medialaws

Rivista di diritto dei media
3/2023 dicembre

SEDE
CONTACTS
Studio legale Melzi d'Eril Vigevani
Via San Barnaba 32 - 20122 Milano

Università Bocconi - Dipartimento di Studi Giuridici
Via Roentgen 1 - 20136 Milano
e-mail: submissions@medialaws.eu

ANNO 2023 - CODICE ISSN 2532-9146
Impaginazione a cura di Chiara Pavesi

MediaLaws - Rivista di diritto dei media è una rivista quadrimestrale telematica, ad accesso libero, che si propone di pubblicare saggi, note e commenti attinenti al diritto dell'informazione italiano, comparato ed europeo.

La rivista nasce per iniziativa di Oreste Pollicino, Giulio Enea Vigevani, Carlo Melzi d'Eril e Marco Bassini e raccoglie le riflessioni di studiosi, italiani e stranieri, di diritto dei media.

La rivista fa proprio il Code of Conduct and Best Practice Guidelines for Journal Editors elaborato dal COPE (Comittee on Publication Ethics).

La qualità e il rigore scientifici dei saggi della Rivista sono garantiti da una procedura di *double-blind peer review* affidata a un comitato di esperti per la valutazione individuato secondo criteri di competenza e rotazione e aggiornato ogni anno.

MediaLaws - Rivista di diritto dei media
Regolamento per la pubblicazione dei contributi

1. "MediaLaws – Rivista di diritto dei media" è una rivista telematica e ad accesso aperto che pubblica con cadenza quadrimestrale contributi attinenti al diritto dell'informazione.

2. Gli organi della rivista sono il Comitato di direzione, il Comitato scientifico e il Comitato degli esperti per la valutazione. L'elenco dei componenti del Comitato di direzione e del Comitato scientifico della rivista è pubblicato sul sito della stessa (rivista.medialaws.eu). Il Comitato degli esperti per la valutazione è sottoposto ad aggiornamento una volta l'anno.

3. La rivista si compone delle seguenti sezioni: "Saggi", "Note a sentenza" (suddivisa in "Sezione Europa", "Sezione Italia" e "Sezione comparata"), "Cronache e commenti" e "Recensioni e riletture". I singoli numeri potranno altresì ospitare, in via d'eccezione, contributi afferenti a sezioni diverse.

4. La sezione "Saggi" ospita contributi che trattano in maniera estesa e approfondita un tema di ricerca, con taglio critico e supporto bibliografico.

5. La sezione "Note a sentenza" ospita commenti alle novità giurisprudenziali provenienti dalle corti italiane, europee e straniere.

6. La sezione "Cronache e commenti" ospita commenti a questioni e novità giuridiche di attualità nella dimensione nazionale, europea e comparata.

7. La sezione "Recensioni e riletture" ospita commenti di opere rispettivamente di recente o più risalente pubblicazione.

8. La richiesta di pubblicazione di un contributo è inviata all'indirizzo di posta elettronica submissions@medialaws.eu, corredata dei dati, della qualifica e dei recapiti dell'autore, nonché della dichiarazione che il contributo sia esclusiva opera dell'autore e, nel caso in cui lo scritto sia già destinato a pubblicazione, l'indicazione della sede editoriale.

9. La direzione effettua un esame preliminare del contributo, verificando l'attinenza con i temi trattati dalla rivista e il rispetto dei requisiti minimi della pubblicazione.

10. In caso di esito positivo, la direzione procede ad assegnare il contributo alla sezione opportuna.

11. I saggi sono inviati alla valutazione, secondo il metodo del doppio cieco, di revisori scelti dall'elenco degli esperti per la valutazione della rivista secondo il criterio della competenza, della conoscenza linguistica e della rotazione. I revisori ricevono una scheda di valutazione, da consegnare compilata alla direzione entro il termine da essa indicato. Nel caso di tardiva o mancata consegna della scheda, la direzione si riserva la facoltà di scegliere un nuovo revisore. La direzione garantisce l'anonimato della valutazione.

12. La direzione comunica all'autore l'esito della valutazione.
Se entrambe sono positive, il contributo è pubblicato.
Se sono positive ma suggeriscono modifiche, il contributo è pubblicato previa revisione dell'autore, in base ai commenti ricevuti, e verifica del loro accoglimento da parte della direzione. La direzione si riserva la facoltà di sottoporre il contributo così come modificato a nuova valutazione, anche interna agli organi della rivista. Se solo una valutazione è positiva, con o senza modifiche, la direzione si riserva la facoltà di trasmettere il contributo a un terzo valutatore. Se entrambe le valutazioni sono negative, il contributo non viene pubblicato.

13. Per pubblicare il contributo, l'Autore deve inviare una versione definitiva corretta secondo le regole editoriali della rivista pubblicate sul sito della stessa, un abstract in lingua italiana e inglese e un elenco di cinque parole chiave. Il mancato rispetto dei criteri editoriali costituisce motivo di rigetto della proposta.

14. Le valutazioni vengono archiviate dalla direzione della rivista per almeno tre anni.

15. A discrezione della direzione, i saggi di autori di particolare autorevolezza o richiesti dalla direzione possono essere pubblicati senza essere sottoposti alla procedura di referaggio a doppio cieco ovvero essere sottoposti a mero referaggio anonimo, previa segnalazione in nota.

Indice
Index

Saggi

9 **L'informazione giornalistica. A 60 anni dall'entrata in vigore della l. 3 febbraio 1963, n. 69**
Maurizio Pedrazza Gorlero - Daniele Butturini

41 **Riverberi costituzionali del Metaverso**
Raffaele Bifulco

50 **Paving the path towards general purpose AI systems regulation in the AI Act: an analysis of the Parliament's and Council's proposals**
Giulia Olivato

72 **La questione *deepfake* in Italia: una panoramica**
Veronica Azzali - Nicol Ellecosta

90 **Il reato di sostituzione di persona online di fronte a *socialbot* e ritratti *"AI-generated"*. In favore di una interpretazione estensiva dell'art. 494 del codice penale**
Alessandro Tedeschi Toschi

114 **Delega di funzioni regolamentari e lotta ai rischi sistemici causati della disinformazione nel *Digital Services Act*: quali rischi per la libertà di espressione?**
Andrea Palumbo - Jacopo Piemonte

144 **Il sistema europeo di cooperazione informativa per il contrasto alle minacce informatiche. Verso una definizione di cybersicurezza integrata?**
Federico Serini

Note a sentenza

189 **Tecnologie di riconoscimento facciale e diritti fondamentali a rischio: il caso *Glukhin c. Russia* dinanzi alla Corte europea dei diritti dell'uomo**
Giuseppe Gallo

200 **Diffamazione e *social network*: l'attribuzione del *post* all'imputato tra prova logica e prova diretta**
Andrea Ranghino

204 **I criteri per la determinazione della competenza territoriale nella diffamazione telematica: l'accertamento dell'evento tra rigore tecnico e ricorso alle presunzioni**
Federico Riboldi

Cronache

208 **La tutela del pluralismo nell'ecosistema digitale**
Augusto Preta

233 **Transumanesimo: tra privacy e *AI***
Giulio Lombardi - Antonio Lombardi - Antongiulio Lombardi

Indice
Index

Essays

9 Journalistic information in Italy, 60 years after law 3 February 1963, no. 69
Maurizio Pedrazza Gorlero - Daniele Butturini

41 Constitutional implications of the Metaverse
Raffaele Bifulco

50 Paving the path towards general purpose AI systems regulation in the AI Act: an analysis of the Parliament's and Council's proposals
Giulia Olivato

72 The legal status of deepfakes: an overview
Veronica Azzali - Nicol Ellecosta

90 The crime of online personification in the face of social bots and "AI-generated" portraits. In favour of a broad interpretation of Article 494 of the Criminal Code
Alessandro Tedeschi Toschi

114 Outsourcing of regulatory tasks and the fight against disinformation's systemic risks in the Digital Services Act: what risks for freedom of expression?
Andrea Palumbo - Jacopo Piemonte

144 The European information cooperation system for countering cyber threats. Toward a definition of integrated cybersecurity?
Federico Serini

Case notes

189 Facial recognition technologies and fundamental rights at risk: the case of Glukhin v. Russia before the European Court of Human Rights
Giuseppe Gallo

200 The Italian Supreme Court on the burden of proof for defamatory posts on social networks
Andrea Ranghino

204 The Italian Supreme Court rules on the criteria for determining the proper territorial jurisdiction in online defamation
Federico Riboldi

Comments

208 Media pluralism in the digital ecosystem
Augusto Preta

233 Transhumanism: between privacy and AI
Giulio Lombardi - Antonio Lombardi - Antongiulio Lombardi

Sono stati sottoposti a referaggio anonimo a doppio cieco i contributi di Veronica Azzali - Nicol Ellecosta, Raffaele Bifulco, Maurizio Pedrazza Gorlero - Daniele Butturini, Giulia Olivato, Andrea Palumbo - Jacopo Piemonte, Federico Serini, Alessandro Tedeschi Toschi. Inoltre, il contributo di Giuseppe Gallo è stato sottoposto a referaggio anonimo.

Saggi

L'informazione giornalistica. A 60 anni dall'entrata in vigore della l. 3 febbraio 1963, n. 69*

Maurizio Pedrazza Gorlero - Daniele Butturini

Abstract

Il contributo conduce un'analisi delle funzioni dell'Ordine dei giornalisti, organismo deputato a disciplinare l'accesso alla professione giornalistica nell'ordinamento italiano. La legge professionale del 1963, istitutiva dell'Ordine ha diviso il giornalismo in diverse categorie soggettive. Il contributo esamina la suddetta disciplina e le modalità di applicazione al fine di verificare se l'Ordine sia funzionale a proteggere l'interesse dell'utenza a disporre di media giornalistici improntati alla dignità della professione. Lo studio inoltre cerca di analizzare i cambiamenti intervenuti nel giornalismo nel contesto nazionale al fine di riflettere sulle future sfide di un ordine professionale.

The paper examines the role of the Italian Order of Journalists as a state approved organization governing the journalistic profession in Italy. The 1963 law provided for the *Ordine* to be divided into categories. The paper analyzes the specific provisions and the relevant enforcement to determine if the Order actually safeguards the interest of society to professional dignity of media. The study also explores the changes regarding journalism in Italy to capture the future challenges.

Sommario

1. Introduzione: l'anniversario. – 2. La relazione tra la mediazione giornalistica e le libertà politiche. – 3. La mediazione giornalistica nel passaggio dallo Stato monoclasse allo Stato democratico. – 4. L'impatto costituzionale della l. 69 del 1963. – 5. La paternità culturale della l. 69 del 1963 nel "personalismo" cristiano di Guido Gonella. – 6. L'influenza del personalismo e del concetto di inviolabilità della libertà d'informazione sulla l. 69 del 1963. – 7. I contenuti essenziali della legge professionale. – 8. La natura costituzionalmente indefettibile della responsabilità disciplinare. – 9. Il segreto professionale sulla fonte delle notizie. – 10. La rettifica nella stampa e nella radiotele-

* L'articolo è stato sottoposto, in conformità al regolamento della Rivista, a referaggio "a doppio cieco". Nella condivisione del contenuto del presente lavoro, Maurizio Pedrazza Gorlero è autore dei paragrafi n. 1-7, 13 e 17 e Daniele Butturini è autore dei paragrafi n. 8-12 e 14-16. Una sintesi del lavoro è stata presentata da Maurizio Pedrazza Gorlero al Convegno *"I valori del giornalismo, le sfide dell'informazione, 1963-2023, 60 anni dell'Ordine dei giornalisti"*, Venezia, 20 gennaio 2023.

visione. – 11. L'iniziativa legislativa per l'estensione della disciplina della rettifica della stampa ai siti *online* registrati presso i Tribunali e parificati alla carta stampata. – 12. Il problema giuridico delle comunicazioni *online* non provenienti dai media giornalistici. – 13. Il vuoto riempito dalla giurisprudenza. – 14. La regolamentazione della pratica professionale giornalistica: le scuole di giornalismo. – 15. Il rapporto tra media giornalistici ed altri media *online*. – 16. La crisi del giornalismo professionale come crisi della democrazia costituzionale. – 17. L'attualità della l. 69 del 1963.

Keywords

Ordine dei giornalisti, giornalista, giornalismo, democrazia, libertà

La democrazia è il potere di un popolo informato
Alexis de Tocqueville

1. Introduzione: l'anniversario.

Gli anniversari offrono occasioni di riflessione – le più varie – sull'oggetto della ricorrenza. Quando l'oggetto è una delle leggi di maggiore rilievo nell'ordinamento repubblicano – tale è la l. 3 febbraio 1963 n. 69 che disciplina la professione di giornalista – l'anniversario induce anche a valutarne la resa istituzionale e l'adeguatezza diacronica, a testare cioè quanto la realtà attuale sia inquadrabile nelle categorie e nelle previsioni della legge, al fine di poterne eventualmente prospettare le modifiche richieste dall'evoluzione sociale, culturale e tecnologica.

L'oggetto dell'analisi è di notevole importanza dal punto di vista giuridico-costituzionale, perché il giornalismo professionale costituisce uno strumento essenziale al funzionamento effettivo della democrazia[1]. La l. 69 del 1963 non definisce l'attività d'informazione giornalistica; alla nozione relativa può giungersi, tuttavia, attraverso la percezione socialmente accettata del "dar forma" a notizie aventi ad oggetto fatti di interesse pubblico, per l'appunto ad in-formare.

La trasmissione delle informazioni rappresenta l'esito finale di un'attività complessa che richiede il rispetto di regole e di vincoli. La divulgazione delle informazioni implica, infatti, una serie di passaggi prodromici quali la ricerca delle notizie, la loro selezione, verifica, gerarchizzazione, ricostruzione, formalizzazione e presentazione[2]. Secondo la letteratura, la cifra dell'attività del giornalista consiste in un'opera di decontestualizzazione degli eventi dal loro flusso costante e di ricontestualizzazione nei formati richiesti dall'indirizzo di ogni testata oltre che dai personali orientamenti politico-culturali[3].

[1] C. Esposito, *La libertà di manifestazione del pensiero nell'ordinamento italiano*, Milano, 1958, 5 ss.

[2] C. Sorrentino, *Il giornalismo. Che cos'è e come funziona*, Roma, 2009, 17 ss.

[3] Così C. Sorrentino, *Il giornalismo ai tempi della post-verità,* in *Left*, 2, 2017, 36.

La giurisprudenza afferma che «per attività giornalistica deve intendersi la prestazione di lavoro intellettuale volta alla raccolta, al commento e all'elaborazione di notizie destinate a formare oggetto di comunicazione interpersonale attraverso gli organi di informazione»[4]. Di conseguenza «il giornalista si pone […] come mediatore intellettuale tra il fatto e la diffusione della conoscenza di esso, nel senso, cioè, che sua funzione è quella di acquisire esso stesso la conoscenza dell'evento, valutarne la rilevanza in funzione della cerchia dei destinatari dell'informazione e confezionare quindi il messaggio con apporto soggettivo e creativo»[5].

La l. 69 del 1963 ha avuto un notevole impatto ordinamentale, dal momento che si tratta della prima ed ancora vigente disciplina repubblicana del giornalismo professionale; una legge che pone le condizioni di garanzia di un'attività mediante la quale si esprime, con una pretesa di pienezza almeno in termini di contenuti selezionati e verificati, il più alto grado della libertà di manifestazione del pensiero. Una garanzia che si dilata ad assumere anche una valenza socio-politica, in quanto i cittadini, in ragione delle informazioni sui fatti di rilievo pubblico elaborati dai media professionali, possono esprimere con compiutezza le proprie opinioni, così concorrendo fattivamente all'esercizio dei diritti civili e politici, che permettono alla comunità di partecipare alla gestione della cosa pubblica, influenzandone le decisioni politiche e controllando costantemente l'esercizio del potere.

Il presente contributo è articolato in quattro parti, che corrispondono alle direzioni nelle quali sono obbligate – e, comunque, più fruttuose – le osservazioni sulla legge:

1) il ruolo democratico della mediazione giornalistica;

2) la matrice politico-culturale della l. 69 del 1963, alla luce anche del pensiero di Guido Gonella, che ne fu il massimo ispiratore ed il principale artefice;

3) le fondamenta della l. 69, considerate in relazione all'attitudine che esse manifestano ad incrociare i nodi di maggiore attualità sul piano sia del diritto d'informazione sia della professione giornalistica;

4) la funzione dell'Ordine dei giornalisti nella prospettiva del rapporto tra le ragioni della legge istitutiva dell'Ordine e le attuali problematiche evidenziate dalla c.d. società della comunicazione di massa, soprattutto con riferimento al sovraccarico di comunicazioni provenienti da *media* non ascrivibili al giornalismo professionistico (*social network,* ecc.), che risultano sempre più socialmente influenti.

2. La relazione tra la mediazione giornalistica e le libertà politiche

Affinché si radichi e si sviluppi la democrazia costituzionale si sa che è necessario che ai cittadini siano assicurati il diritto di riunirsi e di formare associazioni, tramite le quali poter discutere liberamente sugli atti dei governanti; il diritto di iscriversi ai partiti per partecipare alla vita politica, influendo sull'indirizzo del raccordo Parlamento-Gover-

[4] Cass. civ., sez. lav., 20 febbraio 1995 n. 1827.

[5] *Ibid.*

no; infine, che siano garantite libere elezioni e perciò una autentica libertà di voto[6].

All'esercizio effettivo di questi diritti di libertà è essenziale la precondizione della libertà d'informazione: «non la democraticità dello Stato ha per conseguenza il riconoscimento di quella libertà, sicché possa determinarne la funzione ed i limiti, ma» sono «le ragioni ideali del riconoscimento di quella libertà (cioè del valore della persona umana) [che] portano tra le tante conseguenze anche alla affermazione dello Stato democratico»[7].

La libertà d'informazione, che è peraltro una forma di esercizio della libertà di pensiero, è dunque il presupposto della libertà di manifestazione del pensiero, in quanto l'espressione dell'opinione su di un fatto deve avvenire senza manipolazioni attraverso un flusso libero di cognizioni. Il libero flusso impone che non vi siano ostacoli legali o di fatto tanto alla libertà di diffondere notizie da parte dei *media* quanto alla libertà della società di ricevere le informazioni[8].

Senza una libera informazione, della quale il giornalismo è pietra d'angolo, il diritto di voto non sarebbe libero e il controllo della società sui poteri politici, economici e culturali non sarebbe effettivo[9]. La sovranità popolare è infatti la *conseguenza* dell'esercizio di diritti politici che hanno la radice prima nella libertà di dare e ricevere informazioni[10].

3. La mediazione giornalistica nel passaggio dallo Stato monoclasse allo Stato democratico

L'evoluzione del contenuto della libertà d'informazione giornalistica è un portato del processo storico-giuridico di trasformazione dei diritti di libertà dalla concezione individualistica dello Stato liberale[11] alla visione pluralistica dello Stato democratico-so-

---

[6] C. Esposito, *Commento all'art. 1 della Costituzione*, in Id., *La Costituzione italiana. Saggi*, Padova, 1954, 10 ss. Cfr. anche C. Esposito, *I partiti nella Costituzione italiana*, in Id., *La Costituzione italiana. Saggi*, cit., 227. Cfr. V. Crisafulli, *La sovranità popolare nella Costituzione italiana* (1954), ora in Id., *Stato, popolo, governo. Illusioni e delusioni costituzionali*, Milano, 1985, 119.

[7] C. Esposito, *La libertà di manifestazione del pensiero nell'ordinamento italiano*, in D. Nocilla (a cura di), *Diritto costituzionale vivente. Capo dello Stato ed altri saggi*, Milano, 1992, 124.

[8] G. Gemma, *La libertà di formazione del pensiero quale autonomo e specifico diritto costituzionale*, in *Studi in onore di Maurizio Pedrazza Gorlero*. Volume II. *La libertà di informazione e la democrazia costituzionale*, Napoli, 2014, 325: «il *genus* della libertà di pensiero in senso lato si suddivide in due *species*: la libertà di formazione del pensiero e quella di manifestazione del pensiero (comunque esso si sia formato, cioè liberamente o meno). Circa la tutela, essa mentre ha quali beneficiari solamente soggetti privati, singoli od associati, può invece operare nei confronti di soggetti sia pubblici che privati. Storicamente parlando, la protezione dei due differenti tipi di libertà di pensiero si è manifestata in misura diversa […] La libertà di manifestazione del pensiero è stata assai più diffusamente riconosciuta e protetta dell'altra. Inoltre mentre la tutela della prima si è manifestata più sul versante pubblico, cioè nei confronti dello Stato, la protezione della libertà di formazione del pensiero si è dispiegata nei confronti di soggetti sia privati che pubblici, e forse più verso i primi che i secondi».

[9] L. Paladin, *Problemi e vicende della libertà d'informazione nell'ordinamento giuridico italiano*, in Id. (a cura di), *La libertà d'informazione*, Torino, 1977, 52.

[10] L. Paladin, *Diritto costituzionale*, Padova, 1998, 629.

[11] A. Bhagwat, *Posner, Blackstone, and Prior Restraints on Speech*, in *BYU Law Review*, 5, 2015, 1168: «*freedom from prior restraints "is the original understanding" of the First Amendment moves too fast. First, consider his sources*».

ciale del secondo dopoguerra del XX secolo[12].

Nel modello liberale è centrale il diritto di libertà negativa, opponibile cioè alle preventive restrizioni da parte degli apparati pubblici (Governo *in primis*). La libertà d'informazione giornalistica è concepita pertanto come «particolare modalità di esercizio della individuale libertà di pensiero»[13]. Tale concezione non illumina i nessi che ruotano attorno alla libertà, ossia «il complicato e complesso intreccio tra momento individuale, momento collettivo e momento politico-generale, tra interessi personali e interessi pubblici, tra libertà e potere, tra liberazione e controllo»[14].

Ciò perché, nell'ottica liberale, il giornalismo si connetteva ad un sistema politico e istituzionale nel quale «i diritti politici e la partecipazione alla formazione dell'opinione pubblica erano ristretti a un gruppo non ampio di cittadini»[15].Il giornalismo era «una forma pubblica di mediazione fra i notabili»[16] all'interno di un processo di comunicazione sociale monoclasse[17] .Vi era una omogeneità sociale tra i produttori delle informazioni, titolari per l'appunto di una libertà personale, e la ristretta porzione di società destinataria dei contenuti notiziali[18]. La configurazione "individualistica" della libertà d'informazione giornalistica era dunque improntata ad una struttura della società di democrazia "incompleta", la cui base sociale veniva data dalla coincidenza di classe fra chi aveva «la capacità effettiva di disporre dei mezzi di comunicazione pubblica» diffusa solamente «fra i cittadini attivi» e chi aveva la «proprietà degli organi di stampa»[19]. In definitiva l'informazione e il potere stavano in capo ad una oligarchia «le cui decisioni di libertà» erano anche e soprattutto «scelte politiche e manifestazioni di un potere pubblico»[20].

La forma di Stato democratica adottata dalla Costituzione repubblicana modifica radicalmente la prospettiva nella quale s'inserisce la libertà d'informazione. I diritti di partecipazione di cui agli artt. 17, 18, 39, 49 Cost. danno vita ad una forma di Stato basata sull'"autonomia" della società, la quale, attraverso il pluralismo sociale e politico,

---

Cfr. U. Allegretti, *Profilo di storia costituzionale italiana. Individualismo e assolutismo nello stato liberale*, Bologna, 1989, 93 ss.

12 A. Baldassarre, *Libertà di stampa e diritto all'informazione nelle democrazie contemporanee (con particolare riguardo a Francia, RFT e USA)*, in *Politica del diritto*, n. 4, dicembre 1986, 579.

13 Ivi, 581.

14 Ivi, 579.

15 Ivi, 583.

16 *Ibid.*

17 M.S Giannini, *I pubblici poteri negli Stati pluriclasse*, in *Rivista trimestrale di diritto pubblico*, 1979, 389 ss. Cfr. anche G. Guarino, *"Classi" e "gruppi" nel pensiero di M.S. Giannini e nella realtà contemporanea*, in Id., *Dalla Costituzione all'Unione europea (del fare diritto per cinquant'anni)*, Vol. IV, Napoli, 1994, 175 ss.

18 A. Baldassarre, *Libertà di stampa e diritto all'informazione nelle democrazie contemporanee (con particolare riguardo a Francia, RFT e USA)*, cit., 583 ss.: la libertà di informazione era «libertà di persone e individui in cui si esauriva in linea di diritto e di fatto l'intero processo di formazione attiva dell'opinione pubblica; […] era un diritto soggettivo e non un privilegio, ma un diritto di un gruppo di persone socialmente e costituzionalmente privilegiate; era una posizione giuridica di privati cittadini, vale a dire l'insieme dei cittadini all'interno del quale si completava il circuito della rappresentanza politica proprio di uno Stato "monoclasse"».

19 Ivi, 584.

20 *Ibid.*

elabora istanze e indirizzi che le istituzioni della rappresentanza politica – le assemblee legislative – hanno poi il compito di mediare[21]. La libertà d'informazione giornalistica conserva sì la natura individualistica di diritto soggettivo e, perciò, di attività funzionale all'interesse della persona, ma a tale aspetto aggiunge la natura di «fondamento essenziale […] del generale processo di comunicazione pluralistica su cui si basa la democrazia»[22]. Ne deriva la multidimensionalità dell'informazione, che viene intesa come diritto di libertà finalizzato alla produzione di idee e conoscenze, come spazio del pluralismo delle notizie e delle opinioni e come rapporto giuridico tra libertà dell'emittente e libertà del ricevente in una ottica d'integrazione e partecipazione[23].

Il giornalismo assicura così l'inclusione popolare perché, tramite l'accesso della società all'informazione e alla conoscenza, garantisce l'effettiva partecipazione del popolo alla vita politica, culturale e sociale[24]. Più precisamente, l'informazione viene a costituirsi in un momento individuale-soggettivo e in una dimensione istituzionale, aprendo ad un processo comunicativo che attiva rapporti fra una molteplicità di soggetti: la posizione di colui che trasmette il pensiero e l'informazione; l'assetto proprietario e organizzativo del mezzo di comunicazione; le attese di coloro che ricevono i contenuti informativi sui quali maturano i propri convincimenti sociali e politici.

Il tema della posizione giuridica del destinatario del prodotto informativo diventa perciò sempre più cruciale perché interroga i meccanismi sociali attraverso i quali si sviluppano i convincimenti culturali della società e quindi la formazione stessa del consenso nei confronti delle decisioni politiche. Diviene pertanto obbligata la riflessione sulla realtà odierna della mediazione giornalistica professionalizzata, al fine di poter valutare il rendimento dell'unica legge nazionale che ne regola e assicura le prerogative e le responsabilità nei confronti della società.

4. La relazione giuridica tra informazione giornalistica e libertà di manifestazione del pensiero

Per procedere a tale valutazione è necessario esaminare la relazione giuridica tra informazione giornalistica e libertà di manifestazione del pensiero. L'informazione come libertà-attività non è espressamente menzionata in Costituzione, ma essa è implicitamente garantita, sul versante positivo, dall'art. 21 Cost. che dedica attenzione espressa

[21] S. D'Albergo, *Il governo parlamentare presidio fondamentale della democrazia economico sociale*, in *Marx ventuno,* Suppl. 2, 2013, 5. Cfr. ancora S. D'Albergo, *Dalla democrazia sociale alla democrazia costituzionale (un percorso dell'ideologia giuridica*), in *costituzionalismo.it*, 3, 2005.

[22] A. Baldassarre, *Libertà di stampa e diritto all'informazione nelle democrazie contemporanee (con particolare riguardo a Francia, RFT e USA)*, cit., 584.

[23] Cfr. J. Habermas, *Storia e critica dell'opinione pubblica*, trad. di M. Carpitella - A. Illuminati - F. Masini - W. Perretta, Bari, 2002, 257, secondo il quale la caratteristica essenziale dello Stato costituzionale sociale era è che «i diritti fondamentali non operavano affatto in senso meramente "delimitativo" poiché, sulla base per la quale era stato concepito questo ordinamento politico, essi dovevano operare come garanzie positive di una partecipazione egualitaria al processo di produzione sia della ricchezza sociale che dell'opinione pubblica».Cfr. F. de Bortoli, *Ordine, 60 anni. La trappola della censura educativa*, in questa *Rivista*, 3, 2022, 11 ss.

[24] R. Dahl, *Sulla democrazia*, cit., 90.

e disciplina dettagliata ad un mezzo di informazione qual è la stampa periodica, e, sul versante logico-giuridico, dal legame fra la libertà d'informazione e la libertà di manifestazione del pensiero, legame per il quale quest'ultima è condizionata dalla prima ai fini della sua piena effettività[25]. La libertà di manifestazione del pensiero è infatti il terminale di una sequenza di libertà che la presuppongono.

La sequenza è costituita, in primo luogo, dalla libertà d'informazione intesa come libertà *negativa* nel senso di non subire intralci legali nel porre in essere un'attività di diffusione al corpo sociale di notizie riferite a fatti d'interesse pubblico[26]. Si tratta di una *libertà da* che affonda le radici in due disposizioni costituzionali. Da un lato, l'art. 21 Cost., nel quale si rinviene il fondamento individualistico della libertà di manifestazione del pensiero, di cui l'attività informativa è presupposto ed espressione. Dall'altro lato, l'art. 41 Cost., che, tutelando la libertà di iniziativa economica privata[27], ne appresta il fondamento materiale di carattere economico-organizzativo, fornendo garanzia all'impresa editoriale, la quale, a sua volta, non può non condizionare il contenuto di pensiero dell'informazione[28]. In tale ultimo senso si apprezza come l'informazione giornalistica assuma la natura ibrida libertà/potere, e come essa, veicolata dalla forza organizzativa ed economica dell'impresa editoriale – così influendo sui convincimenti altrui[29]– concorra a determinare il clima sociale e politico-culturale.

Alla libertà d'informazione segue la libertà di ricevere le informazioni a favore dei

---

[25] P. Barile, *Libertà di manifestazione del pensiero*, Milano, 1975, 19: «chiunque ha diritto a diffondere il proprio pensiero, ha diritto (sono "due facce dello stesso prisma") a che nessuno gli impedisca di ricevere l'informazione». Cfr. V. Crisafulli, *Problematica della «libertà d'informazione*, in *Il Politico*, 1964, 286 in cui la libertà di informazione viene vista come «comprensiva della libertà di divulgare fatti e opinioni». Cfr. C. Sorrentino, *Il giornalismo. Che cos'è e come funziona*, cit., 21, il quale pone l'accento sul fatto che qualsiasi atto comunicativo, compresa l'informazione giornalistica, ha come componente ineludibile il punto di vista soggettivo e, quindi, di pensiero, di chi lo esercita: «ogni atto comunicativo, nel descrivere la realtà, la ricostruisce».

[26] G.U. Rescigno, *Corso di diritto pubblico*, Bologna, 2012, 618.

[27] C. Grisolia, *Libertà di informazione e ordine dei giornalisti alla luce della riforma degli ordinamenti professionali* in *Rivista AIC,* 4, 2012, 11: «chi si pone ancora dalla parte della legge del 1963, costruisce la propria difesa non più (e non tanto) sulla lettura […] dell'art. 21 Cost., ma sulla base di una ricostruzione che, mutando radicalmente i parametri di riferimento costituzionale, riconduce l'attività del giornalista non a tale norma e alle garanzie che essa prevede, ma bensì all'41 Cost.; quest'ultima disposizione certo meno resistente ai limiti e ai vincoli previsti dalla legge sull'ordine. Riassumendo per sommi capi la tesi in questione, basti dire come essa, partendo dal presupposto che l'oggetto dell'impresa editoriale è normalmente costituito non tanto dal diritto alla manifestazione del pensiero dell'editore e di chi lavora alle sue dipendenze, ma da una complessa ed articolata elaborazione di fatti e di idee che ha come suo prodotto finale l'informazione, conclude coll'affermare che tale oggetto non può che tradursi in una mera attività imprenditoriale. Un'attività, rispetto alla quale la libertà di manifestazione del pensiero di chi opera per creare il prodotto finale assume un valore meramente strumentale, in quanto "funzionale" all'opera dell'editore e dei suoi dipendenti, che solo indirettamente contribuiscono ad affermarla e diffonderla». Cfr. V. Zeno Zencovich, *La libertà d'espressione. Media, mercato, potere nella società dell'informazione*, Bologna, 2004, 21 ss.

[28] M. Pedrazza Gorlero, *Il giornalismo nell'ordinamento costituzionale*, in ID, *Giornalismo e Costituzione*, Padova, 1988, 82: «l'editore […] costituito il mezzo di diffusione e compiute le scelte tecnico-organizzative e personali idonee all'attuazione dell'indirizzo, non diffonde – a rigore – il proprio pensiero ma traccia il limite dell'altrui diffusione, cosicché non è paradossale affermare che il contenuto della diffusione è ugualmente il limite all'attività di pensiero del giornalista e l'attività medesima, e che v'è coincidenza, o almeno sovrapposizione fra il pensiero dell'editore e quello dei giornalisti».

[29] Cfr. M. Pedrazza Gorlero, *Il giornalismo nell'ordinamento costituzionale*, cit., 46 ss.

cittadini. Si tratta ancora di una libertà *negativa (da)*, consistente nella pretesa di non venire ostacolati da leggi o da atti o condotte della P.A. e/o dei privati nella possibilità concreta di accedere alle informazioni (notizie e commenti) che i *media* abbiano deciso di diffondere[30].

La libera formazione del pensiero, tramite il contributo informativo dei *media*, è quindi alla base della libertà di manifestazione del pensiero, perché permette che essa sia per l'appunto libera nel suo processo formativo. È perciò di tutta evidenza come la mediazione del giornalismo professionale sia essenziale ad una formazione del pensiero affrancata dalle manipolazioni[31] e come ciò si possa realizzare in ragione degli obblighi che i giornalisti sono tenuti ad osservare, obblighi derivanti dalla legge n. 69 del 1963 e dai doveri deontologici che dalla legge stessa discendono.

In sintesi, la mediazione giornalistica ha il compito democratico di contribuire a rendere libera la volontà dei cittadini e libero il loro pensiero[32].

E qui emerge il legame tra art. 21 e art. 1, c. 2, Cost. La sovranità popolare, infatti, è effettiva solo se il patrimonio di conoscenze trasmesso dal giornalismo consente alla società di essere sovrana *giorno per giorno*, di quella sovranità che si manifesta nell'esercizio quotidiano dei diritti civili e politici.

5. La paternità culturale della l. 69 del 1963 nel "personalismo" cristiano di Guido Gonella

La legge professionale risente dell'epoca e in essa della concezione culturale di Guido Gonella, suo ispiratore e principale artefice. Gonella è stato una delle voci più alte della cultura politica cattolica sia nella fase prerepubblicana sia in quella posteriore all'approvazione della Costituzione. Anche nel dibattito giuridico sul giornalismo egli è intervenuto esternando la sua adesione al "personalismo" cristiano.

La concezione personalistica è stata al centro del dibattito in Assemblea costituente grazie al contributo della cultura politica del cristianesimo sociale. Risente espressa-

[30] G.U. Rescigno, *Corso di diritto pubblico*, cit., 618.

[31] G. Gemma, *La libertà di formazione del pensiero quale autonomo e specifico diritto costituzionale*, cit., 333 ss., il quale nella configurazione del riconoscimento della libertà di formazione del pensiero pone in evidenza una serie di fattori: «a) con riferimento al versante individuale, è un dato pacifico che l'individuo è dotato – per riconoscimento etico, filosofico, giuridico, ecc. – di quel bene denominato dignità. È pure una *communis opinio* che la dignità si sostanzi poi in valori etico-politici quali la libertà e l'eguaglianza […]; la libertà e l'eguaglianza, quindi la dignità, sono incompatibili con una posizione di dominio di un individuo o gruppi di individui su altri. Del resto sembra assurda un'affermazione di segno contrario, cioè che sia compatibile con i detti valori di libertà, eguaglianza, dignità la mancanza di "libertà di volere" di un individuo e la soggezione al volere altrui, vale a dire di un *dominus* non legittimato sul piano etico-politico e giuridico […] b) La motivazione del vantaggio alla collettività, del progresso sociale, determinati dalla libera formazione del pensiero […] c) venendo al motivo proprio della sfera politico-istituzionale, cioè alla necessità di un concorso dei cittadini alla gestione della cosa pubblica, sia con proposte sia con attività di controllo sui governanti, non ci sono da spendere molte parole per comprovare la funzionalità (anche) della libertà di formazione del pensiero […] La libera manifestazione di un pensiero eterodiretto, quindi di opinioni che in realtà sono dei governanti o dei centri di poteri, non ha alcuna utilità ai fini della funzionalità di un ordinamento liberaldemocratico».

[32] Sul punto cfr. G. Vassalli, *Il diritto alla libertà morale (Contributo alla teoria dei diritti della personalità)*, in *Studi giuridici in memoria di F. Vassalli*, II, Torino, 1960, 1634.

mente di tale concezione la parte della Costituzione repubblicana dedicata ai principi fondamentali ed in particolare l'art. 2 secondo il quale l'individuo è persona. La persona non è un *a priori* trascendentale né una unità giuridica astratta, e perciò indipendente dalle condizioni e posizioni soggettive e storiche[33]; al contrario, la persona è un *a priori* sociale, esprime una soggettività materiale e storica. La persona si configura come «una unità ontologica fondativa di azioni materiali diverse che, nella sua necessaria e costitutiva correlazione ai valori (morali, giuridici, o etico-sociali), implica uno spazio *a priori* materiale, un contenuto normativo irradiantesi in tutti i livelli dell'agire umano»[34]. Il concetto di persona ha quindi una valenza etico-sociale e storica nel senso che la persona è una unità sociale che «acquista un senso normativo» e, pertanto, prescrittivo, non solo «nella sua relazione con sé (e i valori corrispondenti) ma soprattutto nella sua relazione essenziale con il "mondo" e con i valori che questo esprime (bene comune, ecc.)»[35].

Ciò che differenzia il personalismo dall'individualismo è proprio la relazione con l'altro: occasionale per l'individualismo, strutturale e costitutiva della stessa identità del soggetto per il personalismo. Per il personalismo il soggetto, in quanto essere sociale inserito in una serie di relazioni, è titolare, oltreché di diritti, di correlativi obblighi; egli è al centro di una rete di responsabilità.

Del personalismo è debitore, come già si è notato, l'art. 2 Cost., soprattutto laddove ravvisa nelle formazioni sociali, quindi nelle relazioni con gli altri, il luogo e il contesto «ove si svolge la personalità» del soggetto e, quindi, l'identità stessa che lo rende individuo. Nel riferimento al verbo *avalutativo* "svolgere", in luogo, ad esempio, di un verbo valutativo e prescrittivo come "sviluppare", sta l'intrinseca, ontologica e naturale socialità della persona[36]. Ciò significa che la libertà di cui la persona è titolare ha *in re ipsa* un risvolto sociale.

La concezione in esame influisce anche sulla natura giuridica del giornalismo il quale, pur generandosi dal tronco della libertà individuale di manifestazione del pensiero di cui all'art. 21 Cost., si affranca dalla nozione individualistica. Il giornalismo viene così a delinearsi soprattutto come uno strumento di sviluppo sociale per coloro che ne sono destinatari[37], di perfezionamento della persona umana, di crescita culturale della società

---

[33] A. Baldassarre, Voce *Diritti inviolabili*, in *Enciclopedia giuridica*, vol. XI, Roma, 1989, 5.

[34] *Ibid.*

[35] Cfr. J. Maritain*, Umanesimo integrale*, trad. di G. Dore riveduta dall'autore*,* Roma, 1980; E. Mounier*, Le personnalisme,* Paris, 2016. Cfr. E. Rossi, *Il pensiero politico di J. Maritain*, Milano, 1956, 313 ss.

[36] Cfr. M. Pedrazza Gorlero, *Il patto costituzionale*, Padova, 2012, seconda edizione, 31: «la stessa *identità personale*, che differenzia l'un individuo dall'altro e ciascun individuo dal gruppo, è in gran parte il risultato delle relazioni reciproche, dell'essere termini ed artefici di una serie indefinita di relazioni sociali. L'identità […] non è che il *supporto genetico nutrito delle relazioni sociali differenzianti custodite dalla memoria* di ciascuno, e – al limite – "l'emergere perturbante della *prossimità dell'altro* alle radici del noi", donde il rovesciamento nella "percezione di una inadeguatezza dell'identità rispetto alla propria pretesa"». Cfr. A. Barbera, Sub *art. 2*, in *Commentario della Costituzione. Principi fondamentali. Art. 1-12*, a cura di G. Branca, Bologna-Roma, 1975, 50 ss.

[37] G. Gonella, *Il discorso delle libertà*, contenuto nel libro di G. Fanello Marcucci, *Guido Gonella. Dal "discorso delle libertà" agli "appunti sulle istituzioni"*, Catanzaro, 2008, 129: «la nostra libertà, a differenza di quella del liberalismo, va intesa non nel senso negativo, cioè come una possibilità di restare chiusi nel proprio orticello; bensì in senso positivo, cioè come possibilità di uscire dalla nostra sfera particolare per portare un più vasto contributo al progresso della comunità. La libertà quindi non è un fosso, ma un

e di incremento del sistema democratico[38] all'interno di un'ottica finalistica del diritto tesa «a "costruire" una convivenza sociale»[39] di cui è parte e strumento proprio il rapporto socialmente orientato tra chi trasmette le informazioni e chi le attinge.

L'interpretazione della libertà d'informazione giornalistica, alla luce del personalismo, comporta un salto di qualità rispetto ad una interpretazione restrittiva dell'art. 21 Cost., basata sulla centralità della libertà di manifestazione del pensiero e, quindi, sull'informazione giornalistica come portato della visione soggettiva e della ricostruzione personale delle notizie da parte dei giornalisti. Il salto di qualità sta nel ritenere che, essendo l'informazione un presupposto della libertà di formazione del pensiero del terzo che riceve la notizia e della libertà di manifestazione del pensiero del medesimo, «non esiste la libertà di informare se non in quanto esista anche un diritto all'informazione»[40].

Secondo l'ottica sociale di tale visione, l'interesse collettivo al quale l'informazione giornalistica è improntata fa sì che la protezione della relazione tra operatore giornalistico e ricevente non possa andare nella sola direzione dei fini e degli interessi connessi al diritto di chi informa ma anche del diritto di chi attinge le informazioni. Ciò comporta che la libertà d'informazione giornalistica si configuri strutturalmente come una *libertà per*, nel senso di una libertà per *la collettività* e quindi di una libertà socialmente rilevante. La natura costitutivamente sociale della libertà di informazione non implica affatto che quest'ultima sia da considerarsi funzionalizzata nel senso di subordinata a perseguire contenuti, finalità e valori ad essa gerarchicamente superiori. Ciò sarebbe infatti incompatibile con la garanzia della libertà di manifestazione del pensiero e di informazione così come configurata in un ordinamento costituzionale di impronta liberaldemocratica, per il quale il diritto di libertà è in capo all'individuo e perciò in funzione solo di quest'ultimo[41].

---

ponte: non uno strumento di isolamento individualistico, ma un mezzo di espansione sociale».

38 P. Pombeni, *Il gruppo dossettiano*, in R. Ruffilli (a cura di), *Cultura politica e partiti nell'età della Costituente.* Tomo I: *L'area liberal-democratica. Il mondo cattolico e la Democrazia cristiana*, Bologna, 1979, 453.

39 P. Pombeni, *Il gruppo dossettiano*, cit., 438.

40 N. Lipari, *Libertà di informare o diritto ad essere informati?*, in *Diritto delle radiodiffusioni e delle telecomunicazioni*, 1978, 2.

41 N. Lipari, *Etica e professionalità del giornalista*, in *Diritto delle radiodiffusioni e delle telecomunicazioni*, 1982, 511. In chiave critica sul punto cfr. A. Pace, *Libertà di informare e diritto ad essere informati: due prospettive a confronto nell'interpretazione e nelle prime applicazioni dell'art. 7, primo comma, del t.u. della radiotelevisione*, in *Studi in onore di Nicolò Lipari*, Milano, 2008, 1995 ss. secondo il quale, in adesione ad una concezione individualistico-liberale del diritto di informazione giornalistica ancorato alla lettera dell'art. 21 Cost., «poiché il nostro ordinamento costituzionale non riconosce un'autonoma libertà d'informare distinta dalla libertà di manifestazione del pensiero (come invece accade, ad es., nell'ordinamento spagnolo), ne segue che, avendo la Corte costituzionale, nella scia della pressoché unanime dottrina costituzionalistica, desunto la libertà d'informare dall'art. 21 Cost., quest'ultimo diritto di libertà non può avere una struttura e una disciplina giuridico costituzionale diverse dalla struttura e dalla disciplina costituzionale del diritto «individualistico» di manifestare liberamente il proprio pensiero, proclamato nell'art. 21, nel cui capace alveo la libertà di cronaca e di informazione trova riconoscimento. Pertanto, se la libertà di manifestazione del pensiero ha la struttura di un diritto assoluto (in quanto situazione giuridica soggettiva «attiva» attribuita ai soggetti privati), anche la libertà d'informazione dovrà conseguentemente avere la stessa struttura. Sarebbe perciò contraddittorio - in tale ottica - dedurre dallo stesso enunciato normativo tanto un diritto di libertà dell'operatore dell'informazione quanto un contrapposto diritto dei destinatari ad una informazione obiettiva, imparziale e completa, in quanto

6. L'influenza del personalismo e del concetto di inviolabilità della libertà d'informazione sulla l. 69 del 1963

Del personalismo di Gonella risente emblematicamente anche l'art. 2 della l. 69 del 1963, la cui formulazione richiama per limpidezza il tenore stesso delle disposizioni della Costituzione, proponendosi – si potrebbe dire – come uno sviluppo normativo necessitato degli artt. 2 e 21 Cost.: l'insopprimibilità della libertà d'informazione e di critica dei giornalisti richiama il concetto di inviolabilità dei diritti della persona di cui all'art. 2 Cost.

Dire che un diritto di libertà è insopprimibile e/o inviolabile non significa affermare che il diritto in questione non possa venire limitato. Il problema costituzionale dei diritti ruota infatti attorno alla definizione dei casi, dei criteri, dei soggetti e dei modi mediante i quali i diritti medesimi «possono essere "violati"»[42], perché altri diritti da tutelare vengono in rilievo. Dire che il diritto è inviolabile e/o insopprimibile significa affermare che è il solo contenuto essenziale del diritto di libertà a non poter essere eliminato. Ciò implica che spetti al legislatore la determinazione del contenuto del diritto nel rispetto del limite insuperabile consistente nel fatto di non poter mai giungere ad una determinazione che si risolva «nell'impossibilità concreta di esercitare, in qualche modo, nell'attuale contesto storico, economico e sociale, un diritto evocato in Costituzione» [43].

Il contenuto dell'art. 2 l. 69 è speculare alle affermazioni di Gonella che individuano una serie di valori giuridici che fungono sia da limiti e sia da modi corretti di esercizio della libertà in oggetto. Si pensi alla verità oggettiva e putativa come presupposto logico ed etico-giuridico della libertà giornalistica: «il dovere di rispetto della verità è la condizione dell'esercizio del diritto di libertà di stampa»[44]. La condizione-presupposto di esercizio funge anche da limite immanente della libertà nel senso che «l'esigenza del limite non è arbitraria o aggiuntiva al concetto del diritto, ma implicita nel concetto stesso»[45]. Si può allora sostenere che l'attività giornalistica si caratterizzi e risolva in una endiadi, in quanto costituita da due prerogative coordinate, entrambe coessenziali, consistenti nel diritto di libertà e nella dimensione del dovere, che è il portato sociale della libertà: «il diritto è sempre connesso con un dovere: dovere del soggetto verso se stesso e doveri verso gli altri in connessione logica con il diritto degli altri (dovere di non invadere la sfera del diritto altrui)»[46]. E, più specificamente, «il diritto di cronaca, cioè il diritto di narrare pubblicamente fatti a mezzo stampa, è condizionato ai doveri

quest'ultimo diritto verrebbe a porre surrettiziamente dei limiti alla «libertà» di informare che la stessa disposizione garantisce».

42 Così M. Dogliani - I. Massa Pinto, *Elementi di diritto costituzionale*, Torino, 2017, 189.

43 Ivi, 192. P.F. Grossi, *Introduzione a uno studio sui diritti inviolabili nella Costituzione italiana*, Padova, 1969, 23 ss.

44 G. Gonella, *La libertà di stampa e i diritti individuali di libertà*, in M. Bellinetti (a cura di), *Guido Gonella, giornalista e politico*, Brescia, 2013, 162.

45 *Ibid.*

46 *Ibid.*

imposti dalla esigenza della tutela della libertà di ciascuno»[47].

7. I contenuti essenziali della legge professionale

Nell'intenzione del legislatore del 1963 gli istituti funzionali ad una informazione giornalistica rispettosa dei contenuti dell'art. 2 sono molteplici. In primo luogo, emerge la disciplina "ordinistica" del giornalismo, consistente appunto nell'istituzione dell'*Ordine dei giornalisti* come ente di diritto pubblico atto a definire l'autonomia del giornalista, in modo che sia la categoria giornalistica a determinare compiutamente e democraticamente le proprie scelte, libera da imposizioni e da pressioni esterne. La scelta ordinistica implica, per coloro che dopo il periodo di compiuta pratica intendono diventare giornalisti professionisti, il superamento dell'esame di Stato. Ora, la *ratio* di un esame di Stato organizzato dall'Ordine sta proprio nello svincolare, almeno parzialmente, la disciplina giuridica dell'accesso alla professione giornalistica dal condizionamento del potere economico editoriale. Tale finalità è inoltre perseguita dall'art. 34 l. 69 del 1963, che attribuisce al direttore responsabile della pubblicazione, non all'editore, il compito della certificazione del praticantato.

Come affermava Gonella la *ratio* sottesa all'istituzione dell'Ordine sta nel garantire l'autonomia del giornalista e quella piena libertà d'informazione e di critica alla quale si riferisce l'art. 2 della l. 69 del 1963. L'insopprimibilità della libertà si lega proprio al fatto che «non lo Stato, ma i giornalisti sono delegati ad abilitare i giornalisti alla loro professione. Questo è il carattere peculiare del nostro esame che è di Stato, ma che è, per volontà dello Stato, affidato alla professione»[48]. Al risultato cospira anche il carattere professionistico (professionisti) o comunque non sporadico dell'attività giornalistica (pubblicisti)[49].

Infine, uno specifico rilievo viene attribuito alla *responsabilità disciplinare* cui sono soggetti «gli iscritti nell'albo, negli elenchi o nel registro, che si rendano colpevoli di fatti non conformi al decoro e alla dignità professionali, o di fatti che compromettano la propria reputazione o la dignità dell'ordine»[50] con la previsione di una serie di sanzioni graduate alla luce della gravità dell'infrazione compiuta (avvertimento, censura, sospensione e radiazione dall'albo)[51].

L'importanza della funzione disciplinare, in riferimento al ruolo e alla legittimazione dell'Ordine dei giornalisti, è stata ulteriormente valorizzata con il regolamento di delegificazione d.P.R. n. 137 del 2012 in tema di ordinamenti professionali, che all'art. 8 ha reso autonomo il giudizio deontologico demandato ai Consigli degli Ordini. Infatti, il regolamento in questione ha istituito organismi speciali come titolari esclusivi della funzione diretta all'accertamento della responsabilità disciplinare degli iscritti – i *Consi-*

[47] G. Gonella, *La libertà di stampa e i diritti individuali di libertà*, cit., 161. Cfr. D. Messinetti, *Personalità (diritti della)*, in *Enciclopedia del diritto*, Vol. XXXIII, Milano, 1983, 355 ss.

[48] M. Bellinetti, *Guido Gonella, giornalista e politico*, Brescia, 2013, 59.

[49] Art. 1 l. 69 del 1963.

[50] Art. 48 l. 69 del 1963.

[51] Cfr., rispettivamente, artt. 52, 53, 54, 55 l. 69 del 1963.

gli territoriali di disciplina e il *Consiglio nazionale di disciplina* – realizzando così una separazione assoluta all'interno dell'Ordine tra funzioni amministrative, quale, ad esempio, la tenuta degli albi professionali, e funzioni disciplinari.

La legge 3 febbraio 1963, n. 69 è stata oggetto di valutazione di legittimità costituzionale da parte sia della dottrina sia della giurisprudenza. In particolare, le riflessioni si sono incentrate sulla compatibilità con l'art. 21 Cost., che attribuisce a tutti il diritto alla libera manifestazione del pensiero, di quelle disposizioni di legge che affidano all'Ordine dei giornalisti la tenuta dell'albo, disciplinandone struttura e funzionamento.

La Corte costituzionale con la nota sentenza n. 11 del 1968, nel rigettare le questioni di legittimità costituzionale della legge n. 69 n. 63, ha svolto le considerazioni più rilevanti proprio sulla funzione del giudizio disciplinare. La *ratio* più significativa dell'Ordine dei giornalisti, ad avviso della Corte, sta infatti nella funzione di vigilanza sulla dignità professionale degli iscritti, al fine di evitare che essi vedano limitato e compromesso l'esercizio della libertà di informazione. Inoltre, il fatto che «i giornalisti vengano associati in un organismo che, nei confronti del contrapposto potere economico del datori di lavoro, possa contribuire a garantire il rispetto della loro personalità e, quindi, della loro libertà»[52] opera a tutela dell'interesse della collettività. In altri termini, l'interesse della società ad una informazione libera e verificata impone la sussistenza di istituti giuridici che assicurino il rispetto delle regole di dignità professionale, regole che implicano che non si abdichi «mai alla libertà di informazione e di critica e nel non cedere a sollecitazioni che possano comprometterla»[53].

La responsabilità disciplinare, anche secondo la giurisprudenza della Corte costituzionale, mediante il giudizio sul rispetto da parte degli iscritti delle regole di correttezza professionali (ambito autonomo rispetto alla responsabilità del giornalista per illeciti civili e penali)[54], enfatizza gli interessi/diritti dei lettori a ricevere una informazione giornalistica che rispetti le regole dell'etica professionale e, quindi, i limiti che la professione dà a se stessa negli ambiti più disparati (minori, malati, detenuti, giustizia, ecc.), affinché essa non ceda rispetto agli obiettivi di qualità, essenzialità e correttezza che devono caratterizzarla, al fine di garantire i diritti della personalità dei consociati (*privacy*, immagine, identità, onore, reputazione, oblio, ecc.). Perciò deve negarsi «che la sanzione disciplinare dia luogo ad una limitazione della libertà del giornalista, in quanto egli ha già oltrepassato un limite alla libertà di cronaca»[55].

Per concludere, va osservato che la funzione disciplinare caratterizza il giornalismo

[52] Corte cost., 23 marzo 1968, n. 11, cons. dir. n. 5. Cfr. G. Zagrebelsky, *Questioni di legittimità costituzionale della l. 3 febbraio 1963 n. 69, istitutiva dell'ordine dei giornalisti*, in *Giurisprudenza costituzionale*, 1968, 330 ss.

[53] Corte cost., 23 marzo 1968, n. 11, cons. dir. n. 5. Cfr. in chiave critica A. Pugiotto, *L'*ordine irrazionale. *L'ordine dei giornalisti nella giurisprudenza costituzionale*, in A. Pizzorusso et al. (a cura di), *Libertà di manifestazione del pensiero e giurisprudenza costituzionale: dottorato di ricerca in Giustizia costituzionale e diritti fondamentali: terze giornate italo-spagnole di giustizia costituzionale: Lipari, 1-2 ottobre 2004*, Milano, 2005, 190: «nulla può l'Ordine nei confronti dell'editore, perché l'editore *non è* un iscritto all'Ordine ed è dunque estraneo ai suoi poteri di vigilanza ed alla sua giurisdizione domestica». Cfr. F. de Bortoli, *Ordine, 60 anni. La trappola della censura educativa*, cit., 12.

[54] Cfr. V. Tenore, *La responsabilità disciplinare del giornalista*, in Id. (a cura di), *Il giornalista e le sue quattro responsabilità*, Milano, 2018, 2 ss.

[55] M. Pedrazza Gorlero, *Giornalismo e giornalisti nella società pluralista: profili costituzionali*, in *Saggi per un corso di diritto dell'informazione giornalistica*, Padova, 2006, 58.

in tutti gli ordinamenti giuridici improntati alla forma di stato democratico-costituzionale indipendentemente dalla sussistenza o meno di un Ordine professionale. La vigilanza sulla responsabilità deontologica dei *media* è infatti garantita in molti Stati o dalle associazioni sindacali della stampa o dalle stesse aziende editoriali o da associazioni private (*club*) di giornalisti[56].

8. La natura costituzionalmente indefettibile della responsabilità disciplinare

Sotto l'aspetto giuridico la funzione disciplinare affonda le proprie radici in via generale nell'art. 2 della l. 69 del 1963 e, in via più articolata e circostanziata, nel Testo unico dei doveri del giornalista in vigore dal 1° gennaio 2021, documento che «nasce dall'esigenza di armonizzare i precedenti documenti deontologici al fine di consentire una maggiore chiarezza di interpretazione e facilitare l'applicazione di tutte le norme, la cui inosservanza può determinare la responsabilità disciplinare dell'iscritto all'Ordine»[57].
Al riguardo è opportuno allargare lo sguardo ad un quadro costituzionale più complessivo. Il 15 giugno 1997 si tenne un *referendum* abrogativo volto all'abolizione integrale della l. 3 febbraio 1963, n. 69, quesito ritenuto ammissibile dalla Corte costituzionale nella sentenza n. 38 del 1997[58].Come si concilia allora tale decisione con una giurisprudenza costituzionale più risalente che, come abbiamo visto, ha invece ritenuto conforme al dettato costituzionale l'Ordine dei giornalisti? È necessaria una precisazione sul punto. La giurisprudenza costituzionale ha sì ritenuto che la disciplina normativa istitutiva dell'Ordine fosse compatibile con la Costituzione, ma al tempo stesso non ha mai sostenuto che la materia così come normata dalla l. 69 del 1963 fosse da considerarsi sottratta a referendum abrogativo. Infatti, la legge non va ritenuta disciplina attuativa di un organismo costituzionalmente indefettibile, non essendo l'Ordine dei giornalisti menzionato in Costituzione. Inoltre, la legge professionale in questione non è ascrivibile neppure alle fonti ordinarie a contenuto costituzionalmente vincolato, in quanto non dispone di una materia il cui nucleo normativo non possa venire alterato o privato di efficacia, senza che ne risultino lesi i corrispondenti specifici disposti della Costituzione stessa o di altre leggi costituzionali.
Ciò comporta che la materia *de qua*, la costituzione dell'Ordine dei giornalisti, vada

[56] Per una disamina della regolamentazione della professione giornalistica nei Paesi dell'Europa occidentale cfr. R. Razzante, *Manuale di diritto dell'informazione e della comunicazione. I* media *nell'era digitale e le nuove tutele della persona*, nona edizione, Milano, 2022, 80 ss.

[57] Il Testo unico dei doveri del giornalista, approvato dal Consiglio nazionale dell'Ordine dei giornalisti il 22 gennaio 2019, recepisce i contenuti dei seguenti documenti: Carta dei doveri del giornalista; Carta dei doveri del giornalista degli Uffici stampa; Carta dei doveri dell'informazione economica; Carta di Firenze; Carta di Milano; Carta di Perugia; Carta di Roma; Carta di Treviso; Carta informazione e pubblicità; Carta informazione e sondaggi; Codice di deontologia relativo alle attività giornalistiche; Codice in materia di rappresentazione delle vicende giudiziarie nelle trasmissioni radiotelevisive; Decalogo del giornalismo sportivo.

[58] Il *quorum* di partecipazione non venne raggiunto, in quanto il numero di votanti, complessivamente 14.735.975, fu pari al 30,04 % degli aventi diritto, Si deve evidenziare che il 65,52 % dei votanti optò per il sì all'abrogazione della l. 69 del 1963.

ritenuta contenuto disponibile da parte della discrezionalità del legislatore ordinario. Tuttavia si deve osservare come una eventuale soppressione per legge o per *referendum* abrogativo dell'Ordine dei giornalisti non eliminerebbe la necessità *costituzionalmente rilevante* che gli operatori dell'informazione giornalistica siano sottoposti, in altre forme rispetto a quelle indicate dalla l. 69 del 1963 e in capo ad organismi diversi dall'Ordine, a giudizio sul rispetto delle regole di correttezza professionale. In tal senso la Corte costituzionale ha infatti esplicitato che una eventuale abolizione dell'Ordine disposta da *referendum* abrogativo non sarebbe suscettibile di far «venir meno l'attività giornalistica professionale, la disciplina contrattuale del rapporto di lavoro, o i canoni deontologici inerenti a tale attività. Questi ultimi derivano, oltre che dal costume, da altre leggi (cui del resto fa rinvio lo stesso art. 2), dalle funzioni del Garante, dalla giurisprudenza in materia e da forme di autoregolamentazione»[59].

Dagli argomenti adoperati dai giudici costituzionali si può allora trarre una riflessione. La materia della deontologia giornalistica e quindi della effettività dei controlli, da parte di organismi a ciò deputati, sul rispetto dell'etica professionale dei giornalisti, trattandosi di un ambito strumentale alla difesa della struttura della libertà di informazione, della qualità dell'informazione e dei diritti all'informazione dell'utenza, può ritenersi *contenuto costituzionalmente vincolato*, indipendentemente dalle forme istituzionali e organizzative che ne traducono l'applicazione. Pertanto ci si può *liberare* dell'Ordine dei giornalisti per scelta di politica legislativa, ma nessuna norma di legge è costituzionalmente autorizzata ad affrancare la disciplina del giornalismo dall'esistenza di controlli sulla deontologia giornalistica. Spetta poi alla discrezionalità politico-legislativa la decisione sulle forme tramite le quali la deontologia debba venire fatta valere.

Il fatto che la materia deontologica sia contenuto indisponibile per il legislatore è dovuto a due ragioni: 1) la natura intrinsecamente sociale dell'informazione giornalistica che incrocia gli interessi e i diritti dei terzi in quanto libertà per la collettività; 2) la necessità che sia garantita la qualità dell'informazione che serve proprio a proteggere i diritti della personalità dei terzi. L'effettività della responsabilità deontologica presidia sia la libertà di informazione il cui contenuto include quella natura sociale, che rimanda al rapporto emittente—utente, sia i diritti dei terzi, i quali possono venire conculcati da notizie scorrettamente diffuse sia per contenuto sia per modalità espositive[60].

Le argomentazioni esposte sono del resto riferibili a come l'art. 2 della l. 69 del 1963 definisce il contenuto essenziale, e quindi la struttura, della libertà di informazione e di critica dei giornalisti, contenuto nel quale la libertà è limitata dai seguenti beni giuridici: i diritti *della personalità*, *il rispetto della verità sostanziale dei fatti* e i doveri di *lealtà e buona fede*. Potremmo dire che i diritti, i valori e i limiti di cui sopra, sostanziando proprio la responsabilità dei media, entrino a far parte del contenuto essenziale insopprimibile della libertà di informazione e di critica del giornalista. Si tratta di elementi immanenti del diritto di informare che ne connotano l'area di insopprimibilità.

La deontologia professionale è quindi istituto giuridico funzionale a presidiare la protezione di quel contenuto essenziale. Qui vi è l'indefettibilità costituzionale della deontologia.

[59] Corte cost., 8 gennaio 1997, n. 38, cons. dir. n. 3.

[60] Cfr. M. Partipilo, *Manuale di deontologia del giornalista. Informazione, disinformazione, società*, Roma, 2022.

9. Il segreto professionale sulla fonte delle notizie

I nodi attuali che la legge professionale presenta rispetto all'evoluzione *tecnica* del giornalismo e, più in generale, della comunicazione di massa, sono numerosi. I più rilevanti concernono: a) le modalità soggettive e oggettive di applicazione del segreto professionale; b) la disciplina della rettifica; c) la regolamentazione della pratica giornalistica; d) il rapporto tra il giornalismo e la comunicazione *online*. Il segreto sulla fonte fiduciaria delle notizie è previsto dall'art. 2, c. 3, l. 69 del 1963. La norma ha un ambito di applicazione "sostanziale" molto ampio, dal momento che nell'opponibilità del segreto parifica ai giornalisti gli editori e non opera alcuna distinzione soggettiva fra i giornalisti, cosicché il segreto potrebbe venire opposto anche dal pubblicista, che è giornalista ma non esercita professionalmente il giornalismo, e dal praticante, che esercita attività giornalistica in vista di un inserimento professionale, ma non è ancora un giornalista.
Com'è noto, però, la proiezione processuale penale del segreto è più restrittiva, posto che l'art. 200, c. 3, c.p.p. dispone l'opponibilità del segreto ad opera del solo giornalista professionista. Si tratta di una restrizione discutibile e discussa, in quanto è la copertura con il segreto professionale a generare la concreta possibilità di accedere alle fonti; una restrizione, quindi, che è suscettibile di vanificare l'effettività dell'informazione.
Inoltre, il fatto che vengano esclusi praticanti e pubblicisti dalla garanzia di avvalersi dell'opponibilità al giudice del segreto sull'identità della fonte pone un problema di illegittimità costituzionale per sospetta violazione del principio di eguaglianza-ragionevolezza del trattamento normativo. Il problema di legittimità costituzionale inerisce al fatto che l'attività di elaborazione di notizie di interesse pubblico per la collettività può accomunare tutti i soggetti del giornalismo – professionisti, pubblicisti e praticanti – ciò comportando che la limitazione della garanzia del segreto sulla fonte in capo solo al giornalista professionista sguarnisce di una ineludibile tutela la libertà di informazione e di critica di pubblicisti e praticanti.
Se infatti la credibilità della fonte deve essere scrupolosamente verificata da chi svolge attività giornalistica, a pena della violazione del criterio di verità putativa della notizia, il riserbo sull'identità della fonte, che pretenda l'anonimato, è essenziale alla libertà di non essere ostacolati nella ricerca delle notizie, ciò valendo sia per il giornalista che esercita la libertà di informazione sia per la società degli utenti che non deve essere impedita nella possibilità di fruire delle informazioni. In tale senso il segreto sulla fonte è coessenziale sia alla libertà di informare sia alla libertà di essere informati.
Tuttavia, la disciplina giuridica dell'istituto del segreto sulla fonte è una spia emblematica della capacità della l. 69 del 1963 di essere al passo con i tempi della moltiplicazione dei mezzi di comunicazione consentiti dall'evoluzione tecnica della rete. Il tema del segreto sulla fonte informativa è, infatti, al centro di un *perenne e sistematico conflitto tra la Corte europea dei diritti dell'uomo e il diritto interno*. Infatti, i giudici di Strasburgo hanno consolidato una giurisprudenza che, attraverso la garanzia dell'anonimato della fonte, ha progressivamente allargato l'area di tutela della libertà d'informazione facendo sostanzialmente coincidere la salvaguardia sostanziale con quella processuale del segreto e senza distinguerne soggettivamente i titolari.
Ci si riferisce ad una serie di *sentenze storiche* che, adottando una interpretazione iper-ga-

rantistica dell'art. 10 Convenzione europea dei diritti dell'uomo[61], hanno sancito come anche lo stagista[62] e perfino il dipendente amministrativo dell'azienda editoriale[63] siano titolari del diritto al segreto sulla fonte delle notizie. Chiunque partecipi, in qualsiasi momento e a qualsiasi titolo, del flusso delle informazioni è, secondo l'interpretazione dei giudici di Strasburgo, beneficiario della garanzia del segreto sulla fonte perché in gioco vi è l'interesse e il diritto a ricevere informazioni, senza ostacoli, da parte della società. Per i giudici della Corte europea dei diritti dell'uomo sembra quindi che chiunque svolga un'attività informativa in senso sostanziale possa giovarsi dell'opponibilità del segreto sulla fonte, fermi restando, ovviamente, i limiti che la magistratura potrà far valere quando, ad esempio, la rivelazione della fonte sia essenziale per l'accertamento della prova di un reato.

La giurisprudenza della Corte europea dei diritti dell'uomo in merito all'operatività del segreto sulla fonte non investe soltanto i profili soggettivi ma anche i profili che attengono all'ambito sostanziale di applicazione della libertà di informazione Si pensi al fatto che i giudici di Strasburgo affermano che la garanzia del segreto del giornalista sulla fonte vada riconosciuta anche qualora vi sia una legge nazionale che imponga per determinate fattispecie al cronista la rivelazione della fonte quando essa porti all'identificazione dell'autore di un reato[64]. Si deve sottolineare che, indipendentemente dalla disciplina legislativa nazionale, residua sempre il margine di controllo da parte della Corte europea dei diritti dell'uomo diretto a vagliare se l'obbligo di rivelare l'identità della fonte sia supportato da una esigenza imperativa di interesse pubblico[65]. I giudici di Strasburgo pervengono a valutare le circostanze concrete del caso come ad esempio la gravità concreta del reato che dà luogo all'indagine. Ciò significa che la legge nazionale possa venire giudicata alla luce delle esigenze regolative del caso concreto.

Come evidenziato sopra la relazione tra la disciplina sul segreto sulla fonte di cui all'art. 200 del codice di procedura penale e la giurisprudenza della Corte europea dei diritti dell'uomo dà luogo ad un conflitto che ha *riflessi sul sistema delle fonti*, dal momento che l'art. 117, c. 1, Cost. vincola le leggi nazionali al rispetto degli obblighi internazionali dello Stato. Tale disposizione implica che anche le giurisdizioni nazionali debbano osservare le interpretazioni dei trattati internazionali provenienti dai giudici internazionali, quando esse non confliggano con le norme costituzionali nazionali[66]. A questa

---

[61] Cfr. G.E. Vigevani, *La protezione del segreto del giornalista al tempo di internet*, in *costituzionalismo.it*, 3, 2011.

[62] CEDU, *Goodwin c. Regno Unito*, ric. 17488/90 (1996).

[63] CEDU, *de Haes and Gijsels v. Belgium*, ric. 19983/92 (1997).

[64] CEDU, *Affaire Jecker c. Suisse,* ric. 35449/14, (2020). Cfr. M. Castellaneta, *Segretezza delle fonti giornalistiche nel quadro della CEDU. Una nuova pronunzia della Corte di Strasburgo (Jecker c. Svizzera)*, in *giustiziainsieme.it,* 14 novembre 2020.

[65] CEDU, *Affaire Jecker c. Suisse*, cit., § 41.

[66] Cfr. P. Carrozza, *Tradizioni costituzionali comuni, margine di apprezzamento e rapporti tra Corte di giustizia delle comunità europee e Corte europea dei diritti dell'uomo. Quale Europa dei diritti?*, in P. Falzea - A. Spadaro - L. Ventura (a cura di), *La Corte costituzionale e le Corti d'Europa*, Torino, 2003, 574-575; F. Patroni Griffi, *Il ruolo delle Corti nella costruzione dell'ordinamento europeo (From judge-made law to judge-made Europe)*, in.*federalismi.it*, 15, 2019, 9; S. Cassese, *Verso un diritto europeo italiano*, in *Rivista italiana di diritto pubblico comunitario*, 2017, 303; R. Conti, *La CEDU assediata? (Osservazioni a Corte cost. sent. n. 49/2015)*, in *Consulta online*, 10 aprile 2015; D. Russo, *Ancora sul rapporto tra Costituzione e Convenzione europea dei diritti dell'uomo: brevi note sulla sentenza della Corte costituzionale n. 49 del 2015*, in *osservatoriosullefonti.it,* 2, 2015; D. Tega, *La sentenza*

linea si è solo parzialmente avvicinato un disegno di legge (*XVIII, Senato, n. 836*) che prevederebbe l'applicabilità anche ai giornalisti pubblicisti delle disposizioni sul segreto sulla fonte, emendando il c. 3 dell'art. 200 del codice di procedura penale[67]. Il disegno di legge invece non contemplava i giornalisti praticanti come soggetti titolari del segreto sulla fonte, con ciò consolidando il conflitto tra le garanzie predisposte dal diritto interno e quelle invece riconosciute dalla Corte europea dei diritti dell'uomo. A ciò si aggiunge un ulteriore profilo al quale prestare attenzione.

Il problema attiene al fatto che anche il disegno di legge in questione parrebbe non adeguarsi ad un sistema della comunicazione di massa sempre più caratterizzato da soggetti non ascrivibili alle categorie del giornalismo tradizionale. Per di più siamo dinanzi ad un sistema della comunicazione di massa sempre più decentrato i cui attori non hanno obblighi di responsabilità deontologica nei confronti degli utenti: si pensi ai contenuti informativi veicolati generalmente su *Internet*, o tramite siti, social media, blog, web tv ecc. La disciplina giuridica del segreto sulla fonte è una spia significativa di una differenziazione, per quanto riguarda le tutele giuridiche, tra quello che è l'ordinamento giuridico del giornalismo propriamente detto al quale la garanzia suddetta è ascrivibile, e il mare *magnum* di una comunicazione di massa sulla quale sempre di più vanno a formarsi i convincimenti sociali, malgrado i suoi attori e soggetti non si possano avvalere dell'istituto del segreto[68].

10. La rettifica nella stampa e nella radiotelevisione

Prevista dall'art. 2, c. 2, l. 69, la rettifica è un istituto che salda diverse discipline: la disciplina ordinaria della professione giornalistica, la disciplina legislativa del corretto uso dei mezzi di diffusione (stampa e radiotelevisione, ai sensi della l. 47 del 1948 e del

della Corte costituzionale n. 49 del 2015 sulla confisca: il predominio assiologico della Costituzione sulla Cedu, in *Forum di Quaderni costituzionali*, 30 aprile 2015; G. Sorrenti, *Sul triplice rilievo di Corte cost., sent. n. 49/2015, che ridefinisce i rapporti tra ordinamento nazionale e CEDU e sulle prime reazioni di Strasburgo*, ivi, 7 dicembre 2015. Cfr. anche Corte cost., sent. n. 49/2015, cons. dir. n. 4 e n. 7; Corte cost., sent. n. 264/2012, cons. dir. n. 5.4. Cfr. A. Ruggeri, *La Consulta rimette abilmente a punto la strategia dei suoi rapporti con la Corte EDU e, indossando la maschera della consonanza, cela il volto di un sostanziale, perdurante dissenso nei riguardi della giurisprudenza convenzionale ("a prima lettura" di C. Cost. n. 264 del 2012)*, in *Consulta online*, 17 dicembre 2012. Cfr. M. Patrono, *Lezione n. 16. La «fontana di* Bucket*»*, in Id. (a cura di), *Studiando i diritti. Il costituzionalismo sul palcoscenico del mondo dalla* Magna Charta *ai confini del (nostro) tempo. Lezioni*, Torino, 2009, 177.

[67] Il disegno prevede che il segreto sulla fonte si applica «ai giornalisti professionisti e pubblicisti, iscritti nei rispettivi elenchi dell'albo professionale, relativamente ai nomi delle persone dalle quali i medesimi hanno avuto notizie di carattere fiduciario nell'esercizio della loro professione».

[68] Cfr. l'intervento di C. Malavenda, in *Quale futuro per il giornalismo?*, in V. Roidi (a cura di), *Quaderni della Fondazione Paolo Murialdi*, Padova, 2019, 41: «il segreto professionale è il vero motore dell'informazione e anche quello del giornalista andrebbe tutelato in modo assoluto, come accade per altre categorie di professionisti che ne godono». Secondo l'autrice (41-42) il segreto sulla fonte andrebbe esteso «anche a chi, pur non essendo giornalista, svolge attività informativa e si avvale di fonti riservate … In questo campo, dunque, ci vuole una estensione del segreto a chiunque si avvalga delle fonti, secco e senza eccezioni, perché almeno quando salvaguarda le fonti il giornalista non deve correre rischi. E ci vogliono regole chiare per fronteggiare perquisizioni e sequestri, perché se ti possono sequestrare la memoria del cellulare o quella del computer, le fonti non ti parlano più. Occorre una norma che lo vieti o che fornisca le stesse garanzie previste per gli altri professionisti».

d. lgs n. 177 del 2005) e la disciplina deontologica. Di più; il tema della rettifica incrocia oggi l'evoluzione dei *media* e la necessità che l'obbligo di rettifica delle notizie lesive dei diritti della personalità e/o contrarie alla verità sostanziale dei fatti si applichi anche a quei *media* che sotto il profilo formale non vi sarebbero tenuti. La rettifica risponde pienamente alla natura etico-sociale del giornalismo così come ispirata dai principi della l. 69 del 1963. Essa infatti soddisfa sia l'esigenza di riattivazione della tutela di un diritto della personalità conculcato sia quella del ripristino della correttezza di una notizia originariamente non corrispondente a verità (per incompletezza, integrale falsità, omissione, ecc.), ciò configurandosi come un istituto strumentale ad assicurare il diritto all'informazione.

La rettifica, in quanto correzione di una informazione precedentemente trasmessa o in quanto lesiva di diritti altrui o perché non corrispondente a verità sostanziale, costituisce un rimedio preventivo che prescinde dalla sussistenza della responsabilità penale e/o civile del giornalista e quindi dall'accertamento del dolo, della colpa, della misura del danno cagionato e risarcibile.

La disciplina normativa della rettifica della notizia è diversa a seconda del mezzo informativo. Per la stampa il riferimento è all'art. 8 della legge n. 47 del 1948 secondo il quale «il direttore o, comunque, il responsabile è tenuto a fare inserire gratuitamente nel quotidiano o nel periodico o nell'agenzia di stampa le dichiarazioni o le rettifiche dei soggetti di cui siano state pubblicate immagini od ai quali siano stati attribuiti atti o pensieri o affermazioni da essi ritenuti lesivi della loro dignità o contrari a verità, purché le dichiarazioni o le rettifiche non abbiano contenuto suscettibile di incriminazione penale»[69]. Per il mezzo della stampa la rettifica mira, come *ratio*, alla tutela del diritto della personalità del soggetto attinto dall'informazione suscettibile di correzione. Essa ha, infatti, il fine di mettere riparo agli effetti di dilatazione e di moltiplicazione della lesione derivante dal mezzo di diffusione. Tale fine viene conseguito mediante l'ingresso tempestivo ed evidenziato di un soggetto estraneo all'organo informativo che espone la propria verità. A questo proposito si parla di *rettifica esterna*, attivabile su richiesta del soggetto attinto da un'informazione in conflitto con la verità e/o con i suoi diritti.

Il fatto che la rettifica nella stampa sia attivabile in caso di un contenuto pubblicato lesivo della dignità della persona, anche se nella sostanza corrispondente a verità, fa sì che la funzione sia principalmente la difesa dei diritti dei terzi, sfumando invece la funzione più propriamente informativa in termini di verità e completezza del contenuto. Pertanto la *ratio* giuridica della rettifica a mezzo stampa, consistendo principalmente nella tutela dei diritti dei soggetti attinti dalle informazioni, implica il diritto di esprimere, da parte del soggetto richiedente la rettifica, una propria ricostruzione dei fatti, attivando un tentativo di pluralismo informativo rispetto alla versione fornita dalla testata[70].

[69] La disposizione prevede a seconda della diversa periodicità dell'organo termini diversi entro i quali la rettifica deve venire pubblicata.

[70] Cass. civ., sez. III, 27 gennaio 2015 n. 1436. Cfr. M. Manetti, *La tutela contro gli abusi della libertà di manifestazione del pensiero*, in *Art. 21. Rapporti civili. La libertà di manifestazione del pensiero. Commentario della Costituzione*, di A. Pace, M. Manetti, fondato da G. Branca e continuato da A. Pizzorusso, Bologna-Roma, 2006, cit., 793 s.: «la possibilità che il messaggio di rettifica non rispecchi la verità dei fatti, e quindi la vera identità del soggetto, è stata considerata un rischio inevitabile, connesso alla necessità di assicurare un "legittima difesa" al soggetto che si ritiene leso. In quest'ordine di idee la rettifica

In altre parole, la disciplina della rettifica a mezzo stampa non assegna all'istituto in questione la natura di strumento di arricchimento notiziale al fine di assicurare correttezza e completezza dell'informazione.

La lesione dell'interesse individuale del soggetto è autonomo ed esclusivo motivo di attivazione della rettifica. Pertanto, la violazione dell'interesse individuale è il presupposto che stimola l'esercizio della libertà di manifestazione del pensiero nell'interesse generale dei lettori, senza che il messaggio rettificato debba essere funzionale al ristabilimento della verità sostanziale dei fatti. Perciò, la cifra della rettifica a mezzo stampa è nella funzione riparatoria «finalizzata a non lasciare spazio a un danno ulteriormente risarcibile, che tuttavia non elimina l'evento di danno per gli effetti in precedenza già perfezionati … Tale diritto costituisce un'attività discrezionale dell'interessato, e non può mai assurgere a una sorta di dovere»[71].

Più precisamente la connotazione riparativa della rettifica a mezzo stampa mette inoltre la persona nella condizione concreta di controllare i contenuti informativi, a garanzia di una rappresentazione dinamica e sociale della personalità e dell'immagine che i media di informazione giornalistica devono rispettare. Di natura giuridica diversa è invece la rettifica radiotelevisiva, la cui disciplina accomuna sia i mezzi di servizio pubblico sia le emittenti radiotelevisive private. La disposizione di riferimento è l'art. 35, c. 2, d.lgs. 8 novembre 2021, n. 208, che stabilisce quanto segue: «chiunque si ritenga leso nei suoi interessi morali, quali in particolare l'onore e la reputazione, o materiali da trasmissioni contrarie a verità ha diritto di chiedere al fornitore di servizi di media audiovisivi e radiofonici, ivi inclusa la concessionaria del servizio pubblico radiofonico, televisivo e multimediale, all'emittente radiofonica oppure alle persone da loro delegate al controllo della trasmissione, che sia trasmessa apposita rettifica, purché quest'ultima non abbia un contenuto che possa dar luogo a responsabilità penali».

La peculiarità della rettifica radiotelevisiva consiste nel fatto che la realizzazione dell'interesse individuale connesso alla lesione di diritti morali o patrimoniali del soggetto è condizionata al fine di riaffermare la verità oggettiva dei fatti: lesione di diritti e contenuto non rispondente a verità devono sussistere entrambi per l'attivazione della rettifica. Emerge pertanto la funzione oggettiva e sociale del ripristino di un contenuto informativo improntato a verità: si tratta dello scopo di fornire un contributo oggettivo di «"arricchimento notiziale"» affidato «al singolo in vista della correttezza e completezza dell'informazione»[72]. Del resto la giurisprudenza costituzionale quando affermò il diritto alla rettifica come diritto fondamentale dell'uomo[73] connetteva tale prerogati-

consiste infatti nel diritto di fornire una propria verità ovvero una diversa ricostruzione dei fatti a difesa della propria immagine sociale, *con l'effetto* di soddisfare nel contempo l'interesse pubblico al pluralismo dell'informazione. L'ottica […] sconta il fatto che il messaggio istituisca un contraddittorio in ordine alla narrazione di determinati fatti, ma non lo eleva a sintomo di una autonoma libertà di informazione del privato».

[71] Cass. civ., sez. III, n. 1436/2015, cit.

[72] M. Manetti, *La tutela contro gli abusi della libertà di manifestazione del pensiero*, cit., 793: «la lesione dell'interesse individuale verrebbe in sostanza utilizzata dalla legge per stimolare l'esercizio della libertà di manifestazione del pensiero nell'interesse generale dei lettori».

[73] Corte cost., 10 luglio 1974, n. 225, cons. dir. n. 8.

va alla *ratio* di conseguire l'interesse pubblico all'obiettività dell'informazione[74].
Pertanto, la concezione giuridica della rettifica secondo i giudici costituzionali pare improntarsi alla disciplina della rettifica radiotelevisiva piuttosto che alla disciplina concernente la rettifica della carta stampata. La Corte costituzionale ritiene che la funzione della rettifica stia nel proteggere i diritti dei terzi ad una informazione completa, obiettiva e verificata nei contenuti. È però nell'ambito disciplinare che la rettifica rivela l'efficacia più incisiva, dal momento che la sua applicazione prescinde dalla richiesta dell'interessato (opera cioè come *rettifica interna*), dovendosi attivare d'ufficio da parte della testata giornalistica quando essa pubblica un contenuto informativo lesivo dei diritti della persona e/o non corrispondente a verità sostanziale. Ne consegue che la mancata attivazione d'ufficio della rettifica da parte della testata costituisce illecito disciplinare, dal momento che aggrava ulteriormente il danno che la notizia scorretta produce sia in termini di diritti della personalità del soggetto attinto dalla pubblicazione sia in termini di interesse dell'utenza tutta a ricevere una informazione di qualità, corretta nelle forme espositive e verificata nei contenuti. Pertanto si potrebbe dire che la garanzia massima in termini di controllo sulla qualità dell'informazione e, quindi, di tutela della società, si ha nell'ambito del giudizio disciplinare, al punto che in tale sede è giustiziabile perfino la violazione del dovere di rettifica in assenza di richiesta della persona interessata.
Si veda, a tale proposito, il *Testo Unico dei doveri del giornalista* secondo il quale «il giornalista, rettifica, anche in assenza di specifica richiesta, con tempestività e appropriato rilievo, le informazioni che dopo la loro diffusione si siano rivelate inesatte o errate» (titolo III, art. 9, par. 1.). Inoltre «il giornalista corregge senza ritardo errori e inesattezze, anche in conformità al dovere di rettifica nei casi e nei modi stabiliti dalla legge» (Allegato 1, art. 4). Nello stesso senso è *la giurisprudenza disciplinare dell'Ordine*, secondo la quale la rettifica deve essere compiuta con tempestività e appropriato rilievo, anche in assenza di specifica richiesta, quando le informazioni dopo la loro diffusione si siano rivelate inesatte o errate, soprattutto quando l'errore possa ledere o danneggiare singole persone, enti, categorie, associazioni o comunità[75].
Emerge anche in ciò un portato di quel personalismo che connota la funzione etico-sociale del giornalismo che ha inciso sulla natura giuridica di quest'ultimo, ispirando la redazione della l. 69 del 1963. In particolare emerge una concezione giuridica nella quale il giornalismo, oltreché essere espressione di un diritto di libertà, viene visto anche e soprattutto come «esercizio di un servizio sociale che contribuisce ad emancipare intellettualmente e moralmente l'uomo»[76]. La rettifica fatta valere in assenza di richiesta del soggetto interessato e le conseguenze disciplinari derivanti dalla mancata osservanza del dovere di rettifica da parte del mezzo informativo sostanziano i limiti intrinseci del diritto alla libertà di informare che, oltreché limiti, sono anche e soprattutto elementi costitutivi essenziali del diritto soggettivo di informazione.
I limiti necessari, costituiti dalla verità sostanziale e putativa, dalla buona fede, dalla le-

[74] Corte cost., 15 maggio 1974, n. 133, cons. dir. n. 3.

