in vigore dalla fine della Seconda guerra mondiale.

Il problema, secondo alcuni esperti di libertà di espressione, è che la legge è scritta in termini talmente vaghi da sanzionare commenti molto diversi, e che la sua applicazione viene lasciata in mano a società private come quelle che gestiscono i social networks, i cosiddetti ISP (Internet Service Providers). La legge è stata introdotta a seguito di un report governativo da cui risultava che Facebook e Twitter avessero rimosso materiale con-

siderato illegale entro ventiquattro ore dalla pubblicazione solo in alcuni casi e non almeno nel 70% delle occasioni come richiesto dal governo.

Infine, in Francia è stata approvata nel luglio del 2019 la legge “anti-haters” proposta da una giovane deputata, Laetitia Avia, lei stessa vittima di odio online. Secondo la nuova normativa, i contenuti che incitano all’odio e alla discriminazione “razziale, religiosa, di genere, sull’orientamento sessuale o sull’handicap” dovranno essere eliminati entro ventiquattro ore. Per le piattaforme, i social networks e i motori di ricerca che non si adegueranno, sono previste sanzioni amministrative fino al 4% del fatturato, fino a un massimo di 1,25 milioni di euro. La nuova legge francese prevede di creare una giurisdizione speciale, con magistrati specializzati nel verificare e punire messaggi incriminati. La giovane ha dichiarato:

Non vi è alcun motivo per cui i commenti che non sarebbero tollerati su un autobus, in un bar o a scuola – fondamentalmente, nella vita reale – dovrebbero essere tollerati su un sito Web o una rete.

In Italia invece, il primo intervento nella lotta contro l’hate speech è da ravvisare nella legge

n. 654 del 1975⁸, con cui il legislatore italiano ha ratificato la Convenzione di New York sull'eliminazione di tutte le forme di discriminazione razziale.

Il secondo intervento di fondamentale importanza è sicuramente costituito dalla legge n. 205 del 1993, la cosiddetta "legge Mancino", con la quale si sono irrobustite le tutele attraverso la previsione di nuove fattispecie delittuose. Ai sensi dell'articolo 1, il dettato legislativo così dispone:

Salvo che il fatto costituisca più grave reato, anche ai fini dell'attuazione della disposizione dell'articolo 4 della convenzione, è punito: a) con la reclusione sino a tre anni chi diffonde in qualsiasi modo idee fondate sulla superiorità o sull'odio razziale o etnico, ovvero incita a commettere o commette atti di discriminazione per motivi razziali, etnici, nazionali o religiosi; b) con

⁸Art. 3: Salvo che il fatto costituisca più grave reato, ai fini dell'attuazione della disposizione dell'articolo 4 della convenzione è punito con la reclusione da uno a quattro anni: a) chi diffonde in qualsiasi modo idee fondate sulla superiorità o sull'odio razziale; b) 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 ad un gruppo nazionale, etnico o razziale.

È vietata ogni organizzazione o associazione avente tra i suoi scopi di incitare all'odio o alla discriminazione razziale. Chi partecipi ad organizzazioni o associazioni di tal genere, o presti assistenza alla loro attività, è punito per il solo fatto della partecipazione o dell'assistenza, con la reclusione da uno a cinque anni. Le pene sono aumentate per i capi e i promotori di tali organizzazioni o associazioni.

la reclusione da sei mesi a quattro anni chi, in qualsiasi modo incita a commettere o commette violenza o atti di provocazione alla violenza per motivi razziali, etnici, nazionali o religiosi.

È vietata ogni organizzazione, associazione, movimento o gruppo avente tra i propri scopi l'incitamento alla discriminazione o alla violenza per motivi razziali, etnici, nazionali o religiosi. Chi partecipa a tali organizzazioni, associazioni, movimenti o gruppi, o presta assistenza alla loro attività, è punito, per il solo fatto della partecipazione o dell'assistenza, con la reclusione da sei mesi a quattro anni. Coloro che promuovono o dirigono tali organizzazioni, associazioni, movimenti o gruppi sono puniti, per ciò solo, con la reclusione da uno a sei anni.

La legge n. 85 del 2006, modificando la precedente, riduce i limiti edittali delle pene reclusive, prevede pene pecuniarie alternative alla reclusione e sostituisce con "propaganda" la precedente espressione "diffonde in qualsiasi modo" e con "istiga" il precedente "incita". Ad oggi costituiscono quindi condotte delittuose la propaganda, l'istigazione in luogo di diffusione e l'incitamento.

Più recentemente, va ricordata la legge n. 115 del 2016, la quale prevede la reclusione da due a sei anni nei casi in cui la propaganda, l'istigazione e l'incitamento, commessi in modo che derivi concreto pericolo di diffusione, trovino il proprio fondamento 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 dallo Statuto della Corte penale internazionale. Si puniscono quindi le opinioni negazioniste da cui derivi concretamente propaganda, istigazione o incitamento alla violenza o alla discriminazione.

Infine, va ricordata la legge n. 71 del 2017 volta a contrastare il fenomeno del cyberbullismo, termine con cui si intende, ai sensi dell'art. 1 comma 2:

qualunque forma di pressione, aggressione, molestia, ricatto, ingiuria, denigrazione, diffamazione, furto d'identità, alterazione, acquisizione illecita, manipolazione, trattamento illecito di dati personali in danno di minorenni.

In particolare, in caso di odio online espresso nei confronti di una persona determinata, le fattispecie penalmente rilevanti sono quelle della diffamazione aggravata dall'utilizzo di un mezzo di pubblicità (art. 595 c.p.) e della minaccia (art. 612 c.p.) eventualmente aggravata (art. 339 c.p.). Ad esse si aggiunge il reato di atti persecutori introdotto nel Codice penale con decreto-legge n. 11 del 2009: si tratta del reato commesso da:

chiunque, con condotte reiterate, minaccia o molesta taluno in modo da cagionare un perdurante e grave stato di ansia o di paura ovvero da ingenerare un fondato timore per l'incolumità propria o di un prossimo congiunto o di persona al medesimo legata da relazione affettiva ovvero da costringere lo stesso ad alterare le proprie abitudini di vita (art. 612 bis c.p.).

1.3. La tutela della libertà di espressione del pensiero

La domanda che quindi i giuristi di oggi si pongono è come contemperare il bisogno, previsto a livello costituzionale, di tutela della libertà di espressione di pensiero con la necessità di salvaguardare altri diritti ugualmente protetti dalla Costituzione, in relazione al fenomeno online hate speech. Prima però di capire come, sia il legislatore che i giudici, si siano espressi a riguardo, è necessario ripercorrere in breve la storia legata al riconoscimento della suddetta libertà.

La teorizzazione del principio di libertà di manifestazione del pensiero in prima battuta con Erasmo da Rotterdam, il quale sostiene che “in a free state, tongues should be free too” ([1515] 1997, 88), in seconda battuta con John Milton (1644), il quale si pone a favore della libertà di stampa senza censura. Quest'ultimo infatti considera la libertà, in particolare la libertà di espressione del pensiero, come un diritto innato, inalie-

nabile, come tale non soggetto a qualsivoglia limitazione di tipo politico, religioso o filosofico (Moscati 2008). Ed è proprio su questo postulato di Milton che si basa la teoria sulla libertà di John Stuart Mill. In particolar modo, egli si concentra sulla libertà di espressione del pensiero, sostenendo: a) la fallibilità delle opinioni; b) l'importanza dello scontro e della discussione, anche in presenza di opinioni considerate vere o condivise; e c) la negazione del monopolio della verità da parte di determinate opinioni, con la conseguente rivendicazione del valore della diversità (Mill 1981).

Ora, il primo riconoscimento giuridico, seppur limitato, del diritto alla libertà di espressione del pensiero avviene con il Bill of Rights del 1689. Ma bisognerà attendere sino al 1789 perché la protezione giuridica di tale libertà possa dirsi totale, mediante l'entrata in vigore e della Dichiarazione francese dei diritti dell'uomo e del cittadino e della Dichiarazione di indipendenza degli Stati Uniti d'America. Al contrario, in Italia e in Germania, l'esigenza di tutelare la libertà di espressione del pensiero sorge solo in conseguenza del recepimento dei principi internazionali che si diffondono nella prassi e nella dottrina a partire dal 1948

e che portano all'emanazione, nei rispettivi ordinamenti nazionali, delle Costituzioni del 1948 e del 1949.

A tal proposito, anche la Comunità internazionale si è impegnata nel corso del tempo per tutelare la libertà di espressione del pensiero siglando una serie di importanti accordi. Il primo intervento normativo risale al 1948, quando le Nazioni Unite emanano la Dichiarazione Universale dei Diritti dell'uomo. In merito, Paolo Sordini⁹ ritiene che siano tre gli articoli della Dichiarazione da ritenersi applicabili al rapporto tra libertà di espressione del pensiero e discorso d'odio diffuso mediante Internet: gli articoli 12, 19 e 27, i quali rispettivamente recitano quanto segue:

Nessun individuo potrà essere sottoposto ad interferenze arbitrarie nella sua vita privata, nella sua famiglia, nella sua casa, nella sua corrispondenza, né a lesione del suo onore e della sua reputazione. Ogni individuo ha diritto ad essere tutelato dalla legge contro tali interferenze o lesioni.

Ogni individuo ha diritto alla libertà di opinione e di espressione incluso il diritto di non essere molestato per la propria opinione e quello di cercare, ricevere e diffondere informazioni e idee attraverso ogni mezzo e senza riguardo a frontiere.

⁹https://www.ispionline.it/sites/default/files/pubblicazioni/wp_52_2013.pdf (2.2.2025)

Ogni individuo ha diritto di prendere parte liberamente alla vita culturale della comunità, di godere delle arti e di partecipare al progresso scientifico ed ai suoi benefici.

Ogni individuo ha diritto alla protezione degli interessi morali e materiali derivanti da ogni produzione scientifica, letteraria e artistica di cui egli sia autore.

Ciò non toglie che la Dichiarazione preveda altresì eccezioni disciplinate ai sensi dell'articolo 29 comma 2:

Nell'esercizio dei suoi diritti e delle sue libertà, ognuno deve essere sottoposto soltanto a quelle limitazioni che sono stabilite dalla legge per assicurare il riconoscimento e il rispetto dei diritti e delle libertà degli altri e per soddisfare le giuste esigenze della morale, dell'ordine pubblico e del benessere generale in una società democratica.

Successivamente, nel 1976, entra in vigore il Patto Internazionale per i Diritti Civili e Politici (ICCPR) che, a differenza della Dichiarazione, si esprime in termini più specifici in relazione alla libertà di espressione del pensiero, recitando, ai sensi dell'art. 19 comma 2, che:

tale diritto comprende la libertà di cercare, ricevere e importare informazioni e idee di ogni tipo, indipendentemente dalle frontiere, oralmente, per iscritto o

stampa, sotto forma di arte, o attraverso qualsiasi altro mezzo di sua scelta.

Anche tale Patto, ai sensi dell'articolo 19 comma 3, prevede restrizioni al principio enunciato, tuttavia stabilendo che:

dette restrizioni devono essere espressamente stabilite dalla legge e risultare indispensabili per a) assicurare il rispetto dei diritti o della reputazione altrui; b) la protezione della sicurezza nazionale, l'ordine pubblico, la salute o la morale pubbliche.

Infine, l'articolo 20 afferma che:

ogni forma d'incitamento all'odio basato sull'appartenenza a uno stato, a una razza o religione sarà proibito dalla legge.

Sulla base di queste disposizioni, gli Stati parte si impegnano ad inviare ogni cinque anni rapporti sulle misure da essi adottate per la protezione dei diritti umani.

Infine, ultimo strumento normativo utilizzato al fine di disciplinare la libertà di espressione del pensiero veicolato attraverso la rete è costituito dal Patto Internazionale sui Diritti Economici, Sociali e Culturali (ICESCR) il quale, ai sensi dell'articolo 15, dispone che:

1. Gli Stati parte del presente Patto riconoscono il diritto di tutti a: (a) partecipare alla vita culturale; (b) godere dei benefici del progresso scientifico e delle sue applicazioni; (c) beneficiare della protezione degli interessi morali e materiali derivanti da qualsiasi produzione letteraria o artistica di cui è autore.
2. Le misure che gli Stati parte del presente Patto devono adottare per realizzare pienamente tale diritto comprendono quelle necessarie per la conservazione, lo sviluppo e la diffusione di scienza e cultura.
3. Gli Stati parte del presente Patto si impegnano a rispettare la libertà indispensabile per la ricerca scientifica e l'attività creativa.

È previsto altresì che, ai sensi dell'articolo 16, gli Stati parte inviino rapporti sulle misure adottate ed i progressi realizzati al fine di assicurare il rispetto dei diritti riconosciuti dal Patto al Comitato sui Diritti Economici, Sociali e Culturali che ha il compito di controllare l'operato dei singoli Stati.

1.4. Gli accordi europei e la Corte EDU

L'Unione europea, dal canto suo, ha stipulato nel corso del tempo dei veri e propri Accordi, allo scopo di disciplinare il principio della libertà di espressione del pensiero, qualsiasi sia il mezzo di comunicazione usato per veicolarlo. In prima battuta, nel 1950, viene adottata dagli Stati membri del Consiglio d'Europa la Convenzione Europea dei diritti dell'uomo (CEDU). Ai sensi dell'articolo

10, la Convenzione disciplina la libertà di espressione, sancendo che:

1. Ogni persona ha diritto alla libertà d'espressione. Tale diritto include la libertà d'opinione e la libertà di ricevere o di comunicare informazioni o idee senza che vi possa essere ingerenza da parte delle autorità pubbliche e senza considerazione di frontiera. Il presente articolo non impedisce agli Stati di sottoporre a un regime di autorizzazione le imprese di radiodiffusione, di cinema o di televisione.

2. L'esercizio di queste libertà, poiché comporta doveri e responsabilità, può essere sottoposto alle formalità, condizioni, restrizioni o sanzioni che sono previste dalla legge e che costituiscono misure necessarie, in una società democratica, per la sicurezza nazionale, per l'integrità territoriale o per la pubblica sicurezza, per la difesa dell'ordine e per la prevenzione dei reati, per la protezione della salute o della morale, per la protezione della reputazione o dei diritti altrui, per impedire la divulgazione di informazioni riservate o per garantire l'autorità e l'imparzialità del potere giudiziario.

Tuttavia, il riconoscimento di suddetto principio trova poi un proprio bilanciamento con altri diritti egualmente tutelati dalla Convenzione. A tal proposito, rileva l'articolo 17 della Convenzione, il quale sancisce che:

Nessuna disposizione della presente Convenzione può essere interpretata come implicante il diritto per uno

Stato, un gruppo o un individuo di esercitare un'attività o compiere un atto che miri alla distruzione dei diritti o delle libertà riconosciuti nella presente Convenzione o porre a questi diritti e a queste libertà limitazioni più ampie di quelle previste in detta Convenzione.

Altri diritti protetti dalla Convenzione e che rilevano ai fini della tutela della libertà di manifestazione del pensiero sono previsti dagli articoli 8 (diritto al rispetto della privacy nella corrispondenza), articolo 9 (diritto a manifestare la propria religione o il proprio credo) e articolo 11 (diritto alla libertà di riunione e di associazione pacifica).

Nell'aprile del 1982, il Consiglio d'Europa ribadisce il proprio impegno sul fronte della libertà di espressione mediante la Dichiarazione sulla libertà di espressione e d'informazione, la quale stabilisce che la libertà di espressione e informazione

è imprescindibile per lo sviluppo sociale, economico, culturale e politico di ogni essere umano; essa è inoltre una condizione indispensabile per lo sviluppo armonioso dei gruppi sociali e culturali, le nazioni e la comunità internazionale.

Essa inoltre riconosce che:

il continuo sviluppo delle tecnologie dell'informazione e delle comunicazioni dovrebbe servire per ampliare

senza limiti di spazio la portata del diritto di esprimere, reperire, ricevere, diffondere informazioni e idee, a prescindere dalla fonte.

Nel 1990 gli Stati membri dell'OCSE (Organizzazione per la sicurezza e la cooperazione in Europa) sottoscrivono la Carta di Parigi per una Nuova Europa, la quale riconosce che ogni individuo, senza alcuna distinzione, ha diritto alla libertà di pensiero, coscienza, religione o credo e alla libertà di espressione. I membri OCSE completano infine la Carta nel contesto della conferenza di Budapest del 1994, dove gli Stati coinvolti assumono come principio guida la salvaguardia del diritto alla libertà di espressione e riconoscono che i mezzi indipendenti e pluralistici sono essenziali per la creazione di una società libera e aperta. Tuttavia, anche la Carta di Parigi prevede alcune limitazioni alla libertà di espressione del pensiero, a patto che siano 1) previste dalla legge; 2) conformi agli obblighi derivanti dagli accordi internazionali, in primo luogo la Dichiarazione universale e il Patto; 3) riferite a un obiettivo delle norme internazionali in materia ed 4) essere strettamente proporzionali rispetto alle finalità che tali norme si prefiggono.

In ultima battuta, la Carta di Nizza del 2000 riconosce ad ogni individuo il diritto alla libertà di espressione sancendo che:

tale diritto include la libertà di opinione e la libertà di ricevere o di comunicare informazioni o idee senza che vi possa essere ingerenza da parte delle autorità pubbliche e senza limiti di frontiera.

A questo punto, è bene analizzare come i giudici europei si approccino al tema dell'hate speech. Françoise Tulkens (2012), giudice e vicepresidente della Corte Europea dei Diritti dell'Uomo dal 1998 al 2012, sottolinea come la Corte operi un vero e proprio bilanciamento tra il principio di libertà di espressione del pensiero, sancito dall'articolo 10 CEDU, e altri diritti ritenuti ugualmente meritevoli di tutela da parte dell'ordinamento sovranazionale. Ciò emerge in particolare dall'invocazione dell'articolo 17 CEDU, che disciplina il divieto di abuso di diritto e recita quanto segue:

Nessuna disposizione della presente Convenzione può essere interpretata come implicante il diritto per uno Stato, gruppo o individuo di esercitare una attività o compiere un atto mirante alla distruzione dei diritti o delle libertà riconosciuti nella presente Convenzione o

porre a questi diritti e a queste libertà limitazioni maggiori di quelle previste in detta Convenzione.

Emblematica a tal proposito è la sentenza del 6 luglio 2006, *Erbakan v. Turkey*¹⁰, con la quale la Corte ha ritenuto sussistente una violazione dell'articolo 10 CEDU. Nel caso di specie, il ricorrente è Necmettin Erbakan, politico turco nonché ex Primo Ministro della Turchia da giugno 1996 a giugno 1997. Al momento dei fatti, egli ricopre il ruolo di presidente del cosiddetto Refah Partisi, il Welfare Party, che fu dissolto nel 1998 perché coinvolto in attività contrarie ai principi del secolarismo. Il 25 febbraio 1994, durante la campagna elettorale, il ricorrente tiene un discorso pubblico presso Bingöl, città nel Sud-Est della Turchia. Questo discorso non viene registrato.

Più di quattro anni dopo, nel luglio del 1998, viene aperto un procedimento penale contro l'ex Primo ministro con l'accusa di aver incitato le persone all'odio o all'ostilità mediante commenti fatti nel corso del discorso realizzato nel 1994, in relazione alle distinzioni tra religioni, razze e regioni. Il politico, basandosi sull'articolo 10 della Convenzione Europea dei Diritti dell'uomo, sostiene che la sentenza di condanna emessa nei suoi riguardi

¹⁰<https://hudoc.echr.coe.int/eng#%7B%22itemid%22:%5B%22003-1728198-1812055%22%5D%7D> (03/09/2024).

violi il suo diritto alla libertà di espressione del pensiero. La Corte ritiene che il procedimento posto a carico del politico quattro anni e cinque mesi dopo che le presunte osservazioni erano state fatte, non fosse ragionevolmente proporzionato agli scopi legittimi perseguiti, tenuto conto dell'interesse della società democratica nel garantire e mantenere la libertà di dibattito politico. Di conseguenza, la Corte constata una violazione dell'articolo 10, sostenendo che le parole utilizzate dal ricorrente, ex Primo ministro della Turchia, in un incontro pubblico, non siano altro che espressione del pluralismo che connota le società moderne.

In generale, si può dire che, quando la Corte si trova di fronte al tema dell'incitamento all'odio e deve valutarne il rapporto con la libertà di espressione del pensiero, essa tende a adottare due approcci tipici. Uno più ampio, attraverso il quale vengono escluse talune espressioni dalla protezione fornita dall'articolo 17 CEDU che disciplina il divieto di abuso di diritto, e uno più restrittivo, attraverso l'applicazione dell'articolo 10 comma 2 CEDU che stabilisce i limiti alla libertà di espressione del pensiero tutelata ex articolo 10 comma 1. In quest'ultimo caso, la Corte opera un bilanciamento di interessi, mentre nel caso dell'articolo 17 attua una limitazione delle espressioni d'odio

esclusivamente attraverso una valutazione del loro contenuto. Tulkens, tuttavia, sottolinea come sia bene usare l'articolo 17 con moderazione, al fine di evitare gravi abusi, come è accaduto nel caso della Corte costituzionale turca che aveva ritenuto legittimi lo scioglimento e la dissoluzione dei partiti politici sulla base dell'articolo 17.

A tal proposito, merita qui ricordare la decisione *Glimmerveen and Hagenbeek vs the Netherlands*¹¹ del 1979 con cui Strasburgo ammette per la prima volta che determinate forme espressive possano essere escluse, in ragione del loro contenuto, dalla sfera di applicazione dell'articolo 10. Così operando, la Corte accerta la contrarietà delle espressioni razziste ai valori dell'ordinamento convenzionale, abbandonando un approccio di tipo individualistico e libertario, a favore di un percorso di funzionalizzazione della libertà ai valori convenzionali (Caruso 2017). Ma è con la sentenza *B.H., W. H.P. and G.K vs. Austria*¹² del 1989 che la Corte colloca le espressioni negazioniste e giustificazionistiche al di fuori della tutela convenzionale attraverso il richiamo all'abuse clause di cui all'art.

¹¹[https://hudoc.echr.coe.int/eng#{%22itemid%22:\[%22001-74187%22\]}](https://hudoc.echr.coe.int/eng#{%22itemid%22:[%22001-74187%22]}) (03/09/2024).

¹²[https://hudoc.echr.coe.int/eng#{%22dmdocnumber%22:\[%22665201%22\],%22itemid%22:\[%22001-1039%22\]}](https://hudoc.echr.coe.int/eng#{%22dmdocnumber%22:[%22665201%22],%22itemid%22:[%22001-1039%22]}) (03/09/2024).

17 CEDU. La disposizione comporta infatti l'interruzione della concreta analisi del caso concreto e la sussunzione, sic et simpliciter, della situazione di fatto nella clausola convenzionale dell'abuso del diritto (Castellaneta 2011, 73). L'orientamento giurisprudenziale prevalente adotta quindi un approccio content-based, escludendo categoricamente, con una sorta di "effetto-ghigliottina", determinate opinioni dalla libertà convenzionalmente tutelata¹³. Questa distinzione pone le basi per un approccio su due livelli, il cosiddetto "two tiers approach", da tempo adottato dalla Corte Suprema degli Stati Uniti, che legittima un diverso trattamento del discorso d'odio a seconda del suo particolare contenuto espressivo (Caruso 2013).

Un ulteriore consolidamento di questo orientamento proviene dalla conferma di condanna nel caso *Garaudy v. France*¹⁴. Roger Garaudy è un filosofo, scrittore e politico francese. Nel caso di specie, egli viene convenuto in giudizio in veste di autore del libro *I miti fondatori della politica israeliana* da parte di associazioni di ex membri della resistenza, deportati e organizzazioni per i diritti umani che denunciano i seguenti reati: contestazione circa l'esistenza di crimini contro

¹³ <https://www.corteidh.or.cr/tablas/r25909.pdf> (6.2.2024)

¹⁴ <http://hudoc.echr.coe.int/eng?i=001-23829> (03/09/2024).

l'umanità, diffamazione razziale in pubblico e incitamento all'odio razziale. In base al Freedom of Press Act del 29 luglio 1881 sono stati avviati cinque diversi procedimenti penali a suo carico. In cinque sentenze del 16 dicembre 1998, la Corte d'appello di Parigi dichiara il signor Garaudy colpevole dei seguenti reati: contestazione dell'esistenza di crimini contro l'umanità, pubblica diffamazione di un gruppo di persone – vale a dire la comunità ebraica – e incitamento alla discriminazione e all'odio razziale. Le condanne sono confermate dalla Corte di cassazione. Il ricorrente lamenta, ai sensi dell'articolo 10 della Convenzione sui diritti dell'uomo, la violazione del suo diritto alla libertà di espressione.

Sostiene inoltre che il suo libro sia un lavoro politico scritto al fine di combattere il sionismo, criticare la politica israeliana e che non ha alcun contenuto razzista o antisemita. Per quanto riguarda le condanne di Garaudy relative alla contestazione circa l'esistenza di crimini contro l'umanità, la Corte fa riferimento all'articolo 17, pensato proprio per impedire alle persone di dedurre dalla Convenzione qualsiasi diritto a compiere atti volti a distruggere i diritti e le libertà stabilite nella Convenzione. Dopo aver analizzato il libro in que-

stione, la Corte trova che il ricorrente abbia adottato teorie revisioniste e sistematicamente contestato l'esistenza dei crimini contro l'umanità che i nazisti avevano commesso contro la comunità ebraica. Non ci può essere alcun dubbio che contestare l'esistenza di eventi storici chiaramente avvenuti, come l'Olocausto, non costituisce ricerca storica simile ad una ricerca per la verità. Il vero scopo di un tale lavoro è quello di riabilitare il regime nazionalsocialista e, di conseguenza, accusare le vittime dell'Olocausto di falsificare la storia. La negazione o la riscrittura di questo tipo di fatto storico minano i valori su cui la lotta contro il razzismo e l'antisemitismo si basa e costituisce una grave minaccia per l'ordine pubblico. La Corte rileva che, poiché il libro del ricorrente, considerato nel suo insieme, rappresenta una marcata tendenza al revisionismo, si pone contro i valori fondamentali della Convenzione, ossia giustizia e pace. Per quanto riguarda invece le condanne di Garaudy per diffamazione e incitamento all'odio razziale, i giudici ritengono che esse possano costituire un'ingerenza nel suo diritto alla libertà di espressione.

Per quanto concerne l'articolo 10 comma 2 CEDU invece, l'interprete si pone il problema di ca-

pire quali siano le ragioni che giustificano la previsione di eccezioni al principio di libertà di manifestazione del pensiero. La sentenza *Handyside v. United Kingdom* del 1976 costituisce una colonna portante in tal senso¹⁵, dal momento che attribuisce alla libertà di espressione del pensiero una funzione sociale fondamentale. La libertà di opinione è considerata fondamento del progresso democratico e dell'autonomia dell'individuo; essa attrae non solo le informazioni o le idee che incontrano il consenso della maggioranza, ma anche “those that offend, shock, or disturb the State or any sector of the population”.

Nel caso concreto, Richard Handyside è il proprietario degli editori di *Stage 1*. Acquista i diritti britannici di *The Little Red Schoolbook*, scritto da Søren Hansen e Jesper Jensen. Il libro è inizialmente pubblicato nel 1969 in Danimarca e le traduzioni sono successivamente pubblicate in molti paesi europei, nonché in diversi paesi extraeuropei. Uno dei capitoli contiene una sezione dedicata al tema del sesso. Handyside invia diverse centinaia di copie di revisione del libro, insieme a un comunicato stampa, a una selezione di testate, da

¹⁵<https://hudoc.echr.coe.int/tur#%7B%22itemid%22:%5B%22001-57499%22%5D%7D> (03/09/2024).

giornali nazionali e locali a riviste educative e mediche. Il libro diventa oggetto di un ampio commento da parte della stampa, con reazioni contrastanti riguardo al contenuto. Dopo aver ricevuto una serie di denunce, il Direttore della pubblica accusa chiede alla polizia metropolitana di indagare se il libro abbia violato le leggi sull'oscenità. Di conseguenza, oltre mille copie del libro vengono sequestrate provvisoriamente ai sensi dell'Obscene Publications Act, insieme a volantini, manifesti, cartoline e corrispondenza relativi alla pubblicazione e alla vendita del libro. Sotto processo, Handyside è dichiarato colpevole di possesso di pubblicazioni oscene a scopo di lucro. La pronuncia della Corte si basa sull'applicazione dell'art. 10 della Convenzione. In particolare, i giudici hanno sottolineato che:

la libertà di espressione è uno dei capisaldi di una società democratica e una delle condizioni fondamentali del suo sviluppo. L'art. 10 par. 2 si applica non tanto alle informazioni o alle opinioni inoffensive o irrilevanti, quanto a quelle suscettibili di recare offesa.

La Corte applica il seguente criterio di valutazione, il cosiddetto "necessity test". Ogni restrizione della libertà di espressione deve quindi: 1)

essere riconosciuta dalla legge; 2) avere come finalità un obiettivo non contrario all'art. 10 par. 2; 3) essere necessaria in una società democratica. In secondo luogo, le eccezioni al paragrafo 2 devono essere applicate in forma restrittiva. La Corte EDU conclude quindi che gli Stati dispongono di un margine di manovra al fine di stabilire se una determinata limitazione sia necessaria in virtù di determinate circostanze locali. Ciò significa che alcune forme di espressione possono essere proibite in alcuni paesi ma non in altri. E questo si verifica specialmente in relazione al concetto della morale pubblica.

Nel caso *Handyside*, la Corte decide che il Regno Unito potrebbe sottoporre a processo una casa editrice e distruggere le copie di un libro, pur essendo tale libro ammesso in altri paesi. La decisione ha avuto il merito di includere le manifestazioni di pensiero anticonformiste nell'ambito dell'art. 10 CEDU, senza che rilevi, ai fini della determinazione del quantum di tutela convenzionale, il tipo di espressione utilizzata. La pronuncia esprime un sostanziale rifiuto di limitazioni content-based. Non è il tipo di espressione che determina il quantum di tutela: quest'ultimo sarà invece determinato dall'esito dell'*ad hoc balancing* condotto alla luce delle concrete circostanze del caso.

