

C J N

# Diritto Penale Contemporaneo

RIVISTA TRIMESTRALE

REVISTA TRIMESTRAL DE DERECHO PENAL  
A QUARTERLY REVIEW FOR CRIMINAL JUSTICE



4/2021

**LA DISINFORMAZIONE TRA POLITICA E DIRITTO.  
DIMENSIONE ISTITUZIONALE, STRATEGIE PREVENTIVE  
E DINAMICHE PUNITIVE**

*LA DESINFORMACIÓN ENTRE LA POLÍTICA Y EL DERECHO.  
DIMENSIÓN INSTITUCIONAL, ESTRATEGIAS PREVENTIVAS  
Y DINÁMICAS REPRESIVAS*

*DISINFORMATION BETWEEN POLITICS AND LAW.  
INSTITUTIONAL DIMENSION, PREVENTIVE STRATEGIES  
AND PUNITIVE TRENDS*

Il presente contributo, qui pubblicato all'interno di una sezione speciale, costituisce la sezione giuridica della ricerca dal titolo "Come individuare e contrastare operazioni coordinate di disinformazione in Italia. Casi di studio e indicazioni di policy per istituzioni pubbliche e private", condotta nell'A.A. 2020/2021 da ricercatori dell'Università Luiss Guido Carli, della Harvard Kennedy School e della School of Information dell'Università del Michigan. La ricerca è stata realizzata con un contributo dell'Unità di Analisi, Programmazione e Documentazione Storica del Ministero italiano degli Affari Esteri e della Cooperazione Internazionale (MAECI), ai sensi dell'art. 23-bis del d.p.r. n. 18 del 5 gennaio 1967. Le riflessioni contenute in questa ricerca riflettono esclusivamente la visione degli autori e non sono necessariamente rappresentative dell'opinione del MAECI e delle altre istituzioni di ricerca coinvolte. Si ringraziano l'Unità di Analisi, Programmazione e Documentazione Storica del MAECI e gli altri Direttori della ricerca (Irene Pasquetto, Gianni Riotta e Costanza Sciubba Caniglia) per avere consentito la pubblicazione degli scritti in questa sede.

# Punire la disinformazione: il ruolo del diritto penale e delle misure di moderazione dei contenuti delle piattaforme tra pubblico e privato

*Reprimir la desinformación: el uso del derecho penal y las iniciativas de moderación de contenidos de las plataformas entre el sector público y el privado*

*Punishing Disinformation: The Role of Criminal Law and Platforms' Content Moderation Between Public and Private Sectors*

EMANUELE BIRRITTERI

*Assegnista di ricerca in Diritto penale presso l'Università Luiss "Guido Carli"*  
*ebirritteri@luiss.it*

LIBERTÀ DI ESPRESSIONE,  
REATI INFORMATICI, DIRITTO UE,  
DIRITTO PENALE COMPARATO

LIBERTAD DE EXPRESSION,  
DELITOS INFORMÁTICOS, DERECHO UE,  
DIRITTO PENALE COMPARATO

FREEDOM OF EXPRESSION,  
CYBERCRIMES, EU LAW,  
COMPARATIVE CRIMINAL LAW

## ABSTRACTS

Il contributo analizza il tema dell'applicazione di misure sanzionatorie nei confronti di azioni (coordinate e non) di disinformazione, mettendo in rilievo punti di forza e limiti dei meccanismi adottati in tale ambito nei settori pubblico e privato.

La prima parte dello studio, volta a delineare il quadro concettuale di riferimento, si sofferma sulle dinamiche socio-criminologiche delle diverse condotte di diffusione di notizie false online, evidenziando ulteriori tematiche di carattere generale – dai rapporti tra tutela della verità e diritto penale, al rilievo in materia delle attività delle piattaforme private di moderazione dei contenuti immessi in rete dagli utenti – funzionali a una migliore comprensione delle problematiche giuridiche in gioco, nonché delle strategie di contrasto al fenomeno implementate o proposte al riguardo in ambito europolitano, interno e comparatistico, e analizzate nel dettaglio nelle sezioni seconda e terza del lavoro.

L'indagine si conclude con alcune indicazioni di *policy* rivolte ad attori pubblici e privati impegnati nel contrasto alla disinformazione, rilevando come gli strumenti di repressione della strumentalizzazione della rete per la diffusione di informazioni false debbano basarsi su un adeguato coinvolgimento proattivo degli operatori privati, ma nell'ambito di una cornice legislativa pubblicistica che definisca con chiarezza le regole del gioco, assicurando, anche tramite il monitoraggio di autorità pubbliche indipendenti, il rispetto della libertà di espressione dell'individuo e dei minimali principi di garanzia di ciascun paradigma punitivo in grado di incidere, pur nell'ambito di forme di *private enforcement*, su diritti fondamentali.

El artículo analiza la cuestión de la aplicación de sanciones contra las iniciativas de desinformación (coordinadas y no coordinadas), evidenciando los puntos fuertes y las carencias de los mecanismos adoptados en este ámbito en los sectores público y privado.

La primera parte del estudio, destinada a delinear el cuadro conceptual, se ocupa de las dinámicas socio-criminológicas de las diversas conductas de difusión de noticias falsas online, señalando cuestiones adicionales de carácter general – desde la relación entre la protección de la verdad y el derecho penal, hasta la importancia de las actividades de las plataformas privadas para la moderación de los contenidos publicados online por los usuarios – funcionales a una mejor comprensión de las cuestiones jurídicas en juego, así como de las estrategias de lucha contra el fenómeno implementadas o propuestas al respecto en el ámbito europeo, nacional y comparado, y

analizadas en profundidad en las secciones segunda y tercera del trabajo.

El estudio concluye con algunas indicaciones políticas para los actores públicos y privados implicados en la lucha contra la desinformación, señalando que las instrumentos para reprimir la explotación de la web para la difusión de información falsa deben basarse en una correcta participación proactiva de los operadores privados, pero dentro de un cuadro legislativo público que definiera claramente las reglas del juego, asegurando, incluso a partir del control de autoridades públicas independientes, el respeto a la libertad de expresión del individuo y los principios mínimos de garantía de cada paradigma punitivo que pueda incidir, incluso en formas de aplicación privada, en los derechos fundamentales.

---

The paper aims at analysing the enforcement of punitive measures against (coordinated and non-coordinated) actions of misinformation, highlighting strengths and pitfalls of the tools used in this field in public and private sectors.

The first part, outlining the conceptual framework, deals with the criminological dynamics of the various forms of dissemination of untrue news online, highlighting other general issues – from the relationship between the protection of truth and criminal law, to the importance of the activities of private platforms moderating the contents posted online by users – for a better understanding of the legal issues at stake, as well as the strategies implemented or proposed to counter the phenomenon in the European, domestic and comparative contexts (analysed in detail in the second and third sections of the paper).

The final part of the work provides some policy recommendations to public and private actors involved in the fight against disinformation, outlining that the tools to combat the exploitation of the web for the dissemination of false information must be based on an adequate proactive involvement of private operators, but within a public legislative framework that clearly defines the “rules of the game”, ensuring, including through the monitoring of independent public agencies, respect for the freedom of expression of the individual and the minimum guarantees of each punitive model that may affect fundamental rights, even within the context of activities of private enforcement.

## SOMMARIO

I. La punizione della disinformazione tra pubblico e privato: inquadramento del fenomeno. – I.1. Premessa. Informazioni false e condotte illecite online legate all'uso della "parola": alcune note introduttive. – I.2. Tutela della verità e diritto penale: la disinformazione e l'intervento della pena criminale. – I.3. Le "sanzioni" irrogate dalle piattaforme private nell'ambito dell'attività di moderazione dei contenuti. – II. Unione Europea. – II.1. Aspetti di rilievo (sul versante punitivo) delle iniziative eurounitarie. – II.2. Le prospettive future: il *Digital Services Act*. – III. Ordinamenti nazionali e misure di moderazione dei contenuti delle piattaforme private. – III.1. La disciplina italiana: profili di rilevanza penale delle condotte di disinformazione e analisi di alcune proposte di riforma. – III.2. Le "sanzioni" irrogate dalle piattaforme private nell'ambito dell'attività di moderazione dei contenuti: l'esempio di *Facebook* e *Twitter*. – III.3. Il contrasto alla disinformazione nel panorama comparatistico (Francia, Germania, Stati Uniti): alcuni spunti di riflessione. – III.4. Indicazioni di *policy* tra pubblico e privato.

# I. La punizione della disinformazione tra pubblico e privato: inquadramento del fenomeno.

## I.1. *Premessa. Informazioni false e condotte illecite online legate all'uso della "parola": alcune note introduttive.*

Nei capitoli precedenti del presente studio ci si è soffermati sul concetto di disinformazione, avuto riguardo sia alla fenomenologia di tali pratiche allarmanti e alle politiche implementate nello scenario internazionale per prevenire e contrastare simili azioni<sup>1</sup>, sia al tema della responsabilità dei *provider*<sup>2</sup>.

Oggetto di quest'ultima parte dell'indagine è l'analisi degli strumenti sanzionatori attivabili in ambito pubblico e privato quali contromisure rispetto alle condotte (coordinate e non) di *disinformation*.

L'indagine verrà svolta lungo due direttrici teleologiche tra loro necessariamente intrecciate: da un lato, si ricostruirà lo stato dell'arte della "materia", verificando quali siano le principali fattispecie criminose e le misure sanzionatorie private *attualmente* applicabili per contrastare le condotte in parola; dall'altro lato, in prospettiva evolutiva e nell'ottica del suggerimento di alcune indicazioni di *policy* tra pubblico e privato, si intende sottoporre a verifica l'opportunità di strutturare in tale ambito, *pro futuro*, ulteriori e/o diversi meccanismi sanzionatori rispetto a quelli oggi innescabili, interrogandosi su quali siano le migliori tecniche di regolazione punitiva del fenomeno.

Prima di addentrarci in tale disamina, tuttavia, è necessario svolgere alcune considerazioni di carattere socio-criminologico sulle pratiche oggetto di attenzione nell'ambito di questa ricerca<sup>3</sup>.

Come noto, quando ci si riferisce alle condotte *online* di disinformazione, e in particolare alla diffusione in rete delle c.d. *fake news*, ci si imbatte in un fenomeno complesso e suscettibile di assumere in concreto molteplici sfaccettature.

Anzitutto, invero, in letteratura non si registra unanimità di vedute sulla stessa nozione di *fake news*<sup>4</sup>: se, secondo alcuni, può in sostanza definirsi tale la condotta di diffusione di un messaggio in rete che, per il suo contenuto o per il modo in cui esso viene presentato, può dirsi oggettivamente falso – ma solo a condizione che chi lo veicola sia consapevole di tale falsità e agisca animato dalla volontà di perseguire specifiche finalità<sup>5</sup> –, secondo altri, invece, in tale nozione rientrerebbe pure la comunicazione falsa in sé, a prescindere dallo scopo preso di mira dall'agente e anche laddove esso non sia in realtà consapevole del carattere falso della notizia

<sup>1</sup> V. *supra*, Cap. I.

<sup>2</sup> V. *supra*, Cap. II.

<sup>3</sup> Diffusamente sul punto v. la recente indagine monografica di GUERINI (2020a), pp. 20 ss. Per un inquadramento di tali dinamiche si veda anche VISCONTI (2020), pp. 43 ss.

<sup>4</sup> Sull'assenza di una nozione universalmente condivisa di *fake news* v. anche PARK, YOUM (2019), pp. 102 ss.

<sup>5</sup> V. al riguardo, per una ricostruzione, PIZZETTI (2017), p. 49. Si veda anche MATUCCI (2018), p. 9, secondo cui dovrebbe trattarsi di scopi di carattere manipolatorio.



(che “colposamente” contribuisce a far circolare)<sup>6</sup>.

Del resto, la complessità del fenomeno della disinformazione online è legata anche al suo carattere multiforme, essendo molto diversi l’atteggiamento psicologico, le motivazioni e il *background* socio-culturale dei soggetti che, in diversi ambiti, diffondono false notizie in rete.

A volte, infatti, determinati individui, spesso attraverso specifiche “pagine” nell’ambito di *social network*, diffondono notizie false per finalità di mero profitto economico (si pensi, ad esempio, al noto fenomeno del c.d. *clickbait*, ove la notizia falsa funge da “esca” per generare rendite pubblicitarie legate al numero di accessi al sito il cui link, contenente il titolo della notizia falsa, viene condiviso sulla piattaforma)<sup>7</sup>.

In altri casi, invece, le false notizie vengono (dolosamente) condivise e diffuse in rete per finalità, potremmo dire, di carattere ideologico. Lo scopo dell’agente, in tali casi, è di solito quello di orientare e influenzare l’opinione pubblica, spingendola, anche mediante tali azioni fraudolente, ad aderire ad un certo punto di vista o a collocarsi su specifiche “posizioni” rispetto a tematiche sensibili sul piano socio-politico (ad es. il sostegno a un certo partito o candidato, la propria visione su temi legati ai diritti civili, all’immigrazione, alle campagne vaccinali etc.)<sup>8</sup>.

Condotte, queste ultime, che possono essere frutto dell’agire del singolo o dell’azione organizzata di gruppi o potenze straniere interessati, nell’ambito di complicate strategie geopolitiche, ad influenzare il dibattito pubblico e financo l’esito delle elezioni di altri Paesi. Tali collettivi possono peraltro servirsi, oltre che di raffinate metodologie di manipolazione del consenso, di complesse tecniche informatiche (quali l’utilizzo di *bot* automatici, l’interazione coordinata di *fake account* etc.) volte ad aumentare artificiosamente le interazioni sulla notizia falsa e, quindi, la sua visibilità sui *social network*<sup>9</sup>.

In entrambi i casi, peraltro, la notizia falsa, sia essa il frutto dell’azione autonoma di un singolo o di una campagna sistematica di disinformazione, è destinata a circolare rapidamente in rete, anche in considerazione del fatto che essa riguarda sovente temi – come la politica, i flussi migratori, le campagne vaccinali etc. – su cui il confronto sociale tende a polarizzarsi, generandosi quindi centinaia di migliaia di *like*, commenti e condivisioni di contenuti che agevolano la circolazione della notizia falsa, con il relativo effetto di “inquinamento” del dibattito pubblico<sup>10</sup>.

Inoltre, nell’epoca della c.d. *post-truth*, in cui nell’approcciarsi ai predetti temi a contare sono spesso più le emozioni e le sensazioni del singolo che i fatti nella loro oggettività<sup>11</sup>, gli utenti – specie di fronte a notizie che confermano il proprio punto di vista “polarizzato” su una determinata questione sociale o politica – sono particolarmente propensi, secondo un atteggiamento psicologico noto come *confirmation bias*, a diffondere il contenuto oggettivamente falso creato dai predetti individui o gruppi organizzati, pur nella convinzione soggettiva della sua veridicità, così contribuendo alla ulteriore circolazione in rete della notizia falsa<sup>12</sup>.

Meccanismi, questi ultimi, ulteriormente alimentati dallo stesso funzionamento tecnologico degli algoritmi delle piattaforme di interazione sociale, che, analizzando la precedente attività online del soggetto, tendono a proporre all’utente, nella sua *news feed*, contenuti di volta in volta sempre più vicini o “affini” al proprio punto di vista o alle proprie preferenze personali, siano essi generati da altri contatti della rete o da pagine *social*<sup>13</sup>. Del resto, tanto maggiore è la permanenza in rete dell’utente, tanto più aumenterà il profitto della piattaforma, che ha dunque tutto l’interesse a costruire un’esperienza individuale che sia il più possibile *tailor-made*<sup>14</sup>.

Ciò, tuttavia, come da più parti è stato evidenziato, alimenta il fenomeno delle c.d.

<sup>6</sup> Cfr., ad es., PERRONE (2018), p. 5. In una più ampia prospettiva, MATTUCI (2018), cit., p. 11, peraltro, ricorda al riguardo anche la studiata differenza tra «...disinformazione e misinformazione. La prima è l’attività di chi scientemente costruisce una storia falsa per raggiungere uno scopo. La seconda è l’attività di chi, ad esempio, condivide o *re-twit*ta una notizia senza la consapevolezza dolosa di manipolare l’opinione pubblica: è un comportamento, per così dire, “leggero”. Per una efficace tassonomia sul punto v. anche GUERINI (2020b), pp. 11 ss.

<sup>7</sup> Sul fenomeno del *clickbait* in tale ambito v. anche COSTANTINI (2019), p. 63.

<sup>8</sup> In argomento cfr. diffusamente anche MATTUCI (2018), cit., p. 10.

<sup>9</sup> Cfr. anche BASSINI, VIGEVANI (2017), p. 16.

<sup>10</sup> V., per ulteriori riferimenti, anche GUERINI (2020a), cit., pp. 38 ss.

<sup>11</sup> In argomento v., *ex multis*, anche DE SIMONE (2018), pp. 4-5.

<sup>12</sup> V. POLLICINO (2018), p. 81. Per un’analisi dei profili sociologici del fenomeno v. anche CAMPAGNOLI (2020). Va del resto considerato che, come giustamente si è osservato in dottrina, «...è difficile aspettarsi dal disomogeneo pubblico dei social media quella correttezza e senso di responsabilità che di regola si pretende dal giornalista [...]»: così SCAROINA (2018), p. 6. V. anche FUMO, p. 84, il quale rileva come la frattura creatasi «...tra democrazia dell’informazione (resa possibile dai *media* e principalmente da internet) e aristocrazia della conoscenza [...]» costituisca «...il terreno di coltura delle *fake news*».

<sup>13</sup> Diffusamente sul punto MARCHETTI (2020), pp. 29 ss. Cfr. anche PITRUZZELLA (2018), pp. 26 ss.

<sup>14</sup> Su tale fenomeno v. anche ZANON (2018), p. 13.

*eco-chambers*, in cui l'utente, all'interno di un sistema "chiuso", si confronta continuamente con le stesse fonti d'informazione o, comunque, con contenuti espressione di punti di vista tra loro sostanzialmente allineati se non addirittura polarizzati su determinate (e sovente erronee) convinzioni<sup>15</sup>. Una condizione, quest'ultima, che appare dunque lontana da quella ben nota idea del mondo dell'informazione quale c.d. *marketplace of ideas* in cui l'individuo, confrontandosi con una pluralità di fonti e punti di vista, viene messo in condizione di formarsi una opinione corretta – e, quindi, anche di scartare notizie false – potendo svolgere le proprie valutazioni sulla base, per così dire, di un *dataset* realmente completo<sup>16</sup>.

Si tratta, insomma, di un fenomeno particolare adatto, per così dire, ad autoalimentarsi, sia esso o meno il frutto dell'azione ideologicamente orientata di singoli o di gruppi collettivi dediti a condurre campagne sistematiche di disinformazione.

Dalla specifica prospettiva del penalista, peraltro, va tenuto conto del fatto che tali condotte di disinformazione intrecciano sovente le loro sorti con altri comportamenti che, a differenza di quello che come vedremo accade rispetto alle notizie false, assumono contorni maggiormente nitidi di illiceità e che pure hanno trovato in internet uno strumento di notevole e a volte preoccupante amplificazione. Il riferimento è qui, in particolare, ai c.d. *hate speech*<sup>17</sup> e alle condotte di lesione dell'altrui reputazione, la cui dettagliata analisi esula dallo scopo specifico della presente indagine, ma il riferimento ai quali è utile, dalla nostra prospettiva, in considerazione del fatto che la singola notizia falsa è spesso parte di messaggi con elementi e di contesti più ampi, i cui contenuti o il cui scopo ultimo sono (anche) l'istigazione all'odio o l'attacco nei confronti di determinate categorie o rispetto all'onore di singoli individui (spesso avversari politici).

Ben si comprende, quindi, come in tale materia qualsiasi intervento di contrasto in ambito pubblico o privato si muova in un contesto socio-criminologico le cui dinamiche di non agevole interpretazione complicano non poco il compito dei decisori di implementare *policy* che si facciano carico di regolare il fenomeno anche dalla prospettiva del diritto punitivo.

## 1.2. *Tutela della verità e diritto penale: la disinformazione e l'intervento della pena criminale.*

Interrogarsi sul ruolo del diritto penale nel contrasto alle condotte di disinformazione (di singoli o di gruppi nell'ambito di azioni coordinate) rende necessario concentrarsi su due quesiti principali e tra loro interconnessi: a) se, in base al diritto vigente, possa assumere rilievo penale la sola diffusione di una notizia falsa in quanto tale, a prescindere, cioè, dal venire in rilievo di un danno (o di una semplice potenzialità di danno) a interessi altri rispetto alla mera veridicità della notizia in sé; b) se, in prospettiva evolutiva, sia ipotizzabile ed effettivamente legittimabile la criminalizzazione della condotta lesiva del (solo) interesse alla veridicità della notizia; in altri termini, se la verità (dell'informazione in generale e, dalla nostra prospettiva, delle notizie veicolate online) possa assurgere a bene giuridico autonomamente presidabile dal diritto penale.

