IL "LATO OSCURO DELLA RETE": ODIO E PORNOGRAFIA NON CONSENSUALE. RUOLO E RESPONSABILITÀ DEI GESTORI DELLE PIATTAFORME *SOCIAL* OLTRE LA *NET NEUTRALITY*

di Marta Lamanuzzi
(*Assegnista di ricerca presso Università Cattolica del Sacro Cuore di Milano*)

1. La rete è nata e si è inizialmente evoluta come agorà della libertà di espressione, libertà che ha raggiunto il suo apice nei *social network*, piattaforme digitali che, consentendo a chiunque si crei un account – anche sotto falso nome – di produrre contenuti e interagire, hanno costituito la chiave di volta del processo di disintermediazione, ossia del superamento del filtro tipicamente posto dai media "tradizionali" ai contenuti degni di essere offerti al pubblico.

Agli esordi, Internet era inteso come «luogo "altro" rispetto al reale»[1], come banca dati, fonte di contenuti di cui fruire passivamente e catalogo di prodotti da acquistare. La commistione fra virtuale e vita reale inizia, secondo gli studiosi, quando, nel dicembre del 2009, per la prima volta a livello mondiale, piattaforme *social* e blog diventano la destinazione più popolare dei naviganti in rete, superando motori di ricerca, siti di informazione e spazi per lo shopping *online*. Il *web* diventa così luogo della società e non più esterno ad essa e gli utenti non hanno più il ruolo di mero "pubblico" ma assumono anche la veste di autori, creatori della rete, "spettautori"[2], o,

[1] S. Pasta, *Razzismi 2.0. Analisi socio-educativa dell'odio* online, Brescia 2018, 59.
[2] *Ivi*, 62.

secondo la dizione anglosassone, "*prosumer*", ossia, al contempo produttori (*producer*) e consumatori (*consumer*) di contenuti mediatici[3].

Nel corso di tale evoluzione, il «sogno di fluttuante e incondizionata libertà»[4] del *cyberspace*, l'ideale di una rete del tutto avulsa da regole, sede naturale di manifestazione del pensiero ed espressione della più ampia iniziativa economica privata dei *provider*, è entrato via via in conflitto con la realtà degli altri interessi in gioco. Con l'evoluzione tecnologica e social-mediatica, insieme a straordinarie opportunità di conoscenza, interazione e arricchimento, sono infatti emerse nuove modalità di offesa a beni giuridici fondamentali. Fra i fenomeni che hanno dato origine e sostanza a locuzioni quali "*Far Web*"[5] e "lato oscuro della rete"[6], suscitando particolare allarme sociale, vanno senza dubbio annoverati le manifestazioni d'odio e la pornografia non consensuale.

Secondo recenti rilevamenti[7], la pandemia avrebbe portato a una «nuova ondata di discorsi d'odio e di discriminazione» strettamente legata alla propalazione di teorie complottiste che attribuiscono la responsabilità della diffusione del virus (o delle presunte falsità sulla sua esistenza e pericolosità) a determinate categorie di individui, ora i cinesi ora gli ebrei, riscuotendo notevole successo quali risposte semplicistiche e rassicuranti a fronte della complessità e della drammaticità della situazione. Parimenti, ricerche condotte sinora principalmente negli Stati Uniti attestano un notevole incremento, negli ultimi anni, delle offese arrecate attraverso strumenti informatici alla riservatezza sessuale, c.d. *non-consensual pornography*[8].

---

[3] *Prosumer*, crasi fra *producer* e *consumer*, è un termine introdotto da A. Toffler, *The third wave*, New York 1980, per indicare un consumatore che consuma ciò che egli stesso produce.

[4] R. Flor, *Introduction. Cybercrime: finding a balance between freedom and security*, in S. Manacorda (a cura di), *Cybercrime: finding a balance between freedom and security*, Milano 2012, 13. In tal senso anche L. Picotti, *Diritto penale e tecnologie informatiche: una visione d'insieme*, in A. Cadoppi et al. (a cura di), *Cybercrime*, Milano 2019, 41: «da tempo è del resto palese l'infondatezza dell'utopistica o romantica idea di una "rete" quale spazio libero dal diritto, quale porto franco di anarchia e di totale libertà».

[5] M. Grandi, Far Web. *Odio, bufale, bullismo. Il lato oscuro dei* social, Milano 2017.

[6] N. Carr, *Big switch: rewiring the world, from Edison to Google*, New York 2008, trad. it. *Il lato oscuro della Rete: libertà, sicurezza,* privacy, Milano 2008. Il lato oscuro della rete è «popolato da individui e gruppi che, pur nella diversità di accenti e idiomi utilizzati, parlano tutti, salvo qualche rara ma importante eccezione, il linguaggio della violenza, della sopraffazione e dell'annientamento di altri esseri umani». A. Roversi, *L'odio in Rete. Siti ultras, nazifascismo* online, Bologna 2006, 21.

[7] *United Nations Guidance Note on Addressing and Countering COVID-19 related Hate Speech*, liberamente fruibile su *www.digitallibrary.un.org*. Stando ai dati raccolti da una società privata, dal dicembre 2019 al marzo 2020 si è registrato un incremento del 900% dei discorsi d'odio su Twitter verso la Cina e i cinesi, del 200% delle visite su siti con contenuti d'odio, nonché del 70% del cyberbullismo e dell'aggressività verbale fra giovani. *Rising Levels of Hate Speech & Online Toxicity during This Time of Crisis*, ricerca condotta da L1ght, una start-up nel mercato dell'intelligenza artificiale nata nel 2018, *l1ght.com/Toxicity_during_coronavirus_Report-L1ght.pdf*.

[8] A. A. Eaton, H. Jacobs, Y. Ruvalcaba, *2017 nationwide online study of nonconsensual porn victimization and perpetration. A summary report,* 12 giugno 2017, liberamente fruibile su *www.cybercivilrights.org*. A livello europeo, come messo in evidenza dall'*European Institute for Gender Equality*, è da anni avvertita la necessità di

In materia di *hate speech* non esiste un approccio definitorio unitario. In un documento dell'ONU del 2019 vi viene ricondotto «qualsiasi tipo di messaggio, veicolato con discorsi, scritti o comportamenti, che colpisca o usi un linguaggio peggiorativo o discriminatorio nei confronti di una persona o di un gruppo sulla base di una loro peculiare caratteristica, in altre parole, sulla base della religione, dell'etnia, della nazionalità, della razza, del colore, della stirpe, del sesso o di altro fattore identitario»[9].

A livello europeo, la Raccomandazione del Comitato dei Ministri del Consiglio d'Europa del 30 ottobre 1997 vi ricomprende «tutte le forme di espressione che diffondono, incitano, promuovono o giustificano l'odio razziale, la xenofobia, l'antisemitismo o altre forme di odio basate sull'intolleranza» e, sempre nell'ambito del Consiglio d'Europa, il Protocollo addizionale alla Convenzione di Budapest sulla criminalità informatica, firmato a Strasburgo il 28 gennaio 2003, richiama l'espressione per indicare «qualsiasi materiale scritto, immagine o altra rappresentazione di idee o teorie, che sostiene, promuove o incita l'odio, la discriminazione o la violenza, contro qualsiasi individuo o gruppo di individui, sulla base di razza, colore, etnia, religione».

Definizioni incomplete, almeno quelle "europee", nella misura in cui non contemplano l'orientamento sessuale e l'identità di genere fra le caratteristiche "bersaglio", ma che hanno il pregio di attirare l'attenzione sul concetto di pregiudizio, di *bias*, di percezione distorta e aprioristica di una determinata categoria di individui, che sta alla base delle esternazioni offensive o discriminatorie. L'*hate speech* va infatti considerato, anzitutto, un *bias speech*, un discorso fondato sul pregiudizio, così come l'*hate crime* è propriamente un *bias crime*, ossia, come si ricava dalla definizione dell'OSCE, un reato motivato dal pregiudizio (*bias motivation*)[10]. Concetto prezioso, quello di *bias*, per comprendere la genesi sia dell'*hate speech*, anche, a maggior ragione, in questo particolare momento storico in cui la logica del capro espiatorio esercita

intensificare e affinare gli strumenti di monitoraggio del fenomeno. V. *Cyber violence is a growing threat, especially for women and girls*, 19 giugno 2017, *www.eige.europa.eu*.

[9] Traduzione a cura dell'autore. Si riporta la definizione in lingua originale: «*any kind of communication in speech, writing or behaviour, that attacks or uses pejorative or discriminatory language with reference to a person or a group on the basis of who they are, in other words, based on their religion, ethnicity, nationality, race, colour, descent, gender or other identity factor*». *UN Strategy and Plan of Action on Hate Speech*, 18 giugno 2019, *www.un.org*.

[10] «Il reato ispirato dall'odio consiste in un qualunque atto che sia: (1) autonomamente tipizzato da una norma penale ("reato base" - *base offence*) e, in aggiunta, (2) motivato dal pregiudizio basato su una specifica caratteristica della vittima ("motivazione basata sul pregiudizio" - *bias motivation*)». OSCE, *Perseguire giudizialmente i crimini d'odio. Una guida pratica*, Varsavia 2016, 21. Cfr. L. Goisis, *Crimini d'odio. Discriminazione e giustizia penale*, Napoli 2019, 30. «Un *bias crime* accade non perché la vittima è *chi* è, ma piuttosto perché la vittima è *ciò* che è». F. M. Lawrence, *Punishing Hate: Bias Crime under American Law*, Cambridge 1999, 9.

notevole attrattiva[11], sia di talune forme di pornografia non consensuale, e per meglio interpretarne la proliferazione nell'habitat social-mediatico.

Le conseguenze dell'*hate speech*, anche con specifico riferimento al «*Covid-19 related hate speech*», si possono apprezzare in una prospettiva di breve e di lungo termine. In particolare, accresce la vulnerabilità delle categorie bersagliate ai rischi di violenza ed emarginazione politica e sociale, concorre alla radicalizzazione di movimenti estremisti, mina la coesione sociale, la solidarietà e la fiducia – necessarie, fra l'altro, per contrastare efficacemente la diffusione del virus – oltre a rappresentare una minaccia per il godimento dei diritti umani, lo sviluppo sostenibile, la pace e la sicurezza internazionale[12].

Quando si parla di "odio in rete", inoltre, alcuni studiosi distinguono fra condotte volte a offendere una minoranza, incitando alla violenza o alla discriminazione (secondo le citate definizioni di *hate speech*), e condotte di molestia, minaccia, denigrazione *ad personam* finalizzate a interferire negativamente nella vita quotidiana e nel benessere della vittima, riconducendovi fenomeni quali il cyberbullismo, il cyberstalking e l'istigazione al suicidio[13].

Venendo alla *non-consensual pornography*, è bene sottolineare come tale espressione sia preferibile alla locuzione "*revenge porn*", che gode di particolare risonanza mediatica, per varie ragioni. Anzitutto, parlando di *revenge porn* tecnicamente ci si riferisce alla sola cessione o diffusione di immagini o video sessualmente espliciti da parte dell'ex partner della persona raffigurata per motivi di vendetta (c.d. *revenge porn* in senso stretto), escludendo così i casi, per nulla infrequenti, in cui la diffusione di contenuti intimi altrui avviene al di fuori di un (precedente) rapporto di coppia e/o per scherzo, per noia, per leggerezza e non per ritorsione. In secondo luogo, «il termine "*revenge*" rimanda a una forma di retribuzione rispetto a un male ingiusto preventivamente subito, giustificando, in qualche misura, la condotta dell'agente» e comportando il rischio di effetti collaterali «di c.d. "*victim blaming*"», ossia di biasimo della vittima per aver in qualche modo concorso alla propria vittimizzazione[14].

---

[11] Cfr. G. Forti, *Coronavirus, la tentazione del capro espiatorio e le lezioni della storia*, *Il Sole 24 ore*, 9 marzo 2020; Id., *Introduzione. Un'attesa di luce, dalla carità*, in G. Forti (a cura di), *Le regole e la vita. Del buon uso di una crisi, tra letteratura e diritto*, Milano 2020, 12 ss.

[12] *United Nations Guidance Note on Addressing and Countering COVID-19 related Hate Speech*, cit.

[13] G. Ziccardi, *L'odio* online. *Violenza verbale e ossessioni in rete*, Milano 2016, 14; Id., *L'odio e la rete: un'introduzione e alcune possibili linee di ricerca*, in *CD* 2015, 2, 258; M. R. Allegri, *Ubi Social, Ibi Ius. Fondamenti costituzionali dei* social network *e profili giuridici della responsabilità dei* provider, Milano 2018, 176-177. Così anche K. Lumsden, E. Harmer, *Exploring Digital Violence and Discrimination on the Web*, London 2019, 1, in cui la locuzione "*digital violence*" viene utilizzata per indicare «*discrimination, harassment and hate on the Web including flaming, trolling, misogyny, racism and Islamophobia*».

[14] G. M. Caletti, *Libertà e riservatezza sessuale all'epoca di internet. L'art. 612-ter c.p. e l'incriminazione della*

La pornografia non consensuale si atteggia, nell'*id quod plerumque accidit*, quale degenerazione abusiva del *sexting*, ossia come trasmissione o pubblicazione illegittima, in quanto sprovvista del necessario consenso esplicito della persona raffigurata, di contenuti *ab origine* realizzati e inviati dalla stessa, e colpisce in misura statisticamente maggiore e più incisiva le donne rispetto agli uomini[15], connotandosi così, in molti casi, come «*gendered form of sexual abuse*»[16]. Gli stereotipi di genere, infatti, da una parte, si traducono in un aggravamento delle conseguenze lesive per le vittime di sesso femminile, dall'altra, di frequente, influiscono sulla criminogenesi. Quanto al primo profilo, è ancora radicato, anche nelle moderne democrazie occidentali, un doppio standard di valutazione della sessualità: espliciti atteggiamenti sessuali sono per l'uomo motivo di vanto e conferma di virilità, in aderenza al ruolo di genere prescritto, mentre per la donna costituiscono un disvalore, sintomo di impudicizia e ragione di biasimo. Gli stessi retaggi culturali sono altresì, spesso, all'origine del fenomeno, in maniera particolarmente evidente nei casi di *revenge porn* "in senso stretto", quando la condotta è effettivamente tenuta per "vendetta", come «atto di virilità compensatorio»[17], ossia quale reazione all'emancipazione della donna che decide di porre fine al rapporto. Si è osservato come, in tali casi, «la fine della relazione per volontà femminile viene vissuta dagli uomini che commettono *revenge porn* come fonte di frustrazione e di rabbia per il rifiuto da parte dell'oggetto amato, perdita di controllo e di potere; una "evirazione" simbolica, che equivale a una perdita di identità: l'identità dell'uomo, *pater familias*, che per secoli ha garantito al maschio un ruolo sociale e familiare di potere»[18]. Inoltre, ogniqualvolta la diffusione illecita di immagini o video sessualmente espliciti offenda una donna *in quanto donna*, la

pornografia non consensuale, in *RIDPP*, 2019, 4, 2052 ss. L'Autore sottolinea come «quelle che, a prima vista, possono apparire mere disquisizioni terminologiche presentano riflessi anche sul piano politico-criminale. Oltre che a livello culturale, i limiti dell'espressione "*revenge porn*" e l'ambiguità del suo utilizzo in una doppia accezione hanno portato a incertezze di inquadramento normativo. Le giurisdizioni che hanno ritagliato la fattispecie incriminatrice sul "*revenge porn* in senso stretto", infatti, hanno fallito nell'apprestare la giusta tutela alle vittime in relazione a casi dotati della medesima carica offensiva». *Ivi*, 2054.

[15] Cfr. A. A. Eaton, H. Jacobs, Y. Ruvalcaba, *2017 nationwide online study*, cit.

[16] Y. Ruvalcaba, A. A. Eaton, *Nonconsensual pornography among U.S. adults: A sexual scripts framework on victimization, perpetration, and health correlates for women and men*, in *Psychology of Violence*, 2020, 10, 68-78. In tal senso si è affermato che il *revenge porn*, insieme alle altre forme di abuso sessuale basate sull'immagine e la pornografia, «riproduce in modo più ampio l'egemonia di genere» A. Sorgato, *Revenge porn. Aspetti giuridici, informatici e psicologici*, Milano 2020, 228.