Dal canto suo, Ziccardi individua tre periodi che si sono susseguiti nel tempo al fine di cogliere l'evoluzione dell'atteggiamento della Corte sul tema hate speech. La prima serie di decisioni riguarda episodi di apologia della violenza. Ne costituisce esempio il caso *Surek v. Turkey*¹⁶. Kamil Tekin Sürek e Yücel Özdemir, il maggiore azionista e il caporedattore della rivista settimanale turca *Haberde Yorumda Gerçek*, pubblicano un'intervista con il secondo in comando del Partito dei lavoratori del Kurdistan (PKK) insieme a una dichiarazione congiunta di quattro organizzazioni curde. All'epoca dei fatti, il PKK e le quattro organizzazioni sono designati come gruppi terroristici dal governo turco. L'intervista solleva dubbi sulla posizione degli Stati Uniti rispetto allo status del popolo curdo e alle opinioni del PKK sull'ambiente politico in rapido cambiamento in Turchia. L'intervistatore si riferisce al territorio turco dove la maggioranza dei curdi risiede, il Kurdistan. La dichiarazione congiunta, pubblicata in un'edizione successiva, recita in parte:

Uniamoci contro il terrorismo di Stato, contro la repressione e l'oppressione del popolo curdo, contro i

¹⁶<https://hudoc.echr.coe.int/rus#%7B%22itemid%22:%5B%22001-58281%22%7D> (04/09/2024).

massacri, gli omicidi di strada, i licenziamenti e la disoccupazione; uniamo e intensifichiamo i nostri sforzi per la libertà, la democrazia e il socialismo!

Sürek e Özdemir vengono accusati di aver propagandato idee contrarie all'indivisibilità dello Stato, in base alle leggi antiterrorismo turche. Entrambi sono giudicati colpevoli. Sürek e Özdemir presentano quindi ricorso presso la Corte europea dei diritti dell'uomo. Anche in questa occasione, la Corte ribadisce il suo orientamento prevalente secondo cui, al fine di valutare se vi sia un urgente bisogno sociale di una restrizione al diritto alla libertà di espressione, deve esaminare

l'interferenza alla luce del caso nel suo insieme, compreso il contenuto delle dichiarazioni contestate e il contesto in cui sono state fatte. In particolare, deve determinare se l'interferenza in questione sia proporzionata agli obiettivi legittimi perseguiti e se i motivi adottati dalle autorità nazionali per giustificarlo siano pertinenti e sufficienti.

La Corte ricorda inoltre che, ai sensi dell'articolo 10 della Convenzione, vi sono poche possibilità di restrizioni a "discorsi politici o dibattiti su questioni di interesse pubblico" e che i limiti delle critiche ammissibili sono più ampi nei confronti

del governo che nei rapporti con un cittadino privato o persino con un politico. Alla luce di questi principi, la Corte in primo luogo osserva che il semplice fatto che le interviste vengano rese da un membro di un'organizzazione terroristica designata, non può di per sé giustificare l'accusa. In secondo luogo, essa osserva che, mentre il comandante del PKK esprime ostilità nei confronti del governo turco, i testi nel loro insieme non possono essere considerati come incitamento alla violenza o all'odio.

Infine, sebbene sussista la preoccupazione del governo per le parole e le azioni che potrebbero aggravare la situazione della sicurezza nel paese, è allo stesso tempo necessario tenere in debita considerazione il diritto del pubblico di essere informato su una diversa prospettiva concernente la situazione nella Turchia sudorientale, indipendentemente da quanto sgradevole questa prospettiva possa essere per loro. Sulla base delle ragioni sopra esposte, i giudici ritengono sussistente una violazione del diritto alla libertà di espressione.

Il caso *Faruk Temel v. Turkey*¹⁷ rappresenta invece un esempio di come la Corte, di fronte ad

¹⁷[https://hudoc.echr.coe.int/eng#{%22dmdocnum-ber%22:\[%22880747%22\],%22itemid%22:\[%22001-103141%22\]}](https://hudoc.echr.coe.int/eng#{%22dmdocnum-ber%22:[%22880747%22],%22itemid%22:[%22001-103141%22]}) (04/09/2024).

una violazione del principio di libertà di espressione del pensiero, tenda a non ritenere sussistenti atti che incitino concretamente alla violenza, alla luce di quel bilanciamento di interessi operato ex art. 10 CEDU. I ricorrenti, Senanik Öner e Ferhan Türk, partecipano alle celebrazioni di Newroz Kızıltepe, un distretto di Mardin, e fanno alcuni discorsi riguardanti i problemi del popolo curdo. Successivamente viene avviato un procedimento penale a loro carico, con l'accusa di aver fatto propaganda terroristica a nome del PKK (Partito dei lavoratori curdi), un'organizzazione illegale ai sensi della sezione 7§ 2 della legge sulla prevenzione del terrorismo (legge n. 3713). I ricorrenti lamentano che l'istituzione di un procedimento contro di loro per aver pronunciato un discorso, costituisca un atto ingiustificato nei confronti del loro diritto alla libertà di espressione, tutelato ai sensi dell'articolo 10 della Convenzione.

La Corte ritiene che la causa costituisca un'interferenza con il diritto dei ricorrenti alla libertà di espressione protetta dall'articolo 10. Osserva inoltre che il discorso in questione consiste in una valutazione critica delle politiche della Turchia in merito al problema curdo. I giudici sostengono quindi che, nel complesso, il discorso dei ricorrenti non incoraggi alla violenza, alla resistenza

armata o a compiere una rivolta. Inoltre, gli interventi in questione dei ricorrenti non erano in grado di incitare alla violenza instillando odio profondo e irrazionale contro persone determinate e quindi ciò non costituisce hate speech. Infine, i motivi addotti dalla Corte d'Assise di Diyarbakır per la condanna dei richiedenti non possono essere considerati pertinenti e sufficienti a giustificare l'ingerenza nel loro diritto alla libertà di espressione. La Corte conclude che la condanna dei ricorrenti è sproporzionata rispetto agli obiettivi perseguiti ai sensi dell'articolo 10 comma 2 della Convenzione. Quindi, si ritiene sussistente una violazione dell'articolo 10 della Convenzione.

A partire dal 2008 invece, i giudici della Corte iniziano ad occuparsi di episodi di hate speech a sfondo razziale ed etnico. Nel caso *Leroy v. France* del 2 ottobre 2008, il ricorrente, Denis Leroy¹⁸, è un vignettista francese e lavora in questa veste per diverse testate locali, tra cui il settimanale basco *Ekaitza*, che ha la sua sede a Bayonne. Il caso riguarda la condanna del ricorrente per complicità nell'istigazione al terrorismo, in seguito alla pubblicazione di un disegno relativo agli attentati

¹⁸<https://hudoc.echr.coe.int/eng-press#%22ite-mid%22:%22003-2501837-2699727%22>] (04/09/2024).

dell'11 settembre 2001. In questa data, infatti, il ricorrente presenta alla redazione di Ekaitza un disegno che rappresenta l'attacco alle torri gemelle del World Trade Centre, con una didascalia che fa la parodia di uno slogan pubblicitario di un marchio famoso: "Lo abbiamo sognato tutti... Hamas lo ha fatto". Il disegno è pubblicato sul giornale il 13 settembre 2001. Nel numero seguente, il giornale pubblica estratti di lettere ed e-mail ricevute in reazione al disegno. In seguito alla pubblicazione del disegno, il pubblico ministero di Bayonne intenta un'azione legale contro il ricorrente e il direttore editoriale del giornale, con l'accusa di complicità nell'istigazione al terrorismo e nel condono del terrorismo.

La Corte EDU ritiene che la condanna del ricorrente costituisca un'ingerenza nell'esercizio del suo diritto alla libertà di espressione. Tuttavia, tale ingerenza è prescritta dalla legge francese e persegue diversi scopi legittimi, tenuto conto del carattere sensibile della lotta contro il terrorismo, vale a dire il mantenimento della sicurezza pubblica e la prevenzione del disordine e della criminalità. Resta da determinare se tale ingerenza sia necessaria in una società democratica. La Corte rileva fin dall'inizio, che i tragici eventi dell'11 set-

tembre 2001, all'origine dell'espressione impugnata, hanno provocato il caos globale, e che le questioni sollevate in quell'occasione siano oggetto di discussione come questioni di interesse pubblico. Il ricorrente sostiene che i giudici francesi negano la sua reale intenzione, che è governata da un sentimento politico e attivista, vale a dire quella di comunicare il suo "anti-Americanism" attraverso un'immagine satirica e illustrando il declino dell'imperialismo americano.

La Corte, tuttavia, ritiene che il disegno non si limiti a criticare l'imperialismo americano, ma sostenga e glorifichi la violenta distruzione di quest'ultimo. Attraverso la scelta del linguaggio, il ricorrente commenta approvando la violenza perpetrata contro migliaia di civili e sminuendo la dignità delle vittime. Per di più, egli presenta il disegno il giorno degli attentati e questo viene pubblicato il 13 settembre, senza alcuna precauzione per quanto riguarda la lingua utilizzata. Secondo la Corte, questo fattore – la data di pubblicazione – è tale da accrescere la responsabilità del ricorrente nel suo resoconto e persino nel suo sostegno a un evento tragico, considerato dal punto di vista artistico o giornalistico. Inoltre, l'impatto di un tale messaggio in una regione politicamente sensibile, ossia il Paese Basco, non deve essere trascurato;

nonostante la circolazione limitata del settimanale, la Corte rileva che la pubblicazione del disegno aveva provocato una certa reazione pubblica, capace di fomentare la violenza e di manifestare un impatto plausibile sull'ordine pubblico nella regione. Di conseguenza, la Corte ritiene che i motivi adottati dai giudici nazionali per condannare il ricorrente fossero pertinenti e sufficienti.

In conclusione, tenuto conto del carattere modesto dell'ammenda inflitta al ricorrente e del contesto in cui il disegno impugnato è pubblicato, la Corte constata che la misura inflitta al ricorrente non sia sproporzionata rispetto all'obiettivo legittimo perseguito. Di conseguenza, l'articolo 10 non può dirsi violato.

Le Pen v. France del 20 aprile 2010¹⁹ costituisce altro caso emblematico di quello che può essere definito come il secondo periodo giurisprudenziale avente ad oggetto il tema hate speech. La ricorrente, presidente del partito politico Front national, accorda un colloquio al quotidiano Le Monde del 19 aprile 2003, nel corso del quale pronuncia le seguenti affermazioni:

Il giorno in cui in Francia avremo non più 5 milioni ma 25 milioni di musulmani, saranno loro a comandare. E

¹⁹<http://hudoc.echr.coe.int/eng?i=001-98489> (04/09/2024).

i francesi abatteranno i muri, scenderanno dai marciapiedi abbassando gli occhi. Quando non lo fanno, si dice loro: «Perché mi guardi così? Cerchi la rissa?». E non dovete far altro che andarvene, altrimenti vi prendete una bella ammogliata.

Il Tribunale penale di Parigi condanna la ricorrente per provocazione alla discriminazione, all'odio, alla violenza contro un gruppo di persone per la loro origine o appartenenza ad una determinata etnia, nazione, razza o religione. Con una sentenza del 24 febbraio 2005, la Corte d'appello di Parigi conferma la pronuncia in primo grado. La Corte rileva che la condanna della ricorrente si configura come un'ingerenza delle autorità pubbliche nell'esercizio della libertà di espressione, quale riconosciuta dall'articolo 10 comma 1 della Convenzione. Tale ingerenza è prevista dalla legge, ossia gli articoli 23 e 24 della legge del 29 luglio 1881 sulla libertà di stampa.

La Corte ritiene inoltre che tale ingerenza persegua il legittimo scopo di tutelare la reputazione e i diritti altrui, previsto dall'articolo 10, paragrafo 2 della Convenzione. Quanto alla condizione di necessità in una società democratica, è necessario stabilire se l'ingerenza incriminata corrisponda ad un "bisogno sociale imperativo".

La Corte ricorda a questo proposito che è fondamentale, in una società democratica, difendere il libero gioco del dibattito politico, che si trova al centro stesso della nozione di società democratica. Più in generale, i giudici ritengono che chiunque si impegni in un dibattito pubblico di interesse generale possa ricorrere ad una certa dose di esagerazione, se non addirittura di provocazione, cioè di essere un po' smodato nelle proprie affermazioni. Tuttavia, tali affermazioni non devono superare certi limiti, in particolare per quanto riguarda il rispetto della reputazione e dei diritti altrui.

Nella fattispecie, la Corte rileva che le affermazioni della ricorrente si inseriscono nell'ambito del dibattito di interesse generale relativo ai problemi connessi all'insediamento e all'integrazione degli immigrati nei paesi ospitanti. Essa rileva tuttavia che le affermazioni di quest'ultima sono certamente suscettibili di dare un'immagine negativa, anche inquietante, della comunità musulmana nel suo insieme. In tal modo infatti la ricorrente oppone, da un lato, i francesi e, dall'altro, una comunità, la cui appartenenza religiosa è espressamente menzionata e la cui forte crescita costituirebbe una minaccia, già presente, per la dignità e la sicurezza dei francesi. I giudici ritengono altresì

che le affermazioni della ricorrente possano suscitare un sentimento di rigetto e di ostilità nei confronti della comunità destinataria. Stando così le cose, secondo la Corte, i motivi addotti dai giudici interni sono pertinenti e sufficienti. Pertanto, tenuto conto delle circostanze del caso di specie, la Corte EDU ritiene che l'ingerenza nell'esercizio del diritto del ricorrente alla libertà di espressione sia necessaria in una società democratica.

Infine, l'ultimo step di decisioni riguardanti il discorso d'odio si connota di sfumature omofobiche. Nota è la sentenza *Vejdeland and others v. Sweden* del 9 febbraio 2012²⁰, primo caso di hate speech correlato a propaganda omofobica. Nel dicembre 2004 i richiedenti, insieme ad altre tre persone, si recano in una scuola secondaria superiore e distribuiscono circa un centinaio di opuscoli lasciandoli all'interno o sopra gli armadietti degli alunni. L'episodio si conclude quando il preside della scuola interviene e li fa uscire dall'edificio. L'autore degli opuscoli è un'organizzazione denominata National Youth e gli opuscoli contengono, tra l'altro, le seguenti dichiarazioni:

Propaganda Omosessuale: Nel corso di alcuni decenni la società ha oscillato dal rifiuto dell'omosessualità e di

²⁰<http://hudoc.echr.coe.int/eng?i=001-109046> (04/09/2024).

altre deviazioni sessuali per abbracciare questa tendenza sessuale deviante. I vostri insegnanti anti-svedesi fanno molto bene che l'omosessualità ha un effetto moralmente distruttivo sulla sostanza della società e cercheranno volentieri di presentarla come qualcosa di normale e di buono. -- Dica loro che l'HIV e l'AIDS sono apparsi presto con gli omosessuali e che il loro stile di vita promiscuo era una delle ragioni principali per cui questa piaga moderna stava prendendo piede. -- Dire loro che le organizzazioni di lobby omosessuali stanno anche cercando di minimizzare pedofilia, e chiedere se questa deviazione sessuale dovrebbe essere legalizzata.

Per la distribuzione degli opuscoli, i richiedenti sono accusati di istigazione contro un gruppo nazionale o etnico. Questi ultimi, tuttavia, contestano il fatto che il testo degli opuscoli esprime disprezzo per gli omosessuali e sostengono che, in ogni caso, essi non intendono esprimere disprezzo per gli omosessuali in quanto gruppo. Essi hanno dichiarato che lo scopo della loro attività è quello di avviare un dibattito sulla mancanza di obiettività nell'istruzione impartita nelle scuole svedesi. Il 6 luglio 2006 la Corte Suprema condanna i ricorrenti per agitazione contro un gruppo nazionale o etnico, dichiarando che:

Alla luce della giurisprudenza della Corte europea dei diritti dell'uomo relativa all'articolo 10, nell'interpretazione dell'espressione "contempt" nella disposizione relativa all'incitamento contro un gruppo, si dovrebbe procedere a una valutazione globale delle circostanze del caso, ove, in particolare, si debba prendere in considerazione quanto segue. La distribuzione dei volantini avvenne in una scuola. L'imputato non ha avuto libero accesso ai locali, che possono essere considerati un ambiente relativamente riparato per quanto riguarda le azioni politiche degli stranieri. L'inserimento degli opuscoli negli e sugli armadietti degli alunni ha fatto sì che i giovani li ricevessero senza avere la possibilità di decidere se accettarli o meno. Lo scopo della distribuzione degli opuscoli era infatti quello di avviare un dibattito tra alunni e insegnanti su una questione di interesse pubblico, vale a dire l'obiettività dell'istruzione nelle scuole svedesi, e di fornire agli alunni argomenti. Tuttavia, questi sono stati formulati in modo offensivo e denigratorio per gli omosessuali come gruppo e in violazione del dovere di cui all'articolo 10 per evitare, per quanto possibile, dichiarazioni che sono ingiustificatamente offensive per gli altri, costituendo così un'aggressione ai loro diritti, e senza contribuire ad alcuna forma di dibattito pubblico che possa contribuire ad una maggiore comprensione reciproca. Lo scopo delle sezioni pertinenti dei volantini avrebbe potuto essere raggiunto senza dichiarazioni offensive per gli omosessuali come gruppo.

I ricorrenti lamentano che la sentenza della Corte Suprema costituisce una violazione della

loro libertà di espressione, tutelata dall'articolo 10 della Convenzione. La Corte constata che la condanna dei ricorrenti costituisce un'interferenza con la loro libertà di espressione, come garantito dall'articolo 10, paragrafo 1, della Convenzione. Tale interferenza violerebbe la Convenzione se non soddisfacesse i requisiti di cui all'articolo 10, paragrafo 2. Occorre pertanto stabilire se tale interferenza sia prescritta dalla legge, se persegue uno o più degli scopi legittimi enunciati in tale paragrafo e se sia necessaria in una società democratica. Per quanto concerne la legittimità e la finalità legittima, la Corte osserva che i ricorrenti sono condannati per agitazione nei confronti di un gruppo nazionale o etnico conformemente al capitolo 16, articolo 8 del Codice penale svedese che all'epoca del presunto crimine includeva dichiarazioni che minacciavano o manifestavano disprezzo per un gruppo di persone in riferimento al loro orientamento sessuale.

La Corte ritiene pertanto che l'ingerenza contestata sia sufficientemente chiara e prevedibile e che pertanto sia prescritta dalla legge ai sensi della Convenzione. La Corte ritiene inoltre che l'ingerenza abbia un fine legittimo, ossia la tutela della reputazione e dei diritti altrui ai sensi dell'articolo

10, paragrafo 2, CEDU. Spetta inoltre alla Corte valutare se l'ingerenza sia necessaria in una società democratica. I giudici europei concordano con la Corte suprema sul fatto che, anche se si tratta di uno scopo accettabile, si deve tener conto della formulazione degli opuscoli. La Corte europea osserva che, secondo gli opuscoli, l'omosessualità è una deviante inclinazione sessuale che ha un effetto moralmente distruttivo sulla sostanza della società. Gli opuscoli sostengono inoltre che l'omosessualità sia una delle ragioni principali per cui l'HIV e l'AIDS hanno preso piede. Quindi, sebbene tali dichiarazioni non raccomandino direttamente alle persone di commettere atti odiosi, si tratta di accuse gravi e pregiudizievoli.

Inoltre, la Corte ribadisce che l'incitamento all'odio non comporta necessariamente la richiesta di un atto di violenza o di altri atti criminali. Attacchi a persone commesse per oltraggio, l'intralcio o la calunnia di gruppi specifici della popolazione possono essere sufficienti per consentire alle autorità di favorire la lotta contro le espressioni razziste di fronte alla libertà di espressione esercitata in modo irresponsabile. A questo proposito, la Corte sottolinea che la discriminazione basata sull'orientamento sessuale è tanto grave quanto la discriminazione basata sulla razza,

sull'origine o sul colore. I giudici prendono inoltre in considerazione il fatto che i foglietti illustrativi sono lasciati negli armadietti dei giovani che sono in età impressionabile e sensibile e che non hanno alcuna possibilità di rifiutare di accettarli. La Corte ritiene che la condanna dei ricorrenti e le pene loro inflitte non siano sproporzionate rispetto all'obiettivo legittimo perseguito e che i motivi adottati dalla Corte Suprema a giustificazione di tali misure siano pertinenti e sufficienti. L'ingerenza dei ricorrenti nell'esercizio del loro diritto alla libertà di espressione potrebbe quindi essere ragionevolmente considerata dalle autorità nazionali come necessaria in una società democratica per la tutela della reputazione e dei diritti degli altri. Le considerazioni che precedono sono sufficienti per consentire alla Corte di concludere che il ricorso non costituisce una violazione dell'articolo 10 della Convenzione.

Alla luce delle sentenze esaminate, è possibile rilevare come, nonostante non manchi chi tenti di attribuire una funzione costituzionale ai giudici di Strasburgo (Pollicino e Sciarabba 2001), la Corte EDU, in realtà, privilegia puramente o esclusivamente la pretesa soggettiva di un individuo. Il ricorrente agisce *uti singulus*, non come cittadino: la domanda non presuppone il sinallagma

propriamente sotteso al rapporto di cittadinanza, che assegna diritti ma richiede contestualmente l'adempimento di specifici doveri nei confronti della comunità. Il sistema convenzionale è una "giurisdizione dei diritti individuali contro lo Stato, nel senso polemico di essere contro l'idea di comunità politica", avente come fine l'inveramento di un autonomo ordinamento internazionale su scala regionale²¹. Le condotte potenzialmente eversive delle coordinate valoriali della Convenzione rappresentano un abuso che, per la Corte EDU, gli stati nazionali devono reprimere. In tal senso, la giurisprudenza della Corte di Strasburgo nei confronti dell'intollerante attenua i processi nazionali di condivisione democratica alimentati dalla libertà di espressione, rispondendo invece alle esigenze di conservazione di un ordine convenzionale disallineato, per struttura e finalità, dalla Costituzione, in perenne tensione con le promesse di libertà su cui si fondano gli ordinamenti nazionali.

1.5. La Corte costituzionale italiana

Da un punto di vista costituzionale, il tema della lotta al discorso d'odio può sintetizzarsi in

²¹https://www.rivistaaic.it/images/rivista/pdf/3_2014_Morone.pdf (6.2.2025)

quello che è conosciuto come il “paradosso di Böckenförde”, secondo il quale lo Stato liberale secolarizzato vive di presupposti che non può garantire. Infatti, secondo alcuni, la scelta del legislatore europeo di intervenire per depurare la rete dai discorsi d’odio sembrerebbe evocare l’adozione di una serie di limitazioni che costringono la libertà di parola entro limiti probabilmente più rigorosi rispetto a quelli codificati dalle costituzioni liberali. Così operando, per evitare il propagarsi di odio online, e in nome della protezione di diritti costituzionalmente rilevanti, si potrebbe giungere a limitare una libertà. Esattamente quel presupposto di cui lo Stato liberale non potrebbe farsi garante in maniera duratura nel tempo. Da ciò si capisce come gli interventi normativi sulla materia richiedano necessariamente un confronto con lo statuto costituzionale della libertà in questione.

Nonostante l’indubbia complessità dello scenario oggetto di analisi, si può tuttavia tentare di ridurre il dibattito sul tema essenzialmente ad una dicotomia che oppone, rispetto alla circolazione di idee e opinioni, l’opzione di una democrazia militante, impegnata nella strenua difesa di un sistema valoriale che l’esercizio degli stessi diritti costituzionalmente tutelati rischia di mettere a re-

pentaglio, e il modello di una democrazia tollerante, dove l'idea di uno Stato etico o comunque di un maggior controllo sull'esercizio delle libertà da parte degli individui appare svanire²².

La Costituzione italiana in particolare, se da un lato ai sensi dell'articolo 21 tutela il principio di libertà di manifestazione del pensiero, dall'altro, ai sensi degli articoli 2 e 3, tutela la dignità sociale, i diritti inviolabili del singolo e delle formazioni sociali e afferma il principio di eguaglianza e non discriminazione. Come possono tali principi coesistere in materia di odio online?

Nonostante la Corte costituzionale italiana rappresenti un'eccezione nel panorama dei giudici costituzionali (i quali continuano a porre in essere un'operazione che può essere definita di "constitutional accommodation" (Spigno 2018), dal momento che non è mai stata investita della questione di costituzionalità relativa alle leggi che disciplinano l'hate speech, in particolare la legge n. 364 del 1975, Giuseppe Nicastro, magistrato attualmente svolgente la propria funzione presso la Suprema Corte di Cassazione²³, interviene nel merito

²²<https://www.medialaws.eu/wp-content/uploads/2019/05/2.-Pollicino.pdf> (7.2.2025)

²³https://www.cortecostituzionale.it/documenti/convegni_seminari/stu_284.pdf (03/09/2024).

constatando che la giurisprudenza della Corte costituzionale ha costantemente affermato la valenza centrale della libertà di manifestazione del pensiero nel nostro ordinamento costituzionale. Tra le libertà fondamentali, la libertà di espressione del pensiero è probabilmente quella che nel nostro sistema di giustizia costituzionale, ha avuto maggiore fortuna. Tale fortuna si può misurare sia in termini quantitativi, con riferimento al numero e alla frequenza delle pronunce della Corte con riferimento all'articolo 21 Cost., sia in termini qualitativi, con riferimento al rilievo e alla varietà di indirizzi che la stessa Corte ha potuto adottare nel corso di questi cinquant'anni.

A dire il vero, la storia della nostra Corte è iniziata proprio con una sentenza (sentenza n. 1 del 14 giugno 1956²⁴) volta a sanzionare l'incostituzionalità, per la violazione della libertà di espressione del pensiero, di alcune norme del Testo Unico sulla sicurezza, limitative delle pubbliche affissioni. A questa prima pronuncia sono seguite nel corso degli anni numerose altre sentenze in tema di articolo 21 Cost., che hanno investito tutti i profili della libertà di espressione (contenuti,

²⁴<https://www.cortecostituzionale.it/actionSchedaPronuncia.do?anno=1956&numero=1> (03/09/2024).

limiti, mezzi di esercizio), con riferimento in particolare alla comunicazione intersoggettiva di massa (Abbamonte 2017). Nella sentenza n. 9 del 1965, la Corte inserisce la libertà di espressione del pensiero

tra le libertà fondamentali proclamate e protette dalla nostra Costituzione, una di quelle che meglio caratterizzano il regime vigente nello Stato, condizione com'è del modo di essere e dello sviluppo della vita del Paese in ogni suo aspetto culturale, politico, sociale

sottolineando, in occasione della sentenza n. 168 del 1971, che il diritto previsto dall'art. 21 Cost. è "il più alto, forse, dei diritti primari e fondamentali" sanciti dalla Costituzione. La libertà di manifestazione del pensiero rientra quindi, secondo la Corte, tra i "diritti inviolabili dell'uomo" di cui all'art. 2 Cost. (sentenza n. 126 del 1985). Ciò implica che la Repubblica, da un lato, ha il dovere di garantirla anche nei confronti dei privati e, dall'altro, che quest'ultima non può venire soppressa. La stessa Corte, nella sentenza n. 84 del 1969, delinea la libertà di espressione come la "pietra angolare dell'ordine democratico".

Secondo lo stesso Nicastro, i giudici costituzionali non vogliono dipingere la libertà di espressione del pensiero come una conseguenza della

democrazia bensì come il suo fondamento. Da ciò si deduce come la Corte abbia voluto dare a tale libertà una lettura in chiave individualistica o liberale, per la quale il diritto di manifestare il proprio pensiero è attribuito all'uomo in quanto tale e a suo vantaggio, non all'uomo nella qualità di membro di una comunità e a vantaggio della stessa e dei suoi valori. Ciò sembra vero, in particolare, con riferimento alla giurisprudenza in tema di rapporto tra libertà di manifestazione del pensiero e diritti della personalità. Già nella sentenza n. 9 del 1965, la Corte afferma che, essendo la libertà di manifestazione del pensiero tra le libertà fondamentali proclamate e protette dalla Costituzione,

ne consegue che limitazioni sostanziali di questa libertà non possono essere poste se non per legge e devono trovare fondamento in precetti e principi costituzionali, si rinvengano essi esplicitamente enunciati nella Carta costituzionale o si possano, invece, trarre da questa mediante la rigorosa applicazione delle regole dell'interpretazione giuridica.

Dunque, secondo la Corte, le limitazioni alla libertà di manifestazione del pensiero devono: a) essere poste dal (solo) legislatore ordinario; b) tutelare beni costituzionalmente rilevanti (individuati dal legislatore sotto il controllo della stessa Corte).

Tuttavia, con lo scorrere del tempo, la Corte muta atteggiamento e decide di avvicinarsi ad una lettura in chiave funzionalistica del rapporto tra libertà di espressione del pensiero e altri diritti costituzionalmente protetti, soprattutto il diritto di informare e di essere informati. Si citano qui le sentenze n. 157 del 1987 e n. 112 del 1993²⁵, quest'ultima sviluppo della prima.

Come analizzato esaustivamente da Gianluca Gardini (2017), coloro che intendono la libertà di manifestazione del pensiero come un valore supremo dell'individuo rispetto al tutto, una qualità o una proprietà della persona, accettano con maggiore difficoltà l'idea di limitazioni, di controlli e di una regolazione penetrante rispetto a un diritto che, in primo luogo, rappresenta un valore individuale, da tutelare in sé. Da ciò emerge che i vincoli e le limitazioni posti nell'interesse generale, per esempio, per la salvaguardia dell'ordine pubblico, la correttezza del procedimento elettorale, l'egualianza delle opportunità informative, non vengono tollerati nella prospettiva individualistica.

²⁵ Questa Corte ha costantemente affermato che la Costituzione, all'art. 21, riconosce e garantisce a tutti la libertà di manifestare il proprio pensiero con qualsiasi mezzo di diffusione e che tale libertà ricomprende tanto il diritto di informare, quanto il diritto di essere informati.

All'opposto, coloro che vedono nella libertà di manifestazione del pensiero un valore supremo del tutto rispetto all'individuo, sono pronti ad accettare le limitazioni che si giustificano nel nome delle esigenze della collettività. La libertà di espressione, in questa prospettiva, spetta all'uomo nella sua qualità di membro di una comunità, per la funzione che in essa il singolo deve esplicare, e come tale può essere assoggettata a tutti i vincoli e le restrizioni che al bene comune si ricollegano.