Quanto al primo quesito è agevole rilevare come la diffusione di una falsa notizia in quanto tale non assuma rilevanza penale (e più in generale carattere d'illiceità giuridica) di per sé, autonomamente considerata, ma solo nella misura in cui in cui la condotta comunicativa arrechi offesa – a seconda dei casi, nella forma della lesione o della messa in pericolo – a interessi ulteriori e diversi da quello alla mera veridicità dell'informazione/affermazione<sup>18</sup>.

<sup>15</sup> Fenomeno altresì noto come *Filter Bubble*: v. PITRUZZELLA (2018), cit., p. 27, con ampi richiami sul punto anche alla letteratura anglosassone. V. altresì N. ZANON (2018), cit., p. 13.

<sup>16</sup> Su tali aspetti v. anche MATUCCI (2018), cit., pp. 3-4. In merito alla difficoltà, per tali ragioni, di applicare la teoria del *marketplace of ideas* al mondo di internet, v. efficacemente, PITRUZZELLA (2018), cit., p. 36. Circa la necessità di rivisitare la validità di tale metafora si veda anche POLLICINO (2018), cit., p. 79. Sull'origine nell'ambito della giurisprudenza statunitense del concetto di *marketplace of ideas* e sulla correlata idea per cui «...anche le notizie infondate, mendaci o comunque non verificate meriterebbero di circolare, in quanto la loro scarsa attendibilità sarebbe destinata a trovare conferma nell'incapacità di fare presa sull'uditorio [...]», esaltando per converso la qualificazione dell'informazione professionale, v., anche per tutti i relativi riferimenti specie al noto caso *Whitney v. California* (274 U.S. 357, 1927), BASSINI, VIGEVANI (2017), cit., p. 17.

<sup>17</sup> Per un'analisi delle connessioni tra notizie false e discorsi d'odio v. altresì DE SIMONE (2018), cit., pp. 5-6. Per una approfondita disamina del tema dei discorsi d'odio, dal punto di vista giuridico penale, v. la recente indagine monografia di GALLUCCIO (2020).

<sup>18</sup> Con chiarezza, sul punto, anche PERINI (2017), p. 3, la quale rileva come «...Nell'attuale impianto costituzionale, incardinato sul riconoscimento della libertà di manifestazione del pensiero (art. 21 Cost.), la propalazione di un'informazione falsa non assume rilievo penale in quanto

Come meglio vedremo nel prosieguo della trattazione<sup>19</sup>, rispetto, in generale, alle condotte che si declinano nel veicolare delle falsità, tali interessi, a seconda dello specifico ordinamento oggetto d'attenzione, possono consistere, tra i molti, nell'ordine pubblico (con riferimento, ad es. all'istigazione a delinquere), nell'onore individuale (rispetto alla diffamazione), nel patrimonio (si pensi alle falsi informazioni veicolate, si dia il caso, nell'ambito di una truffa), nella protezione dei mercati (avuto riguardo, ad es., ai reati in tema di *market abuse*) etc.

Ciò che qui interessa subito rilevare, però, è che il diritto attualmente vigente non criminalizza in nessun caso la semplice condotta di disinformazione a meno che la comunicazione di una notizia falsa non arrechi pregiudizio a specifici interessi collettivi o individuali, di volta in volta tutelati dalle varie fattispecie incriminatrici che, come si vedrà, in simili contesti possono più di frequente venire in rilievo.

La seconda questione, come anticipato, riguarda invece la possibilità di ipotizzare in ottica futura la criminalizzazione della mera diffusione di notizie false in sé, a prescindere dall'offesa a interessi ulteriori. Si tratta, insomma, di interrogarsi sulla legittimazione della verità quale bene giuridico tutelabile dal diritto penale.

La risposta a questo quesito, peraltro, potrà offrire anche spunti utili a comprendere per quale motivo, *de iure condito*, il diritto penale non si occupi della falsità in quanto tale, ma solo, come detto, quale veicolo di offesa a interessi diversi.

La questione, da leggere necessariamente attraverso le "lenti" dell'art. 21 della Costituzione e della protezione della libertà di espressione, non è di agevole inquadramento.

Già da un punto di vista etico, invero, l'esistenza di un costante dovere morale di verità in capo al singolo, tale da non soffrire alcuna eccezione, è da sempre una questione fortemente controversa e sui cui in filosofia sono state espresse posizioni non univoche<sup>20</sup> (si pensi al noto dilemma della condotta che dovrebbe tenere il soggetto che nasconde in casa sua un individuo inseguito da un assassino, il quale chiede al proprietario dell'immobile se ha incontrato la sua potenziale vittima<sup>21</sup>).

È del resto controversa la stessa possibilità di poter (quantomeno sempre) agevolmente distinguere il falso dal vero, la realtà dei fatti dalle opinioni e dall'approccio soggettivo all'interpretazione di determinate vicende, e ciò specie di fronte ad accadimenti di complesso inquadramento ove il confine tra oggettivo e soggettivo diventa molto sfumato<sup>22</sup>.

Insomma, già da un punto di vista, per così dire, pregiudiziale, la possibilità di criminalizzare la comunicazione di un'affermazione/notizia falsa *tout court* e di imporre quindi per converso, a qualsiasi individuo, un generalizzato dovere di verità, è un'operazione la cui positiva praticabilità è per nulla scontata.

Sul piano più strettamente giuridico, poi, la legittimazione di una simile opzione di politica criminale, come anticipato, non può che essere vagliata alla luce dell'art. 21 della Costituzione.

Nella dottrina costituzionalistica, infatti, come subito diremo si discute sulla possibilità di estendere la garanzia costituzionale relativa alla protezione della libertà di espressione anche alla esternazione di notizie false.

Tralasciamo naturalmente in questa sede le ipotesi in cui la falsa comunicazione dell'individuo leda interessi altri dalla mera veridicità in sé dell'affermazione (abbiamo prima evocato i reati contro l'ordine pubblico, l'onore individuale etc.); in tali casi, invero, anche nella consapevolezza dell'inesistenza in Costituzione di c.d. "diritti tiranni"<sup>23</sup>, la libertà di espressione

*tale*, essendo viceversa decisivo che in ciò si concreti la cifra distintiva di una condotta offensiva di un particolare bene giuridico. In altre parole, al diritto penale non interessa presidiare la verità dell'informazione in sé, ma incidere in senso preventivo rispetto alle potenzialità lesive in grado di sprigionarsi dalla divulgazione o comunicazione dell'informazione falsa *in direzione* dei più diversi oggetti di tutela». In tal senso è del resto unanime la dottrina (non solo) penalistica, tra cui, *ex multis*, v.: COSTANTINI (2019), cit., p. 68; DEL BÒ (2019), p.100; LEHNER (2019), p. 16; MELZI D'ERIL (2017), p. 64; PULITANÒ (2014), p. 90, il quale, rispetto alla dolosa asserzione di fatti non veri, evidenzia che «... un intervento penale potrebbe essere giustificato solo dall'incidenza su interessi "finali", alla cui offesa il falso sia strumentale».

<sup>19</sup> V. *infra* § III.1.

<sup>20</sup> Per una approfondita analisi sul punto, in particolare circa il problema dell'esistenza o meno di un fondamento etico del divieto di mentire, v., per tutti, PADOVANI (2014), pp. 257 ss., il quale evidenzia come non sia stata affatto dimostrata l'esistenza di una legge morale che stabilisca un simile incondizionato e generale divieto. Per una disamina del tema si veda anche DEL BÒ (2019), cit., pp. 93 ss. Sulle relazioni tra morale e dimensione giuridica nella risposta alla domanda sulla possibilità di rendere la verità oggetto di tutela legale v. anche PULITANÒ (2014), cit., p. 86.

<sup>21</sup> In merito a tale dilemma etico v., ancora, anche per tutti i relativi riferimenti, PADOVANI (2014) CIT., p. 271.

<sup>22</sup> Sui controversi confini del concetto di verità v. GULLO (2002), p. 631, il quale peraltro giustamente rileva come spesso in tale materia si registri «...una singola confusione di piani di indagine tra il requisito della verità come connotato intrinseco della libera manifestazione del pensiero e la questione, ben diversa (in quanto dovrebbe attingere il momento soggettivo), dei margini di rilevanza della putatività *ex art. 59 c.p.*». V. anche, sulle difficoltà di definizione del concetto di verità e sulla complessità di esprimere valutazioni assolutiste nell'alternativa tra vero e falso, BASSINI, VIGEVANI (2017), cit., p. 17.

<sup>23</sup> Si veda al riguardo l'insegnamento della Corte Costituzionale, nella nota vicenda ILVA (Corte Cost., sent. 9 aprile 2013, n. 85 in *corteco-*



deve essere bilanciata con altri interessi meritevoli di tutela che possono, pur a certe stringenti condizioni, legittimare la punizione dell'uso della parola che presenti determinati contenuti<sup>24</sup>.

Ci concentreremo, quindi, sulla possibilità di estendere il raggio applicativo – e le correlate garanzie – dell'art. 21 alle affermazioni false in sé, trattandosi naturalmente di una questione la cui risoluzione è logicamente prioritaria rispetto alla risposta al quesito sulla possibilità di criminalizzare la condotta di disinformazione in quanto tale.

Ora, l'indicazione che emerge dall'analisi della letteratura costituzionalistica è che, in sostanza, per stabilire se una condotta comunicativa debba essere protetta ai sensi dell'art. 21 della Cost. non conti tanto il fatto che tale affermazione possa dirsi oggettivamente vera o falsa – ammesso peraltro, come detto, che si sia sempre in grado di orientarsi univocamente rispetto a tale secca alternativa. A risultare decisivo, infatti, è piuttosto l'atteggiamento psicologico del dichiarante, cui andrebbe riconosciuta la garanzia costituzionale anche in caso di «...espressione di fatti obiettivamente errati, qualora in buona fede essi vengano ritenuti veri da parte di chi ne afferma l'esistenza»<sup>25</sup>.

Esula invece dal raggio applicativo dell'art. 21, e non è quindi coperta dalla garanzia costituzionale, l'affermazione di un fatto falso nella consapevolezza soggettiva, da parte dell'agente, della sua non rispondenza al vero – la dichiarazione, in sostanza, oggettivamente e soggettivamente falsa – in quanto la divergenza tra dichiarato e convinzione interna del soggetto non rende quella condotta comunicativa un pensiero “proprio” dell'individuo; invero, a trovare coperta costituzionale in base alla lettera dell'art. 21, si spiega, è la libera esternazione di un “proprio” pensiero, e tale non è quello che scientemente diverge dalla rappresentazione dei fatti interiorizzata dal dichiarante<sup>26</sup>.

Il fatto che tali affermazioni non rientrino nel novero delle condotte comunicative protette dall'art. 21 non significa però – ed è questo un punto essenziale – che esse vadano considerate quale illecite<sup>27</sup> (del resto, si è già detto che *de iure condito* esse non assumono di per sé carattere di rilevanza penale o in generale di illiceità per l'ordinamento).

Ed invero in letteratura è stato chiarito al riguardo che nel sistema di valori delineato dalla Costituzione non è rinvenibile alcun generalizzato dovere esplicito o implicito di verità la cui violazione possa essere, quindi, di per sé legittimamente criminalizzata<sup>28</sup>.

Del resto, si è anche osservato come l'esistenza – quale posizione giuridica ricavabile in via implicita dall'esegesi dell'art. 21 – di un diritto del cittadino non solo di informare, ma anche di “essere informato”, non possa declinarsi nella pretesa *positiva* di ricevere una informazione che abbia un determinato, specifico contenuto o che sia costruita in un determinato modo<sup>29</sup>, la violazione della quale sia quindi eventualmente presidiabile dal diritto penale, ma al più quella di poter accedere a (e entrare in contatto con) ogni mezzo o contenuto di informazione diffuso nel libero dibattito pubblico, senza l'interposizione – nell'ottica, dunque, di una condizione di carattere perlopiù *negativo* – di alcun ostacolo sul punto, in vista della realizzazione di una società pluralista<sup>30</sup>.

*stizionale.it*), secondo cui «...tutti i diritti fondamentali tutelati dalla Costituzione si trovano in rapporto di integrazione reciproca e non è possibile pertanto individuare uno di essi che abbia la prevalenza assoluta sugli altri. La tutela deve essere sempre “sistemica e non frazionata in una serie di norme non coordinate ed in potenziale conflitto tra loro” (sentenza n. 264 del 2012). Se così non fosse, si verificherebbe l'illimitata espansione di uno dei diritti, che diventerebbe “tiranno” nei confronti delle altre situazioni giuridiche riconosciute e protette, che costituiscono, nel loro insieme, espressione della dignità della persona».

<sup>24</sup> V. GALLUCCIO (2020), cit., *passim*. Sul punto si veda, diffusamente, anche PELISSERO (2015), il quale tra l'altro pone in luce il rilievo che assume il contesto politico-sociale di riferimento nell'ambito di tale giudizio di bilanciamento di interessi.

<sup>25</sup> Così ZANON (2018), cit., p. 13. Cfr. altresì CANDIDO (2020), p. 111. Sul punto v. anche SCAROINA (2018), cit., p. 17, in merito alla necessaria tutela delle opinioni anche laddove fondate su fatti “falsi”.

<sup>26</sup> V., con chiarezza, PERRONE (2018), cit., p. 9. Sulla non riconducibilità dell'affermazione subiettivamente falsa alla nozione di espressione di un libero pensiero e, quindi, alla tutela costituzionale, v. altresì MATUCCI (2018), cit., p. 5, la quale in particolare evidenzia che «...la posizione dei costituzionalisti italiani si è espressa in modo pressoché unanime nel senso che il “subiettivamente” falso non è ricompreso nella sfera di tutela di cui all'art. 21 Cost. Ciò non vuol dire, tuttavia, che la manifestazione di tali fatti sia di per sé vietata: vuol dire, appunto, che la loro esternazione non ha copertura costituzionale. La questione è, allora, soltanto se, e quando, si debba (o si possa) fare ricorso alla repressione: la scelta spetta al legislatore». Chiaro al riguardo anche PADOVANI (2014), cit., p. 331, il quale osserva come la libertà di manifestazione del pensiero «...incontra un limite nel fatto che sono da essa escluse le manifestazioni fraudolente e la diffusione di notizie di cui si conosce la falsità, e che quindi per definizione non corrispondono al proprio pensiero». In argomento v. anche DE SIMONE (2018), cit., p. 6.

<sup>27</sup> V. ancora MATUCCI (2018), cit., p. 5, nonché i riferimenti richiamati nella nota precedente. Si veda anche PERRONE (2018), cit., p. 12.

<sup>28</sup> Sul punto, con ampi richiami alla dottrina pubblicistica di riferimento, v. CANDIDO (2020), cit., p. 110, ove si evidenzia l'assenza di una protezione costituzionale dalle affermazioni false in sé.

<sup>29</sup> V. ancora, anche per ulteriori riferimenti, CANDIDO (2020), cit., p. 109, in particolare circa l'impossibilità di ricavare dall'art. 21 un diritto a ricevere informazioni corrette o veritiere. Sul punto v. altresì DE FLAMMINIS (2020), p. 140, e LEHNER (2019), cit., p. 17. Per un'ampia riflessione, in generale, su questi temi si veda anche FOÀ (2020), pp. 248 ss.

<sup>30</sup> Su tale problematica questione, anche per ulteriori riferimenti, si vedano ancora BASSINI, VIGEVANI (2017), cit., p. 18, i quali significativa-

La possibilità di sanzionare la diffusione di notizie false e azioni coordinate di disinformazione, allora, non può che passare dall'applicazione di norme (già esistenti o, eventualmente, da introdurre *ex novo*) a tutela di interessi diversi dalla mera veridicità della notizia in sé, che siano offesi dalla diffusione di tali comunicazioni difformi dal vero.

Sulla adeguatezza dei mezzi sanzionatori oggi applicabili e sulla effettiva praticabilità di opzioni repressive diverse rispetto a quelle esistenti in ottica di riforma (se dal caso, anche alla ricerca di ulteriori beni giuridici lesi dalle azioni di disinformazione in parola), però, torneremo nella parte conclusiva della presente ricerca<sup>31</sup>, essendo dapprima necessario analizzare, nel prosieguo della trattazione, quali illeciti penali possono oggi configurarsi in relazione alle condotte di nostro interesse e i progetti di riforma che hanno interessato la materia negli ultimi anni.

Prima di svolgere tale disamina, però, è necessario effettuare un primo inquadramento anche dell'importante ruolo giocato nel contrasto alle azioni di disinformazione da ciò che definiremo come l'attività di *private enforcement* svolta dalle piattaforme online per moderare i contenuti immessi in rete dagli utenti.

### I.3. *Le “sanzioni” irrogate dalle piattaforme private nell’ambito dell’attività di moderazione dei contenuti.*

Ancor prima dell'attivazione di un eventuale procedimento penale in relazione ai reati configurabili, come vedremo, avuto riguardo alla diffusione di informazioni false in rete, ad assumere fondamentale rilievo nel contrasto alla disinformazione online sono le misure di moderazione dei contenuti applicate dalle stesse piattaforme private.

Dalla nostra prospettiva l'interesse è rivolto, naturalmente, ai principali *social network* (quali soprattutto *Facebook* e *Twitter*), ormai divenuti tra le più importanti arene di confronto e dibattito pubblico e che, per certi versi, hanno consentito ai leader politici – e, in una prospettiva, per così dire, “patologica”, ai singoli e ai gruppi di pressione impegnati in azioni anche coordinate di disinformazione – di instaurare un dialogo diretto con gli elettori e l'opinione pubblica, senza l'ulteriore filtro dei media tradizionali e degli organi di informazione in generale<sup>32</sup>.

Il fatto che simili “territori” privati siano divenuti luoghi in un cui si intrecciano relazioni comunicative di indubbia rilevanza pubblica per la vita sociale e politica di ciascun Paese ha dunque inevitabilmente consegnato un enorme potere nelle mani di tali piattaforme private.

Un potere, a ben vedere, non soltanto di carattere economico, in relazione al sempre più diffuso uso di tali strumenti con i relativi introiti per gli operatori, ma anche, nei fatti, di “regolazione” del dibattito pubblico che si svolge in tali reti sociali.

Infatti, il modo in cui simili operatori costruiscono le condizioni d'uso del servizio e gli standard di comportamento di ciascuna *community* digitale – avuto riguardo, ad esempio, alla diffusione di notizie false o a condotte comunicative di esortazione alla violenza o all'odio – e il funzionamento concreto del correlato *enforcement* per rimuovere i contenuti (o per inibire l'accesso al servizio ad *account* che pongono in essere comportamenti) lesivi di tali “regole del gioco” auto-normate da parte della piattaforma, inevitabilmente vale a costruire una iniziale barriera di contrasto alla disinformazione online<sup>33</sup>.

Si tratta, insomma, nella sostanza delle cose e per quanto in una dimensione privata, della prima linea di intervento contro le notizie false e le azioni (coordinate e non) di disinformazione, anche in considerazione del fatto che tale *enforcement* privato di solito interverrà prima del (o a prescindere dall'esistenza di un) procedimento penale, e secondo tempistiche e modalità decisorie estremamente più brevi e, naturalmente, flessibili di quelle della giurisdizione

mente rilevano come all'idea «...che solo le informazioni che soddisfano un'esigenza informativa meritino circolazione, in quanto funzionali al diritto a essere informati, deve dunque opporsi la ricostruzione, ancora una volta, che vede nell'ecosistema dell'informazione un *marketplace* delle idee in cui la capacità di discernimento dell'utente conduce a sottoporre a un giudizio critico i contenuti informativi».

<sup>31</sup> V. *infra* § III.4.

<sup>32</sup> Efficacemente sul punto, in particolare avuto riguardo al ruolo di *Twitter*, SCAROINA (2018), cit., p. 2, la quale evidenzia come «...i protagonisti dell'attualità hanno ormai eletto questa potenzialmente sconfinata piazza virtuale a strumento di comunicazione privilegiata per illustrare iniziative in materia di giustizia penale, per commentare e comunicare decisioni giudiziarie e iniziative politiche, orientare l'opinione pubblica». Sul tema v. anche CATERINI (2013), pp. 601 ss.

<sup>33</sup> V., per ulteriori riferimenti e richiami, *infra* § III.2.

statale.

Inoltre, le misure sanzionatorie di carattere inibitorio irrogate dalle piattaforme – che a titolo esemplificativo vanno dalla etichettatura o rimozione del contenuto, alla sospensione temporanea per poche ore della possibilità di interagire e condividere contenuti (c.d. modalità d’uso in “solo lettura”), alla sospensione prolungata o definitiva della possibilità per l’*account* autore di violazioni di accedere alla piattaforma etc.<sup>34</sup> – sia per il loro contenuto interdittivo, sia per la rapidità con cui vengono irrogate, possono in molti casi raggiungere un livello di efficacia e deterrenza persino maggiore rispetto a quello delle sanzioni criminali pecuniaria o, nei casi più gravi, detentiva (che arriveranno, se del caso, e ammesso, naturalmente, che il fatto abbia rilevanza penale, dopo anni di complicati processi).