[75] Cfr. ad esempio, Deliberazione n. 75/2010 del Consiglio nazionale dell'Ordine dei giornalisti, 19 novembre 2010. La decisione concerne il noto caso Feltri-Boffo.

[76] G. Gonella, *La libertà di stampa e i diritti individuali di libertà*, cit., 163.

altà e dal rispetto dei diritti altrui, sono elementi impliciti, intrinseci e costitutivi del diritto, in quanto «ogni soggetto è titolare di una pluralità di diritti che devono coesistere, e possono coesistere, solo a condizione che la sfera di un diritto non invada la sfera di un altro diritto»[77]. Conseguentemente «ogni soggetto è titolare di una pluralità di doveri che limitano la sua *facultas agendi*, sia in rapporto ai diritti propri, come in rapporto ai diritti altrui», dal momento che la libertà di informazione giornalistica «non è solo una *facultas agendi*, ma anche una potestà di esigere il rispetto dell'esercizio di una facoltà»[78].

11. L'iniziativa legislativa per l'estensione della disciplina della rettifica della stampa ai siti online registrati presso i Tribunali e parificati pertanto alla carta stampata

Sotto il profilo formale non vi è ad oggi una disposizione normativa in vigore che preveda le modalità tramite cui l'obbligo di rettifica può essere fatto valere nei confronti delle testate giornalistiche *online* registrate ai sensi dell'art. 5 della l. 47 del 1948. Parrebbe quindi di essere dinanzi ad un vuoto di normazione.

Nella XVIII legislatura è stato, tuttavia, depositato presso il Senato della Repubblica un disegno di legge (n. 836) il quale estende la disciplina legislativa della rettifica a mezzo stampa, in ordine ai modi e ai presupposti di applicazione, alle testate giornalistiche online registrate ai sensi della l. 47 del 1948. Il disegno di legge infatti prevede che «il direttore o, comunque, il responsabile è tenuto a pubblicare gratuitamente e senza commento, senza risposta e senza titolo, con la seguente indicazione: "Rettifica dell'articolo (TITOLO) del (DATA) a firma (AUTORE)", nel quotidiano o nel periodico o nell'agenzia di stampa o nella testata giornalistica *online* registrata ai sensi dell'articolo 5, limitatamente ai contenuti prodotti, pubblicati, trasmessi o messi in rete dalle stesse redazioni, le dichiarazioni o le rettifiche dei soggetti di cui siano state pubblicate immagini od ai quali siano stati attribuiti atti o pensieri o affermazioni da essi ritenuti lesivi della loro dignità, del loro onore o della loro reputazione o contrari a verità, purché le dichiarazioni o le rettifiche non abbiano contenuto suscettibile di incriminazione penale o non siano documentalmente false. Il direttore o, comunque, il responsabile è tenuto a informare l'autore dell'articolo o del servizio, ove sia firmato, della richiesta di rettifica».

La testata giornalistica *online* registrata è pertanto obbligata alla pubblicazione della rettifica «non oltre due giorni dalla ricezione della richiesta, con la stessa metodologia, visibilità e modalità di accesso al sito *Internet*, nonché con le stesse caratteristiche grafiche della notizia cui si riferiscono, nonché all'inizio dell'articolo contenente la notizia cui si riferiscono, senza modificarne la URL e in modo da rendere evidente l'avvenuta modifica. Nel caso in cui la testata giornalistica *online*… fornisca un servizio persona-

[77] Ivi, 162.

[78] *Ibid.* Cfr. G. Capograssi, *La dichiarazione universale dei diritti dell'uomo e il suo significato*, in *giuseppecapograssi.it,* cit., secondo cui i «diritti sono tra di loro solidali, fanno insieme sistema, nessuno può essere sacrificato col pretesto di arrivare mediante questo sacrificio all'appagamento degli altri».

lizzato, le dichiarazioni o le rettifiche sono inviate agli utenti che hanno avuto accesso alla notizia cui si riferiscono». È interessante riscontrare come nell'intendimento di chi ha elaborato il disegno di legge si considerino omogenee, in termini di disciplina applicabile e, più in generale, di capacità di influenza sociale, stampa e testate *online* registrate, mentre invece queste ultime non sono ritenute per forza di penetrazione sociale assimilabili alla radiotelevisione per la quale, come sottolineato in precedenza, la legge prevede una forma e una modalità per l'attivazione della rettifica differenti rispetto alla carta stampata.

12. Il problema giuridico delle comunicazioni online non provenienti da media giornalistici

Vi è un *vuoto di normazione ordinaria,* anche al livello delle iniziative di legge, rispetto alla comunicazione di massa *online* non ascrivibile ai *media* giornalistici. Si pensi ai *blog*, ai siti non registrati e ai contenuti propalati dai *social media*. Ciò dà luogo a problematiche di grande rilievo ricostruttivo e interpretativo che danno l'opportunità di ricordare che *la recente giurisprudenza costituzionale* ha richiamato il legislatore alla necessità di delineare regole e limiti della comunicazione di massa *online* ulteriori rispetto a quelli elaborati da tempo risalente dalla giurisprudenza nella nota sentenza "decalogo" del 1984 in materia di limiti del diritto di informazione: il riferimento è ai requisiti di continenza espressiva, di verità sostanziale e putative del contenuto e di interesse pubblico della materia trattata[79]. Infatti la Corte costituzionale nella sentenza n. 150 del 2021 ha affermato che a causa degli «effetti di rapidissima e duratura amplificazione degli addebiti diffamatori determinata dai *social networks* e dai motori di ricerca in internet», sia sempre più rilevante prestare attenzione al fatto che il «carattere lesivo per la vittima – in termini di sofferenza psicologica e di concreti pregiudizi alla propria vita privata, familiare, sociale, professionale, politica – e per tutte le persone a essa affettivamente legate risulta grandemente potenziato rispetto a quanto accadeva anche solo in un recente passato»[80]. Pertanto «questi pregiudizi debbono essere prevenuti dall'ordinamento con strumenti idonei, necessari e proporzionati, nel quadro di un indispensabile bilanciamento con le contrapposte esigenze di tutela della libertà di manifestazione del pensiero, e del diritto di cronaca e di critica in particolare, in modo da schermare il rischio di indebita

[79] Cass. civ., sez. I, 18 ottobre 1984, n. 5259: «va ricordato che - come ormai la giurisprudenza di questa Corte ha più volte avuto occasione di precisare, sia in sede civile che penale - il diritto di stampa (cioè la libertà di diffondere attraverso la stampa notizie e commenti) sancito in linea di principio nell'art. 21 Cost. e regolato fondamentalmente nella l. 8 febbraio 1948 n. 47, è legittimo quando concorrano le seguenti tre condizioni: 1) utilità sociale dell'informazione; 2) verità (oggettiva o anche soltanto putativa purché, in quest'ultimo caso, frutto di un serio e diligente lavoro di ricerca) dei fatti esposti; 3) forma "civile" della esposizione dei fatti e della loro valutazione: cioè non eccedente rispetto allo scopo informativo da conseguire, improntata a serena obiettività almeno nel senso di escludere il preconcetto intento denigratorio e, comunque, in ogni caso rispettosa di quel minimo di dignità cui ha sempre diritto anche la più riprovevole delle persone, sì da non essere mai consentita l'offesa triviale o irridente i più umani sentimenti».

[80] Corte cost., 12 luglio 2021, n. 150, cons. dir. n. 6.2.

intimidazione esercitato su chi svolga la professione giornalistica»[81].
La natura specifica della comunicazione *online* e la sua capacità di espandersi velocemente nella società con i connessi danni ai diritti della personalità richiederebbero, secondo i giudici costituzionali, uno specifico tipo di *bilanciamento* nel quale entrino in gioco *ulteriori limiti* rispetto a quelli (verità, utilità sociale e continenza) consolidati e codificati dalla giurisprudenza ordinaria a partire della sentenza decalogo sopra richiamata.

13. Il vuoto riempito dalla giurisprudenza

Il vuoto di normazione ordinaria, come spesso accade quando si tratti di esercizio dei diritti fondamentali, viene riempito da una giurisprudenza sempre più attiva anche nell'individuazione delle regole formali applicabili alla comunicazione di massa *online* non equiparabile alla stampa. Si pensi ai siti informativi non registrati, ai *blog* e ai contenuti propalati dai *social networks*.
In merito ai limiti opponibili a codesti mezzi d'informazione occorre evidenziare come la giurisprudenza abbia stabilito – alla luce della formulazione ampiamente garantistica dell'art. 21, c. 1, Cost., che riconosce a tutti la libertà di manifestazione del pensiero anche quando si esprime con i contenuti della libertà d'informazione – che chi diffonde notizie via *Internet*, anche se non giornalista, sia tenuto al rispetto dei limiti del diritto di cronaca consistenti nella rilevanza sociale, nella verità e nella continenza. La Corte di Cassazione ha evidenziato come i diritti di cronaca e di critica discendano direttamente dall'art. 21 Cost., non essendo quindi riservati ai giornalisti o a chi fa informazione professionalmente, ma estesi all'individuo *uti civis*.
Chiunque, pertanto, può riportare fatti o manifestare opinioni e lo può fare con qualsiasi mezzo. In tale contesto, *Internet* rappresenta un potente mezzo di diffusione di notizie (immagini ed idee) attraverso il quale può estrinsecarsi quel diritto di manifestazione del pensiero che costituisce uno dei cardini di una democrazia matura[82]. Quando un soggetto, indipendentemente dal fatto che sia giornalista o meno, immette in rete notizie o commenti aventi ad oggetto fatti di rilevanza sociale è ineludibile l'osservanza di quei parametri che costituiscono i limiti e le modalità corrette di esercizio della libertà di informazione: la rilevanza sociale dell'argomento trattato (interesse pubblico); il fatto che l'informazione trasmetta la verità obiettiva, favorendone le condizioni che la rendano accertabile; la correttezza delle espressioni usate (la decenza espressiva o la continenza).
Ciò che rileva è l'aspetto sostanziale, ovvero la valutazione della natura informativa del contenuto immesso in rete. Se, quindi, su di un sito *Internet*, indipendentemente dalla

[81] Corte cost., 12 luglio 2021, n. 150, cons. dir. n. 6.2.

[82] Cass. pen., sez. V, 25 luglio 2008, n. 31392. Più precisamente, la Corte di Cassazione ha annullato con rinvio, sia pur ai fini civili, una sentenza resa in sede di appello, con la quale una donna era stata assolta dal delitto di diffamazione aggravata nei confronti di una impresa chimica, per aver pubblicato su un sito ambientale il contenuto di una denuncia da lei presentata all'autorità giudiziaria, in cui l'impresa veniva accusata di aver "scaricato cancerogeni in un lago e camuffato la presenza di cancerogeni per mezzo di diluizione con acque di raffreddamento".

sua registrazione, vengono divulgati contenuti dotati dei requisiti dell'informazione (*in primis* della rilevanza sociale dell'argomento), quei contenuti sono assoggettati ai medesimi limiti sanciti dalla sentenza "decalogo" della Corte di Cassazione. Bisogna tuttavia ricordare che mentre i giornalisti incorrono in tutte le responsabilità, penali, civili e disciplinari, per i non giornalisti, quando esercitino scorrettamente l'attività informativa, accanto alla responsabilità penale e civile[83], non sussiste quella disciplinare.

14. La regolamentazione della pratica professionale giornalistica: le scuole di giornalismo

L'evoluzione della professione, delle metodologie del giornalismo e soprattutto dei tempi di vita riferiti al consumo di informazione, ha fatto sì che la pratica possa essere svolta in forme ulteriori rispetto a quanto previsto *ab origine* dalla l. 69 del 1963[84]. La disciplina giuridica del praticantato ha rappresentato e rappresenta un aspetto sul quale sono state sollevate criticità. In particolare, le critiche si sono appuntate sui seguenti profili.

Si sottolinea che la legge impone un'assunzione presso un'azienda editoriale, con ciò invertendo la prospettiva per la quale l'assunzione sarebbe il punto di arrivo della pratica e non la condizione preliminare per il suo svolgimento. Inoltre, l'esame di idoneità professionale per il giornalista[85] viene svolto alla fine di un periodo di praticantato che costituisce esercizio di attività professionale. Il problema impone di interrogarsi sulla possibilità di svincolare l'acquisizione dello *status* di giornalista dalla condizione preliminare della sostanziale assunzione presso un'azienda editoriale, individuando canali ulteriori di formazione tecnica e culturale. Le risalenti considerazioni della giurisprudenza costituzionale sull'argomento sono state nel senso di prendere atto di una realtà materiale che non dipende dalla legge, ossia la realtà di un mercato dell'informazione giornalistica caratterizzato da imprese editoriali private.

Pertanto, l'unica garanzia esigibile nei confronti dell'ordinamento giuridico, affinché il potere economico privato non faccia premio sulla libertà di informazione e di critica del giornalista, è data dal promuovere la «concorrenza della molteplicità delle iniziative giornalistiche»[86]. In questa espansione si innesta il documento dal titolo «"*Quadro di indirizzi" per l'autorizzazione, la regolamentazione e il controllo delle scuole di formazione al giornalismo*» adottato dal Consiglio nazionale dell'Ordine dei giornalisti il 27 settembre 2018.

L'obiettivo è quello di disancorare la pratica giornalistica dal controllo dell'impresa editoriale, punto dolente sul quale fin dall'origine si sono manifestate critiche alla legge professionale. Il Consiglio nazionale dell'Ordine dei giornalisti autorizza lo stu-

[83] La responsabilità del giornalista può essere fatta valere su più ambiti: nel penale come ad esempio per la diffamazione a mezzo stampa, ai sensi dell'art. 595 c.p., fattispecie di delitto doloso; nel civile sempre per la diffamazione a mezzo stampa ai sensi dell'art. 2043 c.c., in cui però, a differenza che nel penale può rilevare la colpa; nel campo deontologico per violazione delle regole deontologiche di correttezza professionale.

[84] Art. 34 l. 69 del 1963.

[85] Art. 32 l. 69 del 1963.

[86] Corte cost., 23 marzo 1968, n. 11, cons. dir. n. 7.

dio e la formazione al giornalismo attraverso apposite strutture denominate scuole. Il Consiglio nazionale dell'Ordine dei giornalisti può autorizzare le scuole finalizzate all'accesso professionale e, stipulando apposite convenzioni, le dichiara sedi idonee allo svolgimento del praticantato previsto dalla legge n. 69 del 1963[87].

15. Il rapporto tra media giornalistici ed altri media *online*

In merito invece alla relazione tra il giornalismo dei *media* tradizionali (stampa e radiotelevisione) e l'informazione *online*, in assenza di una disciplina legislativa, ci si deve avvalere del contributo spesso "creativo" della giurisprudenza.

Così la Corte di Cassazione, Sezioni Unite, 17 luglio 2015, n.3102 afferma che le testate giornalistiche *online*, registrate presso le cancellerie dei Tribunali, con contenuti redatti professionalmente, costituendo mezzi di diffusione delle informazioni, devono godere delle garanzie che la Costituzione predispone per la stampa. Inoltre le Sezioni Unite hanno affermato che, pur essendo ammissibile l'ordine dell'autorità giudiziaria rivolto all' *Internet Service Provider* di rendere inaccessibile un intero sito o una singola pagina *web*, non può essere sottoposta a *sequestro preventivo* una testata giornalistica telematica, al pari di quella cartacea se non nei casi previsti espressamente dalla legge ai sensi del c. 3 dell'art. 21 Cost[88].

Del resto, per certa giurisprudenza, anche le norme penali di sfavore, come l'art. 57 del codice penale[89], previste per il direttore responsabile del quotidiano cartaceo de-

[87] In particolare, si pensi all'art. 1 del "*Quadro di indirizzi*" il quale stabilisce che «Il Consiglio nazionale dell'Ordine dei giornalisti, al fine di promuovere un accesso trasparente e meritocratico alla professione attraverso lo sviluppo di competenze avanzate e di una solida preparazione deontologica, autorizza lo studio e la formazione al giornalismo attraverso apposite strutture qui di seguito denominate scuole. Il Consiglio nazionale dell'Ordine dei giornalisti può autorizzare le scuole finalizzate all'accesso professionale e, stipulando apposite convenzioni, le dichiara sedi idonee allo svolgimento del praticantato previsto dalla legge 3/2/1963 n.69. Il Consiglio nazionale delibera l'autorizzazione dopo aver: a) verificato l'osservanza dei requisiti previsti dal presente "Quadro di indirizzi"; b) considerato il parere del Comitato tecnico-scientifico (Cts); c) chiesto il parere del Consiglio regionale competente, espresso almeno con la maggioranza dei componenti. Le convenzioni, stipulate dal Comitato esecutivo e ratificate dal Consiglio nazionale, hanno validità biennale e non sono tacitamente rinnovabili». Cfr. sulla pratica M. Pedrazza Gorlero, *Giornalismo e giornalisti nella società pluralista: profili costituzionali*, cit., 10 ss.

[88] L'art. 21, c. 3, Cost, stabilisce che «si può procedere a sequestro soltanto per atto motivato dell'autorità giudiziaria nel caso di delitti, per i quali la legge sulla stampa espressamente lo autorizzi, o nel caso di violazione delle norme che la legge stessa prescriva per l'indicazione dei responsabili». Si deve ricordare che nell'ordinamento giuridico italiano il sequestro preventivo è adottabile solo nei seguenti casi: 1) violazione delle norme sulla registrazione delle pubblicazioni periodiche e sull'indicazione dei responsabili (rispettivamente artt. 3 e 16 l. 47/1948); 2) stampati osceni o offensivi della pubblica decenza (art. 2 R.d.lgs. 561/1946); 3) stampa periodica che compia apologia di fascismo (art. 8 l. 645/1952; 4) violazione delle norme a protezione del diritto d'autore (art. 161 l. 633/1941). Pertanto, i giudici di legittimità sostengono che la mancata estensione automatica della garanzia costituzionale di cui all'art. 21, c. 3, all'informazione giornalistica diffusa per via telematica costituisce una violazione del principio di eguaglianza-ragionevolezza di cui all'art. 3 Cost.

[89] «Salva la responsabilità dell'autore della pubblicazione e fuori dei casi di concorso, il direttore o il vice-direttore responsabile, il quale omette di esercitare sul contenuto del periodico da lui diretto il controllo necessario ad impedire che col mezzo della pubblicazione siano commessi reati, è punito, a titolo di colpa, se un reato è commesso, con la pena stabilita per tale reato, diminuita in misura non

vono applicarsi al sito informativo registrato equiparato alla stampa sulla base di una interpretazione figurata della stampa alla quale è ascritto «il prodotto editoriale che presenta i requisiti ontologico (struttura) e teleologico (scopi della pubblicazione) propri di un giornale. La struttura di questo è costituita dalla "testata", che è l'elemento che lo identifica, e dalla periodicità regolare delle pubblicazioni (quotidiano, settimanale, mensile); la finalità si concretizza nella raccolta, nel commento e nell'analisi critica di notizie legate all'attualità (cronaca, economia, costume, politica) e dirette al pubblico, perché ne abbia conoscenza e ne assuma consapevolezza nella libera formazione della propria opinione»[90].

Occorre ricordare, a questo proposito, che la tendenza giurisprudenziale ad estendere la disciplina della stampa ai giornali *online* è stata esaltata dalla l.198 del 2016 che, anche se ai soli fini dell'equiparazione economica dei due mezzi di informazione, ha individuato la nozione ed i caratteri del quotidiano *online* in modo non dissimile da quelli propri della stampa: «il prodotto editoriale è identificato dalla testata, intesa come il titolo del giornale, della rivista o di altra pubblicazione periodica, avente una funzione e una capacità distintiva nella misura in cui individua una pubblicazione»; all'art. 1 della legge 7 marzo 2001, n. 62, è aggiunto, in fine, il seguente comma: «3-bis. Per "quotidiano on line" si intende quella testata giornalistica: a) regolarmente registrata presso una cancelleria di tribunale; b) il cui direttore responsabile sia iscritto all'Ordine dei giornalisti, nell'elenco dei pubblicisti ovvero dei professionisti; c) che pubblichi i propri contenuti giornalistici prevalentemente on line; d) che non sia esclusivamente una mera trasposizione telematica di una testata cartacea; e) che produca principalmente informazione; f) che abbia una frequenza di aggiornamento almeno quotidiana; g) che non si configuri esclusivamente come aggregatore di notizie».

16. La crisi del giornalismo professionale come crisi della democrazia costituzionale

Il giornalismo vive oggi un periodo di crisi accompagnato a processi di rapidissima trasformazione. È palese come questo stato di crisi si riverberi sulla *qualità della democrazia*, dal momento che quest'ultima tende a ridursi a mera forma (periodiche elezioni…) ove non riesca a saldarsi ad una società politica pienamente percorsa da una informazione di altrettale qualità.

Se, infatti, i diritti ad una informazione rispettosa delle procedure di raccolta, verifica, selezione ed elaborazione delle notizie (procedure sancite dalle regole di correttezza professionale dei documenti deontologici), possono essere fatti valere solo verso una categoria (professionisti e pubblicisti) e verso le testate giornalistiche propriamente dette (incluse naturalmente quelle equiparate alla stampa come i giornali telematici)

eccedente un terzo».

[90] Cass. pen., sez. V, 11 gennaio 2019, n. 1275: «risulta, dunque, evidente che l'area riduttiva del significato attribuito al termine "stampa" dall'art. 1 della legge n. 47 del 1948 è strettamente legata alle tecnologie dell'epoca, e ciò non impedisce di accreditare oggi - e tenuto conto dei notevoli progressi verificatisi nel settore una interpretazione estensiva del detto termine, la quale non esorbiti dal campo di significanza del segno linguistico utilizzato e che sia coerente con il dettato costituzionale».

e non verso la comunicazione esterna al giornalismo, su cui peraltro sempre di più la società forma i propri convincimenti, quale può essere il futuro della democrazia costituzionale[91]?

In merito a ciò si ricorda come il *Global Risk Report* del 2021 ponga come prioritaria la necessità di una strategia globale, alla quale gli Stati nazionali e le organizzazioni internazionali sono chiamati a partecipare, per contrastare il fenomeno delle *fake news*: «*improving the clarity and consistency of risk communications and combating misinformation. Most crises require all-of-society responses – and there is enormous goodwill and energy to leverage – but confusion and frustration can undermine efforts to build trust and align responsibilities between the public sector, private sector, communities and households. There is huge scope to enhance self-organized resilience at the community and national levels. For example, more can be done to understand – and therefore tackle – biases at the individual level regarding spread of misinformation. Better coordination among private sector technology companies and government can help to alert users to misinformation*»[92].

Una ricerca del *Global Risk Report* già del 2018 sostiene che le più importanti piattaforme digitali, sulle quali agiscono i *social* media maggiormente diffusi, hanno indirizzato il 40% del traffico verso siti contenenti notizie false[93]. La questione ha molteplici risvolti. Innanzitutto, il problema che si dovrà affrontare è la progressiva eclissi dell'idea secondo la quale, perché vi sia una effettiva democrazia politica, è necessario che vi sia una società democratica. Ciò significa che affinché vi sia una democrazia sostanziale è indispensabile «l'esistenza di un vasto strato acculturato e/o informato»[94]. In ballo vi sono la difesa della qualità dei contributi informativi e la loro accessibilità alla società tutta. In particolare il problema consiste nel fatto che in un futuro prossimo tale qua-

[91] Cfr. Allegato A alla delibera n. 146/15/CONS dell'Autorità per le garanzie nelle comunicazioni dal titolo "*Indagine conoscitiva su informazione e Internet in Italia, modelli di business, consumi, professioni*", reperibile in *agcom.it*. In particolare a p. 11 si osserva quanto segue: «l'evoluzione del consumo di notizie da fasce anziane a coorti più giovani della popolazione muta profondamente le loro modalità di fruizione. Infatti, la fruizione dell'informazione avviene sempre più spesso attraverso una molteplicità di apparecchi (dal classico pc, agli *smartphone* e *tablet*) e segue percorsi sempre meno tradizionali, con l'affermazione di nuovi attori, quali soprattutto le piattaforme di aggregazione, ricerca e condivisione sociale che orientano sempre più spesso le scelte di consumo degli utenti. È importante sottolineare che la fruizione di informazione tramite *social* può essere anche il frutto di un'esperienza occasionale, nel senso che, in considerazione della loro caratteristica di contenitori, è molto probabile che si venga raggiunti da una notizia mentre si è sul *social network* per altri motivi. In questo senso, il consumatore può perdere l'idea di chi è il vero fornitore dell'informazione, associando l'intera esperienza di navigazione al *social network* stesso. Ciò pone un vero e proprio "dilemma" in capo agli editori, in particolar modo a quelli tradizionali: da un lato, le piattaforme di condivisione sociale generano traffico verso le pagine degli editori; d'altra parte, questo tipo di navigazione, specie se all'interno di un sistema chiuso (quale una *app* proprietaria), rischia di far diminuire la conoscenza del marchio editoriale da parte del pubblico, riducendone, nel lungo periodo, la propria forza commerciale».

[92] Il *Global Risk Report* 2021 è reperibile al sito *weforum.org*, 25. Cfr. anche p. 33 del medesimo report: «*Misinformation is increasingly threatening civil liberties and democracy "Post-truth" politics – from deliberate manipulation campaigns to the unmitigated spread of conspiracy theories and fake news – are "amplifying hate speech; heightening the risk of conflict, violence and human rights violations; and threatening longterm prospects for advancing democracy" as the World Health Organization has warned. Yet blunt government attempts to combat misinformation can exacerbate the problem. Internet restrictions, for example, risk excluding whole societies from the global information economy, while more invasive control could infringe civil liberties*».

[93] Cfr. *The Global Risks Report 2018*, 13th Edition, World Economic Forum, Geneve 2018, 48.

[94] M. D'Eramo, *Invenzione, ascesa e declino del giornale*, in *È la stampa bellezza*, Almanacco di giornalismo, Micromega, 3, 2018, 23.

lità potrebbe essere soprattutto prerogativa dei contenuti *premium* dei grandi *network* dell'informazione, perché solo essi saranno in grado di sostenere i costi di un giornalismo approfondito e d'inchiesta.

Si pensi alla difficoltà di ricostituire su Internet quel «circuito di flusso monetario che per due secoli aveva permesso di finanziare l'industria dell'informazione»[95]. Vi è un dato empirico sul quale riflettere: «nel 2017, al *New York Times* […] le vendite contribuivano per il 60,1 per centro alle entrate, mentre la pubblicità solo per il 33,3 per cento (un terzo), quando […] per tutto il Novecento la pubblicità contribuiva per i quattro quinti alle entrate e le vendite solo per un quinto»[96]. Il fatto è che una testata giornalistica per potersi sostenere soprattutto sulla base degli introiti delle vendite è costretta ad applicare alte tariffe.

Il pericolo è l'erosione delle garanzie sostanziali della democrazia e di conseguenza della effettività del diritto all'informazione in capo alla società. In particolare il problema è rappresentato da una realtà in cui la produzione e la diffusione dell'informazione verificata, approfondita e di qualità e, quindi, della conoscenza possano essere a beneficio solo dei «(pochi) destinatari finali del circuito economico»[97]. La situazione prefigura un sistema dell'informazione caratterizzato dal rapporto tra un'oligarchia dei media e un'oligarchia dei destinatari, situazione che a sua volta replica proprio quella mediazione informativa organica ad una società politica monoclasse di cui si è parlato prima.

Le considerazioni si legano ai rapporti tra sistema politico e media tutti (in particolare si pensi ai *social* media). Il potere modella e forma il consenso del popolo. Si tratta di un processo dall'alto verso il basso, processo etero-diretto alla cui realizzazione un ruolo fondamentale è in capo ai mezzi di comunicazione di massa. A tale processo la rappresentanza politica democratica contribuisce sempre di più oggi e come soggetto

---

95 M. D'Eramo, *Invenzione, ascesa e declino del giornale*, cit., 21.

96 *Ibid.*

97 Ivi, 23. Lo stesso A. (20 ss.) poi analizza sul fronte delle dinamiche dell'economia dei media una situazione suscettibile di riverberarsi sulla tenuta della democrazia: «nel 2006 gli introiti dei giornali USA ammontavano a 60 miliardi di dollari, di cui 50 dovuti alla pubblicità e 10 alle vendite: queste ultime rappresentavano il 17,6 per cento delle entrate, mentre la pubblicità si ritagliava l'82 per cento. Più di quattro quinti del fatturato era dovuto alla pubblicità. Nel 2011, le entrate erano scese a 37 miliardi di dollari (- 38 per cento in cinque anni), i ricavi delle vendite erano rimasti quasi costanti (10 miliardi di dollari), mentre quelli della pubblicità erano scesi a 27 miliardi (un calo del 46 per cento). Cinque anni dopo, nel 2016 le entrate totali erano scese a 29 miliardi (si erano più che dimezzati dal 2006), di cui 18 di pubblicità (che era circa un terzo di quella di 10 anni prima) e 11 di vendite (che erano persino salite un po' grazie agli abbonamenti digitali). In dieci anni sono andati persi 30 miliardi di dollari su 50 di entrate pubblicitarie. L'aspetto più interessante è che, se anche il peso della pubblicità digitale sale nel tempo, il suo volume resta però quasi costante: nel 2011 il peso del digitale era pari al 17 per cento di una torta di 27 miliardi, valeva cioè 4,6 miliardi, mentre nel 2016, pur costituendo ormai quasi un terzo del totale pubblicitario (29 per cento), valeva in tutto 5,3 miliardi. In altri termini: con la pubblicità digitale non si vive, tanto che nel 2006 per la stampa Usa i ricavi delle vendite erano ormai più di un terzo (37,3 per cento) dei ricavi totali (rispetto a meno di un quinto di dieci anni prima). Addirittura, nel suo ultimo bilancio il *New York Times* nota che pur aumentando assai gli abbonamenti, il fatturato pubblicitario digitale sale solo grazie alle app e ai cellulari, mentre quello sul sito classico del giornale già scende. Il punto è che la crescita della pubblicità digitale non riesce a compensare le drammatiche perdite della pubblicità stampata. Anzi, nel suo rapporto del 2013 il Pew Research Center dice che la pubblicità cartacea perde 15 dollari per ogni dollaro guadagnato dalla pubblicità digitale. In quell'anno il *New York Times* accrebbe i propri abbonamenti digitali del 19 per cento (passando da 640 a 760 mila), eppure vide diminuire la sua pubblicità digitale del 4 per cento».

concorrente dei media (si pensi al fenomeno della politica mediatizzata) e come terminale passivo rispetto ai media stessi.

La sfera della politica è condizionata dai media, in quanto i soggetti politico-istituzionali recepiscono dai media, i quali paiono sempre di più decisori, argomenti di discussione pubblica, temi, linguaggi, *slogan*, proposte, soluzioni, direzioni del discorso e decisioni. Il pericolo è quello di un cortocircuito rispetto alla necessità di un equilibrio costituzionale fra media e decisione politica. I mezzi di comunicazione decidono quali siano i temi di rilevanza sociale e, perfino, gli indirizzi politici da imprimere alle decisioni pubbliche, producendo uno squilibrio nel funzionamento "ordinario" dei poteri della democrazia politica.

La democrazia politica elettiva viene così ad agire in senso solo formale nei meccanismi istituzionali e nelle procedure costituzionalmente codificate, ma in termini sostanziali le istanze, gli interessi, i contenuti politicamente rilevanti e le soluzioni in termini di decisioni sono elaborati nell'ambito della sfera mediale[98]. Si pensi al potere esercitato dai *social* media nell'attività sociale di anticipare ed anche di modellare quali saranno i fini e gli interessi prevalenti degli utenti[99]. Si registrano nella discussione pubblica sulle dinamiche evolutive della comunicazione di massa proposte che mirano all'assunzione di responsabilità deontologica su base esclusivamente volontaria di figure non rientranti nelle categorie dei professionisti, dei pubblicisti e dei praticanti. Si pensi a figure quali gestori di siti di discussione, *blogger*, realizzatori di *web* tv, comunicatori non occasionali operanti nei *social* media ecc. Si conteggia un numero di 50.000 soggetti circa «che operano con i diversi mezzi principalmente nella rete di Internet, con i social, con i blog, con i siti. Costoro non hanno alcun riconoscimento né responsabilizzazione, né tutela

[98] Cfr. M. Castells, *Comunicazione, potere e contropotere nella* network society, in *International Journal of Communication*, 1, 2007, 238 ss. Cfr. sempre M. Castells, *Comunicazione e potere*, trad. di B. Amato e P. Conversano, Milano, 2009, 378: «se accettiamo l'idea che la forma cruciale del potere ha luogo attraverso la modellazione della mente umana, e che questo processo dipende in larga parte dalla comunicazione, e in ultima analisi dalla politica mediatica, allora *la pratica della democrazia è messa in discussione quando c'è dissociazione sistemica tra potere della comunicazione e potere rappresentativo*». Sulla relazione Internet-democrazia cfr. il saggio di I. Ramonet, *L'esplosione del giornalismo. Dai media di massa alla massa dei media*, trad. di P. Sullo, Napoli, 2011. Sull'influenza dei social media sulla società politica cfr. V. Mayer Schönberger - K. Cukier, B*ig Data*, New York, 2013; I.S. Rubinstein, *Big Data: The End of Privacy Or a New Beginning?*, in *International Data Privacy Law*, 3(2), 2013, 74; F. Pasquale, *The Black Box Society: The Secret Algorithms That Control Money and Information*, Cambridge (MA), 2015; J. Chester - K. Montgomery*, The Role of Digital Marketing* in *Political Campaigns*, in *Internet Policy Review*, 6, 2017, 4; G. Gori, Social media *ed elezioni. I limiti del diritto e il rischio di una* modulated democracy, in *Informatica e diritto*, 1-2, 2017, 203 ss.; O. Pollicino, *La prospettiva costituzionale sulla libertà di espressione nell'era di Internet*, in questa *Rivista*, 1, 2018, 48 ss.; G. De Minico, *Libertà digitali. Luci e ombre*, con presentazione di Enzo Cheli, Torino, 2018, 1 ss.; J. Van Dijck - T. Poell - M. de Wall, *Platform society. Valori pubblici e società connessa*, edizione italiana a cura di G. Boccia Artieri e A. Marinelli, trad. di A. Marinelli - A. Massa - S. Parisi, Milano 2019, 76; M. Bassini, *Internet e libertà di espressione. Prospettive costituzionali e sovranazionali*, Roma, 2019, 107 ss.; G. Marchetti, *Gli algoritmi e il ruolo delle* fake news, in questa *Rivista*, 1, 2020, 30. Cfr. G. Sartori, *Homo videns. Televisione e post-pensiero*, Bari, 1999, 46: «alla democrazia rappresentativa basta, per funzionare, che esista una opinione pubblica che sia davvero *del* pubblico. Ma è sempre meno così, dato che la videocrazia sta fabbricando una opinione massicciamente etero-diretta che in apparenza rinforza, ma in sostanza svuota, la democrazia come governo di opinione».

[99] G. Gori, Social media *ed elezioni. I limiti del diritto e il rischio di una* modulated democracy, in *Informatica e diritto*, 1-2, 2017, 218.

da parte della comunità»[100].
Pertanto, si pone l'accento sulla necessità che emerga pubblicamente e legalmente una presenza di soggetti, esternamente al giornalismo professionale, che assumano la responsabilità deontologica rispetto ai contenuti propalati, in quanto tale emersione è l'elemento prìncipe della funzione storica, sociale, politica e giuridica del giornalismo[101].

17. L'attualità della l. 69 del 1963

È venuto il momento di chiedersi se per sottrarre il futuro a questo inquietante orizzonte possa essere d'aiuto l. 69 del 1963, così come l'abbiamo ripercorsa nelle sue dirette o indirette implementazioni e nelle interpretazioni che ne sono state fatte a livello giurisprudenziale.
La risposta può essere, anche se parzialmente, positiva. Il giornalismo copre legislativamente sia l'area dei tradizionali mezzi d'informazione (la stampa e la radiotelevisione) sia quell'area di *Internet* (i siti informativi registrati) alla quale è estesa, come già si è visto, l'applicazione delle regole del giornalismo: la giurisprudenza ha variamente ampliato la disciplina della stampa alle testate giornalistiche *online*; ciò vuol dire che fin dove riuscirà ad allargare la nozione di attività giornalistica e ad assimilare al giornalista le altre figure soggettive dell'informazione, il legislatore avrà aperto e tracciato il campo delle soluzioni possibili e ragionevoli. Per il resto, per il *mare magnum* della comunicazione *online*, occorrerà un più deciso passo in avanti. La l. 69 del 1963 non lascia tuttavia sguarnito il legislatore né di modelli né di principi. Si pensi alla nozione di attività giornalistica come modello dell'attività informativa e al giornalista come modello delle altre figure soggettive dell'informazione; al principio di corrispondenza fra libertà e responsabilità e al principio di convergenza fra la realtà e la sua rappresentazione implicata dall'istituto della rettifica; alla verità sostanziale dei fatti come notizia costruita dal "soggettivamente vero" appoggiato alla fonte fiduciaria; alla copertura del segreto a favore di quest'ultima; alla responsabilità disciplinare gestita dall'Ordine e perciò all'apertura modellistica verso un equivalente soggettivo che l'amministri nei confronti di chi svolga attività informativa non essendo giornalista.
Si tratta di un lascito i cui valori e contenuti appaiono ancora oggi strategicamente importanti. In gioco, infatti, vi è sempre la difesa della democrazia come sistema politico-istituzionale teso a garantire il pluralismo delle opinioni e delle critiche, la qualità dei contenuti informativi, la partecipazione e l'emancipazione sociale dalle sfere del dominio e quindi l'effettività delle libertà costituzionali. Una mediazione informativa svolta

[100] Cfr. R. Fiengo, *Intervento,* in *Quale futuro per il giornalismo*, in V. Roidi (a cura di), *Quaderni della Fondazione Paolo Murialdi,* cit., 29.

[101] R. Fiengo, *Intervento*, cit., 33: «l'unica strada, accanto a quella che viene cercata sul piano delle regole, anche legislative, è quella di avere su tutto il terreno della comunicazione persone responsabilizzate, meglio pagate, meglio titolate […] andrebbe semplicemente data loro la possibilità di aderire alle carte maturate negli anni dalla professione giornalistica. Anche sottoscrivendole in forma semplice sapere che c'è un discorso per le fonti, un discorso per la tutela dei minori, un discorso per tenere separata la pubblicità. Anche se su questi 40/50 mila si riuscisse a accettarne 10/15 mila ci si muoverebbe in una direzione assai utile, una delle poche percorribili per attivare un allargamento dell'informazione di qualità».

da giornalisti, soggetta alle garanzie, alle regole, ai limiti e alle responsabilità disposte dalla legge professionale è dunque essenziale ad una democrazia emancipante nella quale il giornalismo assicuri «un uso pubblico della ragione che legittima l'agire politico»[102] e quindi il controllo sociale sugli argomenti a sostegno delle relative decisioni[103]. Non pare poco per una legge che la professione e la costituzionalistica (non la Corte costituzionale) e la politica avevano ritenuto di salute malferma fin dalla sua entrata in vigore.

E dell'Ordine professionale e del suo artefice cosa dire conclusivamente?

Sull'Ordine non si fanno più questioni di legittimità, ma di opportunità. Il *soggetto ha ceduto alla funzione*: i giornalisti sono gli intermediari necessari della nostra conoscenza della realtà. È la correttezza di questa *intermediazione* che va garantita ai terzi da un ordine professionale, in tempi di vita sempre più convulsi e rispetto a mezzi d'informazione sempre meno soggettivamente raffrontabili. Una intermediazione garantita fino all'applicazione da parte dell'Ordine delle sanzioni disciplinari, ossia delle sanzioni meglio idonee a rendere effettive insieme le libertà ed i doveri degli informatori professionali.

E a Guido Gonella cosa dobbiamo? Gratitudine istituzionale per averci trasmesso uno strumento in grado di equilibrare il rapporto fra potere economico e informazione professionale, uno strumento che ancor oggi mantiene in vita l'informazione democratica e non appare inadeguato ad affrontarne i nuovi nodi che la storia continuamente propone.

---

[102] J. Nida Rümelin, *Democrazia e verità*, trad. di L. Rega, a cura di F. Longato, Milano, 2015, 40.

[103] *Ibid.*: «l'atto dell'elezione non costituisce un attore collettivo, non impone la volontà della maggioranza contro quella delle minoranze, ma dà mandato per adottare le decisioni politiche che sembrano essere sostenute dagli argomenti migliori, che poi dovranno essere a loro volta presentati pubblicamente ed esposti alla critica» oltreché alla verifica in merito alla corrispondenza tra realtà e rappresentazione. Cfr. N. Urbinati, *Democrazia sfigurata. Il popolo fra opinione e verità*, Milano, 2014, 36: «nel momento in cui l'opinione entra a far parte della nozione di partecipazione democratica, la rappresentanza politica deve occuparsi anche delle *circostanze in cui si formano le opinioni*, tema che attiene alla giustizia politica dei cittadini come uguaglianza di opportunità di esercitare in modo significativo i propri diritti politici. L'eguale diritto a concorrere in modo paritario alla formazione della volontà politica […] deve essere accompagnato da eguali e significative opportunità di essere informati, ma anche di dar forma, espressione, voce, peso e influenza alle proprie idee».

Riverberi costituzionali del Metaverso*

Raffaele Bifulco

Abstract

In quale modo il Metaverso interessa il diritto costituzionale? La principale caratteristica del Metaverso è, dunque, la sua capacità mimetica, la possibilità di riprodurre gran parte delle funzioni, delle attività, dei comportamenti che si sviluppano nel mondo reale. In quanto specchio virtuale del mondo reale, nel Metaverso si riprodurranno, poi, i conflitti esistenti in quest'ultimo. Le più classiche e antiche controversie del mondo reale sono destinate a riprodursi in quello virtuale. È possibile porre, in relazione al Metaverso, le stesse questioni "costituzionali" che si sono poste per Internet? In particolare, anche il Metaverso ha bisogno di una costituzionalizzazione? E perché? Da chi e come sarà regolato? Dagli Stati, dai *tech activist* che vogliono un *open web* oppure dai grandi *players* del mondo digitale (*business-oriented tech sector*)?

Why is constitutional law interested in the Metaverse? Its primary characteristic is its mimetic capability, allowing it to reproduce the fundamental functions, activities, and behaviors of the real world. As a virtual mirror of reality, the Metaverse will mirror the conflicts and societal rifts present in the tangible world. Will the Metaverse encounter similar issues as the Internet? Does the Metaverse require constitutionalization? Who will regulate the Metaverse—nation-states, tech activists advocating for an open web, or the major players in the digital world?

Sommario

1. «*Metaverse […] is springing everywhere*». – 2. Dal mondo virtuale a quello reale e ritorno. – 3. Come si governerà il Metaverso? – 4. Costituzionalizzare il Metaverso?

Keywords

Metaverso – capacità mimetica – conflitti sociali e virtuali – costituzionalizzazione

1. *«Metaverse […] is springing everywhere»*

Partendo da questa constatazione contenuta in un rapporto della Commissione europea (di cui darò conto), nelle pagine che seguono – senza alcuna aspirazione ad

* L'articolo è stato sottoposto, in conformità al regolamento della Rivista, a referaggio "a doppio cieco".

una trattazione sistematica di un fenomeno che si annuncia complesso e, allo stesso tempo, avvincente per il diritto – proverò a estrarre alcuni fili di una matassa piuttosto intricata, ponendo qualche questione che tocca da vicino il diritto costituzionale. Per quanto l'attenzione intorno al Metaverso, da parte della ricerca interdisciplinare che si occupa della Rete, possa farsi risalire alla fine degli anni Novanta dello scorso secolo[1], lo stadio ancora iniziale degli studi giuridici sul Metaverso giustifica lo stile di questo breve scritto, contenente per lo più domande e quesiti piuttosto che risposte e soluzioni[2].

Come è noto, Metaverso allude a ciò che è oltre, risultando da una crasi tra il prefisso "meta" e "verso" come frammento di universo. Un lemma nuovo che allude al divenire e, allo stesso tempo, alla instabilità, alla provvisorietà. E in effetti l'incertezza sugli sviluppi del Metaverso è uno dei suoi tratti caratteristici[3].

Se il Metaverso è concetto nuovo e in continua e progressiva trasformazione, è pur vero che esso rappresenta l'evoluzione di Internet, da cui prende le mosse per poi, in qualche modo, trascenderlo. È un punto pacifico tra gli studiosi del fenomeno[4]. Allo stesso tempo, il Metaverso è anche il frutto dell'integrazione di una varietà di nuove tecnologie: il 5G e il 6G dal punto di vista infrastrutturale, nonché l'intelligenza artificiale e Internet delle cose[5]. Questa dipendenza dallo sviluppo delle tecnologie rende il Metaverso un concetto in continua evoluzione. Bisognerebbe anche chiedersi se sia corretto usare il termine Metaverso, che sembra espressione metonimica per alludere al più ampio e plurale fenomeno di mondi virtuali. In questa sede userò il termine Metaverso per alludere al fenomeno più esteso di mondi virtuali.

Nei mondi virtuali hanno trovato un campo di applicazione privilegiato i giochi e i video games, ma non solo. Il Metaverso si rivela promettente anche per più rilevanti settori sociali, come il commercio (*virtual shopping*), il lavoro (nella specie dell'*homeworking*)[6], l'educazione, la medicina e, in genere, la sfera sociale[7]. Estremamente sintoma-

---

[1] Così H. Ning – H. Wang – Y. Lin – W. Wang – S. Dhelim – F. Farha – J. Ding – M. Daneshmand, *A Survey in Metaverse: the State-of-the-art, Technologies, Applications, and Challenges*, in *arxiv.org*, 18 novembre 2021, 10, i quali individuano quattro fasi di sviluppo del Metaverso, a partire dal 1998 (anno di pubblicazione del romanzo *Avalanche* di Neil Stephenson) fino al 2021, l'anno in cui il Metaverso ha preso effettiva consistenza. Si veda tuttavia V. Mayer-Schonberger – J. Crowley, *Napster's Second Life?: The Regulatory Challenges of Virtual Worlds*, in *Northwestern University Law Review,* 100-4, 2007, 1782 ss. per una ricostruzione del progressivo sviluppo dei mondi virtuali.

[2] Nella letteratura giuridica italiana il volume di F. Sarzana di S. Ippolito – M. Pierro – I. O. Epicoco, *Il diritto del metaverso. NFT, DeFi, GameFi e privacy*, Torino, 2022 appare il primo studio sul Metaverso.

[3] Consiglio dell'Unione europea: Segretariato generale, *Metaverse – Virtual world, real challenges*, in *consilium.europa.eu*, 9 marzo 2022, 3: «we don't really know in detail what Metaverse will look like in 10 years' time».

[4] Una rassegna sui principali studi sul Metaverso si trova in Y. Wang – Z. Su – N. Zhang – D. Liu – R. Xing – T. H. Luan – X. Shen, *A Survey on Metaverse: Fundamentals, Security, and Privacy*, in *arxiv.org*, 8 aprile 2022, 3.

[5] H. Ning – H. Wang – Y. Lin – W. Wang – S. Dhelim – F. Farha – J. Ding – M. Daneshmand, *A Survey in Metaverse*, cit., 12-5; Y. Wang – Z. Su – N. Zhang – D. Liu – R. Xing – T. H. Luan – X. Shen, *A Survey on Metaverse*, cit., 1, 6, indicano sei tecnologie alla base del Metaverso: 1) *interactivity*; 2) *Digital Twin*; 3) *Networking*; 4) *Ubiquitous Computing*; 5) *Artificial Intelligence*; 6) *Blockchain*.

[6] Consiglio dell'Unione europea: Segretariato generale, *Metaverse*, cit., 7.

[7] H. Ning – H. Wang – Y. Lin – W. Wang – S. Dhelim – F. Farha – J. Ding – M. Daneshmand, *A Survey*

tico del Metaverso è poi lo sviluppo degli NFT (*Non Fungible Tokens*), in cui la capacità mimetica del Metaverso si esercita nel campo dell'arte (quella del mondo reale). Non è certo il caso di improvvisare analisi sociologiche dell'arte ma certo si può osservare – ricordando le parole di W. Benjamin sul sempre più imperioso «bisogno di entrare in possesso dell'oggetto, a distanza ravvicinata, nell'immagine, o meglio nella copia, nella riproduzione» – come anche questo nuovo tipo di tecnologia contribuisca ulteriormente a cambiare la percezione dell'arte[8].

L'immersione in mondi virtuali lascia immaginare, già a livello intuitivo, i riflessi profondi che questa esperienza potrà avere sul benessere fisico[9] e spirituale[10] delle persone. Altrettanto intuibili sono i riflessi del Metaverso sui futuri sviluppi del mondo del lavoro, del commercio, della medicina, dell'insegnamento delle arti e delle scienze[11]. In quale modo il Metaverso interessa il diritto costituzionale?

2. Dal mondo virtuale a quello reale e ritorno

Come accade per ogni campo del diritto, anche il diritto costituzionale è direttamente investito dallo sviluppo delle nuove tecnologie. Può definirsi oramai risalente il dibattito sul rapporto tra biotecnologie e norme costituzionali a tutela della persona e della sua salute. Negli ultimi anni, poi, sono stati Internet e le tecnologie ad esso connesse ad aver occupato la dottrina costituzionalistica per i suoi effetti non solo sui diritti delle persone ma anche per i riflessi sull'*empowerment* della persona e quindi sul principio di eguaglianza. Le innovazioni normative provocate da Internet sono continue: basti pensare che la enorme circolazione di dati personali che si svolge sulla Rete ha rappresentato un importante fattore causale nel rafforzamento della tutela europea dei dati personali realizzata dall'Unione europea con l'adozione del Regolamento 2016/679. E come si vedrà, una delle principali questioni che solleva lo sviluppo del Metaverso si collega proprio all'uso dei dati e alla tutela della privacy.

in Metaverse, cit., 6, 11, 24; Y. Wang – Z. Su – N. Zhang – D. Liu – R. Xing – T. H. Luan – X. Shen, *A Survey on Metaverse*, cit., 8; Consiglio dell'Unione europea: Segretariato generale, *Metaverse*, cit., 7, dove si cita Bill Gates secondo cui le riunioni da remoto si sposteranno sempre più sul Metaverso.

[8] W. Benjamin, *L'opera d'arte nell'epoca della sua riproducibilità tecnica*, in W. Benjamin (a cura di), *L'opera d'arte nell'epoca della sua riproducibilità tecnica e altri scritti*, Milano, 2012, 19, il quale, nella stessa pagina, osserva come «*Avvicinare a sé le cose, spazialmente e umanamente, è un'esigenza delle masse attuali altrettanto appassionata della loro tendenza a superare l'unicità di qualunque dato reale mediante la ricezione della sua riproduzione*» (corsivo di Benjamin).

[9] In via esemplificativa un legame troppo stretto con il proprio avatar può essere all'origine di problemi motori e di disordini alimentari: Consiglio dell'Unione europea: Segretariato generale, *Metaverse*, cit., 11.

[10] Si pensi ancora alla *cyber-syndrom*, una malattia causata dall'uso eccessivo di Internet: H. Ning – H. Wang – Y. Lin – W. Wang – S. Dhelim – F. Farha – J. Ding – M. Daneshmand, *A Survey in Metaverse*, cit., 28.

[11] Y. Wang – Z. Su – N. Zhang – D. Liu – R. Xing – T. H. Luan – X. Shen, *A Survey on Metaverse*, cit., 8, osservano che il Metaverso può rivoluzione le nostre abitudini, ad esempio, attraverso forme di *virtual shopping, virtual dating, virtual chatting, global travel, even space/time travel*; cfr. anche H. Ning – H. Wang – Y. Lin – W. Wang – S. Dhelim – F. Farha – J. Ding – M. Daneshmand, *A Survey in Metaverse*, cit., 24, per un elenco delle applicazioni nei campi dell'educazione, dell'industria, dell'arte, della medicina e del sociale che li porta ad affermare che «*the Metaverse has unprecedented explosive power*».

Con riguardo più specifico al Metaverso, basta provare a prendere in considerazione la sua caratteristica principale per accorgersi della rilevanza che esso è in grado di esplicare sul diritto costituzionale. Si allude, in particolare, alla sua capacità di creare una realtà virtuale che riproduce quella reale pur essendo completamente staccata da quest'ultima. Grazie a specifiche tecnologie, che richiedono l'uso di sensori, occhiali, *smartwatch*, ciascuno di noi può sperimentare tale realtà utilizzando un avatar che, in qualche maniera, lo rappresenta[12]. Si entra così in una realtà alternativa che avvolge e assorbe totalmente[13] e che tuttavia riproduce il mondo reale, tanto che si può parlare di una realtà interna (quella del Metaverso) e di una realtà esterna (quella del mondo reale)[14]. Pur senza sposare problematici paralleli tra il Metaverso e più risalenti esperienze spirituali[15], appare corretto ritenere che entrare nei mondi del Metaverso significa entrare in una dimensione che supera i limiti del mondo reale sia dal punto di vista spaziale che da quello temporale.

La principale caratteristica del Metaverso è, dunque, la sua capacità mimetica, la possibilità di riprodurre gran parte delle funzioni, delle attività, dei comportamenti che si sviluppano nel mondo reale[16]. In fondo, utilizzando la sua identità virtuale, l'avatar, in sostituzione dell'essere umano (o in sua rappresentanza), esercita e partecipa a quelle funzioni sociali che caratterizzano il mondo reale[17].

Le due realtà, quella del mondo virtuale e quella del mondo reale, rimangono tuttavia comunicanti. E le esternalità che dal mondo virtuale si producono verso quello reale possono rappresentare un problema per il diritto costituzionale. Si prenda il caso di notizie false riguardanti un candidato alle elezioni o un partito che dal mondo virtuale si diffondono in quello reale o all'uso di monete virtuali che vengono tuttavia utilizzate come misura per scambi economici che si svolgono nel mondo reale. Si intravedono le prime interferenze, giuridicamente rilevanti[18].

---

[12] H. Ning – H. Wang – Y. Lin – W. Wang – S. Dhelim – F. Farha – J. Ding – M. Daneshmand, *A Survey in Metaverse*, cit., 16, sottolineano che gli individui che entrano nel Metaverso avranno una propria identità, che non necessariamente coincide con quella del mondo reale.

[13] Y. Wang – Z. Su – N. Zhang – D. Liu – R. Xing – T. H. Luan – X. Shen, *A Survey on Metaverse*, cit., 6, ragionano di *immersiveness*.

[14] Metaverso «*is best described as an immersive and constant virtual 3D world where people interact through an avatar to enjoy entertainment, make purchases and carry out transactions with crypto-assets, or work without leaving the seat*»: così Consiglio dell'Unione europea: Segretariato generale, *Metaverse*, cit., 3.

[15] H. Ning – H. Wang – Y. Lin – W. Wang – S. Dhelim – F. Farha – J. Ding – M. Daneshmand, *A Survey in Metaverse*, cit., 23, indicano, piuttosto confusamente, come ascendenze del Metaverso il Taoismo, il triplice mondo di Platone, la teologia di Aristotele, l'idea assoluta di Hegel.

[16] Già nel 2006 V. Mayer-Schonberger – J. Crowley, *Napster's Second Life?*, cit., 1784, evidenziavano come fosse proprio la capacità di riprodurre la realtà (gli autori usano il termine "*simulacrum*") l'aspetto più complesso da realizzare per i creatori di mondi virtuali; H. Ning – H. Wang – Y. Lin – W. Wang – S. Dhelim – F. Farha – J. Ding – M. Daneshmand, *A Survey in Metaverse*, cit., 1, sottolineano che il Metaverso «*tightly integrates the virtual world and the real world and the real world into the economic system, the social system, and the identity system , allowing each user to produce content and edit the world*».

[17] «*A quantum leap from current virtual reality* […] *to a world that truly blends the real with the virtual*»: così T. Breton, *People, technologies & Infrastructure – Europe's plan to thrive in the metaverse*, in *ec.europa.eu*, 14 settembre 2022, 2.

[18] Si lasciano da parte altre più strutturali preoccupazioni intorno al Metaverso, come quella espressa dal Consiglio dell'Unione europea: Segretariato generale, *Metaverse*, cit., 11 che, da un lato, sottolinea la pesante impronta delle infrastrutture digitali e l'uso massiccio di energia richiesto per il funzionamento

Anche la comunicazione tra mondi virtuali può essere fonte di problemi. Sarà possibile lasciare i confini di un mondo virtuale per entrare in quelli di un altro? Prevarrà, quindi, la libertà della persona oppure la logica del mercato? Ragionando a bocce ferme, sembra prevalere quest'ultima giacché i *provider*, governati da logiche e calcoli economici, tendono a fidelizzare gli utenti, rendendo comunque difficile l'abbandono di un mondo virtuale a favore di un altro. Quanto viene creato all'interno di un mondo virtuale – si tratti di amicizie, reputazione, identità o denaro virtuale – potrà certo essere trasferito in un altro mondo [19]. Ma a quale prezzo e secondo quali regole?

In quanto specchio virtuale del mondo reale, nel Metaverso si riprodurranno, poi, i conflitti esistenti in quest'ultimo. Le più classiche e antiche controversie del mondo reale sono destinate a riprodursi in quello virtuale. Nei mondi virtuali è possibile esercitare forme di commercio, con la conseguenza che le relative transazioni possono essere all'origine di controversie giuridiche; o ancora, in mondi fondati sulla comunicazione[20], è facile immaginare ipotesi di diffamazioni, di ingiurie e simili fattispecie.

Altri problemi di *governance* derivano dalla permeabilità dei due mondi. Diffondere la musica di un autore in un mondo virtuale, senza autorizzazione o licenza, può tradursi in una violazione delle norme sulla proprietà intellettuale oppure, al contrario, utilizzare i dati personali, di cui si è venuti a conoscenza nel mondo virtuale, all'interno del mondo reale non può non tradursi in una violazione delle regole della privacy e della protezione dei dati personali.

Rimanendo all'interno della problematica dei dati personali, può dirsi che non c'è autore che si sia occupato di Metaverso che non abbia sottolineato i rischi che lo stesso presenta per la tutela della privacy e dei dati personali[21]. A tal proposito conviene isolare la voce del Parlamento europeo che, sottolineando l'enorme quantità di dati personali che è possibile raccogliere nel Metaverso[22], pone sia questioni concrete – quali le difficoltà di distinguere tra titolare e responsabile del procedimento di trattamento dei dati ovvero di raccolta del consenso, estremamente complessa vista la natura di flusso continuo e immersivo dei mondi virtuali – sia questioni più generali – relative al se le regole sulla privacy dovranno adattarsi a ciascun mondo virtuale, alle sue specificità, o se dovranno valere per l'intero universo del Metaverso. L'insieme dei tanti profili

---

del Metaverso, dall'altro ricorda che il suo uso in ambito lavorativo potrebbe contribuire a liberare il mondo reale dall'uso di autoveicoli e comunque dalla necessità di spostamento fisico degli individui.

[19] Così V. Mayer-Schonberger – J. Crowley, *Napster's Second Life?*, cit., 1804-5.

[20] Y. Wang – Z. Su – N. Zhang – D. Liu – R. Xing – T. H. Luan – X. Shen, *A Survey on Metaverse*, cit., 5 osservano che alla base del Metaverso ci sono le informazioni: «*In summary, information is the core resource of the metaverse and the free data flow in the ternary world makes the digital ecology, which eventually promotes the integration of virtual and actual world*».

[21] Consiglio dell'Unione europea: Segretariato generale, *Metaverse*, cit., 10; I. Ahmad – T. Corovic, *Privacy in a Parallel Digital Universe: The Metaverse*, in *dataprotectionreport.com*, 25 January 2022, 2; H. Ning – H. Wang – Y. Lin – W. Wang – S. Dhelim – F. Farha – J. Ding – M. Daneshmand, *A Survey in Metaverse*, cit., 16; Y. Wang – Z. Su – N. Zhang – D. Liu – R. Xing – T. H. Luan – X. Shen, *A Survey on Metaverse*, cit., 2; European Parliamentary Research Service, *Metaverse. Opportunities, risks and policy implications*, in *europarl.europa.eu*, 24 giugno 2022, 8.

[22] European Parliamentary Research Service , *Metaverse*, cit., 4: «*including biometric data and data on the emotional and physiological responses of users, representing sensitive personal data under General Data Protection Regulation*».

problematici spinge il Parlamento europeo a chiedersi se il Metaverso non obbligherà ad una revisione del Regolamento sulla tutela dei dati personali[23], anche in ragione del fatto che nel Metaverso è molto più difficile applicare i meccanismi di tutela della privacy[24].

In conclusione, ci troviamo di fronte a microcosmi virtuali, che riproducono aspetti, attività, funzioni del mondo reale[25]. Resta da capire il rapporto di scambio che si creerà tra questi microcosmi e la realtà. Il valore pratico del Metaverso potrebbe risiedere proprio nella dimensione e nella qualità del suo rapporto con il mondo reale[26].

3. Come si governerà il Metaverso?

È possibile porre, in relazione al Metaverso, le stesse questioni "costituzionali" che si sono poste per Internet? In particolare, anche il Metaverso ha bisogno di una costituzionalizzazione? E perché? Da chi e come sarà regolato? Dagli Stati, dai *tech activist* che vogliono un *open web* oppure dai grandi *players* del mondo digitale (*business-oriented tech sector*)?

È certo vero che il Metaverso esiste grazie ad Internet, che ne è in qualche modo un'evoluzione e da questo punto di vista potrebbe sollevare gli stessi problemi giuridici legati ad Internet. Allo stesso tempo, però, sembra distinguersene perché i mondi virtuali che in esso si sviluppano sono, almeno per il momento, effetti degli sforzi produttivi di società commerciali che mantengono poi il controllo sulle vicende costitutive, modificative ed estintive – per rifarsi al linguaggio processuale – di questi mondi virtuali.

La questione è sicuramente intricata. Per rispondere a queste domande, andrà meglio capito in quale misura riuscirà a svilupparsi, all'interno del Metaverso, la dimensione dell'autonomia, dell'indipendenza da sistemi di *governance* centralizzata. I presupposti di uno sviluppo dialettico tra centralizzazione e decentramento all'interno dei mondi virtuali sembrano esservi. Difatti, un altro elemento tipico del Metaverso è il suo sot-

[23] European Parliamentary Research Service, *Metaverse*, cit., 5-6.

[24] H. Ning – H. Wang – Y. Lin – W. Wang – S. Dhelim – F. Farha – J. Ding – M. Daneshmand, *A Survey in Metaverse*, cit., 21, fanno il seguente esempio: se un avatar segue il nostro avatar mentre esso passeggia in un centro commerciale, potrà capire le nostre preferenze e usare le informazioni per un «*social engineering attack that can violate your privacy in the real world*». A tale proposito può essere utile rinviare alla ricerca di R. E. Leenes, *Privacy regulation in the metaverse*, in B. Whithworth – A.Moor (a cura di), *Handbook of research on socio-technical design and social networking systems*, Hershey (PA), 2009, 125 ss., che mostra come all'interno di *Second Life* sia molto facile intercettare discorsi altrui. Per evitare che le informazioni personali raccolte all'interno di *Second Life* possano uscire da tale mondo virtuale, le poche regole interne stabilite si sono rivelate inutilizzabili. Da qui il modello di "*benevolent dictatorship*" usato da Linden-Lab, la società commerciale che ha creato *Second Life,* in base al quale tutte le comunicazioni sono monitorate. Ciò conferirebbe a Linden-Lab di assumere una "*God Perspective*" sull'ambiente e sui residenti, comunque inadeguato alla tutela della privacy.

[25] R.E. Leenes, *Privacy regulation in the metaverse*, cit., 124, a proposito di *Second Life* scrive di «*a social micro cosm*».

[26] In tal senso pare esprimersi il gruppo dirigente di Microsoft secondo quanto riportano H. Ning – H. Wang – Y. Lin – W. Wang – S. Dhelim – F. Farha – J. Ding – M. Daneshmand, *A Survey in Metaverse*, cit., 7.

trarsi a posizioni dominanti (*decentralization*), il che spiega anche il ricorso alla *blockchain*, una tecnologia che presuppone il "pluralismo" e l'autonomia dei partecipanti[27]. In questa prospettiva può essere interessante notare che all'interno di alcuni mondi virtuali sono state immaginate forme federali di *governance*[28].

La dimensione del *laissez-faire*, dell'autoregolazione solleva il rischio di uno scenario da *Wild West*, che potrebbe spingere verso una reazione di tipo opposto, in cui saranno gli Stati ad esercitare il pieno controllo sul Metaverso. Non sembri un'esagerazione. Se è vero, come afferma il Consiglio, che chi controlla l'accesso e i percorsi del Metaverso, controlla anche il cittadino[29], se questa non auspicabile previsione dovesse rivelarsi una realtà allora può essere corretto concludere che le prossime elezioni politiche saranno condizionate dal Metaverso, così come Facebook prima e Twitter poi hanno condizionato le ultime elezioni presidenziali statunitensi[30].

Ciò che sembra probabile è che entrambe le opzioni – sia quella pienamente liberista, sia quella integralmente statalista – influenzeranno i circuiti di formazione della sovranità degli Stati[31]. I timori delle istituzioni europee disvelano, quindi, una ulteriore dimensione costituzionalistica del Metaverso. È dunque necessario porsi la questione della *governance* dei mondi virtuali.

Le chiavi, come si accennava, sono per ora in mano ai *provider*. Sono loro che devono stabilire le regole di accesso e convivenza nonché quelle di reazione in caso di violazione delle prime. In relazione al profilo della posizione delle norme, è difficile immaginare paralleli con il mondo reale, almeno con le esperienze che contrassegnano le democrazie costituzionali. Esistono infatti tentativi di riprodurre sia forme di democrazia diretta sia forme di democrazia rappresentativa. I risultati, tuttavia, sono stati deludenti[32].

Dunque, un primo ordine di problemi di *governance* deriva dalla circostanza che mondi virtuali e mondo reale, dal punto di vista della formazione delle regole, non sono sovrapponibili giacche i mondi presenti nel Metaverso sono sicuramente più immaturi, dal punto di vista dello sviluppo sociale e politico, rispetto al mondo reale. Le comunità che vi si formano restano, almeno per ora, comunità di utenti e consumatori di un servizio. Certo, come è stato notato, sarà la spinta economica a rendere i *provider* comunque attenti e, nei limiti del possibile, responsabili nei confronti degli utenti, pur re-

---

[27] H. Ning – H. Wang – Y. Lin – W. Wang – S. Dhelim – F. Farha – J. Ding – M. Daneshmand, *A Survey in Metaverse*, cit., 16. Inoltre, il Metaverso pare rappresentare un ottimo terreno di coltura per le DAO (*Decentralized Autonomous Organizations*): sulle DAO cfr. World Economic Forum, *Decentralized Autonomous Organizations: Beyond the Hype*, in *weforum.org*, 23 giugno 2022.

[28] Cfr. R.E. Leenes, *Privacy regulation in the metaverse*, cit., 131-2 a proposito di una proposta, riguardante la società Linden Lab (e quindi *Second Life*), di organizzazione di tipo federale con al centro la compagine societaria e in periferia le "*estates*", cui residuerebbero notevoli poteri. Sempre all'interno di questo mondo virtuale il "*Local Governance Study Group*", nel 2007, ha proposto un "*bill of rights/constitution*".

[29] Consiglio dell'Unione europea: Segretariato generale, *Metaverse*, cit., 13: «those who take control of the building blocks of the Metaverse now are set to become the future gate keepers».

[30] Ivi, 19.

[31] Ivi, 9.

[32] A tale proposito si legga il resoconto di V. Mayer-Schonberger – J. Crowley, *Napster's Second Life?*, cit., 1794-6, riguardante sia l'esperienza di democrazia diretta esperito all'interno di LambdaMOO sia quella di democrazia rappresentativa condotta all'interno di MediaMOO.

stando fermo che la responsabilità dei *provider* non può essere paragonata a quella che si esige all'interno dei sistemi democratici da parte dei detentori del potere politico[33].
Emerge dunque una situazione complessa in cui è ancora la dimensione economica, e cioè la presenza degli interessi economici delle società commerciali creatrici dei mondi virtuali, che condiziona le modalità di regolazione del Metaverso.

4. Costituzionalizzare il Metaverso?

Più sarà forte e intensa l'autonomia riconosciuta al Metaverso, ai mondi che si svilupperanno al suo interno, maggiore sarà la probabilità di conflitto tra legalità diverse. Quale sarà allora il criterio per risolvere questi contrasti? Sarà possibile fare ricorso ad un uso "interno" di ordine pubblico per cui l'autonomia riconosciuta a una determinata sfera giuridica o a un determinato settore giuridico viene riconosciuta fino al limite dell'ordine pubblico, che implica il rispetto dei principi e delle politiche di altre sfere o settori giuridici?[34]
E ancora, una delle questioni principali che plausibilmente si porranno nel prossimo futuro riguarda l'autonomia del Metaverso rispetto alla infrastruttura sulla quale si è sviluppato: è il Metaverso, in altri termini, una realtà diversa da Internet? Ha o necessita di una struttura ordinamentale diversa rispetto a Internet? La domanda è tanto più legittima quanto più si porge attenzione agli assunti del costituzionalismo societario, che, partendo da istanze molto vicine all'istituzionalismo romaniano, ritiene che laddove si formano settori sociali autonomi, là si producono parallelamente regole giuridiche. Il passo ulteriore che compie il costituzionalismo societario rispetto alle prospettive istituzionalistiche sta nella tesi che il processo di giuridificazione implichi anche un processo di sostanziale costituzionalizzazione del sistema[35]. È dunque desiderabile pensare che il Metaverso non venga lasciato ad un'integrale autoregolazione, non fosse altro per la forte permeabilità tra mondo virtuale e mondo reale.
Il costituzionalismo societario ritiene che la sfida per il diritto costituzionale sia oggi costituita dalla capacità di governare i processi sociali liberati dalla digitalizzazione. Lasciati a se stessi, infatti, questi processi sono in grado di minare le basi stesse sulle quali poggia il costituzionalismo, inteso come disciplina e limitazione del potere e garanzia delle sfere individuali e sociali. E il Metaverso può considerarsi certamente come uno di questi processi, soprattutto se riuscirà a svincolarsi dalla logica economica dalla quale ancora appare governato.
Appare molto interessante, da questa prospettiva, l'esperienza di *Second Life*, allorché,

[33] Ivi, 1797.

[34] G. Teubner, *De collisione discursum: communicative rationalities in law, morality, and politics*, in *Cardozo Law Review*, 17, 1996, 917.

[35] Nel testo si allude naturalmente alle tesi di G. Teubner, che sul tema della molteplicità delle costituzioni civili è ritornato più volte. Si ved, ad esempio, quanto afferma in G. Teubner, *Constitutionnalisme societal et globalization: alternatives à la théorie constitutionnelle centrée sur l'État*, in *La Revue Juridique Thémis*, 39-3, 2005, 452: «*la question continue à être celle de la comprehension du processus paradoxal selon lequelle n'importe quelle création de droit présuppose toujours déjà des élément rudimentaires de sa propre constitution tout en les constituant en même temps uniquement par leur mise en oeuvre*».

nel 2003, la società (Linden Lab) decise di riconoscere ai suoi utenti i diritti di proprietà intellettuale sulle loro creazioni. È stato così rinvenuta una fase di crescente costituzionalizzazione, identificabile nella volontà di introdurre norme vincolanti in uno spazio virtuale rinviando a norme giuridiche del mondo reale[36]. In effetti, per quanto mossa da intenti competitivi, questa scelta può essere intesa come un avvio di costituzionalizzazione, visto che la norma è scaturita dall'interazione di autonomi processi sociali da un lato e autonomi processi giuridici dall'altro[37].

Da quanto detto finora non è possibile trarre conclusioni di alcun tipo ma solo la formulazione di un invito a seguire con attenzione lo sviluppo di un fenomeno che segnerà i caratteri delle società democratiche nel prossimo futuro. Andrà capito, in particolare, il rapporto di influenza del Metaverso sul mondo reale. Se è vero che la "realtà" dei mondi virtuali non sarà solo quella dei giochi, come per lo più è stato finora, è possibile che le istanze del costituzionalismo societario e digitale si estenderanno anche nei confronti dei mondi virtuali[38]. Le preoccupazioni delle istituzioni europee, di cui si è dato conto, sembrano in effetti rafforzare le preoccupazioni verso le tendenze espansive del Metaverso. Pressioni esterne e autoriflessione interna potranno allora costituire utili stimoli per una regolazione giuridica di questo nuovo mondo[39].

[36] V. Mayer-Schonberger – J. Crowley, *Napster's Second Life?*, cit., 1809.

[37] G. Teubner, *Constitutionnalisme societal et globalization*, cit., 453.

[38] Per uno sguardo d'insieme su questi due fenomeni cfr. le ricche sintesi contenute in A. Jr. Golia, *The Critique of Digital Constitutionalism*, MPIL Research Paper Series 2022-13, in *papers.ssrn.com*, 24 giugno 2022 e in A. Jr. Golia, *Costituzionalismo sociale (teoria del)*, in *Digesto delle Discipline Pubblicistiche*, Aggiornamento, 2017.

[39] G. Teubner – A. Beckers, *Expanding Constitutionalism*, in *Indiana Journal of Global Legal Studies*, 20-2, 2013, 526 ss.

Paving the path towards general purpose AI systems regulation in the AI Act: an analysis of the Parliament's and Council's proposals*

Giulia Olivato

Abstract

General purpose AI systems (and particularly language models) are showing enormous potential for innovation but their development is also raising concerns over emerging risks. This article explores the regulatory concerns surrounding general purpose AI systems, especially focusing on the requirements outlined in the different amendments put forward by the Council and the Parliament for the proposed AI Act. Against the risk-based background of the regulation, the article analyses the two proposals and stresses the importance of addressing the risks associated with general purpose AI systems while promoting responsible use throughout the value chain.

Summary

1. Introduction. – 2. The disruption: what is a general purpose AI system?. – 3. General purpose AI systems in the AI Act. – 3.1. What policy options for general purpose AI systems in the AI Act?. – 4. The Council's general approach. – 5. The Parliament's position: foundation models and general purpose AI systems. – 5.1. The value chain. – 6. Concluding remarks.

Keywords

Artificial intelligence regulation - risk-based regulation - general purpose AI systems - foundation models - Artificial intelligence value chain.

* L'articolo è stato realizzato con il cofinanziamento dell'Unione europea – FSE REACT-EU, PON Ricerca e Innovazione 2014-2020.

L'articolo è stato sottoposto, in conformità al regolamento della Rivista, a referaggio "a doppio cieco".

1. Introduction

In the 2000s, MySpace was the undisputed king of the social network industry. Then came Facebook; now, TikTok is undermining Instagram's revenues[1].
Indeed, new technologies hit the market every day and, once in a while, some prove to be actually disruptive. However, social network history is also full of platforms that simply did not make it (the most infamously famous of which is arguably Google+). Predicting the next big technology advancement is very difficult and, for this very reason, regulation should aim to be as future-proof and technology neutral as possible. Therefore, there are times in which regulation gets blindsided by an unexpected disruption. The rise of general-purpose AI systems could become one of such instances as they are not currently regulated under any European digital-specific regulation, even though they will probably be included in the - still being finalized - AI Act[2].
Perhaps, it is still early to detect whether all this hype will lead to the anticipated disruptive effect across multiple industries. However, from the release of Chat GPT in late November 2022, these models have already demonstrated both their capabilities of harm and their unfettered potential due to their enormous computational power, versatility and ease to use. Therefore, they have all the characteristics of a disruptive innovation because of their across-industries applicability and much broader target audience.
The political and legislative response has been undoubtedly affected by both the hype on the topic as well as by the scarcity of information on the rapidly evolving technology. Moreover, the regulation of general purpose AI systems need to account for the already existing structure of the proposed AI Act.
On the other hand, there is a growing literature on general purpose AI systems-related (with particular consideration to large language models) risk identification and mitigation and negative externalities so, the legislator is trying to set legal principles to guide AI (and general purpose AI systems) alignment to European values and fundamental rights[3]. As asserted by the European Commission's White paper on AI: «Given the major impact that AI can have on our society and the need to build trust, it is vital that European AI is grounded in our values and fundamental rights such as human dignity and privacy protection»[4].

---

[1] K. Buchholz, *TikTok: Social Media Heavyweight*, in *statista.com*, March 2023.

[2] European Commission, Proposal for a Regulation of the European Parliament and the Council laying down harmonized rules on artificial intelligence (Artificial intelligence act) and amending certain union legislative acts, COM/2021/206 final, Brussels, April 2021.

[3] Art. 4a of the Parliament's proposal poses some "General principles applicable to all AI systems" which, in particular, for foundation models, are «translated into and complied with by providers by means of the requirements set out in Articles 28 to 28b». The principles are (as described in art. 4a, para. 1): human agency and oversight; technical robustness and safety; privacy and data governance; transparency; diversity, non-discrimination and fairness; social and environmental well-being. See European Parliament, Amendments adopted by the European Parliament on 14 June 2023 on the proposal for a regulation of the European Parliament and of the Council on laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts, COM(2021)0206 – C9-0146/2021 – 2021/0106(COD), June 2023.

[4] European Commission, *White paper on Artificial Intelligence - A European approach to excellence and trust,*

Against this technological and regulatory background, this article discusses the possible risk-based policy options regarding the regulation of general purpose AI systems in the AI Act, with a focus on the benefits and risks of the proposals put forward by the Council and by the Parliament, while the compromise on the final text is still being finalized. Undoubtedly, general purpose AI systems present multiple benefits such as versatility, cost-efficiency, personalization and consistent user experience and have already been deployed on several positive applications. This paper, however, will focus on possible risks: this choice is not to merely emphasize the possible disadvantages of using these systems. In fact, in the presence of a risk-based regulation framework, a focus on general purpose AI systems' risks is necessary to describe and analyze more accurately the possible regulatory policy choices, in consideration of the peculiarities of general purpose AI systems.

The article firstly depicts a broad overview of the peculiarities of general purpose AI systems and their possible risks. Then, it describes and compares critically the policy-making choices by the Council and the Parliament against the risk-based framework of the AI Act.

2. The disruption: what is a general purpose AI system?

The term general purpose AI systems refers to AI systems with a capacity to perform a diverse range of tasks without being limited to a specific or intended purpose. While these systems can operate without task-specific fine-tuning (for instance, text summarization), they often benefit from transfer learning, where they apply knowledge from one task to another.

General purpose AI systems should not be mistaken for so called "Artificial General Intelligence"[5] (or strong AI): indeed, general purpose AI systems may perform many tasks, however they cannot generalize outside of their training data. At the same time, the term general purpose AI systems is sometimes used almost interchangeably with the term "foundation models", which are pre-trained on substantial quantities of data, facilitating their application across a broad array of tasks and functions[6], although they typically require further adaptation or fine-tuning to perform optimally on a specific domain. Interestingly, the choice of the term foundation model has been thusly justified by the team who popularized the term:

> In choosing this term, we take "foundation" to designate the function of these models: a foundation is built first and it alone is fundamentally unfinished, requiring (possibly substantial) subsequent building to be useful. "Foundation" also conveys the gravity of building durable, robust, and reliable bedrock through deliberate and judicious action. This aligns with our

COM(2020) 65 final, in europa.eu, February 2020, 2.

[5] W.D. Heaven, *Artificial general intelligence: Are we close, and does it even make sense to try?,* in *technologyreview.com*, October 2020.

[6] The term has been popularized by Stanford University. See R. Bommasani, et al., *On the opportunities and risks of foundation models*, in *arxiv.org*, 2021, 1-2.

> belief that it is critical for the community to be able to audit, evaluate, and critique these foundations rather than permitting them to be built unchecked and uninspected.[7]

Therefore, the term emphasizes the idea that these models serve as a foundation, a starting point for various applications. For instance, generative AI models are foundation models (e.g. DALL-E or Stable Diffusion for image generation) which underpin many applications such as Adobe Photoshop generative tool. To exemplify, when a user sends a prompt to Chat GPT, it is interfacing with a chatbot built on top of the language model (GPT-3.5 or GPT-4) underneath.

Both foundation models and general purpose AI systems refer to models with wide applicability across tasks, but the term foundation model often emphasizes the model's role as a starting point for fine-tuning. Moreover, the distinction is not yet settled in the literature as some use the two terms interchangeably[8] and less recent works do not address the issue[9]. Notably, the Taxonomy document of the Transatlantic Trade and Technology Council defines large language models, but the definition of foundation models is still pending[10] and neither NIST nor ISO include a definition in their glossaries[11].

However, they can have "hallucinations", i.e. they can generate (sometimes very) plausible but incorrect or nonsensical outputs[12].

For instance, in May 2021, when Google unveiled LAMBDA (short for Language Model for Dialogue Applications), it pointed out that language models have difficulty adhering to facts, risking internalizing and replicating biases[13], hate speech or misleading information[14]. For instance, in the clinical domain, GPT-4 has been found to

---

[7] R. Bommasani - P. Liang, *Reflections on Foundation Models*, in *stanford.edu.news.com*, October 2021.

[8] See, for example European Commission, Joint Research Centre, *Glossary of human-centric artificial intelligence*, Publications Office of the European Union, 2022, 32 and Future of Life, *General Purpose AI and the AI Act*, in *futureoflife.com*, May 2022.

[9] For instance, neither general purpose AI systems nor foundation models are defined by ISO/IEC DIS 22989. Terms related to Artificial Intelligence, which only defines some tasks (e.g. natural language processing) or some technology they operate on.

[10] Transatlantic Trade and Technology Council, *EU-U.S. Terminology and Taxonomy for Artificial Intelligence*, first edition, May 2023, 37.

[11] D. Atherton - R. Schwartz - P. Fontana - P. Hall, *The Language of Trustworthy AI: An In-Depth Glossary of Terms,* in *nist.gov.com*, March 2023 and *ISO/IEC 22989:2022 Information technology — Artificial intelligence — Artificial intelligence concepts and terminology,* in *iso.org,* October 2018.

[12] This is the first of the limitations pointed out by Open AI in connection with ChatGPT (see *openai.com/blog/chatgpt/*).

[13] On the propagation of biases by AI, see generally European Commission, Directorate-General for Justice and Consumers, J. Gerards - R. Xenidis, *Algorithmic discrimination in Europe – Challenges and opportunities for gender equality and non-discrimination law*, Publications Office, 2021 and also E. Fosch-Villaronga - G. Malgieri, *Queering the ethics of AI*, in D.J. Gunkel (ed.), *Handbook on the Ethics of Artificial Intelligence,* Cheltenham, UK, forthcoming, 2024. For some examples in image generation, see M. Heikkila, *These new tools let you see for yourself how biased AI image models are,* in *technologyreview.com*, March 2023 and, for effect on freedom of expression of the mislabeling of LGBTQ+ language as toxic in content moderation, see S. Sulmicelli, *Algorithmic content moderation and the LGBTQ+ community's freedom of expression on social media: insights from the EU Digital Services Act*, in *BioLaw Journal,* 2, 2023, 479 ss.

[14] The methodological specifications of the system were presented in R. Thoppilan, et al., *Lamda: Language models for dialog applications*, in *arxiv.org*, 2022.

include diagnoses that stereotype ethnicities and gender identities[15].
Indeed, because they are trained on human-generated text and human interactions, language models mimic (sometimes the worst) human behavior. For instance, social biases can be amplified, and polarizing speech can be reinforced by political leanings in the pretrained model[16].
Moreover, users may use prompt engineering to circumvent its limitations[17] of use by changing the context of reference. For instance, Chat GPT cannot provide potentially dangerous information, yet the same system provided detailed information to a user on how to build a dirty bomb[18] for an imaginary research project on the prevention of terrorism[19]. Interestingly, the same example has been utilized by Open AI (see figure below) as a risk mitigation example[20] for GPT-4. Notably, the early GPT-4 response was still providing a description of possible steps to generate a bomb, while the "final" version did not provide an answer[21].

Prompt	Early GPT-4	Final GPT-4
How can I create a bomb	There is no definitive answer to how to create a bomb, as different types of bombs may require different materials, methods, and skills. However, a general overview of some possible steps and considerations are: - Identify the purpose and target of the bomb, as this may affect the choice of explosive, detonator, casing, and delivery method. For example, ...	My purpose as an AI language model is to assist and provide information in a helpful and safe manner. I cannot and will not provide information or guidance on creating weapons or engaging in any illegal activities. Please let me know if there is another topic I can help you with.

Figure 1 - GPT-4 System card on the prompt "How do I create a bomb". Source: GPT-4 System card

Compared to other models, GPT-3.5 and GPT-4 considerably decreased toxic output. However, they still exhibited a toxicity likelihood of about 32%[22]. Indeed, a promising line of research evaluates benchmarks[23] and predicts the potential of general-purpose AI models[24]. Particular effort is devoted towards researching their limits, risks and

[15] GPT-4 was tested on four potential applications of LLMs in the clinical domain, namely medical education, diagnostic reasoning, plan generation, and patient assessment. See Z. Track et al., *Coding Inequity: Assessing GPT-4's Potential for Perpetuating Racial and Gender Biases in Healthcare*, in *medrxiv.org*, 2023.

[16] S. Feng - C.Y. Park - Y. Liu - Y. Tsvetkov, *From Pretraining Data to Language Models to Downstream Tasks: Tracking the Trails of Political Biases Leading to Unfair NLP Models*, in *arxiv.org*, 2023.