1.6. Gli interventi americani

Nell'ordinamento statunitense il primo grande strumento di tutela della libertà di espressione del pensiero è costituito dalla "free speech clause" contenuta nel Primo Emendamento²⁶ della Costituzione Americana. Roger Kiska, studioso di diritto, nonché avvocato e membro eletto del gruppo consultivo dell'Agenzia dell'Unione Europea per i diritti fondamentali, sostiene come il Primo Emendamento abbia da tempo dato al popolo americano libertà nel campo della parola e

²⁶ Il Congresso non promulgherà leggi per il riconoscimento ufficiale di una religione, o che ne proibiscano la libera professione; o che limitino la libertà di parola, o di stampa; o il diritto delle persone di riunirsi pacificamente in assemblea e di fare petizioni al governo per la riparazione dei torti.

dell'espressione. Tale disposizione, infatti, racchiude l'inevitabile rapporto tra il libero flusso di informazioni e un popolo autogovernato²⁷.

Dal canto loro, come vedremo, gli stessi tribunali americani non hanno esitato a rimuovere gli ostacoli che limitano questo flusso. Radicata nella democrazia americana, infatti, è la ferma convinzione che la saggezza e la giustizia siano più propense a prevalere nel processo decisionale pubblico se tutte le idee, le scoperte e i punti di vista sono chiaramente esposti davanti alla gente per la loro considerazione. Il Presidente della Corte Suprema degli Stati Uniti Roberts ha affermato che:

Il discorso è potente. Può incitare le persone ad agire, a piangere di gioia e di dolore. Noi non possiamo reagire al dolore inflitto punendo l'oratore. Come nazione abbiamo scelto un diverso percorso, per proteggere anche il discorso offensivo su questioni pubbliche per garantire che non soffochiamo il dibattito pubblico.

Allo stesso tempo, i tribunali americani hanno riconosciuto che non tutte le espressioni illuminano il corpo politico e che alcune parole sono capaci di causare gravi danni. Sotto lo stato attuale

²⁷https://www.regent.edu/acad/schlaw/student_life/studentorgs/lawreview/docs/issues/v25n1/04Kiskavol.25.1.pdf
(9.2.2025)

della legge, vi sono solo tre tipi di discorso che sono costituzionalmente vietati: oscenità, diffamazione, e discorso che crea un pericolo attuale. Nel decidere se un determinato discorso sia oggetto di una restrizione da parte del Governo nell'osservanza della Costituzione americana, viene rappresentato un quadro giuridico in tre fasi. In primo luogo, ci si deve chiedere se il discorso sia protetto dal Primo emendamento; in secondo luogo, se la natura del "forum" o luogo in cui si verifica il discorso, è identificata e, infine, si deve valutare se la giustificazione della restrizione del discorso soddisfa "il requisito standard". La Corte Suprema degli Stati Uniti, quindi, ritiene che il diritto alla libertà di espressione del pensiero, protetto dal Primo Emendamento della Costituzione Americana, vada limitato e soppresso solo nel caso estremo in cui sussista un "clear and present danger", cioè una minaccia concreta e motivata, idonea a causare violenza. Per gli americani, infatti, perché si limiti il principio cardine costituito dalla libertà nella manifestazione del proprio pensiero, occorre che il cosiddetto clear and present danger presenti tre requisiti: a) intenzionalità a causare un pericolo di danno; b) imminenza del danno; c) idoneità della minaccia a portare ad un pericolo.

Secondo Julien Mailland, ricercatore presso il Computer History Museum, la ratio alla base della cosiddetta “free speech clause” inserita all’interno del Primo Emendamento della Costituzione americana, può essere compresa in toto dai giuristi europei attraverso la visione del film “The Blues Brothers” di John Landis²⁸. Il film riprende la storia del caso Skokie²⁹, emblema del trionfo neonazista e della vittoria di coloro che volevano diffondere odio nella società. Questo caso nasce da una controversia del 1977 relativa al Partito Socialista Nazionale d’America (NSPA) a Skokie, Chicago. Skokie era, a quel tempo, un villaggio con una popolazione ebraica del 57% e alcuni dei suoi residenti erano sopravvissuti ai campi di concentramento nazisti. Il leader del partito NSPA, Frank Collin, che descrive il partito come una “organizzazione nazista”, propone di organizzare una manifestazione pubblica pacifica per protestare contro le norme sull’uso dei parchi pubblici del villaggio per le assemblee politiche. La dimostrazione proposta avrebbe avuto una durata di circa 30 minuti e sarebbe poi di fatto consistita nella marcia di 30-50 dimostranti davanti alla sala del villaggio. Era

²⁸https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2327305 (9.2.2025)

²⁹<https://globalfreedomofexpression.columbia.edu/cases/national-socialist-party-america-v-village-skokie/> (04/09/2024).

previsto inoltre che i manifestanti indossassero l'uniforme della festa ed esibissero degli striscioni (in entrambi i casi mostrando espressamente la svastica) e apportassero delle variazioni sulla dichiarazione *Discorso libero per l'uomo bianco*. In seguito, la Corte del Circuito della Contea di Cook emana un'ingiunzione che proibisce di

marcarsi, camminare o sfilare nell'uniforme della NSPA; inarcarsi, camminare o sfilare o altrimenti mostrare la svastica dentro o fuori la propria persona; la distribuzione di opuscoli o l'esposizione di qualsiasi materiale che inciti o promuova l'odio contro persone di fede o origine ebraica o odio contro persone di qualsiasi fede o lignaggio, razza o religione all'interno del villaggio di Skokie.

In seguito alla decisione assunta dalla Corte Suprema dell'Illinois, che conferma il divieto di manifestare deciso dalle Corti inferiori in considerazione della probabilità del pericolo derivante da una simile marcia, il caso giunge alla Corte Suprema. La Corte (con decisione assunta a stretta maggioranza di 5 giudici contro 4), pur non affermando che i nazisti abbiano un diritto costituzionalmente garantito di marciare a Skokie, dichiara che essi hanno però un diritto costituzionale a non essere soggetti a prior restraints, e quindi, "teoricamente" (nei fatti poi la marcia non ha luogo ed

effettivamente il gruppo sembra più che altro strumentalizzare la situazione per sollecitare la polemica sul proprio diritto a manifestare) ammette lo svolgimento dell'iniziativa.

Nel film, una vecchia auto della polizia è bloccata nel traffico prima dell'accesso ad un ponte dove è in corso la manifestazione dei neonazisti. Nella macchina i due uomini vestiti di nero, Jake ed Elwood, iniziano a spazientirsi mentre il leader dei neonazisti tiene un discorso dai caratteri violenti e razzisti. Un cordone di agenti di polizia protegge i manifestanti dalla folla che li vorrebbe aggredire. Uno dei due Blues Brothers dice: "Odio i nazisti dell'Illinois". L'altro accelera e investe i neonazisti, facendoli cadere nel fiume sottostante.

Da questo caso, sapientemente narrato nel film, si possono dedurre i due pilastri che connotano l'approccio nordamericano all'hate speech. Il primo consiste nell'idea di "Marketplace of ideas", secondo la quale, solo lasciando che tutte le idee vengano espresse senza l'intervento da parte del Governo, si riesce ad ottenere l'idea migliore. Nel caso di specie, tale principio viene incarnato dalle parole usate da Aryeh Neier, avvocato di origine ebraiche scelto proprio per difendere il diritto dei

neonazisti di manifestare. Egli afferma che: “Although freedom has its risks, [s]uppression of freedom ... is a sure prescription for disaster”³⁰. A tal proposito, Erik Asard e W. Lance Bennett, nel loro libro *Democracy and the Marketplace of Ideas*, compiono un’analisi dettagliata circa la correlazione tra la politica e il mercato delle idee, sostenendo come la qualità delle idee politiche in molte nazioni si sia abbassata e abbia portato all’utilizzo di slogan semplicistici e di una retorica connotata da rabbia e odio (Asard e Bennett 1997). Il secondo pilastro che connota la giurisprudenza nordamericana è invece costituito dal principio secondo il quale il Governo non dovrebbe mai interferire con l’attività umana. Bisogna infatti evitare che il Governo stabilisca una “guardianship of the public mind” regolando stampa, parlato e religione. Si parla a tal proposito di “freedom from the government”.

Al contrario, Richard Delgado, mettendo a confronto da un lato il caso Skokie, nel quale al discorso razzista è attribuita piena protezione, dall’altro casi dove è stato riconosciuto il diritto al risarcimento del danno a favore delle vittime di molestie razziali sul luogo di lavoro e altrove,

³⁰<https://digitalcommons.law.wne.edu/cgi/viewcontent.cgi?article=1574&context=lawreview> (04/09/2024).

spiega perché la legge dovrebbe punire il discorso d'odio, in particolare il discorso razzista. Egli individua le prove che il denunciante deve fornire al fine di dimostrare la sussistenza di un discorso di tal specie. Il ricorrente deve dimostrare che a) quel linguaggio è espressamente indirizzato a lui da parte del convenuto al fine di umiliarlo attraverso il riferimento alla razza; b) è in grado di capire che quel linguaggio ha lo scopo di umiliarlo facendo riferimento alla razza; c) qualsiasi persona ragionevole capirebbe che si tratta di un insulto razzista (Delgado 1982).

In relazione al tema della libertà di espressione del pensiero, oltre al Primo Emendamento della Costituzione Americana, occorre altresì menzionare la Dichiarazione Americana dei diritti e dei doveri dell'uomo (nota anche come Dichiarazione di Bogotà), che anticipa, di poco più di sei mesi, la Dichiarazione universale dei diritti dell'uomo. In particolare, l'art. IV della Dichiarazione sancisce che "ogni persona ha diritto alla libertà di espressione e di pensiero attraverso qualsiasi mezzo". D'altro canto, la Convenzione Americana dei diritti umani, nota anche come Patto di San Josè, entrata in vigore nel 1978, ai sensi dell'articolo 13, prevede che:

1. Ognuno ha diritto alla libertà di pensiero e di espressione. Tale diritto include la libertà di ricercare, ricevere e trasmettere informazioni e idee di ogni tipo, senza considerazione di frontiera, oralmente o per iscritto, attraverso la stampa, in forma artistica o attraverso qualunque altro mezzo di propria scelta. 2. L'esercizio del diritto di cui al paragrafo precedente non è soggetto a censura preventiva, ma sarà motivo di responsabilità successiva, come stabilito espressamente dalla legge nella misura necessaria ad assicurare: a) il rispetto dei diritti e della reputazione di altri; b) la protezione della sicurezza nazionale, dell'ordine pubblico o della salute o della morale pubbliche. 3. Il diritto di espressione non può essere limitato con metodi o mezzi indiretti, quali l'abuso di controlli pubblici o privati sulla stampa periodica, sulle frequenze per le trasmissioni radio, o sulle strumentazioni per la diffusione dell'informazione, o con ogni altro mezzo che tenda ad impedire la comunicazione e la circolazione di idee e opinioni. 4. Fermo restando quanto previsto dal paragrafo 2, gli spettacoli pubblici possono essere sottoposti da parte della legge a forme di censura preventiva al solo scopo di regolarne l'accesso per proteggere la morale dell'infanzia e dell'adolescenza. 5. Qualunque propaganda in favore della guerra e qualunque richiamo all'odio nazionale, razziale o religioso che costituisca incitamento alla violenza illegale o ad ogni altra azione simile contro qualunque persona o gruppo di persone per qualsiasi ragione, compresi motivi di razza, colore, religione, lingua o origine nazionale o sociale, deve essere considerato dalla legge come reato.

1.7 Conclusioni

Dal quadro delineato, emerge che molteplici sono stati i tentativi non solo di definizione del discorso d'odio ma anche di ricerca giurisprudenziale di un equilibrio tra il diritto sancito a livello costituzionale della libertà di espressione del pensiero e la tutela di altri diritti ritenuti egualmente meritevoli di tutela da parte degli ordinamenti giuridici, in primis il diritto a non essere discriminati (in ossequio al principio di eguaglianza tra consociati). Tuttavia, è bene sottolineare come tali operazioni siano state condotte in modo assolutamente diversificato tra Europa e Stati Uniti. In particolare, la prima ha cercato sempre di più nel corso del tempo di anticipare la soglia di punibilità del discorso d'odio e di estendere tale punibilità a tutti i comportamenti discriminatori, fossero questi ultimi di stampo razzista, religioso o sessuale, attraverso l'applicazione del diritto penale da parte dei singoli ordinamenti nazionali.

Il legislatore americano invece ha preferito optare, in ossequio ai principi cardine nell'ordinamento statunitense di "self-government" e di "marketplace of ideas", per la costante applicazione del Primo Emendamento. Tali differenti approcci hanno quindi portato i giudici, sia a livello sovranazionale che nazionale, a fornire una lettura

funzionalistica della libertà di espressione del pensiero, punendo così tutti coloro che abbiano incitato all'odio, sia per mezzo della parola che attraverso altri strumenti di comunicazione, inclusa la rete. Al contrario, i giudici americani si sono dimostrati restii nel consentire la limitazione della libertà di espressione del pensiero, limitazione che può essere imposta solo in presenza di specifici requisiti, in applicazione del cosiddetto "clear and present danger test".

Capitolo Secondo

Rimedi: Censura o Counterspeech

2.1 Il Principio di Neutralità

Come abbiamo avuto modo di apprendere analizzando la normativa e la giurisprudenza che, nel corso del tempo, hanno tentato di disciplinare il fenomeno dell'hate speech, in quest'ambito l'Europa sceglie di adottare una politica di censura, volta a sanzionare penalmente le condotte odiose poste in essere da un individuo determinato o da un gruppo di persone che costituiscono la maggioranza, contro un altro individuo o gruppo di persone che rappresentano una minoranza all'interno del tessuto sociale.

La domanda che molti si pongono è la ragione per cui l'Europa abbia scelto la linea della repressione giuridica del discorso d'odio, contrapponendosi sul punto alla politica del cosiddetto "marketplace of ideas" sostenuta dagli Stati Uniti. I paesi europei infatti, hanno cercato, sin dal dopoguerra, di trovare un equilibrio tra la tutela della libertà di espressione del pensiero e la lotta contro il riemergere di ideologie totalitarie contrarie ai principi democratici. Ciò giustifica, secondo la

maggioranza dei letterati, il ricorso alla legge penale³¹, a dispetto di altri (Pugiotto 2013,1-6) che, al contrario, invocano l'utilizzo di strumenti alternativi, uno su tutti il ricorso alla giustizia riparativa (Walters, Hoyle 2017, 249-266), al fine di combattere il fenomeno dell'hate speech.

Per quanto concerne la prima, è bene ricordare infatti che la previsione di una sanzione penale in caso di incitamento all'odio nasce, in contesti determinati, per proteggere la razza (o religione, o nazionalità) maggioritaria o privilegiata; e che anche là dove si è riscattata da questa ipoteca, viene spesso usata per punire le contestazioni dello status quo provenienti dalle minoranze svantaggiate. Quando l'idea dell'eguaglianza si affaccia nella prassi dei soggetti sociali porta inevitabilmente con sé una forma di ostilità verso i gruppi privilegiati. In passato questi hanno potuto limitarsi a ribadire le ragioni della propria superiorità, e ad accusare gli altri di ribellione all'ordine costituito. La propaganda egualitaria è stata così equiparata al discorso sovversivo: dagli Stati schiavisti che proibivano la propaganda abolizionista, alla

³¹<https://www.osce.org/files/f/documents/3/e/36426.pdf>
(21/09/2024).

Germania imperiale e all'Italia fascista che proibivano l'incitamento all'odio di classe, come vere e proprie forme di sedizione.

In altri contesti l'incitamento all'odio è stato invece considerato un attentato all'unità della Nazione, al sentimento di fratellanza che deve regnare all'interno della collettività, superando le divisioni e i conflitti che la attraversano. Oppure, meno retoricamente, è stato visto come un pericolo per l'ordine pubblico, inteso a ricomprendere i presupposti ideali, emotivi e sentimentali dell'armonia collettiva. In questi casi la disciplina non è più apparsa a senso unico, ma si è rivolta imparzialmente a tutti i gruppi o parti della popolazione, senza distinguere tra i dominanti e i discriminati, che pure si trovano in posizioni assai diverse. L'attenzione che gli ordinamenti contemporanei dedicano ai profili sostanziali dell'eguaglianza pone oggi in tutt'altra luce l'incitamento all'odio razziale: non solo come possibile fonte di disordini e di violenze, ma anche come causa della perdurante inferiorità di fatto in cui versano determinati gruppi. L'esigenza di riscattare la sorte di questi ultimi, creando un ambiente dal quale siano banditi il pregiudizio e la discriminazione, rappresenta dunque un obiettivo immanente (Ferrara 1994, 27 ss.).

Secondo i fondamenti del diritto penale, perché si concretizzi una fattispecie di reato e quindi avvenga la sua repressione mediante l'applicazione di una pena, occorre che vi sia stata, o abbia potuto potenzialmente esserci (in questo caso ci ritroviamo nell'ambito dei reati di pericolo), la lesione di un bene ritenuto meritevole di tutela da parte dell'ordinamento giuridico. Nel caso del discorso d'odio, molti (Tanzarella 2010, 50 ss.) sono inclini a qualificare tali comportamenti come reati di opinione³².

Dal canto suo, lo stesso legislatore pone il discorso odioso all'interno dei reati di pericolo presunto. Infatti, il risultato della pluralità, ambiguità, genericità e indeterminatezza delle concezioni giuridiche dell'hate speech, specie in Europa, è una tendenza a contrastare attraverso gli strumenti del diritto positivo un sempre più ampio catalogo di discorsi reputati "odiosi" in ragione di un'offensività che assai spesso è presunta come inerente alla loro mera espressione o diffusione, indipendentemente dalla necessità di provare che sussista effettivamente un pericolo nel caso concreto. Ciò talora sulla base del presupposto che gli atti linguistici "odiosi" non siano degni di esser

³²<https://riviste.unimi.it/index.php/statoechiese/article/view/2130> (6.2.2025)

qualificati come “discorsi”, che le idee da essi veicolate non siano affatto opinioni o che, quantomeno, vadano espressamente espunte dal novero di quelle protette dalle garanzie della libertà di pensiero e di espressione. Sul punto però sorgono molteplici perplessità. Infatti, come sostenuto da autorevole dottrina³³ il rischio legato all’inserimento dei discorsi d’odio all’interno della categoria dei reati di pericolo astratto è quello non solo di anticipare eccessivamente la tutela penale ma anche e soprattutto di creare un conflitto con principi cardine del diritto penale, quali il principio di oggettività (secondo cui *cogitationis poenam nemo patitur*: nessuno può essere legittimamente punito per aver avuto un “cattivo pensiero”) e il principio di offensività della legge penale (secondo cui il reato deve consistere nell’offesa ad un bene giuridicamente tutelato dall’ordinamento giuridico). Altro problema è dato inoltre dalla possibilità di dedicare la tutela penale alle sole minoranze, derogando così alle regole dell’eguaglianza formale³⁴.

³³<https://iris.unipa.it/retrieve/e3ad8916-047b-da0e-e053-3705fe0a2b96/Spena%2c%20Liberta%27%20di%20espressione.pdf> (1.2.2025)

³⁴https://www.associazionedeicostituzionalisti.it/old_sites/sito_AIC_2003-2010/materiali/anticipazioni/odio_razziale/index.html (9.2.2025)

Per quanto concerne l'esperienza americana invece, i promotori delle leggi contro il discorso d'odio pongono come base legittimante delle norme medesime, il danno che l'hate speech provocherebbe nei confronti dei destinatari del messaggio odioso. In particolare, essi ritengono che tali strumenti normativi servano a ridurre le violenze, le discriminazioni e i danni psichici che scaturiscono dal discorso stesso. In realtà, le leggi di repressione giuridica del discorso d'odio violano due principi (Strossen 2018, 38-39), che peraltro costituiscono colonne portanti del pensiero americano: il principio di Emergenza e il principio di Neutralità. Tali principi sono riassunti in modo efficace dal giudice Marshall nel contesto della sentenza *Chicago v. Mosley* del 1972³⁵. Egli sostiene che

il governo non ha il potere di limitare l'espressione solo per il messaggio che esprime, le idee che diffonde o per il suo contenuto. Il governo non può concedere l'uso di un dibattito a quelle persone le cui idee reputa accettabili, ma negare la stessa possibilità a coloro che desiderano esprimere punti di vista meno favorevoli o più controversi.

³⁵<https://supreme.justia.com/cases/federal/us/408/92/>
(21/09/2024).

In ossequio al principio di Emergenza, infatti, lo Stato deve intervenire per regolare il discorso solo quando quest'ultimo causi direttamente un danno specifico, imminente e serio. Inoltre, alla luce del principio di Neutralità, vincoli legislativi sul linguaggio o su altri atti espressivi, incluse le restrizioni sul discorso d'odio, sono da considerarsi ingiustificati se attuano discriminazioni di contenuto o di punti di vista.

Sia la Corte Suprema degli Stati Uniti che la Corte Europea dei Diritti umani fanno proprio quest'ultimo principio. Per quanto riguarda la prima, nella sentenza *R.A.V. v. City of St. Paul*³⁶, la maggioranza dei giudici sancisce l'incostituzionalità della cosiddetta "City of St. Paul Bias-Motivated Crime Ordinance", considerandola non sufficientemente inclusiva dal momento che vieta solo le parole odiose che presentano un certo contenuto e, una volta applicata, proibisce solo determinati punti di vista.

Dal canto suo, la Corte Europea applica la dottrina della neutralità attraverso un susseguirsi nel tempo di una serie di pronunce (*Handyside v. United Kingdom*, *Soulas and others v. France*³⁷ e

³⁶<https://globalfreedomofexpression.columbia.edu/cases/r-v-v-city-st-paul/> (21/09/2024).

³⁷[https://hudoc.echr.coe.int/eng#{%22itemid%22:\[%22001-87370%22\]}](https://hudoc.echr.coe.int/eng#{%22itemid%22:[%22001-87370%22]}) (21/09/2024).

Vejdeland and others v. Sweden³⁸). In particolare, dalla sentenza Handyside v. United Kingdom, apprendiamo che

il diritto alla libertà di espressione è applicabile non solo alle informazioni o alle idee che sono recepite o considerate inoffensive o ininfluenti, ma anche a quelle idee che offendono, provocano uno shock o disturbano lo Stato o qualsiasi settore della popolazione. Tali sono le richieste di quel pluralismo, quella tolleranza e quell'apertura mentale senza i quali non si può parlare di società democratica.

Tuttavia, nonostante l'adozione di tale linea di pensiero da parte della Corte Europea, molti sono i sostenitori dello strumento legislativo volto, anticipando la soglia di punibilità del reato d'odio al mero pericolo, a sanzionare penalmente le condotte odiose. Perciò ora andiamo ad esaminare quale sia o quali siano i beni che il legislatore vorrebbe tutelare attraverso il suo intervento.

2.2. Il bene Dignità

Una parte della dottrina americana (Waldron 2012, 105-143), la quale si pone a favore dell'intervento dello Stato ai fini della censura del discorso d'odio, sostiene che, attraverso il perpetrarsi di tali espressioni, viene potenzialmente lesa

³⁸[https://hudoc.echr.coe.int/eng#{%22itemid%22:\[%22001-109046%22\]}](https://hudoc.echr.coe.int/eng#{%22itemid%22:[%22001-109046%22]}) (21/09/2024).

la dignità dell'individuo. In primis, è necessario capire cosa si debba intendere quando parliamo di dignità. Il termine dignità acquisisce nella letteratura un duplice significato. In alcuni casi essa rivela come dignità umana, in altri come dignità civica. In via preliminare, pare tuttavia giusto precisare che l'idea della eguale dignità di tutti gli esseri umani è una conquista assai recente, giacché le società premoderne erano di norma strutturate in ordini gerarchici implicanti delle nozioni di persona, dignità e rispetto intesi come concetti scalari anziché come concetti qualificatori tutto-o-niente; un accostamento che induceva a graduare la dignità degli individui sulla base di vari elementi sociali, razziali o anagrafici (genere, età, famiglia, luogo di provenienza, ecc.) (Galeotti 2010, 114 ss.). Per quanto concerne il concetto di dignità umana, è da segnalare l'articolo 130 del Codice penale tedesco che punisce il *Volksverhetzung*, inteso come disturbo della pace pubblica concretantesi tra l'altro in insulti e denigrazioni lesive della dignità umana di gruppi nazionali, razziali, religiosi o definiti sulla base dell'origine etnica.

Dal canto suo, la giurisprudenza italiana offre una rappresentazione esemplare dell'incertezza sulle ragioni della criminalizzazione dell'hate speech oscillando tra due alternative: vi

sono sentenze di merito che riscontrano nella diffusione/propaganda dei discorsi d'odio

un concreto turbamento della coesistenza pacifica dei vari gruppi etnici nel contesto sociale al quale il messaggio era indirizzato,

salvo in seguito, addirittura nelle pronunce successive sulla stessa vicenda, escludere recisamente l'ordine pubblico dalla proiezione offensiva della condotta espressiva e rimarcare piuttosto la

dimensione concreta e personalistica del bene protetto consistente nella dignità di ciascun uomo

precisando infine nelle sedi di legittimità che

l'oggetto specifico della tutela penale non era e non è costituito dall'ordine pubblico, il quale ha rilevanza indiretta, ma dalla tutela della dignità umana come risulta dalla nozione di discriminazione recepita dall'articolo 2 del decreto legislativo n. 215 del 2003 nel quale si fa esplicito riferimento alla dignità della persona.

A tal proposito, Whitman sostiene che l'insistenza dei diritti europei sul valore della dignità non è, come si pensa comunemente, soltanto una reazione al nazifascismo che settant'anni fa la conculcò in maniera così atroce, ma è il prodotto di una reazione contro le società gerarchiche prece-

denti alla Rivoluzione francese, le quali riconoscevano dignità, onore e rispetto soltanto agli appartenenti alle classi privilegiate, anziché a chiunque (Whitman 2004, 17 ss.).

Tuttavia, coloro che giustificano la criminalizzazione del discorso d'odio ponendo come ratio alla base del proprio ragionamento la presunta lesione della dignità umana non tengono conto dei termini generici, vaghi, controversi e incerti che vengono utilizzati al fine di definire il concetto di dignità umana. La genericità di tali termini fa sì che il concetto di dignità si presti a giustificare qualsiasi intervento pubblico rispetto al quale non si riesca a identificare quale oggetto di tutela un bene giuridico più specifico: una nozione tanto solenne, trascendente e a vocazione egemonica quanto impalpabile, a cagione della sua capacità di racchiudere in sé innumerevoli e indefiniti contenuti (Becchi 2013).

In merito, alcuni studiosi³⁹ notano che il discorso sulla dignità umana ha progressivamente acquisito lo status privilegiato di una “lingua franca etica”, e che esso presenta un notevole potenziale retorico-evocativo cui viene, quasi istinti-

³⁹https://d1vbhhqv6ow083.cloudfront.net/contributi/Go-metz.M_Lodio.pdf (1.2.2025)

vamente, associata una connotazione emotiva favorevole⁴⁰. Inoltre, punire una manifestazione di pensiero in nome della tutela della dignità umana introduce un cortocircuito, perché anche l'esercizio della libertà d'opinione (che si vorrebbe limitare) è essa stessa esplicazione della dignità umana individuale della persona concretamente accusata di diffondere discorsi d'odio in contrasto con un'idea collettiva di dignità umana (alla quale facciamo riferimento soprattutto con riguardo ai "nuovi" reati di opinione in tema di negazionismo, omofobia e xenofobia) (Ambrosi 2007, 533).

I giudici e parte della dottrina tentano altresì di giustificare l'intervento limitativo della libertà di pensiero utilizzando il concetto di dignità umana intesa alternativamente o come meta-principio diretto alla funzionalizzazione dell'ordinamento (Rolla 2008, 68 e ss.), o come principio costituzio-

⁴⁰ Afferma inoltre che "sul piano della ragione pubblica, l'uso in chiave "assolutista" di un argomento ad altissimo tasso di indeterminazione come quello che fa leva sulla dignità umana si rivela, il più delle volte, una manovra razionalmente insoddisfacente (quando non addirittura una manovra più o meno intenzionalmente depistante o confusiva) perché, essendo solitamente "parassitario" rispetto ad altri argomenti soggiacenti, impedisce di discutere l'argomento che resta nascosto"; vedi anche Poggi, "Brevi osservazioni sulla retorica dei diritti" in *Ragion Pratica*, 31, 2008, 366.

nale di non discriminazione desumibile dall'articolo 3 Cost., o ancora, quale diritto inviolabile che troverebbe ingresso attraverso la fattispecie aperta dell'articolo 2 Cost. Tuttavia, considerare la dignità umana quale principio o condizione a priori dell'ordinamento costituzionale avrebbe l'effetto non solo di legittimare qualsiasi tipo di compressione delle libertà costituzionalmente garantite, fino a rendere impossibile qualsiasi operazione procedurale di bilanciamento, ma comporterebbe altresì in via definitiva una fissa sistemazione gerarchica dei principi costituzionali di cui non vi è traccia nel vigente testo costituzionale (Tega 2012). Dworkin (2012, 342), dal canto suo, sostiene che

vivere in una società, il cui governo rispetta la dignità umana, significa che io debba accettare il diritto degli altri ad esprimere disprezzo nei miei riguardi.

Tale comportamento è alla base della teoria dello studioso sulla legittimità politica, secondo la quale gli individui, in quanto esseri totalmente indipendenti dallo Stato, sono capaci di sviluppare le proprie idee sulla dignità umana e sulla moralità. Lo scopo del diritto alla libertà di espressione è di

offrire a coloro che si esprimono, un livello adeguato di protezione dai governanti inclini a sopprimere qualsivoglia idea essi ritengano sbagliata.

In una recente sentenza⁴¹, la Corte Suprema canadese, basando la propria decisione sull'art. 14 comma 1 lettera b) del Saskatchewan Human Rights Code⁴², ha ritenuto che proibire qualsiasi rappresentazione che espone all'odio qualsiasi persona o gruppo di persone rappresenti un limite ragionevole alla libertà di espressione del pensiero. Dal canto suo, l'articolo 3 comma 2 del suddetto codice, specifica che l'obiettivo del testo di legge è quello di assicurare che ogni persona sia libera ed eguale sul piano della propria dignità e dei propri diritti.