Tutto ciò, beninteso, solleva tutta una serie di problematiche che il giurista non può ignorare e che vanno: dal tema dell’assenza di legittimazione democratica del soggetto (privato) che disciplina le regole di un dibattito che, come detto, assume rilevanza pubblicistica, con tutti i correlati rischi di esercitare forme di censura<sup>35</sup> e di determinare il ben noto c.d. *chilling effect* nell’ambito del libero confronto pubblico; a quello del rispetto, nella costruzione di tali misure sanzionatorie, delle minimali garanzie di qualsivoglia paradigma punitivo anche, per così dire, di carattere privato/disciplinare (ad es. tassatività delle violazioni, divieto di irretrattività etc.); alle garanzie procedurali a tutela del singolo e ai rimedi esperibili rispetto a simili decisioni sulla moderazione dei contenuti, che inevitabilmente finiscono per influire sulla libertà di espressione dell’individuo e anche di soggetti pubblici/politici di primo piano<sup>36</sup> (si pensi, solo per richiamare l’esempio più significativo, alla decisione presa di recente dal Comitato di controllo di *Facebook* sulla sospensione dell’*account* dell’ex Presidente degli Stati Uniti Donald Trump<sup>37</sup>).

Anche su questi temi, allora, avremo modo di soffermarci nell’ambito di tale ricerca, analizzando le *policy* fissate da due tra i più importanti *social network* (*Facebook* e *Twitter*) e rilevanti rispetto al tema delle informazioni false e delle azioni di disinformazione, con speciale attenzione alle misure di *enforcement* e alle sanzioni irrogate da tali piattaforme per le violazioni, da parte dei propri utenti, di tali regole di comportamento e delle condizioni d’uso del servizio offerto.

## II. Unione Europea.

### II.1. *Aspetti di rilievo (sul versante punitivo) delle iniziative eurounitarie.*

Nelle parti precedenti della presente ricerca si sono ampiamente ricostruite le molteplici iniziative intraprese negli ultimi anni in ambito eurounitario per il contrasto alla disinformazione<sup>38</sup>.

In questa sede, lungi dal ripercorrere i contenuti già in precedenza esposti, ci limiteremo ad evidenziare quali elementi di interesse possono desumersi dai documenti e dalle raccomandazioni elaborati in seno all’Unione europea sotto il versante di nostro specifico interesse, relativo alle misure sanzionatorie contro le azioni (coordinate e non) di *disinformation*<sup>39</sup>.

Un primo insegnamento che ci pare possa essere tratto alla luce della prima parte di questo studio è che in nessuno di tali documenti si riscontra un riferimento all’opportunità dell’uso del diritto penale per il contrasto alla diffusione in rete di notizie false<sup>40</sup>. Ciò, ci pare, costituisca ulteriore conferma degli esiti preliminari cui siamo giunti in precedenza circa l’inesistenza,

<sup>34</sup> V. per un’analisi di dettaglio *infra* § III.2.

<sup>35</sup> Su tali rischi, in ambito sia pubblico che privato, v. diffusamente anche PINELLI (2017), pp. 6-7, il quale evidenzia anche i limiti di risposte pubbliche che si muovano in ambito soltanto nazionale, data la natura globale della rete e l’inconveniente della resistenza di Stati “*free rider*” anche a forme di regolazione internazionale.

<sup>36</sup> Torneremo su questi temi *infra* § III.2 e III.4.

<sup>37</sup> Un riepilogo del caso e della decisione adottata dal Comitato può leggersi al [presente link](#).

<sup>38</sup> Cfr. *supra* Cap. I.

<sup>39</sup> Per una disamina delle iniziative europee in tale campo v. pure DE SIMONE (2018), CIT., pp. 6 ss.

<sup>40</sup> In argomento, del resto, la stessa dottrina che ha analizzato nel dettaglio le iniziative eurounitarie in materia evidenzia come «...la tendenza che appare dominante nello scenario europeo della lotta alla disinformazione è quella del *favor* per l’affidamento alle stesse piattaforme digitali del controllo e della censura delle *fake news*, senza interventi pubblici»: così MONTI (2020), p. 295.

già *de iure condito*, di misure repressive da parte dell'ordinamento nei confronti della affermazione falsa in quanto tale e della inopportunità, anche *de iure condendo*, di strutturare reati posti a protezione del bene giuridico della verità dell'informazione *tout court*, dovendosi quindi implementare strategie di contrasto fondate su altre basi<sup>41</sup>.

E invero, ulteriori aspetti di interesse di tali azioni e iniziative europee ci pare attengano, avuto riguardo alla prospettiva sanzionatoria, alla presa d'atto dell'importanza di coinvolgere, nell'azione di contrasto a tali preoccupanti fenomeni, tutti gli *stakeholder* e in particolare le piattaforme private (in tal senso il riferimento è soprattutto al Codice di buone pratiche)<sup>42</sup>.

Ciò, quindi, conferma quanto prima rilevavamo circa il fatto che la moderazione dei contenuti da parte di tali operatori in rete rappresenta il primo livello di contrasto a tali azioni di disinformazione, e come lo studioso del diritto punitivo debba oggi necessariamente occuparsi di fenomeni che, pur svolgendosi in ambito privato, finiscono per delineare un sistema di *enforcement* che, avuto riguardo alla determinazione delle regole di comportamento e all'erogazione di misure sanzionatorie/inibitorie per gli utenti in caso di violazione di tali standard, inevitabilmente influisce in modo a volte anche considerevole sul dibattito pubblico e in generale sul bilanciamento tra libertà di espressione e altri interessi di rilievo costituzionale<sup>43</sup>.

In tale direzione, allora, i documenti analizzati nella prima parte della ricerca offrono indicazioni utili, dalla nostra prospettiva, non solo rispetto ai principi guida cui dovrebbero ispirarsi le *policy* di moderazione dei contenuti implementate dalle piattaforme (che vanno, a titolo esemplificativo, dalla necessità di politiche chiare sulla chiusura degli *account* falsi e sui siti che diffondono pubblicità ingannevoli, alle tecnologie per dare precedenza a notizie autentiche e autorevoli rispetto a quelle false, fino al rilievo delle attività di *fact checking*), ma anche sulle criticità determinate da tali attività di *private enforcement* (dal rischio di legittimare forme di censura privata, alle problematiche in merito al controllo sull'effettiva attuazione da parte delle piattaforme delle misure delineate dal Codice di buone pratiche etc.<sup>44</sup>).

Su tali aspetti, allora, avremo modo di soffermarci nel corso del prosieguo della ricerca<sup>45</sup>. Prima, però, di analizzare i contenuti specifici delle misure pubbliche e private di contrasto alla disinformazione, è necessario tener conto, come subito vedremo, di alcune indicazioni importanti contenute nella proposta di regolamento europeo sui servizi digitali.

## II.2. *Le prospettive future: il Digital Services Act.*

Si è già avuto modo in precedenza di ricostruire genesi e contenuti principali della proposta di regolamento europeo relativo a un mercato unico dei servizi digitali (*Digital Services Act*), data l'importanza che tale proposta di riforma riveste soprattutto in punto di definizione del ruolo dei *provider* nel contrasto alla disinformazione<sup>46</sup>.

In questa sede, allora, ci soffermeremo sui profili di tale proposta che più da vicino interessano gli aspetti sanzionatori rispetto alle condotte di disinformazione online.

La riforma, infatti, riveste particolare importanza, dalla nostra prospettiva, avuto riguardo ai suoi effetti sulle predette attività di *private enforcement* compiute dalle piattaforme online nella misura in cui essa «...fissa standard di trasparenza e responsabilità più rigorosi in relazione alle modalità di moderazione dei contenuti da parte dei fornitori di tali piattaforme, alla pubblicità e ai processi algoritmici. Istituisce obblighi di valutazione dei rischi derivanti dai loro sistemi, allo scopo di sviluppare adeguati strumenti di gestione dei rischi e proteggere l'integrità dei servizi dall'uso di tecniche di manipolazione»<sup>47</sup>.

Oltre a fornire una nozione precisa sia dell'attività di moderazione dei contenuti<sup>48</sup>, sia

<sup>41</sup> V. *supra* § I.2.

<sup>42</sup> Cfr. *supra* Cap. I.

<sup>43</sup> In merito al declinarsi di tali interconnessioni tra libertà di espressione e disinformazione, nonché per una compiuta ricostruzione delle diverse azioni intraprese in ambito internazionale, si veda la disamina di POLLICINO (2020), pp. 175 ss.

<sup>44</sup> Sul rischio, ad es., che in ambito in europeo la disciplina della libertà di informazione venga influenzata dalle logiche del profitto v. LEHNER (2019), cit., p. 14.

<sup>45</sup> V. *supra* § III.1 e ss.

<sup>46</sup> V. *supra* Cap. II.

<sup>47</sup> Cfr. la Relazione introduttiva alla proposta di Regolamento, 3. La relazione e il testo della proposta sono reperibili alla [presente pagina web](#).

<sup>48</sup> L'art. 2, lett. p), della proposta di Regolamento, definisce come "moderazione dei contenuti" «...le attività svolte dai prestatori di servizi intermediari al fine di individuare, identificare e contrastare contenuti illegali e informazioni incompatibili con le loro condizioni generali, forniti dai destinatari del servizio, comprese le misure adottate che incidono sulla disponibilità, sulla visibilità e sull'accessibilità di tali contenuti

delle condizioni generali<sup>49</sup> d'uso del servizio (imponendo in particolare che esse, definiscano i processi e le misure di moderazione secondo un «...linguaggio chiaro e previsto di ambiguità [...]», essendo «...disponibili al pubblico in un formato facilmente accessibile» e nel rispetto, nel corso delle procedure, dei diritti fondamentali dei destinatari<sup>50</sup>), invero, la proposta di Regolamento: istituisce un meccanismo di *notice and action* (art. 14) con regole dettagliate per la gestione delle segnalazioni sulla presenza di contenuti illegali; stabilisce l'obbligo di motivazione (art. 15) in caso di decisione di rimozione dei (o disabilitazione dell'accesso ai) contenuti illegali o lesivi delle condizioni generali d'uso del servizio, con la necessità di precisare l'addebito e di fornire informazioni sui reclami e i mezzi di ricorso esperibili, siano essi interni – rispetto ai quali, peraltro, il Regolamento (art. 17) prevede l'obbligo per la piattaforma di offrire un sistema efficiente di gestione dei reclami rispetto a tutte le misure sanzionatorie applicabili nei confronti dell'utente, dalla rimozione del contenuto fino alla cessazione dell'*account* –, giudiziali o extragiudiziali (questi ultimi dettagliatamente regolati all'art. 18 della proposta); prevede una specifica disciplina per agevolare le segnalazioni attendibili (art. 19) e misure a tutela della trasparenza commerciale e della pubblicità (art. 22 ss.), nonché la necessaria sospensione dell'*account* (art. 20) per i destinatari del servizio che regolarmente forniscono contenuti manifestamente illegali e il diritto di presentazione di un reclamo da parte dei destinatari del servizio (art. 43)<sup>51</sup>.

L'idea, inoltre, è che il regolamento operi parallelamente, sulla base di una strategia integrata, rispetto alle azioni già proposte in ambito euorunitario, in particolare avuto riguardo al Codice di buone pratiche, che andrebbe rivisto e rafforzato anche per integrare le misure del Regolamento<sup>52</sup>.

Va considerato, peraltro, che pur non riferendosi direttamente alle notizie false in sé nella definizione di contenuto illegale<sup>53</sup> (che, del resto, assumono carattere di illiceità solo se offendono interessi altri dalla mera veridicità della notizia), il Regolamento in questione – oltre, come visto, a fornire una precisa base giuridica alle azioni di *enforcement* delle piattaforme nelle attività di moderazione dei contenuti – in diverse sue parti si riferisce alle misure di *compliance* che le piattaforme dovrebbero mettere in campo per agevolare le azioni di contrasto alla disinformazione.

Si pensi, ad esempio, al considerando n. 63 che stabilisce come le piattaforme online di dimensioni molto grandi «...dovrebbero garantire l'accesso del pubblico ai registri della pubblicità visualizzata sulle loro interfacce online per facilitare la vigilanza e la ricerca sui rischi emergenti derivanti dalla distribuzione della pubblicità online, ad esempio in relazione alla pubblicità illegale o alle tecniche di manipolazione e alla disinformazione che hanno ripercussioni negative reali e prevedibili sulla salute pubblica, sulla sicurezza pubblica, sul dibattito civico, sulla partecipazione politica e sull'uguaglianza».

D'interesse sul punto, poi, è anche il considerando n. 68, ove nel ribadire, in sostanza, la differenziazione tra contenuti meramente falsi e illegali, si rileva come un aspetto da prendere in considerazione nella valutazione e gestione del rischio da parte delle piattaforme attenga «...agli eventuali effetti negativi dei rischi sistemici sulle società e sulla democrazia, quali la disinformazione o le attività di manipolazione e abuso», con esplicito richiamo alle azioni anche coordinate di diffusione di informazioni false o fuorvianti. Al riguardo, sempre nel medesimo considerando, la proposta di Regolamento significativamente chiarisce come «...l'adesione a

illegali o di dette informazioni, quali la loro retrocessione o rimozione o la disabilitazione dell'accesso agli stessi, o sulla capacità dei destinatari di fornire tali informazioni, quali la cessazione o la sospensione dell'*account* di un destinatario del servizio».

<sup>49</sup> L'art. 2, lett. q), della proposta di Regolamento, definisce come “condizioni generali” «...tutte le condizioni, le modalità o le specifiche, comunque denominate e indipendentemente dalla loro forma, che disciplinano il rapporto contrattuale tra il prestatore e il destinatario dei servizi intermediari».

<sup>50</sup> Cfr. l'art. 12 della proposta di Regolamento, che, oltre a riferirsi alla necessità di fornire indicazioni sul processo decisionale algoritmico in tale ambito di moderazione dei contenuti e sulla verifica umana di esso da parte della piattaforma, aggiunge tra l'altro al secondo paragrafo come «...i prestatori di servizi intermediari agiscono in modo diligente, obiettivo e proporzionato nell'applicare a far rispettare le restrizioni di cui al paragrafo 1, tenendo debitamente conto dei diritti e degli interessi legittimi di tutte le parti coinvolte, compresi i diritti fondamentali applicabili dei destinatari del servizio sanciti dalla Carta», previsione, quest'ultima, di grande importanza nel definire anche il quadro delle garanzie rispetto a misure che, come abbiamo rilevato, influiscono su libertà di rilievo costituzionale dei cittadini.

<sup>51</sup> Per una sintesi delle principali innovazioni si veda anche la [presente pagina web](#).

<sup>52</sup> V. la Relazione introduttiva alla proposta di Regolamento, cit., 4.

<sup>53</sup> Cfr. il Considerando n. 12 e 68 (su cui cfr. anche *infra*), nonché l'art. 2, lett. g), della proposta di Regolamento, secondo cui costituisce contenuto illegale «...qualsiasi informazione che, di per sé o in relazione ad 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».



un determinato codice di condotta e il suo rispetto da parte di una piattaforma online di dimensioni molto grandi possono essere ritenuti una misura di attenuazione dei rischi adeguata. Il fatto che una piattaforma online rifiuti, senza adeguate spiegazioni, l'invito della Commissione a partecipare all'applicazione di un tale codice di condotta potrebbe essere preso in considerazione, se del caso, nel determinare se la piattaforma online abbia violato gli obblighi stabiliti dal presente regolamento».

Insomma, da un lato il Regolamento si pone nel solco delle misure già intraprese, rendendo il *private enforcement* delle piattaforme, piuttosto che la previsione di misure penali, il vero territorio d'elezione per il contrasto alle azioni di disinformazione<sup>54</sup>; dall'altro, però, stabilisce inevitabilmente una contaminazione di fatto tra *hard* e *soft law*, nella misura in cui rileva come la (ingiustificata) mancata adesione a un Codice di buone pratiche, che per definizione rappresenta un *set* di norme ad adesione meramente volontaria da parte dell'operatore, possa essere valorizzata quale elemento negativo nella valutazione del rispetto da parte della piattaforma degli obblighi del Regolamento – seppur, ci pare si possa evincere dalla lettera del citato considerando, non sia autonomamente sanzionabile.

Un aspetto, quindi, che assumerà un peso anche nella determinazione delle rilevanti sanzioni – basate su efficienti meccanismi ingiunzionali e progressivi, tali da spingere gli operatori ad adeguarsi alle indicazioni delle autorità di *enforcement*<sup>55</sup>, con la possibilità in alcuni casi (art. 56), peraltro, di addivenire a soluzioni negoziate per la definizione del procedimento ispirate, per certi versi, al modello dei *Non Prosecution Agreement* nordamericani – cui, in base a tale proposta, l'operatore potrà essere sottoposto per le violazioni del Regolamento e che rischia dunque, come si è già osservato<sup>56</sup>, di risolversi nella prassi in una obbligatorietà *de facto* per gli operatori dell'adesione ai detti Codici.

Nei confronti, peraltro, delle piattaforme online «di dimensioni molto grandi»<sup>57</sup>, la proposta di Regolamento (art. 26 ss.) stabilisce rilevanti obblighi di *compliance* di autovalutazione e autonormazione per la gestione dei rischi sistemici derivanti dal funzionamento dei loro servizi, tra i quali, per quanto più rileva dalla nostra prospettiva, si richiama «...la manipolazione intenzionale del servizio [...] con ripercussioni negative, effettive o prevedibili, sulla tutela della salute pubblica, dei minori, del dibattito civico, o con effetti reali o prevedibili sui processi elettorali e sulla salute pubblica» (art. 26, par. 1, lett. c).

Spinte anche dalla possibilità di essere sottoposte, in caso di non conformità, al predetto consistente apparato di misure sanzionatorie, allora, tali soggetti saranno obbligati a farsi carico di autovalutare anche il rischio, connesso all'uso dei loro sistemi, legato a tali azioni inquadabili pure nel contesto di operazioni (specie coordinate) di disinformazione e a implementare misure operative, nonché specifici meccanismi di *audit* e di controllo interno, per mitigare la possibilità che si concretizzi un simile uso strumentale delle loro piattaforme.

La logica che permea tale scelta di regolazione è quella di prendere atto che è proprio la stessa organizzazione la cui operatività genera il rischio da mitigare a essere il soggetto nella migliore posizione per contrastarlo<sup>58</sup>. Un approccio, del resto, che permea diverse ulteriori regolazioni – il riferimento è qui soprattutto al GDPR e ai vari sistemi di *corporate criminal liability* – che pure puntano sull'attuazione da parte dell'ente di programmi di *compliance* basati sulla identificazione da parte della stessa persona giuridica, in una logica *tailor-made*, delle

<sup>54</sup> V. anche *supra* § I.3 e *infra* spec. § III.2.

<sup>55</sup> In base agli artt. 38 e ss. della proposta di Regolamento, invero, le responsabilità di *enforcement* della nuova normativa spettano in prima istanza alle autorità nazionali competenti (coordinatori dei servizi digitali) e, in alcuni casi, alla stessa Commissione, e prevedono non solo l'esercizio di rilevanti poteri in termini di imposizione dell'obbligo agli operatori di fornire informazioni e spiegazioni e di effettuare ispezioni e altre azioni di sequestro di informazioni etc., ma anche di ordinare la cessazione di violazioni e imporre misure correttive per far cessare l'inosservanza, nonché di applicare sanzioni pecuniarie per l'inosservanza del regolamento oltre a penalità di mora, disponendo anche misure provvisorie per evitare il rischio di danni gravi. Inoltre, nei casi più gravi, e ove l'operatore non si adegui alle misure previste, l'autorità di *enforcement* potrà anche chiedere alle autorità giudiziarie dello Stato membro di ordinare la restrizione temporanea dell'accesso rispetto ai destinatari del servizio interessati dalla violazione o, solo se ciò non sia fattibile tecnicamente, financo la restrizione dell'accesso all'interfaccia online del prestatore di servizi intermediari sulla quale ha luogo la violazione. In alcuni casi, peraltro, (v. ad es. art. 59 e 60), si possono infliggere sanzioni pecuniarie anche nel caso in cui le piattaforme di dimensioni molto grandi forniscano nel corso del procedimento informazioni inesatte, incomplete o fuorvianti o si rifiutino di rettificarle o di sottoporsi a una ispezione in loco.

<sup>56</sup> V. *supra* Cap. II, § II.3.

<sup>57</sup> Definite all'art. 25 della proposta di Regolamento quali quelle che «...prestano i loro servizi a un numero medio mensile di destinatari attivi del servizio nell'Unione pari o superiore a 45 milioni, calcolato conformemente alla metodologia stabilita negli atti delegati di cui al paragrafo 3».

<sup>58</sup> Su tali scelte di politica criminale e sui relativi rischi v. STELLA (2003), pp. 55 ss., che sul punto evoca una famosa affermazione del tedesco Schunemann, secondo cui in tali casi vi sarebbe il rischio di assegnare alla «capra il compito di giardiniere».

proprie specifiche misure di prevenzione<sup>59</sup> – una tale vicinanza, ad esempio, emerge anche dall'obbligo, previsto dall'art. 32 della proposta di Regolamento per le piattaforme online di dimensioni molto grandi, di nominare uno o più *compliance officer*, quali figure molto simili al DPO in ambito *privacy*, incaricati di monitorare la conformità delle piattaforme al nuovo Regolamento.