[17] M. Hall, J. Hearn, *Revenge pornography: gender, sexuality and motivations*, New York 2018.

[18] A. Sorgato, *op. cit.*, 234. Sulla violenza di genere come reazione all'emancipazione femminile si veda anche, diffusamente, G. Forti, *La tutela della donna dalla cd. violenza di genere. L'intervento sulla relazione affettiva in una prospettiva criminologica "integrata"*, in O. Fumagalli Carulli, A. Sammassimo (a cura di), *Famiglia e matrimonio di fronte al Sinodo. Il punto di vista dei giuristi*, Milano 2015, 42 ss.

condotta assume la connotazione di *hate crime*[19], nella sua autentica accezione di *bias crime*, ossia di crimine fondato sul pregiudizio[20].

Fra le conseguenze lesive più frequenti della pornografia non consensuale, considerato il suo devastante impatto sulla *online reputation* della vittima, si registrano disagi e disturbi emotivi, ripercussioni sulla vita sociale e lavorativa, multivittimizzazione (spesso le vittime divengono oggetto di aggressioni o molestie)[21] e, nei casi più gravi, disturbi dell'alimentazione e autolesionismo fino al suicidio[22].

2. Sebbene l'*hate speech*, l'odio *ad personam* e la *non-consensual pornography* esistessero anche prima della diffusione delle attuali tecnologie digitali e sebbene continuino a esistere anche al di fuori della rete e dei *social network*, è indubbio che la diffusività[23] e la "permanenza" del *web*[24], unite alle logiche di visibilità e indicizzazione dei contenuti, ne amplifichino la portata lesiva. Inoltre, diversi studi attestano come alcune caratteristiche dell'architettura social-mediatica e dell'agire *online* favoriscano meccanismi di disinibizione e deresponsabilizzazione, così come il consolidamento di pregiudizi e la radicalizzazione di gruppi estremisti[25]. La comprensione di tali dinamiche, resa possibile dall'imprescindibile contributo delle scienze empiriche, lungi dal sottendere prospettive di «eccessiva criminalizzazione della rete»[26], è

[19] G. M. Caletti, *Libertà e riservatezza sessuale*, cit., 2055-2056.

[20] Vedi nota n. 10.

[21] Il rischio di multivittimizzazione è particolarmente elevato quando, unitamente ai contenuti sessualmente espliciti, vengono diffuse anche informazioni personali della vittima (nome, indirizzo, numero di telefono, contatto e-mail, riferimenti sui *social network*), fenomeno definito "*doxing*".

[22] Cfr. A. Sorgato, *op. cit.*, 215 ss.

[23] Si parla di "*snowball effect*" per indicare che un contenuto immesso in rete, anche mediante l'invio a una sola persona, può repentinamente diventare virale come una palla di neve che si va via via rapidamente ingrandendo nel rotolare giù dal pendio. Cfr. Cambridge Dictionary, "*A snowball effect*".

[24] L'affermazione «*the Internet never forgets*» allude alla "permanenza" (*permanency*) della rete, ossia al fatto al fatto che, anche qualora si rimuova un contenuto dalla sua sede virtuale originaria, non si avrà mai la certezza di averlo eliminato definitivamente, in quanto il destinatario (o i destinatari), o chiunque sia venuto a contatto con esso, può averlo memorizzato o condiviso con terzi o in un altro ambiente del *web*. Cfr. M. Crockett, *The Internet (Never) Forgets*, in *Science and Technology Law Review*, 2016, 19, 2, 150 ss. Cfr. UNESCO, *Countering Online Hate Speech*, Paris, 2015 e B. Perry, P. Olsson, *Cyberhate: the Globalization of Hate*, in *Info. & Comm. Tech. L.*, 2008, 18, 185 ss.

[25] Gli esperti sottolineano come il *web* sia «un mezzo che, più ancora degli altri mezzi di comunicazione di massa, date le sue caratteristiche peculiari, può diventare uno straordinario amplificatore per messaggi semplificati e tesi apodittiche che sono destinate a suscitare adesione acritica più che a stimolare dibattito, riflessione e partecipazione consapevole», radicalizzando le posizioni e rendendo più difficile il confronto. M. Mensi, P. Falletta, *Il diritto del* Web, 2° ed., Padova 2018, 30. Sul tema, diffusamente M. Santerini, *La mente ostile: forme dell'odio contemporanea*, Milano 2021; Ead., (a cura di), *Nemico innocente: l'incitamento all'odio nell'Europa contemporanea*, Milano 2019.

[26] G. Ziccardi, *op. cit.*, 387.

fondamentale per elaborare una strategia "integrata" di prevenzione e contrasto dei fenomeni in esame[27].

Per fare solo alcuni cenni, molti studiosi riconnettono effetti crescenti di disinibizione, che può diventare "tossica", ossia tradursi in deresponsabilizzazione[28], ai vari "livelli" di anonimato. Il primo livello è costituito dalla *perception*, che ricorre quanto il soggetto, pur agendo con il proprio nome, sente di non poter essere individuato in relazione al contesto di appartenenza (posizione geografica, professione, genere, età, etnia) e dunque si percepisce "non identificabile", in quanto ritiene impossibile per gli interlocutori risalire alle altre informazioni necessarie per "inquadrarlo". Il secondo livello consiste nell'*approved anonymity*, in cui l'utente agisce in forma pseudonima, con un *nickname*, pur avendo utilizzato le proprie generalità per iscriversi alla piattaforma. Il fatto di agire con uno pseudonimo e spesso senza una fotografia aumenta l'effetto disinibitorio, ma l'agente è sempre agevolmente rintracciabile. Ancora, il livello successivo è quello della *disapproved anonymity*, in cui il soggetto ha utilizzato dati falsi per iscriversi alla piattaforma[29]. In questo caso, l'effetto disinibitorio e la deresponsabilizzazione aumentano ulteriormente. Se infatti è vero che è sempre possibile per la polizia postale individuare l'indirizzo IP o il *mac address* da cui l'utente ha agito, egli sa o intuisce che simili indagini vengono condotte solo in casi di una certa gravità. Infine, l'ultimo livello è quello della *full anonymity*, resa possibile solo dall'impiego di procedure sofisticate e dal ricorso a servizi che deviano il traffico e usano processi di crittografia che, se ben utilizzati, garantiscono appunto un anonimato totale e quindi il massimo livello di deresponsabilizzazione[30].

Inoltre, concorre all'effetto disinibitorio la distanza fisica[31], e a volte anche temporale, fra l'azione compiuta *online* e i suoi effetti dannosi, che spesso si

[27] Viene enfatizzata in dottrina la necessità di «prendere piena coscienza del terreno in cui tale libertà si realizza, in modo da calibrare e adeguare la garanzia dei diritti di ciascuno con le dinamiche assai peculiari della realtà virtuale». M. Mensi, P. Falletta, *op. cit.*, 176.

[28] La convinzione di non poter essere "individuati" porta a comportarsi in maniera più disinibita, meno condizionata da convenzioni e norme sociali, il che può essere positivo, in certe situazioni, poiché consente di discutere liberamente di argomenti anche molto intimi, ma negativo in altre, potendosi tradurre in comportamenti aggressivi che il soggetto non terrebbe nelle relazioni *face to face* (c.d. "disinibizione tossica"). P. Wallace, *The Psychology of the Internet*, Cambridge 2016, trad. it. *La psicologia di Internet*, Milano 2016, 139 ss.

[29] Se l'account rimanda chiaramente a un'altra persona (poiché ad es. ne viene utilizzato il nome, il soprannome o una fotografia), la condotta, nel nostro ordinamento, potrebbe integrare il delitto di sostituzione di persona ex art. 494 c.p. Cfr. M. Marraffino, *La sostituzione di persona mediante furto di identità digitale*, in A. Cadoppi et al. (a cura di), *op. cit.*, 307 ss.; S. Corbetta, *Delitti contro la fede pubblica. Falsa creazione di un profilo facebook: è sostituzione di persona*, in *DPP*, 2020, 9, 810 ss.

[30] Uno dei software più utilizzati in tal senso è TOR (*The Onion Router*). Cfr. S. Pasta, *op. cit.*, 78 ss.; G. Ziccardi, *op. cit.*, 97.

[31] Il rapporto fra propensione a infliggere sofferenza e distanza fisica è stato studiato già nel celebre esperimento di psicologia sociale di Stanley Milgram. Cfr. S. Milgram, *Obedience to authority, An Experimental View*, New York 1974, trad. it. *Obbedienza all'autorità*, Torino 2003, 32 ss. Cfr. Z. Bauman, *Modernity and the Holocaust*, New

apprezzano dopo un certo lasso di tempo, a seguito della condivisione "a cascata" del contenuto offensivo. Si aggiunga che, nei casi di offese rivolte contro vittime specifiche, il mancato contatto visivo che caratterizza le interazioni virtuali e la conseguente mancata percezione dei segnali esteriori di disagio riducono le possibilità di una moderazione o cessazione "empatica" dell'offesa[32]. Tali aspetti favoriscono i meccanismi di *moral disengagement*, di disimpegno morale, di neutralizzazione della dissonanza cognitiva fra le proprie azioni e i valori appresi e assimilati dal soggetto, attivando meccanismi quali l'etichettamento eufemistico ("è solo un post, ho messo solo un *like*"), il confronto vantaggioso ("non ho ucciso nessuno") e la deumanizzazione della vittima (che è diffusa, stereotipata o comunque lontana, "invisibile")[33].

Oltre a ciò, la rete, e in particolare molti *social network*, vengono vissuti dagli utenti come "non luogo", come luogo ricreativo, dove le azioni hanno meno importanza rispetto a quelle compiute *offline*, in cui il confine fra lecito e illecito, fra meritevole e riprovevole tende a sbiadire e i criteri di valutazione di un comportamento non sono tanto la verità, l'opportunità, la conformità alla legge e ai valori che fondano il vivere civile, quanto il successo in termini di apprezzamenti e condivisioni. Nell'habitat social-mediatico qualunque idea, qualunque contenuto, purché sensazionalistico, può ottenere popolarità fino a diventare virale e sembra che tale risonanza sia l'unico parametro di giudizio della sua qualità[34].

York 1989, trad. it. *Modernità e Olocausto*, Bologna 1992, 215 ss. e G. Forti, *Prefazione al volume "Io perpetratore, io vittima"*, Torino 2020.

[32] Secondo alcuni studi, l'interazione mediata comporterebbe una ridotta stimolazione dei neuroni specchio e quindi una ridotta capacità di riconoscere le emozioni proprie e altrui, rischiando di tradursi in una sorta di "analfabetismo emotivo". «A caratterizzare molte delle emozioni che si sperimentano nei nuovi media è la loro alterità: pur provandole in prima persona, sono lo specchio di emozioni di altri». I media, infatti, da una parte, favoriscono i processi comunicativi, mettendo in contatto persone che non si trovano nello stesso luogo, dall'altra, sostituiscono l'esperienza diretta dell'incontro tra corpi con una percezione indiretta influenzata dalle caratteristiche del *medium* e del messaggio. Ne conseguono tre effetti: l'assenza di consapevolezza delle proprie emozioni e dei comportamenti a esse associati, la mancata comprensione delle ragioni per cui si prova una certa emozione, l'incapacità di cogliere le emozioni altrui e quindi di reagire adeguatamente alle stesse e ai comportamenti che ne scaturiscono. S. Pasta, *op. cit*., 90 ss.

[33] In altri termini, grazie alle tecniche di disimpegno morale il soggetto può violare i propri standard morali pur mantenendo intatta la propria «*moral integrity*». A Bandura, *Moral Disengagement: How People do Harm and Live with Themselves*, New York 2015, 2. Cfr. Id., *Failures in Self-Regulation: Energy Depletion or Selective Disengagement?*, in *Psychological Inquiry*, 1996, 7, 20-24. Già nei primi anni del Duemila, Bandura aveva usato l'espressione «*death technologies*» per sottolineare come le moderne tecnologie digitali siano pericolose e spersonalizzate, in quanto creano una distanza fisica, e volte anche temporale, fra l'azione e i suoi effetti dannosi, che facilita l'agire deviante. Id., *Selective Moral Disengagement in the Exercise of Moral Agency*, in *Journal of Moral Education*, 31, 2, giugno 2002, 108. Sul punto, si veda anche M. Santerini, *La mente ostile*, cit., 49 ss.

[34] Con il termine post-verità si indica proprio una narrativa, ricorrente nei nuovi media, caratterizzata da un forte appello all'emotività e dalla valorizzazione di credenze diffuse anziché di fatti verificati. Cfr. Enciclopedia Treccani, *Voce "post-verità"*, *www.treccani.it*.

Ancora, la facilità e la velocità dell'azione in rete, unite alla (apparente)[35] gratuità e alla connessione permanente resa possibile da *smartphone* e *tablet*[36], operano a detrimento della ponderazione dei contenuti creati, condivisi e, a maggior ragione, apprezzati, in quanto, secondo diversi studi, attiverebbero quello che Kahneman definisce il "sistema 1", ossia il *Self* intuitivo, che decide rapidamente sulla scorta di impressioni e sensazioni, in luogo del "sistema 2", che è il *Self* razionale, lento e riflessivo[37].

Da ultimo, ma non per importanza, giocano un ruolo tutt'altro che trascurabile nell'amplificazione delle offese, così come nell'alimentare i pregiudizi e nel banalizzare l'odio e la pornografia non consensuale, i meccanismi automatizzati di indicizzazione. I contenuti di incitamento alla violenza e alla discriminazione, intimidatori e sessualmente espliciti, in quanto intrinsecamente sensazionalistici, rischiano di attivare la funzione selettiva dei media, aumentando così il numero degli utenti che li visualizzano e quindi la probabilità che vengano apprezzati e condivisi. Il che, da una parte, concorrendo al "successo mediatico" di quel contenuto, fa leva su un altro meccanismo di *moral disengagement*, la c.d. diffusione della responsabilità: più *like* e condivisioni ha un post meno l'utente si sente "responsabile" ad apprezzarlo o condividerlo a sua volta[38]; dall'altra, favorisce la banalizzazione e la normalizzazione di quel tipo di contenuto[39]: più di frequente compaiono contenuti d'odio e

[35] La gratuità è apparente perché, oltre al costo del dispositivo e della connessione, l'utente "paga" la navigazione e l'iscrizione a *blog* e *social network* in dati personali, che vengono sfruttati economicamente dai *provider*.

[36] Con la disintermediazione, come anticipato, tutti possono creare contenuti e reagire ai contenuti creati da altri utenti e, con l'avvento di *smartphone* e *tablet*, possono farlo sempre e ovunque. Cfr. G. Pitruzzella, *La libertà di informazione nell'era di internet*, in *ML*, 2018, 1, 22.

[37] Daniel Kahneman, premio Nobel per l'economia, ha messo in evidenza come il soggetto si identifichi con quello che in psicologia viene comunemente definito «sistema 2», vale a dire con la parte di Sé razionale, che riflette, ha convinzioni, prende decisioni e opera scelte, ma come, in realtà, spesso, le nostre azioni sono spesso guidate dal c.d. «sistema 1», vale a dire da quel sistema di automatismi, di «impressioni e sensazioni che originano spontaneamente e sono le fonti principali delle convinzioni esplicite e delle scelte deliberate dal sistema 2». D. Kahneman, *Thinking. Fast and Slow*, New York 2011, trad. it. *Pensieri lenti e veloci*, Milano 2012, 21 ss. Cfr. G. Gigerenzer, *Gut Feelings: The Intelligence of the Unconscious*, New York 2007, trad. it. *Decisioni intuitive. Quando si sceglie senza pensarci troppo*, Milano 2009, 16 ss. «La velocità è nemica dell'agire razionale», «nodo problematico per la *Generazione touch*». S. Pasta, *op. cit.*, 67. Si è osservato come «il tempo in Internet è collassato, è un tempo ad alta densità che richiederebbe di poter praticare l'ordine della successione dentro quello della coesistenza: come l'essenza (*Wesen*) della logica hegeliana, in Internet l'istante è carico, stratificato (perché molti messaggi vengono contemporaneamente ricevuti, perché la pagina *web* contiene in sé dei link che solo in un tempo più lungo possono essere percorsi), contiene potenzialmente il tutto». P. C. Rivoltella, *Costruttivismo e pragmatica della comunicazione* on line. *Socialità e didattica in Internet*, Gardolo (TN) 2003, 101.