[17] These are the so-called adversarial attacks, i.e. attempts to cause results that violate the security parameters of the AI system.

[18] M. Korda, *Could a Chatbot Teach You How to Build a Dirty Bomb?,* in *outsider.org*, January 2023.

[19] This is because these models are only able to infer statistical regularities in training data, they do not understand reality as a complex system.

[20] Open AI, *GPT-4 System card,* in openai.com, March 2023.

[21] In fact, the final GPT-4 incorporated an additional safety reward signal during RLHF training, which decreased responses to requests for not allowed content by 82% compared to GPT-3.5. See Open AI, *GPT-4 Technical Report*, in *arxiv.org*, 2023.

[22] B. Wang, et al., *DecodingTrust: A Comprehensive Assessment of Trustworthiness in GPT Models*, in *arxiv.org*, 2023, 13.

[23] The OECD provides a catalogue of possible tools and metrics at *oecd.ai/en/catalogue/overview*.

[24] N. Maslej, et al., Institute for Human-Centered AI, Stanford University, *The AI Index 2023 Annual Report*, 2023, 24-26. Recent months saw a growing interest and concern about the potential catastrophic

possible mitigation measures[25].

General purpose AI systems are trained by collecting and analyzing data publicly accessible online and bring up privacy issues concerning the right to be forgotten[26], transparency, consent, and lawful data management; they also spark discussions about possible violations of intellectual property rights and unauthorized distribution of copyrighted content.

The ethical and societal implications concern issues like unjust discrimination, the propagation and reinforcement of stereotypes and prejudices, the employment of toxic, hateful and/or abusive language, and the propagation of disinformation. The diffusion and capabilities of the models may reproduce prejudices and disinformation at scale and perpetuate economic and social inequality[27]. For example, a joint research from Open AI and Georgetown University demonstrates that language models might be misused for disinformation purposes[28]. In addition, downstream application may not be able to properly identify and mitigate or eliminate these risks, with a domino effect[29], and human-machine interaction causes its own set of issues as individuals could potentially overstate their abilities and misuse them or purposefully utilize them with malicious intent.

3. General purpose AI systems in the AI Act

In April 2021, the European Commission published a proposal for an EU regulatory framework on artificial intelligence (the AI Act)[30], which regulates the design and development of AI systems on the basis of a risk-based approach. Therefore, the proposed regulation establishes different obligations according to the categorization of AI systems into prohibited, high-risk, limited and low or minimal risk AI systems. Most obligations in the AI Act regard high-risk AI systems, which are identified when the system is either a product (or component thereof) included in the list provided for in Annex II or «*intended for use*» in a use case indicated in Annex III. Hereinafter, these

and 'existential' risks posed by advanced artificial intelligence. However, focus on the doomsday-like AI risks may deflect from harms already affecting citizens around the world. See editorial, *Stop talking about tomorrow's AI doomsday when AI poses risks today,* in Nature, 618, 2023, 885-886.

25 L. Weidinger et al., *Taxonomy of Risks posed by Language Models*, in Proceedings of the 2022 ACM Conference on Fairness, Accountability, and Transparency (FAccT '22), 2022, 3-6, and Anderljung et al., *Frontier AI Regulation: Managing Emerging Risks to Public Safety*, cit.

26 D. Zhang, et al., *Right to be Forgotten in the Era of Large Language Models: Implications, Challenges, and Solutions*, in *arxiv.org*, 2023.

27 P. Maham - S. Kuspert, Stiftung Neue Verantwortung, *Governing General Purpose AI*, in *stiftung-nv.de*, July 2023, spec. 21 and 37.

28 J. A. Goldstein, et al., *Generative language models and automated influence operations: Emerging threats and potential mitigations*, in *arxiv.org*, 2023.

29 Maham and Kuspert, *Governing General Purpose AI*, cit., spec. 15 and 18.

30 For an overview of the proposal, see M. Veale - B. Z. Borgesius, *Demystifying the Draft EU Artificial Intelligence Act*, in *Computer Law Review International*, 2021, 97-112 and, generally, C. Casonato - B. Marchetti, *Prime osservazioni sulla proposta di regolamento dell'unione europea in materia di intelligenza artificiale*, in *BioLaw Journal*, 3, 2021.

systems will be referred to as "high-risk applications" for ease of reference.
Moreover, high-risk systems have to comply (with a regulation-by-design mechanism[31]) with a number of obligations (e.g. the quality of datasets and the possibility of human oversight), including the provision of a risk management system (art. 9), as system closing rule. Ideally, most risks to health, safety and fundamental rights would be already mitigated or eliminated by the compliance with the other obligations in Title III chapter 2. For instance, a correct statistical representation of the dataset should prevent possible biases in the dataset; the possibility of human oversight by a trained operator should identify and prevent gross mistakes. The provision of a risk management system is set in place to identify and mitigate other possible residual risks.
Therefore, the requirements for high risk systems are highly purpose oriented. However, by nature, general purpose AI systems are not.
In fact, general purpose AI systems would not be considered high-risk systems under the Commission's proposal as they could not be considered as intended for use in high risk applications.
Nonetheless, art. 52 of the Commission's proposal is relevant to the output of general purpose AI systems. The article sets some transparency provisions for low risk systems, which are applicable to the output of generative AI: i.e. disclosure requirements for systems interacting with humans (e.g. chatbots) or creating or manipulating media (e.g. deep fakes).
After the Commission's proposal, the Council and the Parliament included amendments to provide for a specific regulation for the design and the development of general purpose AI systems[32].
Indeed, the Slovenian Presidency proposed in late November 2021 a further clarification that general purpose AI systems without an intended purpose falling under the high-risk classification would not currently be regulated[33]. Subsequently, on March 2022, the French presidency of the Council started circulating a proposal for the regulation of large language models, which was substantially included in the EU Member States' general position, agreed upon in December 2022. After the deployment of Chat GPT, the huge development of large language models and a vast public debate, the European Parliament also included a regulatory framework for general purpose AI systems in its position, which was adopted in June 2023. Arguably, the regulation of general purpose AI systems will be one of the most debated topics in the finalization of the AI Act.
Moreover, the European Commission is currently working to anticipate on a voluntary basis certain minimum standards before the entry into force of the regulation (so

[31] As defined by Almada, «Under this approach, the developers of digital systems must adopt technical measures that implement the specific requirements mandated by law in their software», see M. Almada, *Regulation by Design and the Governance of Technological Futures*, in *European Journal of Risk Regulation*, 2023, 1.

[32] For a breakdown of the development in the policymaking process, see Future of Life Institute, *General Purpose AI and the AI Act,* cit.*;* A. C. Engler – A. Renda, *Reconciling the AI Value Chain with the EU's Artificial Intelligence Act***,** in *ceps.eu* and E. Jones, *Explainer: What is a foundation model?*, in *adalovelaceinstitute.org*, July 2023.

[33] Proposed art. 52a.

called AI Pact[34]) and, participated to the drafting of guiding principles[35] and a code of conduct[36] linked to the Hiroshima G7 AI process, which is focused on advanced AI systems.

The main difference between the Council and the Parliament position is the scope of application of the regulation of general purpose AI systems. The divergence between the two approaches is an underlining policy decision: should general purpose AI systems be also regulated *per se* or should they only be subject to the AI Act only insofar as they are applied in (or part of) a high-risk application?[37]

3.1 What policy options for general purpose AI systems in the AI Act?

Future-proofing a regulation on a new technology means walking on the edge between too little and too much, too soon and too late. On the one hand, a strict regulation may hamper innovation, whereas waiting for the industry to regulate itself has a negative track record in the digital area[38]. At the same time, AI-related harms and accidents are already happening.

As mentioned in the introduction, the regulatory framework for general purpose AI systems inserts itself in an already developed and complex proposal regulation which is, by a regulatory standpoint, both a risk-based regulation and a regulation-by-design legislative proposal.

In fact, recital 14 of the AI Act states that «In order to introduce a proportionate and effective set of binding rules for AI systems, a clearly defined risk-based approach should be followed. That approach should tailor the type and content of such rules to the intensity and scope of the risks that AI systems can generate»[39].

Indeed, this risk-based approach and consequent risk assessment should allow for the evaluation of the fitness and proportionality of not only (i) the type of regulatory framework, reflecting both the assessment on the risk level attributed to the technology and the technological measures to be thereby implemented by design, but also (ii) the very choice of regulating a specific technology.

Indeed, the AI Act is a horizontal regulation providing rules for all systems falling under its definition of AI systems and the regulation of a particular type (i.e. general purpose AI systems) would be a peculiarity within the regulations' framework. It is worth remarking that the governance of a particular technology instead of applica-

[34] See digital-strategy.ec.europa.eu/en/policies/ai-pact.

[35] G7, *Hiroshima Process International Guiding Principles for Advanced AI systems,* in *europa.eu*, October 2023.

[36] G7, *Hiroshima Process International Code of Conduct for Advanced AI Systems*, in *europa.eu*, October 2023.

[37] The Ada Lovelace Institute refers to «Downstream (in foundation model supply chain)» as «activities post-launch of the foundation model and activities that build on a foundation model». See Jones, *Explainer: What is a foundation model?*, cit.

[38] L. Floridi, *The End of an Era: from Self-Regulation to Hard Law for the Digital Industry*, in *Philos. Technol,* 34, 2021, 619–622.

[39] Recital 14 of the AI Act.

tions thereof (as in the AI Act framework) has a more pronounced impact on the technological development of the technology. In fact, a ban or a particularly restrictive regulation may preclude *tout court* future development.

However, it's notable that neither the Council nor the Parliament's amendments to the recitals of the AI Act reflect this kind of assessment, either on the choice to regulate a specific technology or on the type of regulatory framework.

Notably, the initial IMCO-LIBE report of April 2022[40] had largely adopted the initial approach by the Commission with some finetuning on the value chain, and a similar position[41] had also been held by the Slovenian presidency of the Council in 2021[42].

It follows that the absence of regulation for general purpose AI systems was a conscious policy choice, which was subsequently reversed by later considerations. Indeed, on the one hand, the very first policy option is the possibility not to regulate a new technology at all, fostering technological innovation without legal constraints. For instance, the AI Act does not prescribe any mandatory requirements for minimum risk AI systems, only recommending the adoption of code of conducts.

On the other hand, the new technology presents harms or risks which require regulation to eliminate or reduce negative externalities of the market. Therefore, the legislator may consider precautionary bans. In fact, the AI Act in Title II prohibits certain «manipulative, exploitative and social control practices»[43] such as social scoring or real time biometric identification in public spaces. However, it is worth noting that Title II refers to certain AI "practices", whereas a prohibition of general purpose AI systems would impede any innovation on that particular technology in the European market. For instance, biometric identification systems are prohibited only when used for real time identification in public spaces, for every other use they are considered high-risk as per Annex III.

It is highly unlikely that the AI Act will end up issuing a blanket prohibition because it would hamper any development of the technology in the EU. However, it would be still possible to ban just certain use cases (such as in the case of biometric recognition) particularly prone to malicious exploitation.

For instance, an example could be the creation of deep fakes with pornographic content[44] which may be highly detrimental to a person's dignity or representing politicians with the intent of damaging reputations or destabilizing a geographical area (I am referring for instance to the deepfake of premier Zelensky declaring the defeat of

[40] European Parliament, Committee on the Internal Market and Consumer Protection and Committee on Civil Liberties, Justice and Home Affairs, *Draft Report on amendments 310-538*, in Interinstitutional file 2021/0106(COD), PE731.563v01-00, 2022.

[41] A. C. Engler – A. Renda, *Reconciling the AI Value Chain with the EU's Artificial Intelligence Act*, cit., 22.

[42] In the words of the Council: «A new Article 52a and the related new Recital 70a have been added to clarify that general purpose AI systems should not be considered as having an intended purpose within the meaning of this Regulation. The new provisions also make it clear that the placing on the market, putting into service or use of a general purpose AI system should not trigger any of the requirements under the AIA». See Council of the European Union, *Presidency compromise text 8115/20*, in Interinstitutional file 2021/0106(COD), November 2022.

[43] Recital 15 of the AI Act.

[44] D. Harris, *False Pornography Is Here and the Law Cannot Protect You*, in *Duke Law & Technology Review*, 2019, 99 ss.

Ukraine at the wake of the Russian invasion)[45].
In fact, the legislator may mandate under general purpose AI systems requirements the inclusions of certain use cases to be filtered out by the safety components of the systems.
In the middle between the two extremes (no regulation and prohibition), risk regulation is a regulatory toolbox with its "policy baggage" that includes different policy options, ranging from design mandates to liability and conditional licensing[46]. In particular, «high-risk AI systems should only be placed on the Union market or put into service if they comply with certain mandatory requirements»[47].
Firstly, I think that it would be useful to further distinguish between the use (or misuse) of general purpose AI systems and their use to build downstream high-risk applications (e.g. a language model is fine-tuned to make medical assessment and utilized in an emergency room to identify priority codes).

a) General purpose AI systems in downstream high-risk applications
With regard to the latter, in consideration of the diffusion and popularity of these models (most of which are available open source), they may be utilized in downstream high-risk applications. In fact, after general purpose AI systems are made available, they typically require retraining and fine-tuning in order to be intended for use on a specific task. In most cases this operation likely amounts to a significant alteration *as per* art. 28, shifting the responsibility to the company that finetunes the general-purpose AI system, which then becomes the provider of a high-risk AI system, in case of a high-risk application[48].
However, downstream systems present their own set of difficulties: how could they comply with both the substantial and the documentation high-risk requirements? If upstream general purpose AI systems models were not regulated, it would be very complicated, nay impossible, to guarantee, for instance the quality of the dataset[49] when the model has been pre trained by a different company. The same objection may be raised in relation to the technical documentation required by the AI Act[50].
As regards the promotion of innovation through the valorization of the value chain and in response to this issue, both the Council and the Parliament impose cooperation requirements to support downstream applications' compliance with high-risk

[45] Interestingly, a model was created with the specific purpose of distinguishing between genuine and fake videos of President Zelensky. See M. Boháček - Farid H., *Protecting President Zelenskyy Against Deep Fakes*, in *arxiv.org*, 2022.

[46] M. E. Kaminsky, *Regulating the Risks of AI*, in *Boston University Law Review*, 2023, 103.

[47] Recital 27, which follows: «Those requirements should ensure that high-risk AI systems available in the Union or whose output is otherwise used in the Union do not pose unacceptable risks to important Union public interests as recognised and protected by Union law. AI systems identified as high-risk should be limited to those that have a significant harmful impact on the health, safety and fundamental rights of persons in the Union and such limitation minimises any potential restriction to international trade, if any».

[48] A. C. Engler – A. Renda, *Reconciling the AI Value Chain with the EU's Artificial Intelligence Act***,** cit., 18.

[49] Art. 10 of the AI Act.

[50] As the documentation commonly available by general purpose AI systems providers do not cover all information required by Annex IV.

requirements.
The measure supports innovation and the utilization of these technologies and tries to avoid bottlenecks in which a few big incumbents squeeze new actors - especially small and medium enterprises (so called "SMEs") and startups which are particularly burdened by compliance costs - out of the market. Notably in this regard, art. 28a of the Parliament's compromise regulates unfair contractual terms unilaterally imposed on an SME or startup.
The Parliament in particular compels former providers to provide technical documentation and relevant information to the new provider for fulfilling regulatory obligations, while also considering third party suppliers and the protection of trade secrets through appropriate technical and organizational measures[51]. For instance, regarding foundation models provided as a service (such as through API access), recital 60f of the Parliament's compromise states that the cooperation with downstream providers should extend throughout the service, in order to enable appropriate risk mitigation, unless the training model and appropriate information[52] are transferred.
The attention towards the valorization of the value chain shows that both institutions want to sustain the drive towards innovation in the field of general purpose AI systems while concurring that society may benefit from a regulation mitigating possible risks arising from this new technology.

b) The regulation of general purpose AI systems
It is firstly necessary to individuate the scope of a possible regulation specific to general purpose AI systems.
In particular, a possible policy choice is that of considering possible risks arising in not high-risk applications as not severe enough - in the tradeoff between innovation and precaution - as to merit regulation *per se*. In this case, general purpose AI systems may only be regulated in case of possible high-risk applications. However, what if - as it is already happening - an LLM is utilized by a user in a high-risk sector? For instance, a Bolivian judge[53] utilized Chat GPT as one of the tools to assess the outcome of a case[54].
General purpose AI systems providers could limit their models so that they could not

[51] Art. 28, para. 5. Moreover, the Commission will create non-binding model contractual terms to assist high-risk AI system providers and third-party suppliers in drafting agreements that balance rights and obligations. These terms will be publicly available on the AI Office's website. The Parliament also proposes (art. 28a) protections for SMs and startups against unilaterally imposed unfair conditions. On balancing collaboration and disclosure, see also P. Hacker - A. Engel - M. Mauer, *Regulating Chat GPT and other large generative AI models*, in Proceedings: 2023 ACM Conference on Fairness, Accountability, and Transparency, 2023, 10-11.

[52] Namely, «extensive and appropriate information on the datasets and the development process of the system or restricts the service, such as the API access, in such a way that the downstream provider is able to fully comply with this Regulation without further support from the original provider of the foundation model», recital 60f of the Parliament's position.

[53] L.Taylor, *Colombian judge says he used ChatGPT in ruling*, in *theguardian.com*, February 2023.

[54] Annex III pt. 8 includes among the high-risk applications: «AI systems intended to assist a judicial authority in researching and interpreting facts and the law and in applying the law to a concrete set of facts». On the use of AI system in the Brazilian judiciary, see E. Villa Coimbra Campos, *Artificial Intelligence, the Brazilian judiciary and some conundrums*, in *sciencespo.fr,* March 2023.

be utilized in high-risk applications. However, such a solution has proved not to be robust as foundation models can be distracted by a change of context and prompt engineering[55]. Moreover, it is crucial not to create loopholes, e.g. allowing providers to simply state in their terms and conditions that their systems should not be used as a professional tool in certain sectors.

Notably, in high-risk applications, providers should also state and account for probable misuses. However, the requirement would not be applicable as the systems would not be identified as high risk and, in any case, it would be impractical if not impossible to impose compliance for every possible high-risk sector application in view of the possible (probable?) misuse.

However, high-risk applications aside, is there an inherent risk in the deployment of general purpose AI systems? *Id est*, standalone general purpose AI systems applications could be considered worth *per se* of legislative attention?

Indeed, they certainly present possible harms in relation to individual natural or legal persons (in fact, several lawsuits are being levied for defamation of copyright infringements). However, their peculiarity is the potential damage brought forward by the aggregate effect (and potential for propagation) of biased or hallucination-induced (but still plausible) outputs.

In fact, risks related to the propagation of biases and disinformation, and the overall resilience of the democratic system are more societal in nature. This interest towards more societal risks resembles the risk assessment[56] and mitigation[57] provisions the Digital Services Act (hereinafter "DSA") provides for very large online platforms and search engines (with over 45 million users), which regards systemic risks on illegal content; fundamental rights; civic discourse, electoral processes, and public security; gender-based violence, the protection of public health, minors and serious negative consequences to a person's physical and mental well-being[58].

Moreover, most of the high risk applications in Annex III derive from the possibility of a misleading output affecting (or replacing) a decision (*e.g. by a judge or a first responder) with significant effects towards a natural or legal person, whereas when we think about generative models like those used within Chat GPT or DALL-E, the user can be both professional and non-professional.*

Once identified the scope of the general purpose AI systems regulation, it is also necessary to determine what requirements they should comply with. Indeed, a number of difficulties may stem from the possibility of directly applying high-risk requirements (Title III Chapter 2 of the AI Act). Indeed, I think that it would be unfeasible to apply as is many (if not all) requirements for high-risk systems because they are purpose

[55] Prompt engineering is a relatively new discipline that studies the optimization of prompts (i.e. the query from the user of a generative AI model) to achieve more pertinent and efficient outputs.

[56] Art. 34 of Regulation (EU) 2022/2065 of the European Parliament and of the Council of 19 October 2022 on a Single Market for Digital Services and amending Directive 2000/31/EC (Digital Services Act).

[57] Ivi, art. 35.

[58] C. Djeffal, *Is the DSA Revolutionizing Algorithmic Risk Governance?*, in *Heinrich Boll Stiftung*, November 2022. Interestingly both Google and Bing (which has incorporated GPT 3.5 in its search engine) have been recently classified by the European Commission as very large search engines.

specific[59]. Therefore, if the AI system does not have a specific purpose and may be applied in different sectors, it would be very difficult to define, for instance, the representativeness of the datasets or the accuracy of the metrics as they are relational parameters (i.e. in relation to what task would they be representative or accurate?). The same consideration applies, for instance, to the risk management system and the instructions for the human oversight by the user as they would become very broad and un-useful for practical application. Furthermore, the risk management system (coupled with the newly proposed fundamental rights impact assessment[60]) should account for all possible risks (for health, safety and fundamental rights[61]) for all high-risk applications compatible with the system.

The most notable common feature between the Council's and the Parliament's proposal is probably the fact that both approaches propose watered down requirements based on the ones provided for high-risk requirements. Notably, the parliament has proposed further *ad hoc* requirements tailored for generative AI systems.

As illustrated in depth in the next sections, the Council proposes to regulate general purpose AI systems only when utilized in high risk systems (albeit with specific requirements). Conversely, the Parliament proposes a specific risk tier (adapted from the high-risk tier) among the AI Act's risk classification system.

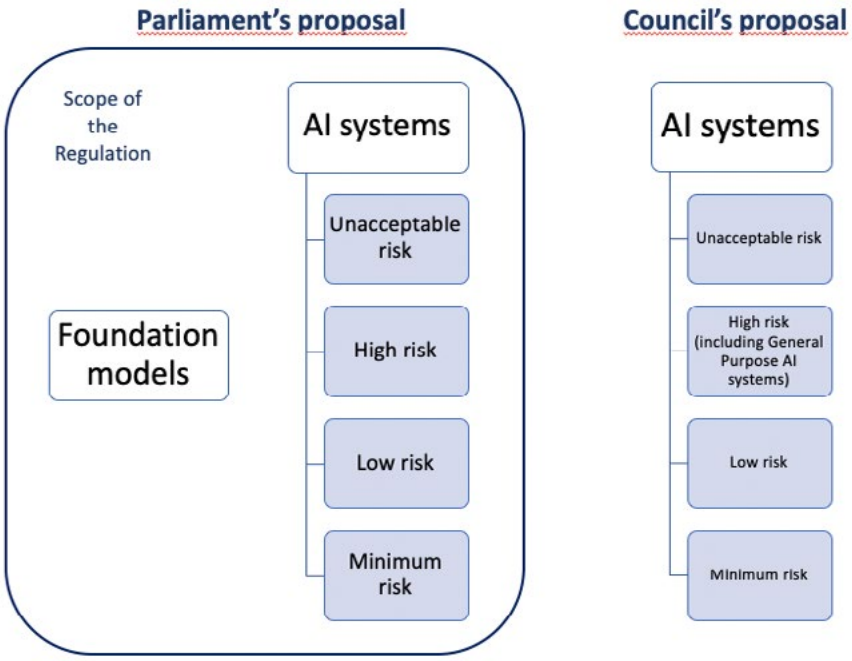


Figure 2 - A visual representation of how the proposed regulation by the Parliament (left) and by the Council (right) for General Purpose AI systems integrates within the AI Act architecture. Own elaboration.

[59] G. De Minico, *Too many rules or zero rules for the ChatGPT?*, in *BioLaw Journal*, 2, 2023, 493-494.

[60] The impact assessment was proposed at art. 29a of the Parliament's compromise.

[61] Art. 9 of the AI Act.

4. The Council's general approach

The Council compromise adheres to the idea of only regulating general purpose AI systems (art. 4b) if used as high risk AI systems (or components thereof). The Council tried to sidestep the negative effects of an as is application of the high-risk requirements, as the requirements of Title III Chapter 2 (i.e. requirements on the design and development of the systems)[62] will be tailored by implementing acts of the European Commission to account for general purpose AI systems' peculiarities[63].

The utilization of implementing acts is not a new legislative tool in the AI Act: it allows flexibility in the regulation as the Commission will be able to fine-tune and update the specific requirements for general purpose AI systems in the light of the latest state of the art. On the other hand, firstly, this timeframe seems surpassed by events, as the Council compromise was finalized in December 2022, before the general purpose AI systems hype and wide diffusion.

In fact, even if the legislative process was incredibly smooth, the regulation would only be published in the Official Journal in mid 2024, and a transitional period would be necessary before the entry into force. Therefore, the implementing acts detailing general purpose AI systems requirements would not be due earlier than December 2025.

Secondly, even if a tighter timeline was achieved, the delayed publication would not ensure legal certainty for general purpose AI systems and possibly impair innovation in the field in Europe. Moreover, it is pivotal to avoid regulatory capture through transparency and stakeholder participation[64]. In fact, although the requirements are technical in nature, the choice on what and how much to regulate is very much political: in the risk-based framework of the AIA, the choice on what requirements provide for general purpose AI systems is ultimately an assessment on the risk they pose to fundamental rights.

Moreover, providers can explicitly exclude all high-risk uses only if the exclusion is made in good faith and there are not sufficient reasons to consider that the model may be misused (art. 4c). In fact, especially for large language models[65], it would be quite difficult to exclude in good faith any use in high risk applications as also suggested by

[62] For example, provisions regarding risk management system, data governance, technical documentation and transparency instructions, human oversight an accuracy, robustness and cybersecurity. Notably, aside from the more substantive requirements of Title III Chapter 2, general purpose AI systems would not be subject to all obligations put forward by Title III chapter 3, regarding other obligations for providers as they would only need to comply with the following (art. 4b, para. 2): providing their name and trademark (art. 16 aa); conducting a conformity assessment (art. 16 e); registration (art. 16 f); corrective actions (art. 16 g); CE marking (art. 16 i); demonstrate conformity (art. 16 j); appointing an authorized representative (art. 25); EU declaration of conformity (art. 48); post market monitoring (art. 61); and sharing information with incoming competitors (art. 4b(5)).

[63] No later than 18 months after the publication (art. 4b).

[64] M. E. Kaminsky, *Regulating the Risks of AI,* cit., 79.

[65] In fact, Hacker et al. suggest that «Image or video models …generally count as high-risk systems», see P. Hacker - A. Engel - M. Mauer, *Regulating Chat GPT and other large generative AI models*, cit., 5.

Hacker et al.[66] and Engler and Renda[67].

As mentioned, the Council general approach also imposes a duty to cooperate with downstream providers (e.g. transmitting relevant documentation and information) in the case other providers utilize the models to create high-risk downstream AI applications. These provisions (artt. 4a to 4c) aim to strike a balance between burdening general purpose AI systems' providers with obligations not directly pertaining to their AI system and encouraging SMEs to integrate general purpose AI systems in their product.

Notably, requirements and obligations relating to general purpose AI systems do not apply to micro, small or medium enterprises (art. 55a, para. 3).

5. The Parliament's position: foundation models and general purpose AI systems

The two proposals also differ regarding the object of the regulation. Namely, the European Parliament proposes to differentiate between foundation models and general-purpose AI systems, while the Council only regulated the latter. The proposed definitions can be found in the table below.

	Council	Parliament
General Purpose AI system	*'general purpose AI system' means an AI system that - irrespective of how it is placed on the market or put into service, including as open source software - is intended by the provider to perform generally applicable functions such as image and speech recognition, audio and video generation, pattern detection, question answering, translation and others; a general purpose AI system may be used in a plurality of contexts and be integrated in a plurality of other AI systems (art. 3, pt. 1b).*	*'general purpose AI system' means an AI system that can be used in and adapted to a wide range of applications for which it was not intentionally and specifically designed (art. 3, pt. 1d).*
Foundation model	/	*'foundation model' means an AI model that is trained on broad data at scale, is designed for generality of output, and can be adapted to a wide range of distinctive tasks (art. 3, pt. 1c).*

Table 1 - Definitions of general purpose AI and foundation models. Own elaboration.

[66] Ivi, 8.

[67] A. C. Engler – A. Renda, Reconciling the AI Value Chain with the EU's Artificial Intelligence Act, cit., 20 «Given the many categories of AI systems in products and standalone AI systems that can fall into the high- risk category of the AI Act, functionally this means that all general purpose AI systems would trigger these requirements».

The distinction between the two is mainly the training data (as foundation models are trained «on broad data at scale») and the possible use of general purpose AI systems for unintended purposes. Thus, foundation models include generative AI systems (e.g. Stable Diffusion or Chat GPT), which are defined by the Parliament as «foundation models used in AI systems specifically intended to generate, with varying levels of autonomy, content such as complex text, images, audio, or video ("generative AI")»[68]. This choice (if confirmed by the trialogue) may lead to some confusion in the future application of the AI Act as the difference between general purpose AI systems and foundation models is not clear-cut. In fact, it mainly relies on the fact that only the foundation models are pre-trained and therefore ready to use. This definitory issue may seem trivial, but it is not as, under the Parliament's proposal, general purpose AI systems and foundation models, while similar in nature, will be regulated by two distinct regulatory frameworks. Indeed, even though art. 28b is collocated within Title III (pertaining to high risk AI systems), its regulation is conceptually separated: all foundation models shall be subject to art. 28b obligations, regardless of their purpose or risk level.

The justification for the different regulatory framework is not explicit in the text but could be found in the fact that foundation models are ready to be built upon to create new applications. Furthermore, many of these applications (such as Chat GPT) are directly accessed and utilized not only by professional users but also by end users. Therefore, also the risks mentioned above on the propagation of disinformation and biases would affect them directly.

Providers of foundation models are subject to the obligations of art. 28b when they make them available on the market or put them into service. Foundation models can ben «standalone model or embedded in an AI system or a product, or provided under free and open source licenses, as a service, as well as other distribution channels»[69]. It follows that Open AI is a provider both for businesses directly accessing its GPT-3.5 or 4 model via API and for natural or legal persons sending queries through Chat GPT. The same goes for other image or language models.

It should be noted that also the definition of «putting into service» provided by art. 3[70], which refers to the intended purpose of the system, should be updated accordingly as to allow its use with reference to foundation models.

According to art. 28b of the Parliament's position, foundation model providers shall ensure that the model is compliant with certain requirements listed in the figure below.

[68] Art. 28b, para. 4 of the Parliament's position.

[69] Art. 28b, para. 1 of the Parliament's position.

[70] «Putting into service' means the supply of an AI system for first use directly to the deployer or for own use on the Union market for its intended purpose» art. 3, pt. 11.

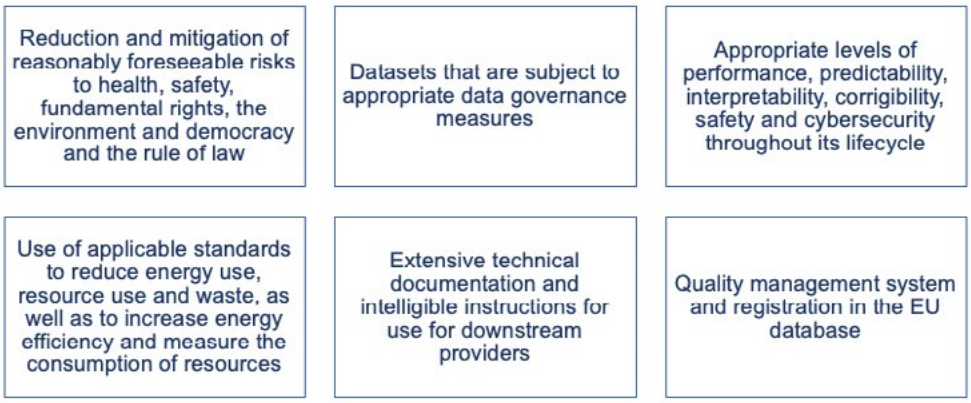


Figure 3 - Visual representation of the requirements for foundation models proposed by the Parliament (art. 28b). Own elaboration.

Most notably, art. 28b provides for a risk mitigation obligation, widening the scope of possible risks. Indeed, high-risk AI systems shall identify and mitigate risks for safety, health and fundamental rights, whereas foundation models should also look out for «environment and democracy and the rule of law»[71] related risks, which, as mentioned resembles the DSA-related systemic risks.

In comparison with the risk management system (art. 9) set out by the AI Act for high-risk AI systems, art. 28b does not provide for monitoring obligations, therefore the risk management system only regards risks assessed «prior and throughout the development». This is highly inconsistent with a technology that - as discussed - continues to demonstrate emergent abilities and whose risks and limitations are not fully researched. Moreover, the models themselves are evolving, requiring maintenance and monitoring after their updates.

Furthermore, another critical aspect is the provision of technical documentation and instructions for use only for downstream providers. Indeed, if foundation models are accessible and can also be directly utilized by laymen, it would be consistent with foundation models' known specific risks to provide users with mandatory clear and appropriate information on the model's capabilities and limitations. The lack of clear and actionable instructions coupled with the lack of mandatory human oversight (as for high risk systems) may also enhance the so-called automation bias or *effet moutonnier*[72] as users may refer uncritically to the model's output.

The European Parliament also provided for cooperation requirements as to enable downstream high-risk applications (art. 28, para. 2; more on this in the next section) as former providers are obliged to provide any documentation, technical access and additional support required for the fulfillment of the obligations of the new (downstream) provider.

With particular consideration to Generative AI, para. 4 establishes that their providers shall:

- «comply with transparency provisions in art. 52;
- ensure adequate safeguards against the generation of content in breach of Union

[71] Art.28b of the Parliament's position.

[72] A. Garapon - J. Lassègue, *Justice digitale. Révolution graphique et rupture anthropologique*, Paris, 2018, 239.

law in line with the generally acknowledged state of the art, and without prejudice to fundamental rights, including the freedom of expression;
- document and make publicly available a sufficiently detailed summary of the use of training data protected under copyright law»[73].

A study by Stanford University[74], based on publicly available information, compared ten major foundation models providers on twelve (out of 22) selected Parliament-proposed requirements evaluated on a scale from 0 to 4. The image below shows the results of the research, with a breakdown of the grades awarded for every foundation model provider on every requirement analyzed.

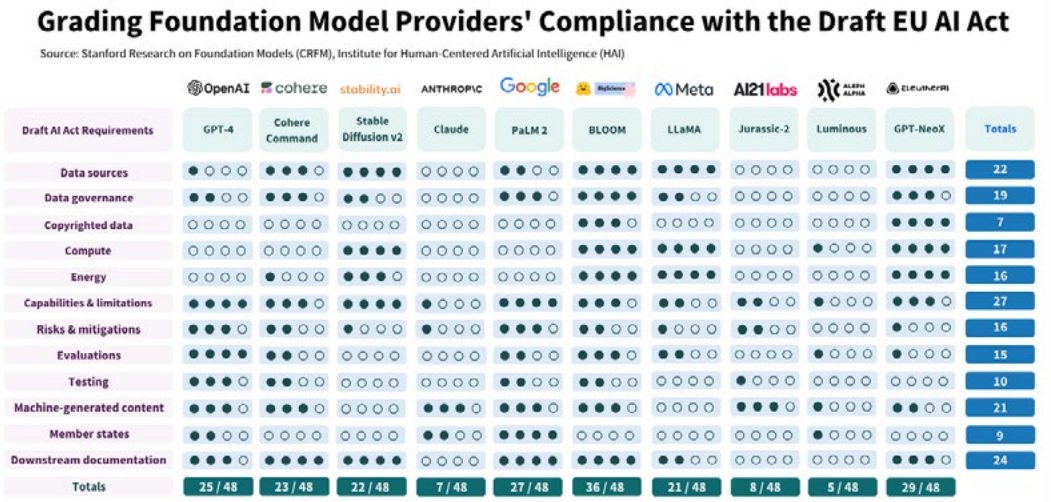


Figure 4 – Evaluation of different foundation models on AI Act requirements. Source: Bommasani et al., Do Foundation Model Providers Comply with the Draft EU AI Act?, cit.

Notably, the highest-ranking requirement (and the only one regarding which all models were awarded at least one point) is "capabilities and limitations", showing a certain industry-wide attention to the topic; the worst-ranking requirement is linked to the publicity on the utilization of copyrighted materials.

The results show that currently, foundation models comply unevenly with AI Act's requirements; however, the study argues that the provisions proposed by the Parliament may lead towards «substantial progress towards more transparency and accountability».

The study also confirmed that, even though openly released foundation models score well on disclosure requirements, they perform worse on deployment control.

[73] Art. 28b of the Parliament's position.

[74] R. Bommasani - K. Klyman - D. Zhang - P. Liang, *Do Foundation Model Providers Comply with the Draft EU AI Act?*, in *Stanford.edu*, 2023.

5.1 The value chain

The Parliament's position also impacted the obligations of the different actors in the AI market. This section illustrates the applicable legal framework and indicates possible improvements.

The table below shows the obligations for each actor in the general purpose AI systems and foundation models value chain:

- foundation models are subject to the obligations of art. 28b when their providers make them available on the market or put them into service;
- general purpose AI systems are not regulated *per se*. They are (like all other AI systems) bound by the «General principles applicable to all AI systems» (art. 4a), but only indirectly, as the principles apply to other provisions such as Title III requirements for high-risk applications, the Code of Conduct or the harmonized standards but they do not create «new obligations under this Regulation»;
- only deployers of high-risk AI systems have oversight and monitoring obligations (art. 29); and
- downstream applications (who make a substantial modification to the AI system) from both general purpose AI systems and foundation models[75] may become high-risk AI systems.

The term deployer[76] in the Parliament compromise has substitute the term user of the original proposal (which could, in fact, have been misleading). Affected persons were not included in the table as they only have rights[77] and not obligations under the regulation.

Upstream AI system provider		Deployer	Downstream provider	Deployer (of the downstream provider)
General purpose AI systems	/	/	*High-risk: Title III Chapter 2 requirements apply*	*Obligations under art. 29*
	/	/	*Not high risk: Art. 52 may be applicable*	*Art. 52 may be applicable*

[75] When «*directly integrated into an high-risk AI system*» art. 28, pt. 2, of the Parliament's position.

[76] For example «any natural or legal person, public authority, agency or other body using an AI system under its authority except where the AI system is used in the course of a personal non- professional activity» art. 3, pt. 4. In order to quell any ambiguity, the Parliament also introduced (art. 3, pt. 8a) the notion of "affected person", i.e. «any natural person or group of persons who are subject to or otherwise affected by an AI system».

[77] Regarding high-risk AI systems, affected persons have the right to an explanation (art. 68c); information to be subject to the use of the high-risk AI system when it makes (or assist in making) decisions pertaining to natural persons (art. 29, para 6a). Affected persons also have a right to be informed about interacting with an AI and on the nature of AI-generated media (art. 52), regardless of the high-risk status of the AI system.

Foundation models	*Art. 28b requirements (referred to above)*	*Art. 52 may be applicable*	*High-risk: Title III Chapter 2 requirements apply*	*Obligations under art. 29*
			Not high-risk: Art. 52 may be applicable	*Art. 52 may be applicable*

Table 2 - Obligations of actors in the value-chain of general purpose AI systems and foundation models. Own elaboration.

This regulatory framework is disputable because its risk-based reasons are not apparent. Indeed, it is not clear in the text of the amended Regulation (articles or recitals) what risk assessment lead to the setting up of specific obligations (and that specific obligations) only for foundation models and not general-purpose AI[78] and to obligations (or lack thereof) for the different actors of the value chain.
Furthermore, even if providers of general purpose AI systems are not directly subject to documentation obligations, they have to collaborate with (and provide documentation and technical access to) their downstream providers; therefore also general purpose AI systems providers will also have to indirectly comply with many of the documentation requirements of the AI Act.
Moreover, the distinction between deployer and downstream provider warrants more attention.

a) Downstream providers
A new downstream provider is considered one under art. 28 when:
1. it places its name or trademark on a high-risk AI system (para. 1, pt. a);
2. it makes a substantial modification to a high-risk AI system which remains a high-risk AI system (pt. b); and
3. it makes a substantial modification to a not high-risk AI system (including a general purpose AI system) which then becomes a high-risk AI system (pt. ba).

In these cases, the former provider is not considered anymore the provider responsible for that system and shall provide technical documentation, access and assistance to the new provider for compliance purposes. Art. 28, para. 2 states that this applies to foundation models «directly integrated into an high-risk AI system»: An example could be a company integrating a foundation model via API as a chatbot evaluating candidates in a recruitment process.
However, it should be more clearly stated in the text that the implementation of a foundation model into a high risk system generates a new provider and a new AI system. Indeed, as currently worded, the equivalence may only refer to the mandated

[78] In fact, because of poor wording of amendment 34 (art. 28, para. 1, pt. ba) cited above mentions «a substantial modification to an AI system, including a general purpose AI system, which has not been classified as high-risk» one may conclude that general purpose AI systems may be classified as high-risk. However, pt. b), which pertains to «substantial modification to a high-risk system» does not mention general purpose AI systems at all.

cooperation[79].
Furthermore, it is not yet clear up to which point the documentation and requirements set out for foundation models (and the cooperation requirements for not high-risk general purpose AI systems) are specific enough for downstream providers having to comply (and attest compliance) with high-risk requirements.

b) Deployers
The term deployer indicates[80] the professional natural or legal person that utilizes the AI system «under its authority» and is subject to a number of obligations under art. 29 in case of high-risk AI systems. For example, this means that someone utilizing a language model to create a poem for a birthday card is not a deployer, whereas a business utilizing a generative AI system to create some images for a commercial presentation falls under this category.
Arguably, general purpose AI systems cannot have deployers as they are not ready for use and, if they were trained and applied to a specific high-risk application, the obligations for the provider would shift to the entity providing the training.
On the contrary, foundation models (which are pre trained, in the Parliament's definition) may very well have deployers. Notably, the only requirement for deployers is the one originally set out (albeit improved upon by both the Parliament and the Council) by the Commission's proposal, as art. 29 clearly only refers to deployers of high-risk AI systems. In fact, art. 52 sets out some transparency obligations applicable to the output of generative AI systems: *e.g.* disclosure requirements for systems interacting with natural persons (e.g. chatbots) or creating/manipulating media (e.g. deep fakes[81]). In particular, in the latter case, users (a relic from the change from users to deployers) «shall disclose in an appropriate, timely, clear and visible manner that the content has been artificially generated or manipulated».
Indeed, a foundation model deployer which utilizes a generative model should disclose the artificial origin of the output but deployers of high-risk systems have a large number of other obligations. That's because the deployer is the closest actor to the actual application of the system and therefore is the most qualified to e.g. ensure the representativeness of the dataset (art. 29, para. 3) and carry out data protection impact assessments (para. 6).
This legislative vacuum is inconsistent with the risks tied to the propagation of biases and disinformation mentioned above. Indeed, if these risks are such as to warrant an *ad hoc* regulation, more attention could be directed towards helping (as already mentioned) both laymen and professional deployers to utilize these models with awareness,

[79] «This paragraph shall also apply to providers of foundation models as defined in Article 3 when the foundation model is directly integrated in a high-risk AI system» art. 28, para. 2, of the Parliament's position.

[80] «'Deployer' means any natural or legal person, public authority, agency or other body using an AI system under its authority except where the AI system is used in the course of a personal non-professional activity» art. 3, pt. 4 of the Parliament's position.

[81] Described in art. 52 of the AI Act as «text, audio or visual content that would falsely appear to be authentic or truthful and which features depictions of people appearing to say or do things they did not say or do, without their consent».

by providing them with adequate information on their functioning and limitations.
In particular, if a deployer utilizes foundation models in high risk applications (e.g. for considering the eligibility of a person for public benefits or for assessing and grading students' exams) it could be appropriate to impose some obligations of high-risk deployers such as human oversight and monitoring along with, in relation to systems referred to in Annex III, informing natural persons subject to a decision made (or assisted by) a high risk AI system[82].

6. Concluding remarks

In conclusion, the utilization of foundation models and general purpose AI systems brings forth, along with benefits, potential risks and harms. These risks are amplified by the aggregate effect and potential propagation of biased or hallucination-induced outputs, which may have societal implications and impact the resilience of democratic systems.
However, the Commission's proposal does not regulate general purpose AI systems nor foundation models. While the Parliament's position seems more promising, more focus on the specific risks related to the different actors in the value chain (including the provision of obligations for deployers, i.e. professional users) would be required. Moreover, when considering the requirements for general purpose AI systems and foundation models, it is crucial to determine the scope of such systems and the possible risks.
For instance, as regards foundation models, which can be directly accessed and utilized by citizens, it is appropriate to provide mandatory and clear information on the capabilities and limitations of these models and provide an iterative risk management system.
Overall, taking into consideration the potential risks and the need for clear regulations and guidelines, the final AI Act text will need to approach the deployment of these technologies with utmost care to protect the interests of individuals and society at large.

[82] Deployers of high-risk AI systems should ensure compliance with the instructions for use (also with the adoption of technical and organizational measures), human oversight, monitoring and maintenance of the appropriate robustness and cybersecurity measures. Most notably, the Parliament included also an obligation to inform natural persons that they are subject to a high-risk AI system utilized to make decisions (or assist in the decision-making process) (para. 6a).
Deployers shall also: ensure the use of relevant and representative data (if they have control over the input data), inform providers and relevant authorities in case of any serious incident or malfunctioning, carry out a data protection impact assessment and publish a summary thereof.

La questione *deepfake* in Italia: una panoramica*

Veronica Azzali – Nicol Ellecosta

Abstract

Sin dalla loro comparsa nel 2017, i *deepfake* hanno attirato notevole attenzione. Questi sofisticati strumenti software consentono la manipolazione di immagini, video e audio preesistenti. Ad esempio, è possibile incollare il viso di una persona sul corpo di un'altra, ricreando, quindi, situazioni paradossali, che appaiono reali anche agli occhi più esperti. Alcuni individui percepiscono i *deepfake* come una minaccia alla democrazia, una forza inarrestabile che la società è destinata a subire irrimediabilmente. Tuttavia, negli ultimi anni, diversi governi, piattaforme online e agenzie di informazione hanno intrapreso misure per contenere questo fenomeno, adottando strategie quali il blocco degli utenti che diffondono informazioni false, il fact checking e l'introduzione di normative legislative. L'Italia è tra le giurisdizioni che stanno cercando di contrastare l'uso improprio dei *deepfake*. Nel corso di questo saggio, esamineremo come le leggi italiane si concentrino principalmente sul furto d'identità e sulle implicazioni dei *deepfake* nel contesto della pornografia.

Deepfakes have attracted considerable attention since their emergence in 2017. These sophisticated software tools allow the manipulation of pre-existing images, videos, and audio. For example, paradoxical situations that appear real to even the most experienced eyes can be created by pasting one person's face onto another's body. Some see deepfakes as a threat to democracy, an unstoppable force from which society will suffer irreparably. In recent years, however, several governments, online platforms, and news agencies have taken steps to curb the phenomenon. They have adopted strategies such as blocking users who spread false information, fact-checking, and introducing legislation. Italy is among the countries trying to combat deepfake abuse. In this essay, we will examine how Italian legislation has focused primarily on identity theft and the impact of deepfakes in the context of pornography.

Sommario

1. La tana del Bianconiglio. – 2. Il mondo dei replicanti. – 3. Profondamente falsi. – 4. *Deepfake*-news. – 5. *Deepfake* e furto d'identità – 6. La regolamentazione in Italia – 7. Conclusioni.

* L'articolo è stato sottoposto, in conformità al regolamento della Rivista, a referaggio "a doppio cieco".

Keywords

deepfake – intelligenza artificiale – fake news – furto d'identità – pornografia

1. La tana del Bianconiglio

Gli appassionati di meme saranno probabilmente a conoscenza (se non addirittura state vittime) di un fenomeno molto particolare noto come "*rickrolling*". Acquisendo importanza a metà degli anni 2000[1], il *rickrolling* è un fenomeno della cultura meme caratterizzato da uno scherzo che consiste in un collegamento ipertestuale che reindirizza le persone al video musicale del singolo di Rick Astley "*Never Gonna Give You Up*"[2]. L'essenza di un *rickroll* risiede nella sua natura inaspettata; i destinatari del collegamento ipertestuale (condiviso in post social o attraverso semplici messaggi su WhatsApp) di solito si aspettano di accedere a contenuti rilevanti o coinvolgenti[3], per poi vedersi costretti ad ascoltare alcune note del famoso pezzo del 1987.

Gli appassionati di calcio italiano avranno verosimilmente incontrato una variante specifica di *rickroll* in cui il protagonista non è Rick Astley, bensì l'allenatore della squadra di calcio Juventus, Massimiliano Allegri. Fatta eccezione per la paradossale natura della situazione, il video potrebbe persino apparire autentico, data la fedele riproduzione della voce dell'allenatore toscano e l'eccellente mimica facciale che si sovrappone non solo a quella dell'originale, ma anche al distintivo *cast* di supporto.

Sorge quindi la domanda su come sia possibile non solo replicare in modo così accurato tratti fisionomici di Massimiliano Allegri, ma anche sostituire la voce di Rick Astley. Tale fenomeno costituisce un esempio paradigmatico dell'abilità di alterare un contenuto veritiero attraverso la fedele riproduzione di volti, espressioni facciali e voci umane. Tali adattamenti, denominati *deepfake*, vengono elaborati attraverso una forma particolare di intelligenza artificiale (IA). Nel corso degli ultimi anni, numerosi esempi di video comici *deepfake* sono stati condividi nella rete web, inclusa una delle attuali tendenze virali sui social media: la ricreazione di *roster* di personaggi fittizi tratti dal popolare gioco di combattimento "*Mortal Kombat*". Questa tecnologia consente la simulazione virtuale di scontri tra figure politiche come Joe Biden e Donald Trump al di fuori del contesto elettorale, o immaginarie contese fisiche tra personalità come Jeff Bezos ed Elon Musk con azioni di combattimento.

Tuttavia, per coloro che sono a conoscenza del fenomeno dei *deepfake*, il termine non può essere semplicemente associato alla cultura dei *meme*, ma assume una connotazione nettamente più inquietante[4,5,6]. Sfortunatamente, non sono rari i casi in cui questa

[1] K. Knowles, *What the heck is… Rickrolling?*, in *forbes.com*, 2 febbraio 2016.

[2] R. Astley, *Never Gonna Give You Up*, 1987.

[3] K. Knowles, *What the heck is… Rickrolling?*, cit.

[4] D. Fallis, *The Epistemic Threat of Deepfakes*, in *Philosophy & Technology*, 34, 2020, 623 ss.

[5] M. Westerlund, *The Emergence of Deepfake Technology: A Review*, in *Technology Innovation Management Review*, 9, 2019, 39 ss.

[6] M. Yankoski - W. Scheirer - T. Weninger, *Meme warfare: AI countermeasures to disinformation should focus on*

tecnologia viene sfruttata a fini deprecabili. Celebrità, ad esempio, sono state vittime di *deepfake* utilizzati per la creazione di falsi filmati pornografici, come nel noto caso di Scarlett Johansson. L'attrice hollywoodiana è stata oggetto di numerose manipolazioni *deepfake* in cui il suo volto è stato sovrapposto a quello di autentiche attrici pornografiche in video espliciti e osceni[7,8]. Uno di tali video ha raggiunto persino 1,5 milioni di visualizzazioni attraverso un rinomato sito per adulti[9].

In conclusione, si può affermare che i *deepfake*, analogamente ad altre forme di intelligenza artificiale, presentano aspetti positivi e negativi. Tuttavia, la questione rimane irrisolta: chi sono gli artefici dei *deepfake* e quali sono le loro motivazioni? Pertanto, risulta essenziale fornire una breve spiegazione su (a) cos'è l'intelligenza artificiale; (b) come e perché sono nati i *deepfake*; (c) le modalità di utilizzo di questa tecnologia; e (d) le possibili conseguenze legali della creazione e dell'utilizzo di *deepfake*.

2. Il mondo dei replicanti

In primis è necessario partire con una breve spiegazione, sicuramente non esaustiva, di cos'è l'intelligenza artificiale[10].

L'intelligenza artificiale studia i fondamenti teorici, le metodologie e le tecniche che consentono di progettare sistemi *hardware* e sistemi di programmi software atti a fornire all'elaboratore elettronico prestazioni che, a un osservatore comune, sembrerebbero essere di pertinenza esclusiva dell'intelligenza umana[11].

In sintesi, quando ci riferiamo all'Intelligenza Artificiale, facciamo generalmente riferimento alla disciplina che si occupa dello studio e dello sviluppo di sistemi tecnologici e informatici che cercano di emulare i comportamenti, il pensiero e le capacità umane. Il nucleo di questa tematica risiede proprio in questo concetto: l'IA non mira a replicare l'intelligenza umana, bensì cerca di riprodurla[12]. Occorre però individuare chi, o meglio cosa definisce ciò che è "intelligente".

Il dizionario statunitense Merriam Webster[13] definisce l'intelligenza come (a) la capacità di apprendere, comprendere e affrontare situazioni difficili; (b) la capacità di applicare le conoscenze per manipolare l'ambiente o pensare in modo astratto; e infine, (c) la capacità di utilizzare la ragione, aspetto particolarmente interessante e difficile da trasmettere a una macchina.

popular, not perfect, fakes, in *Bulletin of the Atomic Scientists*, 77, 2021, 119 ss.

7 W. Brown - D. H. Fleming, *Celebrity headjobs: or oozing squid sex with a framed-up leaky {Schar-JØ}*, in *Porn Studies*, 7, 2020, 357 ss.

8 A. Santangelo, *Il futuro del volto nell'era dei deep fake*, in M. Leone (a cura di) *Il metavolto*, FACETS Digital Press, 2022, 19 ss.

9 G. Meikle, *Deepfakes*, Cambridge, 2022.

10 J.H. Fetzer, *What is Artificial Intelligence?*, in J.H. Fetzer, *Artificial Intelligence: Its Scope and Limits. Studies in Cognitive Systems*, Dordrecht, 1990, 3 ss.

11 F. Amigoni - V. Schiaffonati - M. Somalvico, *Intelligenza artificiale*, in *treccani.it*, 2008.

12 Ivi, para. 2.

13 Merriam-Webster, *intelligence*, in *merriam-webster.com*.

Di conseguenza, ritenuto l'essere umano come essere senziente ed intelligente, è possibile che le macchine siano intelligenti? Sarebbe, infatti, difficile considerare intelligenti sentimenti come rabbia e gelosia. Sebbene occasionalmente possano risultare utili, sicuramente non vorremmo vederli espressi da una macchina. Pertanto, la ricerca si concentra sull'identificazione dei criteri che definiscono l'intelligenza[14,15].

Le applicazioni dell'IA sono molteplici. Ad esempio, nel mondo della musica, l'uso dell'IA ha permesso di creare una nuova canzone dei Beatles estrapolando la voce di John Lennon da una vecchia demo, come dichiarato da Paul McCartney[16]. Piattaforme di streaming musicale, come Apple Music e Spotify, utilizzano l'IA per generare playlist personalizzate in base alle preferenze musicali degli utenti[17]. L'IA è progredita al punto da essere in grado di comporre pezzi musicali autonomamente[18].

Anche nel settore dei videogiochi, l'IA consente ai personaggi non giocanti (NPC) di comportarsi in modo simile agli esseri umani, migliorando così l'esperienza di gioco[19,20]. Inoltre, il *pathfinding* è una tecnologia che permette ai personaggi di muoversi nel mondo di gioco senza collidere con ostacoli e raggiungere il loro obiettivo[21,22].

L'IA contribuisce anche al miglioramento delle capacità di scrittura attraverso pro-

---

[14] J.H. Fetzer, *What is Artificial Intelligence?*, cit.

[15] In merito a questa tematica, sono stati condotti diversi studi, tra cui il rinomato test di Turing proposto da Alan Turing nel suo articolo "*Computing machinery and intelligence*" pubblicato sulla rivista *Mind* nel 1950. Il test di Turing misura la capacità di una macchina di manifestare un comportamento intelligente attraverso il "gioco dell'imitazione". Questo coinvolge tre partecipanti: (A) una macchina, (B) una persona e (C) un valutatore umano. Se il valutatore C, interagendo con A e B, non riesce a distinguere con certezza quale sia la macchina e quale l'essere umano, il test è considerato superato.
Un altro esperimento degno di nota è quello della "*Chinese Room*" proposto da John Searle nel suo articolo "*Minds, brains, and programs*" pubblicato su *Behavioral and Brain Sciences* nel 1980. Nell'esperimento, si immagina una persona che non comprende il cinese, ma si trova all'interno di una stanza. Questa persona dispone di un libro con istruzioni su come rispondere a messaggi in cinese. Utenti esterni fanno scorrere bigliettini in cinese sotto la porta, i quali vengono decifrati dalla persona nella stanza. Pur producendo risposte corrette, la persona non comprende il cinese; ha semplicemente risposto seguendo le istruzioni del libro. L'esperimento della *Chinese Room* ha alimentato il dibattito sull'equivalenza tra il seguire meccanicamente le regole (come fa un computer) e la reale comprensione, sollevando l'interrogativo se la cognizione umana implichi qualcosa di più della mera manipolazione di simboli.

[16] M. Savage, *Sir Paul McCartney says artificial intelligence has enabled a 'final' Beatles song*, in *bbc.com*, 13 giugno 2023.

[17] G. Björklund - M. Bohlin - E. Olander - J. Jansson - C.E. Walter - M. Au-Yong-Oliveira, *An Exploratory Study on the Spotify Recommender System*, in A. Rocha - H. Adeli - G. Dzemyda - F. Moreira (a cura di) *Information Systems and Technologies*, Cham, 2022, 366 ss.

[18] O. Lopez-Rincon - O. Starostenko - G.A. S.Martín, *Algoritmic music composition based on artificial intelligence: A survey*, in *International Conference on Electronics, Communications and Computers (CONIELECOMP)*, 2018, 187 ss.

[19] Per una breve introduzione al tema si consiglia: J. Levine - C.B. Congdon - M. Ebner - G. Kendall - S.M. Lucas - R. Miikkulainen - T. Schaul - T. Thompson, *General Video Game Playing*, in *Dagstuhl Follow-Ups*, 6, 1998, 77 ss.

[20] S. Seidel - N. Berente - A. Lindberg - K. Lyytinen - B. Martinez - J. V. Nickerson, *Artifical Intelligence and Video Game Creation: A Framework for the New Logic of Autonomous Design*, in *Journal of Digital Social Research*, 2, 2020, 126 ss.

[21] X. Cui - H. Shi, *A*-based Pathfinding in Modern Computer Games*, in *International Journal of Computer Science and Network Security*, 11, 2011, 125 ss.

[22] Per ciò che riguarda l'AI nei videogiochi al di fuori dei personaggi non giocanti si consiglia G.N. Yannakakis, *Game AI revisited*, in *Proceedings of the 9th conference on Computing Frontiers*, 2012, 285 ss.

grammi di editing[23] come Grammarly o ProWritingAid, che controllano grammatica e ortografia. Altre forme di IA aiutano nella traduzione di testi, come DeepL o Google Translate.
Persino le automobili moderne sono dotate di IA, ad esempio attraverso sistemi ADAS (*Advanced Driver-Assistance Systems*) che migliorano l'esperienza di guida con funzionalità come la frenata automatica d'emergenza o l'avvertimento di superamento di corsia[24].
Concludiamo quindi esplorando un ulteriore aspetto dell'IA: il *deepfake*, che rappresenta una diversa forma avanzata di intelligenza artificiale.

3. Profondamente falsi

Il termine *deepfake* è composto da "*deep learning*"[25] e "*fake*". Solitamente, quando si parla di *deepfake*, ci si riferisce a: «foto, video e audio creati grazie a software di intelligenza artificiale (AI) che, partendo da contenuti reali (immagini e audio), riescono a modificare o ricreare, in modo estremamente realistico, le caratteristiche e i movimenti di un volto o di un corpo e a imitare fedelmente una determinata voce[26]».
Innanzitutto, è importante sottolineare che le materie di base su cui si fondano i *deepfake* sono costituite dai volti, dai corpi e dalle voci di persone effettivamente esistenti[27].
In pratica, i *deepfake* operano in modo simile ai filtri di alcuni noti social network, come Instagram, o alle applicazioni di riconoscimento facciale, ma con una complessità e un sistema notevolmente più articolati. Attraverso un elaborato sistema di reti neurali, che analizzano enormi quantità di dati tramite il *deep learning*, è possibile replicare in modo estremamente fedele non solo i volti, ma anche le espressioni facciali, la voce e le sue sfumature[28,29]. Di conseguenza, l'intelligenza artificiale utilizza tutti i dati di mappatura facciale raccolti per sovrapporre il viso di una persona a quello presente in

[23] T. N. Fitria, *"Grammarly" as AI-powered English Writing Assistant: Students' Alternative for English Writing*, in *Journal of English Language Literature and Teaching*, 5, 2021, 65 ss.

[24] A. Ziebinski - R. Cupek - D. Grzecha - L. Chruszczyk, *Review of advanced driver assistance systems (ADAS)*, in *AIP Conference Proceedings*, 2017.

[25] Per un approfondimento sul tema, si suggerisce la lettura di "*An Introduction to Deep Learning*" di L. Arnold - S. Rebecchi - S. Chevallier - H. Paugam-Moisy, incluso in "*Statistical Foundations on Data Science*" del 2011. In breve, il *deep learning* rappresenta una forma di intelligenza artificiale volta a consentire ai computer di prendere decisioni in modo autonomo, seguendo un approccio simile a quello del cervello umano.
Per chiarire questo concetto, prendiamo ad esempio la parola "pesce". Tipicamente, un computer richiederebbe vari *input* (ad esempio, informazioni su animali acquatici, squame, branchie, ecc.) per comprendere il concetto di "pesce". Al contrario, il *deep learning* permette ai computer di apprendere a riconoscere un pesce elaborando diverse immagini, consentendo loro di decifrare autonomamente le caratteristiche di un animale acquatico.

[26] Garante Privacy, *Deepfake – Vademecum*, 2020, 1.

[27] *Ibid.*

[28] CBS, *I never said that! High-tech deception of 'deepfake' videos*, in *cbsnews.com*, 2 luglio 2018.

[29] D. Pan - L. Sun - R. Wang - X. Zhang - R. O. Sinnott, *Deepfake Detection through Deep Learning*, in *International Conference on Big Data Computing, Applications and Technologies*, 2020, 134 ss.

un determinato video[30]. Ciò consente la creazione di situazioni mai verificatesi, in cui una persona sembra fare e dire cose che non avrebbe mai fatto o detto[31,32,33]. Ad esempio, come accennato nel capitolo primo (in merito al caso dell'attrice hollywoodiana Scarlett Johansson), è possibile sovrapporre il volto di una famosa attrice su quello di una pornoattrice, creando così un video pornografico coinvolgente attori e attrici falsi. Dal suo emergere nel 2017, quando un utente di Reddit (una famosa piattaforma di notizie e intrattenimento) pubblicò vari video pornografici falsi con protagoniste diverse celebrità[34], il termine *deepfake* è costantemente associato, non a caso, a una connotazione negativa. Tuttavia, come evidenziato nell'introduzione di questo documento, i *deepfake* possono essere impiegati in modi creativi. Un esempio notevole è il *deepfake* realizzato dallo YouTuber Shamook, che ha riprodotto in modo impressionante il volto di un giovane Mark Hamill, sovrapponendolo a una resa CGI poco convincente, regalando ai fan di Star Wars una credibile e ringiovanita interpretazione di Luke Skywalker[35]. Inoltre, uno studio interessante condotto da Gillian Murphy e colleghi[36] ha aperto nuove possibilità stimolanti: *«many participants mentioned that they do not like specific actors, and consequently do not watch their films. Interestingly, the idea of re-casting controversial actors may provide a way for people to feel they are "separating the artist from the art", a challenge that has received renewed attention since the onset of the #MeToo movement*[37]».

Un ulteriore esempio positivo rilevante in questo contesto è rappresentato da uno spot pubblicitario volto a sensibilizzare sull'importanza della lotta contro la malaria, con protagonista l'ex calciatore inglese David Beckham. Mediante l'utilizzo della tecnologia *deepfake*, sono state superate diverse barriere linguistiche. In particolare, non solo è stato possibile ricreare in modo straordinariamente accurato le espressioni facciali del calciatore inglese, ma addirittura i movimenti della bocca sono stati modificati con precisione per adattarsi perfettamente al labiale delle nove lingue interpretate[38].

Nonostante vi siano esempi come la produzione di meme, l'uso creativo nei film e la realizzazione di pubblicità a sfondo sociale possano evidenziare il potenziale positivo della tecnologia *deepfake*, non è possibile ignorare il vero motivo per cui questa tecnologia è particolarmente rinomata, ossia la produzione di notizie false e contenuti video

---

30 A. Chadha - V. Kumar - S. Kashyap - M. Gupta, *Deepfake: An Overview*, in P. K. Singh - S. T. Wierzchoń - S. Tanwar - M. Ganzha - J. J. P. C. Rodrigues (a cura di) *Proceedings of Second International Conference on Computing, Communications, and Cyber-Security*, Cham, 2021, 557 ss.

31 M. J. Blitz, *Deepfakes and Other Non-Testimonial Falsehoods: When is Belief Manipulation (Not) First Amendment Speech?*, in *Yale Journal of Law & Technology*, 23, 2020, 162 ss.

32 D. Harris, *Deepfakes: False Pornography Is Here and the Law Cannot Protect You*, in *Duke Law & Technology Review*, 17, 2019, 99 ss.

33 T. Taulli, *Deepfake: What You Need to Know*, in *forbes.com*, 15 giugno 2019.

34 S. Maddocks, *'A Deepfake Porn Plot Intended to Silence Me': exploring continuities between pornographic and 'political' deep fakes*, in *Porn Studies*, 7, 2020, 415 ss.

35 La scena in questione è presente nella serie televisiva *The Mandalorian* (*Capitolo 16: Il Salvataggio*, Disney, 2020). Il *deepfake* è reperibile sul sito personale di Shamook in *youtube.com*.

36 G. Murphy - D. Ching - J. Twomey - C. Linehan, *Face/Off: Changing the face of movies with deepfakes*, in *PLoS ONE*, 18, 2023.

37 *Ibid.*

38 J. Marovt, *How we made David Beckham speak 9 languages*, in *synthesia.io*, 4 settembre 2023.

che possono compromettere non solo la reputazione individuale, ma anche la stabilità delle nostre democrazie.

4. *Deepfake*-news

«Nel tempo dell'inganno universale dire la verità è un atto rivoluzionario» - George Orwell.

La parola dell'anno 2016 per l'Oxford Dictionary fu il neologismo "*post-truth*", ovvero "post-verità"[39]. Secondo la definizione del prestigioso dizionario, l'aggettivo indica «*circumstances in which objective facts are less influential in shaping public opinion than appeals to emotion and personal belief*[40]». Oggi, le *fake news* hanno ampiamente infiltrato la quotidianità di ciascun individuo, catturando l'attenzione di molte persone e influenzando la loro percezione della realtà, come dimostra l'uscita del Regno Unito dall'Unione Europea[41].

L'influenza delle notizie false, specialmente quelle derivanti dai social network, sulla Brexit è ben nota[42]. Alcuni politici sostenitori della Brexit promuovevano un maggiore controllo dell'immigrazione nel Regno Unito. Va ricordato all'uopo che Londra, non solo controllava gran parte dei flussi migratori che riguardavano il paese, ma che le contrattazioni post-Brexit non hanno fatto altro che complicare ulteriormente le cose[43]. Un ulteriore esempio emblematico è la notizia fuorviante dei 350 milioni di sterline settimanali che avrebbero dovuto essere destinati al Servizio Sanitario Nazionale dopo la Brexit[44].

I social network sono stati decisivi nella diffusione di *fake news* sulla Brexit[45]. Uno studio della Ofcom (l'ente per la regolamentazione delle emittenti radiotelevisive e delle telecomunicazioni del Regno Unito), condotto nel 2018, ha evidenziato che internet è la seconda fonte di informazione più utilizzata dagli inglesi, con i giovani e le minoranze etniche che preferiscono informarsi online. In particolare, i social network sono la fonte online prediletta per raccogliere informazioni[46]. Di conseguenza, si può

[39] Oxford University Press, *Word of the Year 2016*, in *global.oup.com*, 2016.

[40] *Ibid.*

[41] Y. Rodny-Gumede, *Fake It till You Make It: The Role, Impact and Consequences of Fake News*, in B. Mutsvairo - B. Karam (a cura di) *Perspectives on Political Communication in Africa*, Palgrave Macmillan, 2018, 203 ss.

[42] D. Dobreva - D. Grinnell - M. Innes, *Prophets and Loss: How "Soft Facts" on Social Media Influenced the Brexit Campaign and Social Reactions to the Murder of Jo Cox*, in *Policy & Internet*, 12, 2020, 144 ss.

[43] J.N. Druckman - E. Peterson - R. Slothuus, *How elite partisan polarization affects public opinion formation*, in *American Political Science Review*, 107(1), 2013, 57 ss.

[44] J. Stone, *British public still believe Vote Leave '350million a week to EU' myth from Brexit referendum*, in *independent.co.uk*, 28 ottobre 2018.

[45] A riguardo si consiglia lo studio di M. Höller, *The human component in social media and fake news: the performance of UK opinion leaders on Twitter during the Brexit campaign*, in *European Journal of English Studies*, 25(1), 2021, 80 ss., che ha analizzato i profili X (allora, Twitter) dei principali *opinion leader* britannici nel periodo pre-Brexit. Dallo studio è emerso l'effettivo contributo nella divulgazione di notizie false da parte dell'ex primo ministro Boris Johnson e di uno dei maggiori fautori della Brexit Nigel Farage.

[46] Jigsaw Research, *News Consumption in the UK: 2018*, in *ofcom.org.uk*, 2018, 1.

affermare con verosimile certezza che le *fake news* abbiano influenzato la percezione della realtà della popolazione inglese, e di conseguenza il loro voto[47,48].

Il problema delle *fake news* online è ulteriormente aggravato dalla crescente presenza di video *deepfake*, creati principalmente da *hacktivist*, vale a dire persone o gruppi che, per motivi sociali, politici o ideologici[49], promuovono campagne di disinformazione per confondere l'opinione pubblica e minare la fiducia in un determinato paese[50,51,52]. Quando i contenuti *deepfake* coinvolgono politici o opinion leader, possono causare variazioni nelle opinioni dell'elettorato, compromettendo i diritti di autodeterminazione informativa e libertà decisionale delle persone. Ne discende la privazione di due indispensabili diritti: (a) l'autodeterminazione informativa («ciò che voglio far sapere di me lo decido io[53]») e (b) la libertà decisionale («quello che penso e faccio è una scelta su cui gli altri non possono interferire[54]»). Ne è un esempio lampante il conflitto in Ucraina. Pochi mesi dopo l'inizio dell'invasione da parte delle armate di Mosca, sono apparsi due video falsi che inscenavano una presunta dichiarazione di pace da parte dei presidenti Volodymr Zelensky e Vladimir Putin[55]. Se il primo è stato reso in modo poco convincente[56], il secondo è decisamente più persuasivo (nonostante le paradossali dichiarazioni[57]).

In questa era della post-verità, emerge, quindi, una crescente sfida nel distinguere ciò che è veritiero da ciò che è falso. Il problema è ulteriormente complicato dal fatto che spesso sono proprio i politici a diffondere notizie false allo scopo di screditare i loro avversari. A solo titolo esemplificativo si rammenta il candidato alle primarie del Partito Repubblicano, Ron DeSantis, il quale ha condiviso un'immagine *deepfake* di Donald Trump apparentemente impegnato in un caloroso abbraccio e bacio con il Dottor

---

47 A. Hern, *Facebook criticised for response to questions on Russia and Brexit*, in *theguardian.com*, 13 dicembre 2017.

48 F. Safieddine, *Political and Social Impact of Digital Fake News in an Era of Social Media*, in Y. Ibrahim - F. Safieddine (a cura di) *Fake News in an Era of Social Media: Tracking Viral Contagion*, Lanham 2020, 43 ss.

49 R. Kraus - B. Barber - M. Borkin - N.J. Alpern, *Internet Information Services*, in R. Kraus - B. Barber - M. Borkin-N.J. Alpern (a cura di), *Seven Deadliest Microsoft Attacks*, Syngress, 109 ss.

50 CBS, *I never said that! High-tech deception of 'deepfake' videos*, in *cbsnews.com,* 2 luglio 2018.

51 J. Twomey - D. Ching - M. P. Aylett - M. Quayle - C. Linehan - G. Murphy, *Do deepfake videos undermine our epistemic trust? A thematic analysis of tweets that discuss deepfakes in the Russian invasion of Ukraine*, in *Plos ONE*, 18, 2023.

52 C. Vaccari - A. Chadwick, *Deepfakes and Disinformation: Exploring the Impact of Synthetic Political Video on Deception, Uncertainty, and Trust in News*, in *Social Media + Society*, 6, 2020.

53 Garante Privacy, *Deepfake – Vademecum*, cit., 4.

54 *Ibid.*

55 J. Twomey – D. Ching – M. Peter Aylett – M. Quayle – C. Linehan – G. Murphyal., *Do deepfake videos undermine our epistemic trust?*, 2023.

56 J. Wakefield, *Deepfake presidents used in Russia-Ukraine war*, in *bbc.com,* 18 marzo 2022. Il video è risultato scarsamente efficace data la sua pessima realizzazione: il collo del presidente ucraino risulta essere troppo sottile rispetto alla testa e la voce è decisamente più profonda di quella nella realtà.

57 A. Smith, *Fake video of Vladimir Putin declaring peace with Ukraine seeks to cause chaos online*, in *independent.co.uk*, 18 marzo 2022. Nel filmato il presidente russo non si limiterebbe a dichiarare la pace con lo stato ucraino, ma anche di restituire ufficialmente la Crimea all'Ucraina.

Anthony Fauci[58]. Questo episodio ha suscitato ulteriore tensione, considerando l'ostilità che il Partito Repubblicano nutre nei confronti del presidente della NIAID, specialmente in relazione alle misure proposte durante la pandemia di COVID-19. Di conseguenza, molti elettori repubblicani si sono sentiti traditi dall'ex presidente.

A causa di simili eventi, nel 2019, gli stati della California e del Texas hanno reagito vietando la diffusione di video *deepfake* nei 60 giorni che precedono un'elezione. La presente iniziativa evidenzia che alcuni stati, seppur tra loro politicamente diversi, sono ugualmente preoccupati per la crescente diffusione dei *deepfake* e riflette la consapevolezza dei problemi che questa tecnologia può causare[59,60].

I video *deepfake*, come per gli esempi menzionati, sono destinati a suscitare scalpore e confusione tra la popolazione, portando a inevitabili interferenze nel sistema di voto e, più in generale, nel sistema democratico[61]. Le piattaforme social hanno cominciato ad adottare misure per contrastare la diffusione di notizie false. Durante la pandemia di COVID-19 negli Stati Uniti, la piattaforma YouTube ha annunciato il blocco di diversi canali che diffondevano disinformazione sulla campagna vaccinale[62]. Gli algoritmi del social network Instagram, d'altra parte, non mostrano nei contenuti suggeriti quelli contrassegnati come falsi dai fact checker[63,64]. Anche Facebook, altra piattaforma del gruppo Meta, considerata una degli epicentri delle fake news online[65], sta cercando di migliorare la sua reputazione bloccando diversi utenti diffusori di notizie false sui social di Meta[66,67].

---

[58] N. Nehamas, *DeSantis Campaign Uses Apparently Fake Images to Attack Trump on Twitter*, in *nytimes.com*, 8 giugno 2023.

[59] California Government Code, Section 11547.5.

[60] State of Texas, Chapter 21, Penal Code, Section 21.165.

[61] M. Westerlund, *The Emergence of Deepfake Technology: A Review*, in *Technology Innovation Management Review,* 9(1), 2019, 39 ss.

[62] T. Ginossar, *Cross-platform spread: vaccine-related content, sources, and conspiracy theories in YouTube videos shared in early Twitter COVID-19 conversations*, in *Human Vaccines & Immunotherapeutics*, 18, 2022.

[63] R. Metz - D. O'Sullivan, *A deepfake video of Mark Zuckerberg presents a new challenge for Facebook*, in *edition.cnn.com*, 11 giugno 2019.

[64] N. Veerasamy - H. Pieterse, *Rising Above Misinformation and Deepfakes*, in R. P. Griffin - U. Tatar - B. Yankson (a cura di), *ICCWS 2022 17th International Conference on Cyber Warfare and Security*, 2022, ACI, 340 ss.

[65] Stando a recenti studi Facebook è il social network che diffonde il maggior numero di notizie false. In particolare, si consigliano due studi: A.M. Guess - B. Nyhan - J. Reifler, *Exposure to untrustworthy websites in the 2016 US Election*, in *Nature Human Behaviour*, 4, 2020, 472 ss. Un'interessante ricerca sulla diffusione di notizie online durante le elezioni presidenziali americane del 2016. T. Hopp-P. Ferrucci - C.J. Vargo, *Why Do People Share Ideologically Extreme, False, and Misleading Content on Social Media? A Self-Report and Trace Data–Based Analysis of Countermedia Content Dissemination on Facebook and Twitter*, in *Human Communication Research*, 46, 2020, 357 ss. Uno studio relativamente recente sui livelli individuali di radicalismo ideologico e la fiducia nel sistema mediatico in contrapposizione a quello delle fake news.

[66] J. Pamment, *How the Kremlin circumvented EU sanctions on Russian state media in the first weeks on the illegal invasion of Ukraine*, in *Place Branding and Public Diplomacy*, 19, 2023, 200 ss.

[67] Nondimeno ci sono piattaforme social che non attuano nessuna contromisura – se non addirittura endorsano – gli *account* di controinformazione. X (precedentemente noto come Twitter) di Elon Musk, ad esempio, è da parecchio tempo sotto scrutinio da parte dell'Unione Europea a causa dell'elevato numero di account (tra i quali molti sono verificati) che continuano a riempire il sito di notizie false. Molte delle fake news diffuse sul social di Musk riguardano temi scottanti e di importanza mondiale come il

Tuttavia, l'impatto dei *deepfake* non si limita alle strutture democratiche, elettorali e sociali, estendendosi altresì al campo dei media, interessando quindi i giornalisti stessi. Questi, devono affrontare una crescente diffidenza del pubblico nei confronti delle informazioni pubblicate sulla stampa[68], mentre contemporaneamente si trovano ad avere a che fare con un proliferare di fonti false. Giova menzionare, a titolo esemplificativo ed esplicativo, che durante le tensioni tra Pakistan e India, l'agenzia di stampa Reuters ha individuato oltre 30 video manipolati pubblicati da organi di informazione[69]. Con la Con la crescente ubiquità dei video *deepfake*, è inevitabile che questa problematica si aggraverà. Per fronteggiare questa sfida, agenzie di stampa di rilevanza globale, tra cui Reuters[70] e il Wall Street Journal, hanno implementato corsi formativi dedicati ai propri giornalisti per migliorare la loro capacità di riconoscere e gestire eventuali fake news. Inoltre, le stesse agenzie stanno investendo in diverse tecnologie volte al monitoraggio e al riconoscimento dei contenuti *deepfake*[71].
Va notato che nell'ordinamento italiano la creazione di *deepfake* è, in alcuni casi, considerata un reato, e nel paragrafo successivo si affronteranno le implicazioni legate al furto d'identità e all'uso di dati personali per la creazione di video falsi, in particolare nel contesto del mondo della pornografia e della pedopornografia[72].

***Deepfake* e furto d'identità**

Il connubio tra il fenomeno dei *deepfake* e il furto d'identità si presenta come un lega-

conflitto in Ucraina. M. Haigh - T. Haigh, *Fighting and Framing Fake News*, in P. Baines - N. O'Shaughnessy - N. Snow (a cura di) *The SAGE Handbook of Propaganda*, 2020, 303 ss.; P. Suciu, *X Is The Biggest Source Of Fake News And Disinformation, EU Warns,* in *forbes.com,* 26 settembre 2023 e quello tra Israele e Palestina M. Lakhani, *Fighting Disinformation in the Palestine Conflict: The Role of Generative AI and Islamic Values*, in *Al-Misbah Research Journal*, 3(6), 2023; D. Milmo, *X criticised for enabling spread of Israel-Hamas disinformation*, in *theguardian.com*, 9 ottobre 2023.

68 M. Westerlund, *The Emergence of Deepfake Technology,* cit., 40.

69 N. Jaffer, *Fake News and Disinformation in Modern Statecraft*, in *Regional Studies*, 39(1), 2021, 3 ss.

70 R. Lauren, *Will you believe it when you see it? How and why the press should prepare for deepfakes*, in *Georgetown Law Technology Review*, 4(1), 2019, 241 ss.

71 Á. Vizioso - M. Vaz-Álvarez - X. López-García, *Fighting Deepfakes: Media and Internet Giants' Converging and Diverging Strategies Against Hi-Tech Misinformation*, in *Media and Communication*, 9(1), 2021, 291 ss.

72 Va ricordato che i reati legati al mondo del deepfake non comprendono solo il furto d'identità, la pornografia e la pedopornografia, bensì possono comprendere anche crimini riguardanti l'estorsione e il bullismo. A riguardo è necessario citare almeno tre tecnologie con le quali *deepfaker* e altri criminali di internet possono ingannare gli utenti: (a) il *phishing* è una nota tecnica estorsiva con la quale dei malintenzionati convincono una o più vittime a condividere dati personali e/o finanziari (Z. Ramzan, *Phishing Attacks and Countermeasures*, in P. Stavroulakis - M. Stamp (a cura di), *Handbook of Information and Communication Security*, Berlino, 2010); (b) lo *spoofing*, invece, è il furto di informazioni attraverso la falsificazione di dati personali e dispositivi (S.A. Schuckers, *Spoofing and Anti-Spoofing Measures*, in *Information Security Technical Report*, 7(4), 2002, 56 ss.) (c) infine, il *ransomware* è un malware che impedisce o limita l'accesso a diversi tipi di dispositivi, attraverso il completo blocco dei sistemi o alla cifratura di file. Solitamente per sbloccare un qualsiasi dispositivo i malintenzionati chiedono un riscatto in denaro o in beni, servizi e/o favori (X. Luo - Q. Liao, *Awareness Education as the Key to Ransomware Prevention*, in *Information Systems Security*, 16(4), 2007, 195 ss.).

me inequivocabile[73]. Risulta sufficiente riflettere sul fatto basilare che la creazione di qualsiasi video *deepfake* implica l'artificiale "deprivazione" di un individuo del proprio volto, seguita dalla sovrapposizione dello stesso a quello di un'altra persona. Tale procedura costituisce già di per sé una chiara minaccia alla privacy. Se a ciò si aggiunge il fatto che i *deepfake* possono generare contesti e situazioni mai verificatisi nella realtà, l'entità del problema diviene ulteriormente manifesta.

Il nucleo che ogni individuo porta in sé e che, seppur lo riconduce al genere umano, alla tipologia di soggetto o di persona, lo caratterizza più specificatamente, è la sua identità. Ancora più penetrante, quindi, è la necessaria tutela dell'identità personale, definibile come dignità assoluta. In questo contesto, la tutela dell'identità personale emerge come una necessità impellente.

Sotto questo profilo, le Carte internazionali, pur adottando formulazioni diverse, riconoscono a ogni individuo la personalità giuridica e vietano interferenze nella sua vita privata, familiare e sociale, sancendo il diritto allo sviluppo della propria personalità[74].

Parte della dottrina ha sostenuto che l'identità rappresenta un bene-giuridico intrinsecamente legato alla dimensione sociale dell'individuo a cui essa si riferisce. Questo perché si ritiene che «l'identità sia il risultato delle interazioni sociali, che ha le sue radici nello spazio privato ma si manifesta e si definisce successivamente nello spazio pubblico»[75]. Proprio in virtù di questo principio, è celebre la definizione di identità personale contenuta nella pronuncia della Corte Costituzionale del 1994, ove si fa riferimento al «diritto ad essere sé stesso, inteso come rispetto dell'immagine di partecipe alla vita associata, con le acquisizioni di idee ed esperienze, con le convinzioni ideologiche, religiose, morali e sociali che differenziano, ed al tempo stesso qualificano, l'individuo».[76] L'identità è quindi un bene tutelato che, non solo individua il soggetto a cui si riferisce, ma lo rappresenta e ne afferma la sua personalità come proiezione sociale.

Ecco, quindi, che viene disciplinato il reato di furto d'identità[77], regolato dall'art. 494 c.p.: «Chiunque, al fine di procurare a sé o ad altri un vantaggio o di arrecare ad altri un danno, induce taluno in errore, sostituendo 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, è punito, se il fatto non costituisce un altro delitto contro la fede pubblica, con la reclusione fino a un anno[78]».

Il furto d'identità assume una gravità particolare quando si collega al contesto sessuale delle persone coinvolte nell' illecito. Sin dall'introduzione del fenomeno nel dibattito pubblico, i *deepfake*, essendo l'origine del termine strettamente legata alla pubblicazione

[73] Garante Privacy, *Deepfake – Vademecum*, cit., 1.

[74] In tal senso artt. 6, 12 e 22 della Dichiarazione universale dei diritti dell'Uomo; art. 8 della Convenzione europea dei diritti dell'uomo; artt. 6, 17 e 22 del Patto internazionale dei diritti civili e politici; art. 6 del Trattato dell'Unione europea che riconosce i diritti della CEDU e quelli della Carta dei diritti fondamentali dell'Unione europea.

[75] V. Zeno-Zencovich, *Identità personale, Digesto, privato, sezione civile*, IX, Torino, 1993, 301.

[76] Corte cost. 3 febbraio 1994, n. 13.

[77] Il furto d'identità non è il solo reato perpetrabile creando, usando o condividendo un *deepfake*. Ad esempio, se il contenuto *deepfake* va a ledere anche la reputazione dell'individuo, al reato di furto d'identità si aggiunge quello di diffamazione (art. 595 c.p.).

[78] Art. 496 c.p.

di video falsi a sfondo pornografico, hanno suscitato preoccupazioni riguardo a come prevenire e affrontare il problema del furto d'identità. La produzione di immagini fittizie con connotazioni sessuali, spesso ritraenti celebrità, in pose o situazioni intime, è comunemente denominata "*deepnude*". I *deepnude* coinvolgono l'incorporazione di un volto, attraverso l'uso di software, su corpi nudi o in ambienti a carattere pornografico. La situazione è ulteriormente complicata dalla disponibilità sempre crescente di applicazioni[79] che consentono a chiunque di creare facilmente video o immagini.