La proposta di riforma in questione, allora, costituisce sicuramente un significativo passo avanti nella regolazione pubblica del fenomeno della disinformazione, pur, come vedremo nella parte finale della ricerca<sup>60</sup>, lasciando aperti alcuni significativi interrogativi sia avuto riguardo all'effettivo assetto di garanzie per i soggetti destinatari delle misure di moderazione dei contenuti irrogate dalle piattaforme, sia, più in generale, rispetto ai rischi legati all'affidare alle piattaforme il compito di moderare il dibattito pubblico online.

### III. Ordinali nazionali e misure di moderazione dei contenuti delle piattaforme private.

#### III.1. *La disciplina italiana: profili di rilevanza penale delle condotte di disinformazione e analisi di alcune proposte di riforma.*

Come anticipato<sup>61</sup>, nell'ordinamento italiano la diffusione di una notizia falsa nell'ambito di una azione di disinformazione non acquisisce rilevanza penale (o comunque carattere d'illiceità per l'ordinamento) in sé ed autonomamente considerata, ma soltanto ove la condotta comunicativa arrechi offesa – nella forma, a seconda dei casi, della lesione o della messa in pericolo – ad interessi altri rispetto alla mera veridicità della notizia.

La qualificazione giuridico-penale di ciascuna condotta comunicativa posta in essere online, ad ogni modo, non può che dipendere dallo specifico contenuto dell'affermazione veicolata, e potendo tali condotte di disinformazione assumere, naturalmente, i contenuti più vari, i reati eventualmente configurabili di volta in volta in tali contesti saranno necessariamente molteplici e a volte anche estremamente eterogenei tra loro quanto a struttura della fattispecie, interesse tutelato etc. L'elencazione che segue, allora, va interpretata come una indicazione meramente esemplificativa delle figure criminose che, almeno più di frequente, possono risultare integrate rispetto ad azioni (coordinate e non) di disinformazione in rete<sup>62</sup>.

Un primo gruppo di fattispecie, anzitutto, attiene alle ipotesi in cui l'affermazione falsa sia contenuta nel contesto di un discorso più ampio di carattere discriminatorio nei confronti di certe categorie o di vera e propria istigazione a delinquere o apologia di delitto (si pensi agli artt. 414 e ss., 604-*bis*<sup>63</sup> c.p.). Lo studio di tali norme, come anticipato, elusa dallo scopo specifico di tale indagine: del resto, il requisito della falsità della notizia non costituisce un elemento costitutivo di tali reati, che potranno quindi venire in rilievo anche a prescindere dall'oggettiva

<sup>59</sup> Rispetto a tali politiche v., con ampi riferimenti, la recente disamina di GULLO (2020), pp. 244 ss.

<sup>60</sup> V. infra § III.4.

<sup>61</sup> V. supra § I.2.

<sup>62</sup> Per una efficace panoramica delle figure criminose collocabili nell'ambito di nostro interesse si veda anche COSTANTINI (2019), cit., pp. 60 ss.

<sup>63</sup> L'ultimo comma dell'art. 604-*bis* prevede, peraltro, in relazione al contrasto al c.d. fenomeno del negazionismo, una specifica aggravante laddove «...la propaganda ovvero l'istigazione e l'incitamento, commessi in modo che derivi concreto pericolo di diffusione, si fondano in tutto o in parte sulla negazione, sulla minimizzazione in modo grave o sull'apologia della Shoah o dei crimini di genocidio, dei crimini contro l'umanità e dei crimini di guerra, come definiti dagli articoli 6, 7 e 8 dello statuto della Corte penale internazionale», con una previsione che, come noto, ha sollevato e solleva complicate questioni dogmatiche e di politica criminale, che esulano dallo scopo specifico del lavoro e che non possono qui nemmeno sfiorarsi, legate all'ammissibilità di un simile aggravamento del trattamento sanzionatorio connesso anche a ragioni di tutela, per così dire, della veridicità storica dei fatti – per quanto, naturalmente, anche qui ancora una volta non si tuteli l'astratta verità, di per sé, trattandosi di un'aggravante che deve accedere a un fatto di reato lesivo, come detto, di interessi altri. Per una approfondita analisi della questione si veda, in particolare, FRONZA (2012); nonché ID. (2017), pp. 155 ss. Critico, ad es., sulla prospettiva di identificare in tali ambiti l'oggettività giuridica, ad es., nella memoria delle vittime, evidenziando la necessità di ricercare nella protezione delle istituzioni democratiche dalle aggressioni violente i riferimenti di tutela in tale settore, è INSOLERA (2018), p. 123. Sul tema del rapporto tra pena e c.d. leggi memoriali cfr. altresì CARUSO (2018), pp. 166 ss. In argomento, rispetto allo specifico ambito di nostro interesse, v. anche PERINI (2017), cit., p. 5. Come noto, peraltro, la fattispecie è attualmente interessata da un più ampio e noto progetto di riforma – il d.d.l. S.2005, recante “Misure di prevenzione e contrasto della discriminazione e della violenza per motivi fondati sul sesso, sul genere, sull'orientamento sessuale, sull'identità di genere e sulla disabilità”, c.d. d.d.l. Zan, approvato in prima lettura alla Camera il 4 novembre 2020 e attualmente in discussione al Senato –, per un'analisi del quale si rinvia a FILICE (2020).

falsità o meno del contenuto veicolato.

Tuttavia, il riferimento a tali figure criminose è utile in questa sede per almeno due ragioni su tutte: a) da un punto di vista criminologico, come anticipato, simili condotte di propaganda o istigazione si accompagnano spesso alla diffusione di informazioni false costruite ad arte per istigare all'odio nei confronti di certe categorie sociali e sovente per ragioni legate alla promozione di determinate ideologie nell'ambito dello scontro politico; b) dalla prospettiva della tecnica legislativa e della prassi applicativa, si tratta di figure particolarmente esemplificative delle difficoltà proprie delle figure criminose legate alla punizione, come di recente si è evidenziato, della "parola pericolosa", in virtù del difficile bilanciamento, sia in astratto che in concreto, tra libertà di espressione e altri beni giuridici. In questa sede rileva naturalmente l'ordine pubblico, rispetto al quale è stato ben posto in luce quanto sia difficile non solo riempire di contenuto un concetto che rischia sempre di essere difficilmente afferrabile soprattutto se letto in un'ottica meramente ideale (e non materiale, quale, in sostanza, pacifica convivenza sociale della collettività<sup>64</sup>), ma anche, da parte del giudice, accettarne in concreto l'effettiva messa in pericolo di fronte a condotte unicamente comunicative<sup>65</sup>.

Problemi, questi ultimi, che, come subito diremo, necessariamente si pongono anche rispetto all'opportunità/possibilità di sanzionare le condotte di disinformazione facendo leva su altre fattispecie criminose attualmente esistenti o su reati da introdurre *ex novo*.

Un'altra fattispecie, invero, che risulta applicabile rispetto alla diffusione di informazioni false è invero l'art. 656 c.p., che punisce con l'arresto fino a tre mesi o con l'ammenda fino a euro 309 chiunque «... pubblica o diffonde notizie false, esagerate o tendenziose, per le quali possa essere turbato l'ordine pubblico»<sup>66</sup>. Anche in tal caso, quindi, ad essere punita è sì la diffusione, tra l'altro, di una notizia falsa o atta a fornire una rappresentazione alterata della realtà per il modo in cui è presentata<sup>67</sup>, ma solo se e nella misura in cui essa sia idonea ad arrecare turbamento – così come accade, come detto, per le figure di istigazione a delinquere – al controverso bene giuridico dell'ordine pubblico – aspetto, quest'ultimo, che secondo la Corte costituzionale vale a salvare la norma da censure di incostituzionalità per violazione dell'art. 21, dovendo il giudice accertare rigorosamente, e in concreto, l'idoneità dell'affermazione a determinare una turbativa di tale interesse di rilievo per la Carta fondamentale<sup>68</sup>.

Si tratta, peraltro, di una figura contravvenzionale suscettibile di oblazione<sup>69</sup>, con tutti i noti problemi di ineffettività, in punto di scarsa deterrenza di tale modello sanzionatorio (soprattutto, all'evidenza, rispetto ad azioni coordinate di disinformazione), che pure presenterebbe alcuni vantaggi, nella prassi applicativa, in termini di possibilità di imputare l'illecito, a livello soggettivo, indifferentemente a titolo di dolo o colpa<sup>70</sup>.

Collocabile nell'ambito della repressione delle notizie false è anche la contravvenzione di procurato allarme presso l'autorità di cui all'art. 658 c.p., volta a sanzionare chiunque «... annunziando disastri, infortuni o pericoli inesistenti, suscita allarme presso l'autorità o presso enti o persone che esercitano un pubblico servizio». Anche in tal caso si tratta quindi di una contravvenzione, sempre a tutela dell'ordine pubblico, e con tutti i richiamati limiti propri dell'art. 656 c.p., con l'ulteriore difficoltà che in questa ipotesi è necessario accertare il verifi-

<sup>64</sup> Su tali nozioni v., per tutti, DE VERO (1995), pp. 72 ss.

<sup>65</sup> Per un'accurata analisi del tema v. GALLUCCIO (2020), cit., *passim*. Diffusamente sul punto anche PELISSERO (2015), cit., il quale tra l'altro rileva che, nonostante debba condividersi, con riferimento ad es. all'istigazione a delinquere, quella giurisprudenza che richiede l'accertamento di un «... pericolo immediato e concreto di commissione di reati, è anche vero che in questo ambito il pericolo concreto mostra scarse capacità selettive, rimesse al potere discrezionale del giudice, lasciato privo di criteri direttivi controllabili e facilmente condizionabile dal contesto politico-sociale». In merito all'accertamento di tale pericolo si veda, efficacemente, già DE VERO (1993a), pp. 292 ss. Sull'articolazione di tali concetti negli Stati Uniti, in particolare avuto riguardo alla dottrina del c.d. *clear and present danger*, v.: ABBONDANTE (2017), pp. 44 ss., nonché, diffusamente GALLUCCIO (2020), cit., pp. 167 ss.

<sup>66</sup> Sulla fattispecie in questione v. già CHIAROTTI (1964), pp. 514 ss.

<sup>67</sup> Si è osservato, invero, che l'espressione "false, esagerate e tendenziose" va letta come una endiadi, in quanto ciò che conta è che l'espressione presenti una visione alterata della realtà: così Corte Cost., n. 19 del 1962, in *Giur. Cost.*, 1962, 1, 189 ss.

<sup>68</sup> V. Corte Cost., n. 19 del 1962, cit.; nonché, successivamente, Corte Cost., n. 210 del 1975, in *Giust. pen.*, 1976, 1, 353 ss. V. sul tema anche CHIAROTTI (1964), cit., p. 516. In argomento anche DE FLAMMINEIS (2020), cit., p. 145, il quale evidenzia i rischi legati a una visione idealizzata del bene giuridico in tale fattispecie, quale visione etico-politica imposta in un certo contesto storico, e la necessità di agganciare all'ordine pubblico "reale" l'oggetto di tutela in tale ambito.

<sup>69</sup> Per quanto, trattandosi di una ipotesi di oblazione c.d. facoltativa, il giudice, ai sensi dell'art. 162-bis c.p., potrebbe respingere l'istanza avuto riguardo, tra l'altro, alla gravità del fatto.

<sup>70</sup> È stata peraltro depenalizzata, con il d.lgs. n. 8 del 2016, la fattispecie di cui all'art. 661 c.p., oggi invero punita soltanto con una sanzione amministrativa pecuniaria nei confronti di chiunque, pubblicamente, cerca con qualsiasi impostura, anche gratuitamente, di abusare della credulità popolare, sempre, tuttavia, che dal fatto possa derivare un turbamento dell'ordine pubblico. V. altresì COSTANTINI (2017), cit., p. 65.

carsi di un vero e proprio evento costituito dal generarsi dell'allarme presso l'autorità<sup>71</sup>. Una figura, peraltro, che difficilmente risulta adatta a reprimere le notizie false generate con il fine di orientare il dibattito pubblico, anche con azioni coordinate, in una certa direzione, essendo al più in grado di intercettare le condotte di diffusione di notizie false, attraverso vari sensazionalismi, con scopo di lucro economico, per ottenere, ad es., rendite pubblicitarie dall'accesso al sito contenente la falsa notizia.

Lo stesso dicasi per l'eventualità di una falsa notizia diffusa in rete quale artificio/raggiro per truffare uno o più consumatori, inducendoli effettivamente in errore e determinandoli ad effettuare un atto di disposizione patrimoniale per loro pregiudizievole con il conseguimento di un ingiusto profitto in capo all'agente (tale quindi da configurare il reato di truffa ex art. 640 c.p.)<sup>72</sup>.

Di scarsissimo rilievo pratico, per ovvi motivi, è poi il delitto di cui all'art. 265 c.p. (disfattismo politico), volto, tra l'altro, a sanzionare la diffusione di voci o notizie false, esagerate o tendenziose che «...possano destare pubblico allarme o deprimere lo spirito pubblico o altrimenti menomare la resistenza della nazione di fronte al nemico», trattandosi di una fattispecie applicabile soltanto «...in tempo di guerra»<sup>73</sup>. Si tratta, insomma, è il caso di dirlo, di una norma che costituisce soltanto un mero residuo "bellico" del nostro codice penale.

Tuttora applicabile, anche se non con minori difficoltà dei reati fin qui elencati, è inoltre il delitto di cui all'art. 294 c.p. (rubricato "Attentati contro i diritti politici del cittadino") che punisce con la reclusione da uno a cinque anni chiunque «...con violenza, minaccia o inganno impedisce in tutto o in parte l'esercizio di un diritto politico, ovvero determina taluno a esercitarlo in senso difforme dalla sua volontà». Nella dinamica di tale fattispecie, invero, la diffusione di una notizia falsa online o una azione coordinata di disinformazione possono valere ad intercettare il concetto di "inganno" richiamato dalla norma; tuttavia, per l'integrazione dell'illecito non è sufficiente la mera idoneità della condotta a cagionare il risultato descritto, essendo necessario, secondo la struttura di una vera e propria fattispecie di evento<sup>74</sup>, che effettivamente tramite l'inganno si impedisca l'esercizio, sia del caso, del diritto di voto da parte del cittadino o che questi vi dia corso in senso difforme dalla sua reale volontà<sup>75</sup>.

Nella dinamica processuale, pertanto, bisognerà dimostrare – e qui, naturalmente, sta e cade la concreta tenuta pratica della fattispecie – l'esistenza di un effettivo legame causale tra la singola affermazione falsa diffusa online e la decisione ultima da parte del cittadino in merito al proprio diritto politico, presa in senso difforme dalla sua effettiva volontà, che dovrà risultare il "diretto" risultato proprio del recepimento di quella notizia falsa<sup>76</sup>. Una vera *probatio diabolica*, ben si comprende, anche in considerazione della mole di informazioni e notizie che contribuiscono a determinare in un senso o nell'altro la scelta del cittadino nell'esercizio di tale prerogativa.

Applicabile in tali contesti potrà spesso risultare, poi, il reato di sostituzione di persona di cui all'art. 494 c.p., avendo la giurisprudenza di legittimità in più casi chiarito come tale fattispecie possa applicarsi, sussistendone tutti i relativi presupposti<sup>77</sup>, anche, ad es., alla condotta di creazione di falsi *account* online<sup>78</sup>.

Di minore rilievo pratico, poi, alla luce del delineato quadro criminologico di nostro in-

<sup>71</sup> Si è chiarito come tale evento non costituisca una mera condizione psicologica, essendo necessario lo svolgimento di attività concrete e l'attivarsi effettivo di tali autorità: così DE VERO (1993b), pp. 76 ss. Sull'applicazione di tale fattispecie in questo ambito v. anche GUERINI (2020a), cit., p. 150.

<sup>72</sup> V. anche COSTANTINI (2017), cit., p. 67.

<sup>73</sup> In argomento v. ancora COSTANTINI (2017), cit., p. 65.

<sup>74</sup> Sul punto, con richiami alla dottrina pertinente, v., *ex multis*, MAGLIARO, MIGLIORELLI (2018), p. 1098.

<sup>75</sup> V. da ultimo anche Cass. pen., Sez. I, 20 dicembre 2018, n. 16381, in *Pluris*.

<sup>76</sup> V., su tale fattispecie, anche PADOVANI (2014), cit., 15, che richiama l'efficace esempio della persona che chiede un'indicazione sul simbolo del partito che intende suffragare e che riceve dall'autore del reato una dolosa indicazione ingannevole sul punto, viziando il suo voto. Sul punto, e in generale sul rilievo in tale ambito dei c.d. reati elettorali, si rinvia, anche per ulteriori spunti relativi, ad es., alla disciplina del silenzio elettorale, all'indagine di GUERINI (2020a), cit., pp. 151 ss.

<sup>77</sup> In particolare, la norma punisce con la reclusione fino a un anno, se il fatto non costituisce altro delitto contro la fede pubblica, «...chiunque, al fine di procurare a sé o ad altri un vantaggio o di recare ad altri un danno, induce taluno in errore, sostituendo illegittimamente la propria all'altrui persona, o attribuendo a sé o ad altri un falso nome, o un falso stato, ovvero una qualità a cui la legge attribuisce effetti giuridici».

<sup>78</sup> Cfr. Cass. pen., Sez. V, 8 novembre 2007, n. 46674, in *Il diritto dell'informazione e dell'informatica*, 2008, 4-5, 525 ss., con nota di FLICK (2008), secondo cui «...integra il reato di sostituzione di persona (art. 494 cod. pen.), la condotta di colui che crei ed utilizzi un "account" di posta elettronica, attribuendosi falsamente le generalità di un diverso soggetto, inducendo in errore gli utenti della rete "internet" nei confronti dei quali le false generalità siano declinate e con il fine di arrecare danno a un soggetto le cui generalità siano abusivamente spese, subdolamente incluso in una corrispondenza idonea a lederne l'immagine e la dignità». In senso conforme v. da ultimo anche Cass. pen., Sez. V, 22 giugno 2018, n. 42572, in *C.E.D. Cass.*, n. 274008.



teresse, sono le ipotesi in cui la diffusione di una notizia falsa sia idonea a provocare una alterazione del prezzo di merci o di strumenti finanziari<sup>79</sup>: è questo il caso del delitto contro l'economia pubblica di cui all'art. 501 c.p. ("Rialzo e ribasso fraudolento di prezzi sul pubblico mercato o nelle borse di commercio") e dei reati in tema di *market abuse* (art. 185 T.U.F. o art. 2637 c.c., a seconda che la manipolazione del mercato/aggiotaggio riguardi strumenti rispettivamente quotati – ovvero per i quali è stata presentata una richiesta di ammissione alle negoziazioni in un mercato regolamentato – o meno)<sup>80</sup>.

Maggiori potenzialità applicative, in tali contesti, ha infine il delitto di diffamazione (art. 595 c.p.), allorché la diffusione di una notizia falsa, anche nell'ambito di una campagna organizzata e sistematica di disinformazione, si risolve in una illecita lesione dell'onore individuale altrui<sup>81</sup>. Molto spesso, invero, le informazioni false diffuse online riguardano un singolo individuo/soggetto politico, la cui reputazione viene offesa attraverso l'attribuzione, ad es., di fatti falsi di varia natura<sup>82</sup>.

La prassi applicativa dei delitti contro l'onore, peraltro, specie rispetto ai procedimenti penali riguardanti giornalisti, mette in rilievo un'altra delle possibili sfaccettature legata ai richiamati rapporti tra tutela della verità e diritto penale, avuto riguardo, in particolare, al riconoscimento della scriminante del diritto di critica, subordinata ai noti requisiti della pertinenza all'interesse pubblico della notizia, della continenza dell'esposizione e, appunto, per quanto qui maggiormente rileva, della verità del fatto narrato<sup>83</sup>.

Quest'ultimo requisito, invero, è da sempre uno di quelli in tal senso più problematici, dato che la giurisprudenza, nel pretendere una scrupolosa verifica della notizia da parte del professionista dell'informazione e nel negare la sussistenza del requisito della verità in presenza di una non adeguata attività di ricerca di riscontri sulla correttezza del dato diffuso, ha finito per trasformare la diffamazione – come l'autorevole dottrina che meglio ha approfondito tale argomento rileva<sup>84</sup> – in un delitto contestato, nei fatti, secondo parametri propri della responsabilità colposa, nonostante si tratti di un reato punito esclusivamente a titolo di dolo<sup>85</sup>. L'errore (colposo) del giornalista che non verifica correttamente la notizia e che ritiene quindi erroneamente di agire nella sussistenza della scriminante del diritto di cronaca, invero, ai sensi dell'art. 59 c.p., potrebbe sì *non* escludere la responsabilità penale laddove sia dovuto a colpa, ma solo e unicamente se, in aggiunta, si tratti di reati puniti anche a titolo colposo; tuttavia, come detto, non prevenendo la diffamazione anche la punibilità per colpa, l'errore del giornalista – salvi i casi in cui sia integrabile la figura del dolo eventuale – dovrebbe escludere, in difetto (anche, quindi, imputabile a colpa) della rappresentazione della falsità dell'informazione, la possibilità che questi possa rispondere del delitto in parola sotto il versante del diritto criminale<sup>86</sup>.