[38] Sugli effetti di deindividuazione e di deresponsabilizzazione legati alla percezione di agire in massa, si veda anche G. Le Bon, *Psychologie des foules*, Paris, 1985, trad. it. *Psicologia delle folle*, Milano 2019.

[39] Si è parlato, in tal senso, di una «sovraesposizione» dell'odio dovuta all'attuale contesto social-mediatico e, in particolare, alla «pubblicità per interazione» di cui i contenuti sensazionalistici spesso arrivano a godere. A. Spena, *La parola odio. Sovraesposizione, criminalizzazione e interpretazione dello* hate speech, in *Criminalia*, 2017,

pornografici, maggiori saranno gli effetti di assuefazione dell'utente alla violenza verbale e alle offese alla riservatezza sessuale[40].

Inoltre, l'indicizzazione dei contenuti fondata sulla profilazione degli utenti, sfruttata dai gestori a scopo di lucro in quanto consente di personalizzare gli inserti pubblicitari, unitamente alla viralizzazione dei post sensazionalistici, concorre alla radicalizzazione di *bias* e teorie complottiste facendo comparire all'utente contenuti che non fanno altro che confermare le sue credenze, le sue distorsioni cognitive. L'utente si trova così in una sorta di bolla autoreferenziale, definita dagli esperti "*filter bubble*", o, secondo un'altra efficace metafora, in una "*echo chamber*", vale a dire in uno spazio virtuale in cui riecheggiano le opinioni che ha già manifestato o le teorie che ha cercato o condiviso. In questo modo, il rifiuto per il *fact-checking* di *hater* e complottisti è agevolato dalla scarsità di contenuti di segno diverso in cui si imbattono, uscendone invece rafforzate e confermate le loro convinzioni deliranti[41]. A ciò si aggiunge, ancora, il c.d. "effetto gregge", o euristica della socializzazione, ossia «la propensione emotiva e impulsiva degli individui a omologarsi alla pressione sociale del proprio gruppo di riferimento»[42], particolarmente evidente all'interno dei gruppi virtuali.

3. Questo breve *excursus* concorre a spiegare come la rete, e in particolare le piattaforme *social*, per le loro caratteristiche, siano diventate bacino di raccolta e terreno di proliferazione e amplificazione di offese a diritti fondamentali, permettendo così di cogliere una tensione fra almeno tre valori di rilevanza costituzionale o "super" costituzionale: la libertà di espressione, la dignità dell'individuo e la libertà di iniziativa

577 ss.

[40] R. Bortone, F. Cerquozzi, *L'*hate speech *al tempo di Internet*, in *Aggiornamenti sociali*, dicembre 2017, 818 ss. Osservano a tal proposito gli Autori che «a lungo andare ciò produce un effetto di distorsione nell'ecosistema informativo: l'*hate speech* è normalizzato e legittimato, con l'effetto di riprodurre pregiudizi e stereotipi verso le minoranze discriminate». *Ivi*, 824. Cfr. S. Pasta, *op. cit.*, 69; A. Sorgato, *op. cit.*, 223-224; M. Santerini, *La mente ostile*, cit., 56 ss.

[41] In tal senso, si è osservato come «le nuove tecnologie, e soprattutto internet, aiutino le persone ad ascoltare le opinioni di altri individui della stessa mentalità e a isolarsi rispetto a idee differenti», venendosi così a creare «un terreno fertile per la polarizzazione e potenzialmente pericoloso per la democrazia e la pace sociale». C. Sunstein, *Republic.com*, Princeton 2001, trad. it, *Republic.com: cittadini informati o consumatori di informazioni?*, Bologna 2003, 83. Cfr. M. Castells, *Communication Power*, 2° ed., Oxford, 2013. Cfr. E. Pariser, *Filter Bubble: How the New Personalized Web Is Changing What We Read and how We Think*, New York 2011, trad. it. *Il filtro. Quello che internet ci nasconde*, Milano 2012.

[42] S. Asch, *Social Psychology*, New York 1952, trad. it. *Psicologia sociale*, Torino 1963, 523 ss.; Id., *Opinions and social pressure*, in *Readings about the social animal*, 1955, 17-26. Vari studi dimostrano come gli utenti risentano molto della pressione alla conformità e del desiderio di essere popolari, cercando di evitare di lasciare traccia digitale delle proprie opinioni minoritarie. Inoltre, secondo alcune ricerche, i *social network* ridurrebbero notevolmente l'espressione di opinioni reali: chi percepisce di avere un'opinione differente agendo *online* decide di non esprimerla molto più spesso che *offline*. Cfr. S. Pasta, *op. cit.*, 86 ss.

economica dei fornitori di servizi Internet e di *social networking*. Uno «scontro che tocca direttamente le radici del costituzionalismo»[43] e che deve impegnare il giurista in delicatissime attività di bilanciamento nell'ottica di costruire un *web* "costituzionalizzato"[44].

Per quanto concerne il bilanciamento fra tutela della dignità e libertà di espressione, si registra una notevole distanza fra approccio europeo e approccio nordamericano.

L'Unione Europea ha messo in campo un'ampia strategia contro l'*hate speech* e altri contenuti offensivi, combinando, in chiave complementare, strumenti di *hard law* e *soft law*[45]. Fra i primi, che riflettono una logica settoriale, la direttiva 2011/93/UE, che concerne la lotta contro gli abusi sessuali sui minori e la pedopornografia *online*, e la direttiva 2018/1808/UE sui servizi di media audiovisivi, volta, fra l'altro, a contrastare la proliferazione di contenuti nocivi per i minori e l'istigazione all'odio. Fra i secondi, va menzionato il "Codice di condotta per lottare contro le forme illegali di incitamento all'odio *online*" del 2016, su cui si tornerà nel prosieguo.

Con particolare riferimento alla criminalizzazione dei discorsi d'odio, l'OSCE ha sottolineato come essi rappresentino «una grave preoccupazione, poiché possono creare un ambiente favorevole al verificarsi di crimini ispirati dall'odio e, per tale via, alimentare conflitti sociali su più larga scala» e l'*European Commission against Racism and Intolerance* (ECRI), con la Raccomandazione n. 15 del 2012, ha invitato gli Stati membri ad adottare misure penali appropriate ed efficaci per combattere il ricorso, nella sfera pubblica, al discorso d'odio che abbia lo scopo, o ci si possa ragionevolmente attendere abbia l'effetto, di incitare a commettere atti di violenza, di intimidazione, di ostilità o di discriminazione nei confronti delle persone prese di mira, a meno che altre misure meno restrittive possano rivelarsi efficaci e purché il diritto alla libertà di espressione e di opinione sia rispettato.

Nell'ordinamento italiano, ad esempio, sulla scorta di tali indicazioni, l'area di penale rilevanza dell'*hate speech* è stata definita con l'introduzione dell'art. 604-bis c.p., che punisce chi fa propaganda di idee fondate sulla superiorità o sull'odio razziale o etnico, ovvero istiga a commettere o commette atti di discriminazione per motivi razziali, etnici, nazionali o religiosi; chi, in qualsiasi modo, istiga a commettere o commette violenza o atti di provocazione alla violenza per motivi razziali, etnici, nazionali o religiosi, nonché, ancora, chi promuove o partecipa a un'organizzazione, associazione, movimento o gruppo avente tra i propri scopi l'incitamento alla

[43] O. Pollicino, G. De Gregorio, Hate speech*: una prospettiva di diritto comparato*, in *IGDA*, 2019, 4, 421 ss. Cfr. V. Zeno-Zencovich, *Freedom of Expression: A Critical and Comparative Analysis*, New York 2018.

[44] S. Rodotà, *Il diritto di avere diritti*, Roma-Bari 2012.

[45] Cfr. I. Gasparini, *L'odio ai tempi della rete: le politiche europee di contrasto all'*online hate speech, in *Jus*, 2017, 2, 510 ss.

discriminazione o alla violenza per motivi razziali, etnici, nazionali o religiosi. La norma contempla poi, in chiusura, il caso in cui la propaganda ovvero l'istigazione e l'incitamento si fondano in tutto o in parte sulla negazione, sulla minimizzazione in modo grave o sull'apologia della Shoah o di altri crimini previsti dallo statuto della Corte penale internazionale[46].

Nel panorama giuridico statunitense, invece, viene tendenzialmente accordata preminente rilevanza al primo emendamento, posto a tutela della libertà di parola, registrandosi nella giurisprudenza della Corte Suprema la costante affermazione della prevalenza del *free speech* rispetto alla repressione dell'*hate speech* e quindi all'esigenza di tutelare altri interessi come la dignità o la pace sociale. In pochissimi casi, la Corte ha escluso dall'ambito applicativo del primo emendamento lo *speech-act*, ossia quel discorso o quella manifestazione d'odio che comporta il pericolo concreto di un passaggio diretto all'azione (*fighting words*)[47]. In materia di *hate speech*, come in materia di *fake news*, prevale quindi un atteggiamento astensionistico, fondato sull'ideale del *free market of ideas*[48], secondo il quale la rete possiederebbe già gli antidoti rispetto a tali fenomeni, poiché, come un prodotto scadente finisce per soccombere, nelle logiche di mercato, rispetto ai prodotti di qualità, lo stesso varrebbe per le idee offensive e dunque "scadenti"[49]. Tale impostazione, tuttavia, pare non tenere conto dei meccanismi di personalizzazione della rete che, di fatto, impediscono un dibattito virtuale autenticamente pluralistico e quindi paragonabile a un mercato perfettamente concorrenziale[50].

---

[46] Sull'evoluzione della legislazione antiodio in Italia, v., *inter alios*, F. Basile, *Ti odio, "in nome di Dio". L'incriminazione dell'odio e della discriminazione (in particolare, per motivi religiosi) nella legislazione italiana*, in *DPU*, 4 dicembre 2019; M. Pelissero, *Discriminazione, razzismo e il diritto penale fragile*, in *DPP*, 2020, 8, 1018 ss.; A. Spena, *op. cit.*, 577 ss. Vanno inoltre segnalate, da una parte, la nuova strategia contro l'antisemitismo presentata, in data 27 gennaio 2021, da Milena Santerini, Coordinatrice nazionale per la lotta contro l'antisemitismo, coadiuvata da un gruppo tecnico di cui fanno parte, fra gli altri, Gabrio Forti e Giovanni Canzio, dall'altra, la Proposta di legge ZAN e altri, "Modifiche agli articoli 604-bis e 604-ter del codice penale, in materia di violenza o discriminazione per motivi di orientamento sessuale o identità di genere", presentata il 23 marzo 2018.

[47] Sul bilanciamento fra *hate speech* e altre istanze di tutela nella giurisprudenza nordamericana, si veda, diffusamente, A. Galluccio, *Punire la parola pericolosa? Pubblica istigazione, discorso d'odio e libertà d'espressione nell'era di internet*, Milano 2020.

[48] R. Coase, *Markets for goods and Market for ideas*, in *American Economic Review*, 1974. Secondo tale teoria il pluralismo informativo dovrebbe fungere da antidoto a fenomeni di indottrinamento, oligopoli informativi e politicamente orientati. Cfr. M. Bassini, G. E. Vigevani, *Primi appunti su* fake news *e dintorni*, in *ML*, 2017, 1, 18.

[49] Nondimeno, vi sono pensatori nordamericani che guardano con interesse all'approccio europeo e alla necessità di un intervento statale o della comunità internazionale al fine di garantire in maniera uniforme la dignità delle persone, soprattutto di quelle più deboli, e di salvaguardare l'ordine pubblico. *Ex pluribus*, J. Waldron, *Dignity and deformation: The visibility of hate*, in *HLR*, 2010, 123, 1597-1657.

[50] Cfr. M. Cuniberti, *Il contrasto alla disinformazione in rete tra logiche del mercato e (vecchie e nuove) velleità di controllo*, in *Media Education*, 2017, 1, 35; A. Nicita, *È possibile il libero scambio nel mercato della verità?*, in *Il Foglio*, 13 gennaio 2017.

Nondimeno, i grandi fornitori di servizi di *social networking*, aventi per lo più sede negli USA, non si sono finora mostrati del tutto insensibili all'approccio europeo e alla necessità di adottare iniziative, anche in termini di autoregolamentazione, a tutela della dignità degli utenti, se non altro per preservare il proprio capitale reputazionale.

Maggiori affinità si colgono fra strategia europea e nordamericana di contrasto della *non consensual-pornography*. Ad oggi, la diffusione di contenuti sessualmente espliciti senza il consenso della persona raffigurata costituisce reato in più di trenta Stati federati USA[51] e in Europa, in attuazione della Convenzione di Istanbul del 2011 (Convenzione del Consiglio d'Europa sulla prevenzione e la lotta contro la violenza nei confronti delle donne e la violenza domestica) e della cd. "direttiva vittime" (Direttiva 2012/29/UE del Parlamento europeo e del Consiglio del 25 ottobre 2012 che istituisce norme minime in materia di diritti, assistenza e protezione delle vittime di reato), parte integrante della strategia eurounitaria di tutela dei soggetti vulnerabili *offline* come *online*, molti Paesi hanno introdotto fattispecie incriminatrici *ad hoc*. In Italia, ad esempio, con la l. 19 luglio 2019, n. 69, cd. "Codice rosso", è stato introdotto l'art. 612-ter c.p. che punisce chi invia, consegna, cede, pubblica o diffonde senza il consenso delle persone rappresentate immagini o video a contenuto sessualmente esplicito, sia che li abbia realizzati o sottratti (comma 1) sia che li abbia ricevuti o comunque acquisiti (comma 2), purché, in questo secondo caso, abbia agito con il dolo specifico di recare nocumento alla persona offesa[52].

Con riferimento al bilanciamento fra istanze di tutela dell'utente e iniziativa economica privata, la disciplina originaria della rete, improntata a garantire ai *provider* massima libertà, nella leale concorrenza[53], pare a molti obsoleta rispetto all'attuale natura delle piattaforme digitali e, in particolare, dei *social network*, quali «vere e proprie imprese», che però assumono sempre di più «le sembianze di formazioni sociali ove si esprime la personalità dell'individuo»[54]. Si è osservato, inoltre, come l'evoluzione digitale e l'avvento delle piattaforme che ospitano c.d. *user generated content* abbiano trasformato «Internet da uno spazio sconfinato di libera informazione

[51] V. sul punto G. M. Caletti, *"Revenge porn" e tutela penale. Prime riflessioni sulla pornografia non consensuale alla luce delle esperienze angloamericane*, in *DPC, Riv. Trim.*, 2018, 3, 63 ss.

[52] Per un'approfondita analisi della fattispecie si vedano, *inter alios*, G. M. Caletti, *Libertà e riservatezza sessuale*, cit., 2045 ss.; N. Amore, *La tutela penale della riservatezza sessuale nella società digitale. Contesto e contenuto del nuovo "cybercrime" disciplinato dall'art. 612-ter c.p.*, in *LP*, 2020, 1, 1 ss.

[53] «Tutta la storia delle telecomunicazioni degli ultimi anni si può effettivamente ricostruire alla luce del rapporto, talora dialettico, fra regolazione e concorrenza, che connota l'assetto definito a livello nazionale ed europeo». M. Mensi, P. Falletta, *op. cit.*, 11. Cfr. F. Marini Balestra, *Manuale di diritto europeo e nazionale delle comunicazioni elettroniche*, Padova 2016.

[54] Così P. Passaglia, *Internet nella Costituzione italiana: considerazioni introduttive*, in M. Nisticò e P. Passaglia (a cura di), *Internet e Costituzione*, Torino 2014, 37.

a uno spazio controllato da poche grandi *tech-companies*, che controllano e filtrano l'accesso alle informazioni e la loro diffusione attraverso l'utilizzo di algoritmi»[55], sfruttati a fini di lucro. Tale metamorfosi del cyberspazio esige un ripensamento del ruolo dei *provider*, a partire delle piattaforme *social*, nella tutela dei diritti fondamentali degli utenti.