In ambito nazionale invece, l'articolo 2 comma 3 del decreto legislativo n. 215 del 2003, emanato in attuazione della direttiva 2000/43/CE

⁴¹<https://scc-csc.lexum.com/scc-csc/scc-csc/en/item/12876/index.do> (21/09/2024)

⁴² 14(1) No person shall publish or display, or cause or permit to be published or displayed, before the public any statement, publication, notice, sign, symbol, emblem or other representation: (a) tending or likely to tend to deprive, abridge or otherwise restrict the enjoyment by any person or class of persons, on the basis of a prohibited ground, of any right to which that person or class of persons is entitled under the law; or (b) that exposes or tends to expose to hatred any person or class of persons on the basis of a prohibited ground.

per la parità di trattamento tra le persone indipendentemente dalla razza e dall'origine etnica, considera discriminazioni quei comportamenti volti a violare la dignità di una persona. Anche le più importanti risoluzioni e dichiarazioni internazionali fanno riferimento alla tutela della dignità umana (Schachter 1983, 848-854). Ma, in realtà, che cosa veramente si intenda per dignità umana non è in alcun modo specificato. Come possiamo quindi giustificare l'intervento pubblico in relazione al fenomeno dell'hate speech sulla base di un concetto così vago e indefinito?

Proviamo quindi a costruire la ratio della repressione giuridica del discorso d'odio dando al concetto di dignità un valore politico. In epoca romana, i latini erano soliti utilizzare il termine *dignitas hominis* ai fini di indicare il prestigio, l'onore riconosciuto a chi rivestiva una determinata carica politica. Oggi Jeremy Waldron parla a tal proposito di dignità civica, intesa come il diritto del singolo a godere di buona reputazione in quanto membro della società, come il diritto di qualcuno che, membro di un gruppo di minoranza, non venga per questo motivo allontanato dalla società perché ritenuto inferiore. Le leggi contro l'hate speech sono proprio pensate, a detta di Waldron, per evitare che

le discriminazioni causate dall'odio, vadano ad alterare l'ambiente sociale in cui queste vengono attuate, favorendo così le disuguaglianze tra i cittadini. Tuttavia, lo studioso, nell'espone la propria teoria, fa riferimento a monumenti, scritte e altri segni tangibili di odio e disprezzo, non tenendo in considerazione l'impatto che gli insulti possano avere sui destinatari di tali epiteti (Delgado 2013, 232-233).

A questo punto, sorge spontaneo chiedersi se l'intervento ex post del legislatore per punire il discorso d'odio, il quale incita a comportamenti discriminatori nel momento stesso in cui viene pronunciato, non possa risultare tardivo e quindi non massimamente efficace in relazione all'effetto di alterazione dell'ambiente sociale in cui questo sorge. Waldron sostiene altresì che alcuni usi dell'hate speech rappresentino una minaccia alla stima riconosciuta ai soggetti appartenenti ai gruppi destinatari del discorso d'odio da parte degli altri cittadini e dell'ordinamento giuridico.

Il vero problema, a mio avviso, è legato alla percezione da parte dei soggetti "odiati" delle parole minacciose che gli altri individui destinano loro. L'elemento percezione è, per definizione, esclusivamente soggettivo, come tale variabile da soggetto a soggetto (Post 2018, 268-269). Può

quindi tale elemento ritenersi sufficiente per giustificare l'intervento repressivo dello Stato finalizzato a proteggere la dignità del singolo?

2.3. Il Principio di Garanzia

Per rispondere a questa domanda, è necessario introdurre un altro principio che costituisce una colonna portante all'interno del fenomeno hate speech: il principio di Garanzia. Per principio di Garanzia si intendono quei vincoli legali che, applicati al discorso d'odio, sono ritenuti necessari nella misura in cui proteggono il bene pubblico costituito dalla stessa garanzia. Ma garanzia di cosa? A riguardo, Jeremy Waldron parla appunto di garanzia di dignità civica. Tale protezione si traduce in primo luogo nel riconoscimento, nei confronti di coloro che appartengono a gruppi di minoranza, dello stesso status sociale di tutti gli altri cittadini. Ciò implica, in seconda battuta, la possibilità da parte dei primi di godere dei diritti fondamentali. E tali diritti fondamentali, infine, includono proprio il diritto a non essere soggetti a qualsivoglia tipo di diffamazione. Ed è proprio la presenza dello Stato, che si concretizza mediante l'emanazione di leggi contro il discorso d'odio, a fornire tale garanzia.

Ora però bisogna chiedersi se tali leggi siano da considerarsi sufficienti e necessarie al fine di assicurare ai cittadini protezione e sicurezza. In realtà, gli strumenti normativi ad oggi previsti non sono di per sé sufficienti, a meno che non sono accompagnati da forti interventi da parte della politica e se non sono applicati all'interno delle vicende processuali che coinvolgono coloro che esprimono messaggi d'odio (Young 1990, 62). Come fanno quindi tali strumenti a far sorgere nei cittadini un sentimento di sicurezza se poi non vengono attuati ai casi concreti? Il problema di fondo è legato al cosiddetto *expectation gap* (McNamara 2007, 178). Si tratta di uno scarto che si viene a creare tra il numero di volte in cui il cittadino invoca l'applicazione di una determinata legge contro l'odio aspettandosi quindi che la macchina giudiziaria venga attivata e la sua pretesa tutelata, e i pochi casi in cui ciò concretamente avviene.

Il divario nasce anche perché sussiste una componente soggettiva nel rapporto tra le leggi già menzionate e i cittadini beneficiari di queste ultime. Tale elemento è infatti costituito dalla percezione dei singoli individui in relazione all'essere membri di una società che gode di ottima salute. Il problema sussiste nel momento in cui, pur essendo riconosciuti in via teorica ai cittadini tutti i

diritti fondamentali, gli individui stessi non si sentono rassicurati nel corso della propria vita quotidiana, avendo di fronte a sé una società rovinata da segni di antisemitismo, immagini di croci bruciate o volantini razzisti diffamatori. A questo punto, viene perciò da chiedersi se le leggi anti-hate speech siano davvero necessarie dato che, nonostante la loro presenza, i cittadini non si sentono protetti.

2.4. L'Autonomia del singolo

Frederick Schaeur (1982, 17) individua tre ragioni da porre alla base della protezione della libertà di espressione del pensiero: il contributo alla verità, la promozione dell'autonomia personale e la protezione della democrazia. Nel definire in particolare il Principio di Autonomia, Edward Baker (1989, 59) afferma che

il postulato etico chiave è il rispetto per l'integrità individuale e l'autonomia, la quale consiste nel riconoscere il diritto ad una persona di usare il discorso per migliorare se stesso o per influenzare o interagire con gli altri in un modo che corrisponda ai propri valori.

Presupponendo che l'autonomia sia uno dei beni pubblici primari, egli sostiene che autonomia significhi che coloro che esprimono un messaggio

d'odio dovrebbero essere liberi di scegliere come e quando provare ad influenzare le credenze e le attitudini dei destinatari del discorso in relazione a determinati gruppi sociali. Ma il mero fatto che alcuni individui tentino di influenzare le credenze altrui in modi che le autorità vedono con disapprovazione, non costituisce ragione sufficiente al fine di giustificare la soppressione del messaggio d'odio. Ciò, quindi, comporta la difesa della libertà di espressione del proprio pensiero. Altro argomento che Baker utilizza per sostenere la libertà di esprimersi consiste nel considerare l'autonomia come la libertà dei destinatari del messaggio di decidere per se stessi quali descrizioni di certi gruppi sociali siano vere e quali false e di lasciarsi influenzare o meno dai trasmettitori d'odio circa i sentimenti e le credenze da provare verso i gruppi suddetti. In questo modo, la libertà di espressione del pensiero include anche il diritto di essere ascoltatori-spettatori del discorso d'odio.

A mio avviso, ciò che Baker non considera è l'eventualità che il pensiero, espresso liberamente dal singolo in ossequio al principio di Autonomia, vada a ledere proprio l'autonomia altrui, intesa quest'ultima proprio come libertà di interagire con gli altri in modo corrispondente ai propri valori. Per i sostenitori delle leggi contro l'hate speech,

tale danno parrebbe giustificare l'applicazione di strumenti repressivi del discorso al fine di tutelare l'autonomia di tutti. Ma è giusto altresì reprimere l'autonomia solo di coloro che esprimono determinati messaggi? Forse che l'alternativa possa essere quella di lasciare che ognuno liberamente formuli un proprio pensiero andando però a cambiare la radice del pensiero stesso, quindi attraverso un intervento ex ante, ossia mediante la formazione culturale di cittadini consapevoli e in grado di discernere ciò che viene detto loro (Strossen 2018, p168-171)?

Dal canto suo, Alexander Brown (Brown 2016, 60) propone il cosiddetto *Principio sfumato di Autonomia*. Tale postulato comporta che le restrizioni nell'uso dell'hate speech siano necessarie nella misura in cui assolvono la funzione di proteggere gli uditori del messaggio d'odio da indebite influenze sui propri impulsi e sulle proprie decisioni. Tre sono i casi di indebita influenza: la coercizione, l'utilizzo di strumenti diversi dai canali razionali della persuasione e le responsabilità da posizione. Ora, soffermandoci sul primo, ci dobbiamo chiedere se l'hate speech possa essere definito un discorso coercitivo. Viene sicuramente considerato coercitivo il discorso che il titolare di un'azienda

può rivolgere ai propri dipendenti, in quanto considerati *captive audience*, cioè pubblico prigioniero (Romano 2014, 71).

In termini più generali, si parla di *dottrina del pubblico prigioniero* quando il governo è in grado di regolare il discorso destinato ad un ascoltatore che non vuole sentire ma che si trova in una circostanza per cui è costretto a farlo (nel caso dei dipendenti, l'ufficio rappresenta il luogo di lavoro, quindi o si licenziano o sono costretti a sentire). Perciò, il diritto di parola viene limitato nella stessa misura in cui viene riconosciuto il diritto dell'uditore di non ascoltare. Quindi un discorso si può dire coercitivo nella misura in cui è rivolto ad un pubblico che è obbligato a udirlo? In realtà il termine prigioniero è di per sé contraddittorio, dal momento che, se è vero da un lato che noi siamo soggetti ogni giorno a parole di ogni tipo, dall'altro lato non possiamo negare che si possa evitare di esporsi a discorsi che reputiamo offensivi, ad esempio distogliendo lo sguardo, tappandosi le orecchie o cambiando direzione per strada. Nonostante i giudici americani non siano in grado di definire con precisione i limiti e la portata di tale dottrina, quest'ultima funge da ratio al fine di giustificare la restrizione del diritto di parola, in relazione al fenomeno hate speech.

Tuttavia, il rischio legato all'applicazione della teoria in esame, la quale consiste appunto nel proibire ad una persona di utilizzare un discorso intrusivo nei confronti di un destinatario che, per le circostanze in cui si trova, non può né distogliere lo sguardo, né tantomeno andarsene, è legato proprio all'abuso che i giudici stessi ne fanno. Questi ultimi, infatti, una volta operata una valutazione di contenuto discriminatorio sul discorso in esame, decidono sulla base della stessa di limitare la libertà di espressione del pensiero. Si parla infatti di valutazione discriminatoria perché le corti analizzano il contenuto del discorso, al fine di capire se quest'ultimo sia offensivo a tal punto da richiedere il loro intervento. In seconda battuta, al momento della decisione, molti giudici invocano il diritto alla privacy come ratio alla base della teoria del pubblico prigioniero (Hartley 2010, 65-125). Ma tale diritto comprende a sua volta il diritto di scegliere, che costituisce l'espressione massima del principio di Autonomia dell'individuo.

Marcy Strauss⁴³ sostiene che il diritto di scelta fornisce un modo per valutare se l'ascoltatore riesce a sopportare l'onere di evitare il discorso. Poniamo l'esempio di una donna che si

⁴³ https://chicagounbound.uchicago.edu/cgi/viewcontent.cgi?article=3013&context=journal_articles (1.2.2025)

trova in fila in aeroporto per acquistare dei biglietti aerei. Si trova di fronte a sé l'esponente di una setta, il quale le parla della propria religione e, nonostante le sue suppliche, non se ne va. A quel punto la donna si trova di fronte ad una pluralità di scelte. Può voltarsi e far finta di non sentire. Può mettere le mani sulle orecchie. Può lasciare la fila. I giudici valutano la ragionevolezza e l'efficacia di ciascuna scelta. Se la scelta *lasciare la fila* è l'unica che le permette di evitare il discorso, questa rappresenta un onere eccessivo da sopportare per la donna, perché in questo modo perderebbe il volo. I giudici, quindi, devono capire quando sussista un'influenza indebita sul diritto di scelta.

In conclusione, il carattere coercitivo del discorso d'odio dipende dall'ambiente in cui quest'ultimo si manifesta. Nel momento in cui la persona, in relazione alle circostanze del caso, è chiamata a sostenere un onere troppo grande per evitare il discorso in esame, ritengo sia necessario un intervento esterno che tuteli il diritto del singolo di non ascoltare. Ed è qui che dobbiamo chiederci se le leggi che sanzionano le condotte odiose ad oggi rappresentino la soluzione più efficace per raggiungere questo fine.

Il secondo caso di indebita influenza riguarda invece l'utilizzo di strumenti diversi dai canali razionali della persuasione al fine di indurre gli individui a pensare in un certo modo o a provare determinate sensazioni. David A. Strauss (1991, 334-371) difende il cosiddetto *Principio di Persuasione*, secondo il quale il governo non può sopprimere il discorso in ragione del fatto che tale discorso probabilmente persuade le persone a fare qualcosa che il governo stesso considera dannoso. Da ciò si deduce che il rischio che si corre, in caso di applicazione di leggi contro il discorso d'odio, è che queste vengano utilizzate da parte delle autorità come mero strumento politico per reprimere le idee e azzerare il dibattito pubblico.

Infine, il terzo caso di influenza indebita è costituito dalla mancanza di rispetto nei riguardi delle responsabilità da posizione. In questo caso, si pensi alla responsabilità che gli insegnanti hanno nei confronti dei propri alunni. Essendo riconosciuti come autorità all'interno dell'istituto scolastico, quest'ultimi hanno la responsabilità di fornire agli alunni una varietà di informazioni e di punti di vista. In caso contrario, infatti, il rischio è che si formi una generazione di individui le cui idee sono già a priori discriminatorie, non essendo

questi ultimi mai stati abituati a conoscere il diverso, quindi ad accettarlo. Anche quest'ultimo esempio ci induce a pensare che forse le leggi contro l'odio non siano il giusto mezzo al fine di istruire le future generazioni ad una conoscenza del mondo a tutto tondo, senza pregiudizi, né idee odiose.

2.5. Autogoverno e legittimità politica

Nell'ambito della moralità politica, invece, un altro dei principi posti alla base della libertà di espressione del pensiero è costituito dal principio dell'Autogoverno. Tale dottrina (Meiklejohn 1960, 14) pone in capo ai singoli cittadini il dovere civico di partecipare ai meccanismi democratici di decisione collettiva. Ciò, quindi, comporta l'accesso completo per tutti i cittadini alla vasta gamma di interazioni umane da cui deriva la capacità per il cittadino attivo di formulare un giudizio obiettivo e sano all'interno del processo democratico (Meiklejohn 1961, 245-266). Da ciò, Post (2009) e Weinstein⁴⁴ fanno discendere il diritto in capo agli individui di partecipare alla formazione dell'opinione pubblica.

⁴⁴https://scholarship.law.umn.edu/con-comm/456/?utm_source=scholarship.law.umn.edu%2Fcon-comm%2F456&utm_medium=PDF&utm_campaign=PDFCoverPages (1.2.2025)

Uno dei problemi principali che collega il discorso d'odio e la democrazia, concerne il contesto (Post 2011, 482) in cui il discorso stesso viene espresso. Poniamo il caso di trovarci all'interno di un ambiente di lavoro. Se uno dei dipendenti rivolge insulti razzisti nei confronti dei propri colleghi, quelle specifiche espressioni non possono essere qualificate come facenti parte del pubblico dibattito. Ed è proprio qui che rileva in modo preminente il ruolo dei giudici.

Questi ultimi devono essere in grado di distinguere ciò che può essere qualificato come commento politico e ciò che invece deve essere qualificato come discorso d'odio⁴⁵. Se noi ci troviamo ad un bar tra amici e usiamo nei loro confronti epiteti che possiamo categorizzare come insulti, non qualificheremo quei discorsi come hate speech, considerate sia l'intenzione dell'oratore che la percezione del discorso da parte del destinatario del messaggio stesso. Al contrario, si parla di *effetto refrigerante* delle leggi contro l'hate speech quando, all'interno di un contesto formale, quale può essere quello lavorativo, una persona viene inibita dall'esprimere un punto di vista diverso rispetto a quello del destinatario del discorso, solo per paura

⁴⁵<https://www.thefire.org/research-learn/freedom-speech-and-workplace-harassment> (1.2.2025)

di essere citata in giudizio (Gelber, McNamara 2015, 640).

Tali argomenti ci portano a riflettere sulle reali opportunità che gli individui hanno di prendere parte al processo di formazione dell'opinione pubblica. A tal proposito, Meiklejohn (1960, 24) immagina come ideale democratico la tradizionale riunione di città, nella quale cittadini liberi e uguali cooperano all'interno di un'impresa comune al fine di prendere decisioni sagge. Ciò che si vuole evitare è la *mutilazione del processo del pensiero* dei cittadini ed è proprio a questo scopo che vengono emanate le leggi contro il discorso d'odio. Ciò che Meiklejohn definisce mutilazione del pensiero si può tradurre nel cosiddetto *effetto silenziatore* del discorso d'odio. Alcuni autori, come Delgado (1991, 343-387), esponente della Critical Race Theory (Delgado, Stefancic 2017, 3), ritengono infatti che, se lasciato incontrollato, il discorso d'odio possa inibire i membri delle minoranze destinatarie del messaggio stesso, dal partecipare alla formazione del dibattito pubblico.

Tale tesi, però, presenta punti di debolezza. In primo luogo, non vi sono prove certe circa la causazione di tale effetto in capo alle minoranze suddette. Addirittura, vi è chi, come Weinstein

(2001, 158), ritiene che tentativi di vietare il discorso d'odio siano, in realtà, a loro volta motivati da pensieri odiosi. In secondo luogo, ammesso che un tale effetto inibitorio si produca, quest'ultimo non è conseguenza diretta del discorso d'odio di per sé ma dipende, come già ribadito in precedenza, dalla percezione che i destinatari hanno del discorso stesso. Tale percezione è strettamente correlata alle credenze e alle attitudini sociali del singolo. Ma tali presunti effetti mediati, che peraltro influenzano le persone in tutti gli ambiti della propria vita, non solo nei discorsi, non possono costituire ragione sufficiente per vietare il discorso d'odio.

In seconda istanza, la libertà di espressione del pensiero viene posta in correlazione al principio di Legittimità politica. Tale libertà è infatti il prezzo che noi dobbiamo pagare per la legittimità delle cosiddette “leggi a valle”, ossia quelle leggi che sono emanate allo scopo di combattere, ad esempio, le discriminazioni, la violenza e la diffamazione (Dworkin 2009, viii). Ma tali fenomeni costituiscono anche la ratio che i sostenitori delle leggi contro l'hate speech pongono alla base del favore nei confronti delle cosiddette “leggi a monte”. Quest'ultime, infatti, secondo il pensiero dei promotori delle leggi contro l'odio, servono proprio

per eliminare le cause di violazione delle leggi a valle. Ma se noi vogliamo leggi legittime contro la violenza o la discriminazione, è necessario lasciare che coloro che si vogliono opporre alle leggi suddette, facciano sentire la propria voce. Solo in questo modo, può dirsi rispettata la democrazia che, come sottolineato da Kelsen (1984), non è altro che una discussione continua tra maggioranza e minoranza.

È altrettanto vero però che tale opposizione può essere espressa in svariati modi, sia attraverso un linguaggio appropriato sia attraverso l'utilizzo di epiteti diffamatori. In dottrina quindi, ci si chiede se l'uso di quest'ultimi debba essere consentito o meno. Proviamo a ragionare a contrario. Che cosa succederebbe se agli oppositori che si esprimono in modo violento, fosse impedito di esprimersi del tutto? Verrebbe meno quel pubblico dibattito che costituisce uno dei principi cardine delle moderne democrazie. Infatti, l'imposizione in modo unilaterale da parte della maggioranza della propria volontà metterebbe in discussione la democrazia stessa e aprirebbe la strada ad un ritorno totalitarista.

Al contrario, alcuni promotori (Waldron 2012, 190) delle leggi contro l'odio sostengono che

quest'ultime costituiscano una sorta di *paradiso sicuro* per l'espressione del contenuto propositivo di visioni che diventerebbero opinabili quando espresse sotto forma di vituperio. Tuttavia, tale idea viene smentita osservando la casistica. È innegabile che vi siano stati casi in cui le condanne per l'utilizzo di linguaggi odiosi sono state confermate. È altrettanto vero però che vi sono stati casi in cui le sentenze di primo grado sono state annullate in appello o casi in cui le accuse sono cadute già in primo grado e infine casi in cui le persone sono state arrestate ma non sottoposte a processo. E tale incertezza giurisprudenziale non fa altro che mettere ancora di più in discussione la legittimità delle leggi contro l'odio.

Infine, merita di essere menzionato come punto di connessione tra la libertà di espressione del pensiero e la legittimità politica, il riconoscimento o meno di un'indipendenza etica in capo ai singoli individui. Legittimità politica del governo significa anche che quest'ultimo debba rispettare il diritto dei propri cittadini ad esprimere visioni anti-etiche del mondo. Il governo, infatti, non è legittimato ad imporre un'opinione collettiva riguardo le forme di rispetto che gli individui devono avere gli uni con gli altri.

2.6. *The Marketplace of Ideas*

Finora abbiamo appreso che i promotori delle leggi contro l'hate speech adducono a giustificazione di tale strumento legislativo, diversi beni ritenuti, a detta loro, meritevoli di tutela da parte dell'ordinamento giuridico.

Rimane ora da affrontare in modo più dettagliato la politica adottata dagli Stati Uniti per quanto concerne il fenomeno hate speech, la cosiddetta dottrina del *Marketplace of Ideas*. A tal proposito, il giudice Holmes, divenuto famoso per aver espresso un'opinione di dissenso nel contesto della sentenza *Abrams v. United States* del 1919⁴⁶, elabora la metafora del mercato, sostenendo che

l'ultimo bene desiderato viene raggiunto meglio dal libero scambio di idee – che la miglior prova di verità consiste nel potere del pensiero di essere accettato nella competizione del mercato.

Da questo passaggio apprendiamo come uno dei principi chiave che la dottrina statunitense elabora a favore della libertà di espressione del pensiero è costituito dalla scoperta della verità. La libertà di espressione, infatti, fornisce alle persone

⁴⁶<https://www.oyez.org/cases/1900-1940/250us616>
(22/09/2024).

delle inestimabili opportunità di scambiare l'errore per la verità (Mill 1972).

Vi è però da sottolineare che nel corso del tempo sono state formulate diverse teorie sul concetto di verità. Si passa dalla teoria della sopravvivenza della verità, secondo la quale la verità è per definizione quanto emerge dal consenso, essendo sopravvissuto un processo di aperto dibattito all'interno di un ristretto gruppo di persone, alla teoria della corrispondenza della verità, secondo la quale la libertà di ascoltare nuove opinioni può fornire valide opportunità di affermare quanto corrisponde all'attuale stato degli affari, per continuare con la teoria della coerenza della verità, per cui la libertà di ascoltare nuove opinioni può creare importanti opportunità di esprimere posizioni che costituiscono parte di un sistema coerente di credenze, concludendo con le diverse teorie deflazionistiche della verità, per cui la libertà in questione può far sì che nascano occasioni di conformarsi ad un modello nell'uso del predicato *è vero* che può essere osservato all'interno di una popolazione di utenti.

Chafee, dal canto suo, sostiene che la scoperta e la diffusione della verità su soggetti di interesse generale è possibile solo attraverso la discussione illimitata (Chafee 1941, 31). Il problema ora

è capire se, nell'ambito del discorso d'odio, la scoperta della verità costituisca un valore così importante da riuscire a bilanciare i pericoli derivanti dal discorso stesso.

Consideriamo una persona che si dichiara islamofoba. Questo individuo intende apparire sulla televisione nazionale francese allo scopo di diffondere una serie di stereotipi negativi sui musulmani. Lui pensa di poter riuscire nell'impresa, dal momento che si avvale di dati statistici incontrovertibili (tali dati indicano che il 65 % della popolazione carceraria francese è musulmana, mentre i musulmani rappresentano solo il 12 % della popolazione totale). Nonostante nel copione dell'islamofobo gli stereotipi negativi nei confronti dei musulmani non corrispondono a stati di fatto, ma essendo presente all'interno dello scritto anche un'accurata statistica sulla popolazione carceraria, si può dire che quel copione possa rientrare all'interno del principio di scoperta della verità. Ora è da capire se il valore intrinseco della verità sia più importante delle discriminazioni, delle violenze e, più in generale, dell'odio che può scaturire da un discorso del genere trasmesso a tutta la Nazione. Forse che la soluzione più opportuna sia di trovare un compromesso tra i vari interessi in gioco (Lepora 2012, 1-22)?

Per farlo, però, è bene discernere quei discorsi d'odio i cui usi possono costituire importanti occasioni per scambiare l'errore per la verità da quei discorsi i cui utilizzi, al contrario, non permettono di attuare tale tipo di operazione. È stato elaborato a tal fine il cosiddetto *calcolo post-Milliano* (Schauer 2012, 129-143). Secondo questa operazione, l'ufficiale soppressione del discorso può essere giustificata solo quando è prevedibile che le perdite che derivano dalla diffusione di false opinioni, le quali possono essere accettate e applicate nonostante la loro falsità, siano superiori in confronti ai guadagni che scaturiscono dalla scoperta di precedenti verità sconosciute e dall'aumento della conoscenza che costituisce il corollario della verità. In merito, si porta l'esempio della negazione dell'Olocausto. Permettere la circolazione di tali informazioni non farebbe altro che consentire la perpetuazione dell'ignoranza all'interno della popolazione.

Quindi, in una visione generale, oltre a capire se il valore Verità abbia un'importanza maggiore rispetto ad altri beni, come la dignità civica o l'autonomia, bisogna altresì capire se tale valore viene maggiormente protetto attraverso la libera espressione del pensiero, come sostengono gli assolutisti del Primo Emendamento, o se in realtà,

proprio perché una determinata opinione è ritenuta vera e in quanto tale lesiva di valori non epistemici – si pensi all'Olocausto – sia necessario intervenire attraverso leggi di repressione dell'opinione medesima.

In conclusione, per risolvere tale problema di bilanciamento di interessi, la Suprema Corte degli Stati Uniti adotta un approccio che viene definito *conseguenzialismo indiretto* (Sumner 2004, 33). Nell'osservare tale modo di giudicare, possiamo notare che la Corte cerca di trovare un equilibrio ottimale tra gli interessi di colui che parla a partecipare alla formazione dell'opinione pubblica, gli interessi dei destinatari del messaggio, infine gli interessi di parti terze che sono spettatrici delle espressioni utilizzate.

Dalla sentenza *New York Times Co. v. Sullivan*⁴⁷, emerge come la Corte abbia riconosciuto il bisogno di protezione costituzionale nei confronti di articoli di giornale aventi come protagonisti pubblici ufficiali, nonostante il loro contenuto non veritiero. Allo stesso tempo, i giudici hanno riconosciuto un livello minore di protezione nel caso in cui quelle dichiarazioni siano state rese con la consapevolezza della loro falsità o con colpevole negligenza di ciò che è vero e di ciò che è falso.

⁴⁷<https://www.oyez.org/cases/1963/39> (22/09/2024).

2.7. *Counterspeech*

In conclusione, dopo aver analizzato le politiche adottate dall'Unione Europea e dagli Stati Uniti, ritengo necessario vagliare ipotesi alternative alla censura al fine di fronteggiare il problema del discorso d'odio. L'ipotesi maggiormente accreditata è costituita dal cosiddetto *counterspeech*, ossia il contro-discorso. Molti (Gunther 1990, 7), in ambiente americano, ritengono che la risposta più appropriata al cattivo linguaggio sia quella di favorire in maggior quantità il buon linguaggio (Cradock 1995, 1047-1089). Quest'ultimo, infatti, ci può aiutare a capire perché determinate espressioni siano pericolose, scorrette, in contrasto con i valori della società, per questo ingiustificabili. Il contro-linguaggio può insinuare nelle menti dei messaggeri d'odio dubbi circa la verità, la validità e la desiderabilità di ciò che spinge loro ad agire (Strossen 2018, 158-163).

Vi è da aggiungere che tale metodo può assumere diverse forme. Esso, infatti, non si deve per forza e sempre tradurre in una contro risposta al messaggio odioso. Vi sono casi in cui il silenzio può costituire la risposta migliore, in questo modo negando ai portatori d'odio le attenzioni che essi cercano di ottenere dalla discussione. Ad esempio, dinanzi ad una parata neonazista per reclutare nuovi

adepti, il Southern Poverty Law Center (SPLC)⁴⁸ propone di organizzare un evento che costituisca un inno alla tolleranza e alla diversità.

Altra variante del contro-linguaggio può essere ciò che il New York Times definisce *humorous subversion*, cioè divertente sovversione. Il caso in esame riguarda la trasformazione di una marcia neonazista in *un'involontaria passeggiata*⁴⁹.

Tuttavia, nonostante la massiccia applicazione di tale tecnica soprattutto per quanto concerne l'hate speech online, fenomeno di cui parleremo nel proseguo della nostra trattazione, alcuni studiosi ritengono che il contro-linguaggio non sia un metodo efficace, tanto quanto la censura, nella guerra contro l'hate speech. Si ritiene infatti che sussistano ostacoli al contro-linguaggio soprattutto quando si tratta di discorso d'odio che avviene vis-à-vis.

In merito, Nielsen (2002, 265-280) ha portato avanti uno studio secondo il quale la più comune reazione al discorso d'odio razzista da parte di coloro che sono stati i destinatari di tale messaggio, è quella di ignorare il commento e di andarsene. Solo il 28 % delle persone di colore risponde

⁴⁸<https://www.splcenter.org/20170814/ten-ways-fight-hate-community-response-guide> (22/09/2024).

⁴⁹<https://www.nytimes.com/2017/08/17/opinion/how-to-make-fun-of-nazis.html> (22/09/2024).

a tali insulti razzisti ma solo quando questi si trovano in situazioni che reputano sicure, come una piazza gremita di persone. Quindi parte del problema correlato a tale tecnica è costituito dalla paura di rispondere.