Un altro controverso capitolo del rapporto tra verità e diritto penale, insomma.

E senza qui affrontare gli ulteriori temi dell'impossibilità di estendere alle comunicazioni online il regime punitivo previsto per la stampa cartacea tradizionale (si pensi all'art. 57 c.p.), e di alcune problematiche pronunce che si sono registrate in tale ambito rispetto all'alternativa tra applicazione dell'aggravante dell'offesa arrecata con mezzo della stampa ovvero con qualsiasi altro mezzo di pubblicità (dovendosi qualificare come aggravato ai sensi di quest'ultima previsione il fatto diffamatorio verificatosi mediante uso di *social network*, data la peculiare capacità diffusiva di tale strumento e la sua non riconducibilità alle nozioni di stampa fatta propria dalla legge del '48)<sup>87</sup>.

Questo dunque il panorama dei reati attualmente configurabili nell'ordinamento italiano per reprimere le condotte di disinformazione online.

Al netto delle criticità già segnalate rispetto alle singole figure criminose che abbiamo elencato, è necessario a questo punto, prima di richiamare alcuni progetti di riforma che negli

<sup>79</sup> Ove si tratti di strumenti finanziari deve peraltro trattarsi di una alterazione "sensibile" ai sensi dell'art. 185 T.U.F. e dell'art. 2637 c.c..

<sup>80</sup> In argomento v. anche PADOVANI (2014), cit., 12 ss..

<sup>81</sup> Sulla tutela penale dell'onore v., per tutti, l'indagine monografica di GULLO (2013). Di recente in argomento v. anche VISCONTI (2018).

<sup>82</sup> Per un recente inquadramento del tema si veda anche PEZZELLA (2020).

<sup>83</sup> V. da ultimo GULLO (2019), pp. 3928 ss.

<sup>84</sup> V., su tutti, ancora GULLO (2013), cit., pp. 32 ss., anche per tutti i riferimenti bibliografici e giurisprudenziali.

<sup>85</sup> GULLO (2013), cit., p. 32, evidenzia in particolare come l'anomalia dell'indirizzo interpretativo «...che riconosce al giornalista l'esimente putativa a condizione però che questi abbia assolto all'onere di scegliere le fonti con grande oculatezza e che abbia effettuato dunque una accurata verifica della notizia» abbia realizzato «...una torsione interpretativa nell'accertamento degli elementi costitutivi della diffamazione – delitto esclusivamente doloso – che avviene attraverso lo strumentario tipico degli illeciti colposi [...]».

<sup>86</sup> V. ancora GULLO (2013), cit., p. 32.

<sup>87</sup> Per un'aggiornata panoramica di riferimento si rinvia, per tutti, a GULLO (2019), cit., pp. 3919 ss.



ultimi anni hanno interessato la materia, svolgere alcune considerazioni conclusive di carattere generale sul complessivo quadro di contrasto alla disinformazione attualmente in vigore nell'ordinamento italiano e sin qui ricostruito.

La prima riflessione da fare attiene al fatto che nessuna delle norme fin qui richiamate risulta costruita *ad hoc* per reprimere il fenomeno della disinformazione online (frutto di azioni coordinate e non). Ciò naturalmente comporta che nessuna delle fattispecie analizzate risulta pienamente adatta – tanto sul piano oggettivo della descrizione della condotta incriminata, quanto su quello soggettivo delle finalità che devono animare l'agente – ad essere applicata a *tutte* le specifiche declinazioni concrete delle azioni di *disinformation*, che, come abbiamo visto, nelle loro dinamiche concrete sono suscettibili di assumere svariati contenuti e di perseguire finalità anche molto diverse tra loro.

Ciascuna norma, allora, è idonea a regolare soltanto uno dei tantissimi e variegati tasselli del più ampio e intricato mosaico del mondo della disinformazione. Ciò, tra l'altro, in qualche misura non deve stupire, essendo in linea con il carattere di frammentarietà e sussidiarietà del diritto penale, che come visto<sup>88</sup> non tutela mai la veridicità della notizia in quanto tale, reprimendo la diffusione di una informazione falsa solo ove questa risulti di volta in volta lesiva di interessi ulteriori ritenuti meritevoli di protezione.

Altra questione, che per il momento lasciamo in sospeso e che riprenderemo più avanti<sup>89</sup> dopo aver analizzato i più recenti progetti di riforma, è se sia in qualche modo possibile/auspicabile che in ottica futura si costruiscano delle fattispecie ritagliate *ad hoc* sul fenomeno della disinformazione online e a tutela, se del caso, di interessi diversi da quelli protetti dalle figure criminose attualmente vigenti – non essendo, come detto, possibile assumere quale oggetto di tutela la mera veridicità della informazione in sé e per sé.

Una ulteriore considerazione generale da fare rispetto a tale frastagliato impianto repressivo, la quale peraltro rappresenta un aspetto di cui pure bisogna tener conto in ottica di riforma, è che ciò che accomuna gran parte della fattispecie analizzate – a titolo esemplificativo, dall'art. 656 c.p., all'istigazione a delinquere, fino alla diffamazione – pur per ragioni parzialmente diverse a seconda della singola figura in considerazione, è il loro problematico rapporto con la libertà di espressione e il possibile risolversi della loro applicazione concreta in forme di repressione del dissenso democratico (viene sovente evocato al riguardo il fenomeno del c.d. *chilling effect*<sup>90</sup>).

Si pensi, a titolo esemplificativo, alla complicata legittimazione della pena detentiva per i fatti di diffamazione alla luce delle pronunce della Corte EDU<sup>91</sup> e, da ultimo, della Corte Costituzionale<sup>92</sup>, nonché all'ancoraggio dei reati di istigazione a delinquere o di cui all'art. 656 c.p. a un bene giuridico controverso, di consistenza problematicamente “modellabile” e dai contorni sfumati come quello dell'ordine pubblico<sup>93</sup>. Quanto più i contenuti dell'interesse protetto sono incerti e riempibili di significati diversi, invero, tanto più aumenta il rischio che nella prassi i reati volti a colpire condotte comunicative, e posti a protezione di tale interesse, si risolvano in forme più o meno accentuate/esplicite di repressione “politica” degli avversari<sup>94</sup>.

Problemi, questi ultimi, che sono ricorrenti in pressoché tutte le forme di repressione di condotte legittime all'“uso della parola” o, in generale, volendo utilizzare una espressione dalla tassonomia tuttora incerta in letteratura, dei c.d. reati di opinione<sup>95</sup>.

Ad ogni modo, il fatto che da più parti si sia rilevata, a prescindere dagli aspetti di specifico interesse penalistico, la carenza di politiche pubbliche per il contrasto al fenomeno della disinformazione e alle azioni di disinformazione, negli ultimi anni ha spinto il Parlamento a delineare alcuni progetti di riforma che, in alcuni casi, e come spesso accade, hanno puntato proprio sul diritto criminale quale mezzo principale di contrasto a un fenomeno percepito

<sup>88</sup> Cfr. *supra* § I.2.

<sup>89</sup> Cfr. *infra* § III.4.

<sup>90</sup> V. GULLO (2021), pp. 217 ss.

<sup>91</sup> Per una approfondita ricostruzione delle pronunce della Corte europea dei diritti dell'uomo in merito alla proporzione del trattamento sanzionatorio nell'ambito in questione si veda, in particolare, GULLO (2013), cit., pp. 57 ss.

<sup>92</sup> Sul punto si rinvia ancora a GULLO (2021), cit., pp. 217 ss.

<sup>93</sup> Su queste problematiche, proprio rispetto all'art. 656 c.p., v. efficacemente, già ALESSANDRI (1973), pp. 716-717.

<sup>94</sup> Su tali questioni rimane un importante punto di riferimento in letteratura il lavoro monografico di VISCONTI (2008). Diffusamente torna sul tema di recente anche GALLUCCIO (2020), cit., *passim*.

<sup>95</sup> Sull'incertezza di tale nozione, tradizionalmente oscillante tra la riconducibilità ad essa di reati a tutela di interessi individuali o soltanto di reati a tutela di beni collettivi, v.: PELISSERO (2015), cit.; SPENA (2007), pp. 689 ss.

come fonte di allarme sociale per la collettività<sup>96</sup>.

Vediamo dunque gli aspetti di principale interesse, sotto il versante sanzionatorio, dei diversi progetti di riforma che, tra i molti presentati negli ultimi anni nell'ambito di tale materia, più hanno attirato l'attenzione degli interpreti.

Una prima proposta di riforma (A.S. 2688, c.d. d.d.l. Gambaro<sup>97</sup>) ha inteso puntare, anzitutto, sull'introduzione di un nuovo reato, l'art. 656-*bis* c.p., volto a sanzionare la pubblicazione o diffusione «...attraverso piattaforme informatiche destinate alla pubblicazione o diffusione presso il pubblico, con mezzi prevalentemente elettronici o comunque telematici, [di] notizie false, esagerate o tendenziose che riguardano dati o fatti manifestamente infondati o falsi», nonché dei nuovi artt. 265-*bis* e 265-*ter* c.p., rispettivamente dedicati alla repressione di chiunque diffonde o comunica «...voci o notizie false, esagerate o tendenziose, che possono destare pubblico allarme, o svolge comunque un'attività tale da recare nocimento agli interessi pubblici o da fuorviare settori dell'opinione pubblica, anche attraverso campagne con l'utilizzo di piattaforme informatiche destinate alla diffusione online», nonché della diffusione di «...campagne d'odio contro individui o di campagne volte a minare il processo democratico, anche a fini politici»<sup>98</sup>.

Naturalmente la proposta si è da subito attirata le (giuste) critiche della dottrina, non solo per la problematicità, per le ragioni che abbiamo ampiamente evidenziato, della mera repressione della diffusione della falsa notizia in sé, con riferimento a una norma a tutela, di fatto, di un generalizzato obbligo di verità in capo al cittadino (è questo il caso del proposto art. 656-*bis*, depurato dal riferimento all'ordine pubblico), ma anche per le ulteriori norme manifestamente contrarie a requisiti anche soltanto minimi di tassatività della legge penale, con la previsione di precetti estremamente generici e indefiniti e a presidio di beni giuridici di dubbia consistenza (si pensi alle condotte che destano "pubblico allarme", al generico riferimento alla lesione di "interessi pubblici" non meglio definiti, o alle campagne volte a "minare il processo democratico")<sup>99</sup>. Una soluzione, insomma, del tutto non percorribile in quanto lesiva dei minimali principi fondamentali di garanzia del diritto penale.

Diverso è invece il percorso intrapreso da un altro progetto di riforma (A.S. 3001, c.d. d.d.l. Zanda-Filippin<sup>100</sup>) che, sulla scorta, come vedremo, del modello tedesco, lungi da introdurre nuovi illeciti penali, ha inteso proporre, sulla scorta del resto di alcuni contenuti del *Digital Services Act*, l'introduzione di disposizioni, il cui rispetto è presidiato da apposite sanzioni amministrative (art. 7), volte a obbligare i fornitori dei servizi di *social network* a predisporre delle procedure efficaci e trasparenti (art. 2) per la gestione dei reclami relativi a contenuti illeciti, nonché l'obbligo di redigere un dettagliato rapporto (art. 4) sulla gestione dei reclami ricevuti e di rendere disponibile in modo facilmente individuale e accessibile agli utenti la possibilità di far ricorso alle procedure di reclamo, con possibilità, peraltro, di ricorrere anche a meccanismi di autoregolamentazione per la decisione circa la rimozione o la determinazione di misure volte a rendere non raggiungibile il contenuto<sup>101</sup>.

Nella definizione di contenuto illecito (art. 1, comma 3 e 4), peraltro, il d.d.l. si riferisce *sol-*

<sup>96</sup> Sugli aspetti di implementazione di politiche pubbliche nell'ambito del contrasto alla disinformazione si rinvia alle corrispondenti parti precedenti del presente studio (spec. Cap. I).

<sup>97</sup> Cfr. A.S. 2688, recante "Disposizioni per prevenire la manipolazione dell'informazione online, garantire la trasparenza sul web e incentivare l'alfabetizzazione mediatica", reperibile al [presente link](#).

<sup>98</sup> Tra le ulteriori misure della proposta, tra l'altro, si segnalano la previsione di alcuni obblighi di comunicazione (art. 3) e di pubblicazione di rettifiche (art. 4) in capo agli amministratori di piattaforme, misure a tutela del soggetto diffamato (art. 5), nonché la previsione (art. 7) di un problematico dovere generalizzato, per i gestori delle piattaforme, di effettuare un costante monitoraggio dei contenuti diffusi attraverso le stesse, con correlato obbligo (penalmente presidiato) di rimozione dei contenuti non attendibili. Una scelta di politica del diritto, quest'ultima, non solo in controtendenza con le indicazioni legislative rinvenibili sul piano interno ed europeo, ma anche tale da imporre obblighi sostanzialmente inesigibili a tali operatori (per un'analisi di tali problematiche v. *supra* Cap. II).

<sup>99</sup> V. ad es. FUMO (2018), cit., p. 89. È stata correttamente censurata, peraltro, anche l'assoluta sproporzione del trattamento sanzionatorio previsto: così COSTANTINI (2019), cit., 72 ss., cui si rinvia anche per l'analisi di due ulteriori progetti di legge che si sono mossi in una direzione estremamente simile al d.d.l. Gambaro (l'A.S. n. 2689 del 2017, teso a introdurre, in sostanza, una ipotesi speciale di diffamazione per i casi di offesa arrecata in capo agli eletti, e figure volte a sanzionare la diffusione di notizie false online al fine, tra l'altro, "di influenzare e arrecare danno al buon andamento della politica democratica della Repubblica" o "indurre all'odio", e l'A.C. 4552 del 2017, volto ad ampliare il raggio applicativo dell'art. 656 c.p. attraverso, tra l'altro, il riferimento all'utilizzo di strumenti telematici e all'idoneità "ad arrecare danno ingiusto alle persone", con aumento delle pene ivi previste). Disegni di legge, quindi, nei cui confronti possono in definitiva esprimersi le medesime perplessità avanzate rispetto al d.d.l. Gambaro.

<sup>100</sup> Cfr. A.S. 3001, recante "Norme generali in materia di social network e per il contrasto della diffusione su internet di contenuti illeciti e delle fake news", reperibile al [presente link](#).

<sup>101</sup> Per un'analisi di questo e degli altri progetti di riforma qui richiamati si veda anche GUERCIA (2019), pp. 1253 ss. Una dettagliata disamina della proposta può leggersi anche in DE SIMONE (2018), cit., 17 ss.

*tanto* a taluni dei reati primi richiamati – tra gli altri, ad es., l'art. 595 c.p., e gli artt. 414 ss. etc. –, che, come visto, rappresentano parte degli illeciti configurabili nell'ambito delle condotte di disinformazione, senza, quindi, qualificare espressamente quali illecite le notizie false in sé, che non ledano interessi ulteriori presidiati dagli elencati illeciti penali.

L'approccio di tale d.d.l. è dunque quello di far leva più sull'aspetto preventivo piuttosto che su quello repressivo nell'ambito del contrasto alla condivisione di contenuti illeciti e notizie false in rete<sup>102</sup>. Il limite principale di tale proposta, tuttavia, ci pare risieda nel fatto che essa limita l'applicabilità di tali procedure a una lista di illeciti penali che non solo non intercetta tutti i potenziali illeciti penali più di frequente configurabili nell'ambito delle azioni di disinformazione che abbiamo sin qui ricostruito, ma anche poiché, come detto, nessuno di essi è in ogni caso in grado di cogliere al meglio, sotto i versanti oggettivo e soggettivo, le più svariate manifestazioni del fenomeno di nostro interesse. Ferme restando, peraltro, le potenziali criticità – rispetto alle quali, come anticipato, svolgeremo alcune considerazioni nella parte conclusiva della ricerca<sup>103</sup> – connesse all'affidamento esclusivo alle piattaforme private del compito di giudicare della legittimità o meno dei contenuti veicolati in rete.

Ancora diversa, in ultimo, è la scelta di politica legislativa fatta propria da un'altra proposta di riforma (A.C. 4692, c.d. d.d.l. De Girolamo<sup>104</sup>), che, per quanto qui interessa<sup>105</sup>, ha inteso puntare sull'introduzione del divieto, penalmente presidiato (art. 1), di immettere in maniera anonima in internet qualsiasi contenuto, con contestuale obbligo delle piattaforme, la cui violazione comporta anche in tal caso la contestazione di un illecito penale a carico del gestore del *network*, e problematicamente improntato all'applicazione di una pena fissa<sup>106</sup>, di registrare gli utenti tramite, tra l'altro, indirizzo di posta elettronica e codice fiscale.

Al netto della controversa opportunità di introdurre un simile generalizzato (e indiscriminato) divieto di anonimato<sup>107</sup>, e delle criticità legate alla comminatoria di pene fisse<sup>108</sup> e notevolmente sproporzionate rispetto al disvalore del fatto, il progetto in parola, all'evidenza, riguarda in ogni caso un aspetto marginale di un fenomeno, quello delle azioni di disinformazione, ben più ampio e complesso e rispetto al quale, come vedremo nell'ambito della parte conclusiva della ricerca<sup>109</sup>, è necessario implementare ben altre strategie di politica legislativa.

## III.2. *Le “sanzioni” irrogate dalle piattaforme private nell'ambito dell'attività di moderazione dei contenuti: l'esempio di Facebook e Twitter.*

Abbiamo già evocato l'importanza delle misure di moderazione dei contenuti – e delle correlative azioni di *private enforcement* – da parte delle piattaforme private nel contrasto alla disinformazione.

Vediamo allora come si articolano in concreto le *policy* adottate, su tale versante, da due *social network*, *Facebook* e *Twitter*, qui selezionati per la loro ben nota importanza nell'articolazione dell'odierno dibattito sociale e politico.

Prendiamo le mosse dall'analisi delle politiche di *Facebook*. Il noto *social network*, invero, stabilisce nelle proprie condizioni d'uso un generale divieto per l'utente di adottare condotte o

<sup>102</sup> Per un commento alla proposta v. anche GUERINI (2020a), cit., pp. 173 ss.

<sup>103</sup> Cfr. *infra* § III.4.

<sup>104</sup> Cfr. A.C. 4692, recante l'“Introduzione del divieto dell'uso anonimo della rete internet e disposizioni in materia di tutela del diritto all'oblio”, reperibile al [presente link](#).

<sup>105</sup> L'articolo composto da soli due articoli, invero, oltre alle misure in tema di divieto di anonimato, contiene delle previsioni (art. 2) rispetto alla tutela del diritto all'oblio, consentendo ad alcuni soggetti di chiedere la cancellazione di contributi lesivi diffusi in rete, anche rivolgendosi alle autorità pubbliche in caso di rifiuto o omessa cancellazione dei dati.

<sup>106</sup> Da segnalare, invero, l'assoluta sproporzione della pena prevista che, pur limitandosi alla previsione di una ammenda, prevede in tal caso una sanzione (problematicamente, in misura fissa) di 5 milioni di euro (anche nel caso della violazione dell'obbligo di anonimato di cui al primo comma dell'art. 1, peraltro, si prevede una ammenda fissa di 25.000 euro).

<sup>107</sup> MARUCCI (2018), cit., pp. 27–28, ad esempio, rileva come l'anonimato rappresenti pur sempre una garanzia contro indebite interferenze sull'esercizio di diritti di rilievo costituzionale, pur potendosi introdurre limitazioni per legge, ma solo nel rispetto dei principi di proporzionalità e ragionevolezza. Sul rilievo anche storico dell'anonimato, in generale, quale strumento a garanzia della libertà di espressione, e sull'irragionevolezza, peraltro, della previsione di un divieto valido soltanto online, v. DE SIMONE (2018), cit., pp. 37–38. In merito all'inopportunità di sanzionare penalmente la violazione di un divieto di anonimato in rete v. altresì MELZI D'ERIL (2017), cit., p. 65.

<sup>108</sup> V. anche GUERINI (2020a), cit., p. 177.

<sup>109</sup> Cfr. *infra* § III.4.

condividere contenuti, tra l'altro, «...contrari alla legge, ingannevoli, discriminatori o fraudolenti» o «...alle condizioni, agli standard della community e alle altre condizioni e normative applicabili all'uso di Facebook [...]»<sup>110</sup>.

Nell'ambito delle regole relative agli *standard* della *community*, poi, il *network* dettaglia in modo preciso la propria *policy* in materia di notizie false. In particolare, invero, come si legge al punto numero 21<sup>111</sup>, la piattaforma, pur prendendo atto che «...la riduzione della diffusione delle notizie false su Facebook è una responsabilità molto importante» per il *social network*, rileva come «...quella tra notizie false e satira o opinioni è una linea molto sottile. Per questi motivi, non rimuoviamo le notizie false da Facebook, ma ne riduciamo la diffusione in modo significativo mostrandole più in basso nella sezione Notizie»<sup>112</sup>.