Parafrasando il titolo di un importante testo su giovani e nuovi media, *it's complicated*[56], bilanciare tali interessi e istanze di tutela è una delle sfide più ardue in cui il giurista, insieme agli studiosi dei più diversi ambiti disciplinari, è chiamato a cimentarsi. La «rivoluzione cibernetica», cui va riconosciuta «un'importanza strutturale o, se si preferisce, strategica per l'evoluzione del diritto, (...) in quanto rappresenta la frontiera più avanzata dell'innovazione e del cambiamento nell'odierna società globalizzata»[57], deve infatti essere accompagnata da una strategia ponderata e bilanciata, volta a garantire che «Internet non diventi strumento di diffusione e amplificazione delle diseguaglianze, ma continui a essere veicolo di conoscenza, libertà e promozione umana, e che il potenziale creativo delle nuove tecnologie accresca le potenzialità di confronto aperto fra idee, per una conoscenza libera e pluralista»[58].

4.1. Sono diversi gli elementi che hanno favorito la coagulazione del dibattito in tema di contrasto di condotte offensive *online* attorno al ruolo dei gestori delle piattaforme *social*. Fra questi: la loro possibilità "tecnica" di intervento e lo sfruttamento economico, da parte degli stessi, dell'indicizzazione dei contenuti basata sulla profilazione dell'utenza[59]. Nondimeno, qualsivoglia forma di loro coinvolgimento o responsabilizzazione va calibrata considerando l'inevitabile interferenza, da una parte, con la loro libertà d'iniziativa economica privata, di cui si è detto poc'anzi,

[55] M. R. Allegri, *Ubi Social, Ibi Ius*, cit., 214.

[56] D. Boyd, *It's complicated. The social lives of networked teens*, New Haven 2014, trad. it. It's complicated. *La vita sociale degli adolescenti sul* web, Roma 2018.

[57] Cfr. L. Picotti, *Diritto penale e tecnologie informatiche*, cit., 36. In tal senso, la disciplina giuridica della rete «si pone come una sorta di banco di prova per il diritto, che si confronta con l'ineludibile necessità di adottare istituti adeguati alle peculiarità di un fenomeno ormai non più nuovo, ma dalla sconvolgente portata eversiva». M. Mensi, P. Falletta, *op. cit.*, 61.

[58] M. Mensi, P. Falletta, *op. cit.*, 37. In tal senso, si è osservato come «non vi è libertà senza regole, nel mondo *online* come nel mondo *offline*», «proprio nell'ottica di valorizzare internet come strumento e veicolo di libertà e conoscenza, regole e procedure si pongono pertanto come fondamentali strumenti di garanzia e tutela dei diritti nella Rete». *Ivi*, 33.

[59] G. Ziccardi, *op. cit.*, 93. In tal senso, C. Novelli, *La Cassazione su responsabilità del* provider *e contenuto della notifica*, in *DI*, 2019, 50, sottolinea come gli operatori del cyberspazio «se sono in grado di gestire questi dati, utilizzarli a loro convenienza allora hanno anche una posizione privilegiata quando gli si chiede di identificarli e sorvegliarli nelle ipotesi di violazioni di legge».

dall'altra, con la disciplina in materia di *Internet Service Provider* (*ISP*), categoria cui i *social network* appartengono.

Il punto di partenza per delineare il quadro normativo di riferimento è la c.d. direttiva sul commercio elettronico (Direttiva 2000/31/CE), che definisce l'*ISP* come «la persona fisica o giuridica che presta un servizio della società dell'informazione»[60], distinguendo fra prestatori che svolgono attività di mero trasporto (*mere conduit*), prestatori che memorizzano temporaneamente informazioni (*caching*) e prestatori che ospitano e quindi memorizzano durevolmente informazioni (*hosting*). Nell'ambito di tale classificazione, le piattaforme *social*, pur non contemplate espressamente dalla direttiva, si connotano come *host provider*, potendo essere definite, in particolare, «piattaforme basate sulla condivisione di contenuti multimediali creati dagli utenti»[61], (c.d. *user generated content*).

Tornando alla direttiva, per l'attività di *hosting*, quale quella svolta dai *social network*, il *provider* non è responsabile dei contenuti "ospitati" alla duplice condizione che: «a) il prestatore non sia effettivamente al corrente del fatto che l'attività o l'informazione è illecita e, per quanto attiene ad azioni risarcitorie, non sia al corrente di fatti o di circostanze che rendono manifesta l'illegalità dell'attività o dell'informazione; b) non appena al corrente di tali fatti, agisca immediatamente per rimuovere le informazioni o per disabilitarne l'accesso»[62]. Ancora, all'art. 15, viene prescritto che gli Stati membri non impongano ai prestatori «un obbligo generale di sorveglianza sulle informazioni che trasmettono o memorizzano né un obbligo generale di ricercare attivamente fatti o circostanze che indichino la presenza di attività illecite», fatta salva la possibilità di chiedere agli stessi di «informare senza indugio la pubblica autorità competente di presunte attività o informazioni illecite dei destinatari dei loro servizi», di «comunicare alle autorità competenti, a loro richiesta, informazioni che consentano l'identificazione dei destinatari dei loro servizi con cui hanno accordi di memorizzazione dei dati». Alla direttiva nel nostro ordinamento è

[60] Art. 2, lett. b), della Direttiva 2000/31/CE del Parlamento Europeo e del Consiglio dell'8 giugno 2000 relativa a taluni aspetti giuridici dei servizi della società dell'informazione, in particolare il commercio elettronico, nel mercato interno («direttiva sul commercio elettronico»). Più ampia la definizione di *service provider* fornita dall'art. 1, lett. c., della Convenzione del Consiglio d'Europa sulla criminalità informatica, aperta alla firma a Budapest nel 2001: «qualunque entità pubblica o privata che fornisce agli utenti dei propri servizi la possibilità di comunicare attraverso un sistema informatico», nonché «qualunque altra entità che processa o archivia dati informatici per conto di tale servizio di comunicazione o per utenti di tale servizio».

[61] L. Paccagnella, *Sociologia della comunicazione nell'era digitale*, Bologna 2020, 183. In particolare, i siti *social network* vengono solitamente identificati sulla base di tre caratteristiche: la possibilità di costruire e presentare un proprio profilo pubblico o semipubblico, l'articolazione di un elenco di altri utenti con cui si condivide un legame più o meno forte, la possibilità di produrre, interagire o navigare attraverso flussi messi a disposizione dalle proprie connessioni all'interno della piattaforma. *Ivi*, 181.

[62] Art. 14 della direttiva 2000/31/CE, cit.

stata data attuazione con il d.lgs. 9 aprile 2003, n. 70, che, agli artt. 16 e 17, dedicati agli *host provider*, ricalca pedissequamente la disciplina europea.

Fermo restando tale quadro normativo, l'evoluzione tecnologica e l'emersione delle piattaforme *social* hanno portato la giurisprudenza eurounitaria e nazionale al graduale superamento del principio di neutralità della rete, «secondo cui il *provider* (qualunque ne sia la natura) deve rimanere un elemento (appunto) "neutro" dell'infrastruttura tecnologica senza che possa essere obbligato a operare nessun filtraggio dei flussi informativi che la pervadono»[63]. Alle pronunce strettamente ancorate a detto principio[64], infatti, hanno iniziato ad alternarsi, a livello europeo come a livello interno, prospettive differenti, in cui viene messo in luce come molti *host provider*, anziché limitarsi alla *passiva* memorizzazione dei contenuti immessi dagli utenti, partecipano *attivamente* alla loro organizzazione, attraverso operazioni, per lo più automatizzate, di categorizzazione e indicizzazione, aventi scopo di lucro. La gestione algoritmica dei contenuti, fondata sulla profilazione dei fruitori grazie ai dati personali che forniscono e alle attività che svolgono, è infatti uno degli elementi costitutivi del modello di business delle piattaforme *social*[65].

Tale distinzione, di matrice giurisprudenziale, fra *host* attivo e *host* passivo trae fondamento dal considerando n. 42 della direttiva, il quale prevede 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 a una rete di comunicazione sulla quale

---

[63] G. P. Accinni, *Profili di responsabilità penale dell'*hosting provider *"attivo"*, in *AP*, 2017, 2, 8.

[64] La Corte di Giustizia, chiamata a pronunciarsi sulla legittimità di una norma che ingiungeva agli *ISP* di predisporre un sistema di filtraggio a tutela del diritto d'autore in rete, ha ritenuto che tale normativa violasse la direttiva sul commercio elettronico, poiché, lungi dal garantire un giusto bilanciamento fra tutela della proprietà intellettuale e tutela della libertà d'impresa dei *provider*, imponeva la predisposizione di un sistema informatico particolarmente complesso e oneroso rischiando altresì di ledere la libertà di espressione in quanto detto sistema avrebbe potuto rivelarsi inadeguato a distinguere fra attività lecite e illecite portando al blocco di contenuti leciti. Il risultato inoltre sarebbe stato, di fatto, quello di imporre ai gestori un obbligo generalizzato di sorveglianza sulle informazioni ospitate, non previsto dalla direttiva, oltreché materialmente e tecnicamente impossibile da adempiere (se non con l'uso di sistemi automatizzati di dubbia affidabilità). Corte di Giustizia, 24 novembre 2011, *Scarlet Extended SA* contro *Société belge des auteurs, compositeurs et éditeurs SCRL (SABAM)*; Corte di Giustizia, 16 febbraio 2012, *Belgische Vereniging van Auteurs, Componisten en Uitgevers CVBA (SABAM)* contro *Netlog NV*. Sul tema, diffusamente, R. Flor, Social networks *e violazioni penali del diritto d'autore. Quali prospettive per la responsabilità del fornitore del servizio*, in *RTDPE*, 3, 2012, 683 ss.

[65] M. Montanari, *La responsabilità delle piattaforme on-line. Il caso Rosanna Cantone*, in *DI,* 2017, 2, 261, il quale osserva come la produzione e condivisione gratuita di contenuti da parte dei *prosumer* (cfr. § 1) è la «base di un nuovo modello di business basato sull'aggregazione di questi per generare audience e veicolare pubblicità». Cfr. F. Zanovello, *La responsabilità dell'*Internet Service Provider. *Brevi note a Cass. 19 marzo 2019, nn. 7708 e 7709*, in *SI*, 2020, 5, 564, che evidenzia come «il perseguimento del profitto e lo sviluppo tecnologico assurgono a indizi decisivi nel far assumere al *provider* un ruolo attivo, presupponendo che lo stesso adotti una struttura e una visione d'impresa e, soprattutto, ricorra a una certa manipolazione dei dati».

sono trasmesse o temporaneamente memorizzate le informazioni messe a disposizione da terzi al solo scopo di rendere più efficiente la trasmissione» e che «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». La Corte di Giustizia ha infatti osservato come «dal quarantaduesimo considerando della direttiva 2000/31 risulta [...] che le deroghe alla responsabilità (...) riguardano esclusivamente i casi in cui l'attività di prestatore di servizi della società dell'informazione sia di ordine "meramente tecnico, automatico o passivo", con la conseguenza che detto prestatore "non conosce né controlla le informazioni trasmesse o memorizzate"»[66], rimettendo, in ultima analisi, ai giudici nazionali la valutazione della "neutralità" dell'*host*.

Sul punto la giurisprudenza italiana è stata sinora ondivaga[67]. Secondo alcune pronunce [68], servizi di indicizzazione dei contenuti e di offerta di pubblicità personalizzata sulla base della profilazione varrebbero a qualificare l'*host* come attivo; secondo un altro orientamento[69], ove il prestatore del servizio non intervenga in alcun modo sul contenuto caricato dall'utente, ma si limiti a sfruttarne commercialmente la presenza sulla propria piattaforma, lo stesso conserverebbe la sua natura di *host* passivo.

4.2. Tale "creativa" elaborazione giurisprudenziale in materia di *host provider* "attivi" ha aperto la strada alla prospettazione, nel nostro ordinamento, di modelli altrettanto variegati di responsabilità penale delle piattaforme *social* per i contenuti ospitati[70].

Stando a un primo modello, i gestori potrebbero incorrere in responsabilità penale nei soli casi di concorso commissivo nella condotta criminosa dell'utente. A tal proposito, a dimostrazione dell'ampiezza degli scenari possibili, vale la pena di sottolineare come in un caso deciso dalla *Court of Appeal* dello Stato di New York sia

[66] Cfr. Corte Giust. UE, 23 marzo 2010, *Google c. Louis Vuitton*, C-236/08–C-238/08; 24 marzo 2014, *UPC Telekabel Wien Gmbh c. Constantin Film Verleigh*, C-314/12; 16 febbraio 2012, *Sabam c. Netlog*, C-360/10; 24 novembre 2011, *Scarlet Extended c. Sabam*, C-70; 12 luglio 2011, *L'Oreal c. Ebay*, C-324/09.

[67] Cfr. M. Bassini, *La rilettura giurisprudenziale della disciplina sulla responsabilità degli* Internet service provider. *Verso un modello di responsabilità "complessa"?*, in *Federalismi.it.*, 2015, 3, 46-47.

[68] Trib. Milano, 24 febbraio 2010, *caso "Google/Vividown"*; Trib. Roma, 20 ottobre 2011, *RTI contro Choopa*; Trib. Milano, 7 giugno 2011, *RTI contro Italia Online S.R.L*; Trib. Roma, 15 luglio 2016, *RTI contro Megavideo*; Cass. civ., Sez. I, 19 marzo 2019, n. 7708, in *DO*, 2019, 1, 152 ss. Cfr. R. Bocchini, *La responsabilità di Facebook per la mancata rimozione di contenuti illeciti*, in *GI*, 2017, 629 ss.

[69] App. Milano, 7 gennaio 2015, *Mediaset c. Yahoo! Italia*; Cass. pen., Sez. III, 3 febbraio 2014, n. 5107, in *DI*, 2014, 225 ss.

[70] Sul punto, diffusamente, B. Panattoni, *Gli effetti dell'automazione sui modelli di responsabilità: il caso delle piattaforme* online, in *DPC, Riv. trim.*, 2019, 2, 33 ss. Cfr. R. Bartoli, *Brevi considerazioni sulla responsabilità penale dell'*Internet Service Provider, in *DPC*, 2013, 5, 600 ss.

stata prospettata dall'accusa la responsabilità di Facebook per aver concorso in alcuni fatti di attentato avendo attribuito particolare visibilità, attraverso attività automatizzate di indicizzazione, a post di incitamento al terrorismo. Sebbene la Corte abbia deciso, pur con una parziale *dissenting opinion*, di non sposare la tesi accusatoria[71], la vicenda è di indubbio interesse poiché solleva il problema della responsabilità dei *provider* per l'attività svolta da sistemi di intelligenza artificiale che, attivati per scopi leciti, come rendere maggiormente interessante la piattaforma per gli utenti e creare collegamenti funzionali ai loro interessi, una volta programmati, apprendono dal contesto circostante e operano in sostanziale autonomia[72].

Secondo un altro modello, che ha trovato applicazione nella giurisprudenza di legittimità[73], i gestori delle piattaforme possono essere chiamati a rispondere di concorso mediante omissione nei reati commessi dagli utenti non per non averne impedito la realizzazione, bensì per non aver impedito il protrarsi dei loro effetti dannosi (nel caso oggetto della pronuncia, trattandosi di diffamazione, per non aver interrotto la perdurante offesa alla reputazione rimuovendo il contenuto denigratorio). Del resto, pur tralasciando il dibattito dottrinale sull'assimilazione fra impedire un evento e impedire un reato[74], esigere dal *provider* l'impedimento dei reati commessi dall'utente significherebbe onerarlo di un adempimento materialmente e tecnicamente impossibile (se non avvalendosi di sistemi automatizzati di dubbia affidabilità) prima ancora che contrario al divieto di imporre obblighi generalizzati di sorveglianza (ex art. 15 della direttiva sul commercio elettronico, come recepito all'art. 17 del d.lgs. 70/2003). A ciò si aggiunge che il regime delineato dall'art. 57 c.p. non può essere esteso ai gestori, in quanto applicabile al più ai servizi *online* di informazione professionale in virtù dell'equiparabilità funzionale e ontologica degli stessi alla stampa tradizionale[75].