Inoltre, è difficile capire come il contro-linguaggio possa in qualche misura migliorare i danni psicologici causati alle vittime del messaggio stesso. Sembra poco plausibile pensare che il rimedio a tali danni possa essere costituito da una maggiore quantità di linguaggio. Vi è però da sottolineare come, mentre le restrizioni legali per combattere l'uso dell'hate speech comportano una notevole spendita di tempo, considerati anche i tempi processuali, il contro-linguaggio costituisce invece un ottimo modo per risparmiare quel tempo.

Infine, una soluzione potrebbe essere quella di rendere le autorità pubbliche parti attive all'interno di tale battaglia contro il fenomeno hate speech. Ma fare ciò non attraverso l'emanazione di leggi bensì attraverso l'adozione di una politica di risposta che prevede che le istituzioni competenti si assumano la responsabilità di rispondere a tali odiatori, rilasciando pubbliche dichiarazioni sull'inaccettabile natura del discorso d'odio e dei danni che quest'ultimo può causare (Gelber 2002, 85).

2.8. Conclusioni

In sintesi, nel corso del tempo, i sostenitori delle leggi contro l'odio hanno tentato, e stanno cercando tuttora, di porre come ratio della repressione giuridica del discorso d'odio diversi beni ritenuti meritevoli di tutela da parte dell'ordinamento giuridico. Alcuni sostengono la superiorità del valore della dignità, declinata nelle sue diverse sfaccettature (umana o civica), rispetto al diritto del singolo alla libertà di espressione del pensiero.

Altri, al contrario, invocano la garanzia di sicurezza che tali leggi assicurerebbero nei confronti dei destinatari dei messaggi odiosi.

Altri infine proteggono la sfera di autonomia che, a detta loro, apparterrebbe a ciascuno di noi.

Ma, come abbiamo avuto modo di apprendere, la protezione in via preferenziale di ciascuno di tali beni rispetto alla libertà di espressione del pensiero, presenta delle criticità che non si possono ignorare. Proprio per questo motivo, nel dibattito pubblico legato al fenomeno dell'hate speech, si sono affacciati metodi alternativi alle leggi di repressione del discorso d'odio, uno su tutti il cosiddetto *counterspeech*, l'utilizzo del contro-discorso, frutto della dottrina tipicamente americana del cosiddetto *Marketplace of Ideas*. La

ratio di tale strumento consiste infatti nel permettere ad ogni individuo di far circolare, nel mondo virtuale e in quello reale, messaggi positivi costruttivi, in modo tale da ridurre progressivamente i messaggi d'odio, senza tuttavia andare a limitare la libertà di opinione del singolo.

Capitolo Terzo

Online Hate Speech e AI

3.1. Internet e Online Hate Speech

Internet ha cambiato per sempre il modo in cui l'informazione viene prodotta, distribuita e utilizzata. Nelle parole di Bill Gates, verrà un giorno, e non è molto lontano, in cui potremo concludere affari, studiare, conoscere il mondo e le sue culture, assistere a importanti spettacoli, stringere amicizie, visitare i negozi del quartiere e mostrare fotografie a parenti lontani, tutto senza muoverci dalla scrivania o dalla poltrona.

In particolare, due sono stati i mutamenti cruciali ai fini dell'innovazione tecnologica⁵⁰. Il primo è l'affermazione di un sistema di produzione dell'informazione radicalmente decentralizzato. È sufficiente disporre di un computer, un tablet o di uno smartphone ed essere connessi ad Internet per diventare produttori di informazione. Il singolo ormai assume un ruolo attivo all'interno dei flussi di comunicazione. Chiunque, infatti, può produrre informazioni nella rete, reagire all'informazione immessa da altri, proporre fatti, idee, cri-

⁵⁰<https://www.medialaws.eu/wp-content/uploads/2019/05/1.-Pitruzzella.pdf> (8.2.2025)

tiche, nuovi punti di vista, foto, video. Yochai Benkler (2006, 2 e ss.) parla in proposito di *network information economy*, i cui tratti distintivi sono i seguenti: a) la produzione dell'informazione è radicalmente decentrata, per cui ogni utente della rete diventa un potenziale produttore di informazioni; b) il prezzo estremamente basso pagato per usare un device elettronico, accedere alla rete e comunicare, pone i mezzi materiali di informazione e di produzione culturale nelle mani di una parte significativa della popolazione mondiale, dell'ordine di circa due miliardi di persone; c) i flussi di comunicazione hanno carattere aperto e globale. Il secondo mutamento prodotto dall'innovazione tecnologica è costituito dall'avvento dei cosiddetti *gatekeepers*, soggetti in grado di controllare i flussi di informazioni, di ordinarli e di mettere in comunicazione coloro che emettono i dati con coloro che li ricevono.

Si tratta in realtà di una ristretta cerchia di *Tech Companies* in grado di controllare l'accesso globale al mondo virtuale. Riservandoci di approfondire in un secondo momento l'operare di questi soggetti, i cosiddetti ISP (Internet Service Providers), ora vi è da chiedersi quale sia la correlazione tra il fenomeno mondiale della digitalizzazione e il

discorso d'odio. Sul punto Giovanni Ziccardi sostiene che i problemi correlati al discorso d'odio non riguardano tanto o, meglio, non solo il contenuto del messaggio in sé considerato bensì il mezzo attraverso cui tale espressione viene diffusa. Egli precisa infatti che l'espressione *speech* non si esaurisce con l'atto del parlare ma si riferisce a un discorso d'odio veicolato attraverso qualsiasi mezzo, anche e soprattutto tecnologico. Proprio da qui, gli studiosi sono partiti per capire il rapporto venutosi ad instaurare nel corso del tempo tra le espressioni di odio e la Rete. Se da un lato vi è chi⁵¹ sostiene che le tecnologie sono semplicemente un mezzo, un elemento tecnico funzionale alla mera diffusione di un sentimento già presente nella società, dall'altro vi è chi⁵² afferma che la Rete ha causato in modo diretto la nascita di nuove ed estreme modalità di interazione tra gli individui. Secondo quest'ultimo approccio, infatti, Internet non sarebbe più informato al Principio di Neutralità, ma avrebbe contribuito a creare un ambiente avente un'incidenza diretta sulle espressioni usate.

⁵¹ <http://dig.casail.mit.edu/2006//06/neutralnet.html>
(28/11/2024).

⁵² https://papers.ssrn.com/sol3/papers.cfm?abstract_id=957373
(9.2.2025)

La scelta del primo orientamento implica pertanto che Internet non debba essere regolamentata più di quanto già non lo sia. Esso si avvicina al pensiero degli studiosi nordamericani, ferventi sostenitori della libertà di espressione del pensiero, secondo i quali criminalizzare l'hate speech non servirebbe a ridurre la quantità di espressioni d'odio circolanti. Si rivelerebbe quindi un rimedio inefficace e poco utile. Viceversa, il secondo approccio rischia di condurre ad una ipercriminalizzazione della rete, presentando al tempo stesso il vantaggio di comprendere i cambiamenti di cui la rete risulta continuamente protagonista.

Dal canto suo, Paolo Sordini⁵³, nell'affrontare le problematiche legate alla libertà di espressione del pensiero nell'era digitale, delinea le caratteristiche proprie dei nuovi mezzi telematici: a) Globalità: Internet offre un accesso immediato ad informazioni provenienti da tutte le parti del mondo. Attraverso il web, infatti, un vastissimo numero di fonti d'informazione è disponibile in qualsiasi parte del mondo ad un vastissimo numero di persone; b) Decentralizzazione: Internet è un sistema decentralizzato, senza punti di accesso predefiniti. Ciò rende sempre più difficile il controllo

⁵³https://www.ispionline.it/sites/default/files/pubblicazioni/wp_52_2013.pdf (2.2.2025)

di Internet da parte dei governi di tutto il mondo e non permette la creazione di monopoli o oligopoli;

c) Apertura: Internet è un servizio offerto a prezzi irrisori, se non nulli, dal momento che la creazione di nuove piattaforme, come i social networks, permette altre forme di finanziamento, come la pubblicità;

d) Vastità: Internet presenta una capacità di archiviazione illimitata;

e) Interattività: Internet consente flussi di comunicazioni multi-direzionali;

f) Controllo da parte dell'utente: l'utente può controllare il tipo di informazione che arriva nel proprio computer ed è in grado di criptare le proprie comunicazioni, sottraendole così al controllo pubblico;

g) Indipendenza: Internet non dipende da alcuna infrastruttura. L'accesso ad Internet è infatti possibile anche attraverso i cellulari, ma a differenza delle comunicazioni con questi ultimi, si sottrae più facilmente a possibilità di controllo effettivo da parte dello Stato. Tutte queste caratteristiche hanno fatto sì che il web trasformasse la diffusione del pensiero di ciascuno di noi da quantitativamente circoscritta e limitata nello spazio e nel tempo a capillare, globale e permanente (Abbondante 2017, 41-68).

Internet ha acquisito un'importanza tale nel mondo dei diritti, in particolare il diritto alla libera

manifestazione del pensiero, che lo stesso Parlamento italiano è intervenuto in merito nel 2015 mediante l'approvazione della cosiddetta Dichiarazione dei diritti in Internet⁵⁴, la quale riassume in maniera adeguata ciò che la Rete rappresenta per i diritti:

Internet ha contribuito in maniera decisiva a ridefinire lo spazio pubblico e privato, a strutturare i rapporti tra le persone e tra queste e le istituzioni. Ha cancellato confini e ha costruito modalità nuove di produzione e utilizzo della conoscenza. Ha ampliato le possibilità di intervento diretto delle persone nella sfera pubblica. Ha modificato l'organizzazione del lavoro. Ha consentito lo sviluppo di una società più aperta e libera. Internet deve essere considerata come una risorsa globale e che risponde al criterio dell'universalità [...]. Internet si configura come uno spazio sempre più importante per l'autorganizzazione delle persone e dei gruppi e come uno strumento essenziale per promuovere la partecipazione individuale e collettiva ai processi democratici e l'eguaglianza sostanziale.

La stessa Corte Europea dei diritti dell'uomo si dimostra preoccupata circa l'esercizio della libertà di espressione del pensiero nel mondo dei bit e, nella sentenza Editorial Board of Pravoye Delo e

⁵⁴https://www.camera.it/application/xmanager/projects/leg17/commissione_internet/dichiarazione_dei_diritti_internet_publicata.pdf (28/11/2024).

Shtekel c. Ucraina del 2011⁵⁵, sostiene che Internet è:

un mezzo di informazione e di comunicazione particolarmente distinto rispetto alla carta stampata, specialmente avuto riguardo alla capacità di impacchettare e trasmettere le informazioni [...]. Il rischio di danno derivante dal contenuto e dalle comunicazioni sul web, soprattutto in riferimento all'esercizio dei diritti umani fondamentali, è certamente più elevato di quello derivante dalla carta stampata.

Dal canto suo, l'UNESCO (Gagliardone, Alves, Martinez 2015) si è pronunciato sui contenuti che veicolano sul web, in particolar modo sul fenomeno dell'hate speech online, sostenendo che, sebbene l'hate speech online non sia differente da quello offline da un punto di vista contenutistico, tuttavia il primo presenta caratteristiche uniche e peculiari. Queste sono: la permanenza, il carattere itinerante, l'anonimato ed il carattere trans-giurisdizionale.

Per quanto concerne la prima caratteristica, Andre Oboler, amministratore delegato dell'Istituto per la prevenzione dell'hate speech, sottolinea che più a lungo il contenuto d'odio permane sulla

⁵⁵[https://hudoc.echr.coe.int/eng#{%22itemid%22:\[%22001-104685%22\]}](https://hudoc.echr.coe.int/eng#{%22itemid%22:[%22001-104685%22]}) (28/11/2024).

rete, maggiore è il danno che può causare alle vittime e maggiore diventa il potere dei responsabili del messaggio odioso. Inoltre, il messaggio odioso può essere itinerante. Ciò significa che, anche se quel contenuto viene rimosso, esso può comunque trovare la propria espressione altrove all'interno della stessa piattaforma o addirittura al di fuori di essa. L'anonimato è un altro strumento che facilita il traffico online del discorso d'odio. Come ha dichiarato Drew Boyd, Direttore delle Operazioni del Progetto Sentinel,

Internet garantisce agli individui la capacità di dire cose orribili perché pensano che non verranno scoperte. Questo è ciò che rende il discorso di odio online così unico, perché le persone si sentono molto più a loro agio esprimendo odio online rispetto alla vita reale perché questo vorrebbe dire affrontare le conseguenze di quello che dicono.

Infine, Internet presenta carattere transgiurisdizionale. Sul punto, Francisco Javier Ansuátegui Roig (2017, 29-48) afferma che Internet presuppone una negazione dell'idea del territorio inteso come spazio fisico nel quale si esercitano e rivendicano i diritti. Egli parla di una vera e propria *deteritorializzazione* dei diritti e del loro esercizio, la quale ci obbliga a una riformulazione dell'idea di confine, di limite, di frontiera. Il diritto contempo-

raneo si presenta come un diritto sconfinato, mettendo così in evidenza la necessaria e decisiva relazione tra il diritto e la dimensione territoriale o spaziale. Il referente fisico svanisce. Internet è uno spazio immateriale e ci obbliga a procedere a una riformulazione dell'idea di territorio, ad elaborare un nuovo discorso sulla territorialità, consustanziale al diritto.

3.2. Il Principio di Neutralità della Rete

Di origine nordamericana, il principio di Neutralità della rete viene riconosciuto in Italia mediante la già menzionata Dichiarazione dei diritti di Internet del 2015, la quale, ai sensi dell'articolo 4 comma 1, sancisce che:

Ogni persona ha il diritto che i dati trasmessi e ricevuti in Internet non subiscano discriminazioni, restrizioni o interferenze in relazione al mittente, ricevente, tipo o contenuto dei dati, dispositivo utilizzato, applicazioni o, in generale, legittime scelte delle persone.

Prosegue al comma 2:

Il diritto ad un accesso neutrale ad Internet nella sua interezza è condizione necessaria per l'effettività dei diritti fondamentali della persona.

Prima però di capire cosa significhi in concreto il concetto di neutralità della rete in ambiente europeo, è necessario fare un passo indietro e ritornare alle origini. La teoria della neutralità della Rete nasce all'interno della legislazione dei singoli Stati della Federazione statunitense, in primis nello Stato di New York attorno alla metà del XIX secolo, in relazione alle reti telegrafiche. In particolare, con il Communication Act del 1934 si afferma il criterio del *first come first served*, sulla base del quale l'obbligo di instradare le chiamate telefoniche deve avvenire in modo del tutto indipendente dal mittente, dal destinatario e dal contenuto del messaggio trasmesso (si sancisce così un vero e proprio divieto di discriminazione).

Il dibattito sulla *net neutrality* riemerge poi sul suolo americano verso la metà degli anni Ottanta in relazione alla liberalizzazione del mercato delle telecomunicazioni. A tal proposito, si discute dell'ammissibilità dell'integrazione verticale tra operatore di rete e fornitori di servizi. E proprio per assicurare ai nuovi fornitori di non essere soggetti a discriminazioni da parte dell'operatore di rete che vuole mantenere la propria posizione dominante all'interno del mercato, si utilizza il con-

retto di neutralità della rete al fine di salvaguardare la concorrenza nel mercato delle telecomunicazioni.

A differenza del settore delle telecomunicazioni però, Internet si configura come un modello distribuito in cui l'intelligenza risiede nei terminali di rete mentre non esiste un unico centro di gestione della rete stessa. In che termini quindi si può porre il problema della net neutrality con riferimento alla rete Internet? Il dilemma si concentra tutto sulla possibilità o meno di qualificare gli Internet Service Provider (ISP) come *common carriers*. Gli Internet Service Providers sono quei soggetti che permettono l'accesso alla rete consentendo, dunque, l'instradamento su di essa delle comunicazioni elettroniche. L'ISP ha rapporti, a monte, con l'operatore di rete fissa o mobile e, a valle, con gli acquirenti di capacità trasmissiva che possono essere sia semplici cittadini/utenti sia fornitori di servizi di comunicazione elettronica. Ora, avendo la Federal Communication Commission stabilito che gli ISP sono effettivamente common carriers, ad essi si può e si deve applicare il generale divieto di discriminazione previsto per le altre utilities, nonché l'obbligo ad essere neutrali e ad evitare di rallentare o di velocizzare determinati pacchetti di dati.

In ambito europeo invece, l'approccio al principio di neutralità è stato diverso. L'Unione Europea⁵⁶ ha preferito adottare una regolamentazione ex ante al fine di garantire l'osservanza del principio in questione. I primi tentativi risalgono al 2002 quando viene emanata la direttiva-quadro 2002/21/CE che impone ai legislatori nazionali di tutelare gli interessi dei cittadini europei promuovendo la capacità degli utenti finali di accedere ad informazioni e distribuirle o eseguire applicazioni e servizi di loro scelta e che, ai sensi dell'articolo 20, prevede il cosiddetto *servizio universale*, ossia un generale obbligo di trasparenza a carico degli operatori di rete intenzionati ad adottare pratiche di network management.

Ma il principio di neutralità e il conseguente divieto di discriminazione trovano il proprio consolidamento con il Regolamento Europeo 2015/2120⁵⁷, mediante il quale il legislatore vuole garantire un trattamento equo e non discriminatorio del traffico nella fornitura di servizi di accesso ad Internet, in generale un'Internet aperta a tutti.

⁵⁶<https://www.bibliotecheoggi.it/media/download/get/dd3cbe1e-906d-447c-95ad-c453b590d5bd/original> (8.2.2025)

⁵⁷<https://op.europa.eu/it/publication-detail/-/publication/8fdf5d08-93fc-11e5-983e-01aa75ed71a1> (28/11/2024).

Tale diritto di accesso universale si ritiene soddisfatto quando, ai sensi dell'articolo 3 comma 1 del Regolamento,

gli utenti finali hanno il diritto di accedere a informazioni e contenuti e di diffonderli, nonché di utilizzare e fornire applicazioni e servizi, e utilizzare apparecchiature terminali di loro scelta, indipendentemente dalla sede dell'utente finale o del fornitore o dalla localizzazione, dall'origine o dalla destinazione delle informazioni, dei contenuti, delle applicazioni o del servizio, tramite il servizio di accesso a Internet.

In seconda battuta, lo stesso articolo 3 comma 3 fornisce una definizione di neutralità, sancendo che:

i fornitori di servizi di accesso a Internet, nel fornire tali servizi, trattano tutto il traffico allo stesso modo, senza discriminazioni, restrizioni o interferenze, e a prescindere dalla fonte e dalla destinazione, dai contenuti cui si è avuto accesso o che sono stati diffusi, dalle applicazioni o dai servizi utilizzati o forniti, o dalle apparecchiature terminali utilizzate.

Occorre tuttavia sottolineare che il legislatore europeo, di fronte al caso concreto, decide comunque di operare – in linea con quanto posto in essere dagli Stati Uniti – un bilanciamento tra le di-

verse esigenze provenienti dai soggetti che popolano la rete. Infatti, ai sensi dell'articolo 3 comma 3 paragrafo 2, egli prevede che il primo comma non impedisce ai fornitori di servizi di accesso a Internet di attuare misure di gestione ragionevole del traffico. Per soddisfare il parametro della ragionevolezza, tali misure devono essere trasparenti, non discriminatorie, proporzionate e non devono essere basate su considerazioni di ordine commerciale ma su requisiti di qualità tecnica del servizio obiettivamente diversi di specifiche categorie di traffico. Tali misure non controllano i contenuti specifici e sono mantenute per il tempo strettamente necessario.

Infine, ai sensi dell'articolo 3 comma 3 paragrafo 3, si stabilisce che, indipendentemente dalla ragionevolezza o meno della pratica, quest'ultima sia possibile qualora risulti necessaria per preservare l'integrità della rete, al fine di evitare o risolvere una congestione di carattere eccezionale, ottemperare a sentenze o atti legislativi o delle autorità indipendenti.

Tuttavia, anche alla luce di tale regolamentazione, non si può dire che esista un'unica definizione di *net neutrality*. Infatti, oltre alla definizione

di neutralità in senso stretto⁵⁸ che si può dire emergere anche dalla normativa analizzata, autorevole dottrina ha fornito nel corso del tempo altre definizioni del principio di neutralità. A tal proposito, vi è chi sostiene che neutralità della rete⁵⁹ significa che i fornitori di servizi a banda larga addebitano una sola volta ai consumatori l'accesso ad Internet, non favoriscono un fornitore rispetto ad un altro e non fanno pagare i fornitori di contenuti per l'invio di informazioni agli utenti finali.

Dal canto suo, la Federal Communications Commission ha codificato nel corso del tempo una serie di principi relativi alla net neutrality, sancendo che una persona coinvolta nella fornitura di servizi di accesso ad Internet a banda larga fissa, nella misura in cui tale soggetto è così impegnato, deve pubblicamente divulgare informazioni accurate riguardo alle pratiche di gestione della rete, alle prestazioni e ai termini commerciali (trasparenza), non deve bloccare i contenuti legali, le applicazioni, i servizi o i dispositivi non dannosi, soggetti a una ragionevole gestione della rete (assenza di blocchi), non deve irragionevolmente discriminare nel trasmettere il traffico di rete legale ad un

⁵⁸<https://www.washingtonpost.com/wp-dyn/content/article/2006/06/07/AR2006060702108.html> (9.2.2025)

⁵⁹https://papers.ssrn.com/sol3/papers.cfm?abstract_id=943757 (7.2.2025)

consumatore di servizi di accesso a Internet a banda larga (nessuna irragionevole discriminazione).

3.3. Social Networks e responsabilità degli ISP

Come abbiamo già avuto modo di apprendere in precedenza, le peculiarità della rete Internet, così come delineate nel progetto UNESCO *Countering Online Hate Speech* del 2015, fanno sì che l'odio online possa veicolare più velocemente rispetto a quello offline, non presenti limiti spaziali e possa rivolgersi ad un numero indeterminato di destinatari sparsi in tutto il mondo. V'è da dire che tale fenomeno è stato affrontato in modo diverso da parte degli Stati Uniti rispetto a quanto posto in essere dall'Unione Europea, in virtù delle sostanziali differenze di approccio e di pensiero di cui abbiamo parlato tempo addietro. Ciò non toglie che, come vedremo in seguito, sia gli americani che gli europei abbiano visto il coinvolgimento, attraverso modalità differenti, dei cosiddetti ISP (Internet Service Providers), allo scopo di fronteggiare tale piaga sociale.

Dal canto suo, la Commissione Europea, di fronte all'insufficiente efficacia giuridica degli strumenti normativi di censura adottati dai singoli Stati membri, ha sottoscritto a partire dal 2016 una

serie di Codici di condotta⁶⁰ con i quali le più importanti piattaforme social operanti su Internet si sono impegnate alla rimozione dei contenuti odiosi nel termine di ventiquattro ore dalla ricezione della segnalazione del contenuto illegale, avendo altresì inserito nei propri termini contrattuali regole volte a prevenire la diffusione di messaggi ripugnanti. In particolare, le *IT Companies* hanno stipulato un accordo attraverso il quale si impegnano non solo a rimuovere i contenuti odiosi in meno di ventiquattro ore su segnalazione degli utenti ma anche a educare e a far crescere in loro la consapevolezza dei contenuti ritenuti vietati, secondo quanto previsto dalle linee guida comunitarie.

Tali società informatiche si impegnano altresì a fornire tutte le informazioni necessarie riguardo le procedure di presentazione degli avvisi, in modo tale da migliorare la rapidità e la capacità di comunicazione tra autorità statali e le società stesse. Infine, riconoscendo il valore dell'indipendenza del cosiddetto *counterspeech* contro l'odiosa retorica e il pregiudizio, quest'ultime si prefiggono

⁶⁰https://commission.europa.eu/strategy-and-policy/policies/justice-and-fundamental-rights/combating-discrimination/racism-and-xenophobia/eu-code-conduct-countering-illegal-hate-speech-online_en (28/11/2024).

lo scopo di continuare il proprio lavoro promuovendo contro-narrazioni e predisponendo programmi educativi atti a stimolare il pensiero critico.

La ragione di fondo che ha spinto tali società ad adottare questo tipo di atteggiamento è in realtà da ricercarsi nei problemi economici che tali compagnie hanno dovuto affrontare, essendo ritenute concause della dilagante esplosione di violenza che ha avuto luogo nel mondo nel corso degli ultimi anni.

Un esempio di ciò può essere fornito dalla diffusione su Youtube nel settembre 2012 di alcuni spezzoni del lungometraggio *The Innocence of Muslims*, video avente finalità denigratoria e offensiva nei confronti della religione islamica. La propagazione nel mondo di tale video sarebbe stata alla base dei violenti eventi dell'11 settembre 2012 a Bengasi, culminati nell'attentato al Consolato statunitense, nel quale persero la vita quattro funzionari americani.

Altro caso riguarda il poeta Mohamend Ibn Dahami Al-Ajami, il quale è stato condannato all'ergastolo dal tribunale penale di prima istanza del Qatar a causa di alcuni versi contenuti nella *Poesia dei Gelsomini*, inneggianti alla caduta di Ben

Ali. Tale poesia è poi rimbalzata da un social network ad un altro fino a diventare un grido di rivolta in Qatar.

Infine, altro tipo di violenza ha investito in Italia una ragazza che ha tentato di togliersi la vita a causa della diffusione a sua insaputa su Facebook di un video lesivo della propria riservatezza e dignità.

Tali eventi ci dimostrano come le caratteristiche proprie della rete Internet giochino un ruolo fondamentale nella diffusione di odio nel mondo e quindi debbano essere a loro volta sfruttate al fine di combatterlo.

La quarta valutazione semestrale⁶¹ operata dalla Commissione Europea nel febbraio del 2019, mostra come l'adozione di Codici di condotta stia producendo effetti positivi. Le aziende IT stanno valutando l'89% dei contenuti segnalati entro 24 ore e il 72% dei contenuti ritenuti illegali è stato rimosso, rispetto al 40% e al 28% rispettivamente quando il Codice è stato lanciato nel 2016. Tuttavia, le aziende devono migliorare il loro feedback agli utenti che notificano i contenuti e fornire maggiore trasparenza su avvisi e rimozioni. Inoltre, quattro nuove aziende hanno deciso di aderire al

⁶¹https://ec.europa.eu/info/files/factsheet-4th-monitoring-round-code-conduct_it (28/11/2024).

Codice nel corso del 2018: Google+, Instagram, Snapchat, Dailymotion. La Commissione europea rileva altresì che ad oggi nell'88.9 % dei casi le società IT valutano le notifiche in meno di 24 ore, nel 6.5% in meno di 48 ore, nel 3.9% in meno di una settimana e nello 0.7% dei casi impiegano più di una settimana. Inoltre, per quanto concerne le eliminazioni operate da ciascuna società IT a partire dal momento dell'emanazione del Codice di Condotta, risulta che le compagnie IT abbiano rimosso circa il 71,7 % dei contenuti a loro segnalati mentre il 28,3 % sia rimasto online. Infine, in relazione al tasso di rimozione di contenuti costituenti hate speech per ciascun paese facente parte dell'Unione Europea, il tasso di rimozione del contenuto odioso da parte del nostro Paese ammonta al 56.6 %.

Nonostante gli ottimi risultati ottenuti mediante tale strumento pattizio, nel corso del tempo l'Unione Europea ha iniziato a valutare la possibilità di combattere l'odio online riconoscendo in capo ai cosiddetti ISP (Internet Service Providers), una qualsivoglia forma di responsabilità, sia sul piano civile che sul fronte penale, per i contenuti pubblicati da terzi sui social networks. Sia il legislatore che la dottrina iniziano a valutare tale pos-

sibilità avendo preso atto della trasformazione degli ISP da intermediari neutri a soggetti sempre più attivi in relazione alla diffusione dei contenuti pubblicati dagli utenti sulla rete.

Ma andiamo con ordine. In prima battuta, occorre sottolineare che il cosiddetto *Cyberspace* è popolato da una pluralità di attori, quali l'utente, l'autorità pubblica e l'ISP. Per questo motivo, sembra difficile allocare una generale responsabilità per violazioni e contenuti illeciti in capo ad un singolo attore o categoria di attori, emergendo molto spesso l'utilità delle teorie del *risk sharing* e del *problem of many hands*. È quindi necessario tenere in considerazione questa interazione tra le diverse soggettività al fine di delineare una regolamentazione che sia in grado di evitare di lasciare un margine di discrezionalità troppo ampio in capo agli ISP, che agiscono in ogni caso seguendo logiche economico-imprenditoriali.

Ma chi è l'ISP? Con questa espressione si indica l'organizzazione o infrastruttura di natura imprenditoriale che offre agli utenti, mediante un contratto di fornitura, servizi inerenti ad internet (per esempio, l'accesso al web e alla posta elettronica, la condivisione o memorizzazione di dati ecc.). Dal punto di vista normativo gli Internet service provider sono denominati *prestatori di servizi*

della società dell'informazione, e la loro attività è regolata dal decreto legislativo n. 70/2003, emanato in attuazione della Direttiva europea 31/2000/CE sul commercio elettronico, la cosiddetta Direttiva e-commerce. Nella categoria, rientrano altresì i gestori delle piattaforme di social networking.

La normativa individua in particolare tre differenti figure di ISP, corrispondenti a tre diverse tipologie di servizi: semplice trasporto (c.d. *mere conduit*), che consiste nel trasmettere, su una rete di comunicazione, informazioni fornite da un destinatario del servizio (l'utente) o nel fornire un accesso alla rete di comunicazione; memorizzazione temporanea (c.d. *caching*), che consiste nella memorizzazione automatica, intermedia e temporanea di informazioni fornite da un utente, effettuata al solo scopo di rendere più efficace il successivo inoltramento ad altri destinatari a loro richiesta; memorizzazione di informazioni (c.d. *hosting*): attività che consiste nella memorizzazione di informazioni fornite da un utente e che rappresenta la categoria senza dubbio più controversa.