Un'eccezione a tale *policy* generale, legata a intuibili considerazioni circa la tutela della salute pubblica, peraltro, è stata di recente adottata dalla piattaforma in relazione alla diffusione di notizie false relative al virus da Covid-19, avendo la società reso noto di aver rimosso dalla propria rete «...*false claims about COVID-19 vaccines that have been debunked by public health experts*», pur avendo parzialmente (ulteriormente) rivisto tale regola nella misura in cui Facebook ha dichiarato, lo scorso 26 maggio, che «...*in light of ongoing investigations into the origin of COVID-19 and in consultation with public health experts, we will no longer remove the claim that COVID-19 is man-made or manufactured from out apps*»<sup>113</sup>. Ciò, se vogliamo, a testimonianza della delicata (e problematica) natura di qualsiasi attività censoria rispetto alla diffusione di simili contenuti che sovente non assumono di per sé natura illegale per l'ordinamento.

Di recente, inoltre, la piattaforma ha informato l'opinione pubblica circa l'imminente lancio di ulteriori misure organizzative di contrasto alla diffusione di informazioni false, ferma restando la *policy* generale di non rimozione diretta del contenuto falso in sé e per sé. In particolare, la società ha annunciato l'introduzione di un *pop-up* in base al quale, prima di confermare il proprio *like* a una pagina che ha condiviso ripetutamente informazioni qualificate come false da *fact-checkers* indipendenti, l'utente verrà avvisato di tale circostanza con la possibilità di accedere agevolmente alle ulteriori informazioni in merito, potendo solo a quel punto decidere di confermare comunque il proprio *like* e seguire quindi la pagina (con analoghe misure previste per coloro che si apprestano a condividere singoli contenuti qualificati come falsi in base alle previste procedure di *fact checking* da parte del *social network*)<sup>114</sup>.

L'approccio di Facebook, insomma, al netto dell'eccezione sopra richiamata circa le informazioni false relative al covid-19, è quello di non applicare, rispetto alle condotte di disinformazione, la "prima" misura sanzionatoria nell'ambito delle proprie attività di *private enforcement* per la *content moderation*, consistente nella rimozione del contenuto (misura applicabile, ad esempio, tra l'altro, per quelli che promuovano o istighino "seriamente" alla violenza<sup>115</sup>), preferendo piuttosto, come visto, misure organizzative e di funzionamento del sistema che consentano all'utente di prendere decisioni informate o che diano minore visibilità al contenuto falso.

Oltre alla richiamata misura della rimozione del contenuto, poi, la piattaforma applica in generale le più gravi misure interdittive della sospensione o della disabilitazione in modo permanente dell'accesso all'*account* con riferimento agli utenti che abbiano «...violato chiaramente, seriamente o reiteratamente le proprie condizioni o normative, fra cui in particolare gli Standard della community [...]»<sup>116</sup>, informando contestualmente l'utente delle opzioni a sua disposizione per chiedere una revisione della decisione. Tra queste va segnalata la possibilità, una volta esaurite le vie "ordinarie" di ricorso interno nell'ambito delle procedure previste dalla piattaforma, di ricorrere al Comitato per il controllo, un organismo indipendente composto da personalità di alto profilo e di differente estrazione (tra cui, ad es., docenti universitari, un

<sup>110</sup> Cfr. il punto 3.2 delle [Condizioni d'uso](#) del servizio di Facebook.

<sup>111</sup> V. il punto n. 21 nell'ambito della Parte IV degli *Standard della community* di Facebook, consultabili al [presente link](#).

<sup>112</sup> Esplicito, invece, al punto 12 degli *Standard* è il divieto di condotte di incitamento all'odio, ulteriormente dettagliato circa i contenuti tali da integrare una simile condotta.

<sup>113</sup> Entrambe le indicazioni di *policy* da ultimo richiamate sono consultabili al [presente link](#).

<sup>114</sup> Si veda la [pagina web](#) pubblicata il 26 maggio 2021, ove peraltro la piattaforma annuncia anche il proprio «...*effort to include penalties for individual Facebook accounts*», che, oltre, alle pagine, si rendono responsabili di azioni di *misinformation*. Su tali procedure cfr. anche MATUCCI (2018), cit., p. 16. Sul rilievo, peraltro, del ruolo degli algoritmi anche in tali attività di *fact-checking* automatizzate, e sui relativi rischi, v. MARCHETTI (2020), cit., pp. 32 ss. In letteratura, peraltro, attraverso l'analisi di alcuni studi empirici si è osservato come a volte il c.d. *debunking* possa presentare molti limiti non riuscendo nel suo scopo e come addirittura possa non dissuadere gli autori/diffusori di *fake news*, ma al contrario peggiorare la situazione (c.d. *backfire effect*): v. RUFFO, TAMBUSCIO (2020), pp. 73 ss.

<sup>115</sup> Cfr. il punto n. 1 nell'ambito della Parte I degli *Standard della community* di Facebook, consultabili al [presente link](#).

<sup>116</sup> Cfr. il punto 4.2 delle *Condizioni d'uso* del servizio di Facebook, reperibili al già richiamato [link](#).



ex primo ministro, un vincitore di un premio Nobel per la pace etc.), che – rispetto ai vari ricorsi ricevuti dagli utenti e selezionati dal Comitato stesso quali maggiormente significativi o con riferimento ai casi segnalati dalla piattaforma nell’ambito dell’*enforcement* della normativa della *community* sulla moderazione dei contenuti – emetterà in tale ambito decisioni che *Facebook* si impegna in modo vincolante a rispettare «...a meno che l’attuazione di una risoluzione possa violare la legge»<sup>117</sup>.

Venendo invece alle *policy* adottate in tale ambito da *Twitter*, va registrato come tale piattaforma, nell’ambito della definizione delle condizioni d’uso e delle relative norme generali del *social network*, non adotti, come *Facebook*, alcuna esplicita politica di carattere generale in merito alla condotta di diffusione di informazioni false, non ulteriormente connotata in punto di contenuti e caratteristiche specifiche (anche di contesto) della comunicazione.

Questo tuttavia non significa che *Twitter* non si occupi in alcun modo di contrastare la disinformazione. Tutt’altro.

E invero, seguendo una linea diversa che, per comodità esplicativa, potremmo definire un approccio *case by case* o “settoriale”, la politica di *Twitter* si basa sulla determinazione di alcune più puntuali misure di moderazione di contenuti avuto riguardo a specifici ambiti e/o particolari azioni di manipolazione dell’uso dei propri servizi, alcuni dei quali assumo peculiare importanza rispetto alla ricostruita fenomenologia socio-criminologica della disinformazione.

In particolare, per quanto più interessa in questa sede, la piattaforma disciplina un articolato *set* di regole riguardante norme: su condotte che incitano all’odio<sup>118</sup> e sull’esaltazione della violenza<sup>119</sup>; sulle truffe finanziarie<sup>120</sup> e sull’impersonificazione<sup>121</sup>; su account satirici<sup>122</sup>; sulle informazioni ingannevoli in merito al Covid-19<sup>123</sup>; sulla manipolazione della piattaforma<sup>124</sup>;

<sup>117</sup> Si vedano le informazioni e l’atto costitutivo reperibili a [questo link](#). Oltre ai membri del Comitato, in particolare, sono stati nominati diversi *trustee* con il compito di tutelare l’indipendenza dell’*oversightboard* e di garantire l’efficacia del suo operato, giocando anche un ruolo nella complessa procedura di nomina delineata dall’atto costitutivo. Il Comitato, che funziona sulla scorta, per certi versi, del modello della Corte Suprema americana, ha la possibilità di scegliere quali richieste esaminare – cercando di selezionare, come chiarisce l’atto costitutivo, i casi che hanno maggiore potenzialità di «...guidare le decisioni e le normative future» – e svolgere le sue funzioni sulla base della nota regola di *common law* del valore vincolante delle proprie precedenti pronunce.

<sup>118</sup> Prevedendo il divieto di «...promuovere la violenza contro altre persone, attaccarle o minacciarle sulla base di razza, etnia, origine nazionale, casta, orientamento sessuale, genere, identità di genere, religione, età, disabilità o malattia grave», con ulteriori indicazioni di dettaglio sulle tipologie di affermazioni/condotte vietate. Le misure sanzionatorie, poi, vanno dalla richiesta di rimozione del contenuto alla imposizione dell’uso dell’*account* in modalità sola lettura per «...un certo periodo di tempo», fino alla sospensione permanente dell’*account*, possibile sin dal primo esame ove esso mostri «...un’attività fatta soprattutto di comportamenti offensivi, o se è ritenuto responsabile della condivisione di minacce di violenza». Tali informazioni sono reperibili alla [presente pagina web](#).

<sup>119</sup> Sancendo il divieto di «...minacciare di ricorrere alla violenza contro un individuo o un gruppo di persone... [o di] esaltazione della violenza», anche in tal caso con ulteriori specificazioni esemplificative delle tipologie di affermazioni/condotte vietate. Le misure sanzionatorie, poi, vanno qui dalla richiesta di rimozione del contenuto con contestuale blocco dell’accesso all’*account* in attesa dell’avvenuta rimozione, sino alla possibilità di sospendere in modo permanente l’*account* in caso di ripetuta violazione dopo aver ricevuto un avviso. Le informazioni in questione sono reperibili alla [presente pagina web](#).

<sup>120</sup> Con il divieto di «...utilizzare i servizi di Twitter in modo tale da amplificare o sopprimere artificiosamente delle informazioni... [o di] adottare comportamenti che manipolino o perturbino l’esperienza degli utenti», tra cui, ad es., *phishing*, truffe basate su rapporti di fiducia, schemi di *money-flipping* etc. Le misure preventive e sanzionatorie, inoltre, in tal caso prevedono: la possibilità della piattaforma, rilevati livelli di attività sospetti, di procedere a una prova *antispam* (ad. es. tramite la richiesta di un numero di telefono); l’inserimento dell’URL in una *blacklist* di indirizzi ritenuti non sicuri; la richiesta di rimozione di un *tweet*; il blocco temporaneo o permanente, per i casi più gravi, dell’*account* coinvolto. Le informazioni in questione sono reperibili alla [presente pagina web](#).

<sup>121</sup> Vietando agli *account* di fingere di essere altri utenti, organizzazioni o marchi in modo ingannevole o poco chiaro, anche in tal caso con la possibilità di sospendere in via temporanea o permanente l’*account* (cfr. il [presente link](#)).

<sup>122</sup> Obbligando l’utente a chiarire nella denominazione e nella biografia del proprio profilo di non avere alcuna affiliazione con l’oggetto dell’*account*, esplicitandolo in modo chiaro attraverso l’uso di parole quale “parodia”, “*account* falso”, “fan” etc. In relazione all’attività di *enforcement* la piattaforma chiarisce che in tali ipotesi viene data la possibilità al titolare dell’*account* di conformarsi, ferma restando la possibilità di sospensione temporanea o permanente (in caso di ripetizione delle violazioni) dei profili non conformi a tali regole (v. il [presente link](#)).

<sup>123</sup> Prevedendo il divieto di «...utilizzare i servizi di Twitter per condividere informazioni false o ingannevoli sul Covid-19 che potrebbero causare danni», chiarendo che «...perché un contenuto correlato al Covid-19 possa essere etichettato o rimosso, in base a queste norme, deve: presentare un’affermazione di fatto, espressa in termini perentori; essere indiscutibilmente falso o ingannevole, sulla base di fonti autorevoli e ampiamente disponibili; e[d] essere suscettibile di influire sulla sicurezza pubblica o di causare gravi danni», specificandosi altresì alcune tipologie di affermazioni vietate quali, tra le molte richiamate, le indicazioni false o ingannevoli su alcune categorie di persone meno inclini a essere infettate o a sviluppare sintomi avversi sulla base della loro nazionalità o religione, ovvero relative a istruzioni ingannevoli su come effettuare l’autodiagnosi, o inerenti a teorie cospirazioniste rispetto ai vaccini contro la pandemia. Le misure sanzionatorie vanno dalla richiesta di eliminazione del *tweet* con contestuale blocco temporaneo dell’*account* per le violazioni gravi (sempre con la possibilità di blocco o sospensione permanente in caso di violazioni ripetute, con indicazioni di dettaglio in merito alle misure applicabili in base al numero di avvisi di violazione ricevuti dall’utente), fino alla etichettatura del *tweet*, negli altri casi, unitamente, a seconda delle ipotesi, a ulteriori misure, diverse dal semplice avviso agli utenti, quali la riduzione di visibilità del *tweet*, la disattivazione dei Mi piace o della funzione *retweet*, fino all’inserimento di un link con ulteriori spiegazioni o chiarimenti in merito alla notizia. Le informazioni in questione sono reperibili alla [presente pagina web](#).

<sup>124</sup> I divieti, in tal caso, riguardano, tra i molti: l’uso di *account* falsi per ingannare altri utenti; l’amplificazione o l’interruzione artificiosa di conversazioni tramite l’uso di più *account* o il coinvolgimento di altri (richiamandosi qui pratiche quali uso di *account* in sovrapposizione, o di *account* che interagiscono reciprocamente per “gonfiare” l’importanza di alcuni *tweet*; il coordinamento e la creazione di più *account* per



sulle informazioni false in Francia<sup>125</sup>; sull'integrità civica<sup>126</sup>; sui contenuti artificiosi e manipolati<sup>127</sup> e su attività dannose coordinate<sup>128</sup>.

La piattaforma, inoltre, fornisce indicazioni di dettaglio in merito ai principi generali e alle linee guida di riferimento per la propria attività di *enforcement* rispetto all'applicazione di tali *policy* di comportamento della *community*, indicando, ad es., oltre alle "sanzioni" a seconda dei casi applicabili per i contenuti lesivi delle condizioni d'uso – e che, limitandosi alle principali, vanno, a seconda della violazione e della sua gravità, dall'"etichettatura" del *tweet* per informare

creare false interazioni etc.); il "gonfiaggio" artificioso dei propri *follower* o delle interazioni di un *account* tramite condotte quali, ad es., la compravendita del "gonfiaggio" di metriche relative ad *account* o *tweet*, la compravendita di *follower* o interazioni (*retweet*, Mi piace, citazioni etc.), il gonfiaggio reciproco o la vendita di *account* etc. Come per le truffe finanziarie (cfr. *supra* nota n. 120), qui la piattaforma stabilisce l'applicazione di misure di *enforcement* preventive e sanzionatorie, stabilendo la possibilità, rilevati livelli di attività sospetti, di procedere a una prova *antispam* (ad. es. tramite la richiesta di un numero di telefono); di inserimento dell'URL in una *blacklist* di indirizzi ritenuti non sicuri; di procedere alla richiesta di rimozione di un *tweet*, fino al blocco temporaneo o permanente, per i casi più gravi, dell'*account* coinvolto, con la possibilità, in caso di ritenuta violazione delle regole sugli *account* falsi, di richiedere di fornire un documento d'identità ufficiale per ripristinare il profilo. Le informazioni in questione sono reperibili alla [presente pagina web](#).

<sup>125</sup> Indicativa è invece la posizione espressa dalla piattaforma in merito alle recenti novità introdotte dall'ordinamento francese per il contrasto alla disinformazione (su cui v. *infra* il paragrafo successivo). Nella pagina di riferimento (cfr. il [presente link](#)), invece, la piattaforma, da un lato, rinvia alle proprie regole sul divieto di fornire informazioni false sul voto e sulla registrazione per votare (v. *infra* la nota successiva); dall'altro lato, però, nel prendere atto che «...la legge francese richiede che ti forniamo anche un mezzo per segnalare informazioni false in grado di alterare la sincerità del voto o turbare l'ordine pubblico», la piattaforma afferma *apertis verbis*: «...non esaminiamo queste segnalazioni, né prendiamo provvedimenti conseguenti, su base individuale. Tuttavia, utilizziamo tali segnalazioni per spiegare come difendiamo la nostra piattaforma contro la manipolazione». Una presa di posizione, quindi, all'evidenza molto "forte" e con evidenti tratti di problematicità rispetto alla conformità alle indicazioni della normativa francese (su cui cfr. *infra* § III.3 per tutti i pertinenti richiami). Sotto tale profilo, peraltro, da tale affermazione ci pare emerga in modo chiaro quanto rilevavamo in precedenza (cfr. *supra* nel presente paragrafo) circa l'approccio di *Twitter* al tema delle *fake news*: non un generale divieto di diffusione di notizie false in sé e per sé, ma specifici settoriali precetti su notizie false relative a specifici argomenti e interessi sensibili (Covid-19, esercizio del voto etc.), tra cui non pare quindi trovare spazio l'ordine pubblico prima richiamato, della cui problematica nozione si è del resto già detto (cfr. *supra* § III.1.).

<sup>126</sup> Con il divieto di «...utilizzare i servizi di Twitter per manipolare le elezioni o altri processi civici, né per interferirvi. Ciò include la pubblicazione o la condivisione di contenuti che possono inibire la partecipazione o fuorviare le persone in merito a quando, dove o come partecipare a un processo civico», richiamandosi tra le azioni vietate, tra le molte, la diffusione di informazioni ingannevoli sulle procedure di voto o sui requisiti di partecipazione, informazioni ingannevoli sulla chiusura di seggi o sull'attività di forze dell'ordine sul voto o su lunghe code o problemi con le attrezzature o altri disagi sul voto, minacce relative alle sedi di voto o altri luoghi chiave, affermazioni non verificate su brogli elettorali o sui risultati o l'esito di un processo civico tali da interferire con l'attuazione dei risultati dell'elezione (es. rivendicare la vittoria prima della certificazione dei risultati etc.). Le misure sanzionatorie vanno dalla richiesta di eliminazione del *tweet* con contestuale blocco temporaneo dell'*account* per le violazioni di gravità elevata, sempre con la possibilità di blocco o sospensione permanente in caso di violazioni ripetute; da segnalare, peraltro, che in tale ipotesi, come accade per la disinformazione in merito al Covid-19, la piattaforma chiarisce in modo tassativo la tipologia di sanzioni applicabili in base al numero di avvisi di violazione ricevuti dall'utente (dal blocco di 12 ore dopo 2 avvisi alla sospensione permanente in caso di 5 o più avvisi etc.). Applicabile anche in queste ipotesi è poi, per casi diversi da quelli fin qui richiamati, la "sanzione" dell'etichettatura del *tweet*, unitamente, a seconda dei casi, a ulteriori misure, diverse dal semplice avviso agli utenti, quali la riduzione di visibilità del *tweet*, la disattivazione dei Mi piace o della funzione *retweet*, fino all'inserimento di un link con ulteriori spiegazioni o chiarimenti in merito alla notizia. Le informazioni in questione sono reperibili alla [presente pagina web](#).

<sup>127</sup> Sancendo il divieto di «...promuovere in modo ingannevole contenuti multimediali artificiosi o manipolati che potrebbero causare danni con alta probabilità», qualificandosi come tali, tra l'altro, quelli il cui contenuto è stato modificato in modo da alterare composizione, sequenza tempi o inquadratura, o in cui è stata aggiunta o rimossa una qualsiasi informazione visiva o uditiva, o tali da raffigurare una persona reale in modo falsificato o simulato, con ulteriori chiarimenti sulla tipologia di danni che la piattaforma prende in considerazione (tra cui, minacce alla *privacy*, *stalking*, minaccia alla sicurezza fisica, rischio di violenza di massa o di rivolte civili etc.). Le misure sanzionatorie prevedono, in tal caso, il contrassegno del *tweet* che include tali contenuti con l'applicazione di una etichetta per avvisare gli utenti, l'inserimento di un avviso agli utenti prima che condividano il contenuto o clicchino sulla funzione Mi piace, la riduzione della visibilità o l'inserimento di un link per chiarimenti, fino alla possibilità di rimuovere i *tweet* che integrino tutti i criteri prima citati e alla sospensione permanente degli *account* autori di ripetute o gravi violazioni sul punto. Le informazioni in questione sono reperibili a [questo link](#).