---

[71] *Force v. Facebook, Inc.* No. 18-397 (2nd Cir. 2019), su *www.law.justia.com*. I giudici di New York hanno ritenuto applicabile al caso di specie il Titolo 47 U.S. Code, § 230, del 1996, che esclude la responsabilità degli *interactive computer service* per i contenuti ospitati, non potendovisi derogare qualora entrino in gioco elementi di intelligenza artificiale, poiché la creazione di collegamenti e accostamenti fra contenuti e utenti, mediante algoritmi, appartiene alla logica di business della maggioranza di detti servizi.

[72] Cfr. A. Baccin, *Responsabilità penale dell'*internet service provider *e concorso degli algoritmi negli illeciti* online*: il caso force v. Facebook*, in *SP*, 2020, 5, 75 ss.

[73] Cass. pen., Sez. V, 27 dicembre 2016, n. 54946, in *GP*, 1, 2017, con nota di M. Miglio, *I gestori di un sito internet rispondono penalmente per i commenti offensivi pubblicati dagli utenti*; in *QG*, 9 gennaio 2017, con nota di F. Buffa, *Responsabilità del gestore del sito internet*.

[74] Per una ricostruzione delle problematiche sottese all'omesso impedimento di reati altrui si veda L. Bisori, *L'omesso impedimento del reato altrui nella dottrina e giurisprudenza italiane*, in *RIDPP*, 1997, 1339 ss.

[75] Così, *ex pluribus*, Cass. pen., Sez. Un., 29 gennaio 2015, n. 31022, in *DI*, 2015, 1041 ss. e Cass. pen., Sez. V, 1 febbraio 2017, n. 4873, in *FI*, 2017, 4, II, 251 ss. Si vedano diffusamente sul tema G. P. Accinni, *op. cit.*, 4 ss.; G. Corrias Lucente, *Ma i* network providers*, i* service providers *e gli* access providers *rispondono degli illeciti penali commessi da un altro soggetto mediante l'uso degli spazi che lo gestiscono?*, in *GM*, 2004, 2524 ss.; S. Seminara,

È stata quindi "ideata" una responsabilità concorsuale da omesso impedimento degli effetti di un reato altrui. Il fondamento giuridico della posizione di garanzia dell'*ISP* e di tale peculiare obbligo di "impedimento *ex post*" andrebbe ravvisato nell'art. 14, comma 1, lett. b), della direttiva, che prevede che gli *host provider* possano incorrere in responsabilità se, effettivamente a conoscenza di un contenuto illecito nei propri server, omettano di rimuoverlo. Accogliendo questa impostazione, la Cassazione ha ritenuto immune da censure la decisione dei giudici di secondo grado che, ribaltando la sentenza del Tribunale, avevano condannato il gestore di un sito per aver «mantenuto consapevolmente» un articolo, «consentendo che lo stesso esercitasse la [propria] efficacia diffamatoria» da quando ne aveva appreso l'esistenza fino al sequestro dell'ambiente virtuale[76].

Tale indirizzo non è andato esente da critiche in dottrina, anzitutto sotto il profilo della tipicità delle ipotesi di responsabilità penale. Si è osservato, in tal senso, come «nella decisione in analisi si rinviene un'enunciazione di principio (formulazione del paradigma e condanna), ma non emergono le ragioni poste alla base dello stesso: non è chiarito, in altre parole, come si possa concorrere omissivamente in un delitto già consumato, né da dove derivi l'obbligo giuridico di impedire gli effetti della diffamazione»[77]. Non si rinviene, infatti, nel codice penale una previsione che stabilisca un rapporto di equivalenza tra il non interrompere gli effetti di un reato e la sua realizzazione commissiva, sulla falsariga del capoverso dell'art. 40 c.p.[78]. Inoltre, nei particolari settori in cui il legislatore ha previsto espressamente a carico dei *provider* l'obbligo di attivarsi per segnalare ed eventualmente rimuovere determinati contenuti illeciti ha optato, in caso di inadempimento, per la sanzione amministrativa pecuniaria, ritenendola un incentivo adeguato. Si pensi a quanto previsto in materia di tutela del diritto d'autore (art. 1, comma 6, d.l. n. 72/2004, conv. in l. n. 128/2004) e di prevenzione della diffusione di materiale pedopornografico (artt. 14-ter e 14-quater l. n. 269/1998). Pare quanto meno singolare che per la mancata rimozione di un articolo diffamatorio, ancorché in assenza di una specifica previsione di legge o di un

La responsabilità penale degli operatori su internet, in *DI*, 1998, 745 ss.

[76] Cass. pen., Sez. V, 27 dicembre 2016, n. 54946, cit.

[77] A. Ingrassia, *Responsabilità penale degli* internet service provider*: attualità e prospettive*, in *DPP*, 2017, 12, 1626. Cfr. B. Panattoni, *Gli effetti dell'automazione sui modelli di responsabilità*, cit., 44 ss. e V. Pezzella, *La diffamazione. Le nuove frontiere della responsabilità penale e civile e della tutela della* privacy *nell'epoca delle* chat *e dei* social forum, Torino 2020, 824 ss.

[78] A questo proposito, si è osservato come «l'assenza di obblighi giuridici di controllo preventivo (...), e a fortiori di impedimento di reati commessi dagli utenti o da terzi sfruttando i servizi offerti, esclude, in linea di principio, la configurabilità di una responsabilità penale non solo ex art. 40 cpv. c.p., ma anche a titolo di concorso nel reato ex art. 110 c.p., benché un contributo causale, sotto il profilo strettamente tecnico-materiale, sia ravvisabile in forza delle attività e strutture messe a disposizione dall'*ISP*». L. Picotti, *Diritto penale e tecnologie informatiche*, cit., 81. In tal senso, già Id., *Fondamento e limiti della responsabilità penale dei* Service-providers *in Internet*, in *DPP*, 1999, 3, 379 ss.

ordine dell'autorità, il *provider* possa incorrere in responsabilità penale per aver concorso omissivamente nel reato, mentre, per la sua più grave inerzia a fronte del dovere, legislativamente prescritto, di segnalare un video pedopornografico, incorra in una sanzione amministrativa[79].

In conclusione, la complessità dei profili tecnici in gioco e la natura plurisoggettiva dei *provider* sconsigliano il ricorso alla sanzione penale, che per altro comporta il rischio, come si è visto, di vistose deroghe ai paradigmi classici di responsabilità e di scivolamento verso ipotesi di responsabilità oggettiva, "da posizione". In un'ottica *de iure condito*, vanno infatti evitate interpretazioni analogiche delle norme di parte generale e speciale. Si pensi, oltre che alla "creazione" di una responsabilità da omesso impedimento degli effetti di un reato, all'eventuale estensione indiscriminata ai gestori della disciplina delineata per il direttore del giornale all'art. 57 c.p. o all'applicazione agli stessi di norme incriminatrici come l'art. 378 c.p. (favoreggiamento personale). In una prospettiva *de iure condendo*, il ricorso al diritto penale, oltre a dover essere attentamente valutato sotto il profilo della sussidiarietà e della proporzionalità della risposta sanzionatoria, non pare adeguato ed efficace quanto la previsione di una responsabilità amministrativa pecuniaria, idonea a incidere sugli interessi economici dei *provider*.

Tali considerazioni depongono nel senso di promuovere, piuttosto, forme di coinvolgimento dei gestori delle piattaforme *social* fondate sul modello del *notice and take down* e sulla previsione di sanzioni amministrative a presidio degli oneri imposti. Nel nostro ordinamento, ad esempio, in alcuni ambiti, sono previsti obblighi di segnalazione a specifiche autorità (come accennato, qualora il *provider* venga a conoscenza di contenuti pedopornografici deve segnalarli al Centro nazionale per il contrasto della pedopornografia in rete ex artt. 14-ter e 14-quater della l. 3 agosto 1998, n. 269) e oneri di rimozione di contenuti segnalati da privati (ad es. su richiesta di una vittima di cyberbullismo, se ultraquattordicenne, o dei genitori ex. art. 2, della l. 29 maggio 2017, n. 71). In tali discipline di settore, per altro, il legislatore ha preso posizione con riferimento a un altro aspetto controverso della disciplina in materia di *ISP*: «l'effettiva conoscenza» del contenuto illecito che fa insorgere l'obbligo di rimozione, positivizzando un indirizzo giurisprudenziale ormai consolidato. Per garantire una tutela rapida ed effettiva dell'utente offeso da un contenuto illecito, in giurisprudenza è emersa infatti un'interpretazione estensiva che non limita le fonti dell'effettiva conoscenza alla comunicazione da parte dell'autorità, ma vi include anche la segnalazione da parte dell'utente, purché specifica e circostanziata, poiché attendere l'ordine di rimozione dell'autorità potrebbe comportare l'irrimediabile

[79] Cfr. A. Ingrassia, *Responsabilità penale degli* internet service provider, cit., 1626-1627.

pregiudizio di diritti della personalità quali l'immagine, il decoro, la reputazione e la riservatezza[80].

4.3. Se il modello del *notice and take down* si è finora mostrato il migliore in termini di tutela dell'utente, coinvolgimento del *provider* e rispetto della normativa in materia, che vieta obblighi generalizzati di controllo sui contenuti, occorre domandarsi come garantirne la più ampia implementazione ed efficacia.

In merito alle iniziative adottate a livello nazionale, va segnalata la *Netzwerkdurchsetzungsgesetz* (*NetzDG*), legge tedesca che prevede sanzioni amministrative pecuniarie a carico dei *social network* con più di due milioni di utenti registrati in Germania che non adempiano al duplice obbligo di: a) adottare un sistema efficiente di gestione delle segnalazioni degli utenti e di rimozione dei contenuti antigiuridici segnalati; b) redigere un resoconto semestrale sull'attività svolta in tal senso. I contenuti illeciti sono individuati tassativamente richiamando alcune fattispecie di reato, fra cui il pubblico incitamento alla violenza, è previsto un termine di sette giorni per la rimozione dei contenuti, che si riduce a ventiquattro ore per quelli la cui illiceità sia manifesta e, per i casi controversi, il gestore è tenuto a interpellare un organo di autoregolamentazione indipendente appositamente creato[81].

Una simile disciplina, fonte di utili spunti nella direzione di un coinvolgimento "equilibrato" dei gestori dei *social network* nella prevenzione e nel contrasto di odio e pornografia non consensuale, non ha mancato di suscitare perplessità nei commentatori e presenta, indubbiamente, alcuni limiti.

In primo luogo, si è sostenuto che la legge tedesca si ponga in contrasto con gli artt. 3 e 14 della direttiva sul commercio elettronico che sanciscono, rispettivamente, il principio della libera circolazione dei prestatori dei servizi dell'informazione e l'illiceità della condotta inerte del *provider* solo se successiva al momento dell'effettiva conoscenza del contenuto offensivo. Infatti, la previsione, nella *NetzDG*, di termini

[80] La giurisprudenza ha chiarito che la segnalazione da parte del soggetto danneggiato da un contenuto illecito determina l'effettiva conoscenza di tale contenuto in capo al *provider*. Trib. Napoli Nord, 4 novembre 2016, in *DI*, 2017, 2, 243 ss., con nota di M. Montanari, *op. cit*. Così anche Trib. Torino, 7 aprile 2017, n. 1928, in *ML*. Cfr. S. Scarpin, *La responsabilità dell'*Internet Service Provider *per omesso controllo dei contenuti illeciti immessi dagli utenti della rete*, in *DI*, 2018, 3, 263 ss.

[81] La legge definisce i "*provider* di *social network*" come «fornitori di servizi telematici che, con l'intento di trarne profitto, gestiscono piattaforme internet finalizzate a che gli utenti condividano qualsiasi contenuto con altri utenti o lo rendano accessibile al pubblico» (§ 1, comma 2, *NetzDG*). Sono espressamente escluse le piattaforme di carattere giornalistico-editoriale, nonché i contenuti presenti in comunicazioni individuali o avvenute in un gruppo chiuso, si pensi ai servizi Messenger o ai gruppi WhatsApp. J. Rinceanu, *Verso una forma di polizia privata nello spazio digitale? L'inedito ruolo dei* provider *nella disciplina tedesca dei* social network, in *SP*, 11 marzo 2021, 1 ss. Per un'analisi più approfondita dei contenuti della legge si vedano anche, *inter alios*, V. Nardi, *I discorsi d'odio nell'era digitale: quale ruolo per l'*internet service provider*?*, in *DPC*, 7 marzo 2019, 22; S. Braschi, Social media *e responsabilità dell'*Internet Service Provider, in *ML*, 2020, 3, 168 ss.

strettissimi per la rimozione dei contenuti illeciti comporterebbe il rischio per i gestori di incorrere in una responsabilità oggettiva[82]. Tali critiche non sembrano insuperabili. Il contrasto con l'art. 3 pare smentito dall'esplicita promozione, da parte della Commissione europea (Raccomandazione (UE) 2018/334 della Commissione del 1° marzo 2018 sulle misure per contrastare efficacemente i contenuti illegali *online*), del coinvolgimento dei *provider* nel contrasto dei contenuti offensivi e, in particolare, dell'adozione a tal fine di meccanismi di *notice and take down*. Per quanto concerne la ristretta tempistica di azione imposta ai gestori, aspetto cruciale per garantire una tutela immediata alle "vittime della rete" in quanto volta a evitare la viralizzazione dei contenuti offensivi, la soluzione potrebbe essere ricercata nelle dimensioni e nella strutturazione dell'organismo deputato a vagliare le segnalazioni, che dovrebbero essere tali da garantirne l'efficiente funzionamento anzitutto in termini di conoscenza tempestiva dei contenuti segnalati.

Nella dottrina tedesca, d'altro canto, sono stati messi in luce i rischi della *NetzDG* per diversi «diritti umani, appartenenti a vari attori dello spazio digitale»[83], quali «la libertà di opinione, informazione ed espressione artistica» degli utenti, i cui contenuti potrebbero essere rimossi ingiustificatamente [84] senza che sia concesso loro un contraddittorio e senza che siano messi a loro disposizione rimedi giuridici adeguati[85], e il loro «diritto all'autodeterminazione informativa», ossia al controllo sui propri dati personali che i *provider* non solo possono ma devono archiviare in caso di segnalazione[86]. Inoltre sarebbe a rischio la "libertà della professione" dei gestori, la

[82] Cfr. G. Spindler, *Internet Intermediary Liability Reloaded. The New German Act on Responsibility of Social Networks and its (In)Compatibility with European Law*, in *Journal of Intellectual Property, Information Technology and E-Commerce Law*, 2017, 2, 167 ss. Cfr. V. Claussen, Fake-news*, pluralismo informativo e responsabilità in rete*, in *ML*, 2018, 3, 119 ss.

[83] J. Rinceanu, *op. cit.*, 11 ss.

[84] Critico in tal senso B. Schünemann, *Gefährden Fake News die Demokratie, wächst aber im Strafrecht das Rettende auch?*, in *Goltdammer's Archiv für Strafrecht*, 2019, 10, 620 ss.

[85] Il ripristino dei contenuti rimossi (procedura di "*put-back*") richiede attualmente in Germania l'istaurazione di lunghi processi in sede civile, con notevoli oneri economici a carico del ricorrente. Cfr. S. Müller-Franken, *Netzwerkdurchsetzungsgesetz: Selbstbehauptung des Rechts oder erster Schritt in die selbstregulierte Vorzensur? – Verfassungsrechtliche Fragen*, in *AfP, Zeitschrift für das gesamte Medienrecht*, 2018, 1, 1 ss.