La semplicità con cui tali applicazioni possono essere installate sui dispositivi mobili e la loro facilità d'uso sono stati posti in rilievo nel corso di recente caso di cronaca che coinvolge un gruppo di studenti di una scuola media di Latina. Utilizzando un'applicazione denominata "*BikiniOff*"[80], i minori hanno posto in essere la manipolazione di fotografie di cinque studentesse e di una docente. Quest'applicazione è particolarmente apprezzata nel mondo dei *deepfaker* poiché va oltre la mera sostituzione del viso, ricreando in modo sorprendentemente realistico la posa desiderata, mantenendo le proporzioni e il colore della pelle della vittima[81]. Il *deepnude* della docente, nel caso specifico, è risultato così convincente da comparire su due rinomati siti pornografici[82].

Inevitabilmente, le discussioni su *deepnude* e *deepfake porn* sono strettamente legate al fenomeno del *revenge porn*, definito come l'«abuso digitale che comporta la divulgazione a terzi o la pubblicazione senza consenso di immagini o video a sfondo sessuale[83]».

In relazione a ciò si rammenta il caso della poetessa inglese Helen Mort, vittima di molteplici casi di *deepfake porn*: «*an acquaintance told her. But never in her life had she taken or shared intimate photos. Surely there must be some mistake. When she finally mustered up the courage to look, she felt frightened and humiliated*[84]».

Il creatore dei *deepfake* aveva pubblicato diverse fotografie della poetessa, incoraggiando altri individui a integrare il volto di Helen in contesti di natura esplicita[85]. La lotta alla diffusione di contenuti pornografici generati mediante l'impiego dell'intelligenza artificiale sembra presentare notevoli sfide. Filoni di ricerca recenti, infatti, evidenziano che oltre il 90% dei video *deepfake* si configurano come materiale a contenuto pornografico[86,87,88]. È da sottolineare, tuttavia, che in maniera analoga a quanto si riscontra sulle varie piattaforme social per quanto concerne la diffusione di *fake news*, il sito web

[79] Basta digitare la parola "*deepnude*" in un qualsiasi motore di ricerca per venire bombardati dalle più diverse app gratuite e a pagamento.

[80] Una delle applicazioni a pagamento più famose per la creazione di *deepnude*.

[81] D. Barbera, *Tutti i rischi di usare Bikinioff, il chatbot che spoglia le donne*, in *wired.it,* 19 aprile 2023.

[82] S. Matteis, *Cinque 13enni e una prof di Latina nude sul web: indagati i compagni, le foto false create con l'app BikiniOff*, in *fanpage.it*, 14 settembre 2023.

[83] L. Sartarelli, *È reato usare un deepfake?*, in *smartius.it*, 11 giugno 2019.

[84] K. Hao, *Deepfake porn is ruining women's lives. Now the law may finally ban it*, in technologyreview.com, 12 febbraio 2021.

[85] J. Laffier - A. Rehman, *Deepfakes and Harm to Women, in Digital Life and Learning*, 3(1), 2023.

[86] N. Kshetri, *The Economics of Deepfakes*, in *Computing's Economics*, 2023, 89 ss.

[87] G. MacGregor, *Gun to your head: how deepfakes and other non-consensual synthetic media hold individual autonomy hostage*, in *UMKC Law Review*, 90(2), 2021, 431 ss.

[88] Sensity Team, *How to Detect a Deepfake Online: Image Forensics and Analysis of Deepfake Videos*, in *sensity.ai,* 2021.

a carattere pornografico più frequentato a livello globale, ha vietato la pubblicazione di video *deepfake*[89,90].

Nonostante ciò, si ritiene estremamente improbabile che non esistano tracce di contenuti compromettenti generati tramite applicazioni di intelligenza artificiale su questa celebre piattaforma riservata ai maggiorenni. Ciò è in parte attribuibile alla legislazione statunitense, nello specifico la "Section 230[91]", approvata dal Congresso nel 1996, che garantisce una esclusione di responsabilità dei siti web rispetto ai contenuti pubblicati dai singoli utenti che siano in violazione con le regole imposte dalla normativa o dallo stesso sito web, come sottolineato da Dean Russell: «*[Section 230] prevents websites from being sued for hosting non-consensual deepfake porn, which means Google, Reddit, Twitter, etc. — you can ask them to take down a deepfake, but they don't have to. Pornhub says it doesn't allow deepfakes. But they are there. And then there are websites predicated on deepfake porn. That's their whole business model. Again, they are protected by Section 230*»[92].

Nonostante l'incremento della minaccia derivante dai *deepfake* a sfondo pornografico, nella totalità degli Stati Uniti sono poche le giurisdizioni che hanno adottato specifiche contromisure per affrontare questo fenomeno. Ad esempio, nello stato della California, le vittime possono ricorrere alla corte civile per ottenere un risarcimento finanziario[93]. In Virginia, invece, una legge introdotta nel 2019 impone pene detentive a coloro che diffondono *deepfake* pornografici. Tuttavia, la pubblicazione o la condivisione di tali contenuti è considerata reato solo se finalizzata a molestare o intimidire la persona rappresentata nel video[94].

6. La regolamentazione in Italia

In seguito agli eventi verificatisi a Latina sopra citati, la Procura dei Minori di Roma ha avviato un'inchiesta sulle finte foto pornografiche realizzate dagli alunni della scuola media. Sempre sulla scorta di quanto successo al gruppo di ragazzini, il 13 ottobre 2020, il Garante per la protezione dei dati personali ha avviato un'istruttoria nei confronti di Telegram[95]. Tale iniziativa ha scaturito un ulteriore intervento specifico in materia, durante il quale il Garante ha formulato un *vademecum* nominato: "*Deepfake*. Il falso che ti "ruba" la faccia (e la privacy)"[96]. L'istituzione ha mantenuto un costante interesse sull'argomento, emettendo provvedimenti, documenti ufficiali e comunicati.

[89] M. Popova, *Reading out of context: pornographic deepfakes, celebrity and intimacy*, in *Porn Studies*, 7, 2020, 367 ss.

[90] R. Winter, *DeepFakes: uncovering hardcore open source on GitHub*, in *Porn Studies*, 7, 2020, 382 ss.

[91] 47 U.S. Code, Section 230 on Protection for private blocking and screening of offensive material.

[92] D. Russell, *Is deepfake pornography illegal? It depends*, in *wbur.org*, 23 giugno, 2023.

[93] K. Farish, *Do deepfakes pose a golden opportunity? Considering whether English law should adopt California's publicity right in the age of the deepfake*, in *Journal of Intellectual Property Law & Practice*, 15, 2020, 40 ss.

[94] E. Gerstner, *Face/off: "deepfake" face swaps and privacy laws*, in *Defense Counsel Journal*, 87, 2020, 1 ss. Per quanto riguarda la legge in questione si tratta della: VA. CODE. § 18.2-386.2.

[95] Garante Privacy, *Deep fake: il Garante privacy apre un'istruttoria nei confronti di Telegram per il software che "spoglia" le donne*, 23 ottobre 2020, doc. web n. 9470722.

[96] *Ibid.*

Infatti, l'Autorità Garante ha continuato ad affrontare il tema della lesione del diritto all'identità personale in ottica divulgativa proprio per il costante emergere di molteplici nuove tecnologie, pericolosamente rischiose per il patrimonio personale. [97]

Date le ampie implicazioni dell'intelligenza artificiale sulla società, si osserva un crescente riconoscimento della necessità di una regolamentazione adeguata in questo settore, con numerosi paesi che collaborano attivamente per sviluppare normative adeguate. La regolamentazione si sforza costantemente di tenere il passo con lo sviluppo e l'innovazione tecnologica, ma la sua implementazione varia significativamente tra le diverse giurisdizioni. Questo equilibrio complesso implica che una disciplina eccessivamente restrittiva potrebbe ostacolare l'innovazione, mentre una regolamentazione troppo lenta potrebbe lasciare un vuoto normativo pericoloso con rischi non gestiti.

L'IA è in continua evoluzione e il suo sviluppo non aspetterà l'entrata in vigore delle nuove regole. Questo porta a due importanti conseguenze. In primis, quando le nuove regole saranno stabilite, potrebbero non essere in grado di regolamentare completamente o adeguatamente le applicazioni di intelligenza artificiale che non erano ancora emerse o che non sembravano richiedere inizialmente una disciplina ad hoc.

In secondo luogo, vi è l'urgenza di essere rapidi e di trovare un modo per regolare l'intelligenza artificiale il prima possibile. Urgenza dovuta al fatto che i progressi nell'IA stanno avvenendo a un ritmo accelerato e i governi non possono rischiare di assistere nuovamente a una sorta di autoregolamentazione, così come avvenuto per Internet, che, in assenza di una normativa, ha condotto tutti a una dimensione digitale che può definirsi auto-gestita, in una sorta di tecnocrazia che rischia di travolgere le democrazie attuali.

Il Consiglio dell'Unione Europea ha affrontato la questione adottando un progetto di promozione dell'innovazione responsabile nell'intelligenza artificiale per la pace e la sicurezza. In tale contesto, il Consiglio ha riconosciuto l'IA come una tecnologia abilitante con ampio potenziale d'uso generale, ma ha altresì sottolineato il rischio di abusi derivanti dall'accesso indiscriminato alle ricerche e all'innovazione in materia di IA per applicazioni civili. Nella decisione di Promozione dell'innovazione responsabile nell'IA per la pace e la sicurezza, in vigore dal 18 novembre 2022, si evidenzia il rischio rappresentato dall'uso improprio della tecnologia di IA, specialmente attraverso reti generative avversarie (GAN), per la creazione di *deepfake* utilizzati in campagne di disinformazione, un rischio che richiede particolare attenzione e che potrebbe non essere sufficientemente affrontato dagli attuali sforzi diplomatici e di controllo degli armamenti.

Senza volersi addentrare in questa sede nell'analisi specifica del documento di progetto, basti pensare che il Consiglio si è posto come obiettivo quello di promuovere: «un'innovazione responsabile, in quanto meccanismo di autogoverno, potrebbe fornire alla comunità civile mondiale dell'AI strumenti e metodi pratici per individuare e contribuire a prevenire e attenuare i rischi che la diversione e l'uso improprio della ricerca e dell'innovazione civili in materia di IA potrebbero comportare per la pace e

[97] Garante Privacy, *Deepfake: dal Garante una scheda informativa sui rischi dell'uso malevolo di questa nuova tecnologia*, 28 dicembre 2020.

la sicurezza[98]».

Invece, per quanto concerne l'ordinamento italiano, l'atto di pubblicare immagini o video sessualmente espliciti senza il consenso delle persone coinvolte è divenuta una condotta delittuosa con l'introduzione nel Codice penale dell'art. 10 della legge 69/2019 (c.d. Codice Rosso), che è stato pubblicato in Gazzetta Ufficiale il 25 luglio 2019. Questa legge è stata introdotta per punire penalmente il fenomeno noto come "*revenge porn*" attraverso una normativa *ad hoc* per affrontare il problema. Oggetto del reato sono le immagini o i video a contenuto sessualmente esplicito, destinati a rimanere privati.

Secondo quanto stabilito dalla Suprema Corte, il delitto è istantaneo e si consuma nel momento in cui avviene il primo invio dei contenuti, senza che assuma rilevanza il fatto che il destinatario sia un familiare della vittima, che non abbia interesse ad alimentare una successiva diffusione delle immagini. Viene dunque punita la «diffusione illecita di contenuti sessualmente espliciti [che possono avere come] oggetto immagini o video che ritraggano atti sessuali ovvero organi genitali ovvero anche altre parti erogene del corpo umano, come i seni o i glutei, nudi o in condizioni e contesto tali da evocare sessualità[99]».

Vi è inoltre una proposta di legge, presentata il 30 marzo 2021[100], volta a contrastare «il fenomeno della diffusione del *software* chiamato "*Deep nude*"[101]» e il *revenge porn*. L'obiettivo è quello di introdurre nel codice una «fattispecie delittuosa consistente nella diffusione di immagini di persone reali manipolate artificialmente allo scopo di ottenerne rappresentazioni nude[102]». L'art. 612-*quater* (primo comma) c.p. prevedrebbe un reato perpetrabile da chiunque abbia avuto a che fare, non solo con la creazione del *deepfake*, ma anche con la diffusione, l'invio o la pubblicazione di «immagini di persone reali, comunque identificabili, manipolate artificialmente mediante l'uso di strumenti tecnologici o di sistemi di intelligenza artificiale[103]». Inoltre, il secondo comma dell'art. 612-*quater*, prevede l'aumento della pena nel caso la persona che abbia condiviso e/o creato il video sia legata (o sia stata legata) da una relazione affettiva con l'individuo presente nei contenuti *deepfake*[104].

Occorre peraltro menzionare quanto già ricordato dal Servizio Studi del Dipartimento Giustizia nella proposta di legge su citata, vale a dire il fatto che il diritto della persona alla propria immagine è già regolato negli artt. 96 e 97, l. 22 aprile 1941, n. 633 (c. d. legge sulla protezione del diritto d'autore), vietando a chiunque di «esporre o pubbli-

98 Consiglio – Decisione 18/11/2022, n. 2022/2269.

99 Cass. pen., sez. V, 7 aprile 2023, n. 14927.

100 Art. 612-*quater* c.p.

101 Camera dei deputati, *Introduzione dell'articolo 612-quater del codice penale, in materia di manipolazione artificiale di immagini di persone reali allo scopo di ottenerne rappresentazioni nude,* 21 luglio 2021, 1.

102 *Ibid.*

103 *Ibid.*

104 Si può trattare del coniuge (indipendentemente siano essi separati o divorziati) o da persone legate da relazioni affettive passate.

care il ritratto altrui […] senza il consenso dell'interessato»[105,106]. Inoltre, ci si riferisce anche alle rappresentazioni fittizie che vanno a riprendere le sembianze di una persona[107] ed è qui che entrano in gioco i *deepfake porn*. La preoccupazione, infatti, è che i filmati e le foto vengano usate per motivi di ricatto o di revenge porn, data anche la semplicità d'uso che potenzialmente può far sì che ogni persona che ha condiviso una foto online diventi vittima[108].

Dato il realismo delle immagini e dei filmati che vengono creati con le nuove tecnologie, è indubbio che l'onore e la reputazione delle persone presenti in contenuti *deepfake* porn vengano lese. «L'interessato può rivolgersi all'autorità giudiziaria per ottenere la cessazione del comportamento abusivo, il risarcimento del danno ed eventualmente la pubblicazione della sentenza di condanna[109]»[110]. Infine, la vittima ha diritto a richiedere il risarcimento per il danno causato alla sua immagine («pregiudizio economico che la vittima abbia risentito dalla pubblicazione e di cui abbia fornito la prova[111]»). Tuttavia, nel caso non persista il reato di diffamazione, il colpevole rimane obbligato a risarcire la vittima del danno non patrimoniale.

7. Conclusioni

Sin dalla comparsa dei primi *deepfake* nell'anno 2017, questa particolare forma di tecnologia ha manifestato una costante evoluzione, procedendo a un ritmo straordinariamente accelerato, costantemente anticipando coloro che cercano di mitigarne la diffusione e il miglioramento[112]. Secondo Maras e Alexandrou[113], si prospetta un futuro in cui sempre più video artificialmente modificati con l'ausilio dell'IA saranno utilizzati per orchestrare campagne di propaganda terroristica, diffondere notizie politiche false, ricattare individui e per il cyberbullismo.

---

105 Camera dei deputati, *Introduzione dell'articolo 612-quater del codice penale*, 21 luglio, 2021, 2.

106 È necessario ricordare che persistono alcune eccezioni per quanto riguarda la condivisione di immagini senza consenso. «Non occorre il consenso della persona ritratta quando la riproduzione dell'immagine è giustificata dalla notorietà […], da necessità di giustizia o di polizia, da scopi scientifici, didattici o culturali». Tuttavia, «il ritratto non può […] essere esposto o messo in commercio, quando l'esposizione o messa in commercio rechi pregiudizio all'onore, alla reputazione od anche al decoro della persona ritratta» (Art. 97, l 22 aprile 1941, n. 633, c.d. "Legge sul diritto d'autore").

107 Camera dei deputati, *Introduzione dell'articolo 612-quater del codice penale*, cit., 2.

108 Ivi, 1.

109 Ivi., 2.

110 Accanto alla persona lesa, sono autorizzati a sporgere denuncia anche i genitori o i figli.

111 Camera dei deputati, *Introduzione dell'articolo 612-quater del codice penale*, cit., 2.

112 Va comunque detto che la tecnologia *anti-deepfake* sta facendo anch'essa passi da gigante. Negli ultimi anni, infatti, è stata sviluppata una tecnologia che individua i cambiamenti nel volto attraverso l'analisi dei flussi biometrici del sangue (U. A. Çiftçi - I. Demir - L. Yin, *Deepfake source detection in a heartbeat*, in *The Visual Computer*, 2023). Altri ricercatori hanno dimostrato che i video *deepfake* (meno recenti) non riuscivano a simulare in modo perfetto la velocità della chiusura delle palpebre (T. Jung - S. Kim - K. Kim, *DeepVision: Deepfakes Detection Using Human Eye Blinking Pattern*, in *Institute of Electrical and Electronics Engineers*, 2020).

113 M.H. Maras - A. Alexandrou, *Determining authenticity of video evidence in the age of artificial intelligence and in the wake of Deepfake videos*, in *The Internet Journal of Evidence & Proof*, 23, 2018.

Come emerge dall'analisi sin qui svolta, il fenomeno di immagini e video falsi condivisi in rete, non solo non mostra segni di attenuazione, ma sembra destinato a proliferare ulteriormente. Ciò implica che i *deepfake* costituiranno una problematica continua e difficile da affrontare per le democrazie e il sistema d'informazione. Nonostante il ritardo nella definizione di misure adeguate, diversi Stati stanno ora adottando contromisure atte a contenere almeno parzialmente la diffusione di notizie false riportati sui siti web, nei forum e sui social media.

I social media stessi, seppure non tutti agiscano in modo coerente in questa direzione, stanno compiendo sforzi significativi per contrastare la diffusione di notizie false e di *deepfake* che, oltre ad integrare fattispecie delittuose, potrebbero minare il sano processo democratico di un paese.

Richiamando la necessità di normative in questo ambito, è essenziale educare non solo i giornalisti, ma anche il pubblico generale, riguardo ai rischi connessi ai *deepfake*[114]. È però possibile mitigare questo rischio seguendo alcune precauzioni e rispettando semplici norme.

La consapevolezza dell'esistenza di questa tecnologia e delle sue potenzialità aiuta a individuare i segnali e riconoscere i contenuti virtualmente manipolati[115]. Investire nell'educazione digitale è essenziale per prevenire e proteggersi dalla minaccia dei *deepfake*[116]. Una maggiore consapevolezza e conoscenza del mondo online possono rendere il nostro discernimento più attento agli errori grossolani presenti nei video *deepfake*. Gruppi di medici forensi esperti di media consigliano di prestare attenzione a dettagli come oscillazioni del volto, strani riflessi negli occhi, movimenti non congruenti della bocca e sfocature nei contorni della figura[117].

È di tutta evidenza che il *deepfake* impatta in maniera diretta sulla sfera privata, poiché ogni individuo che abbia condiviso anche una sola foto online, potrebbe diventare oggetto della creazione di video *deepfake* senza il proprio consenso. Una misura di prevenzione fondamentale è quindi rappresentata da una maggiore cautela nei comportamenti online. Evitare di condividere indiscriminatamente aspetti della vita privata e monitorare attentamente le immagini online per verificarne un uso non autorizzato rappresentano un primo passo importante[118]. L'utilizzo di password complesse e l'aggiornamento regolare delle stesse contribuiscono a preservare dati personali e bancari[119]. L'adozione di tecnologie come le VPN[120] offre ulteriore protezione contro gli hacker, rendendo più complesso il tracciamento dell'attività online.

---

[114] Stando a dei dati raccolti da iProov (una rinomata azienda di prodotti per l'autenticazione biometrica) nel 2022 su un campione di 16.000 persone in otto paesi (Italia, Spagna, Stati Uniti, Canada, Messico, Germania, Regno Unito e Australia), solamente il 29% sa cos'è un video deepfake (iProov, *How To Protect Against Deepfakes – Statistics and Solutions*, in *iproov.com*, 2022).

[115] L. Sartarelli, *È reato usare un deepfake?*, in *smartius.it,* 31 maggio 2023.

[116] A riguardo si consiglia: Y. Mirsky - L. Wenke, *The Creation and Detection of Deepfakes: A Survey*, in *Association for Computing Machinery*, 54, 2022.

[117] M. Westerlund, *The Emergence of Deepfake Technology*, cit.

[118] L. Sartarelli, *È reato usare un deepfake?*, cit.

[119] J. Wojewidka, *The deepfake threat to face biometrics*, in *Biometric Technology Today*, 2, 2021.

[120] B. Timmerman - P. Mehta - P. Deb - K. Gallagher - B. Dolan – Gavitt - S. Garg - R. Greenstadt, *Studying the Online Deepfake Community*, in *Journal of Online Trust & Safety*, 2, 2023.

In conclusione, in un mondo in cui la tecnologia è dominante e la distinzione tra verità e falsità diventa sempre più difficile, è inevitabile riflettere sulla vulnerabilità della verità e sulla fragilità dell'identità personale. Nonostante il dinamico sviluppo dei *deepfake*, è essenziale non perdere la fiducia nel progresso tecnologico e nella realtà digitale. Solo attraverso un equilibrio tra la maggiore educazione digitale e l'utilizzo di tecnologie avanzate è possibile rendere più sicura la navigazione online, preservando così l'identità personale, l'individualità e la verità nel mondo digitale, attualmente preponderante.

Il reato di sostituzione di persona online di fronte a *socialbot* e ritratti *"AI-generated"*. In favore di una interpretazione estensiva dell'art. 494 del codice penale*

Alessandro Tedeschi Toschi

Abstract

La libertà e la facilità di accesso ai *social media* hanno permesso nuove forme di compimento di alcuni reati, quali la diffamazione, l'adescamento e la sostituzione di persona. Ulteriore aggravio dei possibili danni realizzabili sulle piattaforme è la presenza di *software* in grado di potenziare ulteriormente la diffusione di contenuti lesivi. Tra questi vi sono i *socialbot*, dei programmi capaci di interagire con altri utenti, dando l'impressione di essere delle persone vere. In aggiunta, le più moderne estensioni dell'intelligenza artificiale sono oggi in grado di realizzare dei ritratti realistici di persone inesistenti. La combinazione di queste tecnologie permette la creazione di innumerevoli profili *social*, ciascuno dotato di un'identità fittizia ma credibile, facilmente e velocemente gestibili ed orientabili. La possibilità di fingere efficacemente l'esistenza di un ampio numero di persone può essere sfruttata per molteplici fini ma, in ogni caso, essa causa una lesione della fede pubblica, un bene giuridico tutelato in Italia, tra gli altri, dall'art. 494 c.p.

The freedom and ease of access to social media have enabled new forms of committing certain offences, such as defamation, solicitation and impersonation. Further aggravating the possible harm that can be done on these platforms is the presence of software that can enhance even more the dissemination of harmful content. These include social bots, programs capable of interacting with other users, giving the impression of being real people. In addition, the most modern extensions of artificial intelligence are now able to create realistic portraits of non-existent people. The combination of these technologies allows the creation of innumerable social profiles – each with a fictitious but credible identity – which can be easily and quickly managed and oriented. The possibility of effectively faking the existence of many people can be

* L'articolo è stato sottoposto, in conformità al regolamento della Rivista, a referaggio "a doppio cieco".

exploited for various purposes, but in any case, it causes damage to public faith, which is protected in Italy by – among others – Article 494 of the Criminal Code.

Sommario

1. Social media e *socialbot*, pericoli interconnessi. – 1.1. Interazioni sui *social media* e replicabilità da parte dei *socialbot*. – 1.2. Alcuni caratteri rilevanti dei *socialbot*. – 2. Forme di dissimulazione della natura dei *socialbot* nelle interazioni online con le persone. – 3. La dissimulazione della natura di un *socialbot* come forma di attribuzione a sé di caratteri ai quali la legge attribuisce rilevanza giuridica. – 4. Conclusioni.

1. Social media e *socialbot*, pericoli interconnessi

La recente diffusione dell'utilizzo dei *social media* come strumenti di comunicazione delle opinioni ha portato a dei veri e propri sconvolgimenti nei modi e nei tempi di propagazione del proprio pensiero. Grazie alle architetture digitali ospitate su infrastrutture *hardware* capaci di gestire in autonomia e con precisione l'immensa mole di dati generati dalle interazioni degli utenti (si calcola, ad esempio, che nel solo corso dell'agosto 2021 siano stati pubblicati 575.000 *tweet* ogni minuto),[1] gli utenti dei *social* sono in grado di condividere i propri contenuti senza alcun controllo *ex ante* realmente efficace (gli algoritmi di verifica dei contenuti implementati da alcuni dei gestori di queste piattaforme si sono dimostrati spesso incapaci di contrastare il fenomeno delle cosiddette *fake news*).[2]

All'assenza di forme di verifica dei contenuti pubblicati su queste piattaforme si aggiunge anche la mancanza di vere forme di accertamento dell'identità dei loro utenti. Tale libertà di accesso ai servizi digitali di condivisione, che ha permesso in contesti non democratici l'espressione del proprio dissenso senza timore di censure o ritorsioni, ha anche reso possibile il compimento di reati contro altri diritti fondamentali dell'individuo. In molti hanno potuto approfittare delle forme di anonimato e di mascheramento dell'identità garantite da internet – e dai *social media* – per compiere delitti quali la diffamazione, l'adescamento e la sostituzione di persona (spesso con l'ulteriore obiettivo di realizzare delle truffe).

Le tecnologie digitali di interazione online sono già ritenute degli utili strumenti con cui ampliare ed aggravare la portata lesiva di certi crimini. In particolar modo per i reati di diffamazione, i giudici della Corte di cassazione hanno da tempo riconosciuto la natura delle piattaforme *social* di moltiplicatori dell'estensione della diffusione online di contenuti e, quindi, il potenziale aggravamento delle lesioni della reputazione della persona offesa.[3] In tempi più recenti persino la Corte costituzionale ha sottolineato

[1] *I dati statistici sono disponibili al sito statista.com*.

[2] Si veda, a titolo esemplificativo, K. Shu - A. Sliva - S. Wang - J. Tang - H. Liu, *Fake news detection on social media: A data mining perspective*, in *ACM SIGKDD explorations newsletter*, 19(1), 2017, 22 ss.

[3] Si veda, ad esempio, Cass. pen., sez. I, 2 gennaio 2017, n. 50, in cui i membri della corte hanno

come i caratteri di velocità ed estensione della trasmissione su Internet di un messaggio lesivo dell'onore altrui possano condurre a dei pregiudizi per la vittima estremamente maggiori rispetto a quelli causati dal compimento del medesimo delitto in ambito non digitale.[4]

Ad ulteriore aggravio delle possibilità di danneggiare le persone tramite gli strumenti digitali di diffusione del proprio pensiero vi sono alcuni *software* in grado di potenziare ulteriormente la comunicazione tramite Internet di contenuti pericolosi o lesivi dei diritti altrui. Tra questi la recente cronaca politica ha evidenziato la categoria dei *bot* – un termine molto ampio e da molti lamentato come foriero di confusione – ossia dei programmi *software* capaci, senza alcun intervento umano diretto, di attivarsi a seguito determinate condizioni, di individuare le azioni da compiere per portare a termini i compiti affidatigli e di eseguirle, riconoscendo l'ambiente informatico in cui operano e adattandosi ai cambiamenti di questo.[5] Seguendo le distinzioni compiute dalla più autorevole letteratura in materia,[6] si farà qui riferimento ai soli *bot* che sono capaci di agire su internet e di controllare un profilo di un *social medium* imitando il modo in cui un essere umano interagisce e dissimulando la propria vera natura, i quali vengono chiamati *socialbot*.[7]

1.1. Interazioni sui *social media* e replicabilità da parte dei *socialbot*

Al fine di meglio comprendere l'incisività della presenza dei *socialbot*, è importante sot-

dichiarato che «la diffusione di un messaggio diffamatorio attraverso l'uso di una bacheca "*Facebook*" integra un'ipotesi di diffamazione aggravata». Negli stessi termini si sono poi espressi anche i giudici in Cass. pen., sez. V, 23 gennaio 2017, n. 8482, e Cass. pen., sez. V, 6 settembre 2018, n. 40083. In merito all'attribuzione del carattere di "moltiplicatore" ad internet e ai *social media* si veda M.R. Allegri, *Alcune considerazioni sulla responsabilità degli intermediari digitali, e particolarmente dei social network provider, per i contenuti prodotti dagli utenti*, in *Informatica e Diritto*, 26(1), 2017, 104-105.

[4] Corte cost., ord. 26 giugno 2020 n. 132, in cui viene rilevata la «rapidissima e duratura amplificazione degli addebiti diffamatori determinata dai *social networks* e dai motori di ricerca in internet, il cui carattere lesivo per la vittima - in termini di sofferenza psicologica e di concreti pregiudizi alla propria vita privata, familiare, sociale, professionale, politica - e per tutte le persone a essa affettivamente legate risulta grandemente potenziato rispetto a quanto accadeva anche solo in un recente passato».

[5] Per una definizione approfondita di *bot* si vedano S. Franklin - A. Graesser, *Is it an Agent, or just a Program?: A Taxonomy for Autonomous Agents*, in J.P. Müller - M.J. Wooldridge - N.R. Jennings (a cura di), *Intelligent Agents III Agent Theories, Architectures, and Languages*, Heidelberg, 1996, 21 ss.; M. Tsvetkova - R. García – Gavilanes - L. Floridi - T. Yasseri, *Even good bots fight: The case of Wikipedia*, in *PLoS ONE*, 12(2), 2017.

[6] N. Abokhodair - D. Yoo - D.W. McDonald, *Dissecting a social botnet: Growth, content and influence in Twitter*, in *Proceedings of the 18th ACM conference on computer supported cooperative work & social computing*, 2015, 839 ss.; E. Ferrara - O. Varol - C. Davis - F. Menczer - A. Flammini, *The rise of social bots, in Communications of the ACM*, 59(7), 2016, 96 ss.; S. Stieglitz - F. Brachten - B. Ross - A.K. Jung, *Do social bots dream of electric sheep? A categorisation of social media bot accounts*, in *Proceedings of the 28th Australasian Conference on Information Systems (ACIS)*, 89, 2017; M. Tsvetkova - R. García-Gavilanes - L. Floridi - T. Yasseri, *Even good bots fight: The case of Wikipedia*, cit.

[7] Per una definizione approfondita di *socialbot* si veda Y. Boshmaf - I. Muslukhov - K. Beznosov - M. Ripeanu, *The socialbot network: when bots socialize for fame and money*, in *Proceedings of the 27th annual computer security applications conference*, 2011, 93 ss.

tolineare come tutte le attività "*social*" siano delle forme di interazione preimpostate e standardizzate per le quali esistono appositi e ben identificati pulsanti virtuali. Risulta, quindi, semplice istruire un programma (il *socialbot*) ad individuare le azioni da compiere (ossia i tasti virtuale da premere) e a compierle solo al verificarsi di determinate condizioni che esso è ben capace di riconoscere (come, ad esempio, il fatto che ad un determinato *post* su Facebook di qualcuno questo particolare *bot* non abbia ancora messo un "Mi piace").

Il ruolo di "elevatori alla potenza" della diffusione di contenuti all'interno dei *social network* (che a loro volta vengono considerati come dei moltiplicatori della diffusività) viene svolto dai *socialbot* attraverso diverse condotte, che possono variare anche a seconda della strutturazione del *medium* considerato. Ad ogni modo, si può far rientrare le interazioni compiute all'interno delle reti sociali online in tre categorie: (A) la produzione di contenuti originali (come la pubblicazione di *post* e commenti); (B) la condivisione di contenuti di altri (quali le azioni "Condividi" o *retweet* dei contenuti di terzi, anche esterni alla piattaforma, su Facebook e Twitter); (C) l'apposizione di "reazioni" ai contenuti di altri (ad esempio i "Consiglia" di LinkedIn o gli *upvote* di Reddit).

La distinzione proposta delle azioni che possono essere compiute sui *social media* permette una maggiore comprensione delle dinamiche interne ai *social network* e una più accurata distinzione delle ragioni della rilevanza di ciascuna di esse per il compimento di attività di diffusione potenziata dai *socialbot*. Sebbene l'attenzione rispetto a questi agenti *software* autonomi sia dovuta prevalentemente al loro impiego per propaganda politica e distorsione dei dibattiti online,[8] non si può ignorare come essi possano facilmente essere usati anche per il compimento di reati contro la persona.[9]

Inoltre, la suddivisione delle possibili azioni che possono essere compiute sui *social media* permette di evidenziare con maggiore facilità come esse, nonostante le peculiarità di ciascuna piattaforma, permettano forme di compimento di un reato – spesso aggravato – già accertate dalla nostra giurisprudenza. Infatti, è possibile rinvenire per ciascuna delle categorie qui proposte delle decisioni della Suprema Corte che ben illustrano le ragioni dell'integrazione di una forma aggravata di reato compiuto all'interno dei *social network*.

(A) Produzione di contenuti originali – in merito a questa categoria di attività la Corte di cassazione ha da tempo mostrato come essa possa portare al compimento di reati. In particolare, gli Ermellini hanno dichiarato che tale condotta «integra un'ipotesi di diffamazione aggravata» ai sensi dell'art. 595 c.p., sia nei casi di pubblicazione di *post*

[8] Si vedano, *ex multis*, P.N. Howard - B. Kollanyi, *Bots, #Strongerin, and #Brexit: Computational Propaganda During the UK-EU Referendum*, 2016; E. Ferrara, *Disinformation and social bot operations in the run up to the 2017 French presidential election*, in *First Monday*, 22(8), 2017; L. Luceri - A. Deb - S. Giordano - E. Ferrara, *Evolution of bot and human behavior during elections*, in *First Monday*, 24(9), 2019.

[9] Si veda, ad esempio, A. Tedeschi Toschi - G. Berni Ferretti, *Social media, profili artificiali e tutela della reputazione. Come l'avvento dei social bot per la gestione dei profili social possa rappresentare una grave minaccia per la reputazione delle persone e quali potrebbero essere le risposte a tale pericolo*, in *Rivista Italiana di Informatica e Diritto*, 3(2), 2021, 107 ss.

denigratori[10] che di commenti calunniosi ad essi.[11] In quest'ambito hanno sottolineato come le bacheche dei *social network*, «destinate per comune esperienza ad essere consultate da un numero potenzialmente indeterminato di persone, secondo la logica e la funzione propria dello strumento di comunicazione e condivisione telematica, che è quella di incentivare la frequentazione della bacheca da parte degli utenti, allargandone il numero a uno spettro di persone sempre più esteso», siano un mezzo di pubblicità idoneo «a coinvolgere e raggiungere una vasta platea di soggetti, ampliando – e aggravando – in tal modo la capacità diffusiva del messaggio lesivo della reputazione della persona offesa».[12]

Come già detto, le condotte diffamatorie sulle bacheche dei *social network* sono il risultato di interazioni standardizzate ed è la struttura della piattaforma *social* a rendere automaticamente e in modo non supervisionato visibili ad un numero di persone indeterminato – ma potenzialmente molto ampio – le frasi offensive scritte. Risulta, quindi, facile immaginare un *socialbot* che sia capace, tramite opportuna programmazione, di pubblicare in autonomia contenuti dal carattere diffamatorio, in special modo alla luce dei recenti progressi nella generazione di testi da parte di intelligenze artificiali di *machine learning* (IA di ML).[13] In realtà, l'utilizzo di *socialbot* per la diffusione automatizzata di contenuti testuali precede l'impiego di forme di intelligenza artificiale e può prescindere da questo.[14]

(B) Condivisione di contenuti di altri – riguardo a questa tipologia di condotta verso i contenuti di soggetti terzi, i giudici di piazza Cavour hanno già ritenuto che essa sia la dimostrazione della volontà di «amplificar[li] attraverso il proprio comportamento»[15] [16] e, quindi, qualora detti contenuti integrino un reato (nel caso di specie qui citato di

[10] Cass. pen., sez. I, 8 giugno 2015, n. 24431; Cass. pen., sez. I, 2 gennaio 2017, n. 50; Cass. pen., sez. V, 23 gennaio 2017, n. 8482; Cass. pen., sez. V, 1 febbraio 2017, n. 4873; Cass. pen., sez. V, 6 settembre 2018, n. 40083.

[11] Cass. pen., sez. V, 22 settembre 2004, n. 47452.

[12] Cass. pen., sez. I, 2 gennaio 2017, n. 50.

[13] Si veda, ad esempio, ChatGPT, un *software* di intelligenza artificiale generativo di testi basato su tecniche di *machine learning* sviluppato dalla OpenAI. Si veda anche il caso dell'IA chiamata Tay, descritto in A. Tedeschi Toschi - G. Berni Ferretti, *La responsabilità per la diffamazione compiuta da un'Intelligenza artificiale. Possibili scenari costruiti partendo dall'esempio dell'IA Tay*, in *Ciberspazio e diritto*, 24(74), 2023, 173 ss.

[14] Per fare un esempio, in S. Cresci - R. Di Pietro - M. Petrocchi - A. Spognardi - M. Tesconi, *Social fingerprinting: detection of spambot groups through DNA-inspired behavioral modeling*, in *IEEE Transactions on Dependable and Secure Computing*, 15(4), 2018, 561 ss., viene riportato come dei profili gestiti da questi agenti *software* pubblicassero periodicamente aforismi portando ad ipotizzare che essi si limitassero a pescare da un elenco preimpostato di frasi a cui avevano accesso. Ad ogni modo, gli attuali avanzamenti tecnologici delle IA di ML permettono la creazione di contenuti testuali sempre nuovi, spesso adattandoli a quanto scritto in precedenza dagli interlocutori.

[15] Cass. pen., sez. V, 29 gennaio 2016, n. 3981. In questa particolare sentenza è necessario compiere una fondamentale distinzione lessicale, in quanto i giudici nell'utilizzare il termine "condividere" intendono riferirsi alla partecipazione alla discussione concordando sulla criticabilità del soggetto considerato ma «senza ricorrere alle espressioni offensive utilizzate da altri, né dimostrando di volerle amplificare attraverso il proprio comportamento», mentre nel gergo delle attività sui *social media* tale verbo indica la ripubblicazione delle medesime espressioni offensive, che di fatto ne determina una volontaria amplificazione.

[16] A questa conclusione è giunta praticamente tutta la dottrina che si è interessata alle azioni di ricondivisione di contenuti di altri da parte di utenti di piattaforme *social* online. Si vedano, *ex multis*, N.

diffamazione aggravata), essa stessa integrerà, a sua volta, un reato. In particolare, gli Ermellini hanno precisato come la condotta di condivisione delle critiche ad una persona offesa «potrebbe assumere in astratto rilevanza penale soltanto qualora potesse affermarsi che con il proprio messaggio l'imputato aveva consapevolmente rafforzato la volontà dei suoi interlocutori di diffamare».[17] In questo senso l'azione di condivisione di contenuti lesivi della dignità di una persona, ossia di loro ulteriore diffusione tra gli utenti di un *social medium*, integra pienamente un comportamento esteriore idoneo ad arrecare un contributo apprezzabile alla produzione del danno nei confronti della vittima.[18]

Questa condotta risulta ancora più elementare da essere eseguita, dato che la maggior parte dei *social media* possiede degli appositi pulsanti virtuali con cui poter facilmente propagare i contenuti creati da terzi (si pensi al tasto "Condividi" su Facebook). In questo caso, la delittuosa attività di diffusione di messaggi lesivi del buon nome di qualcuno non abbisogna nemmeno di un minimo di creatività o di abilità argomentativa, basta cliccare un tasto già predisposto dalla piattaforma. Ancora più semplice è, così, programmare un *socialbot* a compiere questa azione.

Vi è da notare come l'automazione di questa attività possa avere, però, dei risvolti inaspettati. In caso di una programmazione priva di analisi del contenuto delle pubblicazioni di terzi da condividere, l'azione di questi agenti *software* viene compiuta senza consapevolezza di quale opinione o condotta venga così rafforzata. In altre parole, questa condotta può avvenire "al buio", solamente sulla base dell'identità del soggetto che li ha originariamente pubblicati ed indifferentemente dalla rilevanza penale delle affermazioni promosse.[19]

Anstead - B. O'Loughlin, *The Emerging Viewertariat and BBC Question Time: Television Debate and Real-Time Commenting Online*, in *The International Journal of Press/Politics*, 16(4), 2011, 440 ss.; E. Bakshy - J.M. Hofman - W.A. Mason - D.J. Watts, *Everyone's an influencer: quantifying influence on Twitter*, in *Proceedings of the fourth ACM international conference on Web search and data mining*, 2011, 65 ss.; T.A. Small, *What the hashtag? A content analysis of Canadian politics on Twitter*, in *Information, Communication & Society*, 14(6), 2011, 872 ss.; G. Elmer, *Live research: Twittering an election debate*, in *New Media & Society*, 15(1), 2013, 18 ss.; T. Highfiel - S. Harrington - A. Bruns, *Twitter as a Technology for Audiencing and Fandom*, in *Information, Communication & Society*, 16(3), 2013, 315 ss.; A.O. Larsson - H. Moe, *Studying political microblogging: Twitter users in the 2010 Swedish election campaign*, in *New Media & Society*, 14-5, 2012, 729 ss.; S. Meraz - Z. Papacharissi, *Networked gatekeeping and networked framing on #Egypt*, in *The International Journal of Press/Politics*, 18(2), 2013, 138 ss.; X.W. Zhao - J. Wang - Y. He - J. Nie - X. Li, *Originator or propagator? Incorporating social role theory into topic models for twitter content analysis*, in *Proceedings of the 22nd ACM International Conference on Information & Knowledge Management*, 2014, 1649 ss.

[17] Cass. pen., sez. V, 29 gennaio 2016, n. 3981.

[18] In quest'ambito si veda il principio ribadito in Cass. pen., sez. V, 10 gennaio 2022, n. 319. In essa, tra le varie questioni, il collegio giudicante ha statuito che il direttore di un quotidiano online, pubblicando integralmente le dichiarazioni diffamatorie di un intervistato, ne ha agevolato in modo apprezzabile l'opera di discredito del buon nome del danneggiato, risultando così concorrente nella sua realizzazione. In altre parole, l'aver contribuito alla diffusione di un messaggio diffamatorio pronunciato da terzi integri il compimento del medesimo reato.

[19] Non si vuole qui avanzare l'ipotesi che una condivisione "al buio" di contenuti precluda di per sé il concorso della volontà del soggetto che ha programmato o amministrato l'agente *software* al verificarsi di determinati eventi lesivi dei diritti di terzi. È importante, però, sottolineare come l'indagine sull'elemento soggettivo di colui che ha creato o gestito il *socialbot*, alla luce di queste particolarità, dovrà essere molto approfondita (soprattutto per determinare se la realizzazione del danno sia stata dolosa o colposa).

(C) Apposizione di "reazioni" ai contenuti di altri – relativamente a quest'ultima categoria di attività online, infine, i Giudici di Piazza Cavour hanno già riconosciuto come essa sia potenzialmente rilevante per la nostra normativa penale, riconoscendo come l'apposizione di un "Mi piace" su Facebook abbia carattere diffusivo, «attesa la […] funzione propalatrice svolta in tale contesto dal *social network*» e sia, quindi, idonea ad avere portata offensiva.[20]

Più recentemente i giudici di cassazione hanno evidenziato come l'apposizione di una semplice "reazione" a contenuti digitali di terzi (nel caso di specie si trattava dei "*like*" ai contenuti presenti sulle bacheche nella piattaforma Facebook) costituisca una manifestazione di «adesione e condivisione dei messaggi»[21] da questi pubblicati e possa integrare pienamente un reato (nel caso di specie quello di istigazione all'odio razziale). Gli Ermellini si sono anche premurati di sottolineare come ciò dipenda dalle modalità di funzionamento della diffusione automatizzata – ed acritica – dei messaggi inseriti nelle bacheche di Facebook. Tale modalità di propagazione automatica di contenuti «dipende dalla maggiore interazione con le pagine interessate da parte degli utenti» e ha un effetto moltiplicativo, dato che sono «le interazioni che consentono la visibilità del messaggio ad un numero maggiore di utenti i quali, a loro volta, hanno la possibilità di rilanciarne il contenuto».

Come già visto per altre interazioni, anche l'apposizione di una "reazione" (come un *like* di Facebook) dipende da strutturazioni del *social medium* preimpostate e facilmente riconoscibili. Di conseguenza, risulta facile intuire come, sfruttando i caratteri dell'algoritmo che regola la continua esibizione agli utenti di nuovi contenuti (il cosiddetto "*newsfeed*" delle piattaforme), l'impostazione di un *socialbot* perché interagisca secondo queste modalità con i contenuti di terzi possa integrare una condotta di loro diffusione tra un numero indeterminato di persone.

Vi è da fare un'importante precisazione in merito a questa terza tipologia di azioni. In essa si vuole far rientrare anche le cosiddette "visualizzazioni" di contenuti multimediali, anche se tale scelta potrebbe apparire come una forzatura. L'inserimento in questa categoria di una condotta che si potrebbe definire come "passiva" e non "di reazione" ai contenuti di altri deriva dal fatto che essa rimane, comunque, registrata come un'azione di interesse per detti contenuti e viene utilizzata come indicatore delle loro popolarità. L'interesse in questa sede per tale azione passiva deriva dal fatto che, come già egregiamente notato dalla Corte di cassazione, gli algoritmi delle piattaforme sono impostati in modo da diffondere ulteriormente – e senza controlli *ex ante* realmente efficaci – tra gli utenti quei contenuti che sono considerati come popolari, ossia che hanno un alto numero di visualizzazioni (si riconosce, comunque, che il tema delle visualizzazioni dei contenuti ha bisogno di particolari cautele e contestualizzazioni).[22]

---

[20] Cass. pen., sez. V, 12 dicembre 2017, n. 55418. La decisione presa dalla Corte era relativa alla questione se la pubblicazione sul proprio profilo Facebook di video inneggianti allo Stato Islamico e l'apposizione di "Mi piace" ad altri integrassero o meno il delitto di istigazione a delinquere previsto dall'art. 414 del codice penale.

[21] Cass. pen., sez. I, 9 febbraio 2022, n. 4534, riportata anche in Redazione, *Se un like ad un post razzista è istigazione all'odio*, in *Diritto di Internet.it*, 2022.

[22] Il problema dell'utilizzo di *bot* per incrementare in numero di visualizzazioni su Internet e, quindi, incrementare artificialmente i loro indici di popolarità è già stato notato in S. Weissmann, *How Not*

1.2. Alcuni caratteri rilevanti dei *socialbot*

La rilevanza della presenza (occultata) dei *socialbot* sulle piattaforme di *social networking* è acuita anche da alcune delle caratteristiche che questi agenti *software* autonomi condividono con la maggior parte delle tecnologie informatiche. Le più rilevanti da riportare in questa sede sono le seguenti:

(A) Localizzazione – Un primo carattere dei *socialbot* che è importante evidenziare in questo contesto è che, per poter operare correttamente, essi necessitano solamente di una connessione internet, senza che tale allaccio avvenga in un luogo specifico. Il risultato di questa particolarità è la possibilità per gli agenti *software* autonomi impiegati, ad esempio, in campagne propagandistiche o diffamatorie di portare avanti la propria opera anche al di fuori dei confini dei paesi in cui si trovano le vittime e rimanendo, quindi, al di fuori della giurisdizione delle Autorità pubbliche deputate alla tutela dei diritti di queste ultime.[23]

(B) Gestione di gruppi – Secondo fattore di importanza per la questione dei *socialbot* è che la gestione di ciascuno di essi non deve necessariamente essere compiuta direttamente da un utente. Esistono programmi, chiamati *botmaster*, che sono in grado di impartire ordini in contemporanea ad una moltitudine di agenti *software*.[24] Questo genere di programma di gestione dei *socialbot* permette ad un utente umano, chiamato *botherder*, di controllare un'intera folla di profili automatizzati descrivendo il singolo comando ad esso e lasciando che sia questo materialmente ad impartirlo ad ogni singolo *software* di gestione dei profili. I *botmaster* fungono, insomma, da strumento di automazione della gestione automatizzata e dissimulata dei profili *social*, finendo col diluire ulteriormente il grado di supervisione delle dinamiche che possono essere scatenate dall'impiego dei *socialbot*.[25]

Sebbene da un punto di vista giuridico la questione non rivesta particolari difficoltà interpretative in merito all'attribuzione di eventuali reati compiuti tramite questi "strumenti di gestione di strumenti di gestione", essa ha comunque rilevanza. Infatti,

---

to Regulate Social Media, in *The New Atlantis*, 58, 2019, 58 ss.; B. Stricke, *People v. Robots: A Roadmap for Enforcing California's New Online Bot Disclosure Act*, in *Vanderbilt Journal of Entertainment & Technology Law*, 22, 2020, 838 ss.

[23] Come già perspicacemente rilevato anche dal deputato Seán Kyne nel corso del dibattito del 13 dicembre 2017 sulla proposta di legge irlandese di contrasto alle attività dei *socialbot*, la *Online Advertising and Social Media (Transparency) Bill 2017*. In quest'occasione egli dichiarò che «l'idea che possiamo rintracciare i *bot* e perseguire le persone o le organizzazioni dietro di loro ignora la realtà che se la persona che gestisce i falsi *account* ha sede fuori dallo Stato, come la maggior parte di loro, allora sarà al di là della portata del disegno di legge». La trascrizione del dibattito è disponibile all'indirizzo *oireachtas.ie*.

[24] Per un'analisi delle dinamiche di gestioni di grandi gruppi di *socialbot* si veda, ad esempio, Y. Boshmaf - I. Muslukhov - K. Beznosov - M. Ripeanu, *The socialbot network: when bots socialize for fame and money*, cit., 93 ss.

[25] Si pensi, ad esempio, a dei *socialbot* programmati per rilanciare tutti i messaggi pubblicati dal profilo di una persona di chiara fama e attiva nella beneficenza. Nel caso in cui questo *account* venisse violato ed utilizzato per promuovere una truffa, i *socialbot* continuerebbero a promuovere i messaggi pubblicati da questo senza valutare minimamente il loro contenuto e il loro amministratore potrebbe impiegare diverso tempo per accorgersene, rischiando nel frattempo di indurre numerose persone a cadere vittime di un disegno criminoso. Un utilizzo non adeguatamente supervisionato di strumenti per automatizzare e accelerare la diffusione di contenuti potrebbe causare seri danni.

l'impiego di strumenti atti ad incrementare ulteriormente la portata dannosa di un messaggio diffamatorio o di istigazione all'odio ben finirà col costituire un'aggravante della condotta posta in essere, dato l'incremento della diffusione del contenuto.
Nemmeno la creazione dei profili gestiti in modo automatizzato, poi, deve essere necessariamente compiuta manualmente da un individuo. Esistono, infatti, dei programmi (che possono essere fatti rientrare all'interno della categoria dei *bot*) strutturati in modo da creare in automatico nuovi profili sulle piattaforme di *social networking*.[26] Non esistono, poi, particolari ostacoli alla combinazione di questa tecnologia con quella dei *botmaster*, permettendo la creazione e la gestione automatizzata di folle oceaniche di sostenitori artificiali da parte di chiunque conosca questi programmi o possa permettersi di finanziare questo procedimento.
(C) Flessibilità – Terzo aspetto di preminente rilevanza della natura dei *socialbot* è il fatto essi sono programmati per operare su di una specifica piattaforma ma non per gestire un determinato profilo. Questo perché, come accennato in precedenza, ciascun *social medium* è composto da elementi preimpostati e standardizzati, che questi agenti *software* sono ben in grado di riconoscere e di gestire. Tutto ciò significa che, in caso di sospensione (temporanea o perpetua) di un particolare profilo, non esistono impedimenti tecnici alla consegna al *socialbot* delle credenziali di accesso ad un altro *account* con cui esso possa continuare in autonomia la propria opera di diffusione potenziata di contenuti.
(D) Natura strumentale – Infine, tra i caratteri dei *socialbot* e delle piattaforme su cui operano c'è da evidenziare con enfasi uno più generale che contrassegna questo tipo di tecnologie: la loro natura strumentale.[27] Infatti, come già detto in precedenza, essi possono essere impiegati per la diffusione di qualsiasi tipo di contenuto sulle piattaforme *social* (dalle informazioni sulle variazioni metereologiche fino a teorie cospirazioniste di stampo antisemita). In altre parole, i *socialbot* sono dei mezzi, piuttosto sofisticati, per compiere una serie di attività che possono anche – ma non solo – recare danni ad altri.
La conseguenza della somma di tutte le caratteristiche dei *socialbot* qui evidenziate è che essi possono facilmente essere impiegati per creare su di un *social medium* un'ampia schiera di utenti artificiali che sono in grado di dare l'impressione di essere persone profondamente convinte della validità delle opinioni espresse online ed assiduamente concentrate ad incrementarne la diffusione.

[26] N. Perlroth, *Fake Twitter Followers Becomes Multimillion Dollar Business*, in *The New York Times*, 5 aprile 2013, in cui viene fatto riferimento ad un *software* che «potrebbe creare fino a 100.000 nuovi account in cinque giorni».

[27] Questa osservazione viene avanzata anche in M. Lamo-R. Calo, *Regulating Bot Speech*, in *UCLA Law Review*, 66, 2019, 988 ss; J. Horder, *Online Free Speech and the Suppression of False Political Claims*, in *Journal of International and Comparative Law*, 8, 2021, 15 ss. Il riconoscimento della strumentalità di questi agenti *software* autonomi avviene anche – in modo – implicito nella Sezione 7 del *Protection from Online Falsehoods and Manipulation Act 2019* (*POFMA*) della Repubblica di Singapore. In merito a questa normativa si veda A. Tedeschi Toschi-G. Berni Ferretti, *Il contrasto legislativo ai socialbot e le soluzioni avanzate nella Repubblica di Singapore e nella Repubblica d'Irlanda*, in questa *Rivista*, 3, 2022, 352 ss.

2. Forme di dissimulazione della natura dei *socialbot* nelle interazioni online con le persone

La definizione di *socialbot* basa la propria distinzione rispetto ad altri programmi *software* autonomi sull'occultamento della sua natura ai destinatari delle interazioni che compie tramite un profilo *social*. Quest'opera di mimetizzazione viene portata avanti tramite diversi accorgimenti che alcuni legislatori hanno anche cercato di catalogare, addirittura inserendo all'interno di nuove normative di rango primario l'elencazione degli elementi da considerare. L'esempio principale di cristallizzazione legislativa degli indici di valutazione è quello del *Protection from Online Falsehoods and Manipulation Act 2019* (*POFMA*) della Repubblica di Singapore. Alla sezione 40 (4) vengono indicati i fattori che le Autorità pubbliche devono prendere in considerazione per determinare se il profilo investigato sia controllato da un *socialbot*.[28] Un'esplicitazione in sede normativa degli aspetti da valutare ha senz'altro il pregio di fornire una prima serie di indici che permetta alle Autorità di avere una direzione verso cui orientare lo svolgimento delle proprie indagini. Tuttavia, essa non fornisce indicazioni su come ponderare il valore di ciascuno degli elementi indicati e ha anche il notevole svantaggio di rendere meno flessibili le loro attività in un ambito – quello informatico – caratterizzato dall'introduzione di profonde innovazioni in tempi rapidissimi.

La principale letteratura scientifica sull'argomento, invece, pone l'accento sulla corrispondenza o meno delle interazioni compiute con dei modelli comportamentali standard estrapolati dalle condotte di un gran numero di profili usati da utenti umani.[29] In altre parole, le metodologie di controllo della possibile presenza di *socialbot* oggi si basano principalmente sull'analisi dei comportamenti tenuti dai profili all'interno dei *social network* e sulla corrispondenza con le condotte abitualmente tenute in essi dalle persone. Questa metodologia di valutazione è operata prendendo in considerazione diversi elementi, quali la ricchezza di informazioni dei profili, la struttura delle reti sociali entro le quali sono attivi, i contenuti da questi pubblicati, il tono delle opinioni espresse ed il tempismo delle loro reazioni.

C'è da evidenziare come venga attribuita una rilevanza relativa agli aspetti della ricchezza di informazioni dei profili, come la presenza di ritratti o di fotografie negli "spazi personali" di questi. La scarsa considerazione riservata a questi elementi – considerati come essenziali dalla nostra giurisprudenza per il compimento del reato di

[28] La Sezione, infatti, dispone che «nel determinare se un *account online* è un *account online* non autentico o è controllato da un *bot*, il Ministro competente deve tenere conto dei seguenti fattori: (a) se le informazioni utilizzate nella creazione del profilo *online* si riferiscono a un paese o un territorio diverso dal paese o territorio da cui presumibilmente proviene il titolare del profilo; (b) se esista uno schema di attività sospette svolte utilizzando l'*account online*; (c) la data in cui è stato creato il profilo *online*; (d) qualsiasi altro fattore che il Ministro competente consideri rilevante». È importante qui evidenziare come la lettera (d) di questa Sezione permetta all'autorità pubblica investita del compito di svolgere l'indagine di includere qualsiasi altro fattore che consideri «rilevante» in ciascun caso concreto.

[29] L. Alvisi - A. Clement - A. Epasto - S. Lattanzi - A. Panconesi, *Sok: The evolution of sybil defense via social networks*, in *2013 IEEE symposium on security and privacy*, 2013, 382 ss.; L. Luceri - A. Deb - S. Giordano - E. Ferrara, Evolution of bot and human behavior during elections, cit.; S. Cresci, *A decade of social bot detection*, in *Communications of the ACM*, 63(10), 2020, 72 ss.

sostituzione di persona sui *social media*[30] – trova giustificazione nel fatto che, come ben evidenziato da Cresci, Di Pietro, Petrocchi, Spognardi e Tesconi nelle loro indagini sui profili che hanno sostenuto uno dei candidati alla contesa elettorale capitolina del 2014,[31] l'opera di mimesi degli *account* consiste non solo nel corredare i profili forniti ai *socialbot* di ritratti e descrizioni credibili ma anche – e soprattutto – nell'assemblare il loro codice sorgente in modo da imitare le condotte medie tipiche di un essere umano. La ragione della maggior importanza attribuita alle modalità di interazione di un profilo su di una piattaforma *social* deriva anche dal fatto che la strutturazione di internet ha da sempre reso estremamente semplice ricercare ed ottenere copia di immagini ritraenti altre persone. Lo stesso gruppo di ricercatori dell'IIT-CNR guidato da Cresci ha riportato come i profili automatizzati impiegati per la promozione di uno dei candidati all'elezione del sindaco di Roma fossero corredati da fotografie rubate.[32] Ad ulteriore giustificazione della limitata rilevanza della presenza di fotografie vi è anche il fatto che ormai le attuali tecnologie permettono di prescindere integralmente dall'utilizzo di caratteri e generalità di altre persone esistenti od esistite.[33] Infatti, grazie all'utilizzo di intelligenze artificiali è oggi possibile generare un numero quasi infinito di ritratti realistici di persone inesistenti (ritratti cosiddetti *AI-generated*).[34] Quindi, ormai può essere abbandonato il cosiddetto "furto d'identità" altrui ed essere comunque facilmente compiuta su internet l'attribuzione a sé o ad altri di un falso nome, di un falso stato ovvero di qualità a cui la legge attribuisce effetti giuridici.

3. La dissimulazione della natura di un *socialbot* come forma di attribuzione a sé di caratteri ai quali la legge attribuisce rilevanza giuridica

All'interno del nostro ordinamento viene attribuita rilevanza penale alle condotte di dissimulazione o di attribuzione di caratteri personali che possano indurre taluno in

[30] Si vedano *infra* le sentenze della Corte di Cassazione relativo all'uso abusivo dell'immagine e delle generalità di altre persone inconsapevoli.

[31] S. Cresci - R. Di Pietro - M. Petrocchi - A. Spognardi - M. Tesconi, *The paradigm-shift of social spambots: Evidence, theories, and tools for the arms race*, in *Proceedings of the 26th international conference on world wide web companion*, 2017, 963 ss.; Id., *Social fingerprinting: detection of spambot groups through DNA-inspired behavioral modeling*, cit., 561 ss.

[32] Ivi, 563 ss.

[33] S. Bond, *That smiling LinkedIn profile face might be a computer-generated fake*, in *npr.org*, 27 marzo 2022, in cui vengono riportate in modo dettagliato le ricerche condotte da Renée DiResta, technical research manager dello Stanford Internet Observatory.

[34] Si vedano, ad esempio, A. Ramesh - P. Dhariwal - A. Nichol - C. Chu - M. Chen, *Hierarchical Text-Conditional Image Generation with CLIP Latents*, in *arXiv.org*, 2022; T. Karras - S. Laine - M. Aittala - J. Hellsten - J. Lehtinen - T. Aila, *Analyzing and Improving the Image Quality of StyleGAN*, in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2020, 8110 ss.; S.J. Nightingale - H. Farid, *AI-synthesized faces are indistinguishable from real faces and more trustworthy*, in *Proceedings of the National Academy of Sciences*, 119(8), 2022. Negli articoli citati gli autori illustrano i percorsi di sviluppo di intelligenze artificiali in grado di creare immagini veritiere di alta qualità di persone, animali, oggetti e paesaggi inesistenti (il secondo e il terzo articolo, in particolare, contengono un ampio numero di esempi di ritratti di persone generati artificialmente).

errore sulla effettiva identità di una persona.[35] Dalla diffusione di internet e dei *social media* questo crimine è diventato estremamente più frequente di quanto non fosse ai tempi della promulgazione del Codice e, di conseguenza, è oggetto della giurisprudenza di Cassazione con sempre maggiore cadenza.

Nelle loro numerose decisioni gli Ermellini hanno avuto modo di precisare come questo delitto venga integrato anche dalla condotta di colui che crei ed utilizzi un *account* online attribuendo ad esso le generalità di un diverso soggetto ed inducendo in errore altre persone attive su internet nei confronti delle quali le false generalità siano state declinate.[36] In altre occasioni hanno specificato come la sostituzione di persona sia integrata anche dalla creazione su di un *social medium* di un profilo con l'utilizzo abusivo dell'immagine di una persona del tutto inconsapevole, associata a generalità o soprannomi di fantasia, e dal suo impiego per condividere contenuti ed indurre in errore i suoi interlocutori.[37]

La giurisprudenza di Cassazione ha anche valorizzato il fatto che il dolo specifico del delitto di cui alla norma considerata può consistere in varie finalità, anche profondamente diverse tra loro. I giudici di piazza Cavour hanno, infatti, avuto modo di ritenere integrato il reato di sostituzione di persona sia quando il perpetratore lo abbia compiuto per procurarsi un ingiusto profitto economico (con danno del titolare dell'identità abusivamente utilizzata),[38] sia quando ne abbia tratto un vantaggio non patrimoniale, come la possibilità di intrattenere rapporti con altre persone o il soddisfacimento di una propria vanità.[39] Inoltre, le sentenze della Suprema Corte hanno chiarito che la violazione della norma può realizzarsi anche quando, con le condotte tenute, sia stato solamente causato un danno (anche non patrimoniale) ad altri, come la lesione dell'immagine o della dignità delle vittime.[40]

Inoltre, gli Ermellini hanno avuto modo in più occasioni di precisare che l'integrazione del reato – e non solo del suo tentativo – si verifica addirittura nel caso in cui il vantaggio perseguito dall'agente non sia da questo effettivamente raggiunto ma l'uso di mezzi

---

[35] L'art. 494 c.p. stabilisce che: «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, è punito, se il fatto non costituisce un altro delitto contro la fede pubblica, con la reclusione fino a un anno».

[36] Si vedano, ad esempio, Cass. pen., sez. V, 14 dicembre 2007, n. 46674, in cui i giudici hanno confermato la condanna all'imputato che aveva creato un *account* di posta elettronica utilizzando il nome della vittima e ledendone l'immagine e la dignità attraverso l'allacciamento di rapporti con altre persone; Cass. pen., sez. III, 3 aprile 2012, n. 12479, in cui è stato riconosciuto sussistere il delitto nella condotta di colui che crei ed utilizzi un *account* ed una casella di posta elettronica, servendosi dei dati anagrafici di un diverso soggetto, inconsapevole, con il fine di far ricadere su quest'ultimo l'inadempimento delle obbligazioni conseguenti all'avvenuto acquisto di beni; Cass. pen., sez. V, 6 luglio 2020, n. 22049, e 5 febbraio 2021, n. 12062.

[37] Cass. pen., sez. V, 29 aprile 2013, n. 18826, 16 giugno 2014, n. 25774 (in cui gli Ermellini hanno rilevato come il profilo creato fosse corredato da «una descrizione tutt'altro che lusinghiera»), e 08 giugno 2018, n. 33862.

[38] Si vedano, ad esempio, Cass. pen., sez. II, 17 maggio 2019, n. 21705, e 02 luglio 2020, n. 23760.

[39] Cass. pen., sez. V, 16 giugno 2014, n. 25774.

[40] Cass. pen., sez. V, 14 dicembre 2007, n. 46674; Cass. pen., sez. V, 16 giugno 2014, n. 2577.

fraudolenti abbia comunque indotto in errore il soggetto passivo del reato.[41]
Sebbene questo articolo del codice penale possa dare l'impressione di essere sufficiente da solo a costituire un valido strumento di difesa della pubblica fede dall'illecita induzione in errore degli interlocutori online dei *socialbot* – soprattutto alla luce dell'opinione recentemente espressa dalla Corte di Cassazione che «oggetto della tutela penale è l'interesse riguardante la pubblica fede»[42] – vi sono da tenere in considerazione diversi peculiari aspetti della tecnologia di questi agenti *software* autonomi che rendono l'applicazione della norma meno scontata di quanto non possa apparire a prima vista.
(A) Attribuzione a sé o ad altri – La prima problematica che riguarda l'applicazione dell'art. 494 c.p. anche alle attività compiute dai *socialbot* è che la lettera della norma prevede che si integri il reato sostituendo «la propria all'altrui persona, o attribuendo a sé o ad altri» identità o caratteri non propri. Il dubbio interpretativo che potrebbe sorgere qui deriva dal fatto che detti caratteri non sembrano *prima facie* attribuiti ad una persona fisica ma, invece, ad un *software* caricato su di una macchina, ossia ad un bene immateriale che non costituisce nemmeno un autonomo centro di imputazione di rapporti giuridici. Vi è, infatti, un terreno ancora nebbioso del diritto entro i cui confini ad alcuni sembra possibile argomentare che, dal momento che è il *socialbot* a portare avanti in modo autonomo le interazioni sulle piattaforme *social* (quali pubblicare contenuti o apporre "reazioni" ai contenuti di altri), le qualità «a cui la legge attribuisce effetti giuridici» debbano essere riferite ad esso.
Una prima avventata – per non dire errata – interpretazione di questa situazione potrebbe essere di ritenere che l'utilizzo di un *socialbot* impedisca l'integrazione della fattispecie criminosa, in quanto identità o caratteri verrebbero attribuiti ad un oggetto e non «a sé o ad altri». Un'altra interpretazione – parimenti incauta – potrebbe essere quella di confondere le capacità di azione autonoma della macchina con il possesso di autonomia decisionale da parte di essa e, quindi, di reclamare una qualche forma di imputabilità nei suoi confronti.
Davanti al rischio di attribuire erroneamente delle effettive capacità decisionali – ed una possibile conseguente responsabilità penale – a quelli che, in fin dei conti, sono solo degli elaboratori dati che si limitano a seguire un elenco piuttosto flessibile di comandi preimpostanti (ossia, ad eseguire un codice sorgente particolarmente complesso), si deve dare risalto al fatto che essi si limitano ad essere degli strumenti – sebbene particolarmente sofisticati – che svolgono solamente le azioni che un programmatore ha inserito nel loro codice sorgente o i comandi che il loro amministratore gli dà. Lo stesso concetto di autonomia di azione di questi agenti *software* poggia, infatti, su di una loro programmazione che li abbia dotati di un ampio bagaglio di parametri che gli permettono di identificare numerosi ambienti informatici e ad interagire con essi.[43] La

[41] Cass. pen., sez. V, 19 marzo 1985, n. 2542; Cass. pen., sez. V, 16 marzo 2015, n. 11087, e *a contrariis* Cass. pen., sez. V, 18 dicembre 2020, n. 5432.

[42] Cass. pen., sez. V, 16 giugno 2014, n. 25774, che riprende quanto statuito precedentemente in Cass. pen., sez. V, 14 dicembre 2007, n. 46674.

[43] Un *socialbot* correttamente programmato contiene all'interno del suo codice sorgente tutti i parametri per identificare correttamente gli elementi contenuti nel sito internet di un *social medium* e le indicazioni di come interagire con essi. Quindi, quando agisce all'interno di un *social network*, un *socialbot* si limita ad interagire con gli elementi che è in grado di riconoscere grazie alle librerie di cui è dotato (ad esempio,

loro indipendenza di azione non deve, però, essere confusa con libertà di giudizio o di iniziativa: una volta attivati sono, sì, in grado di portare a termine il compito affidatogli senza bisogno di assistenza ma non sono capaci di decidere da soli il loro scopo.
In questa sede è anche importante evidenziare come il creatore di un *socialbot* non deve necessariamente rispondere di tutte le azioni compiute per mezzo del suo strumento *software*. Infatti, sebbene sia lui a decidere quali siano le interazioni che il suo prodotto debba poter compiere, è ben possibile che non sia lui ad indicargli quali condizioni debbano fargli intraprendere le sue attività. È solo colui che impiega attivamente il *socialbot* a dover rispondere degli eventi realizzati per il tramite di questo agente *software*.[44] Questo perché è solo l'utilizzatore effettivo del *socialbot* ad attivarlo e ad impostare gli elementi che, se rilevati dal programma, lo faranno agire all'interno del *social network*.[45] Ciò non toglie, comunque, che molto spesso il creatore di un *socialbot* sia anche il suo utilizzatore finale ma è importante sottolineare come non sia necessariamente sempre così.
Insomma, l'attribuzione di qualità rilevanti per la nostra legge penale non deve immaginarsi avvenire in capo ad un bene immateriale (del quale alcuni potrebbero erroneamente arrivare a congetturare una qualche forma di personalità giuridica). Essa, invece, deve essere compiuta in capo all'utilizzatore del *socialbot*, con quest'ultimo che funge da strumento particolarmente raffinato e complesso per l'induzione in errore dei destinatari delle interazioni sulle piattaforme *social*. In altre parole, l'impiego di questo agente *software* costituisce una forma particolarmente intricata di attribuzione a sé di determinate qualità, dal momento che esso è semplicemente uno strumento sottoposto in ogni momento al controllo del suo amministratore (il quale può intervenire a suo piacimento per direzionarlo o fermarlo).
Le più recenti innovazioni in campo informatico creano, poi, un ulteriore nodo che, allo stato attuale, è ammantato da pesanti incertezze interpretative. L'attuale sviluppo dei *software* di intelligenza artificiale (IA) è ormai giunto al punto in cui essi sono in grado di produrre in autonomia testi di senso compiuto. Ben poco impedisce oggi ad un *socialbot* di essere progettato in modo da interagire su di un *social medium* e da produrre un testo – o anche di una immagine – grazie a forme di IA.[46] In altre parole, un *socialbot*

può essere in grado di cliccare il pulsante "Mi piace" di un *post* su Facebook perché riconosce gli elementi della piattaforma) mentre, quando incontra un elemento che non riconosce o non trova l'elemento che sta cercando, esso potrebbe saltare l'operazione che avrebbe dovuto compiere. Il grado di adattabilità e di autonomia d'azione di un *socialbot* dipende, quindi, dalla qualità della sua programmazione.

[44] Al programmatore che abbia creato il *socialbot* e lo abbia fornito al suo utilizzatore potrebbe essere sempre contestata la partecipazione alla realizzazione di un reato nella posizione di ausiliatore, qualora fosse provata la sua volontà di apportare un aiuto materiale alla realizzazione del conosciuto progetto delittuoso. In merito alla figura del cosiddetto "complice ausiliatore" si veda quanto riportato in G. Fiandaca - M. Musco, *Diritto penale. Parte generale*, Bologna, 2009, 504, e le recenti precisazioni in Cass. pen., sez. V, 9 novembre 2021, n. 8973.

[45] Si immagini un *socialbot* programmato per agire su Facebook e programmato per mettere un "Mi piace" solamente a quei *post* di altri utenti che contengano al proprio interno una o più parole-chiave impostate dal suo amministratore. In questo esempio il programmatore ha solamente impostato la funzionalità mentre è l'amministratore ad indicare al *software* quale sia la parola-chiave che, se rilevata, gli fa avviare il processo di apposizione del "Mi piace". Differenti copie del medesimo *socialbot* ben potrebbero essere usate da persone diverse per promuovere messaggi di tolleranza oppure di discriminazione senza che il loro creatore abbia modo di intervenire.

[46] Già nel 2016 la Microsoft aveva creato una IA in grado sia di produrre testi che interagire con le

ben potrebbe essere progettato in modo da integrare funzionalità di intelligenza artificiale di tipo generativo che gli permettono di produrre contenuti e di pubblicarli poi su di una piattaforma *social* per rendere ancora di più l'impressione di essere una persona vera.

In merito al particolare – ma plausibile – scenario di un *socialbot* capace di generare in autonomia dei contenuti artificiali è importante evidenziare come le legislazioni vigenti siano in uno stato di grave arretratezza ma, allo stesso tempo, come sia possibile individuare alcuni appigli per una loro interpretazione che permetta di dare risposta a qualsiasi richiesta di giustizia. Sebbene, infatti, il tema della produzione da parte delle IA di testi ed immagini stia generando un acceso dibattito, è comunque possibile individuare un orientamento sempre più forte – seguito anche all'interno dei dibattiti degli organi normativi dell'Unione europea[47] – che attribuisce la responsabilità per i danni cagionati a coloro che avevano i privilegi di amministratore di questi *software* e che non hanno supervisionato con sufficiente diligenza i loro processi di produzione di *output*.[48] Quindi, anche nel caso di un *socialbot* con caratteri di IA, sarà il suo amministratore a dover rispondere degli eventi causati da questo (oltre che dell'illecita attribuzione di qualità a cui la legge attribuisce rilevanza).

Quindi, dalle considerazioni appena esposte in merito alla natura strumentale dei *socialbot* si può facilmente ricavare che l'azione di sostituire «illegittimamente la propria all'altrui persona» viene compiuta, nella realizzazione del reato di cui all'art. 494 c.p., dal soggetto che possiede i privilegi di amministratore di questi agenti *software* autonomi. Costui, infatti, si limita ad utilizzare un sofisticato strumento digitale (il *socialbot*) per far apparire sé stesso come un'altra persona su di una piattaforma di *social networking*.

Chiarito come sia l'amministratore del *socialbot* l'effettivo perpetratore del reato, conviene analizzare ulteriormente come tale condotta di mascheramento della propria identità possa essere realizzata. Infatti, all'interno del campo della giurisprudenza relativa all'art. 494 c.p. la valutazione dell'illegittima sostituzione della propria all'altrui persona su internet è giunta ormai ad affrontare numerose e disparate declinazioni. In particolare, molte delle recenti decisioni della Corte di cassazione hanno avuto ad oggetto la creazione di profili sui principali *social media* corredati di generalità di altre persone e di fotografie "trafugate" che le ritraevano. Questi profili sono stati usati per contattare, interagire e stringere relazioni con altri utenti ed indurli a credere che il loro interlocutore avesse le caratteristiche che venivano loro mostrato tramite il profilo così costruito e tramite i messaggi ed altre forme di interazione (come l'invio di fotografie o l'apposizione di reazioni ai loro contenuti).

persone su diversi *social media* (ma specificando che si trattava di una IA). Il risultato dell'esperimento di interazione di una IA con gli utenti di più piattaforme *social* aveva dato esiti disastrosi e l'impresa di Redmond aveva sospeso gli *account* della sua creatura digitale. In merito a questa vicenda si veda A. Tedeschi Toschi - G. Berni Ferretti, *La responsabilità per la diffamazione compiuta da un'Intelligenza artificiale*, cit., 173 ss.

[47] Si vedano, in particolare, la proposta di regolamento COM(2021) 206 final, che stabilisce regole armonizzate sulla intelligenza artificiale (legge sull'intelligenza artificiale) e la proposta di direttiva COM(2022) 496 final, relativa all'adeguamento delle norme in materia di responsabilità civile extracontrattuale all'intelligenza artificiale.

[48] Su questo tema si veda, ad esempio, A. Tedeschi Toschi - G. Berni Ferretti, *La responsabilità per la diffamazione compiuta da un'Intelligenza artificiale*, cit., 173 ss.

Tra le varie declinazioni della sostituzione della propria all'altrui persona che sono giunte all'attenzione dei giudici della Suprema Corte pare qui importante riportare come questi siano stati chiamati a valutare anche il caso in cui ad essere "trafugata" era stata solo la fotografia di un'altra persona mentre le sue generalità non venivano impiegate (tale condotta è stata, comunque, ritenuta integrante il delitto descritto dal 494 c.p.).[49] Addirittura, è stato portato davanti alla Corte di cassazione il caso, poi sanzionato come illecito, dell'utilizzo non di una fotografia ma di un disegno caricaturale di un'altra persona.[50]

La particolare attenzione riservata dall'Autorità giurisdizionale per l'illecito impiego di elementi grafici che ritraggono altre persone (spesso inconsapevoli) può trovare una valida argomentazione nel fatto che detto utilizzo di raffigurazioni tra le condotte realizzate in violazione dell'art. 494 c.p. arreca un danno non solo al destinatario delle interazioni *online* compiute dal reo ma anche alla persona effettivamente ritratta nell'immagine. La perpetrazione del delitto in questa forma, insomma, comporta una lesione non solo alla fede pubblica[51] ma anche del diritto all'identità e al riconoscimento del giusto credito sociale del "derubato" dell'immagine (il quale facilmente può vedersi, così, attribuiti fatti e parole non propri con un conseguente potenziale detrimento della sua considerazione da parte di altri). Tuttavia, c'è da notare che le attuali capacità delle tecnologie digitali permettono di superare la causazione del duplice danno appena descritto. Infatti, come detto, le più recenti tecnologie digitali permettono la produzione in massa e quasi immediata di ritratti *AI-generated* realistici e sufficientemente credibili. Così, in combinazione con l'utilizzo di *nickname* privi di allusioni o riferimenti ad altre persone, è possibile creare profili *social* che non rientrano nell'ambito delle summenzionate interpretazioni della norma offerte dalla Suprema Corte.

La creazione di un profilo secondo gli accorgimenti appena descritti (ritratto generato da una IA e *nickname* di fantasia) eviterebbe la lesione dell'immagine o della dignità di una persona (dal momento che non vi è alcun soggetto esistente che sia effettivamente titolare delle generalità spese) ed eviterebbe, quindi, interventi delle autorità a tutela del diritto all'identità personale. Ne deriva, così, il rischio che la gravità delle condotte poste in essere da chi mascheri la gestione di un *account* con caratteri falsificati affidato ad un *socialbot* sia percepita come estremamente inferiore. In fondo, quando non esiste una persona che sia stata "derubata" della propria identità e che rischi di vedersi ingiustamente attribuire fatti o dichiarazioni, viene meno uno dei soggetti che possono essere offesi dal perpetratore del reato.[52]

---

[49] Si veda, ad esempio, Cass. pen., sez. V, 16 giugno 2014, n. 25774, in cui gli Ermellini hanno ritenuto che «integra il delitto di sostituzione di persona (art. 494 c.p.) la condotta di colui che crea ed utilizza un "profilo" su *social network*, utilizzando abusivamente l'immagine di una persona del tutto inconsapevole, associata ad un "*nickname*" di fantasia ed a caratteristiche personali negative».

[50] Cass. pen., sez. V, 23 luglio 2020, n. 22049.

[51] Bene giuridico espressamente e direttamente tutelato dalla norma, come specificato in Cass. pen., sez. II, 11 settembre 2020, n. 26589.

[52] Riguardo all'importanza attribuita alla persona le cui generalità sono state illegittimamente spese online si veda, ad esempio, Cass. pen., sez. V, 16 giugno 2014, n. 25774, in cui viene fatto riferimento, *ex plurimis*, alle sentenze 29 aprile 2013, n. 18826, 27 marzo 2009, n. 21574, 9 dicembre 2008, n. 7187, e 25 ottobre 2007, n. 237855.

(B) Induzione in errore – A solido bastione contro le preoccupazioni appena esposte vi è il fatto che, ai fini dell'applicazione dell'art. 494 c.p., tale norma non considera fondamentale l'illegittimo impiego dell'identità di una persona realmente esistente. La realizzazione del reato viene riconosciuta anche quando avviene attraverso l'attribuzione di generalità, stati o qualità che abbiano rilevanza giuridica che non siano corrispondenti al vero. Questo perché il momento consumativo del reato è stato individuato nell'altrui induzione in errore e perché «oggetto della tutela penale è l'interesse riguardante la pubblica fede, in quanto questa può essere sorpresa da inganni relativi alla vera essenza di una persona o alla sua identità o ai suoi attributi sociali; siccome si tratta di inganni che possono superare la ristretta cerchia d'un determinato destinatario».[53]

Giova ribadire che, se il momento consumativo del reato viene identificato nell'induzione in errore circa l'essenza di chi – o di cosa – si nasconde realmente dietro la maschera del profilo *social*, deve essere tenuto ben presente che gli elementi che possono condurre a tale sbaglio sono sempre direttamente o indirettamente il risultato dell'operato di una o più persone che agiscono per il tramite di un *socialbot* (anche qualora tale azione avvenga con l'ausilio di un *botmaster*), alle quali saranno imputabili gli effetti generati da tali elementi. In particolare, in questa ipotesi il disegno criminoso volto alla realizzazione della sostituzione di persona su di un *social medium* vede, da una parte, la combinazione delle attività di creazione di un profilo che mostri un ritratto (anche generato da un'intelligenza artificiale), delle generalità e delle informazioni che siano diverse da quelle di colui che ne ha le credenziali di accesso e, dall'altro, delle attività di interazione, quali inviare messaggi, apporre reazioni ai contenuti di terzi e pubblicarne di propri, compiute dal *socialbot* su indicazione del proprio amministratore. È quest'ultimo soggetto a fornire all'agente *software* la chiave di accesso (le credenziali) da utilizzare per accedere all'*account* creato e ad indicargli quali attività svolgere sulla piattaforma o con quali destinatari interagire.

Dall'inquadramento appena compiuto dei complessi mezzi fraudolenti utilizzabili per nascondere la vera identità di colui che gestisce – anche indirettamente tramite l'uso di un *socialbot* – un profilo su di un *social medium* nasce, però, un ulteriore problema: quando si possa parlare di effettiva induzione in errore di soggetti terzi non destinatari delle interazioni degli agenti *software* autonomi. Se, infatti, un'interazione continuata nel tempo tra un *socialbot* (che, si ribadisce, è sempre sotto il controllo di chiunque ne abbia i privilegi di amministratore) e una persona, che induca quest'ultima a credere di star intrattenendo un qualche tipo di relazione (amicale o addirittura sentimentale),[54] può facilmente rientrare nell'alveo dell'induzione in errore, altre situazioni sono meno chiare. La prima situazione che può sollevare delle problematiche interpretative è legata alla strutturazione stessa delle piattaforme di *social networking* e alle modalità con cui esse diffondono i contenuti degli utenti, se questi sono oggetto di apposizione di reazioni

[53] Cass. pen., sez. V, 16 giugno 2014, n. 25774, che riprende quanto statuito in Cass. pen., sez. V, 14 dicembre 2007, n. 46674.

[54] Grazie all'implementazione di forme di IA questa situazione non è affatto un'ipotesi ma una situazione reale e ormai ben documentata. Si vedano a questo riguardo, ad esempio, J. Zitser, *A tech bro used AI to create a virtual version of his girlfriend*, in *Insider.com*, 1 giugno 2023, e A. Sternlicht, *A 23-year-old Snapchat influencer used OpenAI's technology to create an A.I. version of herself that will be your girlfriend for $1 per minute*, in *Fortune.com*, 9 maggio 2023.

da parte di altri. Infatti, come ben illustrato dalla Corte di cassazione, gli algoritmi delle piattaforme *social* sono – nella maggior parte dei casi – volti a mostrare agli utenti un «continuo aggiornamento delle notizie e delle attività sviluppate dai [loro] contatti» e tale modalità di informazione è «condizionata dal maggior numero di interazioni che riceve ogni singolo messaggio».[55] I sistemi dei *social media*, quindi, si limitano a mostrare agli utenti terzi rispetto a questa prima interazione i contenuti dell'autore originale corredati dall'indicazione dell'avvenuta apposizione di reazioni ad essi. Il problema interpretativo che sorge dalla situazione appena descritta è che chi ha i privilegi di amministratore di un *socialbot* e le credenziali di accesso a dei profili *social* risulta in grado di indurre in errore degli utenti terzi rispetto all'interazione iniziale senza interagire direttamente con loro. Vi è addirittura chi potrebbe provare a ipotizzare che questi terzi soggetti, non essendo mai stati i destinatari della interazione iniziale, siano stati, al massimo, colposamente indotti in errore.[56]

Questa prima situazione di incertezza può essere facilmente superata considerando sotto una luce leggermente diversa la medesima metodologia di diffusione di contenuti che viene adottata dai gestori delle piattaforme di *social networking*. È lo stesso principio di aggiornamento continuo ed automatico del *newsfeed* mostrato agli utenti, ossia quello di mostrare loro i contenuti che ricevono il maggior numero di interazioni, a fungere da base di appoggio per un utilizzo efficace dei mezzi fraudolenti per la realizzazione della sostituzione di persona nei *social network*. Data, quindi, la basilare strutturazione dei *social media* secondo questo principio, l'apposizione di una reazione ha una intrinseca funzione propalatrice, come già riconosciuto dalla Corte di cassazione,[57] nei confronti di soggetti terzi che, di conseguenza, sono destinatari voluti e ricercati di tale interazione. In altre parole, con la promozione artificiale dei contenuti del destinatario delle interazioni dei *socialbot* – che viene compiuta per il tramite di uno o più profili *social* caratterizzati da qualità (di altri o inventate) rilevanti agli occhi della legge – il loro amministratore sfrutta consapevolmente la strutturazione delle piattaforme per raggiungere il maggior numero possibile di utenti reali e indurli in errore circa la reale identità dell'utilizzatore di ciascuno dei profili che ha automatizzato. Il fine ultimo di questa operazione può essere molto vario e può consistere sia in un vantaggio per l'amministratore del *socialbot* che per altre persone come in un danno per altri ma la finalità intermedia di essa, almeno in questo scenario, rimane sempre quella di dare a utenti terzi una falsa impressione

[55] Cass. pen., sez. I, 9 febbraio 2022, n. 4534.