<sup>128</sup> In tal caso si tratta di una *policy* più complessa, in quanto la piattaforma non detta qui espliciti divieti, rilevando come «...l'attività dannosa coordinata rappresenta il contesto nel quale si muovono i diversi attori, pertanto concentriamo la nostra valutazione dapprima su gruppi, movimenti e campagne, poi sui provvedimenti da adottare nei confronti degli *account* che risultano associati a tali entità». *Twitter* definisce quindi due principale forme di coordinamento (tecnico e sociale), trattandosi di ipotesi relative all'utilizzo delle tecniche di manipolazione della piattaforma (cfr. *supra* nota n. 124) o di interazione tra persone per amplificare o diffondere un determinato messaggio sulla rete, fornendo poi dettagliati chiarimenti sul livello di gravità (basso, moderato, o alto) e sulla tipologia di danno causato da tali gruppi che *Twitter* intende in tali casi prevenire e mitigare, tra cui quello fisico e psicologico e, per quanto più qui rileva, quello informativo, definito come il «... danno che influisce negativamente sulla capacità di una persona di accedere a informazioni fondamentali per l'esercizio dei propri diritti o che perturba in modo significativo la stabilità e/o la sicurezza di un gruppo sociale o della società, compresa la disinformazione medica, ad es. sul COVID-19. Ciò include i comportamenti dannosi contemplati dalle nostre norme su integrità civica, contenuti artificiosi e manipolati e manipolazione della piattaforma e spam». Una volta stabilito che il gruppo, il movimento o la campagna integrano tali requisiti stabiliti da *Twitter* per integrare un'attività dannosa coordinata, le misure sanzionatorie irrogate dalla piattaforma rispetto al collettivo o la campagna in quanto tale vanno dalla limitazione della visibilità dei relativi *tweet*, alla eliminazione della possibilità di raccomandare tali *tweet* o di visualizzare tendenze direttamente collegate a tali gruppi, fino alla sospensione degli *account* utilizzati principalmente per propagare e/o incoraggiare la partecipazione alla identificata attività coordinata dannosa (tutte misure che "scadono" dopo un periodo di tempo definito come "limitato", ma con la possibilità di riapplicarle "manualmente o automaticamente" ove si rilevi la mancata cessazione dell'attività del gruppo). La piattaforma chiarisce, inoltre, che utilizza una "combinazione di tecnologia e di esame umano" per identificare gli *account* associati a tali gruppi, rispetto ai quali risultano applicabili le consuete misure di richiesta di rimozione dei *tweet*, nonché quelle del blocco e della sospensione anche permanente degli *account* ove risultino violate le regole della piattaforma (tra cui naturalmente quelle dettagliatamente ricostruite nelle note precedenti).

gli utenti della violazione, alla richiesta di rimozione del contenuto con contestuale blocco dell'accesso all'*account*, fino alla sospensione temporanea o permanente del profilo –, anche l'importanza della considerazione del contesto della conversazione in cui viene pubblicato il *tweet* “incriminato”, ovvero l'eccezione di interesse pubblico che consente in rari casi di mantenere accessibile un *tweet* lesivo delle regole della piattaforma<sup>129</sup>.

Al contrario, e ciò naturalmente appare molto significativo, poche informazioni sono reperibili sulle modalità interne di gestione delle procedure di reclamo (lo stesso, del resto, vale per *Facebook* ad eccezione delle regole di funzionamento dell'*oversight-board*).

Dal complesso di tali norme dettate da *Twitter* emergono quindi, da un lato, alcuni puntuali divieti di diffusione di informazioni false o ingannevoli in alcuni ambiti settoriali (ad es. pandemia da Covid-19 o elezioni e processi civici), secondo una precisa strategia di contrasto mirato ad alcune specifiche, dettagliate e ben identificate ipotesi di disinformazione ritenute particolarmente insidiose data la rilevanza dei beni collettivi in gioco (salute pubblica, integrità dei processi elettorali etc.); dall'altro, l'indicazione di una serie di divieti e *policy* che mirano a vietare, anche a prescindere dai contenuti veicolati, alcune specifiche modalità fraudolente, e, anche qui, ben individuate, di uso del servizio, volte ad *amplificare artificialmente* la visibilità di certi messaggi e il diffondersi di determinati contenuti (attraverso, ad es., azioni coordinate dannose e forme di coordinamento tecnico o sociale quale l'uso coordinato di *account fake* o altre condotte ingannevoli).

Indicazione, quest'ultima, il cui retroterra “criminologico” è all'evidenza rappresentato (anche) dall'operato di gruppi organizzati che, come visto, risultano impegnati in azioni coordinate di disinformazione.

In definitiva, *Twitter*, pur non dichiarando esplicitamente la propria generale posizione, a differenza di *Facebook*<sup>130</sup>, sull'approccio alla condivisione di notizie false, costruisce un meccanismo di *private enforcement* estremamente articolato nel contrasto ad alcune specifiche (e rilevanti) manifestazioni di tale fenomeno.

Pare emergere implicitamente, tra le righe di tali regole, che la convinzione sia quella dell'impossibilità/inutilità di implementare una *policy generale* sul contrasto alle informazioni false e alla disinformazione in quanto tale, trattandosi del resto, come abbiamo visto, di un fenomeno, che assume molteplici sfaccettature, nel cui ambito si “agitano” interessi e finalità soggettive dell'azione (dis)informativa molto eterogenei tra loro, e rispetto al quale, quindi, mal si adatta un approccio, per così dire, *one size fits all*.

In presenza di un simile stato delle cose, allora, la (non irragionevole) tendenza che pare emergere da tale disamina è quella di implementare misure settore per settore e non già un generalizzato e indiscriminato divieto di condivisione o diffusione di notizie false.

Una indicazione, questa, senz'altro utile anche alla luce di ciò che emerso dallo studio dei profili di rilevanza penale delle condotte di disinformazione, essendosi evidenziato come il diritto penale non si occupi in realtà in nessun caso di reprimere l'affermazione falsa in sé, ma solo quella che rappresenti un veicolo di aggressione a interessi diversi<sup>131</sup>.

Sotto ulteriore versante, l'analisi degli strumenti di *private enforcement* di tali piattaforme nel contrasto alle azioni di disinformazione restituisce un quadro in molti casi frastagliato e non privo di criticità.

Limitandoci ai principali aspetti problematici ci pare di poter ravvisare che essi risiedano, in sintesi: a) nella profonda differenza dell'approccio delle varie piattaforme, pur parimenti centrali per l'articolazione del dibattito pubblico, ai temi qui oggetto d'attenzione, sia in termini di scelta delle misure di contrasto, sia rispetto al livello di dettaglio nell'indicazione delle proprie *policy*; b) nel livello (a volte non sufficiente<sup>132</sup>) di tassatività della descrizione di alcuni precetti, pur trattandosi di misure che di fatto influiscono sulla libertà di espressione del cittadino e sulla sua possibilità di partecipare attivamente al dibattito democratico, di cui tali

<sup>129</sup> Applicabile ad es., tra l'altro, quando l'*account* rappresenta un membro attuale o potenziale di un organo governativo, secondo specifici criteri di bilanciamento tra rischi di danno e interesse pubblico indicati dalla piattaforma. Per la consultazione di questa e altre misure sui principi che guidano l'attività di *private enforcement* di *Twitter* possono consultarsi le “opzioni per l'applicazione delle norme” della piattaforma al [presente link](#), nonché quelle sull'elusione dei divieti imposti dalla piattaforma reperibili alla [presente pagina web](#).

<sup>130</sup> Che, come visto (cfr. *supra* nel presente paragrafo), invece, salve le rilevate eccezioni, afferma chiaramente di non procedere alla rimozione di contenuti falsi.

<sup>131</sup> Cfr. *supra* § I.2.

<sup>132</sup> Si pensi, a mero titolo esemplificativo, ad alcune locuzioni analizzate *supra* nel presente paragrafo e relative, ad es., ai contenuti che “potrebbero causare danni con alta probabilità” o a quelli che risultino falsi “sulla base di fonti autorevoli e ampiamente disponibili”, concetti, ben si comprende, non immediatamente afferrabili o, comunque, univoci.

piattaforme rappresentano ormai una articolazione importante; c) nelle differenze – si pensi alla mancanza in *Twitter* di una figura assimilabile al Comitato per il controllo di *Facebook* – e nella limitata, se non a volte del tutto assente, trasparenza sul funzionamento delle procedure interne di gestione dei reclami sulle decisioni di moderazione dei contenuti; d) nel rischio legato all'affidare *esclusivamente* alle piattaforme private il compito di farsi “arbitri” di dibattiti di primario rilievo per la vita sociale e politica della collettività, con la possibilità di definire liberamente, in assenza di una comune cornice pubblicistica, le “regole del gioco”, con tutti i potenziali riflessi negativi in termini di possibilità di implementare misure e norme ideologicamente orientate e di repressione del dissenso democratico<sup>133</sup>.

Molti e complessi sono allora i temi di cui tener conto nel delineare *pro futuro*, come vedremo<sup>134</sup>, alcune indicazioni di *policy* per operatori pubblici e privati per l'implementazione di misure di contrasto alla disinformazione sul versante sanzionatorio.

### III.3. *Il contrasto alla disinformazione nel panorama comparatistico (Francia, Germania, Stati Uniti): alcuni spunti di riflessione.*

Nelle sezioni precedenti del presente studio si sono già ricostruiti i più significativi contenuti di alcune tra le più importanti misure di contrasto alla disinformazione introdotte (o in discussione) in Francia, Germania e Stati Uniti<sup>135</sup>.

In questa sede, pertanto, lungi dal ripercorrere nuovamente gli aspetti di dettaglio di tali riforme, svolgeremo alcune considerazioni sui profili di interesse di tali percorsi legislativi dalla prospettiva del diritto punitivo.

Rilevante, anzitutto, è la scelta compiuta dall'ordinamento francese – con le leggi nn. 1201 e 1202 del 2018 relative *à la lutte contre la manipulation de l'information* – di introdurre delle sanzioni penali – anche nei confronti di persone giuridiche – per la violazione di specifici obblighi di trasparenza imposti dalla riforma in capo alle piattaforme, nonché quella di consentire all'autorità giudiziaria, nei tre mesi precedenti la consultazione elettorale, di adottare le misure necessarie a far cessare condotte di disinformazione<sup>136</sup>.

Sempre nella direzione di una responsabilizzazione delle piattaforme private si è poi mosso anche il *Netzwerkdurchsetzungsgesetz (NetzDG)* tedesco, con la previsione in capo agli operatori di specifici obblighi – in tal caso, amministrativamente presidiati – di trasparenza in merito al trattamento dei reclami, di rimozione e blocco di contenuti illeciti, di definizione di procedure efficaci e trasparenti per l'implementazione delle azioni di *private enforcement* per la moderazione di contenuti illeciti etc.<sup>137</sup>.

Allo stesso modo, negli Stati Uniti – ove del resto si è chiarito in dottrina come la mera diffusione di informazioni false non sia di per sé autonomamente sanzionabile se non nei limiti della lesione di interessi tutelati altri rispetto alla veridicità della notizia e attraverso l'integrazione di “tradizionali” figure criminose (quali, ad es., la *criminal libel*)<sup>138</sup> – le recenti proposte di riforma si sono sempre mosse, come analizzato nelle parti precedenti del presente studio<sup>139</sup>, nella direzione di introdurre obblighi di *disclosure* o di *facere* in capo alle piattaforme online.

Anche l'analisi comparatistica condotta nell'ambito di questa ricerca, quindi, conferma,

<sup>133</sup> Sull'esistenza di un simile rischio v. anche, tra i molti, MELZI D'ERIL (2019), cit., pp. 66-67, il quale rileva come «...i gestori, anche solo per evitare di incorrere in “guai” giudiziari, eliminerebbero ogni contenuto sensibile, col risultato di creare un appiattimento generale della comunicazione», proponendo quale possibile soluzione alternativa quella «...di porre una responsabilità, solo di tipo risarcitorio, quindi non penale, in capo a terzi, esclusivamente a determinate condizioni, nel caso in cui, ad esempio, costoro non contrastino l'anonimo autore di contenuti illeciti e non collaborino con le persone offese», ovvero quella di estendere al web la disciplina del diritto di replica e rettifica.

<sup>134</sup> Cfr. *infra* III.4.

<sup>135</sup> Cfr. *supra* Capp. I e II.

<sup>136</sup> Cfr., per tutti i puntuali richiami anche alle pertinenti disposizioni, *supra* Cap. II, § III.3. Ampi riferimenti sul punto, anche alla giurisprudenza costituzionale francese, in GUERINI (2020A), cit., pp. 77-78.

<sup>137</sup> Per un'analisi dettagliata si veda, oltre alle già richiamate parti precedenti del presente studio, anche V. CLAUSSEN (2019), pp. 110 ss.

<sup>138</sup> V. in particolare A. PARK, K.H. YOUM (2019), cit., p. 106. Sul punto si vedano anche C. CALVERT *et. al.* (2018), pp. 99 ss. e CHERTOFF (2017). Sulla resistenza, del resto, a forme di intervento di carattere preventivo nell'ambito della regolazione USA sul ruolo dei *provider*, del resto, anche legata all'idea di proteggere internet quale luogo in grado di arricchire il dibattito pubblico di idee diverse, v. pure ABBONDANTE (2017), cit., p. 53. Per un'analisi della evoluzione, nell'ambito dell'ordinamento di *common law* statunitense, della figura della *criminal libel*, si rinvia altresì alla disamina di VISCONTI (2018), cit., pp. 447 ss.

<sup>139</sup> Cfr. *supra* Capp. I e II.

dalla nostra prospettiva, alcune linee generali di tendenza che, insieme alle correlate criticità, abbiamo già avuto modo di evidenziare.

Anzitutto, invero, trova ulteriore riscontro l'assenza di misure penali a tutela della veridicità dell'informazione *tout court*, dato che nessuno degli ordinamenti in parola ha puntato sull'introduzione di misure repressive della mera diffusione di informazioni false in sé.

Ancora, le linee di indirizzo di tali scelte di politica legislativa sembrano confermare come una strada pressoché obbligata, e nella cui direzione si muove del resto anche il *Digital Services Act* in discussione in ambito eurounitario, sia quella della responsabilizzazione delle piattaforme e della messa in campo di misure volte a promuovere la trasparenza, la correttezza e l'efficacia dell'*enforcement* privato relativo alla moderazione dei contenuti (pur con alcune differenze: si pensi, ad esempio, al predetto ruolo della giurisdizione in Francia, in tale ambito, all'approssimarsi di scadenze elettorali).

Tutto questo, naturalmente – e ciò trova riscontro anche nel dibattito generatosi in questi ordinamenti all'indomani di tali riforme<sup>140</sup> – ripropone anche qui alcune delle richiamate criticità legate all'attuazione di tali politiche, tra cui, ad es.: il rischio di legittimare forme (a seconda dei casi, pubbliche o private) di censura rispetto al libero dibattito pubblico; l'implementazione di politiche di auto-normazione e auto-organizzazione che, se limitate e calibrate unicamente su contenuti illegali, possono risultare inefficaci rispetto alle prevenzioni di alcune specifiche azioni di disinformazione; la chiara definizione di un *set* adeguato di garanzie per i soggetti destinatari delle misure sanzionatorie/inibitorie irrogate dalle piattaforme etc.

Si tratta, in definitiva, di una serie di temi e problematiche comuni a qualsiasi misura di regolazione del fenomeno anche sul versante punitivo.

Chiarito il quadro di riferimento anche sul versante comparatistico, non resta allora che delineare, nell'immediato prosieguo della ricerca, i possibili profili evolutivi della materia attraverso alcune indicazioni di *policy* tra pubblico e privato.

### III.4. *Indicazioni di policy tra pubblico e privato.*

Abbiamo già messo in rilievo punti di forza, aspetti problematici e rischi degli strumenti implementati, o proposti in ottica di riforma, in ambito pubblico e privato per la costruzione di misure di contrasto ad operazioni (coordinate e non) di disinformazione, avuto speciale riguardo ai profili sanzionatori della materia.

È giunto quindi il momento di delineare, in conclusione di tale studio, alcune indicazioni di *policy* per attori pubblici e privati, evidenziando linee di indirizzo ed aspetti di cui tener conto nel regolare il fenomeno della disinformazione anche sul versante punitivo.

Tali indicazioni, peraltro, vanno interpretate in continuità con quelle suggerite nell'ambito del precedente capitolo della presente ricerca<sup>141</sup> – e invero il lavoro contiene in calce un prospetto unitario delle *policy* suggerite, che sintetizza in un solo documento gli esiti cui si è giunti nelle varie sezioni dello studio.

Un primo cruciale aspetto, sul versante dei decisori pubblici, attiene al ruolo del diritto penale.

Qui ci sembra che l'analisi condotta abbia chiaramente dimostrato come la via della criminalizzazione della semplice condotta di diffusione (anche in via coordinata) di notizie false in sé e per sé, lesiva unicamente del mero interesse alla veridicità della notizia, non sia un condivisibile percorso di politica criminale<sup>142</sup>.

Del resto, si è visto come non sia possibile identificare – sul piano etico, giuridico, costituzionale – l'esistenza di un generalizzato dovere di veridicità, e di un diritto del cittadino

<sup>140</sup> Si veda ad esempio CLAUSSEN (2019), cit., pp. 119 ss., per una dettagliata disamina di tutte le *legal concerns* sollevate dal NetzDG tedesco, tra cui quella della "*privatization of law enforcement*", ed LEHNER (2019), cit., p. 6, rispetto alle critiche rivolte in Francia alla richiamata riforma, vista da alcuni come un rischio di censura pubblicistica sulla libertà di espressione e di stampa.

<sup>141</sup> Cfr. *supra* Cap. II, § III.4.

<sup>142</sup> Sull'inopportunità dell'utilizzo in tale ambito dello strumento del diritto penale v., tra gli altri, anche FUMO (2018), cit., pp. 86, il quale evidenzia come la repressione penale nel contrasto alle *fake news* sia «...uno strumento sproporzionato e pericoloso. Sproporzionato perché, come è noto, la sanzione penale deve costituire la *extrema ratio*; pericoloso, non solo perché si introduce un principio contrastante con il fondamento dello stato liberale: il diritto di sbagliare e comunque il "diritto di eresia" (e di fare scelte eterodosse, sempre che non danneggino gli altri), ma anche e principalmente perché si accredita, per contrasto, l'idea che esista una verità ufficiale e dunque – surrettiziamente, ma nemmeno troppo – si legittima una prospettiva da Stato etico». Sulla inopportunità per lo Stato di imporre "*vi et armis*" la verità v., efficacemente, anche PERRONE (2018), cit., p. 26.



ad essere informato secondo determinati contenuti o certe modalità, la violazione dei quali sia penalmente presidiabile<sup>143</sup>. Né tantomeno il panorama comparatistico, quantomeno avuto riguardo agli ordinamenti democratici presi in considerazione, restituisce esempi di criminalizzazione della mera diffusione di informazioni false o della condotta di disinformazione in quanto tale<sup>144</sup>.

Se così è, allora, il diritto penale deve continuare a occuparsi delle condotte di diffusione di false informazioni solo se e nella misura in cui esse risultino offensive – a seconda dei casi nelle forme della lesione o della semplice messa in pericolo – di interessi altri rispetto alla mera della veridicità dell’informazione, che di volta in volta, secondo il principio di frammentarietà del diritto criminale, vengano ritenuti meritevoli di protezione penale<sup>145</sup> (si pensi, ad es., all’onore individuale, al regolare andamento dei mercati e, con tutti i *caveat* svolti<sup>146</sup>, all’ordine pubblico/pace sociale etc.).

Non percorribile, altresì, ci sembra la prospettiva, delineata peraltro da alcuni degli analizzati progetti di riforma di recente portati in discussione nel Parlamento italiano<sup>147</sup>, di costruire delle fattispecie penali tarate *ad hoc* sul generale fenomeno della disinformazione e a tutela di (nuovi) beni giuridici, diversi dalla mera veridicità dell’informazione, quali, si dia il caso, il “processo democratico”, non meglio definiti “interessi pubblici” etc.

Ciò, in particolare, per due ragioni su tutte: a) come visto, da un punto di vista socio-criminologico il fenomeno della disinformazione è suscettibile di assumere molteplici sfaccettature, sicché difficilmente può essere efficacemente affrontato sulla base di una unitaria normazione che, anche sul piano punitivo, mai potrebbe cogliere correttamente il variegato novero di comportamenti e finalità soggettive che si agitano in questo campo; b) l’identificazione di (inediti) beni giuridici di natura collettiva dalla dubbia consistenza o dai contenuti incerti e da presidiare penalmente a tutela, ad es., della regolarità di processi civici, del dibattito pubblico etc., finirebbe per riproporre anche in tale contesto le criticità con cui, come visto, da tempo la prassi si confronta, in particolare, nell’ambito dei reati che reprimono condotte comunicative offensive dell’ordine pubblico, con tutti i correlati rischi di rendere tali nuove figure criminose delle forme di repressione del dissenso democratico.

È agevole del resto paventare le storture e le distorsioni applicative cui potrebbe prestarsi l’introduzione di una norma penale che, ad es., punisse la mera diffusione di “notizie false” – anche, eventualmente, nell’ambito di una azione coordinata – con il fine di influire in modo indebito sull’esito di una elezione. L’incertezza sul preciso contenuto del bene giuridico protetto da una simile figura criminosa, e la stessa ricorrente opinabilità, nel dibattito pubblico, della distinzione tra fatti oggettivi e opinioni soggettive, con il conseguente pericolo di una indebita e incontrollabile espansione della discrezionalità del giudice nel riempire di contenuto simili incerte nozioni, potrebbero invero risolversi nella prassi in pericolosi strumenti di punizione del “nemico politico”.