[86] Si è sottolineato come il rischio di violazioni in materia di protezione dei dati personali si acuirebbe fortemente se entrasse in vigore il progetto di "Legge per combattere l'estremismo di destra e i crimini d'odio" (GBRH-E) del 18 febbraio 2020, approvato dal Bundestag il 18 giugno e dal Bundesrat il 3 luglio 2020, che, oltre a prevedere un inasprimento delle disposizioni del codice penale (StGB-E), del codice di procedura penale (StPO-E), della legge federale sulla registrazione dei cittadini (BMG-E), della legge sull'Ufficio federale di polizia giudiziaria (BKAG-E) e della legge sui media telematici (TMG-E), interviene sulla *NetzDG* ampliando il catalogo dei reati, estendendo la punibilità ad atti prodromici alla lesione del bene giuridico e introducendo a carico dei gestori un obbligo di comunicazione all'autorità dei dati personali degli utenti segnalati per aver commesso taluni reati. *Deutscher Bundestag, Drucksache 19/23867, 2.11.2020, https://dipbt.bundestag.de/dip21/btd/19/238/1923867.pdf*. Sui contenuti e sui dubbi di costituzionalità di tale progetto di legge, estesamente, J. Rinceanu, *op. cit.*, 17 ss.

quale, se persegue uno scopo legittimo può essere limitata solo nel rispetto del principio di proporzionalità. Sul punto si è sottolineato come sulla legittimità degli scopi della *NetzDG* «non possono sorgere dubbi», consistendo nella «prevenzione dei rischi per la pubblica sicurezza» e nella «promozione di un cambiamento nella cultura del dibattito in nome della lotta ai crimini d'odio e alle *fake news* nei *social network*». Tuttavia, alcune riserve sono state sollevate con riferimento alla proporzionalità, sia in senso ampio, riferita cioè all'impiego di uno strumento di regolamentazione della rete, sia in senso stretto, per l'eccessiva onerosità delle sanzioni amministrative pecuniarie previste[87].

A ciò si possono aggiungere ulteriori rilievi. Anzitutto, una volta rimosso un contenuto offensivo, anche considerando la diffusività e le permanenza della rete, lo stesso potrebbe essere ripubblicato. In secondo luogo, circoscrivere l'applicazione della normativa alle piattaforme con un certo numero di iscritti, se, da una parte, tiene conto della loro capacità economica di dotarsi di strutture e procedure per la gestione delle segnalazioni, dall'altra, rendendo il loro "habitat" almeno apparentemente più "sicuro", rischia di avere effetti anticoncorrenziali, nel senso di rafforzare ulteriormente l'assetto già pressoché monopolistico che connota i diversi mercati di servizi *online*. Inoltre, le caratteristiche della rete possono conferire notevole risonanza anche a contenuti pubblicati su piattaforme "minori". Ancora, non possono essere ignorati i limiti di una normativa nazionale rispetto all'aterritorialità delle piattaforme digitali[88].

Alle obiezioni riportate si potrebbe rispondere apportando alla disciplina richiamata opportuni correttivi. In primo luogo, se è vero che, accanto a post il cui carattere illecito è autoevidente, si staglia un'ampia zona grigia di contenuti il cui inquadramento richiede una formazione giuridica specialistica, è fondamentale, anche per evitare che vengano eliminati contenuti leciti (cd. *overblocking*), che la gestione delle segnalazioni sia affidata a un organismo *ad hoc*, dotato delle necessarie competenze, le cui dimensioni, per garantirne l'efficienza, dovrebbero variare al

[87] Sul punto K. H. Ladeur, T. Gostomzyk, *Gutachten zur Verfassungsmäßigkeit des Entwurfs eines Gesetzes zur Verbesserung der Rechtsdurchsetzung in sozialen Netzwerken (Netzwerkdurchsetzungsgesetz - NetzDG) i.d.F. vom 16. Mai 2017, BT-Drs. 18/12356, Erstattet auf Ansuchen des bitkom*, 2017, liberamente fruibile su *https://www.bitkom.org/sites/default/files /file/import/NetzDG-Gutachten-GostomzykLadeur.pdf*.

[88] Secondo un'interessante prospettiva, la rete andrebbe considerata «patrimonio comune dell'umanità», qualificazione che avrebbe quali ricadute, fra l'altro, 1) «il divieto di estensione della sovranità nazionale ai beni ritenuti patrimonio comune»; 2) «l'obbligo di assoggettamento dei beni in questione a un regime internazionale di cooperazione per quanto concerne la loro gestione»; 3) «il divieto del loro utilizzo, anche per interessi generali, con modalità tali da arrecare pregiudizio irreparabile all'ambiente» (inteso, in questo caso, come ambiente virtuale) e 4) «l'obbligo di utilizzare tali beni esclusivamente per fini pacifici e, quindi, il divieto di porre in essere attività che siano irrispettose dei principi contenuti nella Carta delle Nazioni Unite relativi al mantenimento della pace e della sicurezza internazionale». G. M. Ruotolo, *Internet (diritto internazionale)*, in *Enciclopedia del diritto. Annali*, Milano 2014, 556 ss.

variare del numero di utenti iscritti alla piattaforma. Al contempo, a tutela della libertà di espressione, sarebbe opportuno prevedere un rapidissimo contraddittorio con l'utente prima di procedere all'eliminazione del contenuto, portandolo così, auspicabilmente, a rimuoverlo di propria iniziativa. Inoltre, avverso la decisione di rimozione l'autore del post dovrebbe avere sempre la possibilità di fare ricorso, attraverso procedure rapide e snelle, rivolgendosi a un'autorità amministrativa indipendente appositamente individuata. Ancora, per evitare la ripubblicazione dei contenuti illeciti eliminati, occorrerebbe accompagnare alla rimozione un divieto di ripubblicazione, presidiato magari dall'impiego di strumenti automatizzati che riescano a individuare e a segnalare quel determinato contenuto, indipendentemente dall'eventuale modifica del nome del file o della sua estensione. La prescrizione di quest'ultimo accorgimento, tuttavia, rischierebbe di cozzare con il divieto di imporre obblighi di controllo generalizzato[89].

Per quanto concerne la libertà di iniziativa economica privata e i profili di proporzionalità, l'obiettivo di una siffatta normativa, ossia il contrasto di condotte illecite, quanto meno di quelle riconducibili alle categorie dell'*hate speech*, dell'odio *ad personam* e della pornografia non consensuale, nonché i vantaggi economici che spesso si ricollegano alla viralizzazione dei contenuti offensivi paiono giustificare l'imposizione ai gestori dell'onere, senza dubbio economicamente gravoso, di adeguarsi alle nuove direttive[90]. Nondimeno, per assicurare la proporzionalità delle sanzioni, occorrerebbe affidarne la commisurazione, oltre che alle valutazioni discrezionali dell'autorità procedente, a sistemi analoghi al meccanismo per quote previsto dal d.lgs. 231/2001[91].

4.4. Tutto ciò premesso, permangono i predetti limiti di una normativa nazionale rispetto all'aterritorialità delle piattaforme digitali e delle offese perpetrate *online*. È lecito quindi domandarsi se una simile disciplina non possa essere adottata a livello internazionale, mediante fonti pattizie (che tuttavia dovrebbero confrontarsi con la distanza fra approccio europeo e nordamericano in tema di *free speech*), o, quanto meno, a livello europeo. In questa seconda ipotesi, gli orizzonti di riforma potrebbero seguire almeno due itinerari.

[89] M. Cuniberti, *L'art. 21 ha settant'anni. Potere e libertà*, in *ML*, 2018, 3, 53.

[90] A. Lang, *Netzwerkdurchsetzungsgesetz und Meinungsfreiheit. Zur Regulierung privater Internet-Intermediäre bei der Bekämpfung von Hassrede*, in *Archiv des Öffentlichen Recht*, 2, 2018, 227 ss.

[91] A tal fine, in Germania, apposite linee guida, emanate dalle autorità amministrative ai sensi del § 4, comma 4, *NetzDG*, dovrebbero garantire una maggiore trasparenza nella commisurazione della sanzione, che deve essere correlata alle dimensioni del fornitore, al numero dei suoi utenti, come pure alle circostanze e alle conseguenze del fatto.

Da una parte, la notevole evoluzione della realtà social-mediatica cui si è assistito nell'ultimo ventennio, come evidenziato da dottrina e giurisprudenza, ha messo in luce una certa obsolescenza della disciplina dettata dalla direttiva sul commercio elettronico, che andrebbe quindi superata, definendo un nuovo quadro normativo europeo in materia maggiormente sensibile alle nuove istanze di tutela dell'utente e di coinvolgimento dei *provider* in tal senso[92]. Detto approccio "normativo" potrebbe tradursi nell'adozione di una nuova direttiva o di un regolamento "*sui generis*", sulla falsariga del *GDPR*[93]. Intervento che si inserirebbe nel solco della competenza concorrente dell'Unione nel settore "spazio di libertà, sicurezza e giustizia" e avrebbe finalità di contrasto di ogni forma di discriminazione, ai sensi dell'art. 10 TFUE, nonché di tutela della dignità e della vita privata, come garantiti dagli artt. 1 e 7 della Carta di Nizza[94].

Dall'altra parte, potrebbe essere incoraggiata la recente tendenza all'autoregolamentazione dei *provider*, attribuendo eventualmente carattere vincolante a linee guida e codici di condotta adottati in autonomia. Si pensi al già citato "Codice di condotta per lottare contro le forme illegali di incitamento all'odio *online*", adottato nel 2016 dalla Commissione europea insieme a Facebook, Microsoft, Twitter e YouTube. Tale strumento di *soft law* prevede che le piattaforme predispongano «procedure chiare ed efficaci per esaminare le segnalazioni riguardanti forme illegali di incitamento all'odio nei servizi da loro offerti» e adottino «regole o orientamenti per la comunità degli utenti volte a precisare che sono vietate la promozione dell'istigazione alla violenza e a comportamenti improntati all'odio», riconoscendo così, fra l'altro, «il ruolo fondamentale delle piattaforme nel contrastare i fenomeni di *hate speech*»[95].

---

[92] Va segnalato che di recente è stato inaugurato un processo di revisione della normativa europea del cyberspazio. In particolare, è stata avanzata una proposta per un Regolamento del Parlamento europeo e del Consiglio relativo a un mercato unico dei servizi digitali (legge sui servizi digitali) e che modifica la direttiva 2000/31/CE, Bruxelles, 15 dicembre 2020, *https://eur-lex.europa.eu*.

[93] Regolamento (UE) 2016/679 del Parlamento europeo e del Consiglio, del 27 aprile 2016, relativo alla protezione delle persone fisiche con riguardo al trattamento dei dati personali, nonché alla libera circolazione di tali dati e che abroga la direttiva 95/46/CE (Regolamento generale sulla protezione dei dati, *General Data Protection Regulation* o *GDPR*). Fra le ragioni che hanno portato i primi commentatori a definire il GDPR un regolamento *sui generis* o una "quasi direttiva", vi è il lasso di tempo di due anni, intercorso fra la sua entrata in vigore (24 maggio 2016) e la sua effettiva applicazione (25 maggio 2018), previsto per consentire agli Stati membri e alle società di adeguarsi alle nuove disposizioni. Cfr. L. Bolognini, E. Pelino, C. Bistolfi (a cura di), *Il Regolamento* privacy *europeo. Commentario alla nuova disciplina sulla protezione dei dati personali*, Milano 2016, 705.

[94] Sul punto, in dottrina si osserva come «l'introduzione da parte del legislatore europeo di uno specifico meccanismo potrebbe risolvere la maggior parte delle incertezze in materia: la previsione e regolamentazione di procedure c.d. di *notice and take down*». B. Panattoni, *Il sistema di controllo successivo: obbligo di rimozione dell'*ISP *e meccanismi di* notice and take down, in *DPC*, 2018, 5, 256.

[95] O. Pollicino, G. De Gregorio, *op. cit*., 430.

I vantaggi dell'approccio "normativo" si apprezzano su diversi fronti. Anzitutto potrebbero essere adottate definizioni comuni, almeno a livello europeo, con riferimento a *hate speech*, principali forme di odio *ad personam* e *non-consensual pornography*[96]. In secondo luogo, potrebbe essere garantita l'armonizzazione dei sistemi di gestione delle segnalazioni e delle tempistiche di rimozione dei contenuti, assicurando così un livello uniforme di tutela dell'utente[97]. Inoltre, lungi dal tradursi in eccessive limitazioni della libertà di espressione, una normativa sovranazionale potrebbe scongiurare rischi di c.d. *collateral censorship*[98], ossia di censura privata da parte delle piattaforme, prevedendo sia che le segnalazioni siano gestite da organismi *ad hoc* dotati delle necessarie competenze, anche giuridiche[99], sia la facoltà di ricorso dell'utente avverso le decisioni di rimozione. Tutto ciò precisato, non vanno ignorate le perplessità della dottrina, che evidenzia le «difficoltà insite nell'armonizzare le diverse "sensibilità" degli ordinamenti giuridici nazionali rispetto a queste tematiche», nonché le presumibili «resistenze di quegli Stati che riterrebbero un intervento

[96] In tal senso, la comunicazione della Commissione europea "Lotta ai contenuti illeciti *online*. Verso una maggiore responsabilizzazione delle piattaforme *online*" (Com. n. 555 del 28 settembre 2017) individua fra gli elementi di debolezza nella direttiva sul commercio elettronico il fatto di non prevedere una definizione normativa di "contenuto illecito", lasciando che a ciò provvedano gli ordinamenti giuridici nazionali, oltre ad alcuni interventi legislativi settoriali adottati a livello di europeo. La già citata proposta di Regolamento relativo a un mercato unico dei servizi digitali propone, quale definizione di "contenuto illegale", all'art. 2, lett. g), «qualsiasi informazione che, di per sé o in relazione a un'attività, tra cui la vendita di prodotti o la prestazione di servizi, non è conforme alle disposizioni normative dell'Unione o di uno Stato membro, indipendentemente dalla natura o dall'oggetto specifico di tali disposizioni». Al considerando 12 viene meglio precisato che «il concetto di "contenuto illegale" dovrebbe essere definito in senso lato e comprendere anche le informazioni riguardanti i contenuti, i prodotti, i servizi e le attività illegali. Tale concetto dovrebbe in particolare intendersi riferito alle informazioni, indipendentemente dalla loro forma, che ai sensi del diritto applicabile sono di per sé illegali, quali l'illecito incitamento all'odio o i contenuti terroristici illegali e i contenuti discriminatori illegali, o che riguardano attività illegali, quali la condivisione di immagini che ritraggono abusi sessuali su minori, la condivisione non consensuale illegale di immagini private, il *cyberstalking*, la vendita di prodotti non conformi o contraffatti, l'utilizzo non autorizzato di materiale protetto dal diritto d'autore o le attività che comportano violazioni della normativa sulla tutela dei consumatori».

[97] In tema di regolazione del *web*, alla luce della natura transnazionale della rete, vi è chi sottolinea come sia di vitale importanza compiere sforzi per armonizzare i vari ordinamenti per mezzo di norme di diritto internazionale, su cui i legislatori nazionali potrebbero intervenire con norme di dettaglio. M. Mensi, P. Falletta, *op. cit.*, 64.

[98] J. M. Balkin, *Free speech and Hostile Environments*, in *CLR*, 1999, 8, 2296.

[99] Solo dotando l'organismo delle necessarie conoscenze e competenze, anche giuridiche, si potrebbero forse superare le comprensibili preoccupazioni di chi, in dottrina, osserva che «è inquietante, in sostanza, l'idea di un privato che verrebbe incaricato di esercitare una sorta di censura per conto dell'ordinamento, avendo i mezzi tecnici ma non quelli culturali per realizzarla». G. Fornasari, *Il ruolo della esigibilità nella definizione della responsabilità penale del* provider, in L. Picotti (a cura di), *Il diritto penale dell'informatica nell'epoca di internet*, Padova 2004, 431.

legislativo dell'Unione esorbitante rispetto ai imiti dettati dal rispetto dei principi di sussidiarietà e proporzionalità»[100].

L'approccio basato sulla *self-regulation*, pur apparendo più duttile e capace di adattarsi all'evoluzione tecnologica [101], non garantirebbe lo stesso livello di armonizzazione delle procedure di *notice and take down* e, quindi, la stessa uniformità nel grado di tutela riconosciuto all'utente e rischierebbe di «risultare inefficace dinanzi alla pressione derivante dagli interessi economici degli operatori del settore»[102].

A ben vedere, le due dimensioni, armonizzazione garantita dall'approccio normativo e duttilità della *self-regulation*, potrebbero trovare aree di contatto laddove la disciplina introdotta a livello sovranazionale si fondasse, al pari del *GDPR*, sul concetto di *accountability*, di «responsabilità fiduciaria» [103], ossia di responsabilizzazione dei gestori attraverso l'imposizione di obblighi di risultato, di *compliance*, lasciando margini di libertà organizzativa sui mezzi per raggiungerli, includendo magari, secondo il modello della *responsive regulation* [104], misure amministrative flessibili che permettano ai *provider*, in sede di prima contestazione, di ottenere una riduzione o la mancata applicazione della sanzione amministrativa in caso di tempestiva revisione delle procedure interne.