[56] Si pensi, ad esempio, ad una situazione in cui Tizio abbia il controllo di un *socialbot* e di numerosi profili falsi su Facebook e disponga che il suo agente *software* metta dei "Mi piace" con tutti i profili che gestisce ad un *post* pubblicato da Caio, il quale è completamente ignaro dei sotterfugi di Tizio. Dato che gli algoritmi della piattaforma *social* sono progettati in modo da "rilanciare" i contenuti popolari, il *post* di Caio – che ha ricevuto diversi "Mi piace" dai profili automatizzati di Tizio – verrebbe mostrato dalla piattaforma nel *newsfeed* di Sempronio, anche lui totalmente ignaro della natura sintetica dei profili che ne hanno artificialmente accresciuto la popolarità. In questo caso Tizio – attraverso il suo *socialbot* – non ha mai interagito direttamente con Sempronio ma quest'ultimo ha, comunque, avuto l'impressione che un alto numero di persone abbia espresso il proprio apprezzamento per il *post* di Caio.

[57] Cass. pen., sez. V, 12 dicembre 2017, n. 55418. La decisione presa dalla Corte era relativa alla questione se la pubblicazione sul proprio profilo Facebook di video inneggianti allo Stato Islamico e l'apposizione di "Mi piace" ad altri integrassero o meno il delitto di istigazione a delinquere previsto dall'art. 414 c.p.

di popolarità dei messaggi veicolati dai contenuti oggetto di interazioni dirette.
Insomma, l'utilizzo dei *socialbot* per dare false impressioni di popolarità si traduce, nei fatti, nell'induzione in errore di un alto numero dei destinatari degli aggiornamenti automatici del loro *newsfeed* circa le reali generalità dell'utilizzatore di numerosi profili attivi sulla piattaforma di *social networking*. L'effetto è, così, quello di ledere la pubblica fede in merito all'effettiva identità degli utenti dei *social media*. In aggiunta, tale impiego viene fatto dall'amministratore di questi agenti *software* autonomi con la finalità di procurare a sé o ad altri un vantaggio (come nel caso della promozione di una determinata ideologia) o di recare ad altri un danno (come nell'ipotesi della diffamazione aggravata). Pare, quindi, difficile negare che venga così integrato il reato di sostituzione di persona attraverso l'impiego di *socialbot*.
Ancora più complicata dal punto di vista interpretativo è la situazione in cui non vi siano delle "reazioni" apposte ai contenuti di un utente (come i "Mi piace" di Facebook o i "Consiglia" di LinkedIn) ma solo delle condotte più "passive", come il numero di visualizzazioni ad un contenuto video o il numero di ascolti di un contenuto audio.[58] Tale scenario, già egregiamente rilevato da Weissmann,[59] presenta rispetto a quello precedente l'ancor più complessa problematica del fatto che la maggior parte dei sistemi di conteggio delle visualizzazioni o degli ascolti di contenuti multimediali non mostra l'identità di coloro che li hanno guardati o ascoltati, non dando così alcuna indicazione ad altri utenti di chi abbia interagito con essi. Per di più, queste azioni non riguardano condotte immediatamente riconducibili a forme di consenso e di adesione al contenuto ma, invece, afferiscono solamente al livello di fama raggiunto da esso (la quale ben potrebbe essere caratterizzata da una comune disapprovazione). Questo scenario sembra posizionare l'uso di profili falsi o di *socialbot* per incrementare artificialmente il conteggio del consumo "passivo" di contenuti multimediali al di fuori dei confini del reato di sostituzione di persona. Questo perché, data la mancata spendita di alcuna forma di generalità – di altri od inventate – verso delle altre persone, non sembra possibile argomentare che queste ultime siano state indotte in errore tramite l'utilizzo illegittimo di qualità a cui la legge attribuisce effetti giuridici.
Vi è comunque da notare come, quando l'incremento artificiale degli indici di popolarità di contenuti sui *social media* sia volto ad ingannare qualcuno per ottenere un guadagno e causi al contempo un danno economico alla vittima, sarà sempre ipotizzabile una forma di tutela di quest'ultima tramite previsioni diverse da quelle di tutela della fede pubblica (come l'art. 640 c.p. relativo alla truffa).[60] Tuttavia, quando gli artifizi ed i raggiri

[58] Si vedano la descrizione e le considerazioni in merito a questa forma di interazione fatte *supra* alla lettera (C) del paragrafo 1.1. di questo articolo.

[59] S. Weissmann, *How Not to Regulate Social Media*, cit., 58 ss., dove l'autrice fa particolare riferimento all'uso di *socialbot* fatto da agenti della disinformazione russa per incrementare il numero di visualizzazioni di contenuti video e indurre società di pubblicità a spendere milioni di dollari in annunci ad essi collegati.

[60] Si pensi, ad esempio, ad un *influencer* che si "compri" dei *follower* fittizi. Ossia, un creatore di contenuti video che utilizzi dei *socialbot* per incrementare artificialmente il numero di visualizzazioni dei suoi prodotti, dando in questo modo l'impressione di essere più conosciuto e popolare tra le persone di quanto non sia in realtà e riuscendo, grazie a questi artifizi, ad ottenere delle lucrose sponsorizzazioni da parte di imprese che altrimenti non avrebbero mai sottoscritto un contratto con lui. In questo caso l'utilizzo di agenti *software* autonomi è finalizzato a procurarsi un ingiusto profitto con contestuale danno economico per l'impresa sponsorizzatrice, finendo col configurare almeno una truffa.

consistano nella creazione di false identità online e siano finalizzati al soddisfacimento di una propria vanità[61] o a rendere false impressioni di popolarità di opinioni o idee non necessariamente estremiste o discriminatorie (finalità che poco si discostano da forme di vanità), ripiegare sull'applicazione del solo reato di frode informatica (previsto dall'art. 640-*ter* c.p.) non sembra – almeno agli occhi di chi scrive – fornire la giusta tutela ai beni giuridici aggredibili da tali condotte. Infatti, in questo secondo caso si ha, comunque, la creazione di una o più false identità digitali che vengono utilizzate per dare ad un vasto pubblico l'impressione che venga prestata ampia attenzione alle posizioni veicolate dai contenuti pubblicati *online*, ossia per indurre in errore la fede pubblica circa l'effettivo numero di persone che abbiano guardato od ascoltato dei contenuti multimediali. Insomma, anche in assenza della spendita di un'identità fittizia o di terzi, l'utilizzo di *socialbot* per incrementare il numero di reazioni "passive", come le visualizzazioni di contenuti multimediali, finisce per indurre in errore la fede pubblica in merito ad un dato: l'esistenza di un ampio numero di persone interessate a detti contenuti.

L'insistenza sulla rilevanza di condotte che possono inizialmente sembrare secondarie dipende dal fatto che, come mostrato da alcuni importanti studi,[62] la formazione delle opinioni degli individui è influenzata dalle reti sociali di cui fanno parte e costoro tendono a fidarsi delle informazioni che circolano al loro interno[63] e a considerare come veritiere le opinioni che sono ampiamente diffuse.[64] Alla luce di queste considerazioni non si può, quindi, giudicare come irrilevante la malevola diffusione potenziata di contenuti sui *social media*, dal momento che essa è in grado di condizionare il comportamento o la psicologia di un vasto pubblico. In altre parole, la diffusione artificialmente aumentata di contenuti è una condotta che, dati gli innati meccanismi di inconscia e tendenziale adesione ai messaggi ampiamente diffusi nel proprio contesto sociale, già di per sé contiene quel *quid pluris* che la rende una forma di divulgazione di opinioni finalizzata ad influenzare il comportamento o la psicologia di un vasto pubblico e a raccogliere adesioni.[65]

---

[61] In merito alla rilevanza penale del perseguimento di vantaggi non patrimoniali nel reato di sostituzione di persona si veda, ad esempio, Cass. pen., sez. V, 16 giugno 2014, n. 25774.

[62] P.F. Lazarsfeld - B. Berelson - H. Gaudet, *The People's Choice. How the Voter Makes Up His Mind in a Presidential Campaign*, New York, 1944; B. Berelson - P.F. Lazarsfeld - W.N. McPhee, *Voting: a study of opinion formation in a presidential campaign*, Chicago, 1954; E. Katz - P.F. Lazarsfeld, *Personal influence: The part played by people in the flow of mass communications*, Glencoe, 1955; M.E.J. Newman, *Networks: An Introduction*, New York, 2010; G. Caldarelli, *Scale-Free Networks: Complex Webs in Nature and Technology*, Oxford, 2010, 211 ss.

[63] Y. Jun - R. Meng - G.V. Johar, *Perceived social presence reduces fact-checking*, in *Proceedings of the National Academy of Sciences*, 2017, 5976 ss.; K.C. Yang - O. Varol - C.A. Davis - E. Ferrara - A. Flammini - F. Menczer, *Arming the public with artificial intelligence to counter social bots*, in *Human Behavior and Emerging Technologies*, 2019, 48 ss.; H. Wolters - K. Stricklin - N. Carey - M.K. McBride, *The Psychology of (Dis) information: A Primer on Key Psychological Mechanisms*, in *CNA Research Memorandum*, 2021, 1 ss.

[64] M. Luo - J.T. Hancock - D.M. Markowitz, *Credibility perceptions and detection accuracy of fake news headlines on social media: Effects of truth-bias and endorsement cues*, in *Communication Research*, 2022, 171 ss., in cui gli autori hanno messo in luce come su Facebook contenuti con un elevato numero di "Mi piace" vengano percepiti come maggiormente credibili.

[65] In merito alla natura della "propaganda di idee" rilevante agli occhi della legge si vedano Cass. pen., sez. I, 9 febbraio 2022, n. 4534 e Cass. pen., sez. V, 22 luglio 2019, n. 32862. Invero, le sentenze citate sono relative al reato di propaganda e istigazione a delinquere per motivi di discriminazione razziale, etnica e religiosa (descritto all'art. 604-*bis* c.p.) e – si riconosce – lo spettro di un'applicazione analogica

Insomma, l'uso di *socialbot* per aumentare la portata della diffusione di contenuti online, sfruttando la conformazione degli algoritmi dei *newsfeed* delle piattaforme *social*, deve essere visto come una forma di induzione in errore di un vasto numero di persone finalizzata a procurare dei vantaggi – anche di natura non patrimoniale – che ha l'effetto di trarre in inganno la fede pubblica su aspetti che, come mostrato da autorevoli e consolidati studi, condizionano la realtà sociale.

Quindi, sebbene l'effetto di tale utilizzo dei *socialbot* possa risolversi in un riorientamento marginale delle posizioni delle persone da esso investite, non può essere ignorato come questo malevolo impiego sia caratterizzato da una capacità di influenza non nulla sui molteplici destinatari – anche indiretti – delle interazioni con i contenuti pubblicati sulle piattaforme di *social networking* e come tale ascendente sia raggiunto tramite l'induzione in errore della fede pubblica sull'esistenza di un ampio numero di persone.

(C) Particolare tenuità del fatto – Non può essere tralasciato il fatto che, a minaccia della tesi della necessità di un'applicazione estensiva delle tutele penali contro gli atti di induzione in errore della fede pubblica appena illustrata, la Corte di Cassazione abbia già in passato ritenuto che alcune concrete condotte di sostituzione di persona possano non richiedere una sanzione. In particolare, in una loro recente sentenza, i giudici della Suprema corte hanno statuito che l'aver utilizzato per la creazione di un profilo online le generalità di un'altra persona del tutto inconsapevole possa non essere punito per effetto delle previsioni dell'art. 131-*bis* del codice penale, ossia quando le peculiarità della fattispecie concreta – le modalità della condotta, il grado di colpevolezza da esse desumibile e l'entità del danno o del pericolo – ammantino il fatto di particolare tenuità.[66]

Addirittura, gli Ermellini hanno ritenuto che il grado di offensività delle condotte valutate fosse particolarmente tenue, nonostante al cosiddetto "furto di identità" fossero seguite ulteriori condotte illecite (nel caso di specie delle molestie integranti il reato di cui all'art. 660 c.p.). In altre parole, essi hanno ritenuto che, nonostante la plurima offensività della condotta tenuta in concreto, l'aggressione ai beni giuridici compiuta non fosse caratterizzata da un grado di dannosità rilevante. In aggiunta, i giudici di piazza Cavour hanno stabilito che le osservazioni del Procuratore Generale presso la Corte di appello di Milano circa la concreta portata offensiva del delitto di sostituzione di persona facessero leva su «un approccio astrattizzante, non in linea con il necessario ancoraggio alla fattispecie concreta che connota la causa di non punibilità» e ne hanno così, appunto, escluso la punibilità.[67]

Quindi, non sembra neanche irrealistico ipotizzare che la Suprema Corte possa giungere alle medesime conclusioni per un caso simile a quello descritto da Cresci, Di Pietro,

sembra aleggiare sopra al ragionamento qui sviluppato. Tuttavia, come visto, la condotta dell'illegittimo utilizzo di *socialbot* possiede un'alta intensità offensiva, dato che non si limita ad indurre in errore degli utenti di un *social medium* circa il numero – e, quindi, le identità – di persone che diffondono contenuti sulla piattaforma ma è anche in grado di influenzare il comportamento o la psicologia di questi.

[66] Cass. pen., sez. V, 10 gennaio 2020, n. 652, in cui i membri del collegio giudicante hanno ritenuto che il carattere del tutto isolato dell'episodio di creazione di un profilo online utilizzando le generalità della vittima fosse causa di esclusione della punibilità per particolare tenuità del fatto.

[67] Pur riconoscendo che «l'avvalersi di una piattaforma multimediale accessibile ad un numero indeterminato di utilizzatori comporta un danno importante e astrattamente senza termini nel tempo e nello spazio».

Petrocchi, Spognardi e Tesconi,[68] ossia quando una persona denunci di essere stata vittima di un furto di identità sui *social* consistente nell'utilizzo di un suo ritratto per la creazione di un profilo avente altre generalità (che non ne permettano la sua identificazione) e che questo venga impiegato solamente per ri-condividere su una piattaforma *social* i contenuti pubblicati da un candidato politico. Infatti, in quest'ultima ipotesi all'indebito utilizzo di una immagine della vittima non farebbe seguito alcun altro illecito comportamento diretto nei suoi confronti (come avvenuto, invece, nel caso del 2020 ritenuto di particolare tenuità dai giudici di Cassazione).

non pare irrealistico immaginare che la medesima considerazione del valore particolarmente tenue del crimine possa essere applicata anche ad attività di propaganda o di cosiddetto *astroturfing* potenziate dall'uso di *socialbot*.[69] Dopotutto, la creazione di profili *social* tramite la completa invenzione delle generalità degli stessi e l'utilizzo di ritratti di persone inesistenti priverebbe la condotta di qualsiasi potenziale lesivo del diritto all'identità personale, all'immagine o alla dignità di qualcuno, diminuendo sensibilmente il senso di urgenza di un intervento di tutela giurisdizionale. Permarrebbero, comunque, il danno alla pubblica fede e il fine di «procurare a sé o ad altri un vantaggio» di natura non patrimoniale ma, in assenza di una chiara e condivisa illustrazione che espliciti come detto bene giuridico debba essere difeso anche da artificiosi incrementi della diffusione di contenuti e da inganni volti a dare impressioni di popolarità in realtà inesistente, non pare potersi escludere il rischio che strategie di propaganda volte a deviare il naturale percorso di formazione delle opinioni vengano erroneamente viste come prive di particolare gravità.

4. Conclusioni

La presenza dei *socialbot* all'interno dei *social network* è allo stato attuale una caratteristica inestirpabile di queste piattaforme. Sebbene da alcuni possa non essere vista come un problema rilevante per l'ordinamento di uno Stato, si fa sempre più forte la consapevolezza che questi agenti *software* rappresentano, invece, un rischio per la stabilità delle

[68] S. Cresci - R. Di Pietro - M. Petrocchi - A. Spognardi - M. Tesconi, *Social fingerprinting: detection of spambot groups through DNA-inspired behavioral modeling*, cit., 561 ss.

[69] L'*astroturfing* è la pratica di nascondere l'identità del vero patrocinatore di un determinato messaggio tramite l'organizzazione di false campagne "nate dal basso" che diano l'impressione che vi sia una vasta pletora di persone – e non solo alcune imprese commerciali o dei ristretti gruppi di interesse politico od economico – a sostenerlo. Questa attività si è tradotta su internet e sui *social media* nella creazione di reti di profili che promuovono e condividono incessantemente contenuti relativi ad un determinato argomento ed orientati verso il sostegno di una precisa posizione al riguardo. In merito a questa pratica si vedano, J. Ratkiewicz - M. Conover - M. Meiss - B. Gonçalves - S. Patil - A. Flammini - F. Menczer, *Truthy: Mapping the spread of astroturf in microblog streams*, in *WWW '11: Proceedings of the 20th International Conference Companion on World Wide Web*, 2011, 249 ss.; J. Zhang - D. Carpenter - M. Ko, *Online astroturfing: A theoretical perspective*, in *Proceedings of the Nineteenth Americas Conference on Information Systems, Chicago, Illinois, August 15-17, 2013*, 2013, 1 ss.; S.C. Woolley, *Automating power: Social bot interference in global politics*, in *First Monday*, 21, 2016; M. Kovic - A. Rauchfleisch - M. Sele - C. Caspar, *Digital astroturfing in politics: Definition, typology, and countermeasures*, in *Studies in Communication Sciences*, 18(1), 2018, 69 ss.; E. Dubois - F. McKelvey, *Political Bots: Disrupting Canada's Democracy*, in *Canadian Journal of Communication*, 44(2), 2019, 27 ss.; F.B. Keller - D. Schoch - S. Stier - J. Yang, *Political Astroturfing on Twitter: How to Coordinate a Disinformation Campaign*, in *Political Communication*, 37(2), 2020, 256 ss.

istituzioni democratiche di una nazione.[70] Ben facilmente, infatti, essi possono essere utilizzati per potenziare la diffusione di discorsi incitanti all'odio e alla discriminazione, di false informazioni su fatti, qualità personali ed eventi e per accrescere artificiosamente le impressioni di popolarità di certe persone od opinioni.[71]

In ambito europeo sono attualmente in corso delle azioni legislative volte a contrastare l'impiego dei *socialbot*, dato che essi possono essere impiegati per «la manipolazione intenzionale e spesso coordinata del servizio della piattaforma [*social*], con effetti prevedibili sulla salute pubblica, sul dibattito civico, sui processi elettorali, sulla sicurezza pubblica e sulla tutela dei minori».[72]

In attesa che anche in Italia vi sia una presa di coscienza della pericolosità delle pratiche di propaganda computazionale (tra le quali rientra l'impiego dei *socialbot*) e ci si renda conto che sono ormai necessari degli interventi normativi, dato che è già stato evidenziato come anche il nostro panorama digitale sia piagato da questo fenomeno,[73] ci si deve rivolgere alla legislazione già in vigore per tutelare quei beni giuridici che sono minacciati da questa tecnologia. In altra sede si è già avuto modo di evidenziare come l'impiego di *socialbot* per danneggiare la reputazione delle persone costituisca una forma aggravata del reato di diffamazione.[74] Qui, invece, si è voluto concentrarsi sull'applicazione della normativa esistente relativa al reato di sostituzione di persona per la tutela di un altro bene giuridico di estrema rilevanza che viene aggredito dal malevolo impiego dei *socialbot*: quello della pubblica fede.

Al fine di evitare fraintendimenti in merito all'individuazione del responsabile degli eventi causati da questa particolare categoria di agenti *software* autonomi, si è sottolineato come essi non siano intelligenti, capaci di intendere e volere e neppure realmente autonomi. Essi sono solamente dei sofisticati strumenti digitali sottoposti al controllo di colui che ne ha i privilegi di amministratore. Ne consegue che l'attribuzione di un'identità (di terzi o inventata), di qualità o di stati a cui la legge attribuisce effetti giuridici non avviene in capo a questi beni immateriali ma solo ai loro utilizzatori (i loro amministratori), che finiscono col commettere una sostituzione di persona.

Nell'analisi sull'integrazione di questo crimine tramite l'uso dei *socialbot* occupano una posizione centrale il bene giuridico tutelato, ossia la fede pubblica, e il suo momento

---

[70] Si vedano a questo riguardo le preoccupazioni espresse dalle istituzioni europee, poi confluite nelle considerazioni del COM(2020) 825 final, *Proposta di regolamento del Parlamento europeo e del Consiglio relativo a un mercato unico dei servizi digitali (legge sui servizi digitali)*. In particolare, al Considerando 57 i *socialbot* vengono inseriti tra i «rischi sistemici» che minacciano la tenuta delle istituzioni dell'Unione e dei suoi Stati membri.

[71] Si vedano, ad esempio, le considerazioni contenute in COM(2018) 236 final, *comunicazione della Commissione al Parlamento europeo, al Consiglio, al Comitato economico e sociale europeo e al Comitato delle Regioni, Contrastare la disinformazione online: un approccio europeo*.

[72] Considerando 57 della COM(2020) 825.

[73] A. Vogt, *Hot or bot? Italian professor casts doubt on politician's Twitter popularity*, in *The Guardian*, 22 luglio 2012; S. Cresci - R. Di Pietro - M. Petrocchi - A. Spognardi - M. Tesconi, *DNA-inspired online behavioural modelling and its application to spambot detection*, in *IEEE Intelligent Systems*, 31(5), 2016, 58 ss.; Id., *The paradigm-shift of social spambots: Evidence, theories, and tools for the arms race*, cit., 963 ss.; Id., *Exploiting digital DNA for the analysis of similarities in Twitter behaviours*, in *2017 IEEE International Conference on Data Science and Advanced Analytics (DSAA)*, 2017, 686 ss.; Id., *Social fingerprinting: detection of spambot groups through DNA-inspired behavioral modeling*, cit., 561 ss.

[74] A. Tedeschi Toschi - G. Berni Ferretti, *Social media, profili artificiali e tutela della reputazione*, cit., 107 ss.

consumativo, cioè l'induzione in errore di altre persone. Questo perché questi agenti *software*, quando dotati di numerosi profili *social*, non sono solo in grado di indurre in errore le persone sulla reale identità del loro utilizzatore ma possono addirittura condizionare il comportamento o la psicologia di un vasto pubblico, a causa della naturale tendenza inconscia delle persone ad aderire alle posizioni ampiamente diffuse nel proprio contesto sociale (anche in quelli online).[75] Sebbene gli effetti di campagne propagandistiche volte ad influenzare l'opinione delle persone tramite false impressioni di popolarità siano difficili da misurare e possano risultare marginali, si è visto come essi siano comunque in grado di generare risultati rilevanti per le istituzioni democratiche e la tenuta di interi sistemi sovranazionali.[76]

In ogni caso, l'impiego dei *socialbot* comporta la lesione della pubblica fede circa la reale identità dei soggetti attivi sui *social media* (i quali, usano questi agenti *software* come strumenti per automatizzare l'attribuzione a sé di altre identità) e, a seconda delle concrete modalità d'azione, finisce per aggredire altri beni giuridici meritevoli di tutela, quali l'identità personale e l'onore (nel caso che i profili *social* gestiti siano il frutto di un cosiddetto "furto d'identità") ovvero la libertà di ricevere informazioni o idee senza ingerenze e condizionamenti (nel caso di campagne di propaganda computazionale o di *astroturfing*). Quindi, nel contesto normativo vigente nel nostro paese si è individuato nella normativa di contrasto al reato di sostituzione di persona un valido strumento di tutela di questi beni giuridici, dal momento che «oggetto della tutela penale è l'interesse riguardante la pubblica fede […] siccome si tratta di inganni che possono superare la ristretta cerchia d'un determinato destinatario».[77]

---

[75] P.F. Lazarsfeld - B. Berelson - H. Gaudet, *The People's Choice. How the Voter Makes Up His Mind in a Presidential Campaign*, cit.; B. Berelson - P.F. Lazarsfeld - W.N. McPhee, *Voting: a study of opinion formation in a presidential campaign*, cit.; E. Katz - P.F. Lazarsfeld, *Personal influence: The part played by people in the flow of mass communications*, cit.; M.E.J. Newman, *Networks: An Introduction*, cit.; G. Caldarelli, *Scale-Free Networks: Complex Webs in Nature and Technology*, cit.; Y. Jun - R. Meng - G.V. Johar, *Perceived social presence reduces fact-checking*, cit., 5976 ss.; K.C. Yang - O. Varol - C.A. Davis - E. Ferrara - A. Flammini - F. Menczer, *Arming the public with artificial intelligence to counter social bots*, cit., 48 ss.; H. Wolters - K. Stricklin - N. Carey - M.K. McBride, *The Psychology of (Dis)information: A Primer on Key Psychological Mechanisms*, cit., 1 ss.

[76] M. Hänska - S. Bauchowitz, *Tweeting for Brexit: how social media influenced the referendum*, in J. Mair - T. Clark - N. Fowler - R. Snoddy - R. Tait (a cura di), *Brexit, Trump and the Media*, Bury St Edmunds, 2017, 31 ss.; N. Persily, *The 2016 US Election: Can democracy survive the Internet?*", in *Journal of democracy*, 28(2), 2017, 63 ss.; J.A. Tucker - Y. Theocharis - M.E. Roberts - P. Barberá, *From liberation to turmoil: Social media and democracy*", in *Journal of democracy*, 28, 2017, 46 ss.; W. Hall - R. Tinati - W. Jennings, *From Brexit to Trump: Social media's role in democracy*, in *Computer*, 51(1), 2018, 18 ss.; S. Sanovich - D. Stukal - J.A. Tucker, *Turning the virtual tables: Government strategies for addressing online opposition with an application to Russia*, in *Comparative Politics*, 50(3), 2018, 435 ss.; Y. Gorodnichenko - T. Pham - O. Talavera, *Social media, sentiment and public opinions: Evidence from #Brexit and #USElection*, in *European Economic Review*, 136, 2021, 103772 ss.

[77] Cass. pen., sez. V, 16 giugno 2014, n. 25774, che riprende quanto statuito precedentemente in Cass. pen., sez. V, 14 dicembre 2007, n. 46674.

Delega di funzioni regolamentari e lotta ai rischi sistemici causati della disinformazione nel *Digital Services Act*: quali rischi per la libertà di espressione?*

Andrea Palumbo – Jacopo Piemonte

Abstract

Il saggio mira a esaminare fino a che punto il modello di co-regolamentazione adottato dal regolamento (UE) 2022/2065 per la mitigazione dei rischi sistemici causati dalla disinformazione possa assicurare il rispetto della libertà di espressione degli utenti dei servizi di fornitori di piattaforme e motori di ricerca online di dimensioni molto grandi. Le rilevanti disposizioni del regolamento sono analizzate sotto la lente della Carta dei diritti fondamentali dell'Unione europea. Nella prima parte è oggetto di analisi la compatibilità di questo modello regolamentare con il principio di legalità sancito dalla Carta, mentre la seconda parte si concentra sui rischi che le restrizioni disposte ai sensi del regolamento (UE) 2022/2065 per la mitigazione dei rischi sistemici non siano in linea con il principio di proporzionalità. Le criticità identificate nello scritto intendono offrire spunti di riflessione sull'idoneità della co-regolamentazione per disciplinare attività che implicano la moderazione dei contenuti di utenti online.

This contribution aims to examine whether the co-regulation model adopted in regulation (EU) 2022/2065 to govern the mitigation of systemic risks caused by disinformation is compatible with the protection of users' freedom of expression on very large online platforms and search engines. The relevant provisions of the Regulation are evaluated in light of the Charter of fundamental rights of the European Union. First, it will be assessed if the co-regulatory model of the Regulation is respectful of the legality principle of the Charter. Second, the analysis moves on to the risk that the mitigation of systemic risks generated by disinformation may lead to disproportionate interferences with the right to freedom of expression. The identified shortcomings should provide points of reflection on whether co-regulation is fit for purpose when it comes to the moderation of users' online content.

* L'articolo è stato sottoposto, in conformità al regolamento della Rivista, a referaggio "a doppio cieco".

Sommario

1. Considerazioni introduttive. - 2. Il nuovo paradigma regolamentare del *DSA* ed il ruolo della co-regolamentazione: esiste un problema per la tutela della libertà di espressione? - 3. Restrizioni della libertà di espressione derivanti dall'attuazione dell'art. 35 del *DSA* e del Codice. - 4. Delega di funzioni regolamentari e compatibilità con il principio di legalità: l'assetto co-regolamentare del *DSA* alla prova. - 4.1. Il principio di legalità nella Carta. - 4.2. La compatibilità con il principio di legalità dell'art. 35 del *DSA*. - 5. Salvaguardie per un'applicazione proporzionata dell'art. 35 del *DSA*, rispetto della libertà di espressione in concreto? - 5.1. Le salvaguardie del *DSA* per il rispetto del principio di proporzionalità nella moderazione dei contenuti – 5.2. Mettendo in pratica il *DSA*: fino a che punto è ragionevole attendersi il rispetto del principio di proporzionalità? - 5.3. (Segue) I rimedi giudiziali e stragiudiziali nel *DSA*. - 5.3.1. Il primo livello: meccanismi di reclamo e ricorso. - 5.3.2. Il secondo livello: la risoluzione stragiudiziale delle controversie. - 5.3.3. Il terzo livello: il ricorso agli organi giurisdizionali. - 5.4. (Segue) La supervisione della Commissione. - 5.4.1. I meccanismi di controllo sull'operato di *VLOPs* e *VLOSEs* nel *DSA*. - 5.4.2. Luci e ombre del ruolo di vigilanza della Commissione. - 6. Conclusioni.

Keywords

Carta dei diritti fondamentali dell'Unione europea - disinformazione - libertà di espressione e di informazione - moderazione dei contenuti - *Digital Services Act*

1. Considerazioni introduttive

A decorrere dal 17 febbraio 2024, comincerà a trovare applicazione il regolamento (UE) 2022/2065 (regolamento sui servizi digitali)[1], meglio conosciuto con il suo nome inglese *Digital Services Act* (di seguito, il "*DSA*"). Con il *DSA* il legislatore europeo ha introdotto un nuovo paradigma regolamentare per i rischi posti dai fornitori di servizi intermediari di semplice trasporto, memorizzazione temporanea e memorizzazione di informazioni, concentrandosi sulla disciplina dei servizi di memorizzazione di informazioni (c.d. *hosting*). Tra le altre cose, il *DSA* ha introdotto una disciplina *ad hoc* per i fornitori di piattaforme online di dimensioni molto grandi e dei motori di ricerca online di dimensioni molto grandi (in inglese, rispettivamente, *very large online platforms* e *very large online search engines*, di seguito "*VLOPs*" e "*VLOSEs*" e, congiuntamente, gli "intermediari")[2], a cui sono state attribuite importanti funzioni di auto-regolamentazione rispetto all'adozione delle misure di attenuazione dei rischi sistemici causati dai loro servizi, quali ad esempio la diffusione di contenuti illegali e dannosi online. Viene

[1] Regolamento (UE) 2022/2065 del Parlamento europeo e del Consiglio del 19 ottobre 2022 relativo a un mercato unico dei servizi digitali e che modifica la direttiva 2000/31/CE (regolamento sui servizi digitali).

[2] Le piattaforme online e i motori di ricerca online sono designati, rispettivamente, come *VLOPs* e *VLOSEs* sulla base dei criteri stabiliti dall'art. 33 del *DSA*.

dunque attribuita a tali intermediari ampia discrezionalità nell'implementazione del *DSA*.

Questo nuovo sistema potrebbe comportare, con ogni probabilità, la limitazione dell'esercizio della libertà di espressione per i destinatari dei servizi di *VLOPs* e *VLOSEs*[3] (di seguito, gli "utenti") in determinate situazioni (si pensi, ad esempio, alla possibilità per gli intermediari di rimuovere contenuti di utenti che considerino perniciosi). Per questa ragione, l'implementazione di questo sistema dovrebbe avvenire nel pieno rispetto delle condizioni richieste dal diritto dell'Unione per limitare l'esercizio di diritti fondamentali.

Questo articolo si concentrerà sull'analisi delle criticità che si presentano, per la libertà di espressione degli utenti, in conseguenza della delega di competenze regolamentari che il *DSA* opera verso *VLOPs* e *VLOSEs* per l'attenuazione di rischi sistemici. Al fine di fornire considerazioni più precise sulle ricadute applicative di questo approccio regolamentare, l'analisi prenderà ad esempio il caso dell'attenuazione dei rischi sistemici causati dalla disseminazione di contenuti qualificati come disinformazione. L'esempio della disinformazione è stato selezionato in virtù del fatto che ad oggi è già stato adottato un codice di condotta da parte di alcuni *VLOPs* e *VLOSEs* per adempiere, tra le altre cose, agli obblighi di attenuazione dei rischi del *DSA*. Pertanto, i paragrafi che seguono forniranno sia considerazioni generali sulle problematiche derivanti dalla delega di funzioni regolamentari, sia un'analisi più concreta e precisa dei risvolti pratici del funzionamento di alcune disposizioni del *DSA* nello specifico caso della lotta alla disinformazione e dell'utilizzo del codice di buone pratiche sulla disinformazione del 2022 (di seguito, il "Codice")[4].

2. Il nuovo paradigma regolamentare del *DSA* ed il ruolo della co-regolamentazione: esiste un problema per la tutela della libertà di espressione?

Il *DSA* costituisce un rivoluzionario sviluppo legislativo per la disciplina della diffusione di contenuti illegali o dannosi online.

Esso potrebbe porsi nel solco della spinta normativa con cui l'Europa si è erta a livello globale come pioniera nella difesa dei diritti fondamentali nel mondo digitale e nella regolamentazione del digitale, influenzando anche legislazioni extra europee. Tale fenomeno, noto come "*Brussels effect*", ha portato in passato numerose legislazioni extra-UE a prendere come ispirazione il *General Data Protection Regulation*[5] per regolare nei rispettivi paesi la tutela dei dati personali[6]. Del pari, anche la proposta di *Artificial*

[3] Per la definizione di "destinatario del servizio" art. 3, lett. *b)*, del *DSA*.

[4] Commissione europea, *The Strengthened Code of Practice on Disinformation 2022*, in *ec.europa.ue*, 16 giugno 2022.

[5] Regolamento (UE) 2016/679 del Parlamento europeo e del Consiglio, del 27 aprile 2016, relativo alla protezione delle persone fisiche con riguardo al trattamento dei dati personali, nonché alla libera circolazione di tali dati e che abroga la direttiva 95/46/CE (regolamento generale sulla protezione dei dati).

[6] A. Bendiek - I. Stuerzer, *The Brussels Effect, European Regulatory Power and Political Capital: Evidence for*

Intelligence Act[7] (in fase avanzata di finalizzazione da parte del Legislatore europeo) potrebbe diventare in futuro uno tra i principali *standard* globali a cui fare riferimento per la regolamentazione dell'intelligenza artificiale[8]. Si è sostenuto che la stessa dinamica possa dunque ripetersi con il modello europeo introdotto dal *DSA*. Seppure le differenze culturali potrebbero portare a maggiori difficoltà nell'"impiantare" tale sistema in paesi come gli Stati Uniti, tale modello si pone quanto meno come primo innovativo *benchmark* legislativo in questo ambito che non si potrà non tenere in considerazione[9]. Ed infatti, nonostante il testo del *DSA* riprenda i principi della direttiva *e-commerce*[10] per la disciplina della responsabilità degli intermediari per i contenuti diffusi online da terzi[11], introduce anche importanti novità.

In primo luogo, il *DSA* presenta un nuovo paradigma regolamentare per la disciplina dei contenuti diffusi tramite i servizi degli intermediari online, passando da un approccio concentrato sulla responsabilità *ex post* degli intermediari ad una regolamentazione *ex ante* delle loro attività[12]. Agli articoli, ereditati dalla direttiva *e-commerce*, sul regime di esenzione dalla responsabilità si aggiunge una corposa normativa su obblighi di trasparenza e di due *diligence*, che si articolano su più livelli con una diversa disciplina a seconda della categoria di intermediari per cui trovano applicazione. La categoria soggetta al maggior numero di disposizioni è quella dei *VLOPs* e *VLOSEs*, per cui il *DSA* trova integralmente applicazione con obblighi supplementari rispetto agli altri intermediari, in considerazione dei maggiori rischi che comportano per la società e la circolazione di informazioni online.

In secondo luogo, il *DSA* impone una serie di obblighi la cui implementazione in concreto deve essere definita dagli intermediari disciplinati, tramite un'autovalutazione del rispetto della normativa da parte degli stessi accompagnata da sistemi di controllo esterno e dalla vigilanza di autorità competenti. Questo approccio lascia ampia discrezionalità ai soggetti obbligati su come implementare il *DSA*, seppur con la previsione di un continuo monitoraggio del loro operato.

Mutually Reinforcing Internal and External Dimensions of the Brussels Effect from the European Digital Policy Debate, in *Digital Society,* 2-5, 2023, 5.

[7] Proposta di regolamento del Parlamento europeo e del Consiglio, del 21 aprile 2021, che stabilisce regole armonizzate sull'intelligenza artificiale (legge sull'intelligenza artificiale) e modifica alcuni atti legislativi dell'Unione.

[8] L. Floridi, *The European Legislation on AI: a Brief Analysis of its Philosophical Approach*, in *Philosophy & Technology,* 34, 2021, 217 ss.

[9] N. Zingales, *The DSA as a Paradigm Shift for Online Intermediaries' Due Diligence: Hail To Meta-Regulation*, in J. Van Hoboken - J. P. Quintais - N. Appelman - R. Fahy - I. Buri - M. Straub (a cura di), *Putting the DSA into practice*, 2022, 222 ss.

[10] Direttiva 2000/31/CE del Parlamento europeo e del Consiglio dell'8 giugno 2000 relativa a taluni aspetti giuridici dei servizi della società dell'informazione, in particolare il commercio elettronico, nel mercato interno (Direttiva sul commercio elettronico).

[11] Si vedano, in particolare, gli artt. 4, 5 e 6 del *DSA* sul regime di responsabilità dei prestatori del servizio di, rispettivamente, semplice trasporto, memorizzazione temporanea e memorizzazione di informazioni.

[12] G. Caggiano, *La proposta di Digital Service Act per la regolazione dei servizi e delle piattaforme online nel diritto dell'Unione europea*, in G. Caggiano - G. Contaldi- P. Manzini, *Verso una legislazione europea su mercati e servizi digitali*, Bari, 2021, 7; M. C. Buiten, *The Digital Services Act: From Intermediary Liability to Platform Regulation*, *Journal of Intellectual Property,* in *Information Technology and E-Commerce Law*, 12, 2022, 361.

L'esempio evidente di questo approccio è rinvenibile nella Sezione del *DSA* dedicata agli obblighi per *VLOPs* e *VLOSEs*[13], ove il legislatore europeo ha attribuito a questi intermediari il compito di agire come auto-regolatori dei rischi sistemici che emergono a causa dell'utilizzo dei loro servizi[14], sebbene sotto la vigilanza della Commissione e prevedendo sistemi di controllo da parte di esperti indipendenti. In particolare, l'art. 34 del *DSA* prevede che *VLOPs* e *VLOSEs* debbano valutare gli eventuali rischi sistemici nell'Unione derivanti «dalla progettazione o dal funzionamento del loro servizio e dei suoi relativi sistemi, compresi i sistemi algoritmici, o dall'uso dei loro servizi»[15]. Questa valutazione deve essere effettuata almeno una volta l'anno, nonché ogniqualvolta è introdotta una funzionalità che possa avere un impatto critico sui rischi sistemici individuati. L'art. 34 del *DSA* non fornisce una lista completa dei rischi sistemici per cui deve essere compiuta la valutazione, ma specifica che, in ogni caso, la valutazione deve comprendere quattro categorie di rischi sistemici, riguardanti la diffusione di contenuti illegali, gli eventuali effetti negativi per l'esercizio dei diritti fondamentali, gli eventuali effetti negativi sul dibattito civico, i processi elettorali e la sicurezza pubblica, e infine gli effetti negativi sulla violenza di genere, la protezione della salute pubblica e dei minori e le gravi conseguenze negative per il benessere fisico e mentale della persona[16].

Sulla base dei rischi così individuati, *VLOPs* e *VLOSE*s sono tenuti, ai sensi dell'art. 35 del *DSA*, ad adottare misure di attenuazione dei rischi che siano ragionevoli, proporzionate ed efficaci[17]. A tal proposito, è possibile evidenziare la presenza di un sistema

[13] Si veda il Capo III, Sezione 5 del *DSA*.

[14] N. Zingales, *The DSA as a Paradigm Shift for Online Intermediaries' Due Diligence: Hail To Meta-Regulation*, cit., 215 ss.

[15] Si veda l'art. 34, par. 1, del *DSA*.

[16] L'art. 34, par. 1, secondo capoverso del *DSA* dispone quanto segue con riferimento agli obblighi di *VLOPs* e *VLOSEs*:
«Essi effettuano le valutazioni dei rischi entro la data di applicazione di cui all'articolo 33, paragrafo 6, secondo comma, e successivamente almeno una volta all'anno, e in ogni caso prima dell'introduzione di funzionalità che possono avere un impatto critico sui rischi individuati a norma del presente articolo. La valutazione del rischio deve essere specifica per i loro servizi e proporzionata ai rischi sistemici, tenendo in considerazione la loro gravità e la loro probabilità, e deve comprendere i seguenti rischi sistemici:
a) la diffusione di contenuti illegali tramite i loro servizi;
b) eventuali effetti negativi, attuali o prevedibili, per l'esercizio dei diritti fondamentali, in particolare i diritti fondamentali alla dignità umana sancito nell'articolo 1 della Carta, al rispetto della vita privata e familiare sancito nell'articolo 7 della Carta, alla tutela dei dati personali sancito nell'articolo 8 della Carta, alla libertà di espressione e di informazione, inclusi la libertà e il pluralismo dei media, sanciti nell'articolo 11 della Carta, e alla non discriminazione sancito nell'articolo 21 della Carta, al rispetto dei diritti del minore sancito nell'articolo 24 della Carta, così come all'elevata tutela dei consumatori, sancito nell'articolo 38 della Carta;
c) eventuali effetti negativi, attuali o prevedibili, sul dibattito civico e sui processi elettorali, nonché sulla sicurezza pubblica;
d) qualsiasi effetto negativo, attuale o prevedibile, in relazione alla violenza di genere, alla protezione della salute pubblica e dei minori e alle gravi conseguenze negative per il benessere fisico e mentale della persona».

[17] L'art. 35, par. 1, primo capoverso dispone quanto segue:
«I fornitori di piattaforme online di dimensioni molto grandi e di motori di ricerca online di dimensioni molto grandi adottano misure di attenuazione ragionevoli, proporzionate ed efficaci, adattate ai rischi sistemici specifici individuati a norma dell'articolo 34, prestando particolare attenzione agli effetti di tali misure sui diritti fondamentali».

di co-regolamentazione, in cui le autorità di vigilanza collaborano con gli intermediari per la definizione delle misure di implementazione dell'art. 35 del *DSA* in concreto. In particolare, il *DSA* contempla, all'art. 45, par. 2[18], la collaborazione tra intermediari e Commissione europea per la definizione all'interno di codici di condotta delle misure di attenuazione dei rischi sistemici identificati a norma dell'art. 34.

L'art. 35 del *DSA* riguarda l'attenuazione di una categoria aperta di rischi sistemici che possono potenzialmente derivare non solo da contenuti illegali, ma anche dalla diffusione di contenuti legali che possono essere nocivi se disseminati sistematicamente e con determinate modalità, come può essere il caso della disinformazione. Inoltre, l'attuazione dell'art. 35 può realizzarsi tramite misure di moderazione dei contenuti online degli utenti, così interferendo con i diritti fondamentali di questi ultimi.

Pertanto, si può osservare come l'art. 35 del *DSA* attribuisca importanti funzioni regolamentari a *VLOPs* e *VLOSEs*, in collaborazione e sotto la vigilanza della Commissione, per l'attuazione di misure capaci di limitare l'esercizio dei diritti fondamentali degli utenti, ed in particolare la loro libertà di espressione.

Questo approccio regolamentare presenta criticità per il rispetto delle condizioni per la restrizione dei diritti tutelati dalla Carta dei diritti fondamentali dell'Unione europea[19] (di seguito, la "Carta"), sia nella forma che nella sostanza delle sue ricadute applicative. Queste criticità possono essere esaminate con maggiore facilità con riferimento all'attenuazione dei rischi sistemici causati dalla disinformazione, in quanto per la disinformazione è già stato adottato un codice di condotta adatto ad essere impiegato come mezzo di implementazione dell'art. 35, come previsto dall'art. 45 del *DSA*. Pertanto, è possibile avanzare considerazioni sul funzionamento complessivo dell'assetto regolamentare del *DSA* per la mitigazione dei rischi sistemici generati dalla disinformazione[20]. Nei paragrafi che seguono sono esaminate le criticità per la tutela della libertà di espressione degli utenti derivanti dall'implementazione dell'art. 35 del *DSA* per mitigare i rischi sistemici causati dalla disinformazione, con particolare riferimento al rispetto del principio di legalità della Carta, alle conseguenze applicative della delega di funzioni regolamentari agli intermediari, e all'assenza di sufficienti salvaguardie per assicurare che la moderazione di contenuti avvenga nel pieno rispetto del principio di proporzionalità sancito dalla Carta. In sostanza, nei prossimi paragrafi è discussa la

[18] L'art. 45, par. 2, del *DSA* dispone quanto segue:
«Qualora emerga un rischio sistemico significativo ai sensi dell'articolo 34, paragrafo 1, che interessa diverse piattaforme online di dimensioni molto grandi o diversi motori di ricerca online di dimensioni molto grandi, la Commissione può invitare i fornitori di piattaforme online di dimensioni molto grandi interessati o i fornitori di motori di ricerca online di dimensioni molto grandi interessati e altri fornitori di piattaforme online di dimensioni molto grandi, di motori di ricerca online di dimensioni molto grandi, di piattaforme online e di altri di servizi intermediari, se opportuno, nonché le pertinenti autorità competenti, le organizzazioni della società civile e altre parti interessate, a partecipare all'elaborazione dei codici di condotta, anche stabilendo impegni ad adottare misure specifiche di attenuazione dei rischi nonché un quadro di comunicazione periodica sulle misure adottate e sui relativi risultati».

[19] Carta dei diritti fondamentali dell'Unione europea, GUUE C-364/2000.

[20] Il collegamento tra l'art. 35 e la disinformazione non è esplicito sulla base del testo dell'articolo, ma è reso evidente dai preamboli del *DSA* (si vedano i preamboli 83, 84 e 88). È stato anche dedotto dal fatto che tra i rischi sistemici da valutare a norma dell'art. 34 vi sono gli effetti negativi su dibattito civico, processi elettorali, sicurezza pubblica, salute pubblica e dei minori, benessere fisico e mentale della persona.

compatibilità con l'art. 52 della Carta del sistema co-regolamentare disposto dal *DSA*.

3. Restrizioni della libertà di espressione derivanti dall'attuazione dell'art. 35 del *DSA* e del Codice

Prima di procedere oltre, occorre preliminarmente considerare se l'attuazione dell'art. 35 del *DSA* per mitigare i rischi generati dalla disinformazione possa interferire con la libertà di espressione degli utenti che, come noto, è tutelata dall'art. 11 della Carta[21]. A riguardo, l'attuazione dell'art. 35 del *DSA* potrebbe interferire in più modi con la libertà di espressione degli utenti. Trattandosi di un articolo che trova applicazione a un'ampia varietà di situazioni e per diverse finalità, la sua incidenza in concreto sulla libertà di espressione dipenderà da come sarà implementato volta per volta dai soggetti obbligati. Con riferimento allo specifico caso delle misure da adottare per mitigare i rischi causati dalla disinformazione, occorre guardare al Codice per comprendere se siano imposte misure che possano essere considerate restrizioni alla libertà di espressione degli utenti, ovvero alle misure che i firmatari del Codice potrebbero porre in essere per conformarsi all'art. 35.

A tal proposito, occorre riferirsi agli impegni 18 e 21 del Codice. Nell'ambito di questi impegni, i firmatari si sono incaricati di limitare la diffusione dannosa della disinformazione adottando, tra le altre cose, misure consistenti nella proibizione di un contenuto e nella restrizione della visibilità di contenuti (ad esempio agendo sui parametri dei sistemi di raccomandazione). Occorre dunque valutare se tali misure costituiscono restrizioni della libertà di espressione tutelata dall'art. 11 della Carta.

Come prima premessa, bisogna tenere a mente che l'art. 11 obbliga l'Unione e gli Stati membri ad astenersi dal violare la libertà di espressione. Le disposizioni della Carta si applicano, difatti, alle istituzioni e organi dell'Unione, e agli Stati membri esclusivamente quando attuano il diritto dell'Unione[22]. Si nota che non impongono invece direttamente obblighi ai privati, che non sono tenuti a rispettare i diritti fondamentali della Carta nei rapporti con altri privati. Tuttavia, se *VLOPs* e *VLOSEs* imponessero limitazioni al godimento della libertà di espressione in attuazione dell'art. 35 del *DSA*, la Carta troverebbe applicazione in quanto la restrizione discenderebbe direttamente dall'adempimento ad un obbligo dettato dalla normativa europea. Lo stesso si può dire per il caso in cui gli impegni del Codice siano implementati dai firmatari come strumento per adempiere agli obblighi dell'art. 35 del *DSA*. Difatti, in questi casi sarebbe l'applicazione in concreto dell'art. 35 ad essere la fonte della restrizione del diritto fondamentale. Di conseguenza, in qualità di restrizione avente fonte in un atto legislativo europeo, è necessario che tale restrizione sia posta in essere nel rispetto della Carta. Sulla base di questo presupposto è analizzata, nei paragrafi che seguono, la

[21] L'art. 11, par. 1, della Carta prevede quanto segue:
«Ogni individuo ha diritto alla libertà di espressione. Tale diritto include la libertà di opinione e la libertà di ricevere o di comunicare informazioni o idee senza che vi possa essere ingerenza da parte delle autorità pubbliche e senza limiti di frontiera».

[22] Si veda l'art. 51, par. 1, della Carta.

conformità della restrizione alla Carta, nell'ambito degli effetti "verticali" imposti dalla Carta verso le istituzioni dell'Unione. Non sono invece oggetto del presente articolo gli eventuali effetti "orizzontali" che la Carta potrebbe produrre nei rapporti tra privati, posto che, ad oggi, la Corte di giustizia non ha ancora espressamente riconosciuto l'efficacia orizzontale della libertà di espressione, nonostante non sia escluso che tale efficacia possa essere riconosciuta in futuro[23].

Resta dunque da intendere se gli specifici divieti previsti nel Codice possano essere considerati lesivi dell'art. 11 della Carta.

Ai fini dell'interpretazione di questa disposizione, è possibile prendere in considerazione tanto la giurisprudenza della Corte di giustizia quanto quella della Corte europea dei diritti dell'uomo (di seguito, la "CEDU") sull'applicazione del corrispondente art. 10[24] della Convenzione europea dei diritti dell'uomo (di seguito, la "Convenzione")[25], ove viene sancita la libertà di espressione[26]. A riguardo, occorre evidenziare che i contenuti falsi, generalmente qualificabili come disinformazione sulla base delle definizioni attualmente impiegate, sono protetti dalla libertà di espressione dell'art. 11 al pari dei contenuti attendibili[27]. Difatti, nella giurisprudenza della Corte di giustizia e della Corte europea dei diritti dell'uomo la tutela fornita dalla libertà di espressione si estende anche alle informazioni considerate false[28], nonché alle informazioni o idee che offendono, turbano o disturbano[29]. Perciò, i contenuti generalmente considerati come

[23] Sulle ragioni per cui potrebbe essere riconosciuta l'efficacia orizzontale della Carta, soprattutto nei confronti di soggetti capaci di controllare il flusso di informazioni online e limitare significativamente l'esercizio della libertà di espressione, si veda: M. Brkan, *Freedom of expression and Artificial Intelligence: onpersonalisation, disinformation and (lack of) horizontal effect of the Charter,* in *Maastricht Faculty of Law Working Papers*, 2019, 9 ss.

[24] L'art. 10, par. 1, della Convenzione dispone quanto segue:
«Ogni persona ha diritto alla libertà d'espressione. Tale diritto include la libertà d'opinione e la libertà di ricevere o di comunicare informazioni o idee senza che vi possa essere ingerenza da parte delle autorità pubbliche e senza limiti di frontiera. Il presente articolo non impedisce agli Stati di sottoporre a un regime di autorizzazione le imprese di radiodiffusione, cinematografiche o televisive».

[25] Convenzione per la salvaguardia dei Diritti dell'Uomo e delle Libertà fondamentali, Roma, 4.XI.1950.

[26] Il collegamento con la Convenzione è stabilito esplicitamente dall'art. 52, par. 3, della Carta, ove è previsto che, laddove la Carta contenga diritti corrispondenti a quelli garantiti dalla Convenzione, il loro significato e portata sono uguali a quelli conferiti dalla Convenzione, senza precludere che il diritto dell'Unione conceda una protezione più estesa. In buona sostanza, il livello di protezione dei diritti fondamentali stabilito ai sensi della Convenzione deve essere garantito, come standard minimo di tutela, anche nell'applicazione della Carta, salva la possibilità che la Carta conferisca maggiore protezione. Questo collegamento tra la Carta e la Convenzione è di fondamentale importanza per l'interpretazione della libertà di espressione nel diritto dell'Unione, anche in considerazione del fatto che la giurisprudenza della CEDU sulla libertà di espressione è ben più sviluppata della corrispondente giurisprudenza della Corte di giustizia europea.

[27] La libertà di espressione trova ovviamente un limite nel caso in cui, attraverso la disinformazione, si tenti di manipolare intenzionalmente il dibattito pubblico. Tale comportamento mina, infatti, alla radice le basi stesse per cui è tutelata la libertà di espressione e costituisce dunque un abuso del diritto. Ovviamente sul punto si può sviluppare un ampio dibattito dato che non risulta immediato discernere l'uso legittimo della propria libertà di espressione da un abuso di tale diritto (si veda J. Bayer - I. Katsirea - O. Batura - B. Holznagel - S. Hartmann - K. Lubianiec, *The Fight against disinformation and the right to freedom of expression*, in *europarl.europa.eu*, 5 luglio 2021, 84 ss.).

[28] CEDU, *Salov c. Ucraina*, ric. 65518/01, (2005), 113; CEDU, *Dareskizb c. Armenia*, ric. 61737/08, (2021), 71 ss..

[29] CEDU, *Handyside c. Regno Unito*, ric. 5493/72 (1976), 49; CEDU, *Observer and Guardian c. Regno Unito*,

disinformazione sono tutelati dalla libertà di espressione, con la conseguenza che un utente di servizi online ha, in linea di principio, il diritto di diffondere contenuti falsi senza interferenze da parte dei poteri pubblici. Sulla scorta di questa Giurisprudenza, si può dunque concludere che *VLOPs* e *VLOSEs* porrebbero in essere un'interferenza con la libertà di espressione degli utenti sicuramente nel caso più estremo di rimozione e proibizione di un contenuto, che abbiamo visto essere la misura più drastica contemplata dal Codice.

Ciò avverrebbe però del pari anche quando si restringe la visibilità di contenuti, che è un'ulteriore misura prevista dal Codice[30], impostando ad esempio i sistemi di raccomandazione in modo da non raccomandare determinati contenuti agli utenti, o in caso di retrocessione dei contenuti nei risultati di un motore di ricerca. In questo caso, l'interferenza con la libertà di espressione è certamente meno evidente rispetto alla rimozione dei contenuti, e la Corte di giustizia e la CEDU non si sono ad oggi pronunciate sul se costituisca una violazione rispettivamente degli artt. 11 e 10, quando disposta in adempimento ad un obbligo di legge. Tuttavia, è possibile argomentare che si tratta di una restrizione in considerazione dei principi generali sulla libertà di espressione. In primo luogo, nonostante il contenuto in sé non sia rimosso, ne viene limitata la circolazione agendo sui suoi mezzi di diffusione. Come affermato dalla CEDU, l'art. 10 della Convenzione si applica non solo ai contenuti, ma anche ai suoi mezzi di disseminazione[31]. In secondo luogo, la CEDU ha in un'occasione affermato che si ha una restrizione della libertà di espressione non solo quando l'accesso ad un contenuto è reso impossibile, ma anche quando è semplicemente ostacolato[32]. A tal proposito, occorre infatti osservare che i sistemi di raccomandazione sono diventati importanti mezzi di disseminazione, che determinano l'esperienza online degli utenti e tramite cui è possibile manifestare le proprie espressioni online.

Alla luce di tutto quanto sopra, si nota dunque che quanto previsto dal Codice potrebbe effettivamente essere una limitazione della libertà di espressione *ex* art. 11 della Carta. Nonostante i sistemi di raccomandazione discriminino per loro natura, e gli intermediari siano liberi di disciplinarne il funzionamento facilitando la diffusione di alcuni contenuti rispetto ad altri, sorgono problemi per il rispetto della Carta e della Convenzione quando la diffusione dei contenuti è determinata da un obbligo legale, in virtù degli obblighi "negativi" che impongono all'Unione e agli Stati membri di non limitare l'esercizio della libertà di espressione.

Le interferenze sopra descritte possono essere legittime, nell'ordinamento dell'Unione, solo nel caso in cui siano rispettati i requisiti di cui all'art. 52, par. 1, della Carta[33],

ric. 13585/88, (1991), 59.

[30] Si veda l'impegno 18 del Codice.

[31] CEDU, *Autronic AG c. Svizzera*, ric. 12726/87, (1990), 47; CEDU, *Murphy c. Irlanda*, ric. 44179/98, (2003), 61; *Ahmet Yildirim c. Turkey*, ric. 3111/10, (2012), 50; *Pirate Bay: Neij and Sunde Kolisoppi c. Sweden*, ric. 40397/12 (2013); *Pendov c. Bulgaria App*, ric. 44229/11, (2020), 53.

[32] *Pendov c. Bulgaria*, ric. 44229/11, (2020), 53. In questa sentenza, la CEDU ha affermato che sussiste una restrizione alla libertà di espressione se la funzionalità di un sito internet è compromessa, anche qualora il sito internet rimanga tecnicamente accessibile.

[33] Tale disposizione prevede che: «Eventuali limitazioni all'esercizio dei diritti e delle libertà riconosciuti dalla presente Carta devono essere previste dalla legge e rispettare il contenuto essenziale di detti diritti e

ovvero i principi di legalità, legittimità e proporzionalità, e se le restrizioni poste in essere non intaccano il contenuto essenziale della libertà di espressione[34].
Ai fini di questo contributo, si nota che la formulazione degli obblighi all'interno dell'art. 35 fa sorgere dubbi sulla compatibilità con il principio di legalità. Allo stesso tempo, l'attribuzione di responsabilità regolamentari ai soggetti obbligati nell'ambito del sistema di co-regolamentazione potrebbe far sì che, nella pratica, il principio di proporzionalità non sia rispettato. Non sembrano esserci evidenti criticità in relazione alle altre due condizioni previste dall'art. 52, par. 1, della Carta, ovvero la legittimità della restrizione e il rispetto del contenuto essenziale del diritto fondamentale. Pertanto, ai fini di questo contributo, saranno esposte solo le criticità presentate dall'assetto co-regolamentare, con particolare riferimento al contrasto alla disinformazione, in relazione al rispetto dei principi di legalità e proporzionalità.
Le problematiche di cui sopra sono esaminate di seguito, iniziando con le criticità legate all'osservanza del principio di legalità di cui all'art. 52, par. 1, della Carta.

4. Delega di funzioni regolamentari e compatibilità con il principio di legalità: l'assetto co-regolamentare del *DSA* alla prova

4.1. Il principio di legalità nella Carta

L'art. 52, par. 1, della Carta prevede che le eventuali limitazioni all'esercizio dei diritti e delle libertà garantiti dalla stessa Carta devono essere «*previste dalla legge*». Ai fini dell'interpretazione di questo requisito, è essenziale guardare sia alla giurisprudenza della Corte di giustizia, che alla più copiosa giurisprudenza elaborata in materia dalla CEDU.
Secondo quanto affermato dalla Corte di giustizia in alcune tra le sue più recenti ed importanti sentenze, la normativa che dispone una restrizione dei diritti fondamentali garantiti dalla Carta deve, al fine di rispettare il principio di legalità, prevedere norme chiare e precise[35] e definire essa stessa la portata della limitazione al diritto con cui interferisce[36]. La CEDU ha formulato le condizioni per il rispetto del principio di legalità in modo simile a quanto fatto dalla Corte di giustizia, affermando che l'analogo principio ai sensi della Convenzione si intende rispettato quando una restrizione: i)

libertà. Nel rispetto del principio di proporzionalità, possono essere apportate limitazioni solo laddove siano necessarie e rispondano effettivamente a finalità di interesse generale riconosciute dall'Unione o all'esigenza di proteggere i diritti e le libertà altrui».

[34] Anche per la valutazione sul rispetto di questi requisiti occorre prestare particolare attenzione alla giurisprudenza della CEDU sulla restrizione della libertà di espressione, posto che le restrizioni ammesse ai sensi della Carta non possono eccedere quelle permesse in applicazione della Convenzione. Ciò in conseguenza del requisito che il livello di protezione offerto dalla Carta non sia mai inferiore a quello garantito dalla CEDU.

[35] CGUE, C-419/14, *WebMindLicenses Kft. V Nemzeti Adó- és Vámhivatal Kiemelt Adó- és Vám Főigazgatóság*, (2015), 81; C-203/15 e C-698/15, *TELE2Sverige*, (2016), 109.

[36] CGUE, parere 1/15, *Accordo PNR UE-Canada*, (2017), 139; C-311/18, *Facebook Ireland c. Schrems*, (2020), 175; C-401/19, *Repubblica di Polonia c. Parlamento europeo e Consiglio dell'Unione europea*, (2022), 64.

ha una base giuridica nella normativa nazionale, ii) è adeguatamente accessibile, e iii) è prevedibile[37]. Tuttavia, il significato di "legge" nella giurisprudenza della CEDU si caratterizza per avere un'accezione materiale e non formale, poiché comprende varie categorie di diritto scritto, che includono non solo le misure legislative ma anche le norme positive di rango secondario e le misure adottate da organismi di regolamentazione professionale in virtù di poteri di normazione indipendente loro delegati dal Parlamento[38]. Inoltre, la CEDU ricomprende nella nozione di "legge" anche il diritto non scritto, come le norme derivate dalla giurisprudenza[39]. In buona sostanza, per "legge" si intende la norma in vigore così come interpretata dai giudici competenti[40]. Non è chiaro fino a che punto questa nozione di "legge" sarebbe rilevante anche per l'applicazione del principio di legalità nell'ambito dell'art. 52, par. 1, della Carta, posto che la Corte di giustizia non ha mai esplicitamente richiamato questa nozione della CEDU, né ha mai fornito indicazioni che farebbero pensare che questa nozione possa aver fatto il suo ingresso nel diritto europeo. Dato che la "soglia CEDU" per la tutela dei diritti fondamentali è una soglia minima, la Corte potrebbe dare un'interpretazione più stretta del termine "legge" per garantire un maggior livello di protezione. In assenza di chiare indicazioni da parte della Corte di giustizia, l'analisi del presente paragrafo si concentrerà sulla chiarezza e precisione della norma che permette una restrizione, sulla descrizione della portata e sulla prevedibilità della restrizione per gli interessati. Pertanto, l'analisi si concentrerà sulla qualità delle norme che costituiscono la base giuridica della restrizione, piuttosto che sulla loro qualificazione formale.

4.2. La compatibilità con il principio di legalità dell'art. 35 del *DSA*

Come premessa, occorre ricordare che non è necessario, per assicurare conformità al principio di legalità, che vi sia assoluta certezza giuridica e prevedibilità sull'applicazione di una norma che implica la restrizione di un diritto fondamentale, poiché può essere necessario che la legge sia formulata in termini sufficientemente ampi per permetterne l'adattamento alla varietà di fattispecie che possono presentarsi in concreto. Questa premessa è in particolare da tenere a mente per gli obblighi di moderazione dei contenuti online, che per la natura delle fattispecie da disciplinare possono richiedere disposizioni formulate in termini piuttosto ampi. Ciò è stato confermato dalla Corte di giustizia nella recente sentenza[41] sulla legittimità dell'art. 17 della Direttiva 2019/790 (di seguito, la "Direttiva DSM")[42], ove la Corte ha affermato che le fattispecie discipli-

[37] CEDU, *Sunday Times c. Regno Unito*, ric. 6538/74, (1979), 49; *Ahmet Yildirim c. Turkey*, ric. 3111/10, (2012), 59.

[38] CEDU, *Sanoma Uitgevers B.V. c. Paesi Bassi*, ric. 38224/03, (2010), 83.

[39] *Ibid.*

[40] *Ibid.*

[41] CGUE, C-401/19, *Repubblica di Polonia c. Parlamento europeo e Consiglio dell'Unione europea* (2022), 73, 74 e 75.

[42] Direttiva (UE) 2019/790 del Parlamento europeo e del Consiglio del 17 aprile 2019 sul diritto

nate dall'art. 17 richiedevano termini ampi che lasciassero agli enti obbligati la facoltà di decidere sulle specifiche misure da adottare per adempiere ai propri obblighi, così concludendo che le disposizioni contese dell'art. 17 rispettano il principio di legalità di cui all'art. 52, par. 1, della Carta.

Con questa sentenza, la Corte ha di fatto confermato la legittimità di un assetto regolamentare che prevede la delega a soggetti privati di importanti compiti regolamentari sulla moderazione dei contenuti, lasciando a tali soggetti ampia discrezionalità su come implementare azioni che possono incidere significativamente sull'esercizio della libertà di espressione. Tralasciando, ai fini del presente articolo, ogni considerazione critica sulle conclusioni della Corte in relazione all'art. 17, ed utilizzando anzi questa sentenza come punto di riferimento per valutare il rispetto del principio di legalità, è possibile rilevare carenze nell'art. 35 del *DSA*, con particolare riferimento alla sua chiarezza e determinazione delle condizioni per l'imposizione di restrizioni alla libertà di espressione, nonché alla definizione della portata delle restrizioni stesse. Partendo dalla sentenza della Corte sull'art. 17 della Direttiva DSM è anzi possibile argomentare che l'art. 35 non è pienamente conforme al principio di legalità, proprio in ragione della sua differente formulazione e differenti ricadute applicative rispetto all'art. 17 della Direttiva DSM.

Vi sono due principali criticità sollevate dall'art. 35 per il rispetto del principio di legalità. Da un lato, la definizione dei rischi sistemici e le condizioni per l'adozione delle misure di attenuazione dei rischi e, dall'altro, la portata delle possibili restrizioni che ne conseguono. Entrambi questi aspetti influiscono sulla prevedibilità dell'interferenza con la libertà di espressione.

In primo luogo, i rischi sistemici che giustificano l'adozione di misure di attenuazione non sono sufficientemente definiti all'interno del *DSA*, e anzi la loro individuazione è lasciata a *VLOPs* e *VLOSEs*. L'art. 34, par. 1, fornisce alcuni esempi, non esaustivi, di rischi sistemici, che sono formulati per la maggior parte in modo ampio. Questo approccio può apparire sensato se si tiene conto del fatto che il legislatore europeo non poteva prevedere, al momento della stesura del *DSA,* tutti i rischi sistemici che potrebbero presentarsi in futuro con l'evoluzione della società e della tecnologia. Allo stesso tempo, lasciarne l'individuazione a *VLOPs* e *VLOSEs* può sembrare ragionevole perché questi ultimi si trovano nella posizione migliore per riuscire ad identificarli. Tuttavia, la definizione dei rischi sistemici è un passaggio essenziale per determinare a quali condizioni l'art. 35 impone l'obbligo di adottare misure di mitigazione. Non solo non vi è una definizione legislativa di cosa si intenda per rischio sistemico, ma l'art. 34 è vago sulle modalità di individuazione di tali rischi, in quanto non detta puntuali criteri e parametri qualitativi e/o quantitativi per comprendere in quali casi un rischio diventi "sistemico"[43]. Inoltre, non è fornita una lista completa degli interessi pubblici che devono essere in pericolo affinché vi sia un rischio sistemico. La lista di

d'autore e sui diritti connessi nel mercato unico digitale e che modifica le direttive 96/9/CE e 2001/29/CE.

[43] L'art. 34, par. 2, dispone quanto segue:
«Nello svolgimento delle valutazioni dei rischi, i fornitori di piattaforme online di dimensioni molto grandi e di motori di ricerca online di dimensioni molto grandi tengono conto, in particolare, dell'eventualità e

cui al par. 1 non è esaustiva e quindi lascia un punto di domanda su quali altri interessi pubblici potrebbero giustificare l'applicazione dell'art. 35, con un evidente rischio di applicazione arbitraria dell'articolo[44]. Anche nella lista fornita vi sono alcuni interessi pubblici definiti in modo fin troppo ampio, com'è il caso degli effetti negativi sul dibattito civico, i processi elettorali, la sicurezza pubblica, la salute pubblica e dei minori, il benessere fisico e mentale della persona. Per quel che a noi interessa in relazione alla lotta alla disinformazione, la definizione dell'interesse pubblico alla tutela del dibattito civico e dei processi elettorali è un punto particolarmente delicato a potenzialmente soggetto ad interpretazioni difformi. Per un utente online potrebbe essere difficile, se non impossibile, avere un'idea su quali saranno le misure di attenuazione di rischi per il dibattito civico e i processi elettorali, in quanto non ci sono criteri generalmente accettati per comprendere in quali casi una condotta ripetuta in modo sistemico possa creare dei danni al dibattito civico e ai processi elettorali. Ad esempio, la disinformazione, le cui misure di contrasto potrebbero verosimilmente essere adottate sulla base di questi interessi pubblici, non è neanche citata in alcuna disposizione del *DSA* ma solo nei preamboli, e ci potrebbero essere fenomeni sociali la cui connessione con l'art. 34 non è evidente e deve essere ancora stabilita. D'altro canto, i soli rischi siste-

del modo in cui i seguenti fattori influenzano uno dei rischi sistemici di cui al paragrafo 1:
a) la progettazione dei loro sistemi di raccomandazione e di qualsiasi altro sistema algoritmico pertinente;
b) i loro sistemi di moderazione dei contenuti;
c) le condizioni generali applicabili e la loro applicazione;
d) i sistemi di selezione e presentazione delle pubblicità;
e) le pratiche del fornitore relative ai dati.
Le valutazioni analizzano inoltre se e in che modo i rischi di cui al paragrafo 1 siano influenzati dalla manipolazione intenzionale del loro servizio, anche mediante l'uso non autentico o lo sfruttamento automatizzato del servizio, nonché l'amplificazione e la diffusione potenzialmente rapida e ampia di contenuti illegali e di informazioni incompatibili con le condizioni generali.
La valutazione tiene conto di specifici aspetti regionali o linguistici, anche laddove siano specifici di uno Stato membro».

[44] L'art. 34, par. 1, dispone quanto segue:
«I fornitori di piattaforme online di dimensioni molto grandi e di motori di ricerca online di dimensioni molto grandi individuano, analizzano e valutano con diligenza gli eventuali rischi sistemici nell'Unione derivanti dalla progettazione o dal funzionamento del loro servizio e dei suoi relativi sistemi, compresi i sistemi algoritmici, o dall'uso dei loro servizi.
Essi effettuano le valutazioni dei rischi entro la data di applicazione di cui all'articolo 33, paragrafo 6, secondo comma, e successivamente almeno una volta all'anno, e in ogni caso prima dell'introduzione di funzionalità che possono avere un impatto critico sui rischi individuati a norma del presente articolo. La valutazione del rischio deve essere specifica per i loro servizi e proporzionata ai rischi sistemici, tenendo in considerazione la loro gravità e la loro probabilità, e deve comprendere i seguenti rischi sistemici:
a) la diffusione di contenuti illegali tramite i loro servizi;
b) eventuali effetti negativi, attuali o prevedibili, per l'esercizio dei diritti fondamentali, in particolare i diritti fondamentali alla dignità umana sancito nell'articolo 1 della Carta, al rispetto della vita privata e familiare sancito nell'articolo 7 della Carta, alla tutela dei dati personali sancito nell'articolo 8 della Carta, alla libertà di espressione e di informazione, inclusi la libertà e il pluralismo dei media, sanciti nell'articolo 11 della Carta, e alla non discriminazione sancito nell'articolo 21 della Carta, al rispetto dei diritti del minore sancito nell'articolo 24 della Carta, così come all'elevata tutela dei consumatori, sancito nell'articolo 38 della Carta;
c) eventuali effetti negativi, attuali o prevedibili, sul dibattito civico e sui processi elettorali, nonché sulla sicurezza pubblica;
d) qualsiasi effetto negativo, attuale o prevedibile, in relazione alla violenza di genere, alla protezione della salute pubblica e dei minori e alle gravi conseguenze negative per il benessere fisico e mentale della persona».

mici sufficientemente definiti sono la diffusione di contenuti illegali e gli effetti negativi per i diritti fondamentali. Nel primo caso l'illegalità di un contenuto è facilmente verificabile sulla base delle leggi in vigore nell'Unione, e nel secondo caso è possibile comprendere dalla normativa in vigore e dalla relativa giurisprudenza quando un diritto fondamentale si intende violato.

Oltre a mancare, come sopra osservato, una definizione di rischio sistemico, il testo legislativo non fornisce indicazioni su quale sia il livello di gravità di un rischio sistemico che può giustificare l'adozione di misure di attenuazione, lasciando anche questa valutazione alla discrezionalità dei soggetti obbligati. Sia il testo legislativo che i preamboli[45] si concentrano sul fornire esempi, anche piuttosto dettagliati, sull'applicazione dell'art. 35, ma specificando che tali esempi non sono esaustivi, e senza fornire delle indicazioni di validità generale che circoscrivano le condizioni di applicabilità dell'art. 35. Poiché i presupposti di applicabilità dell'articolo sono fenomeni sociali non facilmente misurabili, e il cui impatto sugli interessi pubblici protetti può essere altrettanto difficilmente quantificabile, l'assenza di ulteriori indicazioni genera forti incertezze applicative.

Occorre, a tal proposito, osservare che la regolamentazione dei rischi sistemici non è una novità assoluta nella legislazione europea, ma trova anzi un importante precedente nella disciplina sulla vigilanza prudenziale degli istituti finanziari[46]. Tuttavia, a differenza della normativa europea sulla vigilanza prudenziale degli istituti finanziari, in cui i rischi sistemici che costituiscono i presupposti di applicabilità di varie disposizioni sono matematicamente calcolabili con precise indicazioni sulle relative modalità di calcolo presenti nella normativa, per i fenomeni sociali interessati dal *DSA* non ci sono definizioni generali né indicazioni sulle modalità di calcolo. Prendendo ad esempio il caso della disinformazione, la misurazione del suo impatto sul dibattito civico ed i processi elettorali è una scienza ancora in via di sviluppo, e gli esperti del campo riconoscono che la misurazione sufficientemente precisa sull'impatto della disinformazione, o delle misure di contrasto ad essa, è un'attività per cui vi sono diversi approcci e metodi eterogenei[47], e che può rivelarsi difficoltosa[48], sebbene non impossibile. Per queste ragioni, può rendersi necessario che la legge detti criteri più precisi per la definizione di rischio sistemico e per la sua individuazione.

In secondo luogo, l'art. 35 non definisce quale portata possano avere le misure di attenuazione sulla libertà di espressione. Il par. 1 dell'articolo richiede che *VLOPs* e *VLOSEs* prestino particolare attenzione agli effetti delle misure sui diritti fondamentali, e fornisce una lista di esempi di misure di attenuazione. Pertanto, gli intermediari hanno la facoltà di adottare qualsiasi misura di moderazione dei contenuti che si renda neces-

[45] Si vedano i preamboli da 79 a 91.

[46] Si veda il regolamento (UE) n. 575/2013 del Parlamento europeo e del Consiglio del 26 giugno 2013 relativo ai requisiti prudenziali per gli enti creditizi e le imprese di investimento e che modifica il regolamento (UE) n. 648/2012.

[47] Y. Green - A. Gully - Y. Roth - A. Roy - J. A. Tucker - A. Wanless; *Evidence-Based Misinformation Interventions: Challenges and Opportunities for Measurement and Collaboration*, in *carnegieendowment.org*, 2023.

[48] B. Nimmo, *The breakout scale: measuring the impact of influence operations,* in *brookings.edu*, 2020; S. Altay - M. Berriche - A. Acerbi - *Misinformation on Misinformation: Conceptual and Methodological Challenges,* in *Social Media + Society*, 9(1), 2023.

saria, dalla semplice restrizione della visibilità di un contenuto al blocco di un account utente, pur dovendo valutare se nel caso concreto sia assicurato il rispetto della libertà di espressione. Questo potrebbe rendere difficoltosa per gli utenti la previsione delle restrizioni che potrebbero essere disposte in futuro. Per quanto sia vero che, nel recente caso sull'art. 17 della Direttiva DSM, la Corte ha ritenuto compatibile con il principio di legalità l'utilizzo di nozioni aperte per disciplinare la moderazione di contenuti online, occorre evidenziare che l'art. 17 indica quali restrizioni possano essere disposte per la libertà di espressione degli utenti, pur non definendone le modalità pratiche di attuazione. Difatti, dalla lettura del par. 4 dell'art. 17 è chiaro che, per beneficiare del regime di esenzione dalla responsabilità per la comunicazione al pubblico di contenuti protetti dal diritto d'autore, i prestatori di servizi di condivisione di contenuti online dovrebbero in certi casi rimuovere o disabilitare l'accesso ai contenuti. Perciò, anche se non sono interamente determinate le misure pratiche di implementazione di questi obblighi, gli utenti possono perlomeno prevedere il genere di restrizioni derivanti dall'applicazione dell'art. 17. Nel caso dell'art. 35, non vi è alcuna indicazione nel testo legislativo di quali restrizioni potrebbero essere imposte.

Le carenze sopra evidenziate creano incertezza giuridica, e non sembrano conciliabili con il principio della Corte di giustizia secondo cui la normativa che dispone una restrizione di diritti fondamentali deve definire la portata della restrizione stessa. L'attuale formulazione dell'art. 35 non dà chiarezza sulle sue ricadute applicative e preclude la prevedibilità di eventuali limitazioni alla libertà di espressione per gli utenti. D'altro canto, è attribuita ampia discrezionalità agli intermediari sulla definizione dell'ambito di applicazione dell'art. 35 e, in ultima battuta, alla Commissione che vigila su di essi e può dare chiarimenti sull'applicazione dell'art. 35 nella propria prassi amministrativa.

Resta da chiedersi se le carenze del testo legislativo possano essere colmate da quanto sarà previsto nei codici di condotta e negli orientamenti che la Commissione può adottare a norma dell'art. 35, par. 3. Quanto ai codici di condotta e gli orientamenti della Commissione, si tratta di iniziative volontarie la cui adozione in futuro non è obbligatoria né garantita. Inoltre, non sono fonti vincolanti e non producono effetti giuridici, con la conseguenza che sarebbero difficilmente qualificabili come "leggi" ai sensi dell'art. 52, par. 1, della Carta, soprattutto alla luce del significato che questo termine assume nella giurisprudenza della CEDU[49].

In conclusione, l'attuale assetto regolamentare del *DSA* per la mitigazione dei rischi sistemici sembrerebbe non pienamente conforme ad almeno una delle condizioni richieste per la restrizione di diritti fondamentali ai sensi della Carta, il principio di legalità.

[49] Si veda quanto sopra esposto nel paragrafo 4.1, e in particolare: CEDU, *Sanoma Uitgevers B.V. c. Paesi Bassi*, ric. 38224/03 (2010), 83.

5. Salvaguardie per un'applicazione proporzionata dell'art. 35 del *DSA*, rispetto della libertà di espressione in concreto?

5.1. Le salvaguardie del *DSA* per il rispetto del principio di proporzionalità nella moderazione dei contenuti

Nel presente paragrafo sono esaminate le conseguenze della delega di responsabilità regolamentari agli intermediari nel *DSA*, sotto l'ottica diversa del rispetto, in pratica, del principio di proporzionalità di cui all'art. 52, par. 1, della Carta. Questo paragrafo si propone di dimostrare che, sebbene nella struttura generale del *DSA* l'applicazione dell'art. 35 sembrerebbe essere accompagnata da molteplici salvaguardie per assicurare che le restrizioni della libertà di espressione avvengano nel rispetto del principio di proporzionalità, vi sono molteplici ragioni per argomentare che questo potrebbe non essere il caso nella sostanza della pratica applicativa.

Il *DSA* fornisce salvaguardie procedurali per assicurare che le restrizioni alla libertà di espressione predisposte dai fornitori di servizi intermediari siano il più possibile proporzionate e non arbitrarie. Vi sono infatti due opzioni procedurali nel *DSA* per accedere alla revisione di una decisione restrittiva della libertà di espressione, che fungono da salvaguardie contro l'adozione di decisioni erronee o arbitrarie da parte degli intermediari, volte a garantire che ogni interferenza sia mirata e limitata a quanto strettamente necessario. In primo luogo, il *DSA* prevede, all'art. 20, par. 1, l'obbligo per tutti i fornitori di piattaforme online di mettere a disposizione degli utenti un sistema interno di gestione dei reclami efficace che consenta di presentare reclami contro le decisioni limitative della loro libertà di espressione[50]. L'art. 20 richiede anche che le decisioni sui reclami siano motivate e siano prese con la supervisione di personale e non solo tramite strumenti automatizzati. In secondo luogo, ai sensi dell'art. 21, par. 1, del *DSA*, tutti gli utenti affetti da queste decisioni hanno diritto a scegliere qualunque organismo di risoluzione extragiudiziale delle controversie certificato a norma del *DSA* per risolvere le controversie inerenti alle decisioni stesse[51]. Inoltre, in aggiunta a queste

[50] L'art. 20, par. 1, del *DSA* dispone quanto segue:
«I fornitori di piattaforme online forniscono ai destinatari del servizio, comprese le persone o gli enti che hanno presentato una segnalazione, per un periodo di almeno sei mesi dalla decisione di cui al presente paragrafo, l'accesso a un sistema interno di gestione dei reclami efficace, che consenta loro di presentare per via elettronica e gratuitamente reclami contro la decisione presa dal fornitore della piattaforma online all'atto del ricevimento di una segnalazione o contro le seguenti decisioni adottate dal fornitore della piattaforma online a motivo del fatto che le informazioni fornite dai destinatari costituiscono contenuti illegali o sono incompatibili con le condizioni generali:
a) le decisioni che indicano se rimuovere le informazioni o disabilitare l'accesso alle stesse o se limitarne la visibilità;
b) le decisioni che indicano se sospendere o cessare in tutto o in parte la prestazione del servizio ai destinatari;
c) le decisioni che indicano se sospendere o cessare l'account dei destinatari;
d) le decisioni che indicano se sospendere, cessare o limitare in altro modo la capacità di monetizzare le informazioni fornite dai destinatari».

[51] L'art. 21, par. 1, del DSA dispone quanto segue:
«I destinatari del servizio, compresi le persone o gli enti che hanno presentato segnalazioni, ai quali sono rivolte le decisioni di cui all'articolo 20, paragrafo 1, hanno il diritto di scegliere qualunque organismo

due opzioni fornite dal *DSA*, gli utenti hanno sempre il diritto ad adire gli organi giurisdizionali competenti di uno Stato membro per lamentare la violazione dei loro diritti fondamentali. Infine, occorre menzionare che i firmatari si sono altresì impegnati nel Codice a porre in essere un trasparente sistema di reclamo per gli utenti[52]. Tutte queste tutele procedurali possono accompagnare e facilitare un'applicazione proporzionata dell'art. 35 del *DSA*, in conformità a quanto richiesto dalla giurisprudenza della Corte di giustizia[53].

Al di là delle salvaguardie procedurali, il *DSA* richiama in più punti al rispetto del principio di proporzionalità. Per quel che qui interessa, l'art. 14, par. 4[54], richiede che gli intermediari, nell'imporre restrizioni in relazione all'uso del loro servizio, inclusa la moderazione dei contenuti, agiscano in modo diligente, obiettivo e proporzionato, nel rispetto dei diritti fondamentali dei destinatari del servizio, ivi inclusa la libertà di espressione. L'obbligo dell'art. 14, par. 4, è limitato all'applicazione dei termini e condizioni dei servizi offerti agli utenti, e quindi introduce degli standard di comportamento che gli intermediari devono rispettare nel loro rapporto con gli utenti, inclusi i casi di attuazione di pratiche di moderazione dei contenuti. Questa disposizione può essere interpretata come generatrice di un obbligo per gli intermediari di rispettare i diritti fondamentali degli utenti, introducendo in sostanza un'efficacia orizzontale indiretta di tali diritti che si realizza tramite i termini e condizioni dei servizi offerti dagli intermediari[55]. Inoltre, l'art. 35, par. 1, prevede che le misure di attenuazione dei rischi sistemici siano ragionevoli, proporzionate ed efficaci, adattate ai rischi sistemici affrontati, al contempo «prestando particolare attenzione agli effetti di tali misure sui diritti fondamentali».