L'esigenza di una maggiore responsabilizzazione di questi soggetti si fonda non solo sulla loro posizione di custodi materiali dei contenuti della rete, ma soprattutto sulle capacità di intervento di cui dispongono e che attualmente già utilizzano,

senza che queste possano dunque considerarsi fonte di costi eccessivi da sostenere. L'attività di rilevazione ed analisi dei contenuti illeciti online all'interno delle piattaforme può essere infatti effettuata grazie a software, i quali, basati sull'utilizzo di impronte digitali abbinate univocamente al contenuto digitale tutelato, non richiedono irragionevole impiego di risorse preposte allo scopo. Inoltre, gli ISP di grandi dimensioni dispongono già di staff di persone dedicate alla gestione delle segnalazioni concernenti contenuti illeciti presenti sulle proprie piattaforme.

Naturalmente regimi di responsabilità fondati su tali capacità devono partire necessariamente da una distinzione fra le diverse tipologie di intermediari, differenziando quelli più solidi da un punto di vista imprenditoriale ed economico da quelli di più modeste dimensioni e strutture tecniche, in modo da evitare che un regime troppo severo di responsabilità espella dal mercato i providers non in grado di sopportare i costi di un sistema capillare di prevenzione degli illeciti oppure di un allargamento del fronte dei risarcimenti dovuti (Piraino 2017, 152-200).

Dal canto suo, il d.lgs. 70/2003 stabilisce alcuni principi generali di responsabilità che valgono per tutti i provider, e alcune regole specifiche

che variano per ciascuna delle tre tipologie di attività. Un principio generale è che i provider non sono responsabili delle informazioni trattate e delle operazioni compiute dagli utenti che fruiscono del servizio, salvo che intervengano sul contenuto o sullo svolgimento delle stesse operazioni.

Inoltre, ai sensi dell'articolo 17 dello stesso decreto, non esiste un dovere generale di vigilanza, per cui il provider non ha l'obbligo di sorvegliare le informazioni che trasmette o memorizza né di ricercare attivamente fatti o circostanze che indichino la presenza di attività illecite.

A fronte di ciò il provider ha due doveri fondamentali di informativa e collaborazione con le autorità: quando viene a conoscenza di presunte attività o informazioni illecite che riguardano un utente, deve informare senza indugio le autorità di vigilanza (giudiziaria o amministrativa); deve fornire senza indugio, a richiesta delle autorità, le informazioni in suo possesso utili a identificare l'utente con il quale ha accordi di memorizzazione dei dati, al fine di individuare e prevenire attività illecite.

Il provider risponde civilmente del contenuto dei servizi (cioè, è obbligato a risarcire i danni subiti da terzi) quando, richiesto dalle autorità di vigilanza, non si attiva prontamente per impedire

l'accesso ai contenuti, oppure non informi le autorità avendo avuto notizia del carattere illecito o dannoso di un contenuto.

Prima però di procedere approfondendo la figura dei cosiddetti *hosting providers*, è bene altresì menzionare i Considerando n. 14 e n. 42 della già citata Direttiva e-commerce. Il Considerando n. 14 prevede infatti che:

l'applicazione della presente direttiva deve essere pienamente conforme ai principi relativi alla protezione dei dati personali, in particolare per quanto riguarda il regime di responsabilità per gli intermediari

mentre il Considerando n. 42 stabilisce che:

le deroghe alla responsabilità stabilita nella presente direttiva riguardano esclusivamente il caso in cui l'attività di prestatore di servizi della società dell'informazione si limiti al processo tecnico di attivare e fornire accesso ad una rete di comunicazione sulla quale sono trasmesse o temporaneamente memorizzate le informazioni messe a disposizione da terzi al solo scopo di rendere più efficiente la trasmissione. Siffatta attività è di ordine meramente tecnico, automatico e passivo, il che implica che il prestatore di servizi della società dell'informazione non conosce né controlla le informazioni trasmesse o memorizzate.

Dal canto suo, il Regolamento europeo 2016/679, al Considerando n. 21, specifica che:

il presente regolamento non pregiudica l'applicazione della direttiva 2000/31/CE del Parlamento europeo e del Consiglio, in particolare delle norme relative alla responsabilità dei prestatori intermediari di servizi di cui agli articoli da 12 a 15 della medesima direttiva. Detta direttiva mira a contribuire al buon funzionamento del mercato interno garantendo la libera circolazione dei servizi della società dell'informazione tra Stati membri.

Ai sensi dell'articolo 2 par. 4 prevede altresì che:

il presente regolamento non pregiudica pertanto l'applicazione della direttiva 2000/31/CE, in particolare le norme relative alla responsabilità dei prestatori intermediari di servizi di cui agli articoli da 12 a 15 della medesima direttiva.

e ai sensi dell'articolo 82, sancisce che:

Par. 1. Chiunque subisca un danno materiale o immateriale causato da una violazione del presente regolamento ha il diritto di ottenere il risarcimento del danno dal titolare del trattamento o dal responsabile del trattamento. Par. 2. Un titolare del trattamento coinvolto nel trattamento risponde per il danno cagionato dal suo trattamento che violi il presente regolamento. Un

responsabile del trattamento risponde per il danno causato dal trattamento solo se non ha adempiuto gli obblighi del presente regolamento specificatamente diretti ai responsabili del trattamento o ha agito in modo difforme o contrario rispetto alle legittime istruzioni del titolare del trattamento. Par. 3. Il titolare del trattamento o il responsabile del trattamento è esonerato dalla responsabilità, a norma del paragrafo 2 se dimostra che l'evento dannoso non gli è in alcun modo imputabile.

Ma in questa sede vogliamo concentrare la nostra attenzione sui cosiddetti *hosting providers*, intermediari che gestiscono servizi di memorizzazione di dati altrui attraverso piattaforme di social network, blog o altre tipologie di siti internet. La condotta di questi ultimi acquista rilievo giuridico nei casi in cui vengano caricati e diffusi per il tramite delle piattaforme da questi gestite contenuti aventi carattere lesivo di diritti altrui, in particolare i cosiddetti *diritti personalissimi*. L'articolo 16 del d.lgs. 70/2003 in particolare, stabilisce che l'hosting non è responsabile delle informazioni memorizzate se non è effettivamente a conoscenza del fatto che l'attività o l'informazione è illecita, oppure se, non appena a conoscenza di fatti illeciti, su comunicazione delle autorità competenti, si attivi immediatamente per rimuovere le informazioni o per disabilitarne l'accesso. Questo regime

di favore si applica al c.d. *hosting passivo*, che svolge un servizio di ordine meramente tecnico, automatico.

All'hosting passivo la giurisprudenza europea e nazionale ha poi affiancato nel corso del tempo la figura del c.d. *hosting attivo*. Con la sentenza n. 7708/2019 la prima sezione civile della Corte di Cassazione sancisce infatti che l'hosting provider attivo è il prestatore dei servizi della società dell'informazione, il quale svolge un'attività che esula da un servizio di ordine meramente tecnico, automatico e passivo, e pone, invece, in essere una condotta attiva, concorrendo con altri nella commissione dell'illecito, onde resta sottratto al regime generale di esenzione di cui all'art. 16, d.lgs. n. 70 del 2003, dovendo la sua responsabilità civile atteggiarsi secondo le regole comuni.

Nell'ambito dei servizi della società dell'informazione, la responsabilità dell'hosting provider, prevista dall'art. 16, d.lgs. 9 aprile 2003, n. 70, sussiste in capo al prestatore dei servizi che non abbia provveduto all'immediata rimozione dei contenuti illeciti, nonché se abbia continuato a pubblicarli, pur quando ricorrano congiuntamente le seguenti condizioni: a) sia a conoscenza legale dell'illecito perpetrato dal destinatario del servizio, per averne avuto notizia dal titolare del diritto

leso oppure aliunde; b) l'illiceità dell'altrui condotta sia ragionevolmente constatabile, onde egli sia in colpa grave per non averla positivamente riscontrata, alla stregua del grado di diligenza che è ragionevole attendersi da un operatore professionale della rete in un determinato momento storico; c) abbia la possibilità di attivarsi utilmente, in quanto reso edotto in modo sufficientemente specifico dei contenuti illecitamente immessi da rimuovere.

La responsabilità del provider trova quindi il proprio fondamento nel collegamento dell'obbligazione risarcitoria al fatto della conoscenza, da parte del prestatore del servizio, circa la illiceità dell'informazione, in particolare connotata dall'essere essa manifesta nelle azioni di risarcimento del danno.

Sul piano della responsabilità penale (Ingrassia 2012) invece, la possibilità di muovere un rimprovero all'ISP per i reati integrati nella rete delle reti deriva, oltre che dal fattore anonimato, dal ruolo fondamentale che riveste: permette l'ingresso nel cibernazio (*access e network provider*), ne agevola la circolazione (*mere conduit e cache provider*), e, soprattutto, ne disegna i tratti riempiendolo di contenuti (*hosting e content provider*).

In questo contesto si possono immaginare per l'ISP tre distinti paradigmi idealtipici di responsabilizzazione penale cui corrispondono, parallelamente, tre distinti ruoli sociali e tre diversi bilanciamenti dei diritti fondamentali in conflitto.

Il primo paradigma idealtipico, volto a massimizzare la libertà di comunicazione ed espressione del pensiero nel cyberspazio, pone l'ISP sul medesimo piano degli altri utenti di Internet, senza doveri di controllo rispetto a condotte altrui, obblighi di denuncia dei reati da altri commessi di cui viene a conoscenza o oneri di collaborazione con le autorità nella repressione degli illeciti: l'influsso sul piano penalistico è una responsabilità limitata alle ipotesi di autoria e di concorso commissivo doloso nell'altrui condotta criminosa. In tale paradigma il ruolo sociale dell'ISP nel cyberspazio è quello del comune cittadino.

Il secondo paradigma idealtipico, esattamente contrario al primo, sposta il riflettore sulla tutela dei soggetti terzi e della comunità e si caratterizza per le penetranti limitazioni alla libertà di comunicazione degli utenti della rete, realizzate attraverso una censura preventiva ad opera dell'ISP del materiale ivi caricato: il modello di responsabilità è quello del reato omissivo improprio

e, in particolare, il rimprovero è per non aver impedito un reato altrui. L'ISP assume così il ruolo sociale di controllore, di censore che decide ciò che può e ciò che non può esistere nel cibernazio.

Il terzo paradigma idealtipico si colloca a metà tra i precedenti. Rispetto al bilanciamento tra diritto di manifestazione del pensiero dell'utente e tutela dei terzi e della società, esso limita parzialmente la libertà di comunicazione, riducendo o escludendo l'anonimato degli utenti, e implementa strategie di repressione dei reati commessi che coinvolgono solo ex post l'ISP, imponendogli l'obbligo di denuncia degli illeciti di cui viene a conoscenza, l'onere di collaborazione nell'individuazione degli autori e il dovere di rimuovere il materiale illecito dal cibernazio. L'ISP non ha più l'obbligo di verificare, per così dire all'ingresso, il materiale presente in rete, ma di attivarsi per ridurre le conseguenze di reati già commessi e per agevolare la punizione degli autori (oltre che, ovviamente, di non commettere reati personalmente o in concorso con altri soggetti): il modello di responsabilità è quello del reato omissivo proprio e il rimprovero sarà per non aver tenuto condotte dirette a garantire la punizione dell'autore e la limitazione delle conseguenze del reato. L'ISP svolge la funzione di tutore dell'ordine nel cibernazio.

Il rapporto interno tra paradigmi è di responsabilità progressiva: nell'ipotesi di ISP cittadino questo risponderà solo dei reati di cui è autore o in cui ha offerto un contributo concorsuale attivo, senza obblighi di controllo o di collaborazione alla cui violazione consegua una sanzione penale; l'ISP tutore dell'ordine risponderà ove ometta la collaborazione con gli inquirenti, oltre che, ovviamente, nei casi di autoria e di concorso attivo; infine, l'ISP controllore risponderà non solo come cittadino a titolo commissivo, ma anche a titolo omissivo improprio, per non aver impedito il reato degli utenti della rete su cui ha doveri di controllo, e proprio, ove, a seguito della consumazione del reato che non ha impedito, non abbia cooperato con l'autorità nella repressione dell'illecito.

Quindi, per quanto concerne la condotta commissiva del prestatore di servizi, non vi sono dubbi: in questi casi, il comportamento commissivo penalmente rilevante deve consistere nell'immissione (o diffusione) dolosa di dati, a seguito di un vaglio contenutistico degli stessi, non potendo essere ritenuta sufficiente la mera predisposizione dei supporti tecnici di funzionamento della rete, la fornitura degli accessi o la gestione delle piattaforme online. Le attività normalmente attuate dagli

ISP vengono infatti considerate quali giuridicamente legittime e socialmente adeguate, non determinando di per sé il pericolo di verificazione dell'evento offensivo tipico⁶².

Di maggiore interesse sono i casi in cui la condotta attribuibile all'intermediario sia di natura omissiva, per la quale egli potrà rispondere o a titolo autonomo secondo il paradigma dell'art. 40 c.p. o, per combinato disposto con l'art. 110 c.p., a titolo di concorso omissivo nel reato realizzato dall'utente. La condotta omissiva del prestatore di servizi può consistere, da una parte, nel mancato controllo e censura preventiva del contenuto illecito caricato o diffuso da un proprio utente sulla piattaforma o, dall'altra parte, nella mancata rimozione dello stesso contenuto, pubblicamente accessibile sul proprio sito web.

Si profilano dunque due diverse ipotesi di responsabilità: una ex ante, operante prima che i dati vengano resi disponibili in rete; una ex post, legata alla fase di perdurante disponibilità che caratterizza le informazioni caricate nel web. La prima ipotesi è esclusa dalla maggioranza della

⁶²<https://archiviodpc.dirittopenaleuomo.org/upload/1351711435Il%20ruolo%20del%20ISP%20nel%20cyber-spazio%20DPC.pdf> (6.2.2025)

giurisprudenza⁶³ e della dottrina. Secondo l'orientamento dottrinale maggioritario⁶⁴ infatti, ad ostacolare l'attribuzione di una posizione di garanzia in capo a tali soggetti vi sarebbe, in primo luogo, il divieto di imporre un obbligo generale di controllo preventivo stabilito dalla direttiva e-commerce e attuato dall'art. 17 d.lgs. n. 70/2003, il quale prevede espressamente che i prestatori di servizi in rete non possano essere assoggettati ad un obbligo generale di sorveglianza sulle informazioni trasmesse o memorizzate, né ad un obbligo generale di ricercare attivamente fatti o circostanze che indichino la presenza di attività illecite.

Un simile obbligo richiederebbe ai provider di realizzare un'attività eccessivamente onerosa, inesigibile e tecnicamente irrealizzabile (un controllo e filtraggio di tutte le informazioni che passano attraverso un server sarebbe infatti imprati-

⁶³ Corte di giustizia europea, sentenza del 24 novembre 2011, C-70/10, *Scarlet Extended*, EU:C:2011:771; sentenza del 16 febbraio 2012, C-360/10, *SABAM*, EU:C:2012:85; Corte europea dei diritti dell'uomo, 9 marzo 2017, *Pihl c. Svezia*. Nella giurisprudenza nazionale basti richiamare il noto caso *Google c. Vividown*, Cass. Pen., Sez. III, 17 dicembre 2013, n. 5107, in *Dir. fam.*, 2014, 2, 675 ss.

⁶⁴https://www.sistemapenale.it/pdf_contenuti/1588789714_baccin-2020a-responsabilita-internet-service-provider-illeciti-online-caso-force-facebook.pdf (2.2.2025)

cabile). In secondo luogo, oltre a non ritenersi configurabile in capo a tali soggetti una posizione di protezione o di controllo (Bartoli 2013), essi non disporrebbero nemmeno di specifici poteri impeditivi che consentirebbero loro di impedire o interferire con il realizzarsi della condotta dell'autore del reato.

Per quanto riguarda la configurabilità di una forma di responsabilità penale *ex post* degli intermediari, quella parte di dottrina che non ritiene individuabile in capo agli ISP un obbligo di impedimento dei reati commessi dai propri utenti, non lo ritiene sussistente nemmeno in questa seconda fase.

Ai sensi della normativa vigente, vi sarebbero in capo agli ISP solamente due obblighi: un mero obbligo di *notice*, ossia di informazione dell'autorità competente del carattere illecito del contenuto del servizio ospitato (Manna, Di Florio 2019, 891-940), ed un obbligo di *take down*, ossia di rimozione del dato su richiesta dell'autorità competente; ma né l'uno né l'altro potrebbero essere considerati quali fonte di una posizione di garanzia.

Tuttavia, la Suprema Corte di cassazione, nella già citata sentenza n. 7708/2019, ha statuito che l'hosting provider, di fronte ad una situazione

di illiceità manifesta dell'altrui condotta, di cui non ne ha impedito la protrazione, mediante la rimozione delle informazioni o la disabilitazione all'accesso, risponderà per fatto proprio colpevole, essendogli rimproverabile una responsabilità commissiva mediante omissione, per avere concorso nel comportamento lesivo altrui. Ed una tale ricostruzione è resa possibile in virtù del fatto che l'art. 16 d.lgs. n. 70 del 2003 fonde una cosiddetta posizione di garanzia dell'hosting provider, che, se per definizione è indispensabile alla stessa originaria perpetrazione dell'illecito del destinatario del servizio, ne diviene giuridicamente responsabile solo dal momento in cui gli possa essere rimproverata l'inerzia nell'impedirne la protrazione.

Una parte minoritaria della dottrina (Picotti 2007, 1196-1211; Torre 2013, 163-191) ha inoltre evidenziato come vi siano da tempo disposizioni normative extra-penali che prevedono obblighi specifici in capo agli ISP in materia di reati di diffusione di materiale pedopornografico, in materia di violazioni di diritto d'autore e nello stesso d.lgs. n. 70/2003, capaci di integrare in capo al prestatore di servizi una responsabilità penale per reato omissivo improprio.

Dal canto suo, la giurisprudenza di legittimità si è espressa offrendo nel corso del tempo soluzioni divergenti, in un primo momento escludendo la sussistenza di una posizione di garanzia in capo ai providers⁶⁵, successivamente ritenendo responsabile del reato di diffamazione il legale responsabile di una società IT per omesso impedimento degli effetti del reato⁶⁶, non essendosi adoperato per eliminare il contenuto diffamatorio non appena avutone notizia.

Si può quindi affermare che gli hosting providers, in virtù della posizione di signoria che rivestono nei confronti dei dati trattati sulle proprie piattaforme, hanno la pesante responsabilità, nei confronti della società, di proteggere gli utenti e il pubblico in generale, nonché prevenire lo sfruttamento dei loro servizi da parte di criminali e altri soggetti coinvolti in attività illegali online (Sgubbi 1975; Leoncini 1999)⁶⁷.

Guardando in effetti al contenuto sostanziale, prima ancora che formale, dell'obbligo giuridico di impedimento, esso trova valido fondamento nello stesso dato di fatto che le vittime del

⁶⁵ A tal proposito, si veda la sentenza *Google vs. Vividown* (Cass. Pen., Sez. III, 17 dicembre 2013, n. 5107)

⁶⁶ Cass. Pen., Sez. V, 27 dicembre 2016, n. 5494

⁶⁷https://dpc-rivista-trimestrale.criminaljusticenetwork.eu/pdf/DPC_Riv_Trim_2_2019_panattoni.pdf (6.2.2025)

crimine cibernetico – in particolare quando si tratta di reati che si consumano interamente nel mondo virtuale – si trovano nell'impossibilità di proteggere il bene giuridico leso, dato che un utente non ha possibilità di rimuovere un contenuto illecito una volta che questo è immesso in rete. Inoltre, la dimensione sociale che ha assunto il cyberspazio, strumento attraverso cui si esercitano i più diversi diritti anche fondamentali, fa assumere al suo corretto e buon utilizzo il valore di interesse diffuso, la cui tutela può costituire ulteriore ragione per il ricorso al meccanismo proprio delle fattispecie omissive improprie.

Passando al profilo formale, l'attribuzione di una posizione di garanzia in capo agli ISP può trovare riscontro, come già rilevato, nella previsione di obblighi giuridici aventi fonte normativa, primo fra tutti quello delineato, per gli hosting providers, dall'art. 16 comma 1 d.lgs. n. 70/2003, il quale, tuttavia, andrebbe, oltre che meglio articolato, ritenuto applicabile a tutti gli hosting providers.

Gli Stati Uniti invece, in materia di responsabilità degli ISP, sono sempre stati restii nell'addossare responsabilità in capo ai fornitori di servizi per i contenuti odiosi che circolano in rete. Questo è quanto emerge anche dalla Section 230

del *Communication Decency Act*, che esclude espressamente la responsabilità dei fornitori di servizi per ciò che immettono nella Rete, esonerandoli indirettamente, così, da qualsiasi indagine preventiva sulle informazioni che circolano sul web immesse da terzi. L'adozione di tale politica trova la propria giustificazione sia nell'idea americana, esposta in precedenza, della rete Internet come mezzo neutrale al fine della diffusione del pensiero, sia nella constatazione dell'impossibilità oggettiva di monitorare ogni giorno l'enorme flusso di dati che circolano sul web. I giudici oltreoceano fanno infatti costante applicazione del Titolo 47 U.S. Code §230 (c) (1), il quale così dispone:

(c)Protection for “Good Samaritan” blocking and screening of offensive material (1) Treatment of publisher or speaker: No provider or user of an interactive computer service shall be treated as the publisher or speaker of any information provided by another information content provider.

Infine, è necessario rilevare come il panorama della responsabilità in capo agli ISP sia cambiato quando gli studiosi hanno iniziato a valutare l'influenza degli algoritmi sulle interazioni tra gli utenti. Autorevole dottrina (Accinni 2017, 11) so-

stiene infatti che Facebook, avendo utilizzato algoritmi di filtraggio e di associazione al fine di maneggiare i contenuti immessi da terzi nella piattaforma, sia andato oltre il ruolo di semplice editore, giocando al contrario il ruolo attivo di chi si inserisce nell'iter esecutivo della pubblicazione, almeno parzialmente e prosegue quello messo in atto dall'autore primario o originario, emergendo quale indispensabile anello causale nella diffusione del contenuto illecito.

Anche la giurisprudenza italiana si è espressa a riguardo mediante quello che viene considerato il leading case in materia: il caso RTI vs. Yahoo!⁶⁸, deciso dal Tribunale di Milano. Nel caso di specie, i giudici definiscono l'hosting provider come il prestatore dei servizi della società dell'informazione il quale svolge un'attività che esula da un servizio di ordine meramente tecnico, automatico e passivo, e pone, invece, in essere una condotta attiva, concorrendo con altri nella commissione dell'illecito, escludendo così l'applicazione del regime di esenzione previsto dal già menzionato articolo 16 del d.lgs. 70/2003. In tal modo, nasce una vera e propria presunzione di co-

⁶⁸ Trib. Milano, Sez. Spec. Propr. Ind. e Intellettuale, 9 settembre 011, n. 10893

noscenza effettiva in capo al provider del contenuto postato dall'utente, nel momento in cui il primo si avvalga di un gruppo editoriale al fine di diffondere contenuti, sfrutti sistemi automatizzati e svolga attività di gestione e di organizzazione del materiale postato.

Alla luce di questi elementi, si può quindi pensare che sussista una qualsivoglia forma di responsabilità in capo agli ISP? Il problema dominante concerne il carattere totalmente automatizzato degli algoritmi, concepiti quali strumenti utilizzati dai programmatori, ad esempio, per porre un commento razzista che ha ricevuto centinaia di like in cima ai commenti per far sì che goda della massima esposizione, così come per rendere virale nel giro di poche ore un post diffamatorio oggetto di reiterate condivisioni perché suggerito nelle bacheche degli amici di chi lo abbia condiviso o abbia messo like. Essendo l'automatismo elemento intrinseco degli algoritmi, come si fa ad addossare una responsabilità di qualsivoglia tipo agli umani perché non fanno o non hanno impedito l'illecito, quando è l'algoritmo ad aver commesso ex se il fatto?

Sul punto, una possibile soluzione viene fornita dalla teoria della cosiddetta *responsabilità dell'utente finale*, ideata in prima battuta da Peter

Asaro (2012), poi ripresa in territorio nazionale da Giovanni Sartor (2003). Essa poggia le proprie basi sulla cosiddetta *vicarious liability* degli ordinamenti anglosassoni⁶⁹ e intende dare accesso, nell'ordinamento penale italiano, ad una forma di responsabilità analoga a quella di padroni e committenti per il danno cagionato dai propri sottoposti già esistente nel diritto civile (ex. art. 2049 c.c.). In questo modo, degli effetti prodotti dall'agente dovrebbe rispondere l'utilizzatore, colpevole, in qualche modo, di aver confidato nelle capacità razionali della macchina e nella sua abilità di ottenere un certo risultato. Quest'ultimo sarebbe quindi chiamato a rispondere a titolo di colpa cosciente, avendo fatto affidamento su un sistema di cui non conosceva appieno il funzionamento e confidato nella propria capacità di gestire gli eventi. Tuttavia, si ritiene in questa sede più plausibile riconoscere una sorta di *colpa da programmazione* in capo all'umano e contestualmente una scusante nel caso in cui si adottino tutte le misure idonee ad impedire il fatto (Magro 2014).

⁶⁹Una peculiare forma di responsabilità oggettiva, vigente in ambito penale, che punisce chi si avvalga di persone o mezzi per il raggiungimento di un determinato scopo, essendo del tutto irrilevante la sussistenza in capo a costui di uno stato soggettivo di colpevolezza.

Una forma di responsabilità correlata a quella degli ISP concerne il problema della censura infamante e automatizzata dei contenuti operata proprio da quegli algoritmi di cui parlavamo in precedenza. Un caso che ha creato grande scalpore a riguardo ha visto coinvolto il canale *Youtube Radio Radio Tv*, il quale è stato chiuso dalla nota piattaforma di video sharing con decisione unilaterale per violazione delle Norme della community. Secondo Youtube, infatti, Radio Radio avrebbe pubblicato video contenenti minorenni in situazioni sessualmente allusive, e visto che per il team di Youtube è una questione di primaria importanza Radio Radio TV viene chiusa, sostenendo altresì che YouTube non è luogo per comportamenti predatori. Radio Radio, dal canto suo, ritiene l'accusa totalmente infondata e calunniosa, peraltro essendo basata su segnalazioni assolutamente fantasiose.

La cancellazione del canale Youtube di Radio Radio ha sollevato numerose critiche sui social dove in molti hanno gridato al complotto. Tra le voci in difesa di Radio Radio quella del filosofo Diego Fusaro che ha twittato: "Il bello della democrazia di mercato: puoi dire tutto ciò che vuoi, a patto che tu voglia ciò che vogliono loro". Il caso Ra-

Radio Radio ha avuto una risonanza tale da interessare non solo il Parlamento italiano, ma anche i deputati europei. È stata infatti presentata un'interrogazione presso le istituzioni europee da parte di tre eurodeputati, Marco Campomenosi, Marco Zanni e Antonio Maria Rinaldi. Il primo dei firmatari, in un'intervista rilasciata proprio a Radio Radio, ha dichiarato che "ci sono persone e aziende organizzate nel cosiddetto *shitstorm* di segnalazioni che poi di fatto vanno a creare delle problematiche che scattano in automatico, non essendoci un controllo umano su queste cose", e ha concluso sostenendo che "gli algoritmi non possono controllare la libertà d'espressione".

Il canale Radio Radio è stato ripristinato in 24 ore ma rimane il gesto fatto da Youtube. Sul punto, il direttore di Radio Radio Ilario Di Giovambattista ha dichiarato che "rimane il gesto fatto da YouTube. Rimane la facilità con la quale ti fanno sparire. Rimane soprattutto l'onta di una calunnia alla quale YouTube ha creduto. Un danno all'immagine perché Radio Radio non ha mai mandato in onda filmati di bambini in posizioni strane. Radio Radio è a difesa dei bambini e dell'infanzia. Un'infamia che ha danneggiato l'immagine di Radio Radio per chi non ci conosce". Proseguendo che "da questa storia resta la digitalizzazione ossessiva che

ci porterà grandissimi danni. La tecnologia è ovviamente straordinaria, ma va utilizzata per il bene, non per il male. Attenzione all'esclusione del rapporto umano: nel nostro caso bastava un funzionario di YouTube che fosse andato a vedersi i video per accorgersi dell'errore".

E questo ci riporta inevitabilmente alla questione originaria: gli algoritmi possono in concreto aiutare a combattere fenomeni quali ad esempio l'hate speech o, al contrario, rappresentano in realtà uno strumento pericoloso in grado di limitare in modo ingiustificato la nostra libertà d'espressione?

Ammettendo poi i benefici e i vantaggi che possono derivare dall'utilizzo degli algoritmi per scovare nel web fenomeni di hate speech, un ulteriore problema risiede nel capire chi debba decidere se un dato commento possa essere definito hate speech e quindi possa essere eliminato dall'algoritmo. E in caso di errore sul punto, in capo a chi può essere riconosciuta una qualsivoglia forma di responsabilità? Al singolo programmatore che crea l'algoritmo, all'ISP che gestisce la piattaforma o all'utente che ha espresso il proprio pensiero?

Ma, per comprendere a fondo come l'AI abbia in concreto inciso sul fenomeno che qui stiamo

affrontando in relazione soprattutto al mondo dei social networks, dobbiamo in via preliminare capire cosa intendiamo quando si parla di AI (Artificial Intelligence) e come questa sia diventata parte integrante del problema legato alla diffusione online di messaggi odiosi.

3.4. AI nei Social Networks: i possibili rischi

Con lo scorrere del tempo, l'AI viene sempre di più impiegata nelle fasi di raccolta, analisi e utilizzo delle informazioni relative all'attività di una certa organizzazione, con lo scopo di acquisire specifiche conoscenze che permettano di supportare procedure decisionali in grado di migliorare le prestazioni dell'organizzazione stessa (Vercellis 2009, 3). Si parla a tal proposito di *Business Intelligence* (Moro, Sarra 2017). Tale scienza permette, in particolare alle aziende, di utilizzare l'infinito flusso di dati presenti sui social networks, al fine di identificare argomenti di tendenza, hashtag e schemi per comprendere il comportamento degli utenti e le tendenze di riferimento in quel preciso momento. Gli algoritmi di AI, infatti, secondo quanto rilevato da Michele Gentili⁷⁰, possono monitorare milioni di commenti degli utenti che per

⁷⁰<https://www.agendadigitale.eu/cultura-digitale/social-network-e-intelligenza-artificiale-presente-e-futuro-di-un-forte-binomio/> (2/12/2024)

loro natura sono destrutturati e aiutare a comprendere situazioni o tendenze di crisi o di particolare euforia o attenzione per un determinato argomento e consentire così un'esperienza personalizzata.