L'approccio responsivo insito nel principio di *accountability*, che può essere tradotto come principio di responsabilizzazione e obbligo di rendicontazione, è nato in ambito aziendale per indicare i doveri di trasparenza, «intesa come garanzia della completa accessibilità alle informazioni agli utenti», di responsività, «intesa come la capacità di rendere conto di scelte, comportamenti e azioni» e di *compliance*, «intesa come capacità di far rispettare le norme»[105]. Il suo accoglimento da parte del *GDPR* si

---

[100] M. R. Allegri, *op. cit.*, 211.

[101] In tal senso, in dottrina vi è chi individua come soluzione ottimale per garantire il massimo rispetto della libertà in rete la redazione di linee guida orientative per l'autoregolamentazione di *provider* e gestori di *social network*, incentivando «meccanismi collaborativi e di auto-condotta», nell'ottica di trovare il «giusto equilibrio tra efficacia dell'azione di tutela e garanzie di libera iniziativa economica degli operatori stessi e di libertà di espressione degli utenti della rete». In particolare, si propone di «operare su due piani distinti: in senso generale, mediante la redazione di linee guida dirette a tutti gli operatori del settore e, in maniera più mirata, concludendo specifici protocolli di intesa con i soggetti maggiormente significativi (*in primis* i *social network* più diffusi)». M. Mensi, P. Falletta, *op. cit.*, 186-187. In tal senso anche M. Monti, *Privatizzazione della censura e* Internet platforms*: la libertà d'espressione e i nuovi censori dell'agorà digitale*, in *RIID,* 2019, 1, 45.

[102] M. R. Allegri, *op. cit.*, 211. Aggiunge l'Autrice: «solo la periodica valutazione dei risultati eventualmente ottenuti dalle pratiche di autoregolamentazione potrà confermarne la reale efficacia, o piuttosto declassarle a operazioni essenzialmente "cosmetiche" a ridotto impatto».

[103] Cfr. V. Pelligra, *I paradossi della fiducia: scelte razionali e dinamiche interpersonali*, Bologna 2007.

[104] Cfr. J. Braithwaite, P. Pettit, *Not Just Deserts. A Republican Theory of Criminal Justice*, Oxford 1990, 54 ss.; I. Ayres, J. Braithwaite, *Responsive Regulation. Transcending the Deregulation Debate*, Oxford 1992.

[105] M. Iaselli, Privacy*: cosa cambia con il nuovo regolamento europeo*, Assago 2016, 9. Vedi anche C. Bistolfi, *Le obbligazioni di* compliance *in materia di protezione dei dati*, in L. Bolognini, E. Pelino, C. Bistolfi (a cura di), *op. cit.*, 323 ss.; P. Pizzetti, Privacy *e il diritto europeo alla protezione dei dati personali*, Torino 2016, 282 ss.; F. Midiri,

è collocato nel solco di «un notevole cambiamento culturale e di approccio», vale a dire il «passaggio da una concezione prettamente formale di mero adempimento a un approccio sostanziale di tutela dei dati e delle persone stesse»[106], in base al quale ciò che viene richiesto è di adottare adeguate misure di sicurezza, di condurre periodicamente l'analisi dei rischi, nonché di progettare e mettere in atto servizi e programmi volti a consentire la rilevazione di eventuali violazioni. In particolare, il meccanismo di responsabilizzazione che viene in gioco si fonda sull'assunto in base al quale l'affidabilità è la risposta "naturale" all'attribuzione di fiducia. In altri termini, secondo tale teoria, l'affidabilità (*trustworthiness*) è il frutto di una capacità di autoriflessione che emerge nella relazione con l'altro, il quale, dando fiducia (*trustfulness*), stimola una risposta in termini di restituzione e di resistenza alla tentazione dell'opportunismo[107]. Nel caso della prevenzione e del contrasto di *hate speech*, odio *ad personam* e pornografia non consensuale, "dando fiducia" al gestore, ossia conferendogli specifici compiti organizzazione e gestione, si perseguirebbe l'obiettivo di stimolare una sua "rispondenza fiduciaria", ossia di innescare processi virtuosi di partecipazione attenta e operativa nel processo di "costituzionalizzione della rete".

Inoltre, sempre secondo il modello del *GDPR*[108], potrebbe essere incoraggiata l'adozione di codici di condotta, da sottoporre ad accreditamento, al fine di coinvolgere i gestori non solo nel contrasto dei fenomeni afferenti al c.d. "lato oscuro della rete", ma altresì nella definizione delle regole e delle procedure da seguire per raggiungere tale obiettivo[109]. Riprendendo la *NetzDG* (§ 3, comma 6), si tratterebbe di una "autoregolamentazione regolata" ("*Regulierten Selbstregulierung*").

Si è osservato in tal senso come «l'armonizzazione delle fonti e dei rimedi giuridici, per quanto debba partire da una base normativa di matrice sovranazionale e, quindi, in prima istanza europea, non potrà che essere integrata anche da componenti

Il diritto alla protezione dei dati personali: regolazione e tutela, Napoli 2017.

[106] M. Alovisio, F. Di Resta, *Norme* privacy *UE, ecco tutto ciò che bisogna sapere su* accountability *e sicurezza*, in *www.agendadigitale.eu*, 20 giugno 2016, in cui si precisa che, mentre il precedente codice della *privacy* prevedeva «un sistema di responsabilità (…) e una serie di adempimenti formali (informativa, consenso, notificazione al Garante, misure minime e idonee) ma non un approccio di responsabilizzazione», pubbliche amministrazioni e imprese «devono, alla luce del nuovo regolamento in materia di protezione dei dati personali, ripensare attivamente le modalità di gestione e di utilizzo dei dati personali attraverso una loro maggiore responsabilizzazione e adattandosi ai nuovi istituti previsti».

[107] L'Autore spiega come il circolo virtuoso della fiducia abbia il suo fulcro nella decisione dell'agente B di non tradire la fiducia che gli è stata attribuita dall'agente A. Il fatto che l'agente A si fidi dell'agente B, lo ritenga affidabile, non è che una conseguenza naturale. Pertanto, occorre ribaltare la prospettiva tradizionale secondo la quale "ci si fida di chi è affidabile", pervenendo all'opposta conclusione che "è affidabile colui di cui ci si fida". V. Pelligra, *op. cit.*, 189 ss.

[108] Artt. 40 ss. del *GDPR*.

[109] In tal senso, anche M. Cuniberti, *L'art. 21*, cit., 54 ss.

di auto-regolamentazione – come testimoniano le posizioni assunte più volte dalla Commissione europea e dal Consiglio d'Europa in materia – che sono imprescindibili data la natura fluida del *Cyberspace*, il quale non si presta ad essere rigidamente ed esaustivamente disciplinato attraverso gli strumenti e gli schemi del solo diritto pubblico»[110].

4.5. A ben vedere, la filosofia che informa il *GDPR*, imperniata sul principio di "responsabilizzazione", vale a dire sull'imposizione di oneri di *compliance*, con ampi margini di libertà organizzativa in relazione alle misure da adottare per prevenire determinati rischi, non è distante dall'approccio adottato nel nostro ordinamento in materia di responsabilità amministrativa degli enti [111], altro possibile modello di disciplina per i gestori delle piattaforme *social*. Il parallelismo fra le due normative, tuttavia, è vistosamente imperfetto.

Entrambe si fondano su un "*risk-based approach*", imponendo, rispettivamente, al titolare (ed eventualmente al responsabile) dei dati personali e all'ente un ruolo proattivo, ossia quello di adottare, sulla base di un'accurata mappatura dei rischi, tutte le misure tecniche e organizzative, da una parte, e un modello di organizzazione e gestione, dall'altra, idonei a prevenire il rischio di *data breaches* nell'un caso e di commissione di uno dei reati presupposto nell'altro. Alla dimostrazione di aver adottato le cautele idonee a scongiurare il rischio della violazione poi effettivamente verificatasi ambedue le discipline subordinano benefici in termini di esenzione da responsabilità o riduzione delle sanzioni amministrative pecuniarie previste. L'art. 24 del *GDPR* prevede infatti, positivizzando il principio di *accountability*, che, «tenuto conto della natura, dell'ambito di applicazione, del contesto e delle finalità del trattamento, nonché dei rischi aventi probabilità e gravità diverse per i diritti e le libertà delle persone fisiche, il titolare del trattamento mette in atto misure tecniche e organizzative adeguate per garantire, ed essere in grado di dimostrare, che il trattamento è effettuato conformemente» al regolamento e che tali «misure sono

[110] B. Panattoni, *Il sistema di controllo successivo*, cit., 262-263.

[111] Sul tema, *ex pluribus*, M. Romano, *La responsabilità amministrativa degli enti, società o associazioni: profili generali*, in *RS*, 2002, 2-3, 393 ss.; G. De Vero, *La responsabilità penale delle persone giuridiche*, Milano 2008; C. Piergallini, *Paradigmatica dell'autocontrollo penale (dalla funzione alla struttura del "modello organizzativo"* ex *d. lgs. n. 231/2001)*, in *Studi in onore di Mario Romano*, Napoli 2011, 2049 ss.; G. De Simone, *Persone giuridiche e responsabilità da reato. Profili storici, dogmatici e comparatistici*, Pisa 2012; G. Forti, *Uno sguardo ai "piani nobili" del d.lgs. n. 231/2001*, in *RIDPP*, 2012, 1249 ss.; M. Caputo, *Colpevolezza della persona fisica e colpevolezza dell'ente nelle manovre sulla pena delle parti*, in *RIDPP*, 2017, 148 ss.; F. Centonze, *Responsabilità da reato degli enti e* agency problems*. I limiti del d.lgs.vo n. 231 del 2001 e le prospettive di riforma*, in *RIDPP*, 2017, 945 ss.; S. Manacorda, *L'idoneità preventiva dei modelli di organizzazione nella responsabilità da reato degli enti: analisi critica e linee evolutive*, in *RTDPE*, 2017, 107; V. Manes, *Profili e confini dell'illecito para-penale*, in *RIDPP*, 2017, 988 ss.; V. Mongillo, *La responsabilità penale tra individuo ed ente collettivo*, Torino 2018; C. E. Paliero, *La colpa di organizzazione tra responsabilità collettiva e responsabilità individuale*, in *RTDPE*, 2018, 187.

riesaminate e aggiornate qualora necessario». A tal fine, in alcuni casi, il titolare è tenuto a procedere alla cd. "valutazione d'impatto", una procedura, che va dettagliatamente documentata, volta a individuare «l'origine, la natura, la particolarità e la gravità» del rischio per la tutela del diritto alla protezione dei dati[112]. Delle «misure tecniche e organizzative messe in atto» e delle «misure adottate (…) per attenuare il danno subito dagli interessati» viene tenuto conto per valutare il grado di responsabilità del titolare o del responsabile del trattamento e per commisurare la sanzione (art. 83 del *GDPR*). Similmente, la 231, che fa perno sul concetto di "colpa organizzativa", agli articoli 6 e 7, prevede che l'ente vada esente da responsabilità se dimostra, fra l'altro, di aver «adottato ed efficacemente attuato, prima della commissione del fatto, modelli di organizzazione e di gestione idonei a prevenire reati della specie di quello verificatosi». Il rimprovero mosso all'ente consiste quindi nell'aver colposamente omesso di predisporre una struttura aziendale idonea a prevenire la commissione di uno dei reati presupposto espressamente previsti (art. 24 ss.)[113].

Ancora, entrambe le normative prevedono che l'adeguatezza delle misure adottate e il rispetto degli oneri imposti possano essere suffragati dall'adesione a determinati codici di condotta (o "codici etici"), elaborati dalle organizzazioni rappresentative degli enti o dei titolari del trattamento (artt. 6 del d.lgs. 231/2001 e 40 del *GDPR)*. Inoltre, come è prassi che gli enti si servano di appositi organismi, dotati di autonomi poteri di iniziativa e di controllo a cui, come previsto dalla 231, affidano «il compito di vigilare sul funzionamento e l'osservanza dei modelli e di curare il loro aggiornamento» (art. 6), così è prevista dal *GDPR*, talvolta in via facoltativa, talvolta obbligatoriamente, la nomina di un Responsabile della Protezione dei Dati (secondo la denominazione più comune "*DPO*", acronimo dell'inglese "*Data Protection Officer*"), ossia un organo individuale o collegiale chiamato a fornire consulenza in materia di *privacy compliance*[114].

Nondimeno, terminate le analogie, alcuni aspetti cruciali differenziano le due discipline. Per quanto concerne l'ambito applicativo, oltre al fatto che il *GDPR* è un

[112] C*onsiderandum* 84 del *GDPR.* La valutazione d'impatto, sempre consigliata, è obbligatoria, ex art. 35 del *GDPR*, in caso di attività che implichino «una valutazione sistematica e globale di aspetti personali relativi a persone fisiche, basata su un trattamento automatizzato, compresa la profilazione, e sulla quale si fondano decisioni che hanno effetti giuridici o incidono in modo analogo significativamente su dette persone fisiche»; di «trattamento, su larga scala» di dati sensibili e di «sorveglianza sistematica su larga scala di una zona accessibile al pubblico».

[113] Cfr. D. Costa, *I modelli 231 e la* compliance *aziendale sulla tutela dei dati personali. Aspetti comuni e divergenze a quattro anni di distanza dall'entrata in vigore del* GDPR*,* in *GP,* 2020, 5, 1 ss.

[114] La nomina del *DPO* è obbligatoria, a norma dell'art. 37 del *GDPR*, quando il trattamento è effettuato da un'autorità pubblica o da un organismo pubblico, quando il trattamento richieda il monitoraggio regolare e sistematico degli interessati su larga scala e nei casi nei quali le attività principali del titolare o del responsabile consistono nel trattamento, su larga scala, di dati sensibili.

regolamento europeo, mentre la 231 è un normativa nazionale, le previsioni del primo si rivolgono al titolare ed, eventualmente, al responsabile dei dati, sia esso persona fisica o giuridica, mentre la seconda si rivolge alle sole persone giuridiche (ad esclusione di Stato, enti pubblici territoriali, enti pubblici non economici nonché enti che svolgono funzioni di rilievo costituzionale). La differenza più evidente consiste però nel fatto che, diversamente da quanto previsto in materia di responsabilità amministrativa dell'ente, il titolare del trattamento risponde anche se la violazione non è stata commessa a suo vantaggio o nel suo interesse.

A ben vedere, il modello di responsabilità amministrativa dei *provider* delineato dalla *Netzwerkdurchsetzungsgesetz*, sebbene improntato alla medesima logica di responsabilizzazione e prevenzione del rischio, si discosta da entrambi i paradigmi illustrati in quanto prescinde non solo dal vantaggio tratto dal *provider*, ma ancor prima dall'integrazione di condotte illecite da parte degli utenti: una sorta di colpa organizzativa "pura", come quella in cui, secondo recenti proposte di riforma, dovrebbe essere tradotta la responsabilità dell'ente ex d.lgs. 231. In tal senso, in un'ottica *de lege ferenda*, è stato prospettato in dottrina il superamento della *responsabilità per imputazione* dell'ente, auspicando la transizione verso uno schema autonomistico, «una responsabilità "originaria" o "per fatto proprio" dell'ente», ossia una «tipologia di *corporate criminal responsibility* normativamente svincolata dal "fatto" della persona fisica». Si tratterebbe, in altri termini, di creare «un illecito di condotta, nella forma omissiva propria», che sanzioni «la mera inosservanza del dovere di auto-organizzazione» [115], in modo analogo a quanto avviene in altri ordinamenti [116]. Una siffatta responsabilità amministrativa autonoma, di carattere omissivo proprio, per mancato adempimento di doveri di *compliance*, potrebbe essere introdotta a carico dei gestori delle piattaforme *social*, escludendo invece forme di responsabilità penale per concorso commissivo od omissivo di dubbia compatibilità con i principi di tassatività e colpevolezza.

[115] E. Greco, *L'illecito dell'ente dipendente da reato. Analisi strutturale del tipo*, in *RIDPP*, 4, 2123-2124. Cfr. G. De Simone, *Il «fatto di connessione», tra responsabilità individuale e responsabilità corporativa*, in *RTDPE*, 2011, 55; F. D'Alessandro, *Responsabilità da reato degli enti e rischio d'impresa: prospettive di riforma per una* corporate governance *efficace*, in *Corporate governance*, 2020, 31 ss.; F. Centonze, M. Mantovani (a cura di), *La responsabilità «penale» degli enti. Dieci proposte di riforma*, Bologna 2016, 137 ss.