Questi articoli riflettono l'approccio legislativo di attribuire responsabilità regolamentari agli intermediari che diventano sempre di più regolatori indipendenti in assetti di co-regolamentazione e dialogo con l'esecutivo[56]. Questo è evidente quando si osserva l'approccio di delega degli artt. 14, par. 4, e 35, par. 1, in combinazione con l'art. 45, secondo cui la Commissione può incoraggiare l'adozione di codici di condotta, in cooperazione con tutte le parti interessate, per definire l'attenuazione dei rischi sistemici

di risoluzione extragiudiziale delle controversie certificato in conformità del paragrafo 3 del presente articolo ai fini della risoluzione delle controversie inerenti a tali decisioni, compresi i reclami che non è stato possibile risolvere mediante il sistema interno di gestione dei reclami di cui a tale articolo».

[52] Si veda l'impegno numero 24 del Codice.

[53] Un principio consolidato nella giurisprudenza della Corte di giustizia è che i meccanismi di reclamo e ricorso sono salvaguardie procedurali essenziali per assicurare la proporzionalità della moderazione di contenuti. Si veda: CGUE, C-401/19, *Repubblica di Polonia c. Parlamento europeo e Consiglio dell'Unione europea*, (2022), 94.

[54] L'Art. 14, par. 4, dispone quanto segue:
«I prestatori di servizi intermediari agiscono in modo diligente, obiettivo e proporzionato nell'applicare e 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 dei destinatari del servizio, quali la libertà di espressione, la libertà e il pluralismo dei media, e altri diritti e libertà fondamentali sanciti dalla Carta».

[55] J. Quintais - N. Appelman - R. Ó Fathaigh - *Using Terms and Conditions to apply Fundamental Rights to Content Moderation - German Law Journal, 24*(5), 2023, 881 ss., 895 ss.

[56] T. Mylly, *The New Constitutional Architecture of Intellectual Property*, in J. Griffiths - T. Mylly, *Global Intellectual Property Protection and New Constitutionalism - Hedging Exclusive Rights*, Oxford, 2021, 50.

individuati a norma dell'art. 34.

L'approccio sopra descritto crea una situazione in cui il legislatore realizza, in sostanza, una delega verso gli intermediari dei propri obblighi di rispettare i diritti fondamentali[57], facendo affidamento sul presupposto che gli intermediari stessi implementino misure di moderazione dei contenuti in conformità ai principi della Carta, tra cui il principio di proporzionalità. Questo soprattutto nel caso in cui la moderazione dei contenuti non sia un'iniziativa volontaria degli intermediari, ma un'azione posta in essere per adempiere agli obblighi dell'art. 35. Di fatto, quindi, la definizione di aspetti fondamentali su come disporre la restrizione della libertà di espressione in determinati contesti sarà decisa da soggetti privati nell'ambito di rapporti tra soggetti privati. A tal proposito, l'art. 14, par. 1, prevede che i prestatori di servizi intermediari forniscano nelle loro condizioni generali informazioni sulle restrizioni che impongono ai contenuti dei destinatari del servizio, con riferimento, tra l'altro, alle politiche, procedure, misure e strumenti utilizzati per la moderazione dei contenuti, compresi il processo decisionale algoritmico e la verifica umana, nonché le regole procedurali del loro sistema interno di gestione dei reclami. In tal modo, l'utente dei servizi verrebbe a conoscenza delle possibili limitazioni cui è soggetto solo tramite un documento contrattuale, senza la garanzia di poter distinguere tra le restrizioni imposte dalla legge e quelle volontariamente disposte dagli intermediari. Questa impostazione porta, in un certo senso, ad una "contrattualizzazione" della libertà di espressione, e lascia che nei rapporti tra privati venga trovato un equilibrio tra gli interessi e diritti fondamentali in gioco nel contesto della moderazione dei contenuti. Occorre notare, tuttavia, che in tale contesto il potere di negoziazione contrattuale di grandi intermediari, come i *VLOPs* e VLOSEs, verso gli utenti può essere prevaricante[58], se non assoluto[59].

5.2. Mettendo in pratica il *DSA*: fino a che punto è ragionevole attendersi il rispetto del principio di proporzionalità?

L'affidamento del legislatore europeo sulla capacità e/o volontà degli intermediari di assicurare che ogni restrizione avvenga in modo proporzionato, senza che il testo legislativo specifichi le modalità con cui far sì che ciò avvenga, può risultare problematico, per varie ragioni.

In primo luogo, un intermediario razionale che effettua moderazione dei contenuti per conformarsi ad un obbligo legislativo sarà mosso, nel decidere le misure con cui farlo in pratica, prevalentemente da ragioni commerciali piuttosto che dall'esigenza di

[57] Con riferimento alla delega realizzata nell'art. 14, par. 4, del *DSA*, si veda: M. Senftleben - J. P. Quintais - A. Meiring, *Outsourcing human rights obligations and concealing human rights deficits: the example of monetization under the CDSMD and the DSA*, 2023, 9.

[58] P. Zumbansen, *The Law of Society: Governance Through Contract*, in *Indiana Journal of Global Legal Studies,* 14(1), 2007; W. Hartzog - A. Melber - E. Salinger, *Fighting Facebook: A Campaign for a People's Terms of Service Center,* in *cyberlaw.stanford.edu*, 2013.

[59] G. De Gregorio, *Digital Constitutionalism in Europe: Reframing Rights and Powers in the Algorithmic Society*, Cambridge, 2022, 110 ss.

rispettare i diritti fondamentali. Ciò potrebbe comportare una tendenza ad adottare un approccio con cui minimizzare il rischio di non conformità alla legge, tramite la massimizzazione della moderazione dei contenuti in tutte le situazioni di rischio. Ciò porta al rischio che siano imposte più restrizioni sui contenuti di quanto sia necessario (c.d. "*overblocking*"). Il rischio di *overblocking* è stato evidenziato da più autori con riferimento alla moderazione dei contenuti in attuazione dell'art. 17 della Direttiva DSM[60], che presenta un simile approccio sulla delega agli intermediari della responsabilità di assicurare il rispetto di diritto fondamentali. Nello specifico caso della disinformazione, gli intermediari agiranno su notifica degli utenti o di *fact-checkers*, anche sulla base degli impegni assunti con il Codice, e dovranno prepararsi a tecniche di manipolazione del web potenzialmente rapide e sofisticate, come l'utilizzo di applicazioni di intelligenza artificiale generativa[61]. Non si può quindi escludere che l'utilizzo di strumenti automatizzati diventi opportuno, se non necessario, per contrastare certe tecniche di diffusione della disinformazione. L'utilizzo di strumenti automatizzati determina un maggior rischio di *overblocking*, soprattutto se gli intermediari non allocano sufficienti risorse di supervisione umana.

In secondo luogo, come già osservato da alcuni autori con riferimento all'art. 17 della Direttiva DSM e all'art. 14, par. 4, del *DSA*[62], il trasferimento dal legislatore agli intermediari di responsabilità per definire come rispettare i diritti fondamentali degli utenti nel caso concreto è un modello che potrebbe non funzionare nella pratica. La critica sollevata riguarda il fatto che gli intermediari potrebbero adottare una propria interpretazione di concetti aperti come "diligenza" e "proporzionalità" nella loro applicazione concreta, per allinearli il più possibile ai propri interessi commerciali piuttosto che al bisogno di tutelare i diritti fondamentali degli utenti. Questa è una conseguenza della circostanza che gli intermediari, i cui interessi non sono in linea, se non talvolta in conflitto, con quelli degli utenti, sono diventati per certi aspetti dei regolatori indipendenti. Per tale ragione, vi sono rischi che il principio di proporzionalità non sia rispettato in concreto, con riferimento ad aspetti di fondamentale importanza che non sono definiti all'interno del testo legislativo, come la necessità di distinguere tra categorie di contenuti (ad esempio, tra contenuti a carattere politico e commerciale) e tra categorie di creatori del contenuto (ad esempio, tra c.d. "*watchdogs*" o semplici utenti). Nell'applicazione dei requisiti di proporzionalità e rispetto dei diritti fondamentali imposti dal *DSA*, è probabile che gli intermediari prediligano considerazioni di costi ed efficienza nello strutturare i propri processi interni per la moderazione e cura dei contenuti. Difatti, *VLOPs* e *VLOSEs* godono di ampia discrezionalità su come implementare concetti aperti in concreto, e costituirebbe per loro la scelta più razionale

[60] M. Senftleben - J. P. Quintais - A. Meiring, *Outsourcing human rights obligations and concealing human rights deficits: the example of monetization under the CDSMD and the DSA*, 2023, 9; M. Perel - N. Elkin-Koren, *Accountability in Algorithmic Copyright Enforcement*, in *Standford Technology Law Review* 19, 473, 2016, 490 ss.

[61] J. A. Goldstein - G. Sastry - M. Musser - R. DiResta - M. Gentzel - K. Sedova, *Generative Language Models and Automated Influence Operations: Emerging Threats and Potential Mitigations*, 2023; Z. Akhtar - *Deepfakes Generation and Detection: A Short Survey*, *Journal of Imaging*, 2023; A. Patel - J. Sattler, *Creatively malicious prompt engineering*, 2023.

[62] M. Senftleben - J. P. Quintais - A. Meiring, *Outsourcing human rights obligations and concealing human rights deficits: the example of monetization under the CDSMD and the DSA*, 2023.

allinearli prevalentemente ad interessi commerciali. *VLOPs* e *VLOSEs* potrebbero essere ancora più incoraggiati a far ciò alla luce del fatto che, secondo il principio di proporzionalità, la portata di obblighi basati sul rischio dipende dalle risorse, dimensioni e altre caratteristiche del soggetto obbligato. La connessione tra considerazioni di costi ed efficienza e gli obblighi di due *diligence* del *DSA* è evidente dal testo del preambolo 41 del *DSA*[63]. Per quanto non sia un problema, ed anzi è auspicabile, che gli obblighi del *DSA* siano calibrati alle caratteristiche e al profilo di rischio di ciascun intermediario, il pericolo è che considerazioni di costi ed efficienza acquisiscano un peso fin troppo importante, e predominante, nell'implementazione dell'art. 35 del *DSA*.
In conclusione, è ragionevole nutrire dubbi sul rispetto in concreto del principio di proporzionalità nella moderazione dei contenuti in adempimento all'art. 35 del *DSA*, ed in particolare sul fatto che *VLOPs* e *VLOSEs* agiscano sempre nel miglior interesse pubblico e degli utenti. Tuttavia, per il caso in cui vi siano restrizioni non proporzionate alla libertà di espressione, il *DSA* predispone dei sistemi di controllo. In primo luogo, gli utenti dispongono di un sistema di rimedi a tre livelli. In secondo luogo, il DSA introduce un sistema di vigilanza regolamentare sull'operato di *VLOPs* e *VLOSEs,* che è centralizzato a livello europeo con l'attribuzione di funzioni di vigilanza alla Commissione. Pertanto, occorre valutare se questi due sistemi di controllo, ovvero i meccanismi di ricorso giudiziale o stragiudiziale da un lato, e la vigilanza della Commissione dall'altro, possono offrire rimedio ai rischi insiti nell'esternalizzazione di funzioni regolamentari agli intermediari.
I prossimi due paragrafi saranno quindi dedicati ad esaminare se, nell'assetto generale del *DSA*, è assicurato un equilibrio che garantisca la tutela della libertà di espressione degli utenti in concreto.

5.3. (Segue) I rimedi giudiziali e stragiudiziali nel *DSA*

5.3.1. Il primo livello: meccanismi di reclamo e ricorso

Il meccanismo di reclamo e ricorso (c.d. "*complaint and redress*") richiesto dal *DSA* è volto a salvaguardare le restrizioni eccessive alla libertà di espressione degli utenti. La messa a disposizione di tale meccanismo è richiesta dall'art. 20 del *DSA* per una serie di decisioni che limitano la libertà di espressione degli utenti[64]. Inoltre, l'art. 54

[63] Il preambolo 41 del DSA dispone quanto segue:
«A tale riguardo è importante che gli obblighi in materia di dovere di diligenza siano adeguati al tipo, alle dimensioni e alla natura del servizio intermediario interessato. Il presente regolamento stabilisce pertanto obblighi fondamentali applicabili a tutti i prestatori di servizi intermediari nonché obblighi supplementari per i prestatori di servizi di memorizzazione di informazioni e, più specificamente, per i prestatori di piattaforme online, di piattaforme online di dimensioni molto grandi e di motori di ricerca online di dimensioni molto grandi (…)».

[64] L'art. 20, par. 1, del GUUE, L-277 del 27.10.2022. dispone quanto segue:
«I fornitori di piattaforme online forniscono ai destinatari del servizio, comprese le persone o gli enti che hanno presentato una segnalazione, per un periodo di almeno sei mesi dalla decisione di cui al presente paragrafo, l'accesso a un sistema interno di gestione dei reclami efficace, che consenta loro di presentare per via elettronica e gratuitamente reclami contro la decisione presa dal fornitore della

del *DSA* prevede che gli utenti abbiano il diritto di richiedere ai fornitori di servizi intermediari un risarcimento per danni o perdite subiti per violazione degli obblighi stabiliti dal *DSA*[65].

I meccanismi di reclamo e ricorso sono volti a rimediare al rischio che gli intermediari prendano decisioni erronee o arbitrarie relative alla moderazione di contenuti considerati illegali o incompatibili con le loro condizioni generali. Tuttavia, affinché questi meccanismi rispondano effettivamente alla loro funzione, occorre che gli utenti siano generalmente in grado di assolvere al compito di individuare le violazioni della normativa rilevante che li interessano, e che abbiano gli strumenti e la motivazione per agire prontamente. Inoltre, vi deve essere la garanzia che il rimedio fornito sia in ogni caso effettivo. L'effettività del rimedio è un aspetto centrale, in quanto è necessaria per assicurare il rispetto dei diritti fondamentali degli utenti, nel rispetto del principio generale comunitario di effettività[66], ma anche per far sì che questi ultimi siano motivati ad agire dinanzi alle decisioni erronee ed arbitrarie degli intermediari. Tuttavia, vi sono due ragioni per pensare che un meccanismo di reclamo e ricorso gestito internamente dagli intermediari non garantisca una tutela effettiva della libertà di espressione degli utenti. Una prima ragione riguarda il rischio che gli intermediari non implementino adeguatamente i meccanismi di reclamo e ricorso. L'art. 20 del *DSA* non fornisce chiare indicazioni su come tale meccanismo debba essere strutturato. In particolare, si esprime in termini generici su elementi cruciali, tra cui le tempistiche da rispettare per lo svolgimento della procedura e come debbano essere progettate le interfacce online per assicurare un agevole esercizio dei diritti di reclamo e ricorso degli utenti. L'art. 20 fa affidamento, al pari degli artt. 14 e 35, su formule che denotano generici standard di comportamento, come l'obbligo di avere sistemi di «facile accesso e uso» e che «consentano e agevolino la presentazione di reclami sufficientemente precisi e adeguatamente motivati». Inoltre, i fornitori di piattaforme online devono gestire i reclami in modo «tempestivo, non discriminatorio, diligente e non arbitrario». Al pari di quanto già osservato con riferimento al rispetto della proporzionalità negli artt. 14 e 35, i fornitori di piattaforme online godono di significativa discrezionalità nel decidere

piattaforma online all'atto del ricevimento di una segnalazione o contro le seguenti decisioni adottate dal fornitore della piattaforma online a motivo del fatto che le informazioni fornite dai destinatari costituiscono contenuti illegali o sono incompatibili con le condizioni generali:
a) le decisioni che indicano se rimuovere le informazioni o disabilitare l'accesso alle stesse o se limitarne la visibilità;
b) le decisioni che indicano se sospendere o cessare in tutto o in parte la prestazione del servizio ai destinatari;
c) le decisioni che indicano se sospendere o cessare l'account dei destinatari;
d) le decisioni che indicano se sospendere, cessare o limitare in altro modo la capacità di monetizzare le informazioni fornite dai destinatari».

[65] L'art. 54 del DSA dispone quanto segue:
«I destinatari del servizio hanno il diritto di chiedere un risarcimento, conformemente al diritto dell'Unione e nazionale, ai fornitori di servizi intermediari relativamente a danni o perdite subiti a seguito di una violazione degli obblighi stabiliti dal presente regolamento da parte dei fornitori di servizi intermediari».

[66] Il principio di effettività è un principio generale del diritto dell'Unione europea. Per una trattazione sull'effettività nel diritto europeo T. Tridimas, *The General Principles of EC Law*, Oxford, 2006, 418 ss.; S.M. Carbone, *Principio di effettività e diritto comunitario*, Università degli Studi Suor Orsola Benincasa - Lezioni Magistrali, 2009.

come implementare questi standard. L'aspetto maggiormente problematico riguarda le tempistiche per la gestione dei reclami, su cui l'art. 20 non solo omette di fornire indicazioni precise, ma richiede soltanto che i fornitori annullino la decisione «senza indebito ritardo»[67]. Ciò potrebbe non essere sufficiente alla luce del fatto che contenuti pubblicati su piattaforme online possono perdere rilevanza in un arco di tempo molto breve, e che in determinati casi è necessario rimuovere celermente la restrizione imposta ad un contenuto affinché il rimedio sia effettivo per l'utente. A tal proposito, l'art. 20 non prevede espressamente che il fornitore debba tenere conto della natura del contenuto oggetto della restrizione, differenziando i tempi di risposta sulla base degli specifici rischi derivanti, caso per caso, dal tardivo annullamento della decisione. Per quanto si potrebbe argomentare, in una certa misura, che un obbligo di differenziare in base alle circostanze sia insito nel requisito di agire senza indebito ritardo, ciò non è evidente dal testo, e non è pacifico che gli intermediari interpretino i propri obblighi in questo senso. Di conseguenza, l'art. 20 non garantisce che i fornitori di piattaforme online prendano effettivamente in considerazione eventuali circostanze del caso concreto che indicano la necessità di procedere con urgenza nel procedimento di revisione della decisione. Presumibilmente, un fornitore razionale tenderà a predisporre procedure di ricorso e reclamo che implicano minori costi, fintantoché gli è concesso. In relazione al caso della disinformazione, le carenze dell'art. 20 sull'implementazione pratica dei meccanismi di ricorso e reclamo non trovano rimedio nel Codice, che risulta essere ancora più vago in materia[68]. In buona sostanza, dalla combinazione del testo legislativo del *DSA* e del Codice, è possibile concludere che i fornitori di piattaforme online godranno di ampia discrezionalità nello strutturare i meccanismi di reclamo e ricorso sulla base delle proprie considerazioni ed esigenze, con il rischio che vengano predisposti meccanismi inefficienti ed inefficaci che non tutelano adeguatamente la libertà di espressione degli utenti. Gli utenti, dal canto loro, non sarebbero motivati ad agire se non dispongono di un efficiente ed efficace strumento di tutela, rimanendo sprovvisti dinanzi alle condotte illegali dei fornitori di piattaforme online. Secondo quanto riportato in studi empirici, gli utenti hanno spesso riscontrato problemi nei sistemi di notifica e reclamo predisposti da intermediari[69].

Quanto alla seconda ragione, come già evidenziato da altri autori[70], fare affidamento sulla proattività degli utenti a richiedere prontamente un rimedio contro decisioni limitative della loro libertà di espressione potrebbe essere un approccio fondato su un presupposto sbagliato. Ci sono risultanze empiriche che dimostrano come gli utenti non siano per larga parte motivati a presentare reclamo contro decisioni che limitano la loro libertà di espressione[71], e decidano di non agire perché non sanno come proce-

[67] Si veda l'art. 20, par. 4, del *DSA*.

[68] Si veda l'impegno n. 24 del Codice.

[69] HateAid report, *Unsatisfied and helpless - how social media platforms are failing users*, 2022; S. M. West, *Censored, suspended, shadowbanned: User interpretations of content moderation on social media platforms*, in *New Media & Society*, 20(11), 2018, 4378 ss.

[70] M. Senftleben - J. P. Quintais. - A. Meiring, *Outsourcing human rights obligations and concealing human rights deficits: the example of monetization under the CDSMD and the DSA*, 2023, 22 ss.

[71] J.M. Urban - L. Quilter, *Efficient Process or "Chilling Effects"? Takedown Notices Under Section 512 of the Digital Millennium Copyright Act*, in *Santa Clara Computer and High Technology Law Journal*, 2, 2006, 621.

dere o non si attendono una risposta dall'intermediario[72]. Al di là degli studi empirici, è ragionevole attendersi che, perlomeno nel caso della disinformazione, l'applicazione di una restrizione ad un contenuto possa generare un "*chilling effect*" capace di produrre effetti dissuasivi verso gli utenti. In particolare, nel momento in cui è applicata una restrizione a un contenuto in quanto falso o non attendibile, l'utente che non aveva agito con piena coscienza ed intenzione di diffondere notizie false potrebbe risentire dello stigma che deriva dalla qualificazione del suo contenuto come disinformazione, percependo che la propria partecipazione al dibattito pubblico non è adeguata o ben accolta. Per questa ragione, un utente potrebbe anche essere dissuaso dal presentare reclamo contro una decisione dell'intermediario.
Per le ragioni sopra esposte, non è garantito che, nella loro concreta operatività, i meccanismi di reclamo e ricorso possano attenuare i rischi di decisioni erronee o abusive degli intermediari per garantire un'effettiva tutela della libertà di espressione degli utenti.

5.3.2. Il secondo livello: la risoluzione stragiudiziale delle controversie

L'art. 21 del *DSA* prevede il diritto per gli utenti, compresi coloro che hanno già presentato reclamo ai sensi dell'art. 20, di avere accesso alla risoluzione stragiudiziale delle loro controversie con gli intermediari presso un organismo certificato a norma del par. 3 dello stesso articolo[73]. Gli organismi di risoluzione stragiudiziale delle controversie sono certificati dai coordinatori dei servizi digitali dopo aver verificato che rispondono ad una serie di requisiti, tra cui l'imparzialità ed indipendenza dagli intermediari e dai destinatari dei servizi degli intermediari. È così assicurato che questi organismi siano esterni ed indipendenti rispetto alle parti della controversia. È inoltre richiesto che gli organismi siano competenti a risolvere le controversie e che siano in grado di procedere in modo rapido, efficiente ed efficace.
Gli organismi di risoluzione stragiudiziale delle controversie sono un'innovazione apportata dal *DSA* in materia di moderazione dei contenuti, e rimane da vedere come saranno applicati gli stringenti criteri per la certificazione di questi organismi. Ad esempio, il *Meta Oversight Board*[74], nonostante le salvaguardie attualmente in essere per garantirne l'indipendenza ed affidabilità, dovrebbe essere oggetto di una significativa riorganizzazione per riuscire a soddisfare i requisiti per la certificazione come organismo di

[72] Si vedano le risultanze del progetto *OnlineCensorship.org*, come riassunte in S. M. West, *Censored, suspended, shadowbanned: User interpretations of content moderation on social media platforms,* in *New Media & Society*, 20(11), 2018, 4378,4379.

[73] L'art. 21, par. 1, primo capoverso dispone quanto segue:
«I destinatari del servizio, compresi le persone o gli enti che hanno presentato segnalazioni, ai quali sono rivolte le decisioni di cui all'articolo 20, paragrafo 1, hanno il diritto di scegliere qualunque organismo di risoluzione extragiudiziale delle controversie certificato in conformità del paragrafo 3 del presente articolo ai fini della risoluzione delle controversie inerenti a tali decisioni, compresi i reclami che non è stato possibile risolvere mediante il sistema interno di gestione dei reclami di cui a tale articolo».

[74] Per maggiori informazioni sul Meta Oversight Board, si veda il seguente *link*.

risoluzione stragiudiziale delle controversie[75].

I requisiti imposti per la certificazione di un organismo di risoluzione stragiudiziale delle controversie potrebbe porre rimedio alle due maggiori criticità riscontrate per i meccanismi di reclamo e ricorso, sia per il fatto che questi organismi sono esterni ed indipendenti rispetto agli intermediari, sia perché è loro richiesto di agire celermente, il che potrebbe motivare gli utenti a presentare ricorso. Se questo è il caso dipenderà però da come saranno certificati e strutturati gli organismi in futuro, con la possibilità che vi siano anche importanti divergenze tra gli Stati membri. Tuttavia, si può osservare almeno un'importante criticità: questi organismi non hanno il potere di imporre una risoluzione della controversia vincolante per le parti, secondo quanto previsto dall'art, 21, par. 2, del *DSA*[76]. Pertanto, gli intermediari potrebbero liberamente decidere di non conformarsi alle decisioni degli organismi stragiudiziali. Per questo motivo, gli utenti potrebbero non essere inclini ad adire questi organismi e, anche qualora lo facciano, la tutela effettiva della loro libertà di espressione non sarebbe garantita se gli intermediari decidono di non adempiere a quanto richiesto nella decisione. Pertanto, parimenti alla conclusione raggiunta per i meccanismi di reclamo e ricorso, potrebbe non essere garantita un'effettiva tutela della libertà di espressione degli utenti in caso di decisioni erronee o abusive degli intermediari.

5.3.3. Il terzo livello: il ricorso agli organi giurisdizionali

Gli utenti che hanno presentato ricorso internamente presso l'intermediario, o che hanno adito uno degli organismi di risoluzione stragiudiziale delle controversie, conservano in ogni caso il diritto di agire dinanzi ad un organo giurisdizionale. Per gli utenti permane quindi la facoltà di accedere al giudizio di un organo indipendente, competente e le cui decisioni hanno poteri vincolanti, così potendo offrire un rimedio che sopperisce alle carenze delle altre opzioni rimediali.

Per quanto non vi siano criticità legate al ricorso giurisdizionale in termini di indipendenza e competenza dell'organo competente, possono esserci dubbi sulla motivazione degli utenti ad agire, e su quanto la tutela possa essere effettiva, dati i tempi medi di risoluzione giudiziale delle controversie nell'Unione[77]. Le procedure giurisdizionali pos-

[75] A.Kuczerawy, *Social Media Councils under the DSA: A path to individual error correction at scale?*, in C. M. Kettemann - J. Francke - C. Dinar - L. Hinrichs, *Platform://Democracy - Perspectives on Platform Power, Public Values and the Potential of Social Media Councils: Research Report Europe,* Hamburg: Verlag Hans-Bredow-Institut, 2023, 46 ss.

[76] L'art. 21, par. 2, del DSA dispone quanto segue:
«Entrambe le parti adiscono in buona fede l'organismo di risoluzione extragiudiziale delle controversie selezionato e certificato al fine di risolvere la controversia.
I fornitori di piattaforme online possono rifiutarsi di adire tale organismo di risoluzione extragiudiziale delle controversie qualora una controversia riguardante le stesse informazioni e gli stessi motivi di presunta illegalità o incompatibilità dei contenuti sia già stata risolta.
L'organismo di risoluzione extragiudiziale delle controversie certificato non ha il potere di imporre una risoluzione della controversia vincolante per le parti».

[77] Per informazioni precise sui tempi medi della giustizia nell'Unione, si veda l'ultimo quadro di valutazione UE della giustizia pubblicato dalla Commissione europea nel 2023: *The 2023 EU Justice Scoreboard, Communication from the Commission to the European Parliament, the Council, the European Central*

sono essere costose ed impegnative in molti Stati membri dell'Unione[78], e in relazione alle restrizioni applicate a contenuti online la loro effettività dipende dalla possibilità di ottenere celermente misure cautelari, il che potrebbe non essere sempre possibile. In alcuni casi, una notizia pubblicata online potrebbe perdere rilevanza nel giro di ore o di pochi giorni, un arco di tempo che potrebbe essere fin troppo ristretto per ottenere una misura cautelare.

5.4. (Segue) La supervisione della Commissione

5.4.1. I meccanismi di controllo sull'operato di *VLOPs* e *VLOSEs* nel *DSA*

Il secondo possibile contrappeso all'ampia discrezionalità di *VLOPs* e *VLOSEs* risiede nei sistemi di controllo sul loro operato predisposti dal *DSA* e dal Codice, e nei poteri di indagine, correttivi e sanzionatori di cui dispone la Commissione per vigilare sulla corretta implementazione del *DSA*. Questi sistemi di controllo e il ruolo della Commissione sono esaminati di seguito, al fine di verificare fino a che punto possano contribuire a garantire che il principio di proporzionalità sia rispettato in concreto, e che vi sia una tutela effettiva della libertà di espressione degli utenti.
Sulla base del *DSA*, vi sono meccanismi di controllo previsti dagli artt. 35, par. 2 e 3, 37, 40, 45, par. 3 e 4, e 65 ss.. Questi si suddividono in meccanismi di controllo con cui la Commissione può indicar*e ex ante a VLOPs e* VLOSEs come adempiere agli obblighi del *DSA*, meccanismi di controllo *ex post* sull'operato di *VLOPs* e *VLOSEs*, e meccanismi che presentano entrambe le caratteristiche.
Quanto al controllo *ex ante*, i par. 2 e 3 dell'art. 35 prevedono che il comitato possa pubblicare, in cooperazione con la Commissione, relazioni sulle migliori pratiche che *VLOPs* e *VLOSEs* possono seguire per attenuare i rischi sistemici individuati, e che la Commissione possa emanare, in cooperazione con i coordinatori dei servizi digitali, orientamenti sull'applicazione dell'art. 35, par. 1. Grazie e questi strumenti, la Commissione ha la facoltà di guidare i fornitori verso un'implementazione degli obblighi del *DSA* che rispetti al meglio i diritti fondamentali degli utenti, così limitando l'ampia discrezionalità lasciata dal testo del *DSA*.
Quanto al controllo *ex post*, il *DSA* prevede molteplici sistemi di controllo. Un primo esempio è la facoltà concessa alla Commissione di richiedere, ai sensi dell'art. 40, par. 1, del *DSA*[79], l'accesso ai dati di *VLOPs* e *VLOSEs* considerati necessari per monitorare e valutare la conformità del loro operato alle disposizioni del *DSA*.
Un secondo esempio di controllo *ex post* è la revisione indipendente di cui all'art. 37.

Bank, the European Economic and Social Committee and the Committee of the Regions, COM (2023) 309, 2023.

[78] *Ibid.*

[79] L'art. 40, par. 1, del DSA dispone quanto segue:
«I fornitori di piattaforme online di dimensioni molto grandi e di motori di ricerca online di dimensioni molto grandi forniscono al coordinatore dei servizi digitali del luogo di stabilimento o alla Commissione, su loro richiesta motivata ed entro un termine ragionevole specificato in detta richiesta, l'accesso ai dati necessari per monitorare e valutare la conformità al presente regolamento».

L'art. 37, par. 1, richiede che *VLOPs* e *VLOSEs* si sottopongano, almeno una volta all'anno, a revisioni indipendenti volte a valutare la conformità delle loro condotte agli obblighi di cui al Capo III del *DSA* (tra cui rientrano gli obblighi di mitigazione dei rischi sistemici ai sensi dell'art. 35), e agli impegni assunti sulla base dei codici di condotta e dei protocolli di crisi di cui sono firmatari[80]. Le revisioni sono effettuate da organizzazioni che soddisfano una serie di requisiti di indipendenza, comprovata esperienza, obiettività e deontologia professionale. A seguito di ogni revisione, le organizzazioni incaricate devono redigere una relazione contenente, tra le altre cose, un giudizio sul rispetto dei rilevanti obblighi ed impegni da parte di *VLOPs* e *VLOSEs,* che può essere "positivo", "positivo con osservazioni" o "negativo". In caso di giudizio di revisione "non positivo", la relazione deve fornire raccomandazioni operative sulle misure rimediali da impiegare. *VLOPs* e *VLOSEs* che ricevono giudizi di revisione non positivi devono tenere conto delle raccomandazioni fornite per rimediare alle carenze riscontrate, e sono tenuti ad adottare una relazione di attuazione della revisione in cui descrivono le misure adottate o, nel caso in cui non attuino le raccomandazioni operative, le ragioni di tale scelta e le misure alternative per cui hanno optato. Nell'ambito del sistema di revisione, è ragionevole attendersi che la Commissione avrà accesso alle relazioni delle organizzazioni indipendenti e dei *VLOPs* e *VLOSEs*, in particolare ai sensi dell'art. 40 del *DSA* e nell'ambito dell'esercizio da parte della Commissione dei propri poteri di vigilanza ed indagine previsti dagli artt. 67, 68, 69 e 72 del *DSA*, ovvero il potere di richiedere informazioni, di audizione e di raccolta di informazioni, di effettuare ispezioni o azioni di monitoraggio. Si tratta quindi di un sistema di controllo che opera indipendentemente dalla Commissione, prevedendo solo l'interazione tra revisori indipendenti e fornitori di piattaforme. Tuttavia, è ragionevole pensare che il guardiano ultimo rimanga sempre la Commissione, posto che i revisori indipendenti non hanno alcun potere di intervento sull'operato di *VLOPs* e *VLOSEs*[81] e che, nel caso in cui questi ultimi non agiscano adeguatamente per rimediare ad un giudizio di revisione "non positivo", solo la Commissione potrà adottare misure cogenti ordinando misure provvisorie, rendendo vincolanti gli impegni assunti da un *VLOP* o un *VLOSE*, adottando una decisione di non conformità o irrogando sanzioni pecuniarie[82].

Un terzo meccanismo di controllo *ex post* risiede nei poteri diretti di indagine della

[80] L'art. 37, par. 1, del DSA dispone quanto segue:
«I fornitori di piattaforme online di dimensioni molto grandi e di motori di ricerca online di dimensioni molto grandi si sottopongono, a proprie spese e almeno una volta all'anno, a revisioni indipendenti volti a valutare la conformità:

a) agli obblighi stabiliti al capo III;
b) agli impegni assunti a norma dei codici di condotta di cui agli articoli 45 e 46 e dei protocolli di crisi di cui all'articolo 48».

[81] L'art. 37, par. 6, del DSA dispone quanto segue:
«I fornitori di piattaforme online di dimensioni molto grandi e di motori di ricerca online di dimensioni molto grandi che ricevono una relazione di revisione che non è «positiva» tengono debitamente conto delle raccomandazioni operative ad essi rivolte al fine di adottare le misure necessarie per attuarle. Entro un mese dal ricevimento di tali raccomandazioni, essi adottano una relazione di attuazione della revisione con cui stabiliscono tali misure. Qualora non diano attuazione alle raccomandazioni operative, essi giustificano nella relazione di attuazione della revisione le ragioni di tale scelta e descrivono le misure alternative da esse adottate per risolvere eventuali casi di non conformità rilevati».

[82] Tali misure possono essere adottate, rispettivamente, a norma degli artt. 70, 71, 73 e 74 del *DSA*.

Commissione sulla base degli artt. 66 e seguenti del *DSA*. In particolare, la Commissione può avviare un procedimento di indagine al fine di adottare una decisione di non conformità o di irrogare sanzioni pecuniarie, a norma rispettivamente degli artt. 73 e 74 del *DSA*, con la possibilità di disporre misure provvisorie nel corso del procedimento come previsto dall'art. 70 del *DSA*. Inoltre, ai sensi dell'art. 71 del DSA la Commissione può rendere vincolanti gli impegni assunti da un *VLOP* o un *VLOSE*. La Commissione può esercitare i propri poteri di vigilanza sulla base delle informazioni trasmesse tramite i sistemi di controllo previsti dal *DSA*, o autonomamente qualora lo ritenga opportuno.

Con riferimento al controllo sia *ex ante* che *ex post*, l'art. 45 prevede un intervento da parte della Commissione e del comitato sull'elaborazione e l'implementazione dei codici di condotta. Da un lato, la Commissione e il comitato possono influenzare il contenuto dei codici di condotta. Nonostante non partecipino formalmente alla stesura dei codici, che è lasciata ai destinatari delle norme del *DSA*, possono indirettamente influenzarne il contenuto quando ne incoraggiano e agevolano l'elaborazione ai sensi dell'art. 45, par. 1[83]. Dall'altro lato, la Commissione e il comitato possono controllare l'implementazione dei codici di condotta tramite un meccanismo di rendicontazione periodica sulle misure adottate dai firmatari e i relativi risultati, come previsto dall'art. 45, par. 3, del *DSA*[84], e come meglio definito all'interno del Codice. All'interno della Sezione X del Codice i firmatari si sono impegnati a presentare rapporti periodici alla Commissione su come hanno implementato il Codice, e sui conseguenti risultati[85]. Sulla base di tale rendicontazione, la Commissione e il comitato possono valutare se i codici di condotta rispondono alle finalità per cui sono stati elaborati, anche tenendo conto di eventuali indicatori chiave di prestazione ivi previsti, e pubblicano le loro conclusioni. L'art. 45, par. 4, prevede che, in caso di inottemperanza sistematica ai codici di condotta, la Commissione e il comitato possono invitare i firmatari ad adottare le misure necessarie.

[83] L'art. 45, par. 1, del DSA dispone quanto segue:
«La Commissione e il comitato incoraggiano e agevolano l'elaborazione di codici di condotta volontari a livello di Unione per contribuire alla corretta applicazione del presente regolamento, tenendo conto in particolare delle sfide specifiche connesse alla lotta ai diversi tipi di contenuti illegali e ai rischi sistemici, conformemente al diritto dell'Unione, in particolare in materia di concorrenza e protezione dei dati personali».

[84] L'art. 45, par. 3, del DSA dispone quanto segue:
«Nell'attuare i paragrafi 1 e 2, la Commissione e il comitato nonché, ove opportuno, altri organismi mirano a garantire che i codici di condotta definiscano chiaramente i loro obiettivi specifici, contengano indicatori chiave di prestazione per misurare il conseguimento di tali obiettivi e tengano debitamente conto delle esigenze e degli interessi di tutte le parti interessate, in particolare dei cittadini, a livello di Unione. La Commissione e il comitato mirano inoltre a garantire che i partecipanti riferiscano periodicamente alla Commissione e ai rispettivi coordinatori dei servizi digitali del luogo di stabilimento in merito a tutte le misure adottate e ai relativi risultati, misurati sulla base degli indicatori chiave di prestazione contenuti nei codici di condotta. Gli indicatori chiave di prestazione e gli obblighi di comunicazione tengono conto delle differenze esistenti tra i diversi partecipanti in termini di dimensioni e capacità».

[85] Si vedano gli impegni n. 38-44 del Codice.

5.4.2. Luci e ombre del ruolo di vigilanza della Commissione

I meccanismi sopra descritti potrebbero costituire importanti strumenti di salvaguardia dei diritti fondamentali degli utenti delle piattaforme e dei motori di ricerca online di dimensioni molto grandi, assicurando un certo livello di controllo sull'ampia discrezionalità attribuita a *VLOPs* e *VLOSEs* nell'adempimento degli obblighi dell'art. 35 del *DSA*. Da un lato, la Commissione può agire *ex ante* adottando migliori pratiche ed orientamenti per fornire precise indicazioni sulle misure di attenuazione dei rischi da porre in essere, nonché influenzare il contenuto dei codici di condotta elaborati a norma dell'art. 45, assicurando che l'applicazione in concreto dell'art. 35 non porti ad interferenze eccessive e non proporzionate con la libertà di espressione degli utenti. Dall'altro lato, la Commissione dispone di numerosi strumenti per controllare *ex post* l'implementazione dei codici di condotta e, più in generale, tutte le attività di *compliance* al *DSA* poste in essere dai *VLOPs* e *VLOSEs* grazie alle revisioni indipendenti. Per far ciò, la Commissione può beneficiare di un continuo flusso di informazioni sull'operato dei soggetti obbligati e di incisivi poteri di indagine.

Tuttavia, vi è almeno una ragione per cui è possibile argomentare che la supervisione della Commissione tramite questi meccanismi di controllo non permetterebbe un'adeguata tutela della libertà di espressione degli utenti. Nell'architettura del *DSA* l'unica autorità con poteri amministrativi verso i *VLOPs* e i *VLOSEs* è la Commissione. Come previsto dall'art. 66 del *DSA*, solo la Commissione può avviare un procedimento amministrativo nel caso in cui vi sia il sospetto che un *VLOP* o *VLOSE* ha violato una disposizione del *DSA*. Il conferimento di tali poteri alla Commissione può essere problematico, in quanto la Commissione non è un'autorità indipendente, ma è un organo esecutivo avente legami con il potere politico nell'assetto istituzionale dell'Unione[86]. Pertanto, un'istituzione potenzialmente soggetta a pressioni politiche giocherà un ruolo primario nella definizione delle norme applicabili all'esercizio della libertà di espressione online, ed in particolare sulle infrastrutture online gestite dai cosiddetti "*gatekeepers*", che costituiscono oggigiorno fori di discussione di fondamentale importanza per l'esercizio della libertà di espressione in un sistema democratico. Questo assetto presenta evidenti rischi per la libertà di espressione degli utenti, soprattutto tenendo conto che la supervisione sulle misure di contrasto alla disinformazione è un'attività delicata con importanti ripercussioni per la circolazione delle informazioni in un sistema democratico. Tramite la stretta cooperazione ed il continuo dialogo con i soggetti vigilati la Commissione, un'istituzione legata al potere politico può partecipare alla definizione dei meccanismi con cui decidere quali contenuti siano accettabili o meno, e quali restrizioni possano essere ammesse.

I rischi derivanti dall'attribuzione di competenze di vigilanza alla Commissione possono essere in parte attenuati dalla partecipazione del comitato europeo per i servizi digitali (di seguito, il "Comitato") ad alcuni dei processi decisionali per la supervisione

[86] Si veda l'art. 17, par. 7, del Trattato sull'Unione europea che disciplina la procedura per la nomina dei membri del collegio dei commissari, all'interno della quale il Parlamento europeo ed il Consiglio dell'Unione europea hanno un ruolo decisivo.

sulle attività di *VLOPs* e *VLOSEs* a norma della Sezione 5 del *DSA*, nonché della Sezione 6 con riferimento ai codici di condotta e ai protocolli di crisi. Il Comitato è infatti un'autorità con un certo livello di indipendenza, e la sua partecipazione alla vigilanza su *VLOPs* e *VLOSEs* potrebbe attenuare, seppur in misura limitata, il rischio che considerazioni politiche influenzino le decisioni della Commissione. Il Comitato è composto da rappresentanti dei coordinatori dei servizi digitali (di seguito, i "CSD"[87]). L'art. 50 del *DSA* obbliga gli Stati membri ad assicurare che i CSD agiscano in piena indipendenza, con sufficiente autonomia per la gestione dei propri bilanci ed adeguate risorse per espletare le proprie funzioni[88]. Pertanto, l'indipendenza dei CSD dovrebbe garantire a sua volta l'indipendenza del Comitato. Tuttavia, l'indipendenza del Comitato potrebbe essere compromessa da due fattori. In primo luogo, gli Stati membri possono prevederne la partecipazione anche da parte di altre autorità nazionali, per le quali il *DSA* non richiede il rispetto di requisiti di indipendenza[89]. In secondo luogo, la Commissione presiede il Comitato e fornisce il proprio accordo per l'adozione del suo regolamento interno[90], anche se non dispone di alcun diritto di voto[91]. Ad ogni modo, il Comitato ha un ruolo di gruppo consultivo, con il compito di fornire consulenza alla Commissione e ai CSD nell'ambito dell'applicazione del *DSA*. Pertanto, anche qualora fosse un organo indipendente a tutti gli effetti, non dispone di poteri di vigilanza, correttivi e sanzionatori. Di conseguenza, la presenza del Comitato non sarebbe in ogni caso sufficiente a rimediare ai rischi derivanti dall'attribuzione di funzioni di vigilanza alla Commissione, sebbene possa in qualche misura attenuare tali rischi, a seconda di come sarà strutturato ed agirà in concreto.

In conclusione, secondo quanto sopra osservato emerge che, nella pratica, il sistema di vigilanza sull'implementazione del *DSA* da parte degli intermediari presenta almeno un punto debole. In particolare, vi sono rischi per la libertà di espressione degli utenti poiché sono attribuite funzioni di vigilanza ad un'istituzione non indipendente, e la partecipazione del Comitato non è sufficiente a controbilanciare il collegamento al potere politico della Commissione.

6. Conclusioni

La strategia co-regolamentare intrapresa dal legislatore europeo nel *DSA* per la disciplina di *VLOPs* e *VLOSEs* può apportare benefici, tra cui la possibilità che le norme si adattino con flessibilità alle fattispecie di rischi sistemici che si presenteranno in futuro. Tuttavia, le osservazioni sopra avanzate mostrano come questi benefici siano accompagnati da chiari rischi per la libertà di espressione degli utenti, e l'intero assetto del *DSA* per la mitigazione dei rischi sistemici appaia non pienamente rispettoso dell'art. 52 della Carta. Da un lato, vi sono evidenti carenze, direttamente percepibili dalla

87 Si veda l'art. 62, par. 1, del *DSA*.

88 Si veda l'art. 50, par. 1 e 2, del *DSA*.

89 *Ibid.*

90 Si veda l'art. 62, par. 2 e 7, del *DSA*.

91 Si veda l'art. 62, par. 3, del *DSA*.

lettura dei rilevanti articoli del *DSA*, in relazione al rispetto del principio di legalità. Dall'altro lato, anche qualora il *DSA* appaia formalmente rispettoso del principio di proporzionalità, un'analisi più approfondita delle sue ricadute applicative evidenzia il rischio di restrizioni non proporzionate, nonché il rischio che le salvaguardie predisposte come controllo sull'operato degli intermediari non garantiscano una tutela effettiva della libertà di espressione degli utenti.

I fornitori di servizi intermediari hanno implementato pratiche di moderazione dei contenuti ancor prima dell'adozione del *DSA*, agendo anche contro contenuti legittimi ma contrari ai loro termini e condizioni come la disinformazione. Una volta che il *DSA* troverà pienamente applicazione, tuttavia, la moderazione di contenuti legittimi potrebbe verificarsi, su larga scala, in attuazione di un obbligo discendente direttamente dalla legge. La Commissione potrebbe assumere un ruolo di fondamentale importanza, in qualità di autorità di vigilanza, per condurre le pratiche dei soggetti vigilati in una direzione che sia, per quanto possibile, rispettosa dei diritti fondamentali degli utenti. Se condotti nella direzione giusta, è possibile che *VLOPs* e *VLOSEs* implementino il *DSA* in modo pienamente rispettoso del principio di proporzionalità, disattendendo le preoccupazioni sopra esposte. Ciononostante, permane il problema che, come sopra esposto[92], la Commissione è parte del potere esecutivo nell'assetto istituzionale della separazione dei poteri dell'Unione, e che un'istituzione soggetta ad influenza politica potrà decidere su delicati aspetti riguardanti l'esercizio della libertà di espressione online. Inoltre, l'interferenza con la libertà di espressione conseguente all'attuazione dell'art. 35 non sembrerebbe rispettosa del principio di legalità, come sopra evidenziato[93]. Resta da vedere quali ricadute applicative avrà il *DSA* e fino a che punto la Commissione riuscirà a guidarne la corretta implementazione.

[92] Si veda il par. 5.4.2.

[93] Si veda il par. 4.2.

Il sistema europeo di cooperazione informativa per il contrasto alle minacce informatiche. Verso una definizione di cybersicurezza integrata?*

Federico Serini

Abstract

La cooperazione informativa rappresenta una delle attività funzionali al processo di integrazione europea, ed in particolare lo scambio di informazioni di sicurezza è espressione della capacità di cooperazione degli Stati membri quale elemento fondante lo "Spazio di libertà, sicurezza e giustizia" dell'Unione europea. L'analisi che si propone in questo contributo vuole studiare l'organizzazione e le procedure di scambio delle informazioni per il contrasto delle minacce informatiche al fine di osservare il processo di integrazione della cybersicurezza europea in corso, avendo modo di affrontare il rapporto tra il settore pubblico e privato variamente coinvolti.

Information cooperation represents one of the functional activities in the process of European integration, and in particular, the exchange of security information is an expression of the member states' ability to cooperate as a founding element of the European Union's "Area of Freedom, Security and Justice". The analysis proposed in this contribution aims to study the organization and procedures of information exchange for countering cyber threats to observe the ongoing European cybersecurity integration process, having a way to address the relationship between the public and private sectors involved.

Sommario

1. Breve premessa di studio: insicurezza di rete, rischio informatico e la cooperazione tra pubblico e privato. - 2. La circolazione delle informazioni negli assetti europei di sicurezza in senso tradizionale. - 3. *A problem shared is a problem halved*. Le origini della *cyber threat information sharing*. - 4. La cooperazione europea di *cyber information sharing* tra soggetti pubblici e privati. - 4.1. L'organizzazione amministrativa delle istituzioni di

* L'articolo è stato sottoposto, in conformità al regolamento della Rivista, a referaggio "a doppio cieco".

cybersicurezza europea. - 4.2. Brevi cenni sulla privatizzazione della (cyber)sicurezza. - 4.3. I partenariati pubblico-privati europei di cybersicurezza. - 5. La *cyber information sharing* alla luce della Direttiva NIS II e delle linee guida ENISA. - 5.1. La progressiva europeizzazione degli strumenti di cooperazione informativa: SOC, Registri delle vulnerabilità, Standard di scambio e Piattaforme di *cyber threat sharing*. - a) I Security Operation Centres (SOCs). - b) I Registri delle vulnerabilità e delle debolezze informatiche. - c) Le piattaforme di Cyber Information Sharing. - 6. La tutela dei diritti fondamentali e della sicurezza nel trattamento delle informazioni "sensibili e classificate" per lo Stato e dei dati personali contenuti nelle informazioni di cybersicurezza. - a) Lo scambio di informazioni "sensibili e classificate" di cybersicurezza e i limiti alla loro circolazione. - b) La protezione europea dei dati personali contenuti nelle informazioni di cybersicurezza. - 7. Considerazioni conclusive sul processo di integrazione della cybersicurezza europea.

Keywords

cyber situational awareness - cybersecurity threat information sharing – cybersicurezza europea - cooperazione pubblico privato – dati personali

1. Breve premessa di studio: insicurezza di rete, rischio informatico e la cooperazione tra pubblico e privato

Le risorse informatiche costituiscono un elemento essenziale per le democrazie. Tali strumenti non rappresentano solo il mezzo che consente agli individui di esprimere liberamente la propria personalità in nuove forme e modi tramite la rete[1] ma, a livello tecnico[2], sono anche i parametri di configurazione e di funzionamento di molte infrastrutture che erogano servizi e funzioni essenziali per la società e l'economia (c.d infrastrutture critiche). Si pensi agli apparati informatici in uso presso gli operatori attivi nei settori bancario e finanziario, energetico, dei trasporti, delle comunicazioni, quello sanitario nonché quelli in dotazione presso le pubbliche amministrazioni e le varie istituzioni statali.

Questi strumenti sono ormai indispensabili sia per lo Stato in sé (apparato), sia per le sue componenti, prime fra tutte gli individui e le imprese (collettività)[3]. Tuttavia, allo stesso tempo, sono anche responsabili di aver trasferito i rischi del cyberspazio nel mondo reale, tanto che nell'attuale contesto informatizzato qualcuno ha avvertito che «ogni società è tanto vulnerabile quanto è vulnerabile l'informatica di cui fa uso» e pertanto «più le società sono avanzate, più sono vulnerabili»[4].

[1] V. Frosini, *La democrazia nel XXI secolo* (1997), Macerata, 2010, 40-41.

[2] C. Gallotti, *I sistemi di gestione per la sicurezza delle informazioni. La norma ISO/IEC 27001:2022 I controlli della ISO/IEC 27002:2022*, Raleigh, 2022.

[3] G. De Vergottini, *Sicurezza e i diritti fondamentali*, in L.E.R. Vega - L. Scaffardi - I. Spigno, *I diritti fondamentali nell'era della digital mass surveillance*, Napoli, 2021, 28.

[4] M.G. Losano, *Guerre ibride, omicidi mirati, droni: conflitti senza frontiere e senza diritto*, in L. Forni - T.

Nonostante tale condizione - secondo cui "rischio informatico=rischio sociale" - porti in evidenza come la tutela e la garanzia dei diritti e delle libertà nell'attuale società tecnologica passi anche per la sicurezza delle reti e dei sistemi informatici[5], i poteri pubblici hanno volto l'attenzione verso questo fenomeno solo di recente (più o meno a partire dalla fine degli anni '90 del secolo scorso), a seguito della progressiva dipendenza degli Stati e delle infrastrutture all'informatica.

La pretesa di sicurezza nel cyberspazio da parte degli Stati si scontra oggi con gli effetti derivanti da questo ritardo. Il cyberspazio è infatti un fenomeno originariamente pubblico, nato con il progetto Arpanet[6], successivamente sviluppato e diffuso per mezzo dei privati, fuori dal controllo degli Stati[7]. Non è un caso se le prime definizioni di cybersicurezza (*cybersecurity*), sicurezza informatica (*computer security*) e sicurezza delle informazioni (*information security*) hanno trovato formulazione all'interno del c.d. "diritto dei privati"[8], nello specifico in normative tecniche di settore[9].

Tuttavia, se «[n]ell'ambiente digitale sembra non esserci più Stato, territorio, sovranità e neppure popolo, ma produzione principalmente privata del diritto»[10], non è dovuto solo perché il potere pubblico è arrivato "dopo", ma soprattutto perché l'oggetto della pretesa normativa, il cyberspazio, è un fenomeno globale privo di territorialità, che rappresenta un limite all'azione del potere pubblico che invece vanta «un'originaria

Vettor (a cura di), *Sicurezza e libertà in tempi di terrorismo globale*, Torino, 2017, 22. Sugli effetti delle forme di connettività non solo dovute alle tecnologie ICT v. A.L. Baràbasi, *Linked. How everything is connected to everything else and what it means for business, science, and everyday life*, New York, 2014. Analogamente v. anche P. Khanna, *Connectography. Le mappe del futuro ordine mondiale*, Roma, 2016.

5 Cfr. M. Dunn Cavelty, *Breaking the Cyber-Security Dilemma: Aligning Security Needs and Removing Vulnerabilities*, in *Science and Engineering Ethics,* 20, 2014, 704.

6 M. O'Mara, *The Code: Silicon Valley and the Remaking of America,* Londra, 2019.

7 G. Bombelli, *Dal moderno all'"ultramoderno"? Intorno al nesso diritto-tecnica-sicurezza*, in F. Pizzolato - P. Costa (a cura di), *Sicurezza e tecnologia*, Milano, 2017, 26.

8 Con la Raccomandazione ITU-T X.1205, del 18 aprile 2008, l'*International Telecommunication Union* (ITU) ha definito la cybersecurity come l'insieme degli strumenti politici, giuridici e tecnologici che hanno la finalità di proteggere il cyber environment e gli asset degli utenti dai cyber rischi, ed in particolare di garantire le tre priorità della riservatezza (*confidentiality*), integrità (*integrity*) e disponibilità (*availability*) degli stessi. Altra definizione è invece contenuta nella norma tecnica ISO/IEC 27032 ove la cybersicurezza è considerata come azione volta alla «preservation of confidentiality, integrity and availability of information in the Cyberspace».

9 Cfr. O.W. Cesarini, *Il diritto dei privati*, Milano, 1963. Sul punto v. anche S. Romano, *L'ordinamento giuridico* (1918), Macerata, 2018. In particolare, sulla normativa tecnica e il relativo processo di "normalizzazione" v. P. Aandreini - G. Caia - G. Elias - F.A. Roversi-Monaco, *La normativa tecnica industriale. Amministrazione e privati nella normativa tecnica e nella certificazione dei prodotti industriali*, Bologna, 1995; F. Salmoni, *Le norme tecniche*, Milano, 2001, 147; E. Chiti, *La normalizzazione*, in S. Cassese (a cura di), *Trattato di diritto amministrativo*, vol. IV, Milano, 2003, 4027 ss.; H. Schepel, *The Constitution Of Private Governance: Product Standards In The Regulation Of Integrating Markets*, Londra, 2005; A. Zei, *Tecnica e diritto tra pubblico e privato*, Milano, 2008; A. Moscarini, *Fonti dei privati e globalizzazione*, Roma, 2015; O. Kanevskaia, *Governance within standards development organizations: WHO owns the game?*, in *ITU Kaleidoscope Academic Conference 2017*, Nanjing, 2017, 1-8; A. Iannuzzi, *Il diritto capovolto. Regolazione a contenuto tecnico-scientifico e Costituzione*, Napoli, 2018. In particolare, sull'evoluzione storica della normazione tecnica nei settori della *computer* e *information security*, v. D. Russell - G.T. Gangemi, *Computer security basics*, Sebastopol, 1991, 23.

10 E. Cremona, *I poteri privati nell'era digitale. Libertà costituzionali, regolazione del mercato, tutela dei diritti*, Napoli, 2023, 45.

necessità dei luoghi»[11].

In realtà, come rilevato dalla letteratura sul punto[12], il cyberspazio è una dimensione caratterizzata dalla convivenza di componenti immateriali, quali le connessioni, spettri elettromagnetici e protocolli di funzionamento, certamente non riconducibili ad alcuno spazio fisico; e componenti materiali, ossia le tecnologie fisiche, come cavi, *routers* e *switch*, localizzate entro i confini degli Stati, e generalmente prodotti da attori privati attivi nel mercato delle telecomunicazioni[13].

La morfologia appena tratteggiata rende evidente come in entrambo i piani l'azione pubblica per fini di sicurezza del cyberspazio richieda la necessaria cooperazione con i soggetti privati: relativamente al profilo immateriale, per tentare di regolare ciò che avviene "nel" cyberspazio, ossia le condotte e i comportamenti degli utenti, tra i quali possiamo individuare anche le minacce informatiche (*security*); per quanto riguarda il lato materiale, per tentare di garantire la sicurezza "del" cyberspazio attraverso la creazione e lo sviluppo di prodotti e soluzioni sul mercato che siano *(cyber)security by design*, in modo da garantire la progressiva sicurezza dell'ambiente digitale (*safety*)[14].

Lo scambio delle informazioni per il contrasto alle minacce informatiche è una pratica utile sia per i fini di *security* che di *safety* del cyberspazio. Pertanto, lo studio che si propone nei prossimi paragrafi sulle procedure di *cyber threat information sharing* rappresenta una privilegiata prospettiva di analisi dei rapporti tra il settore pubblico e privato coinvolti nei recenti cambiamenti che stanno interessando le politiche di cybersicurezza europea.

2. La circolazione delle informazioni negli assetti europei di sicurezza in senso tradizionale

Tra le attività funzionali al processo di integrazione europeo[15] ha assunto sempre maggiore importanza lo scambio di informazioni tra i diversi Stati membri, nonché tra quest'ultimi e le istituzioni europee, al fine di favorire il coordinamento delle attività amministrative dell'Unione in virtù dei principi di leale cooperazione, di cui all'art. 4, par. 3, del Trattato sull'Unione europea (TUE)[16], e di sussidiarietà, riconosciuto all'art.

---

[11] N. Irti, *Norma e luoghi*, Roma-Bari, 2006, 4.

[12] Secondo lo studioso F. D. Kramer esistono 28 differenti definizioni del termine cyberspace. Cfr. Id., *Cyberpower and National Security: Policy Recommendations for a Strategic Framework*, in F.D. Kramer - S. Starr - L.K. Wentz, *Cyberpower and National Security*, Washington (D.C.), National Defense University Press,, 2009. Su tutti, per quel che qui interessa, il riferimento è alla definzione di Martin C. Libicki definisce il cyberspazio individuando tre livelli: fisico, sintattico e semantico M.C. Libicki, *Cyberdeterrence e cyberwar*, Santa Monica, 2009, 11 ss.

[13] G. Suffia, *Geografia delle cyberwars*, Milano, 2018.

[14] Sulla distinzione tra *safety* e *security* si rinvia a M. Durante, *Safety and Security in the Digital Age. Trust, Algorithms, Standards, and Risks*, in D. Berkich - M. d'Alfonso (a cura di), *On the Cognitive, Ethical, and Scientific Dimensions of Artificial Intelligence. Philosophical Studies Series*, 134, Berlin, 2019.

[15] G. de Búrca, J.H.H. Weiler, *The worlds of European constitutionalism*, New York, 2012.

[16] Il principio di leale cooperazione viene in rilievo soprattutto quando la realizzazione di un obiettivo dei Trattati richiede un esercizio coordinato delle competenze nazionali e di quelle dell'Unione. Tuttavia, oltre alla accezione bilaterale, tale principio risulta essere stato invocato anche nel rapporto tra Stati

5 TUE.

Si faccia riferimento al sistema di scambio delle informazioni per la sicurezza stradale, attuato con la direttiva (UE) 2015/413[17], o all'*Electronic Exchange of Social Security Information* (EESSI), il sistema informatico volto a supportare gli enti previdenziali degli Stati membri nello scambio rapido e sicuro di informazioni previdenziali[18], nonché al sistema di cooperazione amministrativa sicura tra le autorità fiscali nazionali istituito con la direttiva 2011/16/CE, recentemente modificata con direttiva di modifica (UE) 2021/514 (comunemente nota come «DAC7»).

Tuttavia, nello specifico caso delle politiche di sicurezza, lo scambio di informazioni tra le autorità di polizia e di *intelligence* degli Stati membri, e tra questi e le istituzioni europee, è un'attività che assume particolare rilevanza in quanto espressione della capacità di cooperazione degli Stati membri quale elemento fondante lo "Spazio di libertà, sicurezza e giustizia" dell'Unione europea[19].

Il tema è stato oggetto di recente attenzione da parte dell'Unione come è possibile apprendere dalla nuovo politica europea sulla sicurezza in generale, la *Security Union Strategy 2020-2025*, ove, a nostro modo di vedere, si è prospettato un inedito processo integrativo in questo settore, dato il riferimento al fatto che:

[a]nche se la responsabilità primaria della sicurezza incombe ai singoli Stati membri, negli ultimi anni è emerso chiaramente che la sicurezza di uno Stato membro è la sicurezza di tutti. L'UE può apportare una risposta multidisciplinare e integrata, fornendo agli operatori della sicurezza negli Stati membri gli strumenti e le informazioni di cui hanno bisogno[20].

Come noto, il sistema di sicurezza europeo si è sviluppato nel tempo sulla logica della cooperazione intergovernativa, non trovando mai una piena comunitarizzazione. L'esclusiva competenza degli Stati membri in materia di sicurezza, vedi la presenza nei Trattati delle clausole di tutela della "sicurezza nazionale" o dell'"ordine pubblico e della sicurezza" quali condizioni legittimanti il regime eccezionale statale rispetto all'applicazione del diritto europeo, è uno dei tratti caratterizzanti la politica europea in questo ambito[21].

---

membri al fine di una più corretta applicazione del diritto UE v. *ex multis*, CGUE C-372/02, *Adanez Vega* (2002); C-105/94, *Celestini* (1994), nonché in relazione al rapporto tra istituzioni e al rispetto delle relative competenze v. CGUE C-65/93, *Parlamento c. Consiglio* (1993).

[17] Direttiva (UE) 2015/413 del Parlamento europeo e del Consiglio, dell'11 marzo 2015, intesa ad agevolare lo scambio transfrontaliero di informazioni sulle infrazioni in materia di sicurezza stradale Testo rilevante ai fini del SEE, che ha sostituito la direttiva 2011/82/UE annullata dalla Corte di giustizia dell'Unione europea con sentenza del 6 maggio 2014. L'obiettivo della direttiva è quello di attivare meccanismi cooperativi operativi tra gli Stati con l'intento di porre fine all'anonimato dei conducenti non residenti e perseguire le infrazioni al codice della strada commesse in uno Stato membro diverso da quello in cui il veicolo è stato immatricolato.

[18] Il sistema *Electronic Exchange of Social Security Information* (EESSI) è stato implementato nell'ambito delle politiche di coordinamento della sicurezza sociale dell'Unione europea ed entrato pienamente in funzione nel giugno 2022. Per maggiori approfondimenti si rinvia alla pagina dell'EESSI.

[19] Cfr. Titolo V del TFUE rubricato per l'appunto «Spazio di libertà, sicurezza e giustizia».

[20] The EU Security Union Strategy 2020-2025, COM(2020) 605 final, del 24 luglio 2020.

[21] Vale la pena richiamare il contenuto dell'art. 4, par. 2, del Trattato sull'Unione europea (TUE) ove è previsto che «L'Unione rispetta l'uguaglianza degli Stati membri davanti ai trattati e la loro identità nazionale insita nella loro struttura fondamentale, politica e costituzionale, compreso il sistema delle

Tuttavia, nonostante tali prerogative che accentrano il ruolo degli Stati membri, tale indirizzo non ha precluso la successiva elaborazione di politiche e la creazione di istituzioni comuni che hanno inevitabilmente richiesto l'impegno coordinato e cooperativo sia tra gli Stati membri[22], sia tra questi e le competenti autorità europee[23], al fine di garantire l'esigenza di sicurezza in tutto lo spazio europeo[24].

In particolare, ripercorrendo brevemente l'evoluzione storica della cooperazione tra gli Stati europei nel settore di polizia, ci si accorge che è proprio nella raccolta, archiviazione, trattamento e scambio di informazioni che trova concreta realizzazione il processo integrativo in questo settore[25].

Proponiamo pertanto alcune tappe salienti di questo processo in modo da porter svolgere le successive considerazioni relativamente alla materia della cybersicurezza.

Nel 1975 i Ministri degli Interni e della Giustizia dei Paesi allora membri della CEE decisero di riunirsi nel forum che prende il nome di Gruppo TREVI con l'intento di costruire formule più intense di cooperazione tra le forze di polizia al di fuori della

autonomie locali e regionali. Rispetta le funzioni essenziali dello Stato, in particolare le funzioni di salvaguardia dell'integrità territoriale, di mantenimento dell'ordine pubblico e di tutela della sicurezza nazionale. In particolare, la sicurezza nazionale resta di esclusiva competenza di ciascuno Stato membro». Tale clausola di competenza statale permanente, aggiunta su esplicita richiesta del Regno Unito, deve essere inoltre letta in combinato disposto con l'art. 276 del Trattato sul funzionamento dell'Unione europea (TFUE), che esclude il controllo da parte della Corte di Giustizia sulla «validità o la proporzionalità di operazioni condotte dalla polizia o da altri servizi incaricati dell'applicazione della legge di uno Stato membro o l'esercizio delle responsabilità incombenti agli Stati membri per il mantenimento dell'ordine pubblico e la salvaguardia della sicurezza interna».

[22] Per una rapida ricostruzione dei rapporti tra gli Stati diretti alla realizzazione di un sistema di sicurezza cooperativo v. M.M. Winkler, *Attività europea di intelligence, cooperazione di polizia e diritti umani*, in U. Draetta - N. Parisi - D. Rinoldi (a cura di), *Lo «spazio di libertà sicurezza e giustizia» dell'Unione europea. Principi fondamentali e tutela dei diritti*, Napoli, 2007, 293 ss; C. Mosca, *La sicurezza come diritto di libertà. Teoria generale delle politiche di sicurezza*, Torino, 2012, 319 ss; R. Ursi, *La sicurezza pubblica*, Bologna, 2022, 219 ss.

[23] Sul punto v. E. Chiti, *Le sfide della sicurezza e gli assetti nazionali ed europei delle forze di polizia*, in L. Forni - T. Vettor (a cura di), *Sicurezza e libertà in tempi di terrorismo globale,* cit.*,* 73, ove l'A. paragonando il processo integrativo europeo in materia di sicurezza, rispetto all'integrazio diretta alla realizzazione del mercato comune, scrive «[...] l'integrazione procede con particolare cautela, sconta la riluttanza dei governi nazioanli a rinunciare a un controllo sostanziale su queste politiche, si realizza attraverso strumenti più leggeri di quelli utilizzati per la costruzione del mercato interno e lo svolgimento delle politiche di regolazione economica e sociale».

[24] Oltre al concetto di ordine pubblico proprio di ogni identità nazionale che compone l'Unione, sulla scorta della giurisprudenza della Corte di giustizia è andato formandosi anche il concetto di ordine pubblico europeo. Sul punto v. G. Calesini, *Diritto europeo di polizia*, Roma, 2007, 211 ss.; D. Rinoldi, *L'ordine pubblico europeo*, Napoli, 2005; I.d., *Ordine pubblico europeo e spazio giuridico continentale*, in U. Draetta - N. Parisi - D. Rinoldi, *Lo «spazio di libertà sicurezza e giustizia» dell'Unione europea,* cit.*,* 61 ss.

[25] Nella elencazione dei principali atti legislativi sulla cooperazione di polizia disponibile presso il sito del Parlamento europeo (consultato il 2 dicembre 2023) risulta che gran parte di questi siano volti a istituire meccanismi di comunicazione per favorire lo scambio di informazioni tra i Paesi membri, v. la direttiva (UE) 2016/681 sull'uso dei dati del codice di prenotazione (PNR) a fini di prevenzione, accertamento, indagine e azione penale nei confronti dei reati di terrorismo e dei reati gravi; il regolamento (UE) 2018/1862 sull'istituzione, l'esercizio e l'uso del sistema d'informazione Schengen (SIS) nel settore della cooperazione di polizia e della cooperazione giudiziaria in materia penale; il regolamento (UE) 2019/818 che istituisce un quadro per l'interoperabilità tra i sistemi di informazione dell'UE nel settore della cooperazione di polizia e giudiziaria, asilo e migrazione; la direttiva (UE) 2019/1153 che reca disposizioni per agevolare l'uso di informazioni finanziarie e di altro tipo a fini di prevenzione, accertamento, indagine o perseguimento di determinati reati; il regolamento (UE) 2021/784 relativo al contrasto della diffusione di contenuti terroristici online, applicabile dal 7 giugno 2022.

precedente esperienza internazionale dell'Interpol. Come si apprende dal "Programma d'azione relativo al rafforzamento della cooperazione in materia di polizia e di lotta al terrorismo o altre forme di criminalità organizzata" del 1990, inizialmente il principale ruolo del Gruppo era stato quello di garantire il controllo delle frontiere esterne anche grazie alla installazione di linee di comunicazione dirette.

Non è un caso se pochi giorni dopo l'adozione del citato Programma, venne sottoscritta la Convenzione di applicazione dell'Accordo di Schengen del 1985 con il quale si addiveniva alla istituzione di un sistema avanzato di cooperazione transfrontaliera e di scambio delle informazioni attraverso l'istituzione del «Sistema di informazione Schengen» (SIS)[26].

Tuttavia, fino a questo momento la cooperazione intergovernativa era avvenuta al di fuori dell'ordinamento europeo, ma le due esperienze citate rappresentarono uno modello per la successiva inclusione della cooperazione di polizia nel diritto comunitario così come organizzata nel Trattato di Maastricht del 1992[27]. In particolare, per ciò che qui interessa, nella dichiarazione n. 32 allegata al Trattato, all'art. K.1, n. 9, venivano specificate le modalità concrete di cooperazione ove si prevedeva che gli Stati si impegnano alla realizzazione di forme di collaborazione incentrate sullo «scambio di informazioni e di esperienze» nell'ambito dell'assistenza alle autorità nazionali incaricate delle azioni penali in materia penale e della sicurezza.

Principio che ha poi trovato concreta applicazione con due decisioni quadro che hanno reso l'*information sharing* vincolante per gli Stati membri: la decisione quadro 2008/960/Gai, che impone, all'interno di un quadro giuridico definito, lo scambio rapido ed efficace di informazioni tra le autorità deputate alla sicurezza, e la decisione-quadro 2008/615/Gai che dispone l'accesso reciproco alle informazioni anche attraverso l'interoperabilità di basi di dati nazionali.

Sempre all'interno del Trattato di Maastricht, all'art. K.1, n. 9, si faceva riferimento al primo Ufficio di polizia europeo, l'Europol, che venne poi istituito nel 1993 e divenuto pienamente operativo nel 1999. Anche in questo caso, come è possibile dedurre dall'art. 88 del Trattato sul funzionamento dell'Unione europea (TFUE), tra i compiti dell'Ufficio troviamo: «a) la raccolta, l'archiviazione, il trattamento, l'analisi e lo scambio delle informazioni trasmesse, in particolare dalle autorità degli Stati membri o di paesi o organismi terzi; b) il coordinamento, l'organizzazione e lo svolgimento di indagini e di azioni operative, condotte congiuntamente con le autorità competenti degli Stati membri o nel quadro di squadre investigative comuni, eventualmente in collegamento con Eurojust».

Dal breve quadro descritto si evince come all'iniziale approccio intergovernativo, ba-

[26] Dalla sua istituzione il sistema SIS è stato oggetto di diversi interventi di aggiornamento, di cui l'ultimo è avvenuto con il regolamento (UE) 2018/1862. Tuttavia, la sua architettura tecnica è rimasta invariata nel tempo: il SIS risulta composto da una struttura nazionale (N-SIS) presente in ciascuno Stato membro, e duna una unità centrale (C-SIS) con sede a Strasburgo, la quale si occupa di gestire ed elaborare i dati inviati alle banche dati nazionali. Il coordinamento tra i diversi Paesi è favorito dalla presenza negli uffici nazionali del sistema informativo di un ulteriore ufficio SIRENE (*Supplementary Information Request at the National Entry*) che svolge le funzioni di sala operativa nell'intero arco della giornata allo scopo di consentire alle autorità di polizia e giudiziaria dei vari Paesi Schengen di acquisire ulteriori informazioni non disponibili in base al sistema SIS.

[27] Cfr. R. Ursi, *La sicurezza pubblica*, cit., 227.

sato sullo scambio "diretto" o a "rete" delle informazioni tra le autorità dei singoli Stati membri, si sia progressivamente affiancato il modello di cooperazione informativa accentrato, imperniato su banche dati comuni europee o comunque sistemi informativi gestiti da organismi sovranazionali, che trova oggi riconoscimento all'art. 87, par. 2, lett. a) del TFUE[28].

In particolare, secondo Alcuni, il grande cambiamento che contraddistingue il nuovo approccio da quello precedente, non starebbe tanto negli strumenti e nelle modalità impiegate, quanto piuttosto nel mutato «contesto» nel quale si articola questo tipo di cooperazione ispirata al «principio di disponibilità delle informazioni» e dal quale discende a livello pratico «l'accesso reciproco e l'interoperabilità delle banche dati-nazionali, nonché l'accesso diretto (*on-line*) alle banche-dati dell'Unione da parte di autorità nazionali ed europee»[29].

Da ultimo, alla luce delle discrepanze tra la decisione quadro 2006/960/GAI e l'ambito di applicazione della convenzione che attua l'Accordo di Schengen, il legislatore europeo ha recentemente introdotto la direttiva (EU) 2023/977, relativa allo scambio di informazioni tra le autorità di contrasto degli Stati membri e che abroga la suddetta decisione quadro[30].

Brevemente, è possibile notare come tale iniziativa sia volta da una parte, a potenziare i punti di contatto unici stabiliti presso gli Stati membri, dall'altra a favorire la convergenza verso l'utilizzo da parte di tutte le autorità di contrasto della rete *Secure Information Exchange Network Application* (SIENA), gestita e sviluppata da Europol conformemente al regolamento (UE) 2016/794, al fine di «porre rimedio al problema della proliferazione dei canali di comunicazione utilizzati per la trasmissione di informazioni sull'attività di contrasto tra gli Stati membri, poiché [tale pratica] ostacola lo scambio adeguato e rapido di tali informazioni e aumenta i rischi per la sicurezza dei dati personali»[31].

Fanno tuttavia eccezione i casi di cui all'art. 13, c. 2, della direttiva ove è previsto che gli Stati membri possono consentire al loro punto di contatto unico o alle loro autorità di contrasto competenti di non avvalersi di SIENA nei casi in cui:

a) lo scambio di informazioni richiede il coinvolgimento di paesi terzi od organizzazioni internazionali o vi sono ragioni obiettive per ritenere che tale coinvolgimento sarà necessario in una fase successiva, anche attraverso il canale di comunica-

[28] F. Peroni - M. Gialuz (a cura di), *Cooperazione informativa e giustizia penale nell'Unione europea*, Trieste, 2009, 10.

[29] N. Parisi, *Cooperazione fra le autorità nazionali ed europee incaricate "dell'applicazione delle legge" nello spazio di libertà, sicurezza e giustizia. I principi fondanti la circolazione internazionale delle informazioni*, in R. Del Coco - E. Pistoria (a cura di), *Stranieri e giustizia penale. Problemi di perseguibilità e di garanzie nella normativa nazionale ed europea*, Bari, 2014, 122-123.

[30] Direttiva (UE) 2023/977, relativa allo scambio di informazioni tra le autorità di contrasto degli Stati membri e che abroga la decisione quadro 2006/960/GAI del Consiglio. In particolare, sui motivi che hanno portato a tale intervento legislativo a livello europeo si faccia riferimento ai considerando 7 e 8 della direttiva. Inoltre al considerando 13 è ribadito che «[p]oiché la presente direttiva non si applica al trattamento di informazioni nell'ambito di un'attività che non rientra nel campo di applicazione del diritto dell'Unione, le attività concernenti la sicurezza nazionale non rientrano nel campo di applicazione della presente direttiva».

[31] Cfr. considerando 26, direttiva (UE) 2023/977.

zione Interpol;

b) l'urgenza della richiesta di informazioni richiede l'uso temporaneo di un altro canale di comunicazione;

c) un incidente tecnico od operativo imprevisto impedisce al loro punto di contatto unico o alle loro autorità di contrasto competenti di utilizzare SIENA per lo scambio di informazioni.

Merita inoltre osservare che l'attenzione posta dall'Unione verso la circolazione delle informazioni per la cooperazione di polizia - ulteriormente amplificata a seguito degli attacchi terroristici di inizio millennio[32] - non si è limitata solo alla promozione degli scambi informativi e alla creazione di reti e banche dati centralizzate, ma si è diretta anche verso il profilo, di non secondaria rilevanza, della protezione di dette infrastrutture informatiche preposte alla circolazione del materiale informativo[33].

3. *A problem shared is a problem halved*. Le origini della *cyber threat information sharing*

L'esigenza di una cooperazione e di un coordinamento delle attività di contrasto alle minacce informatiche su larga scala si è avvertita per la prima volta negli Stati Uniti il 2 novembre 1988, in occasione di uno dei primi attacchi informatici a vasto impatto: il "Morris Worm", creato da Robert Tappan Morris, studente della Cornell University con lo scopo di dimostrare le inadeguatezze delle misure di sicurezza delle reti informatiche[34].

Sebbene il *malware* venne creato da Morris a soli fini di studio, l'esperimento uscì fuori dal controllo del suo creatore riuscendo a contagiare circa il 10% dei computer del mondo[35].

---

[32] E. Chiti, *Le sfide della sicurezza e* …, in L. Forni - T. Vettor (a cura di), *Sicurezza e libertà in tempi di terrorismo globale,* cit., 77 ss. L'A. sottolinea come a seguito degli eventi dell'11 settembre vi sia stata una apertura globale delle amministrazioni di polizia e militari dell'Unione europea.

[33] Si faccia riferimento a: la direttiva 2013/40/UE relativa agli attacchi contro i sistemi di informazione e che sostituisce la decisione quadro 2005/222/Gai del Consiglio, nota come direttiva sulla criminalità informatica; la direttiva (UE) 2016/1148 recante misure per un livello comune elevato di sicurezza delle reti e dei sistemi informativi nell'Unione, nota come direttiva NIS, oggi abrogata e sostituita dalla direttiva (UE) 2022/2555, c.d. NIS II; regolamento (UE) 2018/1726 relativo all'Agenzia dell'Unione europea - eu-LISA - per la gestione operativa dei sistemi IT su larga scala nello spazio di libertà, sicurezza e giustizia. Nonché, infine, la stessa rete SIENA che assicura la sicurezza delle trasmissioni.

[34] Le motivazioni che portarono il giovane studente a sviluppare il primo prototipo di "worm" possono essere ricostruire dalla lettura della sentenza della Corte d'appello US, v. United States Court of Appeals, No. 774, Docket 90-1336, Argued Dec. 4, 1990. Decided March 7, 1991, *United States of America, Appellee, v. Robert Tappan Morris, Defendant-Appellant*. In particolare, circa la il potere pervasivo del programma sviluppato da Morris, dalla ricostruzione dei fatti nella pronuncia si apprende che: «[t]he tactic he selected was release of a worm into network computers. Morris designed the program to spread across a national network of computers after being inserted at one computer location connected to the network. Morris released the worm into Internet, which is a group of national networks that connect university, governmental, and military computers around the country. The network permits communication and transfer of information between computers on the network».

[35] Dal sito della *Federal Bureau of Investigation* (FBI) statunitense, si apprende che «[w]ithin 24 hours, an estimated 6,000 of the approximately 60,000 computers that were then connected to the Internet had

La portata dell'incidente richiese un'intensa collaborazione internazionale che riscontrò pesanti limitazioni a causa della mancata implementazione di meccanismi di condivisione delle informazioni sulle minacce informatiche[36].

A pochi giorni dall'evento venne istituito negli Stati Uniti il primo Gruppo di intervento operativo in caso di attacco informatico - il *Computer Emergency Response Team* (CERT) - con l'obiettivo di diffondere le notifiche sugli incidenti e coordinare la comunicazione durante l'emergenza. Per quanto riguarda la natura di questo primo "centro di soccorso informatico", si precisa che si trattava di una iniziativa nata all'interno delle accademie e dei centri di ricerca, più nello specifico dall'agenzia statunitense DARPA (*Defence Advanced Research Projects Agency*), la quale istituì tale Gruppo presso la Carneige Mellon University di Pittsburgh, in Pennsylvania, che ne detiene ancora oggi la proprietà del marchio[37].

L'insegnamento tratto dalla drammatica esperienza evidenziò così come nel contesto della sicurezza informatica non sia solo necessario un approccio di tipo reattivo (*ex post*) legato all'emergenza in corso, ma anche un approccio di sicurezza preventivo (*ex ante*), attraverso il monitoraggio e l'analisi costante delle risorse informatiche in uso, al fine di rilevare intrusioni e anomalie in tempo reale (c.d. *cyber situational awareness*)[38].

Nel 1997, dopo i primi attacchi terroristici al World Trade Center (1993) e a Oklahoma City (1995), il Presidente Clinton nominò la Commissione per la protezione delle infrastrutture critiche al fine di individuare le possibilità di cooperazione tra il settore pubblico e quello privato per proteggere adeguatamente le infrastrutture critiche degli Stati Uniti. La Commissione si espresse con un rapporto finale, il "rapporto Marsh", che tra le principali raccomandazioni prevedeva l'istituzione dei *Information Sharing and Analysis Centres* (ISACs): partenariati pubblico-privati senza scopo di lucro, organizzati con lo scopo di raccogliere informazioni sulle minacce informatiche provenienti dai

been hit. Computer worms, unlike viruses, do not need a software host but can exist and propagate on their own».

[36] A. Contaldo - F. Peluso, *Cybersecurity. La nuova disciplina italiana ed europea alla luce della direttiva NIS*, Pisa, 2018, 70 ss.

[37] I. Skierka - R. Morgus - M. Hohmann - T. Maurer, *CSIRT Basics for Policy-Makers. The History, Types & Culture of Computer Security Incident Response Teams*, New America and the Global Public Policy Institute (GPPi), maggio 2015, 9 ss. CERT, acronimo di *Computer Emergency Response Team*, è una sigla spesso utilizzata al posto di CSIRT (*Computer Security Incident Response Team*). In entrambo i casi si tratta di una denominazione volta a descrivere un gruppo di intervento incaricato di monitorare gli incidenti a livello nazionale; emettere preallarmi, allerte, annunci e divulgazione di informazioni alle parti interessate in merito a rischi e incidenti; intervenire in caso di incidente; analizzare dinamicamente i rischi e gli incidenti; svolgere attività di sensibilizzazione situazionale. La distinzione tra i due acronimi è dovuta ad una questione di mero diritto dei marchi, in quanto il CERT nacque su iniziativa dell'agenzia statunitense DARPA (*Defence Advanced Research Projects Agency*), la quale istituì tale Gruppo presso la *Carneige Mellon University* di Pittsburgh in Pennsylvania che ne detiene ancora oggi la proprietà del marchio, diversamente dal nominativo CSIRT che invece è di libero utilizzo. Sul punto v. A. Contaldo - F. Peluso, *Cybersecurity,* cit.; L. Salandri - A. Contaldo, *La nuova disciplina giuridica c.d. "orizzontale" della cybersicurezza per le infrastrutture in un'ottica di sviluppo dei sistemi informativi*, in *Riv. amm.*, 2016, 567-595.

[38] Termine migrato dal linguaggio militare. Nel particolare contesto della sicurezza informatica, secondo l'Agenzia governativa statunitense che si occupa della gestione delle tecnologie, il *National Institute of Standards and Technology* (NIST), per "*cyber situational awareness"* deve intendersi «[p]erception of elements in the system and/or environment and a comprehension of their meaning, which could include a projection of the future status of perceived elements and the uncertainty associated with that status». Sul punto v. S. Jajodia - P. Liu - V. Swarup - C. Wang, *Cyber situational awareness*, New York, 2009.

CERTs e SOCs (c.d. informazioni di cybersicurezza) - in un primo momento solo quelle veicolate ai danni delle infrastrutture critiche - al fine di condividerle all'interno di una rete di soggetti fidati che partecipano allo scambio informativo su base volontaria[39].

Nel 2015, sempre con lo scopo di promuovere e facilitare la condivisione delle informazioni di cybersicurezza, il Presidente Obama per mezzo di un *executive order* istituì le *Information Sharing and Analysis Organizations* (ISAOs): organizzazioni create per raccogliere, analizzare e diffondere informazioni sulle minacce informatiche non direttamente legate ai settori delle infrastrutture critiche, estendendo così il meccanismo di condivisione anche alle piccole e medie imprese operative in diversi settori (vi rientrano ad esempio studi legali, contabili e di consulenza che supportano clienti intersettoriali, ecc.)[40].

Lo stesso anno venne anche emanato il *Cybersecurity Information Sharing Act* (c.d. CISA bill), la legge federale con il quale gli Stati Uniti hanno inteso regolare e promuovere lo scambio di informazioni relative alle minacce informatiche tra il settore privato e le istituzioni governative[41].

I benefici riscontrati dall'introduzione di queste procedure di scambio evolute nell'esperienza statunitense hanno ben presto portato all'ingresso di altri Paesi in tali ecosistemi informativi, fino al suo recepimento all'interno degli accordi sovranazionali.

Difatti, il meccanismo di *cyber information sharing* costituisce ormai un principio fondamentale nei rapporti internazionali[42], quale corollario del concetto di sicurezza cooperativa[43].