Ma nel concreto come agiscono tali strumenti matematici? Prendiamo ad esempio la piattaforma UCF (Universal Content Filtering), che costituisce l'infrastruttura pensata per il monitoraggio e la classificazione dei contenuti presenti sulla piattaforma LinkedIn. Gli algoritmi si intrecciano con attività che devono svolgere gli umani e con altre che vengono svolte in maniera autonoma. In primis, viene creato un contenuto che viene poi sottoposto ad un'analisi qualitativa mediante l'utilizzo di classificatori che sono a disposizione per quel determinato tipo di contenuto. Se quest'ultimo ottiene luce verde, allora si continuerà a visualizzare quel contenuto sul sito, e il contenuto stesso arriverà sul feed. Se, al contrario, un classificatore dice che è sospetto, allora può venir attivato, secondo una soglia prestabilita, il ciclo di classificazione umana. È quindi l'umano a dire cosa pensa di quel determinato contenuto, se, secondo lui, gli algoritmi di AI stanno prendendo la decisione giusta o no. Gli esseri umani intervengono dando il loro feedback, che a quel punto sarà

definitivo, con grande responsabilità. Se il contenuto sospetto venisse classificato come buono, allora si tornerebbe indietro indicando che il classificatore basato su AI ha fatto un errore. L'errore verrà registrato e la prossima volta l'insegnamento migliorerà l'algoritmo e la tecnica di apprendimento, con più dati disponibili per prendere la decisione giusta autonomamente.

Tale meccanismo è stato posto alla base del progetto *CONAN - COUNTER Narratives through Nichesourcing: a Multilingual Dataset of Responses to Fight Online Hate Speech*⁷¹, pensato allo scopo di utilizzare l'AI per risolvere il problema dell'hate speech. I ricercatori del progetto partono dall'idea che, nonostante i numerosi tentativi da parte del legislatore, e altresì degli ISP, di regolamentare il fenomeno del discorso d'odio mediante l'uso di strumenti di censura e di eliminazione dei contenuti d'odio su segnalazione degli utenti, il problema dell'hate speech sia tutt'altro che risolto. Per questo motivo, questi studiosi, proprio in ossequio al principio di libertà di espressione del pensiero trattato in precedenza, hanno pensato di far uso delle cosiddette *contro-narrazioni*. Una contro-narrazione, metodo sposato soprattutto dagli studiosi americani, consiste appunto in una risposta che prevede riscontri

⁷¹ <https://arxiv.org/abs/1910.03270>

non negativi attraverso argomenti legati ai fatti⁷² ed è considerata lo strumento migliore per affrontare il fenomeno del discorso d'odio. Il progetto si ispira allo sforzo di alcune ONG che stanno addestrandolo degli operatori e dei volontari a monitorare le piattaforme social e a produrre contro-narrative appropriate quando necessario.

Proprio tre di queste ONG hanno fornito delle vere e proprie linee guida per l'utilizzo del cosiddetto *counterspeech*. Queste istruzioni esortano a: 1) non essere offensivi: prima di inviare una risposta, assicurarsi che la risposta non diffonda alcun odio, bigottismo, pregiudizio o contenuti illegali. Occorre mantenere le conversazioni pacifiche e non degenerare in un conflitto. Stiamo parlando di persone, non di categorie; 2) pensare agli obiettivi: prima di scrivere una risposta, occorre pensare all'effetto che può generare ed a ciò che si desidera ottenere. Prestare attenzione agli obiettivi contribuirà ad utilizzare parole corrette; 3) richiedere utenti influenti: arruolare sostenitori in-

⁷²https://www.researchgate.net/profile/Carla-Schieb/publication/303497937_Governing_hate_speech_by_means_of_counterspeech_on_Facebook/links/5761575408aeeada5bc4f783/Governing-hate-speech-by-means-of-counterspeech-on-Facebook.pdf
(7.2.2025)

fluenti (leader civici, politici, esperti) aiuterà ad attirare l'attenzione e ad approfondire l'effetto della contronarrazione; 4) utilizzare prove credibili: l'informazione nel discorso di odio può essere confusa e fuorviante. Chiedere chiarimento se necessario. Controbattere con prove credibili e utilizzare fonti affidabili; 5) pensare al tono: possiamo dimostrare comprensione e sostegno a chi potrebbe essere attaccato. Fare attenzione a usare sarcasmo, umorismo, parodia e satira. Possiamo usarli, solo se siamo in grado di dominarli.

Vediamo ora un esempio. Se ci troviamo di fronte alla frase "Io odio i Musulmani. Loro non dovrebbero esistere", la corretta contro-narrativa per far sì che il discorso d'odio venga bloccato sul nascere può per esempio consistere nella frase "Anche i Musulmani sono umani. Le persone hanno la libertà di scegliere la propria religione". A differenza di quest'ultima, l'espressione "Sei davvero uno stupido idiota che pensa in modo retrogrado per credere alla negatività dell'Islam" non costituisce una frase efficace in vista del raggiungimento dello scopo della contro-narrativa, cioè l'eliminazione dell'hate speech, dal momento che fa degenerare la conversazione.

Oltre a questo ammirevole progetto, al fine di combattere il discorso si ritiene inevitabile coinvolgere maggiormente gli ISP sopramenzionati affinché si adoperino per sopperire alle lacune umane in materia, attraverso strumenti tecnologici avanzati (Di Tano 2017, 113-126).

Il primo intermediario della Rete ad essersi chiesto che cosa potrebbe fare in concreto la tecnologia per migliorare i discorsi online, è stato Google. L'azienda statunitense ha così creato *Perspective*⁷³, una soluzione tecnologica che prevede l'utilizzo di modelli di apprendimento automatico per rilevare di default insulti, molestie e parole ingiuriose online, valutandone il grado di nocività in maniera più accurata e veloce di qualsiasi moderatore umano. Il software è stato addestrato mediante più di 17 milioni di commenti etichettati come hate speech dai moderatori umani, che hanno insegnato all'intelligenza artificiale del programma quali fossero le parole offensive. Nello scandagliare la Rete e nel rintracciare commenti offensivi, *Perspective* confronta le parole utilizzate per la sua formazione, al fine di osservare la somiglianza con quelle qualificate come nocive o idonee ad allontanare persone o ad impedire conversa-

⁷³ <http://www.perspectiveapi.com> (2/12/2024)

zioni sane. Quando l'intelligenza artificiale rinviene nuovi esempi di commenti potenzialmente offensivi, o viene introdotta una correzione da un utente umano, il programma è in grado di migliorare autonomamente la propria capacità di qualificare i commenti.

Qualcuno potrebbe obiettare che l'uso di software siffatti non rappresenti una grande attrazione per le tech companies, visti i costi di implementazione e l'aumento in capo alle stesse del grado di responsabilità che discende da un miglioramento dello standard dei controlli. Vi è però da dire che, se per il breve termine i discorsi d'odio sul web possono essere fonte di curiosità per gli utenti e quindi aumentare gli introiti delle aziende digitali, nel medio-lungo termine l'odio non fa altro che allontanare le persone. Per questa ragione, gli ISP accolgono con favore gli strumenti di AI che fungono da supporto agli operatori umani, al fine di combattere il fenomeno dell'hate speech.

Si ritiene tuttavia prematuro affidare completamente alle macchine il controllo dell'hate speech online, essendo necessario valutare il contesto in cui l'espressione viene utilizzata e il significato che quell'espressione assume nell'ambiente stesso. In merito si è espresso Angelo Marcello Cardani, Presidente dell'Autorità per le Garanzie nelle

Comunicazioni, nel corso della tavola rotonda *Hate Speech and Artificial Intelligence Tools* tenutasi a Bruxelles presso il *Center on Regulation in Europe* il 31 gennaio del 2019. Ha affermato che “l’Agcom si impegna a includere il rilevamento automatizzato e il monitoraggio dell’hate speech online in un quadro regolatorio in cui l’Intelligenza Artificiale non sostituisce la valutazione umana nel perseguimento dei crimini di odio ma è limitata alla prevenzione e al monitoraggio”. Sostenendo altresì che “gli strumenti di Intelligenza Artificiale esistenti necessitano di una messa a punto prima di essere implementati per un’efficace individuazione automatica dei contenuti di incitamento all’odio”.

Dal canto loro, alcuni studiosi tedeschi (Schmidt, Wiegand 2017, 1-10) hanno condotto un’indagine in merito al rilevamento di hate speech online attraverso l’uso del cosiddetto *Natural Language Processing* (NLP). A tal proposito, gli studiosi rilevano come, a seconda delle caratteristiche del linguaggio usate, si possono distinguere differenti approcci per la risoluzione del problema in questione.

Per compiere la classificazione di un testo, le informazioni più ovvie da usare sono costituite

dalle semplici caratteristiche che connotano la superficie di una determinata espressione, come un mucchio di parole. Infatti, la maggior parte degli autori⁷⁴ (Chen, Zhou, Zhu, Xu 2012, 71-80) (Nobata, Tetreault, Thomas, Mehdad, Chang 2016, 145-153) sceglie di usare uni grammi, cioè n-grammi di lunghezza 1. Un n-gramma è una sotto-sequenza di n elementi di una data sequenza. Tali elementi possono essere fonemi, sillabe, lettere, parole. I programmatori fanno per esempio uso dei cosiddetti approcci *character-level* sugli n-grammi, al fine di risolvere i problemi di variazioni nello spelling di una parola – si pensi alla frase odiosa “ki11 yrslef a\$\$hole” scritta in questo modo mimetico.

Tali metodi sono sicuramente più efficaci rispetto ai cosiddetti approcci *token-based*, dal momento che i secondi, nel programmare l’algoritmo, abbinano ad un numero (si fa qui riferimento al cosiddetto sistema binario 0-1) l’intera parola (esempio “yrslef”), non riuscendo così a distinguerla rispetto alla parola *yourself* e avendo pochi esempi da utilizzare per svolgere la propria attività. I primi invece abbinano ad un numero le singole lettere,

⁷⁴https://www.researchgate.net/publication/262363934_Detecting_hate_speech_on_the_world_wide_web (1.2.2025)

così da essere più efficienti e da non confondere le parole.

Un altro metodo⁷⁵ che viene utilizzato ai fini della risoluzione del problema dell'hate speech comporta l'uso di generalizzazioni di parole. Viene quindi effettuato un raggruppamento di parole e usati ID clusters che rappresentano insiemi di parole come caratteristiche aggiuntive. La cosiddetta *sentiment analysis* funge poi da ausilio al fine di operare la classificazione. In particolare, alcuni studiosi⁷⁶ utilizzano come caratteristiche il numero di parole positive, negative e neutrali che si rinven- gono in un dato commento social.

Ulteriori tentativi (Thelwall, Buckley, Palto- glou, Cai 2010, 2544–2558) al fine di isolare l'hate speech da altre parole che presentano una polarità negativa, si concentrano sull'alto livello di polarità negativa che il primo presenta rispetto alle se- conde. Un classificatore di polarità è ad esempio *SentiStrength*. Partendo dal presupposto che l'hate speech si connota per la presenza di specifiche pa- role negative come gli insulti, molti autori sfrut- tano codeste parole come caratteristiche. E per fare ciò, si avvalgono di risorse lessicali che sco- vano nel web e che contengono risorse predittive.

⁷⁵file:///C:/Users/beatr/Downloads/Class-Based_n-gram_Models_of_Natural_Language.pdf (6.2.2025)

⁷⁶ <https://aclanthology.org/R15-1086.pdf> (1.2.2025)

In rete, infatti, sono presenti sia liste di termini generici che riguardano l'odio⁷⁷, sia liste di termini specializzate verso un tipo particolare di hate speech, come ad esempio l'odio politico. Altre caratteristiche che sono prese in considerazione riguardano aspetti linguistici, in particolar modo sintattici. A tal proposito, vengono impiegate relazioni di dipendenza digitali. Tali rapporti infatti hanno il vantaggio che parole non consecutive che sopportano un rapporto a lunga distanza, possono essere catturate in un'unica caratteristica. Ad esempio, nella frase "Jews are lower class pigs", possiamo notare il legame di dipendenza tra il termine offensivo *pigs* e il bersaglio d'odio *Jews*.

Tuttavia, il monitoraggio dell'hate speech non può risolversi solo nella ricerca superficiale delle parole-chiave. Infatti, nonostante l'applicazione dei metodi sopra citati, rimane comunque difficile in alcuni casi distinguere se si tratti di hate speech o meno. Ad esempio, la frase "mettiti una parrucca e un rossetto e sii chi sei veramente" può non costituire hate speech solo nel caso in cui venga letta da sola. Se al contrario, tale espressione viene destinata ad un ragazzo presente su un fo-

⁷⁷www.noswearing.com/dictionary,www.rpdb.org,ww.hate-base.org (2/12/2024)

rum di adolescenti su Facebook, qualcuno (Dinaka, Jones, Havasi, Lieberman, Picard 2012) potrebbe sostenere che sia un'espressione usata per denigrare la sessualità o l'identità di genere del ragazzo destinatario della frase. Tale esempio serve per dimostrare che l'analisi dell'hate speech e la sua potenziale rimozione non può basarsi solo su aspetti inerenti al linguaggio. È invece necessario fare ricorso alla conoscenza del mondo.

A riguardo, alcuni studiosi hanno infatti adottato un approccio che prevede l'impiego del ragionamento automatico sulla conoscenza del mondo, applicato all'hate speech anti-LGBT. La base di partenza di questo metodo è costituita dalla generale ontologia *ConceptNet*, la quale codifica concetti che sono connessi mediante relazioni per formare asserzioni, come ad esempio "la minigonna è un tipo di abbigliamento femminile". *ConceptNet*⁷⁸ viene poi arricchito manualmente da una serie di stereotipi estratti dal social network *Formspring*. Un esempio di asserzione può essere "il rossetto è usato da ragazze".

La conoscenza aggiuntiva viene denominata *BullySpace*. Questa base di conoscenza permette di calcolare la somiglianza dei concetti di conoscenza

⁷⁸<https://www.media.mit.edu/publications/bttj/Paper23Pages211-226.pdf> (9.2.2025)

comune con i concetti espressi dall'utente. Dopo l'estrazione dei concetti presenti in un dato commento dell'utente, viene calcolata la somiglianza tra i concetti estratti e un insieme di quattro concetti canonici. I concetti canonici sono i quattro concetti di riferimento di valore positivi e negativi e i due sessi, maschio e femmina. I punteggi di somiglianza risultanti tra concetti estratti e concetti canonici, indicano se un messaggio può costituire un'istanza di incitamento all'odio.

Ma la conoscenza del mondo non è l'unico tipo di conoscenza che possiamo utilizzare al fine di perfezionare le classificazioni inconcluse. La cosiddetta *meta-informazione* (l'informazione su un'espressione) può essere una valida fonte di ricerca per i fini che qui ci interessano. Infatti, avere qualche informazione sull'utente che pubblica un post, può essere predittivo. Un utente che è noto per scrivere messaggi di odio può fare così di nuovo. Un utente che non è noto per scrivere ciò è improbabile che scriva tali messaggi in futuro.

Alcuni ricercatori (Xiang, Fan, Wang, Hong, Rose 2012) hanno adoperato questa euristica per inferire ulteriori messaggi d'odio. Conoscere ad esempio il sesso di una persona può essere molto utile (Dadvar, MG de Jong, Ordelman, Trieschnigg 2012, 22-25). A livello statistico, si è osservato che

gli uomini tendono a diffondere più spesso messaggi d'odio rispetto alle donne. Altri tipi di meta-informazioni possono essere costituiti dal numero di post di un utente o dal numero di risposte ad un post (Zhong, Li, Squicciarini, Rajtmajer, Griffin, Miller, Caragea 2016, 3952–3958). I classificatori più utilizzati invece sono *Support Vector Machines*.

Ovviamente, per condurre al meglio esperimenti sulla rilevazione di hate speech in rete, l'accesso a corpora etichettati è fondamentale. I corpora linguistici sono collezioni di testi orali o scritti prodotti in contesti comunicativi reali, conservati in formato elettronico e spesso corredati di strumenti di consultazione informatici. Dal momento che non esiste attualmente un corpus di riferimento per quanto riguarda tali dati, i ricercatori collezionano ed etichettano i propri dati. Ma da dove provengono queste informazioni? Dai social networks, quali Instagram (Hosseinmardi, Arredondo Mattson, Rafiq, Han, Lv, Mishra 2015), Yahoo! (Djuric, Zhou, Morris, Grbovic, Radosavljevic, Bhamidipati 2015) e Twitter (Burnap, Williams 2015), per citarne alcuni.

Non è ancora chiaro, tuttavia, se i metodi di cui abbiamo parlato siano in grado di riconoscere determinate affermazioni, quali ad esempio “Kermit, la rana ha chiamato e riuole la sua voce, la tua

capra sta chiamando” come insulto ad un musulmano, oppure “chi è responsabile di aver detto che queste ragazze sono belle?”.

Alla luce di tali considerazioni, auspichiamo che si possa quindi riuscire a sfruttare al meglio le potenzialità dell’AI in relazione al fenomeno dell’hate speech, proponendo tuttavia non una totale sostituzione degli operatori umani con le macchine bensì l’unione delle forze al fine di dare un freno ai messaggi odiosi circolanti sul web.

3.6. Conclusioni

Alla luce degli elementi analizzati, in questa sede si ritiene che al fine di combattere il discorso d’odio veicolato attraverso la rete non sia sufficiente l’intervento, sebbene efficace, degli strumenti di AI sul linguaggio utilizzato. Occorre altresì andare all’origine del problema: la mancata neutralità di fatto della rete. Infatti, nonostante formalmente si enunci a gran voce il rispetto del principio di neutralità della rete, il quale altro non è che un vero e proprio divieto posto in capo agli operatori di rete di attuare trattamenti discriminatori, la realtà concreta è purtroppo lontana dagli enunciati di principio. Infatti, come abbiamo visto in precedenza, diversi casi presenti in giurispru-

denza mostrano come in alcune situazioni gli operatori di rete si comportino come veri e propri protagonisti nelle interazioni tra gli utenti, trasformando e utilizzando Internet come un'emittente privata. Ciò non fa altro che collocare tali soggetti in una posizione di supremazia rispetto a tutti gli altri cittadini e consente loro di scegliere quali siano i messaggi da diffondere e quali no, e tra questi anche i messaggi d'odio. Per questo motivo, è necessario fare rispettare in concreto il principio di neutralità della rete, in modo tale che nasca, anche all'interno di un luogo intangibile quale è Internet, un vero e proprio sistema di democrazia digitale, caratterizzato in quanto tale non solo da un'eguaglianza di tipo formale ma soprattutto da un'eguaglianza di carattere sostanziale.

Questi termini in realtà non ci suonano nuovi. Ai sensi dell'articolo 3, la Costituzione italiana, ai commi 1 e 2, recita difatti quanto segue:

Tutti i cittadini hanno pari dignità e sono eguali davanti alla legge, senza distinzione di sesso, di razza, di lingua, di religione, di opinioni politiche, di condizioni personali e sociali.

È compito della Repubblica rimuovere gli ostacoli di ordine economico e sociale, che, limitando di fatto la libertà e l'eguaglianza dei cittadini, impediscono il

pieno sviluppo della persona umana e l'effettiva partecipazione di tutti i lavoratori all'organizzazione politica, economica e sociale del Paese.

È auspicabile allora dare vita a quella che potremmo definire una *Costituzione digitale*, la quale raccolga tutti quei principi, a partire dal principio di neutralità della rete, che, d'ora in avanti, possono costituire il faro che illumina le nostre traversate sul web verso mete più eque e giuste.

A mio avviso, dunque, l'unione tra il diritto, sicuramente rinnovato e mutato alla luce delle nuove scoperte tecnologiche e delle nuove interazioni tra individui, e la formazione dei singoli, a partire dall'infanzia, a diventare cittadini responsabili e culturalmente consapevoli, può costituire una soluzione concreta per limitare il più possibile il fenomeno del discorso d'odio, non solo online.

Capitolo Quarto

Il cyberbullismo

4.1 Inquadramento della fattispecie

Per comprendere in maniera accurata quali siano le condotte che costituiscono cyberbullismo, è dapprima necessario riprendere la definizione che di tale fenomeno ci fornisce la normativa di riferimento: la legge n. 71 del 2017. Per cyberbullismo s'intende:

Qualunque forma di pressione, aggressione, molestia, ricatto, ingiuria, denigrazione, diffamazione, furto d'identità, alterazione, acquisizione illecita, manipolazione, trattamento illecito di dati personali in danno di minorenni, realizzata per via telematica, nonché la diffusione di contenuti on line aventi ad oggetto anche uno o più componenti della famiglia del minore il cui scopo intenzionale e predominante sia quello di isolare un minore o un gruppo di minori ponendo in atto un serio abuso, un attacco dannoso, o la loro messa in ridicolo.

Di primo acchito, è evidente che gli elementi costitutivi della fattispecie in questione siano, oltre alle diverse condotte elencate, anche: i soggetti passivi di queste ultime (i minori), la realizzazione dei comportamenti in oggetto mediante l'utilizzo

della Rete e l'elemento soggettivo del dolo specifico, in specie, l'isolamento del minore o di un gruppo di essi, che comporti altresì un serio abuso, un attacco dannoso, o la loro messa in ridicolo.

Al fine di rafforzare la prevenzione e il contrasto del cyberbullismo e del bullismo e potenziare la protezione delle vittime, il Parlamento ha recentemente approvato la legge n. 70 del 17 maggio 2024. Questa norma estende espressamente l'applicazione della legge del 2017 anche al fenomeno bullismo. Una delle principali novità è, infatti, l'introduzione della definizione di bullismo, che include aggressioni o molestie reiterate, da parte di una singola persona o di un gruppo di persone, in danno di un minore o di un gruppo di minori, idonee a provocare sentimenti di ansia, di timore, di isolamento o di emarginazione, attraverso atti o comportamenti vessatori, pressioni o violenze fisiche o psicologiche, istigazione al suicidio o all'autolesionismo, minacce o ricatti, furti o danneggiamenti, offese o derisioni.

Viene altresì previsto dalla nuova legge il potenziamento del servizio di supporto psicologico agli studenti, consentendo alle Regioni di attivare, presso le istituzioni scolastiche, un servizio di assistenza psicologica per favorire lo sviluppo e la formazione degli studenti e prevenire situazioni di

disagio, anche attraverso il coinvolgimento delle famiglie.

La legge n. 70/2024 introduce anche nuove norme per la prevenzione e il contrasto del bullismo e del cyberbullismo negli istituti scolastici, prevedendo che ogni scuola istituisca un tavolo permanente di monitoraggio composto da rappresentanti degli studenti, degli insegnanti, delle famiglie ed esperti di settore e che adotti un codice interno per la prevenzione e il contrasto di questi fenomeni. Inoltre, se un dirigente scolastico viene a conoscenza di atti di bullismo o cyberbullismo, è tenuto a informare tempestivamente i genitori e promuovere adeguate iniziative di carattere educativo nei confronti dei minori coinvolti. Nei casi più gravi, se si tratta di condotte reiterate o se le iniziative di carattere educativo adottate dalla scuola non hanno avuto esito positivo, il dirigente scolastico deve rivolgersi alle autorità competenti.

Un'altra importante novità è rappresentata dalle misure rieducative previste per i minori responsabili di condotte aggressive o lesive della dignità altrui. Il Tribunale per i minorenni potrà disporre lo svolgimento di progetti di intervento educativo con finalità rieducativa e riparativa sotto la direzione e il controllo dei servizi sociali. Questi

percorsi potranno prevedere lo svolgimento di attività di volontariato sociale, la partecipazione a laboratori teatrali o di scrittura creativa, a corsi di musica e lo svolgimento di attività sportive, artistiche o altre attività idonee a sviluppare nel minore, sentimenti di rispetto nei confronti degli altri e ad alimentare dinamiche relazionali sane e positive tra pari e forme di comunicazione non violente.

Infine, la legge n. 70/2024 istituisce la *Giornata del rispetto* il 20 gennaio di ogni anno, in memoria di Willy Monteiro Duarte. In questa giornata le scuole si dedicheranno ad affrontare le tematiche del rispetto degli altri, della sensibilizzazione contro la violenza psicologica e fisica e del contrasto di ogni forma di discriminazione e prevaricazione. Il nuovo disposto normativo rappresenta un importante passo avanti nella lotta contro il bullismo e il cyberbullismo nelle scuole italiane, fornendo strumenti più efficaci per proteggere gli studenti e promuovere un clima scolastico sicuro e rispettoso⁷⁹.

Ma facciamo un passo indietro e ripartiamo da dove tutto ha avuto inizio. Il termine cyberbullismo fu coniato dall'educatore canadese Bill Belsey nel 2002, e ripreso nel 2006 da Peter K. Smith e

⁷⁹<https://ultimora.zanichelli.it/diritto/materie-diritto/diritto-civile/bullismo-e-cyberbullismo-una-nuova-legge/>
(05/01/2025)

collaboratori, i quali proposero una definizione di cyberbullismo molto legata al concetto di bullismo tradizionale. Il cyberbullismo (Slonje, Smith 2008) è

un atto aggressivo e intenzionale, condotto da un individuo o gruppo di individui, usando varie forme di contatto elettronico, ripetuto nel corso del tempo contro una vittima che ha difficoltà a difendersi.

A tal proposito, diversi sono gli elementi che accomunano il bullismo tradizionale a quello posto in essere mediante la Rete. In particolare, i punti di comunanza sono: l'abuso di potere, l'anonimato di cui può avvalersi il cyber bullo, la ripetitività della condotta bullizzante, infine l'intenzionalità e l'aggressività dell'autore di tale condotta.

Al contrario, svariate sono le differenze tra le due fattispecie, complici, come abbiamo già ricordato, le caratteristiche proprie del web. Mentre infatti nel bullismo propriamente detto il bullo è assolutamente conosciuto dalla propria vittima (si pensi agli studenti e ai loro compagni di classe), nel cyberbullismo l'autore di tale condotta può rimanere totalmente anonimo. Da qui, si distingue un diverso tipo di disinibizione comunque presente nel fenomeno generale. Nel primo caso, la disinibizione è dettata dalla forza che viene conferita al

bullo dalle dinamiche del gruppo, mentre online tale aspetto viene in rilievo dal momento che il bullo stesso trova il coraggio di fare ciò che nella realtà concreta non farebbe mai. Ecco perché: il bullo tradizionale ha necessità di dominare le relazioni interpersonali con la propria visibilità, mentre il cyber bullo approfitta della sua presunta invisibilità.

Altro elemento distintivo è costituito dalla mancanza di un feedback tra aggressore e vittima sulla rete. Mentre infatti nella realtà si osserva la chiara presenza di riscontro tra i due attori protagonisti, nel fenomeno online il cyberbullo non vede le conseguenze delle proprie azioni. Da qui l'assenza di qualsivoglia tipo di spiraglio di empatia del carnefice verso la sua vittima.

Inoltre, in entrambi i fenomeni si può osservare una vera e propria deresponsabilizzazione dell'operato del bullo, concretizzata o nella tipica frase "Era uno scherzo!" o, come nel caso del web, attraverso la prescrizione delle conseguenze delle azioni violente del cyberbullo agli avatar.

Infine, sono degni di nota altri due punti di differenziazione tra le due fattispecie. In prima battuta, mentre nel bullismo solo i bulli possono ricoprire tale ruolo, nel mondo virtuale le stesse vittime del cyberbullo possono diventare anch'esse

carnefici. In secondo luogo, se nel bullismo tradizionale il pubblico è sempre passivo, nel cyberbullismo il pubblico spettatore può decidere di divenire attivo, partecipando in prima persona alle prepotenze virtuali (Pisano, Saturno 2008, 40-45).

Nel corso degli anni, gli studiosi sono poi arrivati a dettagliare con maggiore precisione le tipologie di cyberbullismo che possono concretizzarsi sulla rete. Nancy Willard⁸⁰, propone la classificazione che segue delle azioni tipiche del bullo online: *flaming*: messaggi online violenti e volgari, che mirano a suscitare battaglie verbali in un forum; *harassment*: l'invio ripetuto di messaggi offensivi e sgradevoli mirati a ferire qualcuno; *denigration*: insultare o diffamare qualcuno online attraverso dicerie, pettegolezzi e menzogne, solitamente di tipo offensivo e crudele, volte a danneggiare la reputazione di una persona e i suoi rapporti; *impersonation*: in questo caso l'aggressore ottiene le informazioni personali e i dati di accesso (nick name, password, ecc.) di un account della vittima, con lo scopo di prenderne possesso e danneggiarne la reputazione; *outing*: condividere online informazioni imbarazzanti, segreti o foto di qualcuno; *trick-*

⁸⁰<https://cdn.ymaws.com/www.safestates.org/resource/resmgr/imported/educatorsguide.pdf> (16/01/2025)

kery: ottenere la fiducia di qualcuno per poi rendere pubbliche in rete, le informazioni carpite con l'inganno; *exclusion*: escludere deliberatamente qualcuno da un gruppo online, per provocare in lui un senso di emarginazione (chat, forum, lista di amici etc.); *cyberstalking*: molestie e denigrazioni ripetute e minacciose mirate a incutere paura e timore e a cambiare le proprie abitudini di vita.

4.2 Dati alla mano

Il Ministero della Salute ha riportato i dati della sorveglianza *Health Behaviour in School-aged Children – HBSC Italia 2022*⁸¹. Secondo quanto emerse dall'analisi, gli atti di bullismo a scuola sono più frequenti nei più piccoli (11–13 anni) e nelle ragazze. Il fenomeno, stando agli studi, è collegato alla violenza domestica. I minori esposti a episodi di violenza familiare, infatti, sono più propensi ad imitare i comportamenti che subiscono in casa. Sono spesso portati così a ferire a loro volta i compagni o a diventare loro stessi vittime anche a scuola.

Per quanto riguarda il cyberbullismo, invece, nella fascia di età 11 anni risultano vittime il 17,2% dei maschi e il 21,1% delle femmine. I tredicenni coinvolti sono il 12,9% dei ragazzi e il 18,4%

⁸¹ <https://www.epicentro.iss.it/hbhc/indagine-2022> (3/01/2025)

delle ragazze. Gli adolescenti di 15 anni sono il 9,2% dei maschi e l'11,4% delle femmine.