[116] Si vedano, ad esempio, nell'ordinamento britannico, il *Corporate Manslaughter and Corporate Homicide Act* e il *Bribery Act*. Cfr., sul tema, R. D. Ormerod, R. Taylor, *The Corporate Manslaughter and Corporate Homicide Act*, in *Criminal Law Review*, 2007, 589 ss.; A. Pinto, M. Evans, *Corporate Criminal Liability*, London 2013, 317 ss.; C. Wells, *Corporate Responsibility and Compliance Programs in the United Kingdom*, in S. Manacorda, F. Centonze, G. Forti (a cura di), *Preventing Corporate Corruption. The Anti-Bribery Compliance Model*, Heidelberg-New York 2014, 505 ss.

5. Per concludere, le "variabili usurpatrici" della rete, quali la possibilità di agire in forma anonima e le logiche di visibilità governate dagli algoritmi, oltre a deporre nel senso di un necessario coinvolgimento dei *provider* e, in particolare, dei fornitori che ospitano *user generated content*, nel contrasto di fenomeni particolarmente offensivi, come l'*hate speech*, l'odio *ad personam* e la pornografia non consensuale, esigono una riflessione sulla necessità di una loro regolazione.

Per quanto concerne l'anonimato, se, da una parte, gli studi di psicologia attestano che, nelle sue varie declinazioni, possa produrre effetti di deresponsabilizzazione favorendo i meccanismi di disimpegno morale, dall'altra, «sembrerebbe offrire benefici non irrilevanti anche dal punto di vista dell'autonomia dei gruppi, consentendo alle minoranze (di genere, di ceto, di etnia, di orientamento sessuale) di esprimere critiche, rivendicare pretese e organizzare forme di mobilitazione a un grado di intensità altrimenti impossibile» con effetti positivi di «redistribuzione del potere sociale»[117].

Nondimeno, vi è chi propone, quale mediazione fra le diverse istanze in gioco, di contrastare i due livelli più avanzati di anonimato, la *disapproved anonymity* e la *full anonymity*, consentendo invece un "anonimato controllato"[118], mediante l'imposizione alle piattaforme digitali dell'obbligo di chiedere agli utenti di identificarsi in fase di registrazione, garantendo la scrupolosa tutela dei dati personali così forniti. In altri termini, fatta salva la prerogativa dello pseudonimato, ossia di esprimersi e interagire in rete attraverso uno pseudonimo, si tratterebbe di prescrivere ai gestori, quanto meno delle piattaforme *social*, di richiedere agli utenti, quale condizione per registrarsi, di fornire le proprie generalità (già alcuni *social network* hanno iniziato a pretendere che gli utenti con profili "sospetti" si identifichino mediante fotografia o documento)[119]. Tale misura, in caso di commissione di fatti illeciti, consentirebbe di identificare l'utente in modo più agevole e sicuro di quanto non avvenga attraverso le indagini di polizia postale su indirizzi IP e *mac address* e potrebbe avere effetti deterrenti rispetto alla produzione e diffusione di disinformazione e contenuti offensivi[120]. Come si è osservato, «si tratterebbe ovviamente di una transizione irta di

[117] M. Cuniberti, *Democrazie, dissenso politico e tutela dell'anonimato*, in *AP*, 2014, 2. Cfr. F. De Simone, *'Fake news', 'post truth', 'hate speech': nuovi fenomeni sociali alla prova del diritto penale*, in *AP*, 2018, 1.

[118] Così G. M. Riccio, *Diritto all'anonimato e responsabilità civile del* provider, in L. Nivarra, V. Ricciuto (a cura di), *Internet e il diritto dei privati. Persona e proprietà intellettuale nelle reti telematiche*, Torino 2002.

[119] G. Resta, *Anonimato, responsabilità, identificazione: prospettive di diritto comparato*, in *DI*, 2014, 2, 204-205. L'Autore osserva altresì che «l'anonimato non è sempre sinonimo di redistribuzione del potere sociale, salvaguardia del dissenso politico, sfida alle costrizioni poste dai vincoli sociali e dalle condizioni di contesto. Esso può anche costituire, per via dell'assottigliamento delle norme sociali che governano il discorso nominativo, uno strumento di diffamazione a basso costo, *harassment* sessuale, incitazione all'odio razziale e ideologico». *Ivi*, 184.

[120] M. Fumo, *Bufale elettroniche, repressione penale e democrazia*, in *ML*, 2018, 1, 87.

difficoltà complicata dalla presenza di autorità di controllo rispondenti a governi illiberali, ma è in ogni caso un passaggio su cui interrogarsi nel momento in cui il mondo digitale è ormai diventato un'estensione di quello reale (ricreandone da ultimo anche la componente sociale) e, conseguentemente, non pare illogico estendere anche ad esso quelle forme di controllo sulle persone già riconosciute e accettate *offline*»[121].

Venendo agli algoritmi fondati sulla profilazione degli utenti, che presiedono le logiche di visibilità dei contenuti in rete, essi concorrono, come illustrato, alla distorsione dell'opinione pubblica, alla viralizzazione dei contenuti sensazionalistici, quali i contenuti di odio e sessualmente espliciti, alla conferma dei pregiudizi e alla radicalizzazione di gruppi estremisti[122]. A coloro che li ritengono parte integrante del modello di business dei *provider* e, in particolare, delle piattaforme *social*, riconducendoli alla libera iniziativa economica privata, vi è chi ribatte che tale libertà vada «armonizzata con l'"utilità sociale", e ponendo come suoi limiti invalicabili la tutela della "sicurezza", della "libertà" e della "dignità umana"»[123]. In tal senso, «il ricorso all'algoritmo non può divenire una forma di deresponsabilizzazione dei soggetti che lo adoperano»[124].

Pertanto, da un parte, vi è chi propone un impiego virtuoso degli algoritmi, per compensarne la portata "usurpatrice", prospettandone ad esempio l'utilizzo per individuare contenuti offensivi, verificarli, segnalarli o direttamente rimuoverli in maniera automatizzata[125]. Un'idea interessante, ma che andrebbe sondata con cautela,

---

[121] A. Mantelero, *La responsabilità* on-line*: il controllo nella prospettiva dell'impresa*, in *DI,* 2010, 3, 421.

[122] «Quello in analisi è, cioè, un processo economico che, assorbendo a fini pubblicitari in piattaforme private il dialogo e l'espressione libera dei cittadini produce una mercificazione di ogni interazione comunicativa: tanto quella di elevato valore sociale, quanto quella dell'insignificante e quotidiano chiacchiericcio, quanto, infine quella che genera lesione delle sfere giuridiche altrui». M. Montanari, *op. cit.*, 262.

[123] M. Cuniberti, *L'art. 21*, cit., 46-47.

[124] S. Rodotà, *op. cit.*, 403. Il filoso argentino Benasayag ha coniato l'espressione «governamentalità algoritmica» per indicare che «la vita degli individui e delle società è orientata e strutturata da macchine». Lo studioso spiega che «gli algoritmi funzionano a partire da quelle micro-informazioni raccolte in massa nel mondo digitale che, messe insieme e correlate tra loro, determinano i profili (…). Il modello epistemologico su cui si fonda questa governamentalità è quello dell'essere umano modulare, simile a quello posto dalla biologia molecolare. Si riduce l'individuo non alle sue decisioni ma ai suoi micro-comportamenti. La biologia non si interessa alla vita stessa, ma alle particelle elementari che si associano tra loro. Nel caso della governamentalità algoritmica è la stessa cosa: si inserisce nella macchina una quantità enorme di micro-atti di diverso tipo – dove la persona passa le sue vacanze, quali percorsi segue, quali siti consulta su internet, chi chiama, ecc. – in base ai quali è possibile stabilire profili virtuali. (…) La governamentalità, quindi, non si rivolge più alla persona (…). Tutti i dati sono pre-individuali, perché le persone sono assimilate tra loro e poi assimilate ai loro profili». «Gli adepti dell'intelligenza artificiale sono convinti che l'individuo sia una macchina, un ricettacolo di diversi micro-comportamenti, senza un senso complessivo. Non ci sono più né individui, né popolazioni, né comunità: esistono profili e avatar virtuali». M. Benasayag, *La Tyrannie des algorithmes*, Paris 2019, trad. it. *La tirannia dell'algoritmo*, Milano 2020, 81-84. Cfr. M. Mezza, *Il contagio dell'algoritmo. Le Idi di marzo della pandemia*, Roma 2020.

[125] G. Ziccardi, *op. cit.*, 87-88. Osserva l'Autore: «l'aspetto tecnico assume, quando si discute di odio *online*, un ruolo essenziale», «conoscere l'architettura della rete è fondamentale per capire i percorsi dei messaggi,

in quanto il rischio di arbitrarietà insito nell'utilizzo di algoritmi è emerso, ad esempio, con riferimento agli algoritmi predittivi in materia di imputabilità e di pericolosità sociale [126]. In particolare, elementi che potrebbero rendere difficile, in quanto connotata da esiti incerti, tale funzione dell'algoritmo sono la complessità del linguaggio umano, con le sue infinite sfumature[127], e il possibile impiego di strategie *ad hoc*, o addirittura di analoghe tecnologie, in grado di aggirare il filtro algoritmico.

Dall'altra, ci si potrebbe interrogare sulla sostenibilità economica per i gestori del superamento dei meccanismi di indicizzazione fondati sulla profilazione degli utenti e sul carattere sensazionalistico dei contenuti [128]. Se, infatti, vista la mole di informazioni che circola in rete, è indubbiamente agevole per i naviganti venire a contatto con contenuti e pubblicità che incontrano i loro interessi e che riflettono le preferenze che hanno già espresso, è altrettanto evidente come tale sistema possa contribuire alla viralizzazione di contenuti offensivi nonché all'infantilizzazione dell'utente e alla sua chiusura in bolle autoreferenziali, con conseguenti effetti distorsivi e rischio di radicalizzazione di gruppi di *hater* e complottisti. La neutralizzazione di tali meccanismi ci restituirebbe forse una rete più pluralistica e quindi più democratica[129]. In alternativa, per superare i molti *bias* che, come emerso

---

eventuali questioni di giurisdizione, le potenzialità lesive, l'influenza dei trend o dell'attualità sul danno arrecato, le possibilità di intervento in tempo reale o in un secondo momento, la capacità di rimozione e, quindi, le tecniche di difesa». In tal senso, «un aspetto tecnico interessante riguarda la capacità della rete di autoregolamentarsi e la possibilità dei *provider* di pensare a soluzioni automatizzate, o parzialmente automatizzate, per intervenire sui contenuti che transitano».

[126] Cfr., *ex pluribus*, M. Bertolino, *Problematiche neuroscientifiche tra fallacie cognitive e prove di imputabilità e di pericolosità sociale*, in *DPP*, 2020, 1, 40 ss. e L. Maldonato, *Algoritmi predittivi e discrezionalità del giudice: una nuova sfida per la giustizia penale*, in *DPC, Riv. trim.*, 2019, 2, 401 ss.

[127] F. Antinucci, *L'algoritmo al potere. Vita quotidiana ai tempi di* google, Bari 2009, 43.

[128] Sul punto M. Cuniberti, *L'art. 21*, cit., 54 ss.

[129] La proposta di Regolamento relativo a un mercato unico dei servizi digitali cui si è già fatto riferimento, al considerando 62, partendo dalla consapevolezza che «un elemento essenziale dell'attività di una piattaforma *online* di dimensioni molto grandi consiste nel modo in cui le informazioni sono messe in ordine di priorità e presentate nella sua interfaccia *online* per facilitare e ottimizzare l'accesso alle stesse da parte dei destinatari del servizio (...), suggerendo, classificando e mettendo in ordine di priorità le informazioni in base ad algoritmi», sottolinea come tale sistema svolga «un ruolo importante nell'amplificazione di determinati messaggi, nella diffusione virale delle informazioni e nella sollecitazione del comportamento *online*». «Le piattaforme *online* di dimensioni molto grandi dovrebbero pertanto provvedere affinché i destinatari siano adeguatamente informati e possano influenzare le informazioni che vengono loro presentate. Esse dovrebbero indicare chiaramente i principali parametri di tali sistemi di raccomandazione in modo facilmente comprensibile per far sì che i destinatari comprendano la modalità con cui le informazioni loro presentate vengono messe in ordine di priorità. Esse dovrebbero inoltre adoperarsi affinché i destinatari dispongano di opzioni alternative per i principali parametri, comprese opzioni non basate sulla profilazione del destinatario». A integrazione di tale proposta normativa, il Garante europeo per la protezione dei dati personali raccomanda che sia prescritto alle piattaforme di prevedere, quale "impostazione predefinita", l'organizzazione dei contenuti *non* basata sulla profilazione e di consentire all'utente di personalizzare i criteri di indicizzazione dei post. European Data Protection Supervisor,

da recenti studi[130], inficiano gli algoritmi, nella letteratura sociologico-digitale si propone di «sottoporre queste tecnologie alla base delle piattaforme digitali a un sistema di *auditing*, di consulenza, che non ne valuti solo l'efficacia funzionale, ma anche le conseguenze sociali del loro funzionamento»[131].

Ancora, per contrastare la velocità delle interazioni in rete, quale variabile usurpatrice che porta a un agire poco ponderato, si potrebbe prevedere l'introduzione di filtri (eventualmente basati su un sistema di parole chiave) che impongano all'utente una pausa per riflettere sull'opportunità dei contenuti che sta pubblicando o diffondendo, in modo da indurlo a un *quid pluris* di riflessione[132].

Quello che è certo è che solo un approccio interdisciplinare, che tenga in considerazioni i diversi interessi in gioco e le caratteristiche tecniche della rete, riflettendo, nella sua complessità, quella delle problematiche cui si è fatto cenno, potrà offrire realistiche prospettive di efficacia e di giustizia[133].

Opinion 1-2021 on the Proposal for a Digital Services Act, 10 febbraio 2021, *https://sites.les.univr.it/cybercrime/wp-content/uploads/2020/05/21-02-100pinion_on_digital_ services_act _en.pdf, 17*.

[130] Cfr. S. U. Noble, *Algorithms of Oppression: How Search Engines Reinforce Racism*, New York 2018; B. Easter, *'Feminist brevity in light of masculine long-windedness': Code, space, and online misogyny*, in *Feminist Media Studies*, 2018, 4, 675 ss.; A. Massanari, *#Gamergate and the Fappening: How Reddit's algorithm, governance, and culture support toxic technocultures*, in *New Media and Society*, 2017, 19(3), 329 ss.; C. Miller, *McInnes, Molyneux, and 4chan: Investigating pathways to the alt-right. Southern Poverty Law Center*, su *https://www.splcenter.org/20180419/mcinnes-molyneux-and-4chan-investigating-pathwaysalt-right*, 19 aprile 2018.

[131] D. Bennato, *Se (anche) l'algoritmo è sessista: ecco perché Instagram preferisce la pelle femminile nuda*, 14 settembre 2020, *www.agendadigitale.eu*.

[132] A. Visconti, *Alcune considerazioni criminologiche e politico-criminali sulle cd. 'fake-news'*, in *Jus*, 2020, 68-69.

[133] È di fatto percepita in dottrina la necessità di «nuove e articolate discipline giuridiche relative agli obblighi (e alle responsabilità) incombenti sugli *ISP*, parametrate all'importanza, al valore economico e all'estensione dei servizi e delle attività che svolgono, da rispettare in termini generali, non solo a seguito di singole specifiche richieste delle Autorità, di fronte a già avvenute violazioni concrete, come oggi sostanzialmente avviene. Per presidiare effettivamente i diritti e gli interessi giuridici di più elevata importanza è ovviamente indispensabile anche il ricorso a sanzioni (non solo penali) proporzionate, efficaci e dissuasive, che sia gli Stati, sia gli organismi e le giurisdizioni sovranazionali (...) possano irrogare o quantomeno richiedere». Cfr. L. Picotti, *Diritto penale e tecnologie informatiche*, cit., 89.