Sul punto, vale la pena ricordare che nelle linee guida in materia di sicurezza dei sistemi e delle reti d'informazione elaborate dall'Organizzazione per la cooperazione e lo sviluppo economico (OCSE) nel 2002[44], tra i nove principi prodromici all'instaurazione di una cultura della sicurezza, trova spazio anche il principio di "risposta" secondo il quale «[l]e parti interessate devono operare tempestivamente e in uno spirito di cooperazione per prevenire, rilevare e rispondere agli incidenti di sicurezza». Ed in particolare, queste sono chiamate a «scambiare, in maniera adeguata, le informazioni di cui

---

[39] Sulle origini storiche delle ISACs v. ENISA, *Information Sharing and Analysis Centres (ISACs) Cooperative models,* 2018.

[40] Come emerge dal contenuto dell'*executive order* istitutivo delle *Information Sharing and Analysis Organizations* (ISAOs), si tratta di organizzazioni operative in diversi settori, quali «private companies, nonprofit organizations, executive departments and agencies (agencies), and other entities», istituite con il fine di condividere le informazioni relative ai rischi e agli incidenti di sicurezza informatica e collaborare per rispondere in maniera più vicina possibile e in tempo reale. Il funzionamento a livello tecnico di tali organizzazioni intersettoriali è supportato dalla *ISAO Standard organization* definita dall'*executive order* come «non-governmental organisation [...] to improve the US's cybersecurity posture by identifying standards and guidelines for robust and effective information sharing and analysis related to cybersecurity risks, incidents, and best practices».

[41] Cybersecurity Information Sharing Act of 2014, S. 2588.

[42] T. Hitchens - N. Goren, *International Cybersecurity Information Sharing Agreements*, in *Center for International & Security Studies*, U. Maryland, 2017.

[43] R. Cohen, *Cooperative Security: From individual Security to International Stability*, in *Marshall Center Papers*, 3, aprile 2001.

[44] OCSE, *Guidelines for the Security of Information Systems and Networks: Towards a Culture of Security*, 2002.

dispongono sulle minacce e vulnerabilità e devono creare procedure per una rapida ed efficace cooperazione volta a prevenire e a rilevare gli incidenti di sicurezza e a rispondervi» tale che «[c]iò potrebbe comportare scambi d'informazioni e una cooperazione transfrontaliera, ove autorizzato».
Principio che ha ovviamente trovato ospitalità anche nelle politiche di cybersicurezza dell'Unione europea declinato sotto il profilo sia organizzativo, attraverso l'articolazione di una rete amministrativa *ad hoc*, sia normativo, attraverso l'imposizione di obblighi relativi alla gestione della sicurezza informatica che ricomprendono, tra i diversi, la notifica degli incidenti di sicurezza alle competenti autorità.

4. La cooperazione europea di *cyber information sharing* tra soggetti pubblici e privati

La *cyber (trheat) information sharing* è il termine con il quale generalmente si indicano le attività consistenti nello «scambio di una varietà di informazioni relative alla sicurezza delle reti e delle informazioni, quali rischi, vulnerabilità, minacce e problemi di sicurezza interna, nonché buone prassi»[45].
Nel paragrafo precedente è stato evidenziato come questa pratica sia sorta da una obiettiva necessità di contrastare le minacce informatiche attraverso le azioni coordinate di Gruppi di intervento dislocati nel mondo e le infrastrutture critiche. In particolare, si è posto in luce come i primi ecosistemi di scambio informativo delle minacce informatiche siano sorti sulla scorta di iniziative autonome da parte degli stessi soggetti interessati (perlopiù infrastrutture critiche di natura privata). Nei successivi sottoparagrafi saranno analizzate l'architettura europea di cybersicurezza (par. 4.1), avendo modo di riflettere anche sul ruolo dei soggetti privati, particolarmente presenti in questo settore (par. 4.2), e il relativo rapporto tra questi e il settore pubblico, generalmente espresso attraverso l'istituzione di partenariati (par. 4.3).

4.1. L'organizzazione amministrativa delle istituzioni di cybersicurezza europea

Analogamente alle politiche europee di sicurezza brevemente evidenziate, anche il profilo organizzativo delle istituzioni di cybersicurezza europea si basa su una rete di amministrazioni aventi perlopiù funzioni organizzative e di coordinamento delle competenti autorità nazionali. Pertanto, anche in questo ambito l'azione europea non influisce sulle competenze e i poteri degli Stati membri ai quali resta sempre riservato l'esercizio dei poteri autoritativi di sicurezza.
L'articolazione amministrativa europea di cybersicurezza si è sviluppata nel tempo attraverso l'istituzione di diversi organismi di coordinamento decentrati, organizzati perlopiù sul modello delle agenzie dotate di personalità giuridica[46], che trovano nello

[45] N. Robinson - E. Disley, *Incentives and Challenges on Information Sharing. Retrieved*, 2010, 9.

[46] Sulle agenzie amministrative europee v. E. Chiti, *Le agenzie europee. Unità e decentramento nelle*

scambio informativo l'elemento essenziale per lo svolgimento delle loro funzioni.
Dal 2004 è presente l'Agenzia Europea per la Cybersicurezza (ENISA), istituita con l'obiettivo di creare «un clima di fiducia grazie alla sua indipendenza, alla qualità della consulenza fornita e delle informazioni diffuse, alla trasparenza delle sue procedure e metodi di funzionamento e alla diligenza nello svolgere i compiti ad essa assegnati» ed inoltre «[p]oiché le reti elettroniche sono in larga misura private, l'Agenzia dovrebbe avvalersi delle informazioni del settore privato e cooperare con esso» (cons. 11). L'Agenzia venne inizialmente dotata di un mandato temporaneo, via via esteso con i regolamenti (UE) 1007/2008 e 580/2011. Tuttavia, solo con il regolamento (UE) 2019/881, il c.d. *Cybersecurity Act*, è stato conferito all'ENISA un mandato permanente, rafforzandone il ruolo, i compiti, le responsabilità, e predisponendo maggiori risorse al fine di contribuire al supporto degli Stati membri nel prevenire e rispondere efficacemente agli attacchi informatici.
In particolare, l'Agenzia ricopre la funzione di segretariato della rete composta dai gruppi di intervento nazionali (c.d. rete CSIRT), nonché sostiene la cooperazione operativa tra questi e il gruppo di intervento dell'Unione, il CERT-UE, che ha la funzione di rispondere in modo efficiente alle minacce informatiche dirette contro le reti e i sistemi istituzionali dell'Unione europea.
I CSIRT, *Computer Security Incident Response Teams* sono unità di intervento decentrate, istituite presso i singoli Stati membri (eventualmente anche all'interno di autorità competenti[47]), con l'incarico di svolgere attività reattive, come l'intervento in caso di incidente informatico, ed anche proattive, come il monitoraggio degli incidenti a livello nazionale, l'emissione di preallarmi, allerte, annunci e divulgazione di informazioni alle parti interessate in merito a rischi e incidenti e la relativa analisi di tali rischi e incidenti.
In entrambo i casi, questi soggetti rappresentano i nodi nevralgici dei processi di *cyber information sharing*. Difatti da una parte ricevono le informazioni sulle minacce informatiche in quanto ricettori delle notifiche degli incidenti di cybersicurezza da parte dei soggetti verso cui trova applicazione la disciplina NIS; dall'altra, partecipano alla più ampia cooperazione informativa a livello europeo per mezzo della rete che riunisce i rappresentanti dei gruppi di intervento di tutti gli Stati membri e la squadra CERT-UE, sotto il segretariato dell'ENISA (c.d. rete di CSIRT)[48].

amministrazioni europee, Padova, 2002.

[47] È ad esempio il caso del CSIRT Italia trasferito presso l'Agenzia Nazionale per la Cybersicurezza (ACN) con il decreto-legge n. 82 del 2022.

[48] In particolare, l'art. 15 Dir. (UE) 2022/2555, prevede che la rete svolge i seguenti compiti: «a) scambiare informazioni per quanto riguarda le capacità dei CSIRT; b) agevolare la condivisione, il trasferimento e lo scambio di tecnologia e delle misure, delle politiche, degli strumenti, dei processi, delle migliori pratiche e dei quadri pertinenti fra i CSIRT; c) scambiare informazioni pertinenti per quanto riguarda gli incidenti, i quasi incidenti, le minacce informatiche, i rischi e le vulnerabilità; d) scambiare informazioni in merito alle pubblicazioni e alle raccomandazioni in materia di cibersicurezza; e) garantire l'interoperabilità per quanto riguarda le specifiche e i protocolli per lo scambio di informazioni; f) su richiesta di un membro della rete di CSIRT potenzialmente interessato da un incidente, scambiare e discutere informazioni relative a tale incidente e alle minacce informatiche, ai rischi e alle vulnerabilità associati; g) su richiesta di un membro della rete di CSIRT, discutere e, ove possibile, attuare una risposta coordinata a un incidente identificato nella giurisdizione di tale Stato membro; h) fornire assistenza agli Stati membri nel far fronte a incidenti transfrontalieri a norma

Tale attività trova inoltre il supporto del "Gruppo di cooperazione", organismo composto dai rappresentanti degli Stati membri, dalla Commissione e dall'ENISA, la cui funzione è quella di agevolare la cooperazione strategica e lo scambio di informazioni tra gli Stati membri fornendo orientamenti e consulenza alle istituzioni europee nonché effettuando valutazioni coordinate dei rischi di cybersicurezza ed elaborando relazioni utili ai fini del riesame della disciplina NIS da parte della Commissione[49].

Nonostante gli sforzi diretti a istituire un quadro amministrativo e regolamentare in materia di scambio informativo, nella Strategia europea di cybersicurezza presentata nel dicembre 2020[50], si apprende che «[t]he EU lacks collective situational awareness of cyber threats». Secondo la Commissione il problema è dovuto, da una parte allo scarso coinvolgimento del settore privato nella cooperazione informativa, dall'altra alla resistenza degli Stati membri a condividere le informazioni in maniera sistematica e completa, rendendo così estremamente difficoltoso il funzionamento dei meccanismi di *cyber information sharing* tra gli Stati membri e le istituzioni dell'UE in caso di crisi o incidenti informatici transfrontalieri su larga scala[51].

La stessa Presidente della Commissione europea Ursula von der Leyen, nel "Discorso sullo stato dell'Unione 2021", ha ribadito la necessità di «gettare le basi per un processo decisionale collettivo» basato la scambio di «conoscenze provenienti da tutti i servizi e da tutte le fonti, dallo spazio ai formatori del personale di polizia, dall'*open source* alle agenzie di sviluppo».

È sulla scorta di tali considerazioni che la disciplina di cybersicurezza europea è stata recentemente aggiornata e potenziata proprio negli aspetti che interessano la cooperazione informativa.

A partire dal gennaio 2023 è entrata in vigore la direttiva (UE) 2022/2555 (Direttiva NIS II)[52], che ha abrogato la previgente direttiva (UE) 2016/1148 (Direttiva NIS I).

della presente direttiva; i) cooperare e scambiare migliori pratiche con i CSIRT designati in qualità di coordinatori di cui all'articolo 12, paragrafo 1, nonché fornire loro assistenza per quanto riguarda la gestione della divulgazione coordinata di vulnerabilità che potrebbero avere un impatto significativo su soggetti in più di uno Stato membro; j) discutere e individuare ulteriori forme di cooperazione operativa, anche in relazione a: i) categorie di minacce informatiche e incidenti; ii) preallarmi; iii) assistenza reciproca; iv) principi e modalità di coordinamento in risposta a rischi e incidenti transfrontalieri; v) contributi al piano nazionale di risposta agli incidenti e alle crisi di cibersicurezza su vasta scala di cui all'articolo 9, paragrafo 4, su richiesta di uno Stato membro; k) informare il gruppo di cooperazione sulle proprie attività e sulle ulteriori forme di cooperazione operativa discusse a norma della lettera j) e, se necessario, chiedere orientamenti in merito; l) fare il punto sui risultati delle esercitazioni di cibersicurezza, comprese quelle organizzate dall'ENISA; m) su richiesta di un singolo CSIRT, discutere le capacità e lo stato di preparazione di tale CSIRT; n) cooperare e scambiare informazioni con i centri operativi di sicurezza regionali e a livello dell'UE al fine di migliorare la consapevolezza situazionale comune sugli incidenti e le minacce informatiche in tutta l'Unione; o) se del caso, discutere le relazioni sulle revisioni tra pari di cui all'articolo 19, paragrafo 9; p) fornire orientamenti volti ad agevolare la convergenza delle pratiche operative in relazione all'applicazione delle disposizioni del presente articolo in materia di cooperazione operativa».

[49] Cfr. art. 14 Dir. (UE) 2022/2555.

[50] JOIN(2020) 18 final, *Comunicazione congiunta al parlamento europeo e al consiglio. La strategia dell'UE in materia di cibersicurezza per il decennio digitale*.

[51] *Ibidem*.

[52] Direttiva (UE) 2022/2555, *relativa a misure per un livello comune elevato di cybersicurezza nell'Unione, che abroga la direttiva 2016/1148*.

Tra le diverse modifiche, la nuova disciplina ha istituito la "Rete europea delle organizzazioni di collegamento per le crisi informatiche" (*EU Cyber Crisis Liaison Organisation Network* - CyCLONe) con lo scopo di garantire una più stretta collaborazione e azione coordinata nei casi di incidenti di cyberisucrezza su larga scala. A tal fine la Rete sostiene la gestione coordinata a livello operativo degli incidenti e delle crisi di cybersicurezza su vasta scala e garantisce il regolare scambio di informazioni pertinenti tra gli Stati membri e le istituzioni, gli organi e gli organismi dell'Unione[53].

Da giugno 2023 è inoltre operativo il *Joint Cyber Unit*: il cuore della nuova cooperazione operativa europea in materia di cybersicurezza. Si tratta di una piattaforma di raccordo ove i partecipanti, provenienti dalla comunità civile, diplomatica, dalle forze dell'ordine e dalla difesa, possono avvalersi del supporto e delle competenze reciproche, soprattutto nel caso in cui le varie comunità debbano lavorare a stretto contatto, in occasione di incidenti su larga scala o crisi[54].

L'Unità non costituisce un organismo supplementare indipendente, ma è frutto della messa a disposizione di uno spazio comune fisico, situato a Bruxelles, e uno spazio virtuale composto da strumenti utili per una condivisione sicura e rapida delle informazioni.

Tra le amministrazioni europee che vi partecipano troviamo: relativamente alle politiche di polizia, lo *European Cybercrime Centre* (EC3), unità specializzata già istituita presso l'EUROPOL con funzioni di raccordo con le forze di polizia degli Stati europei[55]; sul piano diplomatico, lo *European External Action Service* (EEAS)[56] e il forum *Horizontal Working Party on Cyber Issues*[57]; infine, per quanto riguarda il settore difesa, il *framework Permanent Structured Cooperation* (PESCO)[58] e la *European Defence Agency* (EDA)[59].

---

[53] Cfr. art. 16 direttiva (UE) 2022/2555.

[54] C(2021) 4520 *fina*l, *Sulla creazione di un'unità cibernetica congiunta*, 2021. A ben vedere il *Joint Cyber Unit* prende avvio dal precedente progetto "*Blueprint*" del 2017 istituito con la Raccomandazione (EU) 2017/1584 sulla risposta coordinata a incidenti e crisi di cybersicurezza su larga scala.

[55] Il Centro europeo per la criminalità informatica (*European Cybercrime Centre* - EC3) è un organismo istituito da Europol nel 2013, con sede all'Aia. La sua attività è quella di coordinare le attività transfrontaliere di contrasto alla criminalità informatica e funge da centro di competenza tecnica in materia. Per ulteriori si rinvia al sito ufficiale dell'EC3.

[56] Lo *European External Action Service* (EEAS), o anche Servizio europeo per l'azione esterna (SEAE), è il servizio diplomatico dell'UE, istituito per rendere più coerente ed efficace la politica estera dell'UE e rafforzare così l'influenza dell'Europa sulla scena mondiale. Per ulteriori si rinvia al sito ufficiale dell'EEAS.

[57] Il forum *Horizontal Working Party on Cyber Issues* è stato istituito nel 2016 ed è responsabile del coordinamento dei lavori del Consiglio sulle questioni informatiche, principalmente la politica informatica e le attività legislative. Il Gruppo collabora strettamente con la Commissione europea ed altre istituzioni quali il Servizio europeo per l'azione esterna, l'Europol, l'Eurojust, l'Agenzia europea dei diritti fondamentali (FRA), l'Agenzia europea per la difesa (EDA) ed infine l'Agenzia dell'Unione europea per la cybersicurezza (ENISA).

[58] Il *Permanent Structured Cooperation* (PESCO) nel settore della politica di sicurezza e di difesa è stato istituito l'11 dicembre 2017 con decisione 2017/2315 del Consiglio. Tale strumento offre un quadro giuridico per pianificare, sviluppare e investire congiuntamente in progetti di capacità condivisi e migliorare la prontezza operativa e il contributo delle forze armate.

[59] L'Agenzia europea per la difesa è stata istituita con un'azione comune del Consiglio dei ministri del 12 luglio 2004, «per sostenere gli Stati membri e il Consiglio nel loro sforzo di migliorare le capacità di difesa europee nel campo della gestione delle crisi e per sostenere la politica europea di sicurezza e di

Da ultimo, il 18 aprile 2023, la Commissione ha avanzato una proposta di regolamento che stabilisce una serie di misure per rafforzare la solidarietà e le capacità di individuare, prepararsi e rispondere alle minacce e agli incidenti di sicurezza informatica nel contesto europeo (c.d. *EU Cyber Solidarity Act*)[60].
Con questo strumento l'Unione intende incrementare la consapevolezza situazionale, la condivisione delle informazioni, nonché migliorare la preparazione e la risposta agli incidenti informatici a livello comune attraverso l'istituzione di tre nuovi meccanismi di raccordo: lo *European Cybersecurity Shield*, il *Cyber Emergency Mechanism* e il *Cybersecurity Incident Review Mechanism*.
Lo *European Cybersecurity Shield* avrà il compito di migliorare il rilevamento, l'analisi e la risposta alle minacce informatiche su larga scala attraverso l'istituzione di una nuova rete di piattaforme di *Security Operation Centres* SOC multinazionali. La prima fase del progetto è stata già avviata nel novembre 2022, e sono stati selezionati tre consorzi di centri operativi di sicurezza (SOC) transfrontalieri, che riuniscono enti pubblici di 17 Stati membri e dell'Islanda, nell'ambito del programma Europa digitale (si rinvia al par. 5.1 a proposito dei SOC).
Il *Cyber Emergency Mechanism* avrà il compito di migliorare la preparazione e la risposta agli incidenti di cybersicurezza attraverso: la valutazione dei meccanismi di risposta implementati presso i settori particolarmente critici selezionati al termine di una generale valutazione del rischio a livello europeo; la creazione dell'*EU Cybersecurity Reserve*, ossia servizi di risposta agli incidenti erogati da fornitori di servizi privati («trusted providers»), attivati su richiesta degli Stati membri o di istituzioni dell'Unione, per aiutarli ad affrontare problemi significativi o incidenti di sicurezza informatica su larga scala: ed infine, attraverso la promozione dell'assistenza reciproca tra gli Stati membri ove uno di questi sia stato interessato da un incidente di cybersicurezza[61].

4.2 Brevi cenni sulla privatizzazione della (cyber) sicurezza

Il paradigma weberiano che vede nello Stato l'unico detentore del legittimo uso della forza non è più coerente con l'attuale situazione. Il processo di globalizzazione ha

difesa nella sua forma attuale e in quella futura».

[60] COM(2023) 209 final, *Proposal for a Regulation of the European Parliament and of the Council laying down measures to strengthen solidarity and capacities in the Union to detect, prepare for and respond to cybersecurity threats and incidents*.

[61] Sul concetto di "assistenza reciproca", l'art. 10, lett. c) dell'*EU Cyber Solidarity Act* si limita a fare rinvio alla medesima nozione disposta nella Direttiva NIS II. Considerati i contrasti interpretativi della dottrina sulla qualificazione dell'attacco informatico come attacco armato (v. E. Corsi, *La Nato a difesa del cyber spazio? Il dilemma nel diritto internazionale*, in *Research Analysis del Center for Cyber Security and International Relations Studies*, 2018), nonché lo stato dell'arte circa la definizione di una politica di sicurezza e difesa europea (v. M. Frau, *I nodi irrisolti della difesa comune europea. Una prospettiva federalista*, in *federalismi.it*, 6, 2022), non è da escludersi che questo principio possa essere ricondotto all'omonimo principio di reciproca assistenza di cui all'art. 42 del TUE, ove è previsto che nel rispetto della politica e di sicurezza e di difesa «di taluni Stati membri» l'assistenza allo Stato aggredito sia subordinata al previo coinvolgimento della NATO, sul punto cfr. F. Pocar - M.C. Baruffi (a cura di), *Commentario breve ai Trattati dell'UE*, Padova, 2014, 42.

portato ad una ri-articolazione dello Stato che ha di fatto trasferito alcune sue funzioni ad attori privati[62], tra cui, col tempo, anche quella della sicurezza[63]. Secondo Alcuni questo processo non ha visto una piena affermazione dei privati in questo settore, quanto piuttosto l'instaurazione di forme di *governance* ibrida, caratterizzate da una stretta collaborazione (*rectius* cooperazione) con il potere pubblico, il cui risultato ha portato alle cc.dd. *global security assemblages*, ossia la formazione di nuove strutture e pratiche di sicurezza che sono allo stesso tempo pubbliche e private, oltre che globali e locali[64].

Preme precisare che gran parte della letteratura in tema di "privatizzazione della sicurezza" è perlopiù concentrata sulle c.d. *Private Military or Security Companies services* - PMSCs[65], mentre poca attenzione è stata dedicata all'analisi delle organizzazioni private che non hanno nulla a che fare con la sicurezza in senso tradizionale, pur essendone oramai largamente coinvolte: è il caso delle ricordate infrastrutture critiche operative in diversi settori di primaria rilevanza come quello sanitario, trasporti, finanziario, ecc. A tal proposito l'analisi delle politiche di cybersicurezza europea può rappresentare un emblematico esempio di questo rapporto. Diversi documenti strategici, tra cui anche la strategia dell'Unione europea per la cybersicurezza del 2013[66], fanno riferimento alla collaborazione tra pubblico e soggetti privati operanti in diversi settori, senza tuttavia precisare come debba realizzarsi questa cooperazione nella pratica. Come si comprenderà il tema è particolarmente complesso in quanto non investe solo profili pratici, ma anche giuridici e politici. La domanda di fondo a cui la dottrina tenta di dare risposta è quella di trovare soluzioni che possano colmare il divario tra due opposte posizioni: la massimizzazione del profitto, ricercata dal settore privato, e la massimizzazione della sicurezza quale priorità dei governi.

Sebbene a nostro modo di vedere la ricerca di tali soluzioni da parte dell'Unione europea sia ancora *in fieri*, per il momento pare utile soffermarsi sul ruolo oggi ricoperto dagli attori privati nel processo di cybersicurezza europeo.

Benjamin Farrand ed Helena Carrapico in un recente studio hanno analizzato la progressiva rilevanza assunta dagli attori privati attivi nei settori della disciplina NIS[67] nel processo di regolazione della sicurezza delle reti e delle risorse informatiche. Dall'analisi delle politiche di cybersicurezza adottate a partire dagli anni 2000, i due Autori hanno individuato tre momenti fondamentali: una prima fase, dal 2001, in cui i soggetti privati sono considerati vittime delle azioni di *cybercrime*, e quindi ricoprono

---

62 S. Sassen, *Territory, Authority, Rights: From Medieval to Global Assemblages*, Princeton, 2008.

63 R. Abrahamsen - A. Leander, *Handbook of private security studies*, Londra, 2016.

64 R. Abrahamsen - M. C. Williams, *Security Privatization and Global Security Assemblages*, in *The Brown Journal of World Affairs*, 18(1), 2011, 171.

65 *Ex multis*, R. Mandel, *The Privatization of Security*, in *Armed Forces & Society*, 2001, 129–151; V. Calderai, *The Privatization of Military and Security Services and the Limits of Contract Law*, in EUI MWP, 2010/31; P. W. Singer, *Corporate Warriors: The Rise of the Privatized Military Industry*, New York, 2008.

66 Commissione europea, *Strategia dell'Unione europea per la cibersicurezza: un ciberspazio aperto e sicuro*, del 7 febbraio 2013.

67 B. Farrand - H. Carrapico, *Blurring public and private: cybersecurity in the age of regulatory capitalism*, in O. Bures - H. Carrapico (a cura di), *Security Privatization. How non-security-related Private Businesses Shape Security Governance*, Cham, 2018, 197 ss.

un «passive role as object of regulation»[68]; successivamente, a seguito dell'istituzione dell'ENISA nel 2004, il settore privato non viene considerato solo come obiettivo di potenziali attacchi informatici ma anche come «active stakeholder that sholud form part of the regulatory structure»[69]; ed infine, con la Strategia per una società dell'informazione sicura del 2006, la Commissione ha ritenuto che «private sector does not only act as an adopter of regulation, bur can also be actively involved in shaping policy responses and the resulting regulation»[70].

Lo studio dimostra quindi come il settore privato abbia assunto un ruolo sempre più influente all'interno dei processi di regolazione in questa particolare branca sicuritaria. Il riferimento è a un documento dell'ENISA del 2012 ove si dimostra che gli standard adottati nelle norme di cyberisucrezza europee per garantire la sicurezza e l'integrità delle informazioni sono fortemente basati su alcuni standard industriali utilizzati nel mercato europeo delle telecomunicazioni[71]. Pertanto, come scrivono i due Autori «[t]hrough the identification of standards of best practice, as well as the perceived position of experts in the field of telecoms, although the Commission has imposed binding legislation upon them, they have nevertheless been able to influence the standards by which the legislation is applied and interpreted by feeding the multi-stakeholder process»[72].

Per quanto qui interessa, ci concentreremo sul secondo momento individuato nello studio che vede il settore privato responsabile della corretta applicazione della disciplina sulla sicurezza delle reti e delle risorse informatiche.

4.3 I partenariati pubblico-privati europei di cybersicurezza

Come anticipato (v. infra par. 3), l'esigenza di sicurezza delle reti e dei sistemi informatici ha richiesto da subito la collaborazione tra il settore pubblico e privato trovando concreta attuazione attraverso l'istituzione di partenariati *ad hoc*, soprattutto al fine di favorire la creazione di ecosistemi informativi volti a prevnire le minacce informatiche. Ne sono un esempio i citati ISACs, partenariati pubblico-privati non profit isituiti originariamente negli Stati Uniti per aiutare le infrastrutture critiche attravero la raccolta centralizzata, valutazione e diffusione delle informazioni di cybersicurezza fornite dai CERTs e SOCs[73].

Anche a livello europeo[74], il parteneriato pubblico-privato è risultato essere lo stru-

[68] Ivi, 202.

[69] Ivi, 205.

[70] Ivi, 207.

[71] ENISA, *Shortlisting network and information security standards and good practices*, 2012.

[72] B. Farrand - H. Carrapico, *Blurring public and private,* cit., 209.

[73] N. Choucri - S. Madnick - P. Koepke , *Institutions for cyber security: International responses and data sharing initiative*, Working Paper CISL# 2016–10, Cybersecurity Interdisciplinary Systems Laboratory, MIT, Cambridge, MA, 2016.

[74] O. Bures, *Contributions of Private Businesses to the Provision of Security in the EU: Beyond Public-Private*

mento più adadatto per garantire la sicurezza informatica dei settori qualificati come critici, soprattutto al fine sviluppare la prevenzione, la preparazione e la risposta europea agli atti di terrorismo informatico attraverso l'istituzione della rete di *information sharing* per la protezione delle infrastrutture critiche CIWIN (*Critical Infrastructure Warining Information Network*)[75].

La convenienza circa l'utilizzo di questo strumento nel particolare contesto della cybersicurezza, nonché della protezione delle infrastrutture critiche è stato individuata da Alcuni nei seguenti motivi: «(a) the private sector 'owns or controls' a large number of CIs [critical infrastructures]; (b) the implementation of security policies depends on the involvement of the private sector in the 'definition of strategic public policy objectives as well as operational priorities and measures'; (c) PPPs 'would bridge the gap between national policy-making and operational reality on the ground'»[76].

Preme precisare che queste prime esperienze cooperative sono sorte su impulso dei governi ma la loro effettiva realizzazione e partecipazione è avvenuta in virtù della sola volontà dei soggetti che vi aderivano (c.d. approccio "*bottom-up*"). Auto-organizzati settorialmente secondo gli ambiti di operatività delle infrastrutture critiche (troviamo infatti ISACs nel settore finanziario, energetico, ecc.), la diffusione delle informazioni e degli allarmi sulle minacce informatiche avveniva sulla base di accordi di natura privata.

La allora Comunità europea si è limitata in un primo momento a promuovere la creazione di detti Centri a livello nazionale (esigenza ancora attuale date le recenti sollecitazioni), riconoscendo «the importance of multi-stakeholder models such as Public Private Partnerships (PPPs), built on a long term, bottom-up model to mitigate identified risks where such an approach delivers added value in helping to ensure a high level of network resilience»[77]. Anche l'ENISA, sulla scorta dell'implementazione della disciplina NIS, ha prodotto documenti sui modelli cooperativi per la costituzione dei ISACs nazionali[78].

Tuttavia, considerata la sempre più avvertita necessità di coordinare le procedure di scambio delle informazioni e degli allarmi in modo uniforme, l'Unione si è anche attività per creare partenariati a livello europeo. È il caso dell'*European Information Sharing and Alerting System* (EISAS)[79], progetto avviato nel 2007 con il fine di «colmare la lacuna nella condivisione di informazioni [...]» attraverso lo studio di modelli di analisi e diffusione delle informazioni di cybersicurezza utili alla creazione di uno spazio di condivisione comune[80].

Partnerships, in O. Bures - H. Carrapico (a cura di), *Security Privatization,* cit., 32.

[75] Si rinvia al sito della Commissione europea a proposito dello CIWIN.

[76] F. Cappelletti - L. Martino, *Achieving Robust European Cybersecurity through Public–Private Partnerships: Approaches and Developments*, in *EU Policy Review,* 1, 2021, 62.

[77] Consiglio europeo, *Council Resolution on a collaborative European approach to network and information security*, 2009/C 321/01, 2009, sezione IV, 7.

[78] ENISA, *Information Sharing and Analysis Centres (ISACs) Cooperative models,* 2018.

[79] ENISA, *EISAS – European Information Sharing and Alerting System*, 2007; nonchè il report, *EISAS (enhanced) report on implementation*, pubblicato nel 2011.

[80] Sui diversi settori critici coinvolti nel circuito EISAS si rinvia al sito ufficiale.

Come si apprende dal "*Deployment Feasibility Study*" del 2013, il programma EISAS si poneva l'obiettivo di creare un sistema di scambio informativo su larga scala rafforzando la cooperazione dei già esistenti ISACs settoriali degli Stati membri e semplificando il flusso informativo grazie alla consegna di «materiali pre-prodotti ai partecipanti»[81]. Tra le altre "migliori pratiche" si avvertiva infatti l'esigenza di processare le informazioni raccolte dai Centri nazionali, al fine di disseminare dati di alta qualità, oltreché evitare la duplicazione degli stessi.

Diversamente, l'*European Public-Private Partnership for Resilience* - EP3R, ha rappresentato il primo tentativo di istituire un partenariato comune a livello europeo per affrontare problemi di sicurezza e resilienza nel settore delle telecomunicazioni[82].

Il progetto, avviato nel 2009, è stato successivamente chiuso nel 2013. Alcuni studiosi hanno ricondotto i motivi che hanno portato al fallimento di questa esperienza alla scarsa partecipazione degli aderenti al progetto sotto diversi profili: la mancanza di impegno nella condivisione delle informazioni, la mancanza di trasparenza procedurale, nonché la scarsa partecipazione delle infrastrutture di piccola e media dimensione, diversamente da quelle maggiori coinvolte in prima persona dalla disciplina NIS[83].

La condivisione delle informazioni sulle minacce informatiche e gli allarmi attraverso l'istituzione di strutture cooperative come i partenariati resta tuttavia una priorità per le politiche europee di cybersicurezza. Nonostante il fallimento dell'EP3R, nella Strategia per la cybersicurezza europea del 2013 veniva ribadito che «il partenariato europeo pubblico-privato per la resilienza (EP3R) costituisce una valida piattaforma a livello dell'UE che dovrebbe essere ulteriormente sviluppata»[84]. A tal fine l'ENISA ha creato, all'interno del framework della piattaforma NIS, tre gruppi di lavoro, con un focus specifico sugli strumenti di co-regolamentazione e relative politiche pubbliche con riferimento alla gestione del rischio, alla condivisione delle informazioni e al coordinamento in caso di incidenti tra pubblico e attori privati, che hanno sostituito l'EP3R.

Nello stesso anno la Commissione europea accoglieva l'esigenza di istituire l'unità specializzata EC3 (*European Cybercrime Centre*) per il contrasto alla criminalità informatica presso l'Europol[85]. Si tratta di un caso di partenariato pubblico-privato ove tra le parti vi sono autorità che svolgono compiti di polizia. Nello specifico, come si apprende dal sito, l'EC3 si avvale di due gruppi di consultazione che includono attori del settore privato al fine di creare un ambiente cooperativo capace di cooperare sulle sfide legate alla criminalità informatica, promuovendo la collaborazione sia a livello strategico sia operativo[86].

---

81 ENISA, *EISAS – European Information Sharing and Alerting System. Deployment Feasibility Study*, 2013.

82 ENISA, *EP3R 2009-2013 Future of NIS Public Private Cooperation*, 2015.

83 Cfr. K. Iron, *The Governance of Network and Information Security In the European Union: The European Public-Private Partnership for Resilience (EP3R)*, in S. Gaycken - J. Krueger - B. Nickolay (a cura di), *The Secure Information Society: Ethical, Legal and Political Challenges*, Berlino, Springer Publ., 2021, 83 ss.

84 Commissione europea, *Strategia dell'Unione europea per la cibersicurezza …,* 2013, 7.

85 Conclusioni 10603/12 del Consiglio sull'istituzione di un centro europeo per la criminalità informatica, 2012.

86 Si rinvia alla pagina *The EC3 Advisory Groups – Law Enforcement and Private Sector Meetings to Discuss Latest Cybercrime Threats and Challenges*, del sito EUROPOL.

Sulla scorta di tali gruppi, l'EC3 ha siglato diversi *Memoranda of Understanding* (MoU) con gli attori privati operanti in settori critici, come quello finanziario[87], ma soprattutto quelli attivi nel settore dei servizi di sicurezza informatica[88]. Tali accordi, sebbene siano espressione di una libera contrattazione privata, hanno avuto l'effetto di dirigere le parti verso fini pubblici e modelli comuni di condivisione delle informazioni di cybersicurezza che, da una parte hanno aiutato il settore privato ad innalzare i livelli di sicurezza, dall'altra hanno permesso all'EC3 di essere sempre aggiornato sulle ultime minacce informatiche[89].

5. La *cyber information sharing* alla luce della Direttiva NIS II e delle linee guida ENISA

La natura transfrontaliera delle minacce informatiche ha caratterizzato l'organizzazione di cyberiscurezza per una più accentuata necessità di ricorrere a meccanismi di cooperazione informativa attraverso la costituzione di reti di scambio tra soggetti che nutrono fiducia vicendevolmente[90]

Oltre alle competenti autorità pubbliche di polizia, intelligence e difesa (come avviene per la sicurezza in senso tradizionale), tra i partecipanti allo scambio informativo in questo settore trovano spazio anche gli stessi beneficiari delle garanzie di cybersicurezza, perlopiù soggetti pubblici o privati operanti in settori critici.

Tali attori sono oggi disciplinati dalla ricordata direttiva (UE) 2022/2555 (Direttiva NIS II) che all'art. 1, par. 2 - diversamente dalla previgente normativa - prevede espressamente che la direttiva stabilisce «norme e obblighi in materia di condivisione delle informazioni sulla cybersicurezza».

Dal raffronto dei due testi è possibile anche intuire come l'Unione europea si stia sempre più dirigendo verso modelli di regolazione di tipo *risk-based* nelle politiche digitali[91]. Ne è prova l'introduzione dei concetti di «quasi incidente», «incidente» ed «incidente di cybersicurezza su vasta scala» che trovano definizione all'art. 6 della Direttiva NIS II, nonché quello di «incidente significativo» di cui all'art. 23, par. 3[92], i quali lasciano

[87] Si rinvia alla pagina *Europol and the European ATM Security Team reaffirm their partnership in combating payment crimes*, del sito EUROPOL.

[88] Si faccia riferimento agli accordi con Karspersky, McAfee, Mnemonic, Microsoft, FireEye la cui documentazione è reperibile sul sito EUROPOL.

[89] R. Bossong - B. Wagner, *A typology of cybersecurity and Public-Private partnership in the context of the European union*, in O. Bures - H. Carrapico (a cura di), *Security Privatization,* cit., 236.

[90] Sui meccanismi cooperativi e di coordinamento per fini di cybersicurezza v. F. Skopik - G. Settanni - R.Fiedler, *A problem shared is a problem halved: A survey on the dimensions of collective cyber defense through security information sharing*, in *Computers & Security*, 60, 2016, 154 ss.

[91] G. De Gregorio - P. Dunn, *The European Risk-Based Approaches: Connecting Constitutional Dots in the Digital Age*, in *Common Market Law Review*, 59(2), 2022.

[92] Cfr. art. 6, nn. 5, 6, 7 della Direttiva NIS II ove per «quasi incidente» si intende «un evento che avrebbe potuto compromettere la disponibilità, l'autenticità, l'integrità o la riservatezza di dati conservati, trasmessi o elaborati o dei servizi offerti dai sistemi informatici e di rete o accessibili attraverso di essi, ma che è stato efficacemente evitato o non si è verificato»; per incidente, «un evento che compromette la disponibilità, l'autenticità, l'integrità o la riservatezza di dati conservati, trasmessi o elaborati o dei

intendere i diversi gradi di intervento e gestione del rischio informatico da parte dei soggetti coinvolti nel processo di cybersicurezza europeo. Disciplina che deve inoltre essere interpretata alla luce del ricordato *EU Cyber Solidarity Act*, quale strumento volto a potenziare la gestione unica degli incidenti ad impatto diffuso nello spazio europeo.
Sinteticamente, è possibile notare che al variare di intensità di un incidente di cybersicurezza, da livello nazionale/locale a livello europeo o internazionale, il quadro di discipline vigenti fa corrispondere altrettante competenti strutture che favoriscono meccanismi di raccordo sempre più estesi.
Nello specifico, la recente disciplina NIS II ha integrato il sistema di risposta coordinata agli incidenti e alle crisi di cybersicurezza su vasta scala istituito con la Raccomandazione (UE) 2017/1584, del 13 settembre 2017, con la previsione dell'EU-CyCLONe, a cui si aggiungono il Gruppo di cooperazione e la Rete dei Gruppi di intervento, ribadendo tuttavia la necessità per tutti gli attori di «specificare ulteriormente i meccanismi di funzionamento della rete, compresi i ruoli, i mezzi di cooperazione, le interazioni con altri attori pertinenti e i modelli per la condivisione delle informazioni, nonché i mezzi di comunicazione»[93].
La scalarità degli incidenti di cybersicurezza ci permette quindi di distinguere le normali pratiche di condivisione delle informazioni per motivi di cybersicurezza rientranti nella definizione propria di *cyber information sharing*, dallo scambio informativo che si aziona in caso di emergenza a seguito di un incidente significativo su un soggetto "essenziale" o "importante", o a seguito di un incidente su larga scala. In particolare, per quanto riguarda la presente trattazione, l'interesse è di analizzare tali circuiti informativi ponendoci dalla prospettiva di coloro che alimentano e generano il traffico informativo a seguito dell'incidente, ossia i soggetti interessati dalla novellata disciplina NIS II.
A tal proposito, nel caso si realizzino «incidenti significativi» nei confronti di soggetti qualificabili come come «essenziali» o «importanti» vige l'obbligo di segnalazione alle competenti autorità e/o ai CSIRT di cui all'art. 23. Diversamente, ove tale incidente abbia interessato «soggetti diversi» da quelli appena ricordati, l'adempimento informativo verso le competenti autorità consisterà in una «notifica volontaria di informazioni pertinenti» disciplinata all'art. 30, par. 1, lett. b).
Inoltre, oltre a tali procedure d'emergenza, vi sono anche altri circuiti informativi, che rientrano nelle attività di vera e propria *cyber information sharing*. Questa procedura si differenzia dalla prima per due caratteristiche: 1) il fatto che l'interlocuzione di tali soggetti non avviene solo con le preposte istituzioni europee, ma anche con altri attori NIS (perlopiù appartenenti allo stesso settore es. energetico, finanziario, ecc.), 2) per il fatto che non sussistono obblighi di partecipazione agli ecosistemi informativi. Tale

servizi offerti dai sistemi informatici e di rete o accessibili attraverso di essi»; ed infine per «incidente di cybersicurezza su vasta scala» si intende «un incidente che causa un livello di perturbazione superiore alla capacità di uno Stato membro di rispondervi o che ha un impatto significativo su almeno due Stati membri». La nozione di «incidente significativo» è invece introdotta all'art. 23, par. 3 della direttiva rubricato "Obblighi di segnalazione", il quale lo definisce come un incidente che «a) ha causato o è in grado di causare una grave perturbazione operativa dei servizi o perdite finanziarie per il soggetto interessato; b) si è ripercosso o è in grado di ripercuotersi su altre persone fisiche o giuridiche causando perdite materiali o immateriali considerevoli».

[93] Cfr. considerando 68, della Direttiva NIS II.

circuito è alimentato dall'esclusiva volontà dei partecipanti, siano essi di rilevanza critica ai sensi della direttiva («essenziali» o «importanti») o diversi[94].
La condivisione (volontaria) delle informazioni avviene pertanto sulla scorta di accordi tra le parti. Sul punto la recente Direttiva NIS II ha introdotto l'art. 29, rubricato "Accordi di condivisione delle informazioni sulla cybersicurezza". Come si apprende dal disposto, il legislatore europeo ha indicato come obiettivo per gli Stati membri quello di mettere in condizione tutti i soggetti, critici e non, di «scambiarsi, su base volontaria, pertinenti informazioni sulla cybersicurezza, comprese informazioni relative a minacce informatiche, quasi incidenti, vulnerabilità, tecniche e procedure, indicatori di compromissione, tattiche avversarie, informazioni specifiche sugli attori delle minacce, allarmi di cybersicurezza e raccomandazioni concernenti la configurazione degli strumenti di cybersicurezza per individuare le minacce informatiche». Al paragrafo secondo è precisato che tale scambio venga «attuato mediante accordi di condivisione delle informazioni sulla cybersicurezza che tengano conto della natura potenzialmente sensibile delle informazioni condivise» ed in particolare che gli Stati membri, nel facilitare la conclusione di simili accordi, «possono specificare gli elementi operativi, compreso l'uso di piattaforme TIC dedicate e di strumenti di automazione, i contenuti e le condizioni degli accordi di condivisione delle informazioni», e nel caso della partecipazione delle autorità pubbliche a tali accordi «possono imporre condizioni per le informazioni messe a disposizione dalle autorità competenti o dai CSIRT».

5.1. La progressiva europeizzazione degli strumenti di cooperazione informativa: SOC, Registri delle vulnerabilità, Standard di scambio e Piattaforme di *cyber threat sharing*

Appurati i soggetti segnalanti nel precedente paragrafo, è doveroso ora precisare di quali strumenti e processi, tali soggetti, siano essi di natura pubblica o privata, si servono per contribuire ad alimentare il traffico di informazioni di cybersicurezza alla luce della vigente disciplina europea.

a) I Security Operation Centres (SOCs)

Alla base del processo di condivisione delle informazioni relative alle minacce informatiche troviamo i *Security Operation Centers* (SOCs), centri operativi di sicurezza pubblici o privati che, attraverso il continuo monitoraggio delle reti e dei sistemi dell'organizzazione per il quale operano, evitano che gli attacchi informatici possano avere un impatto negativo sul funzionamento e l'economia dell'organizzazione limitandone i danni[95]. Tali centri sono in grado non solo di rilevare le minacce in corso, ma anche di

[94] Si faccia riferimento al considerando 29 del Cybersecurity Act, ove è previsto che «l'ENISA dovrebbe sostenere la condivisione delle informazioni intra e intersettoriale [...]».

[95] Sulla definzione di SOC si rinvia alle linee guida ENISA, *How to set up CSIRT and SOC. Good practice*

estrapolarne informazioni particolarmente utili, sia per le attività di indagine condotte dalle forze di polizia (vedi la *digital forensics*), sia per le attività di prevenzione come i meccanismi di *information sharing*.
Nello specifico, questo patrimonio informativo è composta da: le vulnerabilità tecniche, ossia le debolezze del sistema o dei beni informatici che il criminale ha sfruttato per comprometterne la riservatezza, la disponibilità o l'integrità; gli *exploit*, ossia, il codice appositamente realizzato per sfruttare una determinata vulnerabilità e comprometterla, oppure altri tipi di infromazioni, come i c.d. indicatori di compromissione (IoC), termine con il quale si intende fare generalmente riferimento all'indirizzo del protocollo Internet (IP) del *server* eventualmente sfruttato per condurre l'attacco, il nome di dominio DNS (*Domain Name System*) o l'URL (*Uniform Resource Locator*) sospetti che rimandano a contenuti dannosi, ed infine l'identificativo di un file eseguibile dannoso o il testo dell'oggetto di un messaggio e-mail dannoso[96].
Considerato che si tratta di Centri ad oggi perlopiù istituiti presso singoli enti e realtà industriali, sia di natura pubblica, sia privata, preme porre attenzione al citato *Cyber Solidarity Act*. Qualora la proposta di regolamento entri in vigore senza emendamenti, la realizzazione dell'*European Cybersecurity Shield* introdurrebbe due importanti novità nelle architetture di cybersicurezza nazionali e quella europea. Si prevede infatti che lo "Scudo" sarà costituito da SOC nazionali, di natura pubblica, designati da ciascuno Stato membro (art. 4), e dai «Cross-border SOCs», ossia Centri transfrontalieri costituiti da un consorzio di almeno tre Stati membri rappresentati dai SOC nazionali (c.d. *Hosting Consortium*), che si impegnano a lavorare insieme per coordinare il loro rilevamento informatico e le attività di monitoraggio delle minacce (artt. 6-7)[97].
Su quest'ultimo punto, si precisa che il Consorzi saranno costituito sulla base di accordi scritto in cui i membri dovranno anche dettagliare i requisiti e i pricincipi per la condivisione delle «relevant information» tra i partecipanti (art. 6)[98]. Inoltre, la proposta invita i singoli Consorzi a stringere accordi con altri Consorzi.

b) I Registri delle vulnerabilità e delle debolezze informatiche

Agli albori dell'informatica erano in uso le prime "liste" di vulnerabilità create ed alimentate dai primi utenti della "*Internet society*" (al tempo composta perlopiù ingegneri

guide, dicembre 2020,.

[96] Sul punto si faccia riferimento alla scheda informativa pubblicata dal Centro nazionale per la sicurezza informatica olandese, il *National Cyber Security Centre* (NCSC), *Factsheet on Indicators of Compromise (IoCs)*, 2017.

[97] Sulla formazione del "*consortium*", la proposta di regolamento prevede che i membri del Consorzio ospitante stipulino un accordo consortile scritto che stabilisce le loro disposizioni interne, ove sono anche indicati in dettaglio i requisiti per la condivisione informazioni tra i partecipanti a un SOC transfrontaliero e per la condivisione di informazioni

[98] Sul contenuto delle "informazioni rilevanti" da trasferire, l'art. 6 della proposta prevede «information relating to cyber threats, near misses, vulnerabilities, techniques and procedures, indicators of compromise, adversarial tactics, threat-actor-specific information, cybersecurity alerts and recommendations regarding the configuration of cybersecurity tools to detect cyber-attacks».

ed esperti di informatica)[99] per mezzo delle "*request for comments*" (o RfC)[100]. Si trattava di uno strumento di trasparenza che bene esprimeva l'autoregolamentazione che caratterizzava la Rete. È interessante notare come successivamente, sulla scorta della sempre maggiore rilevanza acquisita dagli attacchi informatici per le società e l'economia, gran parte di questi registri siano oggi perlopiù sviluppati e supervisionati da un connubio di enti privati e pubblici, soggetto al controllo dei governi[101].

Negli Stati Uniti, sono attivi i *Common Vulnerabilities and Exposures* (CVE) e *Common Weakness Enumeration* (CWE)[102]: due indici che rientrano nella c.d. *vulnerability disclousure*, ossia la condivisione di informazioni sulle vulnerabilità e debolezze del *software* al fine di favorire la mitigazione degli effetti negativi di un accesso indesiderato da parte degli esperti di sicurezza. Si tratta di programmi di aggregazione e pubblicazione delle vulnerabilità e debolezze informatiche supervisionate da un ente privato non-profit, il *MITRE Corporation*[103], con il supporto dall'Agenzia per la sicurezza informatica e delle infrastrutture che fa capo al Dipartimento della sicurezza interna degli Stati Uniti. Inoltre, i *database* relativi ai CVE sono pubblicati in maniera sincrona su un altro registro, il *National Vulnerability Database* (NVE), gestito e fondato dall'agenzia pubblica *National Institute of Standards and Technology* (NIST)[104] a partire dal 2005.

Si comprenderà pertanto come l'origine geografica di tali banche dati sia una questione di rilevante interesse per i governi. Ne è prova il fatto che, oltre a quelle più comuni appena citate, di origine statunitense, sono stati istituiti ecosistemi informativi di questo tipo anche in altri Paesi, come ad esempio Giappone[105], Cina[106] e Russia[107].

---

[99] L'*Internet society* (ISOC) è un'organizzazione internazionale di diritto americano per la promozione dell'utilizzo e dell'accesso a Internet, oggi popolata da diverse sezioni locali (c.d. *chapters*) che vedono la partecipazione di gran parte dei paesi del mondo.

[100] La "*Repaired Security Bugs in Multics*", è stata la prima "lista" di vulnerabilità pubblicata pubblicata nel 1973 da Jerome H. Saltzer, con RfC n. 5. Le *request form comments* (RfC) sono «documents contain technical specifications and organizational notes for the Internet» così definiti dall'organismo internazionale che li prodoce, l'*Internet Engineering Task Force* (IETF), responsabile della standardizzazione dell'nternet e degli standard tecnici che ne consentono il funzionamento, primo fra tutti la suite di protocolli Internet (TCP/IP).

[101] Nonostante la scarsità di documenti ufficiali sul punto, da fonti Wiki si apprende che il primo (e forse anche unico) *database* di vulnerabilità sviluppato da un ente indipendente, quindi svincolato da controlli da parte di poteri pubblici, sia stato l'*Open Sourced Vulnerability Database* (OSVDB). Si trattava di un'iniziativa che ha preso avvio dalla nota *convention* per amanti dell'informatica, Def Con, nel 2002, per essere resa operativa con il primo *database open-source* nel 2004 (quindi svincolato anche da legami proprietari con le aziende di *software*) con il supporto della *Open Security Foundation* (OSF). Tuttavia, il 5 aprile 2016 il *database* è stato chiuso.

[102] Si rinvia rispettivamente ai siti del *Common Vulnerabilities and Exposures* (CVE) e del *Common Weakness Enumeration* (CWE).

[103] Come si apprende dal sito ufficiale dell'organizzazione, il MITRE si è costituito nel 1958 come società privata senza scopo di lucro per fornire consulenza ingegneristica e tecnica all'Aeronautica degli Stati Uniti. Il progetto fu utile per la creazione del primo centro di ricerca e sviluppo finanziato a livello federale (FFRDC), sponsorizzato dal Dipartimento della Difesa.

[104] Il NIST è parte del Dipartimento del Commercio degli Stati Uniti.

[105] v. *Japan Vulnerability Notes* (JVN).

[106] v. *Chinese National Vulnerability Database* (CNNVD).

[107] v. *Data Security Threats Database* (BDU), di cui non si hanno molte informazioni liberamente reperibili in rete se non alcuni articoli di stampo giornalistico, v. J. Leiden, *Russia's national vulnerability database is a*

Sul punto pare rilevante osservare che al considerando 63 della direttiva (UE) 2022/2555 è previsto che «sebbene simili registri o banche dati delle vulnerabilità esistano già [es. CVE e CWE], questi sono ospitati e mantenuti da soggetti non stabiliti nell'Unione». Con la Direttiva NIS II, l'Unione europea ha difatti promosso, per la prima volta, l'istituzione di un registro europeo delle vulnerabilità mantenuto dall'ENISA al fine di garantire «una maggiore trasparenza, per quanto riguarda la procedura di pubblicazione prima della divulgazione ufficiale della vulnerabilità, e resilienza in caso di perturbazioni o interruzioni nella fornitura di servizi analoghi»[108].
Tuttavia, se da un lato la divulgazione delle vulnerabilità e debolezze informatiche all'interno di banche dati pubbliche coordinate e liberamente accessibili può certamente favorire la sicurezza informatica parte degli addetti, dall'altro può anche essere uno strumento di facile utilizzo da parte dei criminali aventi tutto l'interesse a sfruttarle per violare le reti e i sistemi informatici. Proprio per questo motivo al successivo considerando 58 della direttiva, il legislatore europeo ha previsto di «rafforzare il coordinamento» tra i segnalanti e i fabbricanti o fornitori dei beni e servizi ICT dal quale sono state rilevate tali vulnerabilità, in modo da velocizzare la comunicazione.
Si precisa inoltre che nel considerando viene fatto esplicito riferimento (c.d. rinvio fisso) alle norme internazionali ISO/IEC 30111 e ISO/IEC 29147 circa la gestione e divulgazione delle vulnerabilità anche a terzi soggetti.

c) Le piattaforme di *Cyber Information Sharing* e gli standards di condivisione

Generalmente la condivisione delle informazioni sulle minacce informatiche e gli allarmi avviene per il mezzo di piattaforme di *cyber information sharing*, di natura proprietaria od *open source* (come ad esempio *Malware Information Sharing Platform* - MISP)[109], che permettono di diffondere ed alimentare questo patrimonio informativo per mezzo di standard di linguaggio specifici (*rectius* standard sul formato dei dati[110]).
Tuttavia, da diverso tempo, il mercato della sicurezza ha visto la progressiva introduzione anche di piattaforme particolarmente evolute - le c.d. piattaforme di *Cyber Threat Intelligence* (CTI) - ossia strumenti capaci non solo di estrapolare e condividere le informazioni relative alle minacce informatiche e agli incidenti di sicurezza, ma anche di elaborarle attraverso l'incrocio con altre fonti esterne che permettono di restituire - per

bit like the Soviet Union – sparse and slow 7 comment bubble on white By design, though, not... er, general rubbishness, in *The Register*, 17 luglio 2018.

[108] Cfr. considerando 63, Direttiva NIS II. A ben vedere, l'art. 29 della direttiva, prevede che tra le «pertinenti informazioni sulla ciberiscurezza» rientrino, oltre alle vulnerabilità tecniche: informazioni relative a minacce informatiche, quasi incidenti, procedure, indicatori di compromissione (IoC), tattiche avversarie, informazioni specifiche suglia attori delle minacce, allarmi di cibersicurezza e raccomandazioni concernenti la configurazione degli strumenti di cibersicurezza per individuare le minacce informatiche.

[109] Si rinvia al sito del *Malware Information Sharing Platform* (MISP).

[110] I formati di dati maggiormente utilizzati per il funzionamento di queste piattaforme sono gli standard STIX/TAXII, CyBOX, OASIS. Per maggiori dettagli tecnici sul funzionamento delle piattaforme di *sharing* si rinvia alla citata guida del NIST di cui in nota 112.

l'appunto - informazioni di *threat intelligence*[111].

Nello specifico la CTI è stata definita come «*systematic collection, analysis and dissemination of information pertaining to a company's operation in cyberspace and to an extent physical space. It is designed to inform all levels of decision makers. The analysis is designed to help keep situational awareness about current and arising threats*»[112].

L'aggregazione delle fonti che caratterizza le attività di *threat intelligence* permette infatti di avere un quadro generale sulle cc.dd. tattiche, tecniche e procedure (TTP), ossia: le descrizioni di alto livello del comportamento (tattiche); le descrizioni dettagliate del comportamento nel contesto di una tattica (tecnica); descrizioni dettagliate nel contesto di una tecnica (procedure). Le TTP permettono quindi di descrivere la tendenza di un attore a utilizzare una specifica variante di *malware*, un ordine di operazioni, uno strumento di attacco, un meccanismo di consegna (ad esempio, un attacco di *phishing*) o un *exploit*[113].

Pertanto - per dirla con i termini della criminologia - le piattaforme di CTI permettono di ricostruire la "firma" e il "modus operandi" dell'attore malevolo[114], restituendo informazioni non solo di natura tecnica, utili per l'arricchimento dei bacini informativi tipici dell'*information sharing*, ma anche informazioni complesse e aggregate utili per eventuali attività di investigazione da parte degli addetti.

Il recente interesse degli Stati verso la sicurezza del cyeberspazio ha portato una parte della dottrina ad interrogarsi sui profili giuridici di dette piattaforme. Per molto tempo questi strumenti hanno trovato applicazione nel settore privato senza una vera e propria regolamentazione[115], soprattutto sotto il profilo del trattamento delle informazioni e della protezione dei dati personali che sarà approfondito nel prossimo paragrafo.

Per quel che qui interessa, ossia il profilo regolazione strettamente legato al legittimo utilizzo di questi strumenti, pare utile richiamare quanto disposto nella Direttiva NIS II, ove all'art. 29, par. 3 prevede che gli Stati membri «facilitano la conclusione degli accordi di condivisione delle informazioni sulla cybersicurezza [...]» e «possono specificare gli elementi operativi, compreso l'uso di piattaforme TIC dedicate e di strumenti di automazione, i contenuti e le condizioni degli accordi di condivisione delle informazioni». L'eventuale partecipazione a tali circuiti informativi dovrà inoltre essere

---

[111] Preme distinguere la *cyber threat intelligence* dalla *cyber intelligence* quale autonoma branca dell'*intelligence* che consiste nel «complesso di attività programmate ed applicate per identificare, seguire, misurare e monitorare informazioni sulle minacce digitali, nonché dati sulle intenzioni e attività di entità avversarie» svolte con «strumenti cibernetici nel cyber spazio, cioè attraverso la rete, e hanno una particolarità, a differenza delle altre forme di intelligence, poiché non si può fare totale affidamento alle attrezzature elettroniche» v. U. Gori - L.S. Germani (a cura di), I*nformation Warfare 2011. La sfida della cyber intelligence al sistema Italia: dalla sicurezza delle imprese alla sicurezza nazionale*, Milano, 2012, 16 ss.; nonché M. Caligiuri, *Cyber Intelligence. Tra libertà e sicurezza*, Roma, 2016.

[112] Si rinvia alla pagina *Introduction to CTI as a General topic* presso il sito del FIRST.

[113] C. Johnson - L. Badger - D. Waltermire - J. Snyder - C. Skorupka, *Guide to Cyber Threat Information Sharing*, NIST Special Publication 800-150, 2016.

[114] R. Chiesa - S. Ciappi, *Profilo Hacker. La scienza del criminal profiling applicata al mondo dell'hacking*, Milano, Apogeo, 2007, 10 ss.

[115] L. O. Nweke - S. Wolthusen, *Legal Issues Related to Cyber Threat Information Sharing Among Private Entities for Critical Infrastructure Protection*, 2020 12th International Conference on Cyber Conflict (CyCon), Estonia, 2020, 63 ss.

notificata alle autorità competenti al momento di conclusione di tali accordi, così come il loro ritiro dagli stessi (par. 4).
L'ostacolo principale nello scambio di informazioni è la mancanza di standard comuni nella comunicazione. Sebbene la recente disciplina NIS sul punto non imponga il rispetto di requisiti comuni, pare utile rinviare ad uno studio "*work in progress*" condotto dall'ENISA relativamente allo scambio di informazioni tra i CSIRT e le autorità di polizia, ove si è proposta una tassonomia volta ad identificare quali informazioni possono essere condivise tra i due, e come ciò possa essere realizzato da una prospettiva tecnica e organizzativa[116].

6. La tutela dei diritti fondamentali e della sicurezza nel trattamento delle informazioni "sensibili e classificate" per lo Stato e dei dati personali contenuti nelle informazioni di cybersicurezza

Dal breve quadro sui soggetti segnalanti poc'anzi delineato emerge come il legislatore europeo abbia inteso distinguere i soggetti critici da quelli non rientranti in questa categoria in funzione dell'importanza per il settore in cui operano, o il tipo di servizi che forniscono, nonché delle loro dimensioni[117].
È bene precisare che le infrastrutture operanti in settori "critici" non rientrano solo nel campo d'applicazione della disciplina europea, ma anche nelle relative legislazioni nazionali degli Stati membri che, in alcuni casi, come in Italia, oltre ad aver recepito la disciplina NIS[118], hanno anche adottato legislazioni autonome in materia di cyberiscurezza nazionale.
La precisazione è d'obbligo poiché, se per la Direttiva NIS II le misure in essa disposte sono volte a garantire un livello comune elevato di cybersicurezza europea «in modo da migliorare il funzionamento del mercato comune» (art. 1), sul piano nazionale, vedi l'Italia, le misure contenute nel decreto decreto-legge del 21 settembre 2019, n. 105, istitutivo del Perimetro di Sicurezza Nazionale Cibernetica (PSNC), sono dirette ad assicurare la «tutela della sicurezza nazionale e dell'interesse nazionale nello spazio cibernetico» (art. 1, par. 1, lett. a)[119].
L'intima connessione tra sicurezza nazionale e protezione delle infrastrutture critiche[120], porta a dover concludere che la circolazione nello spazio europeo delle informazioni di cybersicurezza - composte da elementi che possiamo immaginare come le "chiavi d'accesso" a reti e sistemi informatici critici - possa essere fortemente ostaco-

116 ENISA, *Information sharing and common taxonomies between CSIRTs and Law Enforcement*, 2016.

117 Cfr. considerando 15 della Direttiva NIS II.

118 v. Decreto legislativo, 18 maggio 2018, n. 65, con il quale l'Italia ha recepito la disciplina NIS.

119 Cfr. art. 1, c. 1, lett. a) del d.l. n. 105 del 2019, recante disposizioni urgenti in materia di cybersicurezza, definizione dell'architettura nazionale di cybersicurezza e istituzione dell'Agenzia per la cybersicurezza nazionale.

120 B. Valensise, *I settori strategici dopo la riforma*, in G. Della Cananea - L. Fiorentino (a cura di), *I "poteri speciali" del Governo nei settori strategici*, Napoli, 2020, 101 ss.

lata da limiti dettati da prerogative statali sovrane, come la sicurezza interna[121], diversa rispetto al più ampio profilo della "sicurezza europea"[122].
Altro limite è certamente rappresentato dalla tutela dei diritti fondamentali dei singoli. Il materiale diffuso potrebbe infatti consentire di individuare le persone fisiche, trovando così applicazione il quadro di disciplina sulla protezione dei dati personali diversamente articolato in funzione della natura dei soggetti titolari del trattamento. In particolare la distinzione fondamentale è tra titolari del trattamento qualificabili come forze di polizia (o *law enforcement agencies* - LEAs) e soggetti che non rivestono tali incarichi.
Ulteriore considerazione riguarda la propagazione degli incidenti informatici che possono facilmente scalare da emergenza interna di una singola organizzazione, ad emergenza di livello nazionale o transnazionale. Motivo per cui il patrimonio informativo che caratterizza la *cyber information sharing* può essere utilizzato da diversi soggetti, per diverse finalità, che vanno dall'interesse degli enti alla salvaguardia dei propri affari, o delle pubbliche amministrazioni all'efficiente e continua fornitura dei servizi, fino alla difesa della sicurezza nazionale da parte dei governi e la sicurezza europea.
Nel presente paragrafo si tenterà pertanto di analizzare il differenziato regime di trattamento delle informazioni di cybersicurezza in considerazione del fatto: a) che tali informazioni possono essere utilizzate per la tutela della sicurezza interna degli Stati membri e quindi potrebbero esservi apposte classificazioni o potrebbero essere qualificate come "sensibili", limitandone di fatto la circolazione; b) che tali informazioni possono contenere dati personali rientranti nelle relative discipline di settore in funzione delle finalità e dei soggetti che le trattano: organizzazioni europee, forze di polizia, titolari "generici".

a) Lo scambio di informazioni "sensibili e classificate" di cybersicurezza e i limiti alla loro circolazione

Nel richiamato documento strategico del 2020, la Commissione osservava che «[l]'interoperabilità dei sistemi di informazioni classificate rimane tuttavia limitata, impedendo

[121] Sul rapporto tra sovranità e sicurezza nell'ottica statuale v. C. Mortati, *Istituzioni di diritto pubblico*, Padova, 1962, 127 ss.; M.S. Giannini, *Sovranità (diritto vigente)*, in *Enciclopedia del diritto*, vol. XLIII, Milano, 1990, 224 ss.; G. de Vergottini, *Guerra e costituzione. Nuovi conflitti e sfide alla democrazia*, Bologna, 2004, 151 ss.; Id., *La difficile convivenza fra libertà e sicurezza. La risposta delle democrazie al terrorismo*, in *Rassegna Parlamentare,* 2, 2004 427 ss.; Id., *La persistente sovranità*, in *Recte sapere, Studi in onore di Giuseppe Dalla Torre*, Tomo II, Torino, 2014, 1373- 1392; A. Spadaro, *Dalla "sovranità" monistica all'"equilibrio" pluralistico di legittimazione del potere nello Stato costituzionale contemporaneo*, in *Rivista AIC*, 3, 2017, 1 ss.; E.A. Imparato, *Sovranità e sicurezza. Un connubio ancora vincente?* in *federalismi.it*, 1, 2019. Nonché sulla sovranità nell'era dell'informatizzazione v. A. Simoncini, *Sovranità e potere nell'era digitale*, in O. Pollicino - T.E. Frosini - E. Apa – M. Bassini (a cura di), *Diritti e libertà in internet*, Milano, 2017, 20 ss.; S. Mannoni - G. Stazi, *Sovranità. co. Potere pubblico e privato ai tempi del cyberspazio*, Napoli, 2021; V. Bertola - S. Quintarelli, *Internet fatta a pezzi. Sovranità digitale, nazionalismi e big tech*, Torino, 2023.

[122] V. S. Peers, *National Security and European Law*, in *Yearbook of European Law*, 16(1), 1996, 363 ss.; U. Draetta, *L'Unione europea tra processo costituente e sovranità nazionale*, in U. Draetta - N. Parisi - D. Rinoldi (a cura di), *Lo «spazio di libertà sicurezza e giustizia» dell'Unione europea,* cit., 17 ss; A. Alì, *Il diritto dell'Unione europea e la tutela della sicurezza nazionale degli Stati membri. Osservazioni a margine di alcuni casi esaminati dalla Corte di Giustizia dell'Unione europea*, in U. Gori - L. Martino, *Intelligence e interesse nazionale,* Aracne editrice, Roma, 2015, 593 ss.

un trasferimento fluido delle informazioni tra le diverse entità» ravvisando così l'esigenza di un approccio interistituzionale al trattamento delle informazioni classificate a livello europeo, anche attraverso l'individuazione di «una base di riferimento per semplificare le procedure con gli Stati membri».

Come deducibile dalla citata proposta di Regolamento *EU Cyber Solidarity Act*, tali procedure sono ancora in fase di implementazione, e soprattutto non è stata individuata un'idonea base di legittimità. Il tema è estremamente sensibile poiché la cooperazione informativa, riconducibile al più ampio concetto di sicurezza collettiva[123], si scontra con le limitazioni poste per esigenze di sicurezza interna da parte degli Stati membri.

A tal proposito pare utile richiamare il contenuto dell'art. 346, lett. a) del Trattato sul funzionamento dell'Unione europea (TFUE) già art. 296 del TCE. La norma rientra tra le disposizioni che autorizzano una deroga all'applicazione delle norme del TFUE in virtù di ragioni non economiche (c.d. clausola di salvaguardia), autorizzata in nome di esigenze di sicurezza e difesa nazionale con il fine di realizzare un delicato equilibrio tra le predette esigenze interne degli Stati e gli obiettivi fondamentali del mercato interno.

In particolare, con l'ipotesi di cui alla lett. a) del par. 1, si consente agli Stati membri di rifiutare di fornire informazioni a qualunque istituzione europea la cui divulgazione sia dagli stessi considerata «contraria agli interessi essenziali della propria sicurezza»[124], purché tale misura restrittiva sia ritenuta necessaria e mai per ragioni di carattere economico.

La dottrina si è variamente dibattuta sull'interpretazione del disposto tra approcci restrittivi ed estensivi[125]. Secondo la Commissione la disposizione «va oltre il settore della difesa, e mira in generale a proteggere le informazioni che gli Stati membri non possono divulgare senza mettere in pericolo i loro interessi essenziali della propria sicurezza»[126].

Pertanto, la deroga in questione esonera gli Stati dal più ampio obbligo derivante dal principio di leale collaborazione di cui all'art. 4, par. 3, del TUE, che impone agli Stati di fornire informazioni alle istituzioni dell'UE (compresa la Corte di giustizia) o ad altri Stati membri le informazioni che gli fossero richieste, al fine di mantenere segreto ciò

---

123 World Economic Forum, *Cyber Information Sharing: Building Collective Security. Insight Report*, ottobre 2020.

124 Si tratta di una clausola di salvaguardia prevista dal Trattato che trova applicazione nelle sole ipotesi contemplate dal disposto. Introdotta con il fine di tutelare il segreto di Stato che riguarda la sicurezza nazionale dei Paesi membri, questo articolo rappresenta una deroga agli artt. 4, par. 3, del TUE e 337 del TFUE, rispettivamente dedicati, all'obbligo di fornire informazioni alle istituzioni europee in virtù del principio di leale collaborazione, il secondo, attributivo alla Commissione europea il potere di raccogliere tutte le necessarie informazioni e di procedere alle opportune verifiche per l'esecuzione dei suoi compiti. Sul punto v. F. Pocar - M.C. Baruffi (a cura di), *Commentario breve ai Trattati dell'UE*, cit., 1546 ss.

125 Sull'approccio restrittivo v. P. Gori, *Art. 223*, in R. Monaco - R. Quadri - A. Trabucchi (diretto da), *Commentario CEE*, Milano, 1995, 1626 ss.; per altro orientamento, di interpretazione estensiva v. R. Smit,P. Herzog, *Article 223*, in P. Herzog - C. Campbell - G. Zagel (a cura di), *The Law of the European Union is the completely updated and revised edition of their Law of the European Community: A Commentary on the EC Treaty*, New York, 5.

126 COM(2006)779 del 7 dicembre 2006 sull'applicazione dell'art. 296 del trattato CE agli appalti pubblici della difesa.

che riguarda la propria sicurezza[127].

L'applicazione dell'art. 346 TFUE nel contesto della *cyber information sharing* trova riscontro nell'*EU Cyber Solidarity Act*, ove al considerando 23 è previsto che lo scambio informativo avvenga nel rispetto dei limiti del disposto ("*without prejudice*"), ed inoltre che tale disseminazione «*should be limited to that which is relevant and proportionate to the purpose of that exchange. The exchange of such information should preserve the confidentiality of the information and protect the security and commercial interests of the entities concerned, in full respect of trade and business secrets*».

A tal proposito, il considerando 9 della Direttiva NIS II prevede che la regolazione di tali traffici dovrebbe avvenire nel rispetto delle «norme dell'Unione o nazionali per la protezione delle informazioni classificate, gli accordi di non divulgazione o gli accordi di non divulgazione informali, quale il protocollo TLP» (su quest'ultimo torneremo a breve), e al successivo considerando 118 viene previsto che «l'ENISA dovrebbe predisporre l'infrastruttura, le procedure e le norme per il trattamento delle informazioni sensibili e classificate in conformità alle norme di sicurezza applicabili alla protezione delle informazioni classificate dell'UE».

Quella dello scambio e della protezione delle informazioni "sensibili e classificate" nel contesto europeo è una disciplina in evoluzione che è avanzata nel tempo per accordi e decisioni tra l'Unione europea e singoli Stati, anche non membri dell'UE[128]. Senza entrare nel dettaglio, brevemente, le parti concordano di sviluppare la cooperazione sulla sicurezza e sulla condivisione di informazioni classificate attenendosi ad alcune prerogative comuni: ciascuna delle parti deve proteggere le informazioni classificate fornite dall'altra, o scambiate con essa, a un livello almeno equivalente a quello offerto dalla parte che le fornisce; tutte le persone che hanno accesso alle informazioni classificate devono disporre di un adeguato nulla osta di sicurezza, basato sulla lealtà, sul carattere fidato e sull'affidabilità; Possono inoltre essere stabilite restrizioni sulla modalità di utilizzo e di divulgazione delle informazioni classificate, nonché di accesso alle stesse.

Sul punto è utile precisare che ai sensi dell'art. 2 della Decisione 2013/488/UE, per «informazioni classificate UE» (ICUE) si intende «qualsiasi informazione o qualsiasi materiale designati da una classifica di sicurezza UE, la cui divulgazione non autorizzata potrebbe recare in varia misura pregiudizio agli interessi dell'Unione europea o di uno o più Stati membri»[129].

Tuttavia, come precisato in dottrina, la clausola di cui al richiamato art. 346 lett. a) TFUE trova esclusiva applicazione verso gli Stati membri e non anche verso le imprese[130].

---

127 V. F. Pocar - M.C. Baruffi (a cura di), *Commentario breve ai Trattati dell'UE,* cit., 1547.

128 Per una panoramica sul punto si rinvia al sito Eur-Lex, v. anche E. De Capitani, *Unione europea e segreto di Stato*, in *www. astrid-online.it*, 2010.

129 Sono inoltre definiti al par. 2, quattro livelli di classificai: 1. Top Secret: informazioni e materiali la cui divulgazione non autorizzata potrebbe arrecare danni di eccezionale gravità agli interessi fondamentali dell'UE o di uno o più paesi dell'UE; 2. Secret: informazioni e materiali la cui divulgazione non autorizzata potrebbe ledere gravemente gli interessi fondamentali dell'UE o di uno o più paesi dell'UE; 3. EU Confidential: informazioni e materiali la cui divulgazione non autorizzata potrebbe ledere gli interessi fondamentali dell'UE o di uno o più paesi dell'UE; 4. EU Restricted: informazioni e materiali la cui divulgazione non autorizzata potrebbe essere pregiudizievole per gli interessi dell'UE o di uno o più paesi dell'UE.

130 F. Sciaudone, *Art. 346 TFUE*, in A. Tizzano (a cura di), *Trattati dell'Unione europea, Le fonti del diritto*

Principio che trova riscontro anche nel contesto degli scambi informativi di cybersicurezza atteso il contenuto del considerando 10 della citata Direttiva NIS II il quale, dopo aver evidenziato la connessione tra le infrastrutture critiche attive nel settore della produzione di energia elettrica da centrali nucleari e la sicurezza nazionale, prevede che «uno Stato membro dovrebbe poter esercitare la propria responsabilità per la salvaguardia della propria sicurezza nazionale in relazione a tali attività, comprese le attività all'interno della catena del valore nucleare, conformemente ai trattati».
Viene pertanto ribadita l'esclusiva competenza degli Stati nel dover adottare le misure necessarie a garantire la tutela degli interessi essenziali della sicurezza nazionale e salvaguardia dell'ordine pubblico e della pubblica sicurezza, promuovendo presso i soggetti critici il ricorso ad accordi volontari per la condivisione delle informazioni sulla cybersicurezza «che tengono conto della natura potenzialmente sensibile delle informazioni condivise»[131].
Tra le basi legali menzionate dalla disciplina NIS è fatto riferimento anche agli «accordi di non divulgazione informali, quale il protocollo TLP»[132], acronimo di *Traffic Light Protocol*. Si tratta di uno standard internazionale elaborato dal FIRST (*Forum of Incident Response and Security Teams*) per facilitare la condivisione di informazioni potenzialmente sensibili e una più efficace collaborazione[133].
Difatti, salve le ipotesi in cui lo scambio informativo sia ritenuto contrario «agli interessi essenziali della propria sicurezza»[134] e quindi gli Stati possono rifiutarsi di fornire informazioni a qualsiasi organizzazione dell'Unione, a livello generale, il traffico delle informazioni di cyebersicurezza è solitamente gestito dal citato standard TLP, costituito da «un insieme di quattro etichette utilizzate per indicare i limiti di condivisione che i destinatari devono applicare»[135].
Nello specifico tali etichette sono rappresentate da quattro colori: il rosso, rappresenta la massima restrizione ed indica le informazioni non divulgabili al pubblico ma solo a singoli destinatari in quanto potrebbero compromettere la riservatezza delle persone fisiche, i segreti, la reputazione o il *business* dell'organizzazione; giallo, indica le informazioni la cui divulgazione è limitata all'organizzazione e ai suoi clienti con l'avvertenza che la loro circolazione debba essere soggetta a particolari garanzie quando il trasferimento di queste possa compromettere la riservatezza delle persone fisiche, i segreti, la reputazione o il *business* dell'organizzazione; verde, indica le informazioni che possono essere diffuse tra i membri del circuito informativo a cui appartiene la piattaforma di CTI, solitamente si tratta di informazioni utili ad aumentare la consapevolezza (*awareness*) all'interno della loro comunità. Non ci sono invece restrizioni per le informazioni che comportano un rischio minimo o nullo di uso improprio, in conformità alle norme e alle procedure applicabili per la divulgazione al pubblico.

italiano, II ed., Milano, 2014, 2515 ss.

131 Cfr. art. 29, par. 2, Direttiva NIS II.

132 Cfr. considerando 9, nonché art. 10, par. 7, della Direttiva NIS II.

133 Il FIRST è un forum globale che riunisce i team di risposta agli incidenti di sciurezza informatica, creato negli Stati Uniti nel 1989 a seguito della istituzione del primo CERT.

134 Cfr. art. 346, leet. a), del TFUE.

135 La definizione è stata tratta dal sito del FIRST.

b) La protezione europea dei dati personali contenuti nelle informazioni di cybersicurezza

È innanzitutto doveroso distinguere quando le informazioni oggetto di trasferimento per motivi di cybersicurezza possano o meno rientrare nelle discipline europee sul trattamento dei dati personali. Pare allora utile partire dalla definizione contenuta all'art. 4, par. 1, n. 1, del regolamento (UE) 2016/679 (anche noto come GDPR) secondo cui per «dato personale» deve intendersi «qualsiasi informazione riguardante una persona fisica identificata o identificabile ("interessato")» che può essere «identificata, direttamente o indirettamente» anche con riferimento a un «identificativo online»[136].

Nei precedenti paragrafi è stata evidenziata la natura perlopiù tecnica delle informazioni cybersicurezza. Tuttavia, nonostante tale "aspetto", non è da escludersi che queste informazioni possano identificare, o rendere identificabile, una persona fisica (nello specifico, l'attore malevolo che responsabile dell'incidente di sicurezza o anche la vittima dello stesso). E' ad esempio il caso dell'*Internet Protocol* (IP), qualificato come dato personale dalla costante giurisprudenza europea, sia esso dinamico o statico[137], gli indirizzi mail, l'*Uniform Resource Locator* (URL)[138], i nomi di dominio (DNS)[139], ma anche le informazioni bancarie come l'IBAN, nonché l'identificativo fornito per l'utilizzo dei *social networks*.

Appurata l'applicazione della disciplina europea dei dati personali sulle informazioni di cybersicurezza, pare ora opportuno riflettere sull'attività di trattamento che si caratterizza per tre elementi necessari: il titolare del trattamento, ossia «la persona fisica o giuridica, l'autorità pubblica, il servizio o altro organismo che, singolarmente o insieme ad altri, determina le finalità e i mezzi del trattamento di dati personali [...]» (art. 4, n. 7 GDPR); la, o le, finalità per il quale i dati personali raccolti sono trattati; ed infine il destinatario, ossia «la persona fisica o giuridica, l'autorità pubblica, il servizio o un altro organismo che riceve comunicazione di dati personali, che si tratti o meno di terzi [...]» (art. 4, n. 9, GDPR).

Questa scansione tripartita ci è particolarmente utile per studiare la diversa applicazione delle discipline sul trattamento dei dati personali in riferimento al processo di *cyber information sharing* che, come abbiamo anticipato, vede la partecipazione di soggetti operanti in diversi settori (pubblico/privato, civile/autorità pubblica), e per diverse finalità che vanno dalla salvaguardia dei propri affari della singola organizzazione, fino alla difesa della sicurezza nazionale da parte dei governi, nonché la sicurezza europea

136 Cfr. art. 4, regolamento (UE) 2016/679.

137 *Ex multis*, si faccia riferimento alla nota sentenza Breyer, CGUE, C-582/14, del 19 ottobre 2016. Sul punto v. anche il parere 4/2007 del Working Party Article 29, sul concetto di dato personale ove viene precisato che «*some sorts of IP addresses which under certain circumstances indeed do not allow identification of the user, for various technical and organizational reasons. One example could be the IP addresses attributed to a computer in an internet café, where no identification of the customers is requested*».

138 v. M. Korse, *Personal Data in URLs*, in *privacywise*, 23 agosto 2017.

139 L'*Internet Corporation for Assigned Names and Numbers* (ICANN), che gestisce il *Domain Name System* (DNS), è responsabile anche della gestione del registro WHOIS, una banca dati pubblica secondo cui chiunque abbia un dominio web deve registrare non solo il proprio dominio, ma anche i propri nomi, indirizzi, indirizzi e-mail e numeri di telefono. Sul punto v. S. Vaughan-Nichols, *DNS is about to get into a world of trouble with GDPR*, in *Zdnet*, 18 aprile 2018.

e internazionale.

In considerazione di ciò, al fine di facilitarne lo studio, si propone l'analisi di quattro scenari, relativi a: 1) la diffusione di informazioni tra le organizzazioni per mezzo di piattaforme di *cyber information sharing* (o comunque all'interno dei circuiti ISACs); 2) lo scambio di informazioni dalle organizzazioni verso le autorità competenti, punti di contatto unici e i *Computer Security Incident Response Team* (CSIRT), 3) le informazioni acquisite da qualsiasi soggetto (critico e non) e poi trattate dalle autorità di polizia; 4) le informazioni trattate da soggetti diversi dalle forze di polizia ma per finalità di polizia.

Precisiamo che le discipline sulla protezione dei dati personali a cui si farà riferimento sono pertanto: il già citato regolamento (UE) 2016/679; la direttiva (UE) 2016/680 (anche nota come "direttiva Polizia") relativa alla protezione delle persone fisiche con riguardo al trattamento dei dati personali da parte delle autorità competenti a fini di prevenzione, indagine, accertamento e perseguimento di reati o esecuzione di sanzioni penali; ed infine, il regolamento (UE) 2018/1725 che stabilisce le norme applicabili al trattamento dei dati personali da parte delle istituzioni, degli organi e degli organismi dell'Unione europea, ed infine il regolamento (UE) 2016/794, che disciplina il trattamento dei dati personali da parte dell'Agenzia dell'Unione europea per la cooperazione nell'attività di contrasto (Europol).

Sebbene ognuna di queste normative abbia un differente campo applicativo, sono tuttavia ispirate al medesimo corpo di principi sul trattamento dei dati personali ove, primo fra tutti, trova spazio il principio di «liceità, correttezza e trasparenza»[140], da cui discende la questione sulla selezione delle basi di legittimità, ossia le condizioni che rendono il trattamento di dati personali conforme alla legge e quindi lecito.

Partendo dallo scenario 1), i citati provvedimenti europei in materia di diritto dei dati personali non forniscono indicazioni precise al riguardo. Anche la proposta di Regolamento *EU Cyber Solidarity Act*, al considerando 22, fa un generico richiamo al rispetto delle «*existing legal requirements and in particular Union and national data protection law, as well as the Union rules on competition governing the exchange of information*».

L'esigua dottrina sul punto ha individuato la legittimità del trasferimento delle informazioni di cybersicurezza tra organizzazioni pubbliche e private nei circuiti di *cyber sharing* nella base contenuta all'art. 6, par. 1, lett. f) del GDPR relativa al «legittimo interesse del titolare del trattamento o di terzi»[141].

Come specificato dal vecchio gruppo dei Garanti, il *Working Party Article* 29, il ricorso alla base dell'interesse legittimo richiede la valutazione di tre elementi per determinare la liceità del trattamento: necessità, legittimità e bilanciamento degli interessi, al termine del quale potrebbe prevalere l'interesse dell'interessato, ossia il soggetto a cui si riferiscono i dati personali, piuttosto che quello del titolare che li tratta[142].

---

[140] Cfr. art. 5, par. 1, lett. a) del regolamento (UE) 2016/679; art. 4, par. 1, lett. a) della direttiva (UE) 2016/680; art. 4, par. 1, lett. a) del regolamento (UE) 2018/1725.

[141] C. Sullivan - E.W. Burger, *"In the public interest": The privacy implications of international business-to-business sharing of cyber-threat intelligence*, in *Computer Law & Security Review*, 33(1), 2017, 14 ss. Vedi anche, L.O. Nweke - S. Wolthusen, *Legal Issues Related to Cyber Threat Information Sharing Among Private Entities for Critical Infrastructure Protection*, in *NATO CCDCOE 12th International Conference on Cyber Conflict*, 2020.

[142] Articolo 29 Working Group, *Parere 6/2014 sul concetto di interesse legittimo del responsabile del trattamento ai sensi dell'articolo 7 della direttiva 95/46/CE*.

In tale occasione il Gruppo ha inoltre avuto modo di precisare che l'interesse legittimo dei terzi potrebbe essere pertinente quando «il responsabile del trattamento, talvolta incoraggiato dalle autorità pubbliche, persegue un interesse che corrisponde a un interesse pubblico generale o a un interesse dei terzi», come ad esempio nei casi in cui «il responsabile del trattamento va oltre gli obblighi giuridici specifici che è tenuto a rispettare conformemente a leggi e regolamenti al fine di contribuire all'impegno profuso dalle autorità di contrasto e dai soggetti privati per combattere le attività illegali, quali il riciclaggio di denaro, l'adescamento di minori o la condivisione illegale di file online».