Le politiche pubbliche in tale ambito, allora, anche sul versante del contrasto “punitivo” alla disinformazione, non possono che essere rivolte – sulla scorta di alcuni dei modelli analizzati – all’introduzione di una cornice legislativa per la responsabilizzazione delle piattaforme, che, oltre al versante degli obblighi di trasparenza/*disclosure* e di *risk assessment e management*<sup>148</sup>, si muova nella direzione di delineare un preciso *set* di regole per la definizione delle regole d’uso del servizio e per la correlata implementazione, da parte degli operatori, delle attività di *private enforcement* in merito alla moderazione dei contenuti illeciti e lesivi di tali *policy*. Al riguardo, peraltro, appare condivisibile la scelta del *Digital Services Act* di limitare l’imposizione degli obblighi di *facere* più stringenti alle piattaforme di rilevanti dimensioni.

In tale ottica, poi, bisognerà aver cura di risolvere alcune delle principali criticità connesse a simili forme di regolazione e che abbiamo avuto modo di evidenziare nel corso dell’indagine.

Un primo aspetto, allora, sarà legato a una esplicita presa di posizione circa la necessità, da parte delle piattaforme, di rispettare i fondamentali principi di garanzia sanciti dalle Carte

<sup>143</sup> Cfr. *supra* § I.2.

<sup>144</sup> Diversa è invece, ad es., la situazione che caratterizza alcuni ordinamenti improntati all’attuazione di politiche di stampo schiettamente autoritario: v., sul punto, con ampia e approfondita disamina specie riguardo alle Repubbliche asiatiche, GUERINI (2020a), cit., pp. 83 ss.

<sup>145</sup> In argomento si vedano le conclusioni cui perviene l’indagine di GUERINI (2020a), cit., p. 209, in merito alla possibilità di ricavare un limitato spazio di intervento penalistico nell’ambito del settore elettorale, ad es. rispetto alle condotte di creazione di cc.dd. *deepfake*.

<sup>146</sup> Cfr. *supra* § III.1.

<sup>147</sup> Cfr. *supra* § III.1.

<sup>148</sup> Circa l’opportunità di promuovere in tale ambito l’adozione di strumenti di *compliance*, anche sulla scorta dell’esperienza applicativa del d.lgs. n. 231 del 2001 in tema di responsabilità da reato degli enti, v. DE SIMONE (2018), cit., p. 30.



europee dei diritti (su tutti, naturalmente, il diritto alla libertà di espressione) nel disciplinare le regole d'utilizzo del proprio servizio.

Necessaria, inoltre, ci sembra in tal senso la definizione in via legislativa dei principi generali cui dovrebbe ispirarsi la trasparente e corretta definizione, da parte degli operatori, non solo delle procedure di gestione dei reclami, ma anche delle singole violazioni, riguardanti sia i contenuti illeciti, sia quelli semplicemente lesivi delle condizioni d'uso del servizio, così come le relative misure sanzionatorie/interdittive di carattere provvisorio o definitivo (dall'etichettatura o rimozione del contenuto, al blocco temporaneo dell'uso del servizio, fino alla sospensione temporanea o permanente dell'*account*).

Nel definire tali violazioni e le relative sanzioni, in particolare, tale cornice legislativa pubblicistica dovrebbe delineare alcuni principi guida per obbligare le piattaforme a garantire, in tale attività di (auto)normazione e nella correlata azione di *enforcement*, un nucleo minimo e inderogabile di garanzie *sostanziali e procedurali* proprie di ciascun paradigma punitivo moderno, tra cui, ad es.: il principio di legalità delle violazioni e delle misure sanzionatorie/interdittive, con i relativi corollari della irretroattività, della tassatività/precisione delle previsioni punitive, e del divieto di analogia, definendo altresì con chiarezza i soggetti titolari della potestà di dettare tali regole; il principio di proporzionalità del trattamento sanzionatorio rispetto alla concreta gravità della violazione; il divieto di responsabilità oggettiva e l'affermazione del principio di colpevolezza, con la necessità di specificare l'elemento soggettivo (dolo o colpa) necessario per integrare la violazione; il diritto al contraddittorio e la garanzia di terzietà e indipendenza del soggetti deputati a irrogare la sanzione e a decidere sui connessi reclami; il diritto di richiedere il riesame della decisione etc.

Dovrebbe, peraltro, trattarsi di un articolato minimo di garanzie, con la possibilità per il soggetto privato di rafforzare tali previsioni prevedendo ulteriori forme di tutela per l'utente.

Ferma restando, poi, l'inopportunità, per le evidenziate ragioni, di imporre delle *policy* generali circa le condotte di mera disinformazione e diffusione di contenuti falsi, sarebbe da valutare anche l'introduzione di singole misure settoriali volte ad uniformare l'approccio delle piattaforme circa alcune specifiche *modalità d'uso* dei servizi, che possono a volte rivestire un peculiare rilievo anche rispetto al tema delle azioni di disinformazione (si pensi, ad es., all'opportunità di imporre ai *network* la previsione di un divieto per gli utenti di porre in essere azioni, coordinate e non, volte all'utilizzo di tecniche informatiche ingannevoli per generare in modo fittizio maggiore interazione rispetto a certi messaggi veicolati online, aumentandone fraudolentemente la visibilità).

Come autorevolmente evidenziato<sup>149</sup>, poi, è necessario che in tale ambito si preveda l'intervento di un'autorità pubblica indipendente non direttamente in ciascuna singola decisione sanzionatoria/inibitoria implementata dai *network* – data la mole di contenuti diffusi in rete sarebbe del resto complicato affidare tale compito a un'unica autorità nazionale o aumentare ulteriormente il carico di lavoro dell'autorità giudiziaria –, ma quantomeno in sede di riesame della decisione contestata dall'utente, adottando una decisione il cui contenuto sia vincolante per l'operatore privato. Autorità pubblica indipendente, poi, che, anche sulla scorta del modello delineato dal *Digital Services Act* nell'ottica di un coordinamento internazionale<sup>150</sup>, vigili sul rispetto di tale cornice legislativa, applicando sanzioni di natura pecuniaria e ingiunzionale<sup>151</sup> nei confronti delle piattaforme che non rispettano gli obblighi in parola. Si dovrebbero inoltre prevedere anche dei meccanismi di c.d. *cooperative compliance* per la risoluzione dei procedimenti – si pensi, ad esempio, a un possibile periodo di “messa alla prova”, con l'assunzione di impegni precisi di carattere riparatorio da parte dell'ente privato, il cui positivo esito dia luogo all'estinzione dell'illecito – che, tuttavia, per non disincentivare le piattaforme a conformarsi *ex ante* a tali normative, ci pare dovrebbero essere accessibili soltanto gli enti che ancor prima

<sup>149</sup> Si veda in particolare la proposta di PITRUZZELLA (2018), cit., p. 45. Per un'ampia indagine sul punto si veda anche il volume di PITRUZZELLA *et al.* (2017), *passim*. Condivide la proposta anche PERRONE (2018), cit., 28. Sulla ricorrenza in tale ambito di tutti gli elementi che giustificano la creazione di un'autorità garante e sulla possibilità di estendere sul punto le prerogative dell'AGCOM, altresì introducendo un modello sanzionatorio-punitivo sulla scorta del caso tedesco, v. GUERINI (2020a), cit., pp. 207-208. Critico, invece, rispetto alla prospettiva di introdurre un'autorità di controllo in tale settore è ZANON (2018), cit., 16.

<sup>150</sup> Sui limiti, in tale ambito, di risposte sul piano soltanto nazionale v. PINELLI (2017), cit., pp. 6-7. In generale sull'importanza, rispetto a fenomeni globali (quali, ad es., la corruzione) di risposte coordinate e armonizzate nello scenario internazionale v., per tutti, SEVERINO (2019), p. 19.

<sup>151</sup> La dottrina che si è occupata del contiguo tema della diffamazione, tra l'altro, rileva – nel contesto di una più ampia proposta di riforma – la necessità di far leva in certi casi anche su sanzioni di carattere interdittivo di serio effetto dissuasivo che, in quel contesto, colpiscono i colpevoli nell'esercizio della loro attività: così GULLO (2013), cit., p. 200 ss. Cfr. anche GULLO (2016), pp. 49 ss.

della contestazione della violazione abbiano mostrato un fattivo impegno nel conformarsi ai detti obblighi legali<sup>152</sup>.

Il complesso di interventi qui delineato, del resto, appare necessario rispetto ad attività di *enforcement* che, come visto, influiscono in modo significativo sull'articolarsi del dibattito pubblico e sull'esercizio della libertà di espressione del cittadino, in modo da evitare i rischi di legittimare forme di censura pubbliche o private.

Il *set* di regole qui delineato, infine, ci sembra valga a costituire anche un utile parametro di riferimento per le *policy* implementabili, anche spontaneamente, da parte degli operatori privati.

A prescindere dall'emanazione di una simile cornice pubblicistica, invero, è auspicabile che le piattaforme private si adeguino alle predette indicazioni, costruendo un sistema di *enforcement* privato rispettoso dei predetti principi di garanzia e assicurando un elevato livello di trasparenza e dettaglio, a differenza di quanto oggi accade, anche nel rendere pubbliche le modalità di funzionamento e le specifiche fasi delle procedure interne di applicazione delle misure sanzionatorie/inibitorie e per la gestione dei correlati reclami da parte degli utenti.

Solo in tal modo, in definitiva, sarà possibile realizzare una efficace *partnership* pubblico-privato, nel contrasto alle operazioni di disinformazione, che assicuri un equo bilanciamento tra tutti gli interessi in gioco.

---

## Bibliografia

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

Alessandri, Alberto (1973): "Osservazioni sulle notizie false, esagerate o tendenziose", in *Rivista italiana di diritto e procedura penale*, 1973, pp. 708-724.

Bassini, Marco, Vigevani, Giulio Enea (2017): "Primi appunti su *fake news* e dintorni", *Rivista di Diritto dei Media*, 1, pp. 11-22.

Birritteri, Emanuele (2021): "La responsabilità da reati ambientali degli enti collettivi: profili dogmatici e tecniche di prevenzione", *Diritto penale contemporaneo – Rivista trimestrale*, 1, pp. 290-314.

Calvert, Clay *et. al.* (2018): "*Fake News and the First Amendment: Reconciling a Disconnect Between Theory and Doctrine*", *University of Cincinnati Law Review*, 1, pp. 99- 138.

Campagnoli, Maria Novella (2020): "Informazione, *social network* & diritto: dalle *fake news* all'*hate speech online*. Risvolti sociologici, profili giuridici, interventi normativi" (Milano, Key Editore).

Candido, Alessandro (2020): "Libertà di informazione e democrazia ai tempi delle *fake news*", *Federalismi*, 20, pp. 105-122.

Caruso, Corrado (2018): "Memoria, verità, pena. È ancora il tempo della libertà di espressione?", *Diritto penale contemporaneo – Rivista trimestrale*, 4, pp. 166-174.

Caterini, Mario (2013): "Criminalità, politica e *mass media*", *Politica del diritto*, 4, pp. 601-623.

Chertoff, Michael (2017): "*Fake News and the First Amendment*", *Harvard Law Review Blog*, 10 novembre 2017.

<sup>152</sup> Sul punto – per ulteriori riferimenti bibliografici e per una più approfondita analisi relativa pur al diverso aspetto della declinazione di tale tema nel settore della disciplina della responsabilità da reato degli enti, ma con considerazioni generali che possono ritenersi valide anche in questo ambito – sia consentito il rinvio a BIRITTERI (2021), pp. 304 ss.

Chiarotti, Franco (1964); voce “Diffusione o pubblicazione di notizie false o tendenziose”, *Enciclopedia del diritto*, vol. XII (Milano, Giuffrè), pp. 514-518.

Claussen, Victor (2019): “*Fighting hate speech and fake news. The Network Enforcement Act (NetzDG) in Germany in the context of European legislation*”, *Rivista di diritto dei media*, 6, pp. 110-136.

Costantini, Anna (2019): “Istanze di criminalizzazione delle *fake news* al confine tra tutela penale della verità e repressione del dissenso”, *Diritto penale contemporaneo – Rivista Trimestrale*, 2, pp. 60-80.

De Flammineis, Siro (2020): “Diritto penale, beni giuridici collettivi nella sfida delle *fake news*: principio di offensività ed emergenze”, *Sistema penale*, 6, pp. 131-146.

De Simone, Federica (2018): “*Fake news, post truth, hate speech*: nuovi fenomeni sociali alla prova del diritto penale”, *Archivio penale*, 1, pp. 1-49.

De Vero, Giancarlo (1993a): voce “Istigazione a delinquere e a disobbedire alle leggi”, *Digesto delle discipline penalistiche*, (Torino, Utet Giuridica), pp. 292-303.

De Vero, Giancarlo (1993b): voce “Inosservanza di provvedimenti di polizia e manifestazioni sediziose e pericolose (Contravvenzioni)”, *Digesto delle discipline penalistiche*, (Torino, Utet Giuridica), pp. 76-87.

De Vero, Giancarlo (1995): voce “Ordine pubblico (delitti contro)”, *Digesto delle discipline penalistiche*, (Torino, Utet Giuridica), pp. 72-96.

Del Bò, Corrado (2019): “La protezione dal falso e la tutela del vero tra filosofia e diritto”, *Governare la paura*, 4, pp. 93-104.

Filice, Fabrizio (2020): “Il disegno di legge in materia di omo-lesbo-bi-transfobia e abilismo. L'analisi delle nuove fattispecie incriminatrici. Verso un diritto penale antidiscriminatorio?”, *Questione giustizia*, 26 novembre 2020, pp. 1-19.

Flick, Caterina (2008): “Falsa identità su Internet e tutela penale della fede pubblica degli utenti e della persona”, *Il diritto dell'informazione e dell'informatica*, 4-5, pp. 526-543.

Foà, Sergio (2020): “Pubblici poteri e contrasto alle *fake news*. Verso l'effettività dei diritti atletici?”, *Federalismi*, 11, pp. 247-264.

Fronza, Emanuela (2012): “Il negazionismo come reato”, (Milano, Giuffrè).

Fumo, Maurizio (2018): “Bufale elettroniche, repressione penale e democrazia”, *Rivista di Diritto dei Media*, 1, pp. 83-91.

Galluccio, Alessandra (2020): “Punire la parola pericolosa? Pubblica istigazione, discorso d'odio e libertà di espressione nell'era di *internet*” (Milano, Giuffrè).

Guercia, Pierluigi (2019), “I progetti di legge sulle *fake news* e la disciplina tedesca a confronto”, in Cadoppi, Alberto, Canestrari, Stefano, Manna, Adelmo, Papa, Michele (a cura di): “*Cybercrime*”, (Milano, Utet Giuridica), pp. 1253-1271.

Guerini, Tommaso (2020a): “*Fake news* e diritto penale. La manipolazione del consenso nelle democrazie liberali” (Torino, Giappichelli).

Guerini, Tommaso (2020b): “La tutela penale della libertà di manifestazione del pensiero nell'epoca delle *fake news* e delle infodemie”, *Discrimen*, 16 giugno 2020, pp. 1-66.

- Gullo, Antonio (2002): “Brevi note sulla verità come limite del diritto di cronaca”, *Diritto e formazione*, 5, pp. 629-636.
- Gullo, Antonio (2013): “Diffamazione e legittimazione dell'intervento penale. Contributo a una riforma dei delitti contro l'onore”, (Roma, Aracne).
- Gullo, Antonio (2016): “La tela di Penelope”, *Diritto penale contemporaneo - Rivista trimestrale*, 1, pp. 31-51.
- Gullo, Antonio (2019): *Sub art. 595*, in Padovani, Tullio (a cura di): “Codice Penale” (Milano, Giuffrè), pp. 3907-3940.
- Gullo, Antonio (2020): “I modelli organizzativi”, in Lattanzi, Giorgio, Severino Paola (a cura di): “Responsabilità da reato degli enti, vol. I, Diritto sostanziale”, (Torino, Giappichelli), pp. 241-288.
- Insolera, Gaetano (2018): “Quale memoria per quale diritto penale?”, *Diritto penale contemporaneo - Rivista trimestrale*, 4, pp. 118-127.
- Lehner, Eva (2019): “Fake news e democrazia”, *Rivista di Diritto dei Media*, 1, pp. 93-122.
- Magliaro, Letizio, Migliorelli, Andrea (2018): *Sub art. 294 c.p.*, in Cadoppi, Alberto, Canestrari, Stefano, Veneziani Paolo (diretto da): “Codice penale commentato” (Torino, Giappichelli), pp. 1097-1099.
- Marchetti, Gloria (2020): “Le fake news e il ruolo degli algoritmi”, *Rivista di Diritto dei Media*, 1, pp. 29-36.
- Matucci, Giuditta (2018): “Informazione online e dovere di solidarietà. Le fake news tra educazione e responsabilità”, *Rivista AIC*, 1, pp. 1-32.
- Melzi d'Eril, Carlo (2017): “Fake news e responsabilità: paradigmi classici e tendenze incriminatrici”, *Rivista di Diritto dei Media*, 1, pp. 60-67.
- Monti, Matteo (2020): “La disinformazione online, la crisi del rapporto pubblico-esperti e il rischio della privatizzazione della censura nelle azioni dell'Unione Europea (Code of practice on disinformation)”, *Federalismi*, 11, pp. 281-305.
- Padovani, Tullio (2014): “Menzogna e diritto penale” (Pisa, Pisa University Press).
- Park, Alan, Youm, Kyu Ho (2019): “Fake news from a legal perspective: the United States and South Korea compared”, *Southwestern Journal of International Law*, 25.1, pp. 100-119.
- Pelissero, Marco (2015): “La parola pericolosa. Il confine incerto del controllo penale del dissenso”, *Questione giustizia*, 4.
- Perini, Chiara (2017): “Fake news e post-verità tra diritto penale e politica criminale”, *Diritto penale contemporaneo*, 20 dicembre 2017, 3, pp. 1-14.
- Perrone, Roberto (2018): “Fake news e libertà di manifestazione del pensiero: brevi coordinate in tema di tutela costituzionale del falso”, *Nomos*, 2, pp. 1-30.
- Pezzella, Vincenzo (2020): “La diffamazione: le nuove frontiere della responsabilità penale e civile e della tutela della privacy nell'epoca dei social, delle fake news e degli hate speeches” (Milano, Utet Giuridica).
- Pinelli, Cesare (2017): “Postverità verità e libertà di manifestazione del pensiero”, *Rivista di Diritto dei Media*, 1, pp. 1-7.

- Pitruzzella, Giovanni (2018): “La libertà di informazione nell’era di Internet”, *Rivista di Diritto dei Media*, 1, pp. 1-28.
- Pitruzzella, Giovanni, Pollicino, Oreste, Quintarelli, Stefano (2017): “Parole e Potere: Libertà d’espressione, *hate speech* e *fake news* (Milano, Egea).
- Pizzetti, Francesco (2017): “*Fake news* e allarme sociale: responsabilità, non censura”, *Rivista di Diritto dei Media*, 1, pp. 48-59.
- Pollicino, Oreste (2018): “La prospettiva costituzionale sulla libertà di espressione nell’era di Internet”, *Rivista di Diritto dei Media*, 1, pp. 48-82.
- Pollicino, Oreste (2020): “*The european approach to disinformation: comparing supranational and national measures*”, *Annuario di diritto comparato e di studi legislativi*, pp. 175-212.
- Pulitanò, Domenico (2014): “Cura della verità e diritto penale”, in Forti, Gabrio, Varraso, Gianluca, Caputo, Matteo (a cura di): “Verità del precetto e della sanzione penale alla prova del processo” (Napoli, Jovene), pp. 63-98.
- Ruffo, Giancarlo, Tambuscio, Marcella (2020): “Capire la diffusione della disinformazione e come contrastarla”, *Federalismi*, 11, pp. 72-84.
- Scaroina, Elisa (2018): “Giustizia penale e comunicazione nell’era di *Twitter* tra controllo democratico e tutela dell’onore”, *Archivio penale*, 3, pp. 1-28.
- Severino, Paola (2019): “Strategie di contrasto alla corruzione nel panorama interno e internazionale. *Best practice* ed esigenze di armonizzazione”, *Rassegna dell’Arma dei Carabinieri*, 2, pp. 13-26.
- Spena, Alessandro (2007): “Libertà di espressione e reati di opinione”, *Rivista italiana di diritto e procedura penale*, 2-3, pp. 689-738.
- Stella, Federico (2003): “La costruzione giuridica della scienza: sicurezza e salute negli ambienti di lavoro”, *Rivista italiana di diritto e procedura penale*, 1-2, pp. 55-70.
- Visconti, Arianna (2018) “Reputazione, dignità, onore. Confini penalistici e prospettive politico-criminali”, (Torino, Giappichelli).
- Visconti, Arianna (2020): “Alcune considerazioni criminologiche e politico-criminali sulle c.d. *Fake News*”, *Jus*, 1, pp. 43-71.
- Visconti, Costantino (2008): “Aspetti penalistici del discorso pubblico”, (Torino, Giappichelli).
- Zanon, Nicolò (2018): “*Fake news* e diffusione dei social media: abbiamo bisogno un’Autorità Pubblica della Verità?”, *Rivista di Diritto dei Media*, 1, pp. 12-17.