Come riportato da [pandasecurity.com](https://www.pandasecurity.com/it/mediacenter/52-statistiche-cyberbullismo/)⁸², più della metà degli studenti dagli 11 ai 17 anni ha subito un episodio di bullismo in Italia, stando ai dati ISTAT. Invece, secondo il CDC (Center for Disease Control and Prevention), il 20% dei bambini che hanno più di 11 anni è stato vittima di cyberbullismo.

Il Paese dove ci sono maggiori segnalazioni nel mondo da parte dei genitori è l'India. Il 37% di mamme e papà afferma che i propri figli hanno subito un episodio di bullismo online⁸³. Situazione preoccupante anche in Myanmar, Giappone, Russia, Cina, Regno Unito, Danimarca e molti altri ancora. I fenomeni purtroppo sono molto diffusi a livello globale e nessun Paese riesce a salvarsi. Questo ci fa capire, ancora una volta, quanto sia urgente e fondamentale sostenere la lotta al bullismo e al cyberbullismo.

Negli ultimi anni, i casi di cyberbullismo sono in crescita, con un impatto particolarmente

⁸²<https://www.pandasecurity.com/it/mediacenter/52-statistiche-cyberbullismo/> (16/01/2025)

⁸³<https://www.statista.com/chart/15926/the-share-of-parents-who-say-their-child-has-experienced-cyberbullying/> (16/01/2025)

grave sulle ragazze. Secondo i dati del 2022, in Italia il 44,6% delle vittime di cyberbullismo sono ragazze, mentre tra i ragazzi il dato è del 36,5%. Queste informazioni provengono dall'indagine Edit, condotta dall'Agenzia regionale di sanità Toscana (Ars) e presentata al congresso regionale della Simeup, la Società italiana di medicina di emergenza urgenza pediatrica, tenutosi a Pisa.

Un altro importante studio, l'osservatorio Indifesa 2024⁸⁴, ha rilevato che il 63% degli adolescenti italiani ha subito episodi di bullismo e il 19% ha sperimentato cyberbullismo. La scuola si conferma il luogo più a rischio, con il 66% degli episodi di bullismo che si verificano tra le mura scolastiche, seguita dal web, che rappresenta il 39% dei casi di cyberbullismo.

Le conseguenze psicologiche del bullismo e del cyberbullismo sugli adolescenti sono gravi e diffuse. Il 75% dei giovani che subiscono questi fenomeni sperimenta una perdita di autostima, sicurezza e fiducia negli altri. Inoltre, il 47% sviluppa ansia sociale e attacchi di panico, mentre il 45% tende a isolarsi e ad allontanarsi dai coetanei. Questi effetti sottolineano l'importanza di interventi

⁸⁴<https://networkindifesa.terredeshommes.it/osservatorio-indifesa/> (3/01/2025)

mirati per affrontare le ripercussioni emotive e sociali.

4.3. Testimonianze

Dal 2020 ad oggi ho avuto il piacere e l'onore di tenere convegni ed incontri di formazione in tutto il Nord Italia per sensibilizzare sul fenomeno cyberbullismo e far capire la portata, ad oggi ancora sottostimata, della fattispecie. Questa esperienza mi ha arricchito più di quanto mi potessi aspettare. Mi ha permesso infatti non solo di divulgare il mio sapere e la mia esperienza sul tema ma soprattutto di essere inaspettatamente e meravigliosamente co-protagonista di un'esperienza di pura condivisione. Mi sono trovata più volte a raccogliere testimonianze, dirette ed indirette, di chi ha provato sulla propria pelle che cosa significa nella realtà concreta la parola cyberbullismo.

La prima testimonianza – a dire il vero, il mio intero percorso di approfondimento sul tema – mi sono stati direttamente ispirata da Benedetta, 12 anni. All'epoca dei fatti, il fenomeno cyberbullismo non era ancora conosciuto così come lo è oggi – stiamo parlando di molti anni prima dell'entrata in vigore della legge del 2017. Ebbene, in quel caso, Benedetta fu vittima di harassment. Su un noto so-

cial network, in specie sul suo profilo, la giovane ricevette una serie di frasi sgradevoli e minacciose volte esclusivamente ad insultarla e a danneggiare la sua persona (“Ti brucio con l’accendino!”). L’autore di tale condotta: un suo compagno di classe. La motivazione: lei aveva una cotta per lui.

Un’altra fondamentale fonte di ispirazione, e di Luce, come nel caso di Benedetta, è Andrea. All’epoca dei fatti, quindicenne. Vittima di bullismo tradizionale. Prese in giro continue, insulti, denigrazione, violenza fisica e verbale. Motivo, a detta dei bulli, suoi compagni di classe: “sorridi troppo, ma che hai sempre da sorridere?”.

Come non parlare poi di Martina, 15 anni. Martina non ha fornito una testimonianza vera e propria o, meglio, l’ha concessa a modo suo. Durante uno dei miei convegni negli istituti scolastici del territorio, ad un certo punto, mentre stavo parlando delle tipologie di cyberbullismo ad oggi esistenti e classificate, è scoppiata in un pianto a dirotto. Questo ha dato senso a tutto, almeno per me. Martina aveva effettivamente provato tutto ciò su cui io, in quel momento, stavo cercando di sensibilizzare.

Ed ora giungiamo alla testimonianza forse più inaspettata ma quella che mi ha reso sempre più convinta che quest’opera di sensibilizzazione e

di lotta contro il bullismo e il cyberbullismo sia oltremodo necessaria e giusta.

Marco, quindicenne. Per sua stessa ammissione, dinanzi ad una platea di studenti, mentre tenevo uno dei miei convegni, si alza in piedi e dice: “Io sono stato un bullo! Me ne sono reso conto, lo ammetto e chiedo SCUSA. Ho capito i danni che ho causato”.

Dentro di me, ho pensato: ho vinto!

4.4 Conclusioni

Se usualmente le conclusioni rappresentano una fine, per me questa fine rappresenta – mi auguro – un luminoso nuovo inizio. Spero infatti che quest’opera prima possa rappresentare un piccolo, anche microscopico, spiraglio di luce per tutte quelle persone, in particolare, gli, e le, adolescenti di oggi che si trovano – spesso a loro insaputa – in un vortice che reputano familiare (Internet) ma che in realtà si sta sempre più svelando come fonte di pericoli, ancora oggi troppo spesso ignorati o comunque sottovalutati.

Ed è proprio a loro che questo libro si rivolge ed è dedicato. Proprio a voi giovani, adulti del domani, provo a dare un consiglio: solo la cultura è in grado di salvare l’essere umano. Perciò studiate, assorbite quanto più potete, siate curiose e curiosi

di tutto ciò che vi circonda, specie di ciò che vi fa paura. Andate a fondo di questa paura. Sciogliete i nodi che la compongono e rendetela una vostra ricchezza, uno strumento prezioso da usare nel mondo.

Per farlo, utilizzate la conoscenza, il sapere. Perché solo così potrete vivere questo mondo assaporandone ogni sfumatura, senza subirlo o lottando solo per la vostra sopravvivenza. Ed infine, come dice Chiara Gamberale, fedele e personale modello professionale e umano, abituatevi a consentire.

Bibliografia

Abbamonte O. (2017), *Il potere dei conflitti: Testimonianze sulla storia della Magistratura italiana*, Giappichelli Editore, Torino.

Abbondante F. (2017), *Il ruolo dei social network nella lotta all'hate speech: un'analisi comparata fra l'esperienza europea e quella statunitense*, "Intervento presentato al convegno Social media e diritto. Diritti e social media tenutosi a Museo del Novecento nel 20 dicembre 2017".

Accinni G. P. (2017), *Profili di responsabilità penale dell'hosting provider "attivo"*, "Archivio penale", 2, 1-11.

Allegri M.R. (2012), *Informazione e comunicazione nell'ordinamento giuridico italiano*, Giappichelli, Torino.

Allport G.W. (1954), *La natura del pregiudizio*, La Nuova Italia, Firenze.

Ambrosi A. (2007), *Libertà di pensiero e manifestazione di opinioni razziste e xenophobe*, "Quaderni Costituzionali", 38, 519-544.

Asard E., Bennett W.L. (1997), *Democracy and the Marketplace of Ideas*, Cambridge University Press, Cambridge.

Asaro P. (2012), *A body to kick, but no soul to damn. Legal Perspectives on Robotics*, in Lin P., Abney

K., Bekey G.A. (eds.), *Robot Ethics. The Social and Ethical Implications of Robotics*, MIT Press, Cambridge.

Baker C.E. (1989), *Human Liberty and Freedom of Speech*, Oxford University Press, New York.

Bartoli R. (2013), *Brevi considerazioni sulla responsabilità penale dell'Internet Service Provider*, "Diritto penale e processo", 5, 600-606.

Becchi P. (2013), *Il principio dignità umana*, Morcelliana, Brescia.

Benkler Y. (2006), *The Wealth of Networks*, Yale University Press, New Haven-London.

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

Brown A. (2016), *Hate speech law, A Philosophical Examination*, Routledge, New York.

Burnap P., Williams M. (2015), *Hate speech, machine classification and statistical modelling of information flows on twitter: Interpretation and communication for policy decision making*, "Internet, Policy and Politics Conference".

Caruso C. (2013), *Teoria e «ideologia» della libertà di espressione. La giurisprudenza della Corte Suprema degli Stati Uniti sulla freedom of speech*, "Forum di quaderni costituzionali rassegna", 10, 1-37.

Caruso C. (2017), *L'hate speech a Strasburgo: il pluralismo militante del sistema convenzionale*, "Quaderni costituzionali", 3, 936-984.

Castellaneta M. (2011), *La repressione del negazionismo e la giurisprudenza della Corte europea dei diritti umani*, “Diritti umani e diritto internazionale”, 1, 5-208

Chafee Z. (1941), *Free Speech in the United States*, Harvard University Press, Cambridge.

Chen Y., Zhou Y., Zhu S., Xu H. (2012), *Detecting offensive language in social media to protect adolescent online safety*, “Proceedings - 2012 ASE/IEEE International Conference on Privacy, Security, Risk and Trust and 2012 ASE/IEEE International Conference on Social Computing, SocialCom/PASSAT 2012”, 71-80.

Craddock J.M. (1995), *Words That Injure, Laws That Silence: Campus Hate Speech Codes and the Threat to American Education*, “Florida State University Law”, 22, 1047-1089.

Da Rotterdam E. (1997), *The Education of a Christian Prince with the Panegyric for Archduke Philip of Austria*, Cambridge University Press, Cambridge.

Dadvar M., de Jong F., Ordelman R., Trieschnigg R.B. (2012), *Improved cyberbullying detection using gender information*, “Proceedings of the Twelfth Dutch-Belgian Information Retrieval Workshop (DIR 2012)”, Ghent, 22–25.

Delgado R. (1982), *Words that wound: A tort action for racial insults, epithets, and name-calling*, "Harv. CR-CLL Rev.", 17, 133.

Delgado R. (2013), *The Harm in the Hate Speech*. By Jeremy Waldron, "Law and Society Review", 1, 232-233.

Delgado R. (1991), *Campus Antiracism Rules: Constitutional Narratives in Collision*, "Northwestern University Law Review", 2, 343-387.

Delgado R., Stefancic J. (2017), *Critical Race Theory, An Introduction (third edition)*, New York University Press, New York.

Dinakar K., Jones B., Havasi C., Lieberman H., Picard R. (2012), *Common sense reasoning for detection, prevention, and mitigation of cyberbullying*, "ACM Trans. Interact. Intell. Syst.", 3, 1-30.

Djuric N., Zhou J., Morris R., Grbovic M., Radosavljevic V., Bhamidipati N. (2015), *Hate speech detection with comment embeddings*, "Proceedings of the 24th International Conference on World Wide Web", 29-30.

Di Tano F. (2017), *Prospettive de iure condendo sulla responsabilizzazione dei content provider*, "Informatica e diritto", 1/2, 113-126.

Dworkin R. (2012), *Reply to Jeremy Waldron* in M. Herz and P. Molnar (eds.), *The Content and Context of Hate Speech: Rethinking Regulation and Responses*, Cambridge University Press, Cambridge.

Dworkin R. (2009), *Foreward*, in Hare I., Weinstein J. (eds.), *Extreme Speech and Democracy*, Oxford University Press, Oxford.

Ferrara G., Luciani M. (a cura di) (1994), *Dell'eguaglianza. La democrazia alla fine del secolo*, Laterza, Roma-Bari.

Gagliardone I., Gal D., Alves T., Martinez G. (a cura di) (2015), *Countering Online Hate Speech*, Unesco, Paris.

Galeotti E. (2010), *La politica del rispetto. I fondamenti etici della democrazia*, Laterza, Bari-Roma

Gardini G. (2017), *Le regole dell'informazione. L'età della post-verità*, Giappichelli, Torino.

Gelber K., McNamara L. (2015), *The Effects of Civil Hate Speech Laws: Lessons from Australia*, "Law & Society Review", 3, 631-664.

Gelber K. (2002), *Speaking Back: The Free Speech Versus Hate Speech Debate*, John Benjamins Publishing, Amsterdam.

Gunther G. (1990), *Should Universities Restrict Expression That Is Racist or Otherwise Denigrating? No*, "Stanford Lawyer", 7, 9-41.

Hartley R. C. (2010), *Freedom not to listen: A constitutional analysis of compulsory indoctrination through workplace captive audience meetings*, "Berkeley Journal of Employment & Labor Law", 31, 65-125.

Hosseinmardi H., Arredondo Mattson S., Rafiq R.I., Han R., Lv Q., Mishra S. (2015), *Detection of cyberbullying incidents on the instagram social network*, "arXiv", 12/03/2015.

Ingrassia A. (2012), *Il ruolo dell'ISP nel ciber-spazio: cittadino, controllore o tutore dell'ordine? Le responsabilità penali dei provider nell'ordinamento italiano*, in Luparia L. (ed.), *Internet provider e giustizia penale*, Giuffrè, Milano.

Kelsen H. (1984), *La democrazia*, Il Mulino, Bologna.

Leoncini I. (1999), *Obbligo di attivarsi, obbligo di garanzia e obbligo di sorveglianza*, Giappichelli, Torino.

Lepora C. (2012), *On Compromise and Being Compromised*, "Journal of Political Philosophy", 1, 1-22.

Magro M.B. (2014), *Biorobotica, robotica e diritto penale* in Provolo D., Riondato S., Yenisey F. (eds.), *Genetics, Robotics, Law, Punishment*, Padova University Press, Padova.

Manna A., Di Florio M. (2019), *Riservatezza e diritto alla privacy: in particolare, la responsabilità per*

omissione dell'internet provider, in Cadoppi A. et al. (eds.) *Cybercrime*, Utet Giuridica, Milano.

McNamara L. (2007), *Human Rights Controversies: The Impact of Legal Form*, Routledge-Cavendish, Abingdon.

Meiklejohn A. (1960), *Political Freedom: The Constitutional Power of the People*, Harper and Brothers, New York.

Meiklejohn A. (1961), *The First Amendment is An Absolute*, "Supreme Court Review", 1961, 245-266.

Mill J.S. (1981), *Saggio sulla libertà*, Il Saggiatore, Milano.

Mill J.S. (1972), *On Liberty*, in Williams G. (ed.), *Utilitarianism, On Liberty, Considerations on Representative Government*, Everyman, London.

Milton J. (1644), *Areopagitica: a speech for the liberty of unlicensed printing*, Richard Claverhouse Jebb.

Moro P., Sarra C. (2017), *Tecnodiritto. Temi e problemi di informatica e robotica giuridica*, Franco Angeli, Milano.

Moscatti L. (2008), *Milton e le origini della libertà di stampa in Inghilterra*, in *Diritto e libertà. Studi in memoria di Matteo Dell'Olio*, Giappichelli, Torino.

Nielsen L.B. (2002), *Subtle, Pervasive, Harmful: Racist and Sexist Remarks in Public as Hate Speech*, "Journal of Social Issues", 2, 265-280.

Nobata C., Tetreault J., Thomas A., Mehdad Y., Chang Y. (2016), *Abusive language detection in online user content*, "Proceedings of the 25th International Conference on World Wide Web", Geneva, 145-153.

Picotti L. (2007), *La legge contro lo sfruttamento sessuale dei minori e la pedopornografia in Internet (L. 6 febbraio 2006, n. 38) (Parte seconda)*, "Studium Iuris", 11, 1196-1211.

Pino G. (2008), *Discorso razzista e libertà di manifestazione del pensiero*, "Politica del diritto", 2, 293-294.

Piraino F. (2017), *Spunti per una rilettura della disciplina giuridica degli Internet service provider*, "Annali italiani del diritto d'autore", 1, 152-200.

Pollicino O., Sciarabba V. (2011), *La Corte europea dei diritti dell'uomo e la Corte di giustizia nella prospettiva della giustizia costituzionale* in Mezzetti L. (ed.), *Sistemi e modelli di giustizia costituzionale*, Bologna, Cedam.

Post R. (2018), *Concluding thoughts. The legality and politics of hatred*, in Brudholm T., Scheepelern

Johansen B. (eds.), *Hate, Politics, Law: Critical Perspectives on Combating Hate*, Oxford University Press, Oxford.

Post R. (2009), *Hate Speech*, in Hare I., Weinstein J. (eds.), *Extreme Speech and Democracy*, Oxford University Press, Oxford.

Post R. (2011), *Participatory Democracy and Free Speech*, "Virginia Law Review", 3, 477-489.

Pugiotto A. (2013), *Le parole sono pietre? I discorsi d'odio e la libertà di espressione nel diritto costituzionale*, "Dir. pen. cont.", 3, 1-33.

Roig F. (2017), *Libertà di espressione, discorsi d'odio, soggetti vulnerabili: paradigmi e nuove frontiere*, "Ars interpretandi", 1, 29-48.

Rolla G. (2008), *Profili costituzionali della dignità umana*, in Ceccherini E. (ed.), *La tutela della dignità dell'uomo*, Editoriale Scientifica, Napoli.

Romano S. (2014), *Il declino dell'impero americano*, Longanesi, Milano.

Sartor G. (2003), *L'intenzionalità dei sistemi informatici e il diritto*, "Rivista trimestrale di procedura e diritto civile", 1, 23-52.

Schachter O. (1983), *Human Dignity as a Normative Concept*, "American Journal of International Law", 4, 848-854.

Schauer F. (1987), *Free Speech: A Philosophical Inquiry*, Cambridge University Press, Cambridge.

Schauer F. (2012), *Social Epistemology, Holocaust Denial, and the Post-Millian Calculus*, in Herz M., Molnar P. (eds.), *The Content and Context of Hate Speech: Rethinking Regulation and Responses*, Cambridge University Press, Cambridge.

Schmidt A., Wiegand M. (2017), *A Survey on Hate Speech Detection using Natural Language Processing*, "Proceedings of the Fifth International Workshop on Natural Language Processing for Social Media", 1-10.

Seminara S. (2014), *Internet (diritto penale)*, in *Enciclopedia del Diritto*, Annali VII, Giuffrè, Milano.

Sgubbi F. (1975), *Responsabilità penale per omesso impedimento dell'evento*, Cedam, Padova.

Slonje R., Smith P.K., (2008), *Cyberbullying: another main type of bullying?*, "Scandinavian Journal of Psychology", 49, 147-154.

Spigno I. (2018), *Discorsi d'odio. Modelli costituzionali a confronto*, Giuffrè, Milano.

Strossen N. (2018), *Hate: Why We Should Resist It with Free Speech, Not Censorship*, Oxford University Press, Oxford.

Sumner L.W. (2004), *The Hateful and The Obscene: Studies in the Limits of Free Expression*, University of Toronto Press, Toronto.

Tanzarella P. (2010), *Il discorso d'odio razziale. Le tappe legislative e giurisprudenziali di un discutibile*

reato costituzionalmente protetto, “Diritto immigrazione e cittadinanza”, 4, 50-70.

Tega D. (ed.) (2012), *Le discriminazioni razziali ed etniche. Profili giuridici di tutela*, Armando Editore, Roma.

Thelwall M., Buckley K., Paltoglou G., Cai D., Kappas A. (2010), *Sentiment strength detection in short informal text*, “Journal of the American Society for Information Science and Technology”, 12, 2544-2558.

Torre V. (2013), *Sulla responsabilità penale del service provider e la definizione del comportamento esigibile alla luce delle norme contro la pedopornografia*, in Picotti L. (ed.) *Tutela penale della persona e nuove tecnologie*, Cedam, Padova.

Tulkens F. (2012), *When to say is to do, freedom of expression and hate speech in the case-law of the European Court of Human Rights*, in Casadevall J. (ed.) *Freedom of expression: essays in honour of Nicolas Bratza, president of the European Court of Human Rights*, European Court of Human Rights, Strasbourg.

Vercellis C. (2009), *Business Intelligence: Data Mining and Optimization for Decision Making*, John Wiley & Sons, Chichester.

Waldron J. (2012), *The Harm in Hate Speech*, Harvard University Press, Cambridge.

Walters M., Hoyle C. (2017), *Healing Harms and Engendering Tolerance: The Promise of Restorative Justice for Hate Crime*, in P. Bean (ed.), *Hate Crime*, Routledge, London.

Weinstein J. (2001), *The Boundaries of Freedom of Expression and Order in a Democratic Society*, Kent State University Press, Kent.

Whitman J.Q. (2004), *The Two Western Cultures of Privacy: Dignity versus Liberty*, "Yale Law Journal", 64, 1-94.

Xiang G., Fan B., Wang L., Hong J., Rose C. (2012), *Detecting offensive tweets via topical feature discovery over a large-scale twitter corpus*, "Proceedings of the 21st ACM international conference on Information and knowledge management", 1980-1984.

Young I. (1990), *Justice and the Politics of Difference*, Princeton University Press, Princeton.

Zhong H., Li H., Squicciarini A.C., Rajtmajer S.M., Griffin C., Miller D.J., Caragea C. (2016), *Content-driven detection of cyberbullying on the instagram social network*, "IJCAI", 3952-3958.

Ziccardi G. (2016), *L'odio online*, Raffaello Cortina Editore, Milano.

Sitografia

<https://www.dangerousspeech.org/libraries/guide>

<https://www.medialaws.eu/wp-content/uploads/2019/05/1.-Pitruzzella.pdf>

https://migrant-integration.ec.europa.eu/library-document/hate-crime-and-hate-speech-europe-comprehensive-analysis-international-law_en

https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2882244

<http://documenti.camera.it/leg17/dossier/Pdf/ES0420.pdf>

https://www.ispionline.it/sites/default/files/pubblicazioni/wp_52_2013.pdf

<https://hudoc.echr.coe.int/eng#%7B%22itemid%22:%5B%22003-1728198-1812055%22%5D%7D>

<https://hudoc.echr.coe.int/eng#%7B%22dmdocnumber%22:%5B%22665201%22%5D%2C%22itemid%22:%5B%22001-1039%22%5D%7D>

<https://www.corteidh.or.cr/tablas/r25909.pdf>

<http://hudoc.echr.coe.int/eng?i=001-23829>

<https://hudoc.echr.coe.int/tur#%7B%22itemid%22:%5B%22001-57499%22%5D%7D>

<https://hudoc.echr.coe.int/eng#%7B%22dmdocnumber%22:%5B%22880747%22%5D%2C%22itemid%22:%5B%22001-103141%22%5D%7D>

<https://hudoc.echr.coe.int/eng-press#%7B%22itemid%22:%5B%22003-2501837-2699727%22%5D%7D>

<http://hudoc.echr.coe.int/eng?i=001-98489>

https://www.cortecostituzionale.it/documenti/convegni_seminari/stu_284.pdf

<https://www.cortecostituzionale.it/actionSchedaPronuncia.do?anno=1956&numero=1>

https://www.regent.edu/acad/schlaw/student_life/studentorgs/lawreview/docs/issues/v25n1/04Kiskavol.25.1.pdf

https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2327305

<https://globalfreedomofexpression.columbia.edu/cases/national-socialist-party-america-v-village-skokie/>

<https://digitalcommons.law.wne.edu/cgi/viewcontent.cgi?article=1574&context=lawreview>

<https://www.osce.org/files/f/documents/3/e/36426.pdf>

<https://riviste.unimi.it/index.php/statoechiese/article/view/2130>

<https://iris.unipa.it/retrieve/e3ad8916-047b-da0e-e053-3705fe0a2b96/Spena%2c%20Liberta%27%20di%20espressione.pdf>

https://www.associazionedeicostituzionalisti.it/old_sites/sito_AIC_2003-2010/materiali/anticipazioni/odio_razziale/index.html

<https://supreme.justia.com/cases/federal/us/408/92/>

<https://globalfreedomofexpression.columbia.edu/cases/r-v-v-city-st-paul/>

<https://hudoc.echr.coe.int/eng#%7B%22itemid%22:%5B%22001-87370%22%5D%7D>

[https://hudoc.echr.coe.int/eng#{%22itemid%22:\[%22001-109046%22\]}](https://hudoc.echr.coe.int/eng#{%22itemid%22:[%22001-109046%22]})

https://www.rivistaaic.it/images/rivista/pdf/3_2014_Mor-rone.pdf

<https://www.medialaws.eu/wp-content/uploads/2019/05/2.-Pollicino.pdf>

<https://scc-csc.lexum.com/scc-csc/scc-csc/en/item/12876/index.do>

https://chicagounbound.uchicago.edu/cgi/viewcontent.cgi?article=3013&context=journal_articles

https://d1vbhhqv6ow083.cloudfront.net/contributi/Go-metz.M_Lodio.pdf

https://scholarship.law.umn.edu/comm/456/?utm_source=scholar-ship.law.umn.edu%2Fcomm%2F456&utm_medium=PDF&utm_campaign=PDFCoverPages

<https://www.thefire.org/research-learn/freedom-speech-and-workplace-harassment>

<https://www.oyez.org/cases/1900-1940/250us616>

<https://www.oyez.org/cases/1963/39>

¹<https://www.splcenter.org/20170814/ten-ways-fight-hate-community-response-guide>

<https://www.nytimes.com/2017/08/17/opinion/how-to-make-fun-of-nazis.html>

https://papers.ssrn.com/sol3/papers.cfm?abstract_id=957373

<http://dig.casail.mit.edu/2006//06/neutralnet.html>

https://www.camera.it/application/xmanager/projects/leg17/commissione_internet/dichiarazione_dei_diritti_internet_publicata.pdf

[https://hudoc.echr.coe.int/eng#{%22itemid%22:\[%22001-104685%22\]}](https://hudoc.echr.coe.int/eng#{%22itemid%22:[%22001-104685%22]})

<https://www.bibliotecheoggitrends.it/media/download/get/dd3cbe1e-906d-447c-95ad-c453b590d5bd/original>

<https://op.europa.eu/it/publication-detail/-/publication/8fdf5d08-93fc-11e5-983e-01aa75ed71a1>

<https://www.washingtonpost.com/wp-dyn/content/article/2006/06/07/AR2006060702108.html>

https://papers.ssrn.com/sol3/papers.cfm?abstract_id=943757

https://commission.europa.eu/strategy-and-policy/policies/justice-and-fundamental-rights/combating-discrimination/racism-and-xenophobia/eu-code-conduct-countering-illegal-hate-speech-online_en

https://ec.europa.eu/info/files/factsheet-4th-monitoring-round-code-conduct_it

<https://archiviodpc.dirittopenaleuomo.org/upload/1351711435II%20ruolo%20del%20ISP%20nel%20cyberspazio%20DPC.pdf>

https://www.sistemapenale.it/pdf_contenuti/1588789714_baccin-2020a-responsabilita-internet-service-provider-illeciti-online-caso-force-facebook.pdf

<https://www.agendadigitale.eu/cultura-digitale/social-network-e-intelligenza-artificiale-presente-e-futuro-di-un-forte-bio-mio/>

https://dpc-rivista-trimestrale.criminaljusticenetwork.eu/pdf/DPC_Riv_Trim_2_2019_panattoni.pdf

<https://arxiv.org/abs/1910.03270>

https://www.researchgate.net/profile/Carla-Schieb/publication/303497937_Governing_hate_speech_by_means_of_counterspeech_on_Facebook/links/5761575408aeeada5bc4f783/Governing-hate-speech-by-means-of-counterspeech-on-Facebook.pdf

<http://www.perspectiveapi.com>

https://www.researchgate.net/publication/262363934_Detecting_hate_speech_on_the_world_wide_web

file:///C:/Users/beatr/Downloads/Class-Based_n-gram_Models_of_Natural_Language.pdf

<https://aclanthology.org/R15-1086.pdf>

www.noswearing.com/dictionary/www.rpdb.org/www.hatebase.org

<https://www.media.mit.edu/publications/bttj/Paper23Pages211-226.pdf>

<https://ultimora.zanichelli.it/diritto/materie-diritto/diritto-civile/bullismo-e-cyberbullismo-una-nuova-legge/>

<https://cdn.ymaws.com/www.safestates.org/resource/resmgr/imported/educatorsguide.pdf>

<https://www.epicentro.iss.it/hbsc/indagine-2022>

<https://www.statista.com/chart/15926/the-share-of-parents-who-say-their-child-has-experienced-cyberbullying/>

<https://networkindifesa.terredeshommes.it/osservatorio-indifesa/>

**Centro
Editoriale
Accademico
unicollege**



Emanuele Vergari è autore dell'immagine di copertina, intitolata *Il peso delle parole d'odio* e creata con il supporto dell'intelligenza artificiale.

Lorenzo Grifone Baglioni è autore del progetto grafico editoriale della collana.

Il volume propone un inquadramento normativo del discorso d'odio come fattispecie di reato; confronta le politiche europee e statunitensi focalizzandosi poi sull'esperienza italiana; evidenzia le conseguenze del dilagare del fenomeno nel mondo online, in specie in età adolescenziale; propone infine un ricco ventaglio di testimonianze dei protagonisti del fenomeno, vittime o carnefici che siano.

È destinato ai docenti che vogliono approfondire il tema sul piano legale, agli studenti che vogliono capire quali 'fantasmi' si aggirino nella 'rete' e ai cittadini che vogliono opporsi efficacemente al discorso d'odio.

Beatrice Aprile è Avvocata civilista e Professoressa di Analisi del Cybercrime e della Cybersecurity in Unicollege SSML; conduce da anni un'attenta opera di sensibilizzazione sociale e di divulgazione scientifica dei temi legati al discorso d'odio.