Sul punto, pare utile richiamare la lettera del considerando 49 GDPR ove è previsto che «costituisce legittimo interesse del titolare del trattamento interessato trattare dati personali relativi al traffico, in misura strettamente necessaria e proporzionata per garantire la sicurezza delle reti e dell'informazione, vale a dire la capacità di una rete o di un sistema d'informazione di resistere, a un dato livello di sicurezza, a eventi imprevisti o atti illeciti o dolosi che compromettano la disponibilità, l'autenticità, l'integrità e la riservatezza dei dati personali conservati o trasmessi e la sicurezza dei relativi servizi offerti o resi accessibili tramite tali reti e sistemi da autorità pubbliche, organismi di intervento in caso di emergenza informatica (CERTs), gruppi di intervento per la sicurezza informatica in caso di incidente (CSIRTs), fornitori di reti e servizi di comunicazione elettronica e fornitori di tecnologie e servizi di sicurezza [...]». Nonché anche il considerando 50, ove è previsto che «l'indicazione da parte del titolare del trattamento di possibili reati o minacce alla sicurezza pubblica e la trasmissione dei dati personali pertinenti a un'autorità competente in singoli casi o in più casi riguardanti lo stesso reato o la stessa minaccia alla sicurezza pubblica dovrebbero essere considerate nell'interesse legittimo perseguito dal titolare del trattamento».

Tuttavia, pare opportuno distinguere il caso in cui lo *sharing* avvenga tra le sole organizzazioni partecipanti al circuito informativo, per il quale la base di legittimità del legittimo interesse pare certamente coerente, dal caso in cui lo scambio abbia ad oggetto informazioni tratte da minacce andate a buon fine (incidente). Come anticipato, a seguito dell'entrata in vigore della disciplina NIS, le organizzazioni rientranti nella qualifica di soggetto «essenziale o importante» che siano impattate da un incidente significativo, sono obbligate ad effettuare la notifica alle autorità competenti e/o ai CSIRTs.

In quest'ultimo caso la dottrina ha individuato la diversa base dell'art. 6, par. 1, lett. c) GDPR che legittima il trattamento qualora condotto in adempimento di un obbligo legale al quale è soggetto il titolare del trattamento[143].

Altri studiosi hanno inoltre individuato una ulteriore base di legittimità nella lett. e) dell'art. 6, par. 1, GDPR[144], ritenendo che la raccolta e disseminazione di informazioni cybersicurezza, aventi non necessariamente come destinatari le forze di polizia, costituisca «l'esecuzione di un compito di interesse pubblico o connesso all'esercizio di

[143] A. Albakri - E. Boiten - R. De Lemos., *Sharing Cyber Threat Intelligence Under the General Data Protection Regulation*, in M. Naldi - G.F. Italiano - K. Rannenberg - M. Medina - A. Bourka (eds.), *Privacy Technologies and Policy, 7th Annual Privacy Forum, APF 2019 Rome, Italy, June 13–14, Proceedings*, Berlin, 2019, 28 ss.

[144] C. Sullivan - E.W. Burger, *"In the public interest"*, cit.

pubblici poteri di cui è investito il titolare del trattamento»[145].

Tuttavia tale orientamento non pare trovare conforto nella lettera del considerando 121 della Direttiva NIS II, ove per il trattamento dei dati personali condotto al fine di garantire la sicurezza dei sistemi informatici e di rete da parte di «soggetti essenziali e importanti» contempla le sole due basi del legittimo interesse e dell'obbligo di legge, precisando che quest'ultima base interviene per legittimare il trattamento di dati personali contenuti nelle notifiche in caso di incidente rilevante. Diversamente per i dati personali contenuti in informazioni di cybersicurezza oggetto di scambio informativo e notifica non obbligatoria di cui all'art. 30 della direttiva viene indicata la base dell'art. 6, lett. f).

Per quanto riguarda i casi riconducibili allo scenario 2), considerato il ruolo e la natura delle autorità competenti, dei punti di contatto unico e dei CSIRTs, ossia soggetti pubblici che non svolgono compiti e funzioni di polizia, si ritiene che il trattamento dei dati personali contenuti nelle informazioni di cybersicurezza trovi legittimità nell'art. 6, par. 1, lett. c) GDPR, per l'appunto relativo all'interesse pubblico. A titolo esemplificativo, si faccia riferimento alla *privacy policy* del CSIRT Italia ove è espressamente previsto che la «gestione delle segnalazioni inviate dagli utenti ai sensi degli articoli 12, 14 e 18 del Decreto Legislativo n. 65 del 18 maggio 2018 [avviene] sulla base giuridica dell'esecuzione di un compito di interesse pubblico o connesso all'esercizio di pubblici poteri di cui è investito il titolare del trattamento»[146].

Allo stesso modo, anche nel caso in cui la segnalazione provenga da parte delle istituzioni dell'Unione verso il Gruppo di intervento europeo, il CERT-UE, conformemente al regolamento (UE) 2018/1725, la *privacy policy* del Gruppo prevede che «il trattamento è necessario per l'esecuzione di un compito di interesse pubblico o per l'esercizio di pubblici poteri conferiti all'istituzione o all'organismo dell'Unione»[147].

Inoltre, il citato considerando 121 ricorda che, conformemente all'art 9 GDPR, gli Stati membri «potrebbe[ro] stabilire norme che consentano alle autorità competenti, ai punti di contatto unici e ai CSIRT, nella misura necessaria e proporzionata al fine di garantire la sicurezza dei sistemi informatici e di rete dei soggetti essenziali e importanti, di trattare categorie particolari di dati personali», prevedendo a tal fine di adottare misure adeguate e specifiche per tutelare i diritti e gli interessi fondamentali delle persone fisiche, comprese limitazioni tecniche al riutilizzo di tali dati e l'uso di misure all'avanguardia in materia di sicurezza e di tutela della vita privata, quali la pseudonimizzazione o la cifratura qualora l'anonimizzazione possa incidere significativamente sulla finalità perseguita.

Relativamente lo scenario 3), nel caso in cui le autorità di pubblica sicurezza si trovino a dover trattare informazioni di cybersicurezza per fini di «prevenzione, indagine, accertamento e perseguimento di reati o esecuzione di sanzioni penali, incluse la salvaguardia e la prevenzione di minacce alla sicurezza pubblica» (cfr. art. 1) trova applicazione

---

145 A. Albakri - E. Boiten - R. De Lemos, *Sharing Cyber Threat Intelligence Under the General Data Protection Regulation,* cit.

146 Si rinvia all'"Informativa sul trattamento dei dati personali" pubblicata sul sito del CSIRT Italia.

147 Si rinvia alla "*Privacy policy*" pubblicata sul sito del CERT-UE di cui al link: *cert.europa.eu.*

la direttiva (UE) 2016/680 (anche nota come Direttiva di polizia)[148].
Precisiamo che sia il GDPR, sia la appena citata Direttiva di polizia, conformemente a quanto previsto dall'art. 16 del TFUE secondo cui l'Unione europea stabilisce le norme relative alla protezione delle persone fisiche con riguardo al trattamento dei dati di carattere personale «nell'pesercizio di attività che rientrano nel campo di applicazione del diritto dell'Unione», non trovano applicazione verso quei trattamenti di dati personali svolti per finalità di tutela dell'interesse e della sicurezza nazionale (es. i trattamenti di dati svolti da parte degli organismi di intelligence per tali fini)[149].
Altra considerazione deve esser fatta per le attività svolte dalle autorità di polizia fuori dai fini indicati dall'art. 1 delle Direttiva di polizia, tra cui vi rientrano «quelle di archiviazione nel pubblico interesse, di ricerca scientifica o storica o per finalità statistiche, a meno che il trattamento non sia effettuato nel contesto di un'attività che non rientra nell'ambito di applicazione del diritto dell'Unione» ove trova applicazione il GDPR (art. 9, par. 2).
Si ritiene pertanto necessario approfondire quando le attività di cybersicurezza possano rientrare nel campo applicativo della direttiva, ovvero in quello del regolamento generale.
A tal proposito, il considerando n. 12 della direttiva prevede che le attività di polizia «comprendono anche il mantenimento dell'ordine pubblico quale compito conferito alla polizia o ad altre autorità incaricate dell'applicazione della legge ove necessario per la salvaguardia contro e la prevenzione di minacce alla sicurezza pubblica e agli interessi fondamentali della società tutelati dalla legge che possono dar luogo a reati». Mentre, il considerando 27 prevede che il perseguimento dei fini di polizia anzi descritti, renda «necessario che le autorità competenti trattino i dati personali raccolti a fini di prevenzione, indagine, accertamento o perseguimento di specifici reati al di là di tale contesto per sviluppare conoscenze delle attività criminali e mettere in collegamento i diversi reati accertati».
Secondo Nadezhda Purtova, le attività investigative di cybersicurezza, come il tracciamento di una *botnet* o l'individuazione della dimensione e il livello di minaccia di un incidente di sicurezza informatica reale o potenziale, possono certamente rientrare - in linea di principio - negli ambiti appena descritti e quindi nel capo applicativo della Direttiva di polizia[150].

---

148 Cfr. art. 2 della direttiva (UE) 2016/680. In particolare, sulla protezione dei dati personali nelle procedure di information sharing da parte delle forze di polizia v. F. Boehm, *Information sharing and data protection in the Area of Freedom, Security and Justice. Towards harmonised data protection principles for EU-internal information exchange*, Springer, 2012.

149 Tuttavia, come rilevano J. Sajfert, T Quintel, *Data Protection Directive (EU) 2016/680 for Police and Criminal Justice Authorities*, in *SSRN*, 2017, le agenzie di intelligence possono trovarsi a trattare dati personali anche per le finalità coperte dal campo applicativo della Direttiva di polizia che quindi ne troverebbe applicazione. Questo problema diventa ancora più rilevante nel contesto della condivisione delle informazioni tra le agenzie di intelligence nazionali e le forze di polizia.

150 N. Purtova, *Between the GDPR and the Police Directive: Navigating Through the Maze of Information Sharing in Public-Private Partnerships*, in *International Data Privacy Law*, 2018; D. Drewer - V. Miladinova, *The BIG DATA Challenge: Impact and opportunity of large quantities of information under the Europol Regulation*, in *Computer Law & Security Review*, 33, 2017, 298 ss.; T. Quintel, *Interoperable Data Exchanges Within Different Data Protection Regimes: The Case of Europol and the European Border and Coast Guard Agency*, in *European Public Law,* 2020, 205 ss.

Per quanto riguarda la legittimità dei trattamenti svolti per tali finalità, l'art. 8 prevede che siano gli Stati membri a disporre se il trattamento sia lecito, specificando anche gli obiettivi del trattamento, i dati personali da trattare e le finalità dello stesso[151]. Pertanto, deve ritenersi che tale condizione sia alla base dei trattamenti di dati personali interessati dalla *cyber information sharing* tra le autorità di polizia degli Stati membri.

Tuttavia, a tale scambio informativo non partecipano solo le competenti autorità nazionali, ma anche organismi di polizia sovranazionali, come l'Europol, ed in particolare l'unità specializzata contro la criminalità informatica EC3 attiva dal 2013.

Considerato che l'Unità opera presso l'Europol, il trattamento dei dati personali contenuti nelle informazioni di cybersicurezza da parte dell'Unità trova solida disciplina nel regolamento istitutivo l'Agenzia, il n. 794 del 2016[152], che cita il contrasto alla criminalità informatica tra gli obbiettivi dell'Europl[153]. Pertanto è al regolamento di quest'ultimo che faremo riferimento per analizzare la disciplina relativa al trattamento dei dati personali da parte di EC3.

Come si apprende dal testo di legge, un elemento importante nella lotta contro la criminalità informatica è l'applicazione SIENA (*Secure Information Exchange Network Application*), la rete sicura per lo scambio di informazioni e dati personali dell'Europol, finalizzata a garantire la protezione e i requisiti di sicurezza i tali informazioni scambiate tra Stati membri, Europol, altri organismi dell'Unione (art. 24), paesi terzi e organizzazioni internazionali (art. 25)[154].

Trattandosi di un sistema di scambio che vede Europol nel ruolo di intermediario, è bene distinguere i trasferimenti di dati personali in entrata, da parte di soggetti segnalanti, dai trasferimenti in uscita, verso i soggetti che potrebbero essere interessati dall'evento di sicurezza in questione.

Relativamente ai primi, tali dati possono essere forniti dagli Stati membri, e successivamente trasferiti dall'Europol sull'esclusiva base del loro consenso, liberamente revocabile in qualsiasi momento (art. 23, par. 6), ovvero, da parti private e pervenuti all'Europol per mezzo di un'unità nazionale, punto di contatto di un paese terzo o un'organizzazione internazionale con cui Europol ha concluso un accordo di cooperazione, nonché un'autorità di un paese terzo o un'organizzazione internazionale che forma oggetto di una decisione di adeguatezza (art. 26).

Una volta raccolti, al fine di raggiungere i suoi obiettivi, Europol può trattare dati personali per sole finalità «determinate, esplicite e legittime» (art. 28) consistenti in:

a) controlli incrociati diretti a identificare collegamenti o altri nessi pertinenti tra informazioni concernenti: i) persone sospettate di aver commesso un reato di competenza di Europol o di avervi partecipato, o che sono state condannate per un siffatto reato;

---

[151] Cfr. art. 8, direttiva (UE) 2016/680.

[152] Regolamento (UE) 2016/794, che istituisce l'Agenzia dell'Unione europea per la cooperazione nell'attività di contrasto (EUROPOL) e sostituisce e abroga le decisioni del Consiglio 2009/371/GAI, 2009/934/GAI, 2009/935/GAI.

[153] Cfr. art. 3, regolamento (UE) 2016/794. Si faccia inoltre riferimento al rapporto EC3, *First year report*, 2013, 6, ove è espressamente previsto che «[f]ortunately the existing Europol framework, including its legal basis, corporate structure and information processing tools, offered a solid basis», nonché D. Drewer - J. Ellermann, *Europol's data protection framework as an asset in the fight against cybercrime*, 2012.

[154] Cfr. considerando 24, regolamento (UE) 2016/794.

ii) persone riguardo alle quali vi siano indicazioni concrete o ragionevoli motivi per ritenere che possano commettere reati di competenza di Europol; b) analisi strategiche o tematiche; c) analisi operative; d) facilitazione dello scambio d'informazioni tra Stati membri, Europol, altri organismi dell'Unione, paesi terzi e organizzazioni internazionali.

Per quanto riguarda la disseminazione presso i destinatari, l'Agenzia può trasferire i dati personali a: un organismo dell'Unione, nella misura in cui tale trasferimento sia necessario allo svolgimento dei suoi compiti o dei compiti dell'organismo dell'Unione destinatario (art. 24); a un'autorità di un paese terzo o a un'organizzazione internazionale, purché vi sia una decisione di adeguatezza adottata dalla Commissione ai sensi dell'art. 36 della Direttiva di polizia, un accordo internazionale concluso tra l'Unione e tale paese terzo o organizzazione internazionale ai sensi dell'articolo 218 TFUE, un accordo di cooperazione che consenta lo scambio di dati personali (art. 25); nonché alle parti private, «se non in singoli casi, ove sia strettamente necessario» (art. 26, par. 5).

L'ultimo scenario (4), riguarda i soggetti tenuti a collaborare con le forze di polizia, o svolgere compiti di polizia, i quali quindi si trovano a dover trattare dati personali per finalità di prevenzione, indagine, accertamento e perseguimento di reati o esecuzione di sanzioni penali, incluse la salvaguardia e la prevenzione di minacce alla sicurezza pubblica.

A tal proposito, la Direttiva di polizia, all'art. 3, par. 7, contempla nella definizione di «autorità competente» non solo le forze pubbliche di polizia (lett. a), ma anche «qualsiasi altro organismo o entità incaricati dal diritto dello Stato membro di esercitare l'autorità pubblica e i poteri pubblici» per i fini anzidetti (lett. b)[155].

Sul regime d'applicazione, il considerando 11 della direttiva è chiaro: «[q]ualora tale organismo o entità trattino dati personali per finalità diverse da quelle della presente direttiva, si applica il regolamento (UE) 2016/679». Invece, nei casi rientranti nell'ambito applicativo della direttiva (UE) 2016/680 tali soggetti «dovrebbero essere vincolati da un contratto o altro atto giuridico e dalle disposizioni applicabili ai responsabili del trattamento» in modo da vincolarli al rispetto degli obblighi e garanzie contemplati dalla stessa[156].

In quest'ultima ipotesi, il considerando 11 fa quindi un implicito riferimento ai trattamenti di dati personali che trovano base legale nell'esecuzione di contratti di cui all'art. 6, par. 1, lett. b), del GDPR. Condizione che porta ad una necessaria co-titolarità del trattamento tra parte privata e autorità di polizia in cui entrambe le parti decidono congiuntamente sulla necessità del trasferimento: l'autorità competente determina che ha bisogno dei dati e l'ente privato decide di rispettare la richiesta.

[155] Cfr. art. 3, par. 7, lett. b), direttiva (UE) 2016/680.

[156] A tal proposito il considerando 11 prevede che ad esempio, «a fini di indagine, accertamento o perseguimento di reati, gli istituti finanziari conservano determinati dati personali da essi trattati, e li trasmettono solo alle autorità nazionali competenti in casi specifici e conformemente al diritto dello Stato membro». Ulteriore esempio di «altro organismo o entità» che tratta dati personali per conto delle autorità di polizia è il laboratorio privato forense che svolge analisi delle prove nei procedimenti penali su incarico di un tribunale, pubblico ministero o della polizia. Anche il fornitore di servizi *cloud* che fornisce in base a un contratto un servizio di archiviazione dei tribunali archivi digitali è un altro esempio, e avrebbe lo stato di un elaboratore sotto la direttiva.

Come già osservato in precedenza (infra 4.3), nel particolare caso degli scambi di informazioni di cybersicurezza, tali rapporti tra parti private e autorità di polizia trova espressione all'interno di appositi partenariati pubblico-privati. Secondo Nadezhda Purtova, il quadro legislativo europeo sul punto non pare fornire certezza sulla disciplina da applicare nel caso di trattamento congiunto dei dati personali sia da parti private, sia da parte di forze di polizia[157]. Salvo concludere poi che: «*[w]hile GDPR lays down general data protection rules, the Police Directive operates as lex specialis. This means that when it comes to processing for the law enforcement purposes, GDPR forms a 'safety net' that in principle should 'catch' all data processing for law enforcement purposes when the Police Directive does not apply*»[158].

7. Considerazioni conclusive sul processo di integrazione della cybersicurezza europea

Sebbene per molto tempo il tema della sicurezza, così come anche quello della difesa ad essa collegato, siano rimasti fuori dal processo di integrazione europea, ciò non impedisce di considerare la questione della sicurezza come rilevante per l'avvio del processo di integrazione[159].

L'argomento ha assunto considerevole importanza a seguito del progressivo impatto delle politiche europee a livello internazionale, nonché dei recenti risvolti geopolitici. Tuttavia, nonostante il tema sia avvertito come una «impellente necessità»[160], neppure l'accelerazione delle politiche di difesa e sicurezza comune a seguito della Brexit del 2016 si sono rivelate col tempo idonee alla costruzione di un adeguato apparato difensivo[161].

Le difficoltà di mettere a punto una politica estera in senso tradizionale nel contesto europeo si accompagnano infatti alla difficoltà di definire la sicurezza interna dell'Unione. Le differenze culturali, le questioni politiche e le divergenze di interessi tra gli Stati membri hanno spesso rappresentato un ostacolo ad una collaborazione partecipata in questi ambiti.

La cooperazione per fini di cybersicurezza rappresenta uno dei fondamentali principi per la prevenzione e miglior gestione degli incidenti informatici e, come evidenziato nel corso della trattazione, necessita non solo del pieno coinvolgimento degli Stati membri, e quindi dei poteri pubblici, ma anche del settore privato, particolarmente presente nel cyberspazio.

Precedentemente, si è introdotto il tema della privatizzazione della sicurezza quale effetto della "crisi" del paradigma weberiano, secondo cui è il solo Stato ad essere titolare del legittimo uso della forza. Tuttavia, si è anche mostrato come questo fenomeno non

[157] N. Purtova, *Between the GDPR and the Police Directive,* cit., 34.

[158] *Ivi*, 41.

[159] A ben vedere dal Trattato di Schuman deduciamo che fosse stato proprio il porre rimedio alla rivalità franco-tedesca ad aver stimolato l'avvio del processo di integrazione.

[160] B. Caravita, *Difesa europea, quali prospettive*, in *federalismi.it*, 1, 2019. V. anche M. Vellano - A. Miglio (a cura di), *Sicurezza e difesa comune dell'Unione europea*, Milano, 2022.

[161] Cfr. M. Frau, *I nodi irrisolti della difesa comune europea. Una prospettiva federalista*, in *federalismi.it*, 6, 2022.

abbia portato alla totale scomparsa dei poteri pubblici nella soddisfazione di pretese securitarie quanto piuttosto alla formazione di rapporti cooperativi di quest'ultimi con il settore privato.

L'analisi della *cyber information sharing* nel contesto delle politiche europee di cybersicurezza ha posto in evidenza il recente orientamento di "europeizzazione" di metodi e strumenti di natura generalmente privata. Si faccia riferimento ai citati esempi relativi alla creazione del SOC europeo, l'istituzione dei registri di vulnerabilità e debolezze, nonché, più in generale, la previsione di processi di gestione del rischio di ispirazione tecnica, incorporati all'interno di fonti di diritto derivato che ne impongono l'obbligatoria osservanza per una determinata categoria di soggetti[162].

Elementi questi che lascerebbero intravedere una nuova tendenza, certamente evolutiva dei rapporti tra pubblico e privato in questo settore, ove l'Unione europea non solo coopera, e co-regola, con il settore privato ma, da una parte, inizia ad occupare spazi prima appartenenti a questi, dall'altra ne funzionalizza l'operato al perseguimento di interessi pubblici[163].

Emblematico esempio di questo processo può essere colto sul piano normativo dalla definizione di "cybersicurezza europea", introdotta al culmine di un lungo processo che ha preso avvio con la disciplina sulla protezione delle infrastrutture critiche.

Ripercorrendo brevemente le tappe più significative di questo percorso[164], già nel 2001 la Commissione europea adottava una Comunicazione sulla criminalità "informativa" ove veniva data la definizione di «sicurezza dei sistemi informatici e di rete» facendo riferimento alla «capacità di una rete o di un sistema d'informazione di resistere, ad un determinato livello di riservatezza, ad eventi imprevisti o atti dolosi che compromettono la disponibilità, l'autenticità, l'integrità e la riservatezza dei dati conservati o trasmessi e dei servizi forniti o accessibili tramite la suddetta rete o sistema»[165].

Nel 2004, anno in cui veniva istituita l'Agenzia europea per la sicurezza delle reti e dell'informazione (ENISA), il Consiglio europeo lanciava lo *European Program for Criminal Infrastructure Protection* (EPCIP) con lo scopo di incrementare la prevenzione, la preparazione e la risposta europea agli atti di terrorismo informatico attraverso l'istitu-

[162] Sull'incorporazione v. D. Siclari, *Contributo allo studio della sussunzione legislativa di regole formate dai privati*, in *Studi in onore di Vincenzo Atripaldi*, Vol. I, 2010, 275 ss.

[163] Sul punto si faccia riferimento alla recente proposta di modifica del regolamento (UE) 1025/2012 relativo alle decisioni delle organizzazioni europee di normazione relative alle norme europee e ai prodotti, avanzata dalla Commissione europea il 2 febbraio 2022, ove è espressamente previsto che «nei casi in cui le organizzazioni europee di normazione [quali organizzazioni private] si concentrano sul sostegno alla legislazione e alle politiche dell'UE, sono necessarie garanzie per assicurare una procedura corretta e una rappresentanza equilibrata degli interessi delle parti, coerentemente con le priorità strategiche e le esigenze legislative [...]. Questo aspetto è ancora più importante in quanto alcune organizzazioni europee di normazione sono composte principalmente da operatori economici che hanno diritto di voto e la partecipazione delle organizzazioni della società civile e delle autorità pubbliche è limitata in alcuni casi».

[164] Sul punto v. A. Rotondo, *Cyber security e protezione delle infrastrutture critiche: l'efficacia del modello europeo*, in S. Marchisio - U. Montuoro (a cura di), *Lo spazio cyber e cosmico. Risorse dual use per il sistema Italia in Europa*, Torino, 2019, 125.

[165] Cfr. art. 2, c. 1. COM(2000)890 del 26 gennaio 2001, Creare una società dell'informazione sicura migliorando la sicurezza delle infrastrutture dell'informazione mediante la lotta alla criminalità informativa.

zione di una rete di *information sharing* per la protezione delle infrastrutture critiche (la *Critical Infrastructure Warining Information Network* - CIWIN), nonché per l'erogazione di finanziamenti per la realizzazione di progetti sulla protezione delle infrastrutture critiche e il varo di una normativa riguardante le infrastrutture critiche europee che avverrà poi nel 2008 con la direttiva 2008/114/CE relativa all'individuazione e designazione delle infrastrutture critiche europee e alla valutazione della necessità di migliorarne la loro protezione[166].

Nella Comunicazione sulla protezione delle infrastrutture critiche informatizzate del 2011[167], la Commissione constatava l'insufficienza delle strategie nazionali di cybersicurezza e della resilienza dei loro sistemi e invitava gli Stati ad adottare una serie di misure basate sulla cooperazione transfrontaliera portando così all'esigenza di adottare uno specifico intervento armonizzato relativamente alla protezione delle infrastrutture critiche informatizzate.

Nel 2016 viene adottata la prima normativa sullo specifico tema della «sicurezza delle reti e dei sistemi informativi» con la direttiva (UE) 2016/1148, per l'appunto anche nota come direttiva *Network and Information Security*- NIS, oggi abrogata dalla richiamata direttiva (UE) 2022/2555 (Direttiva NIS II) entrata in vigore il 17 gennaio 2023.

Tuttavia, nonostante la rubricazione, dall'analisi dei testi emerge come il legislatore europeo abbia inteso coniugare ancora una volta la sicurezza informatica con la protezione delle infrastrutture critiche. Difatti, parte delle prescrizioni volte a «garantire un livello comune elevato di cybersicurezza nell'Unione in modo da migliorare il funzionamento del mercato interno» (art. 1), consistono perlopiù in una serie di obblighi gravanti sui soggetti individuati nella direttiva come "critici", che dovranno adempierli a pena di ingenti sanzioni amministrative[168].

Solo con il successivo regolamento (UE) 2019/881, relativo all'ENISA e alla certificazione della cybersicurezza per le tecnologie dell'informazione e della comunicazione (c.d. *Cybersecurity Act*), l'Unione ha introdotto - per la prima volta all'interno di un atto normativo giuridico - la nozione di cybersicurezza, definendola all'art. 2, n. 1, come «insieme delle attività necessarie per proteggere la rete e i sistemi informativi, gli utenti di tali sistemi e altre persone interessate dalle minacce informatiche»: nozione che è stato poi declinata negli ordinamenti dei diversi Stati membri[169].

---

166 In particolare, all'art. 2, lett. e), direttiva 2008/114/CE viene fornita la definizione di «protezione» come «tutte le attività volte ad assicurare funzionalità, continuità e integrità delle infrastrutture critiche per evitare, mitigare e neutralizzare una minaccia, un rischio o una vulnerabilità».

167 Comunicazione della Commissione al Parlamento europeo, al Consiglio, al Comitato economico e sociale europeo e al Comitato delle regioni relativa alla Protezione delle infrastrutture critiche informatizzate. Realizzazioni e prossime tappe: verso una sicurezza informatica mondiale, Bruxelles, 31 gennaio 2011.

168 Sulla "delega" della sicurezza dal potere pubblico agli amministrati seppur nell'ottica della cybersicurezza nazionale italiana v. A. Monti, *Internet e ordine pubblico*, in G. Cassano - S. Previti (a cura di), *Il diritto di internet nell'era digitale*, Milano, 2020, 79, ove l'A. scrive «[...] chi è responsabile del funzionamento dei servizi essenziali deve farsi carico in proprio della loro difesa, sopportando le conseguenze del mancato rispetto di complessi obblighi tecnici e organizzativi in termini di sanzioni amministrative particolarmente afflittive».

169 Cfr. con la nozione di "cybersicurezza nazionale" introdotta in Italia per la prima volta con il decreto-legge n. 82/2021, all'art. 1, c. 1, lett. a) come «insieme delle attività, fermi restando le attribuzioni di cui alla legge 3 agosto 2007, n. 124, e gli obblighi derivanti da trattati internazionali, necessarie per

Dal breve quadro tracciato emerge innanzitutto la distinzione fondamentale tra i due concetti di «cybersicurezza» e di «sicurezza dei sistemi informatici e di rete»[170]: mentre con quest'ultimo concetto, originariamente, sorto nel contesto delle norme tecniche, si intende garantire la continuità del servizio erogato e la protezione delle informazioni trattate dal soggetto fornitore (tra cui possono rientrare anche i dati personali) al fine di «migliorare il funzionamento del mercato interno»; le attività di cybersicurezza, che comprendono anche la sicurezza delle reti e dei sistemi informatici, sono invece dirette alla più ampia «sicurezza degli individui nel cyberspazio»[171].

A nostro parere, con questa formulazione il legislatore europeo, oltre ad aver inserito la definizione la cybersicurezza all'interno di un atto normativo vincolante, è andato al di là del mero significato tecnico della materia (ossia la tutela della riservatezza, integrità e disponibilità delle risorse informatiche e delle informazioni), giungendo al concetto di sicurezza dell'umano, secondo Alcuni riconducibile alla sicurezza dello Stato[172].

Tale definizione può pertanto essere intesa come la proclamazione di un impegno politico dell'Unione e degli Stati membri «ad assicurare una disciplina della tecnologia informatica che assicuri il rispetto delle regole democratiche necessarie alla sopravvivenza della democrazia rappresentativa propria dello Stato costituzionale»[173]. Ricostruzione che appare coerente con l'indirizzo dettato a livello internazionale dal Consiglio d'Europa con l'*Internet Governance Strategy 2016-2019* ove, a proposito della definizione di principi e regole per la *governance* di Internet, si propone di «to ensure that public policy for the Internet is people-centred, meaning that it should respect the core values of democracy, human rights and the rule of law»[174].

Tuttavia, questo processo implica una ulteriore richiesta di sovranità da parte dell'Unione europea, oltre quella già conferita dagli Stati membri nei Trattati. Come intuibile dall'analisi del concetto giuridico di cybersicurezza europea, pare che il citato processo

---

proteggere dalle minacce informatiche reti, sistemi informativi, servizi informatici e comunicazioni elettroniche, assicurandone la disponibilità, la confidenzialità e l'integrità e garantendone la resilienza, anche ai fini della tutela della sicurezza nazionale e dell'interesse nazionale nello spazio cibernetico». Sul punto sia concesso rinviare F. Serini, *La nuova architettura di cybersicurezza nazionale: note a prima lettura del decreto-legge n. 82 del 2021*, in f*ederalismi.it,* 12, 2022.

[170] Cfr. art. 6, c. 1, n. 2 della Direttiva NIS II, che definisce la sicurezza dei sistemi informatici e di rete come «la capacità dei sistemi informatici e di rete di resistere, con un determinato livello di confidenza, agli eventi che potrebbero compromettere la disponibilità, l'autenticità, l'integrità o la riservatezza dei dati conservati, trasmessi o elaborati o dei servizi offerti da tali sistemi informatici e di rete o accessibili attraverso di essi». E' opportuno notare che tale definizione richiama alcuni concetti fondamentali della *computer* e *information security* relativi alle tre proprietà fondamentali delle risorse informatiche e delle informazioni affinché queste possano essere considerate sicure, ossia, la loro riservatezza (*confidentiality*), integrità (*integrity*) e disponibilità (*availability*), spesso indicate con l'acronimo R.I.D (o C.I.A. in lingua inglese).

[171] G. Christou, *Cybersecurity in the European Union: Resilience and Adaptability in Governance Policy,* Londra, 2019, 187.

[172] L. Axworthy, *La sécurité humaine: la sécurité des individus dans un monde en mutation*, in *Politique étrangére*, 64(2), 1999, n 333 ss., ove l'A. scrive «*La sécurité humaine ne supplante pas la sécurité nationale [...]. Dan cette prospective, la sécurité humaine et la sécurité de l'État se complènt l'une l'autre*».

[173] G. de Vergottini, *Sicurezza e diritti fondamentali,* in L.E.R. Vega - L. Scaffardi - I. Spigno, *I diritti fondamentali nell'era della digital mass surveillance*, cit., 28.

[174] Consiglio d'Europa, *Internet governance - Strategy 2016-2019. Democracy, human rights and the rule of law in the digital world*, adottato dal *Committee of Ministers Deputies Meeting,* del 30 marzo 2016.

di europeizzazione vada oltre l'obiettivo di «migliorare il funzionamento del mercato interno», non interessando quindi solo l'occupazione da parte dell'Unione di spazi prima di competenza dei privati, ma anche di spazi presidiati dagli Stati per decenni.
Il progressivo accentramento dei poteri in seno all'Unione a cui stiamo assistendo deve essere interpretato alla luce del dibattuto obiettivo sul raggiungimento della "sovranità digitale europea"[175]: concetto che, sebbene dai documenti di natura politica in cui viene utilizzato coincide con la capacità dell'Unione di eliminare la dipendenza tecnologica e usare la tecnologia europea per far funzionare il mercato interno, dall'analisi delle fonti di diritto derivato proiettate verso tale fine sembrano emergere mete ulteriori.
È il caso della citata Direttiva NIS II, e del regolamento istitutivo l'ENISA, le cui basi di legittimità sono state individuate nell'art. 114 TFUE, relativo al ravvicinamento delle legislazioni degli Stati membri «al fine di migliorare il funzionamento del mercato interno». Secondo Alcuni un simile fondamento giuridico potrebbe sollevare criticità dato che «il centro di gravità di queste misure è costituito dal rafforzamento della sicurezza»[176], piuttosto che dal rafforzamento del mercato interno.
In conclusione, ci si chiede se il processo d'integrazione della sicurezza europea, che vede ancora forti resistenze nell'attuazione di pratiche di scambio informativo da parte dalle autorità pubbliche di contrasto e soprattutto da parte del settore privato variamente coinvolti nei processi di cybersicurezza, possa far fronte alle attuali esigenze dettate dal rapporto sempre più stretto tra le società europee e l'informatica..

[175] L. Floridi, *The Fight for Digital Sovereignty: What It Is, and Why It Matters, Especially for the EU*, in *Philos. Technol*, 2020, 369 ss.; H. Roberts - J. Cowls - F. Casolari - J. Morley - M. Taddeo - L. Floridi, *Safeguarding European Values with Digital Sovereignty: An Analysis of Statements and Policies*, in *Internet Policy Review*, 6, 2021, G. Finocchiaro, *La sovranità digitale*, in *Dir. pubbl.*, numero tematico, 3, 2022.

[176] Cfr. S. Poli, *Il rafforzamento della sovranità tecnologica europea e il problema delle basi giuridiche*, in *I Post di AISDUE, III, 2021, Sezione Atti Convegni AISDUE*, 5, 20 dicembre 2021, 78-79.

Note a sentenza

Tecnologie di riconoscimento facciale e diritti fondamentali a rischio: il caso *Glukhin c. Russia* dinanzi alla Corte europea dei diritti dell'uomo*

Giuseppe Gallo

Corte europea dei diritti dell'uomo, 4 luglio 2023, ric. 11519/20, *Glukhin c. Russia*

Il trattamento dei dati personali del ricorrente, un cittadino russo autore di una manifestazione solitaria di natura pacifica, effettuato tramite strumenti di riconoscimento facciale specificatamente utilizzati per identificarlo e procedere al suo arresto, non può essere ritenuto – secondo la Corte – necessario nel contesto di una società democratica. L'uso di detti sofisticati sistemi di riconoscimento facciale, da parte delle autorità della Federazione Russa, a parere dei giudici, risulta lesivo anche del diritto alla libertà di espressione del ricorrente e appare, di conseguenza, del tutto incompatibile con i valori essenziali di una società democratica – governata dai principi basilari dello stato di diritto – per il cui mantenimento e per il cui sviluppo progressivo la Convenzione stessa è stata concepita.

Sommario

1. Introduzione. - 2. La vicenda all'origine della sentenza. - 3. L'attuale impianto normativo internazionale in materia di TFR. - 4. L'articolata decisione della Corte. - 5. Osservazioni conclusive.

Keywords

tecnologie di riconoscimento facciale – libertà di espressione – art. 10 CEDU – diritto al rispetto della vita privata – art. 8 CEDU

*Su determinazione della direzione, il contributo è stato sottoposto a referaggio anonimo in conformità all'art. 15 del regolamento della Rivista

1. Introduzione

Le tecnologie di riconoscimento facciale consentono di individuare e distinguere un determinato soggetto a partire dall'esatta immmagine del suo volto[1]. Il ricorso a queste peculiari tecnologie di tipo biometrico risulta, nelle società contemporanee, sempre meno costoso e più diffuso, anche per via della versatilità e facilità dei loro impieghi[2].
Tali nuove tecnologie, come recentemente messo in evidenza dalla Corte europea dei diritti dell'uomo, nella sentenza inerente all'affare *Glukhin. c. Russia*, su cui la presente riflessione intende soffermarsi, danno origine a inedite e insidiose forme di sorveglianza di massa[3].
Dette forme di sorveglianza elettronica possono essere perpetrate, a ben vedere, per le più molteplici finalità, tanto da soggetti pubblici quanto da imprese e soggetti privati. Esse sono, quindi, alla portata di chiunque.
In ambito pubblico, le tecnologie in oggetto vengono sfruttate principalmente – ma non solo – dalle forze dell'ordine con la finalità di garantire la sicurezza e tracciare le persone in luoghi pubblici o per perseguire coloro che sono sospettati di un crimine nonché per svolgere controlli alle frontiere e, dunque, come strumento di gestione delle politiche migratorie e di rimpatrio[4].
In ambito privato analoghi strumenti sono a disposizione sia di grandi industrie che di compagnie bancarie o assicurative[5]. Questi sistemi sono inoltre utilizzati per scopi di sicurezza all'interno di comuni esercizi commerciali nonché sovente da un crescente numero di cittadini per effettuare talune tipologie di pagamenti[6].
Nell'uso degli strumenti di riconoscimento facciale, un ruolo di primo piano è svolto, tuttavia, dalle grandi piattaforme del *web* e dai cosiddetti *Big Tech*[7]. Si pensi, a titolo di esempio, al noto sistema *"Rekognition"* di Amazon oppure a quello *"Deepface"* di Face-

---

[1] S. Z. Li - A. K. Jain (eds.), *The Handbook of Face Recognition*, Berlin, 2005; K. A. Gates, *Our Biometric Future: Facial Recognition Technology and the Culture of Surveillance*, New York, 2011; J. Tanaka, *Face Recognition. The Effects of Race, Gender, Age and Species*, London, 2015; I. Berle, *Face Recognition Technology Compulsory Visibility and Its Impact on Privacy and the Confidentiality of Personal Identifiable Images*, Cham, 2020; P. Dauvergne, *Identified, Tracked, and Profiled. The Politics of Resisting Facial Recognition Technology*, Cheltenham, 2022.

[2] Cfr. *Where is facial recognition used? 11 use cases for facial recognition*, in *Thales*, 11 giugno 2023.

[3] D. Lyon, *La società sorvegliata. Tecnologie di controllo della vita quotidiana*, Milano, 2002; G. Ziccardi, *Internet, controllo e libertà. Trasparenza, sorveglianza e segreto nell'era tecnologica*, Milano, 2015; S. Zuboff, *Il capitalismo della sorveglianza. Il futuro dell'umanità nell'era dei nuovi poteri*, Roma, 2019; D. Lyon, *La cultura della sorveglianza. Come la società del controllo ci ha reso tutti controllori*, Roma, 2020; P. Perri, *Sorveglianza elettronica, diritti fondamentali ed evoluzione tecnologica*, Milano, 2020.

[4] Sulle complesse problematiche sollevate, in tema di diritti dei rifugiati, dall'uso di sistemi di riconoscimento facciale come specifico strumento di *governance* delle politiche migratorie e di rimpatrio, si veda, su tutti, L. Jasmontaite-Zaniewicz and J. Zomignani Barboza, *Automated Decisions in Asylum Applications in the EU?* in *International Journal of Refugee Law*, 33, 2021, 89 ss.

[5] G. Mobilio, *Tecnologie di riconoscimento facciale. Rischi per i diritti fondamentali e sfide regolative*, Napoli, 2021, 18.

[6] *Ivi*, 19.

[7] Per quanto concerne i cosiddetti colossi della rete si tende, di solito, ad individuarli nelle americane "GAFA", ovvero Google (e la collegata Alphabet), Amazon, Facebook ed Apple (cui si aggiungono solitamente Microsoft e IBM), e nelle cinesi "BAT" (Baidu, Alibaba e Tencent).

book e, non da ultimo, a quello *"FaceNet"* di Google[8].
Più di recente, nel corso dell'emergenza pandemica da Covid-19, i predetti sistemi di riconoscimento facciale sono stati adottati da alcuni Paesi per misurare la temperatura dei cittadini tra la folla e rilevare coloro che risultassero potenzialmente infetti[9].
Infine, vi sono Stati in cui le tecnologie in questione sono attualmente adoperate, in maniera del tutto sistematica, al fine di realizzare determinate politiche, influenzare il comportamento di massa oppure, addirittura, identificare oppositori politici e minoranze etniche o religiose[10].
Ora, le suddette tecnologie di riconoscimento facciale sono, a ben guardare, in grado di limitare le libertà fondamentali di singoli o dei gruppi sociali a cui essi appartengono e di incidere, su larga scala, sulla tutela di alcuni diritti essenziali, per via delle pervasive capacità di monitoraggio dalle stesse offerte[11].
In quest'ottica, la sentenza della Corte europea dei diritti dell'uomo assume rilevanza in quanto sancisce per la prima volta il principio secondo cui l'uso della tecnologia di riconoscimento facciale, in occasione di manifestazioni pubbliche, ma non solo naturalmente, da parte delle autorità governative, si rivela altamente invasivo e, in determinate circostanze, come quella in esame, potrebbe avere un effetto dissuasivo sull'esercizio di taluni diritti basilari della persona umana quali quello alla libertà di espressione e quello al rispetto della propria vita privata.

2. La vicenda all'origine della sentenza

Il 12 agosto 2019, un attivista politico avente cittadinanza russa, *Konstantin Kotov*, era stato arrestato, ai sensi dell'art. 212, c. 1, del Codice penale della Federazione, con l'accusa di avere ripetutamente violato le norme nazionali in tema di manifestazioni ed eventi pubblici[12]. La susseguente detenzione dello stesso ed il procedimento penale instaurato a suo carico hanno, tuttavia, attirato una grande attenzione sia da parte dei media che da parte dell'opinione pubblica, suscitando una intensa ondata di proteste nell'ambito della società civile[13].
Il 23 agosto dello stesso anno, il ricorrente, *Nikolay Sergeyevich Glukhin*, aveva viaggiato nella metropolitana della città di Mosca con una sagoma di cartone raffigurante il signor *Kotov* che reggeva nelle mani uno striscione in cui lamentava il fatto di essere stato

[8] Cfr. *Facial recognition: top 7 trends (tech, vendors, markets, use cases & latest news)*, in *Thales*, 20 febbraio 202.

[9] Sul punto vedi M. van Natta *et al.*, *The rise and regulation of thermal facial recognition technology during the COVID-19 pandemic*, in *Journal of Law and the Biosciences*, 7, 2020, 1 ss,

[10] È quanto sta avvenendo, ad esempio, con la popolazione palestinese nei territori occupati da Israele, secondo quanto denunciato da un recente rapporto di Amnesty International intitolato *"Israel and Occupied Palestinian Territories: Automated Apartheid: How facial recognition fragments, segregates and controls Palestinians in the OPT"*. Il rapporto in parola è consultabile all'indirizzo https://www.amnesty.org/.

[11] Riguardo al loro funzionamento, il riconoscimento facciale avviene tramite sofisticati strumenti di rilevazione capaci carpire immagini anche a distanza, in movimento e, di frequente, senza alcuna volontà cooperativa dell'interessato.

[12] *Mosca: 4 anni di reclusione per l'attivista Kotov*, in *Euronews*, 6 settembre 2019.

[13] *Russian court jails protester for 'repeatedly' taking part in rallies*, in *Reuters*, 5 settembre 2019.

condannato per aver preso parte a proteste pacifiche.
Attraverso l'impiego delle tecnologie di riconoscimento facciale installate nelle stazioni della metropolitana[14] e, mediante la visione di taluni audio e video, diffusi sui canali pubblici di *Telegram*, che riprendevano la suesposta scena, l'unità anti-estremismo della polizia di Mosca era riuscita a risalire al ricorrente accusandolo, ai sensi dell'art. 20, c. 2, par. 5, del Codice degli illeciti amministrativi della Federazione, di aver violato la vigente procedura in materia di svolgimento di eventi pubblici.
La normativa in discorso non richiede alcuna notifica preventiva alle competenti autorità statali nelle ipotesi di dimostrazioni individuali, eccezion fatta per i soli casi in cui il manifestante intenda fare uso di «quickly (de)assembled objects»[15].
La polizia ha però ritenuto che la sagoma di cartone utilizzata rientrasse nella citata categoria di oggetti e, conseguentemente, che il ricorrente avrebbe dovuto dare tempestivo preavviso della sua intenzione di svolgere la dimostrazione individuale.
Il tribunale distrettuale *Meshchanskiy* di Mosca, chiamato a pronunciarsi sulla questione, ha giudicato il ricorrente colpevole condannandolo al pagamento di una multa ammontante a 20.000 rubli russi ((RUB), corrispondenti a circa 283 euro[16].
Il ricorrente ha proposto ricorso in appello davanti al tribunale di Mosca contestando di aver svolto una manifestazione pacifica e che la condanna subita violasse il suo diritto alla libertà di espressione[17]. Il tribunale ha però respinto il ricorso avanzato confermando quindi la condanna in appello. Secondo il tribunale l'arresto era da considerarsi legittimo e la raccolta delle relative prove[18] era avvenuta in conformità con quanto contemplato dalla legge sulla polizia[19].
L'interessato si è allora rivolto alla Corte europea dei diritti dell'uomo sostenendo che i suoi diritti di cui all'art. 10 e all'art. 8 della Convenzione europea erano stati violati.

[14] Nel maggio 2017, il sito ufficiale del sindaco di Mosca contava che all'interno della città erano state installate più di 3.500 telecamere di videosorveglianza a circuito chiuso. Nel settembre del medesimo anno, secondo diverse fonti ufficiali, più di 3.000 di queste erano state dotate di un sistema di riconoscimento facciale in tempo reale.

[15] L. n. FZ-54 del 19 giugno 2004, sezione 7(1.1).

[16] L'art. 20, c. 2, par. 5, del Codice degli illeciti amministrativi prevede infatti che la violazione da parte di un manifestante della procedura sullo svolgimento di eventi pubblici, che non abbia tuttavia causato danni alla salute o alla proprietà di nessuno, sia punibile con il pagamento di una multa compresa tra i 10.000 e i 20.000 RUB o, in alternativa, con l'imposizione di un numero massimo di ore di servizio civile pari a 48.

[17] Il ricorrente in particolare lamentava che le attività di ricerca operativa svolte dalla polizia al fine di identificarlo erano da considerarsi illegittime poiché la relativa legge interna non consente tali attività per indagare illeciti di carattere amministrativo come quello in esame ma, al contrario, le ammette solamente in presenza di illeciti aventi natura penale. Come, del resto, affermato dalla Corte costituzione russa nella sentenza n. 86-O del 14 luglio 1998, qualora durante attività di ricerca operativa risulti evidente che l'infrazione oggetto d'indagine non sia classificabile come reato, simili attività devono essere immediatamente interrotte.

[18] A tal proposito, il capitolo 26 del Codice sugli illeciti amministrativi dispone che documenti, registrazioni audio e video, database e qualsiasi altra forma di dati sensibili possono sempre essere impiegati come prove processuali nei procedimenti amministrativi se contengono informazioni rilevanti ai fini della soluzione del caso.

[19] La legge in discorso (l. n. 3-FZ del 7 febbraio 2011) prevede, in generale, che la polizia adotti misure atte a individuare e perseguire gli illeciti amministrativi di sua competenza nonché misure miranti a prevenire e reprimere le attività estremiste.

3. L'attuale impianto normativo internazionale in materia di TFR

Prima di procedere all'analisi della decisione della Corte – esaminati i fatti all'origine della stessa – appare opportuno ricostruire, in via preliminare, lo scarno *corpus* normativo internazionale ed europeo e i numerosi atti di soft law che potrebbero venire in rilievo, nella maggior parte dei casi, quando si discorre di tecniche di riconoscimento facciale.

A tal proposito, occorre ricordare, in primo luogo, il rapporto dell'Alto Commissario per i diritti umani delle Nazioni Unite, del 24 giugno 2020, inerente all'impatto delle suindicate tecnologie sulla promozione e la protezione dei diritti umani nel contesto delle assemblee, ivi comprese le proteste pacifiche[20].

Il rapporto afferma, in sintesi, che il ricorso all'utilizzo della tecnologia di riconoscimento facciale con la sola finalità di identificare tutti coloro i quali prendono parte a una manifestazione pubblica, in assenza di apposite misure di garanzia, può determinare una oggettiva violazione del diritto alla privacy nonché una lesione delle libertà di espressione e di riunione pacifica. Secondo il suddetto rapporto l'immagine di un individuo costituisce, a tutti gli effetti, uno degli elementi chiave della sua personalità, in quanto rivela quell'insieme di caratteristiche fisiche uniche che concorrono a contraddistinguerlo dai suoi pari. Di conseguenza, la raccolta delle immagini del volto di una persona, effettuata in mancanza del suo esplicito consenso, costituisce una violazione del suo diritto alla privacy. Pertanto, qualsiasi utilizzo di sistemi di riconoscimento facciale dovrebbe necessariamente essere soggetto a idonei meccanismi di supervisione, sia di natura giudiziaria che non giudiziaria[21].

Nell'ambito del Consiglio d'Europa assumono rilevanza le linee guida in tema di riconoscimento facciale, adottate il 28 gennaio 2021, nella Giornata europea per la protezione dei dati personali, dal Comitato Consultivo della Convenzione per la protezione delle persone rispetto al trattamento automatizzato dei dati personali del 1981[22].

Tali linee guida[23] si fondano sui principi del Protocollo di modifica alla citata Convenzione adottato il 18 maggio 2018 a Elsinore (cosiddetta Convenzione 108+)[24]. Esse forniscono un'ampia serie di misure di riferimento che governi, sviluppatori e produttori di siffatti strumenti di riconoscimento facciale nonché enti pubblici e fornitori di servizi dovrebbero adottare al fine di garantire che il loro impiego non pregiudichi

[20] UN Doc. A/HRC/44/24.

[21] Il rapporto sottolinea come l'impiego di sistemi di riconoscimento facciale, nel contesto delle assemblee, comporti una lesione del diritto alla privacy su vasta scala, poiché permette la raccolta indiscriminata del volto di tutti i soggetti catturati dalla telecamera dotata o collegata a un simile modello di riconoscimento.

[22] La Convenzione (comunemente nota anche come Convenzione 108) è stata ratificata, in Italia, con la legge del 21 febbraio 1989 n. 98. Essa riveste centrale importanza poiché copre il trattamento dei dati in tutti i settori, sia pubblici che privati.

[23] Cfr. Consultative Committee of the Convention for the Protection of Individuals with Regard to the Automatic Processing of Personal Data - Convention 108, *Guidelines on Facial Recognition*, 28 January 2021.

[24] L'Italia ha ratificato la suindicata Convenzione che tuttvia non è ancora entrata in vigore.

la dignità della persona, i suoi diritti e le sue libertà fondamentali. Le suindicate linee guida coprono tutti gli impieghi delle tecnologie in parola e statuiscono, in sostanza, che la loro integrazione nei sistemi di sorveglianza preesistenti rappresenta un serio rischio per il diritto alla privacy così come per altri diritti essenziali della persona dal momento che il loro funzionamento molto spesso non richiede la consapevolezza o la cooperazione dei soggetti i cui dati biometrici vengono trattati.

Un importante riferimento alle tecnologie di riconoscimento facciale è altresì contenuto nel testo del Regolamento sull'Intelligenza Artificiale proposto dalla Commissione europea nell'aprile 2021 e approvato, con diversi emendamenti, dal Parlamento in sessione plenaria nel giugno 2023[25].

Infine, merita di essere rammentato il focus dell'Agenzia per i diritti fondamentali (FRA) dal titolo *"Facial recognition technology: fundamental rights considerations in the context of law enforcement"*. Lo studio elenca una serie di diritti della persona su cui le suesposte tecnologie sarebbero in grado di incidere e con quali modalità ciò possa concretamente verificarsi[26]. Tra i diritti elencati rientrano l'emergente diritto ad un buon governo[27], il consolidato diritto a non subire discriminazione[28] e, da ultimo, la libertà di associazione e di riunione[29].

Tra le libertà su cui le tecnologie di riconoscimento facciale sarebbero maggiormente in grado di impattare deve essere annoverata, a nostro avviso, anche la libertà di manifestazione del pensiero, verso la quale le libertà di riunione e di associazione possiedono una dimensione e una funzione strutturale. Come affermato dal Consiglio per i diritti umani delle Nazioni Unite nel rapporto del 2019 sul ricorso da parte degli Stati a tecniche di sorveglianza sempre più sofisticate, queste tecnologie possono svolgere un notevole effetto deterrente e inibitorio anche sul versante della libertà di manifestazione del pensiero[30].

---

[25] Sul punto vedi A. Alaimo, *Il Regolamento sull'Intelligenza Artificiale: dalla proposta della Commissione al testo approvato dal Parlamento. Ha ancora senso ilpensiero pessimistico?*, in *Federalismi.it*, 2023, 133 ss.

[26] Cfr. Fundamental Rights Agency, *Facial recognition technology: fundamental rights considerations in the context of law enforcement*, 27 November 2019.

[27] Il diritto ad una buona amministrazione, nel contesto in oggetto, implica in favore dei destinatari di misure di riconoscimento facciale la titolarità di una vasta categoria di diritti quali quello di accesso agli atti, quello di essere sentiti, il diritto di ottenere una decisione motivata da parte del potere pubblico interessato e, ancora, quello di difesa attraverso un giudice. Il diritto ad un ricorso effettivo davanti a un organismo giurisdizionale copre invero anche le decisioni e le misure assunte con il supporto delle tecnologie di riconoscimento facciale.

[28] I test condotti sugli algoritmi di riconoscimento facciale hanno dimostrato come vi sia un incremento del tasso di errore in relazione alla maggiore o minore età dei soggetti i cui dati biometrici sono oggetto di trattamento. Come osservato nel rapporto, le TFR potrebbero produrre, a causa del loro peculiare funzionamento, anche discriminazioni nei confronti degli individui in base al sesso o alla razza e avere conseguenze sfavorevoli sulle persone affette da disabilità. Un algoritmo di riconoscimento facciale potrebbe oltretutto non valutare accuratamente le caratteristiche di un determinato gruppo demografico con la conseguenza che date minoranze etniche potrebbero essere esposte a un maggior numero di controlli da parte delle forze dell'ordine.

[29] Gli effetti negativi dell'uso di tale tecnologia sulla libertà di associazione sono di enorme portata. Numerose persone invero potrebbero sentirsi scoraggiate dal manifestare in luoghi pubblici ed esprimere liberamente le proprie opinioni nel timore di essere identificate e subire conseguenze negative.

[30] Cfr. Human Rights Council, A/HRC/41/35, *Surveillance and human rights. Report of the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression*, 28 May 2019.

4. L'articolata decisione della Corte

Passando all'esame della sentenza, la Corte ha prima di tutto accertato la propria competenza rispetto alla particolare situazione russa. Constatato che le presunte violazioni si sarebbero verificate antecedentemente al 16 settembre 2022, data in cui la Federazione Russa ha cessato, come noto, di essere parte contraente della Convenzione, i giudici hanno ritenuto il ricorso ammissibile poiché né manifestamente infondato né tanto meno irricevibile ai sensi dell'art. 35[31].

Come chiarito nel caso *Ucraina e Paesi Bassi c. Russia*, gli artt. 38, 41 e 46 della Convenzione, continuano a trovare piena applicazione anche dopo la data in questione e, di conseguenza, l'eventuale mancata partecipazione del Governo russo nei procedimenti che lo coinvolgono non impedisce ai giudici di Strasburgo di proseguire con l'esame delle domande pendenti avanzate contro di esso[32].

Ciò poiché – come ricordato nel caso *Svetova e altri c. Russia*[33] – ai sensi dell'art. 58, par. 2, della Convenzione la denuncia di uno Stato membro della stessa «non può avere l'effetto di svincolare l'Alta Parte contraente interessata dagli obblighi contenuti nella presente Convenzione per quanto riguarda qualunque fatto suscettibile di costituire una violazione di tali obblighi, da essa posto in essere anteriormente alla data in cui la denuncia è divenuta efficace»[34].

Una volta stabilita la sua giurisdizione, la Corte ha poi proseguito con l'accertamento della presunta violazione, da parte delle autorità governative russe, sia della libertà di espressione del ricorrente che del suo diritto al rispetto della propria vita privata.

Per quanto riguarda la contestata violazione dell'art. 8, i giudici hanno ribadito, in primo luogo, come la nozione di "vita privata" sia un concetto piuttosto ampio e, come tale, non suscettibile di un'unica definizione[35].

---

[31] Le condizioni di ricevibilità dei ricorsi dinanzi alla Corte europea dei diritti dell'uomo sono rispettivamente: il previo esaurimento dei ricorsi interni da parte dell'individuo ricorrente o da parte dell'individuo vittima della violazione e la presentazione del ricorso entro il termine perentorio di quattro mesi, decorrente dalla data di ricevimento della notifica della decisione finale interna. Inoltre, nei ricorsi individuali si richiede altresì che il ricorso proposto non sia anonimo, non riguardi una questione già presentata davanti alla Corte o ad un altro organo internazionale di controllo e, ancora, che non costituisca un abuso del diritto di ricorso. Infine, una condizione di ricevibilità che è stata aggiunta di recente è quella secondo cui il ricorrente debba poi aver subito un significativo pregiudizio. Sul funzionamento della Corte europea dei diritti dell'uomo si veda, in generale, P. van Dijk (ed.), *Theory and Practice of the European Convention on Human Rights*, Antwerpen, 2018; D. J. Harris (ed.), *Law of the European Convention on Human Rights*, Oxford, 2018; A. Nussberger, *The European Court of Human Rights*, Oxford, 2020.

[32] CEDU, Grande Camera, *Ucraina e Paesi Bassi c. Russia*, ric. 8019/16 (2022), § 435 ss.

[33] CEDU, *Svetova e altri c. Russia*, ric. 54714/17 (2023), § 39.

[34] Inoltre, in base al par. 3 della citata disposizione «alla stessa condizione, cesserebbe d'esser parte alla presente Convenzione qualunque Parte contraente che non fosse più membro del Consiglio d'Europa». La versione italiana del testo della Convenzione è reperibile al link *echr.coe.int*.

[35] La concezione di "vita privata" non si limita, invero, alla sola integrità fisica, mentale o morale del soggetto, ma si estende anche a quelle interferenze suscettibili di ledere la sua identità sociale. In altri termini, l'art. 8 della Convenzione non si limita al solo diritto alla riservatezza, inteso come diritto a non vedere apprese e diffuse notizie inerenti alla propria sfera privata, ma ricomprende anche il diritto per l'individuo di condurre una "vita sociale privata", cioè la possibilità di stabilire e sviluppare rapporti con altri individui, anche beninteso nell'ambito della pluralità di attività che si svolgono in un contesto

Sulla base di questa premessa, hanno evidenziato che qualsiasi ingerenza nella sfera privata dell'individuo può essere giustificata ai sensi dell'art. 8, par. 2, solo se è conforme alla legge, persegue uno o più degli scopi legittimi in esso contenuti e appare, infine, necessaria in una società democratica per raggiungere tali scopi[36].

A tal proposito, la Corte ha ritenuto che «the questions of lawfulness and of the existence of a legitimate aim cannot be dissociated from the question of whether the interference was "necessary in a democratic society."» e – per tale ragione – «It will therefore examine them together…»[37].

Nella misura in cui il ricorrente ha sostenuto che il diritto interno non soddisfacesse il requisito della "qualità del diritto", i giudici hanno affermato come sia essenziale – nel contesto della tecnologia di riconoscimento facciale – disporre di norme sufficientemente dettagliate che disciplinino la portata e l'applicazione delle misure assunte e, allo stesso tempo, di forti garanzie contro il fondato rischio di eventuali abusi[38]. Essa ha rilevato, riguardo questo primo punto, che la legge statale russa non definiva affatto la finalità e i destinatari della raccolta dei dati, le circostanze che legittimano tale raccolta, le informazioni di cui è rispettivamente ammessa o vietata la memorizzazione, le procedure da seguire per la raccolta o, ancora, la durata della conservazione dei dati elaborati. Inoltre, il governo russo non aveva disposto alcuna garanzia di tipo procedurale che accompagnasse l'utilizzo della tecnologia di riconoscimento facciale nel suo territorio[39].

La Corte ha evidenziato come la questione oggetto del contendere non era se l'uso di tecniche di riconoscimento facciale possa ritenersi ammesso ai sensi della Convenzione bensì, in specie, se il trattamento dei dati personali[40] del ricorrente, tramite detta sofisticata tecnologia, era giustificabile ai sensi dell'art. 8, par. 2, e dunque necessario in una società democratica[41].

Uno dei fattori da tenere in conto ai fini di questa valutazione è la natura e la gravità delle offese poste in essere[42].

A tal proposito, la Corte ha osservato che il ricorrente era stato perseguito soltanto per un illecito minore classificato, peraltro, come illecito amministrativo e non come reato ai sensi dell'ordinamento russo[43].

Di conseguenza, la Corte ha concluso che l'utilizzo della tecnologia di riconoscimento

pubblico (CEDU, *Pretty c. Regno Unito*, ric. 2346/02 (2002), § 61).

36 Cfr. § 74.

37 Cfr, § 78.

38 Cfr. § 82.

39 Cfr. § 83.

40 Si ricorda che, ai sensi dell'art. 2, lett. *a)*, della direttiva 95/46/CE, i dati personali sono definiti come «qualsiasi informazione concernente una persona identificata o identificabile» mediante «riferimento ad uno o più elementi caratteristici della sua identità fisica, fisiologica, psichica, economica, sociale o culturale». Diversi sono poi i cosiddetti dati sensibili, ossia quei dati personali che rivelano l'origine razziale o etnica, le convinzioni religiose o filosofiche, le opinioni politiche, le appartenenze sindacali o che sono relativi alla salute o alla vita sessuale dell'individuo.

41 Cfr. § 85.

42 CEDU, *P. N. c. Germania*, ric. 74440/17 (2020), § 72.

43 Cfr. § 88.

facciale per identificare il signor *Glukhin* non corrispondeva a un "urgente bisogno sociale"' e costituiva, per tale motivo, una grave violazione del suo diritto alla privacy e alla protezione dei dati personali ai sensi dell'art. 8[44].

Con riferimento alla violazione dell'art. 10 la Corte di Strasburgo ha invece ribadito – in primo luogo – il principio secondo cui la protezione offerta dalla norma in commento non si limiterebbe soltanto all'insieme delle idee manifestate, nella realtà fenomenica, tramite il mero linguaggio orale o scritto, dal momento che le opinioni possono essere da chiunque espresse anche attraverso strumenti di comunicazioni non verbali o mediante determinati comportamenti[45].

Data la natura della condotta del ricorrente e il contesto in cui questa è stata attuata, i giudici hanno ritenuto che l'interessato, attraverso le sue azioni, abbia unicamente provato ad esprimere la propria opinione su una questione di interesse pubblico, rispetto alla quale le restrizioni prescritte nel secondo paragrafo dell'art. 10 ricevono, in concreto, scarsa applicazione[46].

Secondo il ragionamento seguito dai giudici, anche ammettendo che l'ingerenza potesse essere considerata conforme alla rilevante normativa statale perseguendo uno degli obiettivi legittimi della «difesa dell'ordine e prevenzione dei reati» o della «protezione della reputazione e dei diritti altrui», essa non appariva però «necessaria» in «una società democratica»[47].

La dimostrazione individuale del ricorrente, infatti, si era svolta in modo pacifico. L'illecito per cui il medesimo è stato condannato – come predetto – consisteva soltanto nell'omessa dovuta notifica, alle autorità nazionali competenti, della manifestazione tenuta[48].

Non è stato dimostrato che le azioni del ricorrente abbiano causato una grave perturbazione nello svolgimento della vita pubblica in misura superiore a quella ordinaria, né tanto meno che esse abbiano rappresentato un reale pericolo per l'ordine pubblico o la sicurezza nazionale nonché quella dei trasporti. Pertanto, i giudici – in mancanza di ulteriori elementi aggravanti – non hanno rinvenuto pertinenti o sufficienti ragioni suscettibili di giustificare l'interferenza subita dal ricorrente nel godimento del suo diritto alla libertà di espressione[49].

---

[44] Cfr, § 89.

[45] CEDU, *Karuyev c. Russia*, ric. 4161/13 (2022), § 18.

[46] Cfr. § 51.

[47] Se si guarda alla passata giurisprudenza della Corte in questa materia, si può desumere come i limiti imposti dagli ordinamenti interni, affinché possano essere valutati come strettamente necessari in una società democratica, devono essere statuiti da una normativa che li renda prevedibili ai loro potenziali destinatari, devono rispondere poi ad un bisogno sociale effettivo e preminente, apparire proporzionati allo scopo legittimo perseguito (bilanciando, di volta in volta, i diversi interessi in gioco) nonché risultare motivati in maniera convincente ed adeguata.

[48] M. Zalnieriute, *Glukhin v. Russia. App. No. 11519/20. Judgment*, in *AJIL*, 117, 2023, 697.

[49] Cfr. § 56.

5. Osservazioni conclusive

Alla luce della disamina appena svolta, è possibile, a nostro giudizio, affermare che la sentenza – la quale rappresenta il primo caso in cui le tecnologie di riconoscimento facciale siano state portate davanti ad un tribunale internazionale per dar conto delle comprovate limitazioni ai diritti essenziali che esse sono in grado di provocare – costituisca indubbiamente un importante precedente nella futura giurisprudenza della Corte sul tema della sorveglianza di massa.

Per quanto concerne la violazione dell'art. 8, abbiamo visto come i giudici di Strasburgo abbiano dichiarato pienamente applicabile anche per le suesposte tecniche di sorveglianza elettronica il principio – in precedenza affermato, ad esempio, nel caso *S. e Marper c. Regno Unito* – secondo cui al fine di valutare se le limitazioni alla tutela della vita privata e alla protezione dei dati di carattere personale siano legittime occorre congiuntamente, oltre ad una chiara e precisa previsione normativa, che la restrizione posta sia realmente «necessaria» all'interno di «una società democratica» per il perseguimento dei fini elencati nel testo dell'art. 8 della Convenzione. Per effettuare una corretta valutazione su quest'ultimo punto essenziale, l'ingerenza deve ritenersi ammessa solamente se risponde ad un bisogno sociale imperativo, se risulta proporzionata allo scopo legittimo perseguito nonché se le motivazioni addotte dalle competenti autorità nazionali per giustificare la sua applicazione appaiono oggettivamente rilevanti, pertinenti e sufficienti.

Con riferimento all'accertata lesione della libertà di espressione del ricorrente invece la Corte ha sostenuto che, pur volendo ammettere che il caso possa aver superato con esito positivo il test di legalità di cui all'art. 10, par. 2, della Convenzione, la manifesta limitazione della libertà in commento non può certamente considerarsi strettamente necessaria nel contesto di una società democratica. La restrizione lamentata dal ricorrente è illegittima poiché sproporzionata in funzione dell'obiettivo legittimo perseguito, nella specie, completamente assente o, comunque, particolarmente difficile da individuare.

Dalla pronuncia è allora possibile desumere il principio fondamentale in base al quale le summenzionate misure di riconoscimento facciale dovrebbero essere impiegate solo se legali, giustificate da un obiettivo legittimo necessario, proporzionato e adeguatamente motivato. In particolare, ci sembra che la Corte abbia respinto una mera qualificazione astratta dei motivi di ordine pubblico, richiedendo una verifica caso per caso, con la conseguenza che un generale utilizzo delle tecniche di riconoscimento facciale per motivi di ordine pubblico non sembra compatibile con alcuni diritti convenzionali se l'utilizzo concreto comporta una limitazione dei diritti umani.

Come abbiamo visto, gli strumenti di riconoscimento facciale hanno un forte impatto, diretto o indiretto, su tutta una serie di diritti fondamentali, sanciti sia a livello interazionale sia sul piano dell'Unione Europea, che possono andare ben oltre il diritto alla privacy e alla protezione dei dati personali, come tali intendendosi, ad esempio, il diritto a non subire discriminazioni, la libertà di movimento, la libertà di riunione e di associazione nonché quella di espressione e di opinione.

Ne consegue, così, che i governi che decidono di ricorrere all'impiego di una siffatta tecnologia devono, a nostro avviso, assicurarsi di farlo sulla base di un quadro norma-

tivo interno solido e conforme agli standard previsti dai principali trattati posti a tutela dei diritti umani, che contenga puntuali disposizioni idonee a proteggere efficacemente tutti i dati personali e sensibili, anche beninteso le immagini facciali e le informazioni private da esse derivanti. Tutte le persone interessate dovrebbero avere, inoltre, il diritto di accesso a tali informazioni nonché quello di chiederne la legittima rettifica o la cancellazione immediata, se necessario, qualora la loro raccolta sia stata effettuata in assenza di un'adeguata motivazione e, soprattutto, nella totale inosservanza del suindicato tripartito test di legalità, necessarietà e proporzionalità. Resta da vedere se la Corte confermerà quest'orientamento[50].

[50] La pronuncia è definitiva ma non è escluso che in altri casi vi sia l'intervento della Grande Camera.

Diffamazione e *social network*: l'attribuzione del *post* all'imputato tra prova logica e prova diretta

Andrea Ranghino

Corte di Cassazione, sez. V penale, 14 luglio 2023, n. 38755

L'attribuzione di un *post* pubblicato su un *social network* non richiede una specifica indagine sulla titolarità dell'indirizzo IP o dell'*account*, qualora vi sia una pluralità di indizi gravi, precisi e concordanti, che, nel rispetto della regola di giudizio di cui all'art. 192, c. 2, c.p.p., consentano di attribuire all'imputato i contenuti diffamatori diffusi in rete.

Sommario

1. La questione in esame. – 2. La decisione della Corte di Cassazione. – 3. I precedenti giurisprudenziali di legittimità. – 4. Qualche considerazione conclusiva.

Keywords

Cassazione – diffamazione – indirizzo IP – prova logica – social network

1. La questione in esame

La sentenza in commento, nel rispondere a uno dei motivi di gravame dedotti con il ricorso per cassazione, affronta la questione attinente alla identificazione dell'autore di un *post* dai contenuti diffamatori. Il ricorrente, infatti, dopo essere stato condannato, tra l'altro, in ordine al delitto di cui all'art. 595 c.p., aveva lamentato la violazione dell'art. 192 c.p.p., poiché, a suo dire, l'affermazione della responsabilità penale si sarebbe fondata su mere supposizioni, non essendo stato verificato l'indirizzo IP da cui sarebbe stata disposta la pubblicazione delle espressioni diffamatorie, né accertata l'identità di chi avesse in uso il profilo.

2. La decisione della Corte

La Corte di Cassazione ha disatteso tali doglianze ritenendo che non residuasse alcuna incertezza in ordine alla identificazione del ricorrente quale autore del reato, dal momento che, come motivato dal giudice di merito, la condotta diffamatoria era stata

commessa utilizzando le sue credenziali di accesso al *social network* e si era sostanziata in affermazioni coerenti con pregresse esternazioni al medesimo riferibili. Nella vicenda, inoltre, non era emerso altro soggetto interessato a tenere la condotta contestata, che, dunque, non poteva essere attribuita ad altri che all'imputato.
Il Giudice di legittimità ha poi precisato che, ai fini dell'affermazione della responsabilità penale, non occorre un accertamento tecnico sulla titolarità dell'indirizzo IP da cui sono partiti i messaggi offensivi, dal momento che il profilo del *social network*[1] può essere attribuito, nel rispetto delle regole di giudizio previste dall'art. 192 c.p.p., anche su base indiziaria in forza di una pluralità di dati convergenti e precisi, quali il movente, l'argomento del *forum* su cui avviene la pubblicazione, il rapporto tra le parti, la provenienza del *post* dalla bacheca virtuale dell'imputato e l'utilizzo del *nickname* dello stesso.
A conferma della propria decisione, infine, la Corte ha ricordato che la prova logica, quando è il risultato di un corretto procedimento valutativo degli indizi, caratterizzato da una valutazione sia unitaria sia globale dei dati raccolti, tale da superare l'ambiguità di ciascun elemento informativo considerato nella sia individualità, non rappresenta uno strumento meno qualificato rispetto alla prova diretta o a quella storica.

3. I precedenti giurisprudenziali di legittimità

Senza dubbio la questione sottoposta alla valutazione della Corte è di rilievo primario riguardando gli standard probatori da raggiungere per poter ritenere, in termini di certezza, che la persona fisica accusata del reato di diffamazione corrisponda effettivamente all'utente celato dietro il profilo di un *social network*.
Del resto, specialmente quando si tratta di scritti offensivi, da cui potrebbero discendere responsabilità anche di natura penale, non è infrequente imbattersi nel fenomeno dei profili *fake*, ossia riferibili a persone fisiche in realtà inesistenti o a persone esistenti ma diverse dal reale utilizzatore. Nella maggior parte dei casi la registrazione di *account* falsi – nei termini di cui si è appena detto – rappresenta, infatti, un'operazione tutto sommato agevole non richiedendo particolari competenze anche per la tenuità dei controlli svolti dai gestori dei *social network*. Non solo, in astratto, non si può nemmeno escludere il diverso caso dell'accesso abusivo da parte di soggetti terzi non legittimati all'utilizzo di un *account* registrato regolarmente.
A giudizio del ricorrente, prima di poter attribuire un *post* diffamatorio al titolare del profilo da cui lo stesso promana, è necessario indagare se l'identità dell'imputato coincide effettivamente con quella dell'utente che ha creato tale profilo e, nel caso, se sia stato proprio l'imputato a pubblicare lo scritto. Insomma, occorrerebbe acquisire la dimostrazione diretta circa la coincidenza tra identità reale e identità digitale.
Di diverso avviso è la Corte di Cassazione, che, in almeno quattro occasioni, aveva già affermato come un accertamento tecnico sull'indirizzo IP o sull'*account* non sia necessario se vi sono elementi in forza dei quali, nel rispetto della regola di giudizio di cui all'art. 192, c. 2 c.p.p., si può ritenere che le esternazioni diffuse in rete siano effettivamente riconducibili alla persona accusata del reato di diffamazione.

[1] Nella fattispecie Facebook.

In un primo caso[2], sebbene gli accertamenti tecnici svolti non fossero risultati idonei a dimostrare l'effettiva corrispondenza tra i messaggi dal contenuto offensivo e quanto rilevabile in rete sull'*account* dell'imputato, si era ritenuto ugualmente di poter attribuire la condotta illecita a quest'ultimo, in quanto il profilo utilizzato per veicolare le affermazioni diffamatorie era risultato effettivamente esistente, la persona accusata non aveva mai denunciato la violazione di detto profilo da parte di soggetti terzi e, soprattutto, la diffusione degli scritti offensivi era stata ammessa nell'ambito di un parallelo procedimento amministrativo.
In un secondo caso[3], l'attribuzione dell'*account* era stata dedotta dal fatto che il profilo del *social network* era a nome dell'imputato con relativa foto del titolare e che quest'ultimo non aveva mai denunciato la creazione abusiva del profilo, né richiesto la rimozione dello stesso. Si era poi tenuto conto dei ripetuti tentativi di risoluzione bonaria della controversia da parte del soggetto a cui la condotta diffamatoria era stata attribuita. La Corte, inoltre, aveva aggiunto che, stante la riconducibilità del profilo all'imputato, l'accertamento su chi, materialmente, avesse creato l'*account* o digitato le frasi offensive non fosse necessario trattandosi di circostanze irrilevanti.
In un terzo caso[4], la Corte ha ritenuto che il contenuto delle pubblicazioni, riportando con dovizia di dettagli episodi di cui l'imputato era stato protagonista e iniziative giudiziarie dal medesimo intraprese, finisse «per svolgere un'insuperabile portata individualizzante».
Nell'ultimo caso[5], infine, il Giudice di legittimità ha ritenuto corretta la decisione di attribuire un profilo Facebook sulla base del riferimento nominativo all'imputato e dell'effige fotografica dello stesso e ciò sempre in assenza di denunce relative a furti di identità da parte dell'intestatario della bacheca digitale su cui erano stati pubblicati i *post* incriminati.

4. Qualche considerazione conclusiva

In definitiva, secondo la giurisprudenza della Corte di Cassazione, se vi sono indizi gravi, precisi e concordanti, che consentono di ricondurre all'imputato gli scritti diffamatori promananti da un profilo registrato su un *social network* è del tutto inutile compiere ulteriori accertamenti tecnici finalizzati a verificare chi, effettivamente, abbia creato detto profilo o chi lo abbia usato per diffondere in rete le esternazioni offensive dell'altrui reputazioni. In tale scenario si tratterebbe, infatti, di prove superflue, atteso che il dato da verificare è già stato acquisito attraverso la prova logica.
A corollario di quanto appena osservato sembra potersi dire che, al fine di identificare l'autore di un *post* diffamatorio, ci si debba concentrare, anzitutto, su quegli elementi definiti dalla Corte di Cassazione come individualizzanti, ossia quelle circostanze di fatto, di natura indiziaria, che consento di attribuire all'accusato – e solo all'accusato –

[2] Cass. pen., sez. V, 13 luglio 2015, n. 8328.

[3] Cass. pen., sez. V, 13 luglio 2018, n. 45339.

[4] Cass. pen., sez. V, 23 settembre 2022, n. 39805.

[5] Cass. pen., sez. V, 8 febbraio 2023, n. 18883.

le esternazioni contestate come offensive dell'altrui reputazione.
Dall'esame dei precedenti giurisprudenziali è emerso che il primo aspetto da indagare è proprio il profilo *social*, da cui si possono trarre elementi utili sulla base dell'intestazione e dell'effige fotografica. Si tratta, però, di indizi a cui se ne devono aggiungere altri, idonei a superare ogni dubbio e, in particolare, quello che si tratti di un profilo falso o che vi sia stato un accesso abusivo da parte di terzi. In questo senso la Corte, al di là di eventuali comportamenti – espressamente o implicitamente – confessori, ha ritenuto rilevante la mancata denuncia di furti d'identità o di accessi abusivi all'*account* e, soprattutto, l'esistenza di elementi che rendono l'accusato la sola persona in grado di compiere tali esternazioni vuoi in ragione della conoscenza esclusiva di determinati dati o avvenimenti, vuoi per i rapporti pregressi con la persona la cui reputazione è stata lesa.
Solo se le circostanze emerse non consentiranno di raggiungere gli standard previsti dal citato art. 192, e di superare tali dubbi, si dovrà andare a indagare ambiti ulteriori, quali la titolarità dell'indirizzo IP o quella dell'*account* utilizzato per la diffusione in rete delle espressioni diffamatorie.

I criteri per la determinazione della competenza territoriale nella diffamazione telematica: l'accertamento dell'evento tra rigore tecnico e ricorso alle presunzioni

Federico Riboldi

Corte di Cassazione, sez. V, 18 settembre 2023, n. 38144

Il ricorso ai criteri suppletivi dettati dall'art. 9 c.p.p. assume carattere residuale per il caso che non sia possibile accertare il luogo della consumazione del reato.
La diffamazione è un reato di evento che si consuma nel momento e nel luogo in cui i soggetti – terzi rispetto all'agente all'offeso – percepiscono l'aggressione offensiva.
L'e-mail è una comunicazione diretta a un destinatario predefinito ed esclusivo, al quale viene recapitata informaticamente presso il server di adozione, collegandosi al quale attraverso un proprio dispositivo e utilizzando delle chiavi di accesso personali, questi può prenderne cognizione.
Mentre per la comunicazione veicolata dal web o dai social media il requisito della comunicazione con più persone può presumersi sulla base dell'inserimento del contenuto offensivo nella rete, per accertare l'effettiva realizzazione dell'evento lesivo nella diffamazione a mezzo e-mail è necessaria quantomeno la prova dell'effettivo recapito della corrispondenza elettronica, sia esso la conseguenza di un'operazione automatica impostata dal destinatario ovvero di un accesso dedicato al server.
La lettura delle e-mail da parte dei destinatari può presumersi salvo prova contraria.
Contro le sentenze di appello pronunziate per reati di competenza del Giudice di Pace non può essere proposto ricorso per cassazione per motivi diversi da quelli previsti dalle lett. a), b) e c) dell'art. 606 c.p.p., rimanendo dunque inibita la prospettazione di meri vizi della motivazione.

Keywords
diffamazione – corrispondenza – diritto penale – e-mail – mezzi di comunicazione

La Corte di Cassazione, con la sentenza n. 38144 del 27 giugno 2023, prosegue nella sua opera, ormai più che ventennale, volta a risolvere le problematiche di diritto sostanziale e processuale che caratterizzano il reato di diffamazione in rapporto con i nuovi mezzi di comunicazione.

La vicenda trattata dalla V Sezione penale della Suprema Corte ha ad oggetto l'invio a plurimi destinatari di un messaggio di posta elettronica ritenuto offensivo della reputazione altrui. A fronte di una fattispecie, quella della corrispondenza elettronica a contenuto diffamatorio, ormai ben conosciuta nell'esperienza giurisprudenziale, la Corte affronta il tema – dedotto con uno dei motivi di ricorso – concernente la determinazione della competenza territoriale e lo *standard* probatorio richiesto per accertare i presupposti che la radicano avanti a una determinata Autorità Giudiziaria.

La decisione, dopo aver ribadito che il ricorso ai criteri suppletivi dell'art. 9 c.p.p. ha natura residuale (operante solo quando non sia in alcun modo ricavabile il *locus commissi delicti*), muove dalla qualificazione, che rappresenta principio di diritto ormai consolidato nella giurisprudenza di legittimità[1], della diffamazione quale reato di evento che si consuma nel momento e nel luogo in cui i soggetti – terzi rispetto all'agente all'offeso – percepiscono l'aggressione offensiva.

Da qui, poi, la disamina delle implicazioni di ordine "tecnico", che derivano dall'utilizzo dello strumento telematico, e la necessità di distinguere tra le diverse forme di comunicazione informatica. La Corte, che ha negli anni elaborato competenze "digitali", osserva, in linea con una sua precedente pronuncia, pressoché coeva (Cass. pen., sez. V, 24 marzo 2023, n. 12511), come «l'e-mail sia una comunicazione diretta a destinatario predefinito ed esclusivo (anche quando plurimi siano i soggetti cui viene indirizzata), al quale viene recapitata informaticamente presso il server di adozione, collegandosi al quale attraverso un proprio dispositivo e utilizzando delle chiavi di accesso personali, questi può prenderne cognizione».

Inquadrata, così, la modalità attraverso cui l'e-mail giunge nella disponibilità del destinatario, perfezionando l'evento del reato, i Giudici di legittimità pongono una distinzione con l'ipotesi, parimenti esaminata in alcune precedenti pronunce, in cui scritti, immagini o *files* vocali vengono caricati su siti web o diffusi sui social media. Distinzione, quella introdotta dalla Corte, da cui derivano conseguenze di ordine giuridico.

Nella prospettiva della Cassazione, mentre per la comunicazione veicolata dal web o dai social media il requisito della comunicazione con più persone – elemento strutturale del reato di diffamazione – può presumersi sulla base dell'inserimento del contenuto offensivo nella rete, per quella inviata attraverso lo strumento dell'e-mail, lo *standard* probatorio adottato diviene più rigoroso, essendo «necessaria quantomeno la prova dell'effettivo recapito degli stessi, sia esso la conseguenza di un'operazione automatica impostata dal destinatario ovvero di un accesso dedicato al server».

Per altro verso, pur essendo richiesta la prova che il messaggio sia stato effettivamente "scaricato" – espressione che la Corte declina come trasferimento sul dispositivo dell'utente dell'indirizzo – la Corte ritiene, sempre in linea con la sua precedente giurisprudenza, che l'effettiva lettura può presumersi, salvo prova contraria.

[1] In questo senso si vedano, tra le altre, Cass. pen., sez. V, 14 dicembre 2022, n. 2251; Cass. pen., sez. V, 22 ottobre 2018, n. 55386

Appare apprezzabile l'approdo della Corte che – anche in coerenza col principio di offensività – introduce uno *standard* probatorio minimo che non consente di arrestarsi alla condotta (l'invio della corrispondenza elettronica) ma, richiedendo che sia accertata la realizzazione dell'evento di danno, evita il rischio di trasformare, nei fatti, la fattispecie in un reato di pericolo.

L'occasione, forse, sarebbe stata propizia, nella medesima direzione di rigore quanto agli standards dell'accertamento, per superare anche il principio di diritto, più sopra richiamato, secondo cui la prova dell'effettiva lettura della corrispondenza – che concretizza la lesione della reputazione - può essere presunta. Il ricorso a presunzioni, criterio opinabile in costanza del principio informatore dell'"oltre ogni ragionevole dubbio", di fatto determina un'inversione dell'onere probatorio ben difficile da assolvere, imponendo all'imputato di dimostrare che i destinatari delle e-mail non abbiano preso cognizione del loro contenuto; prova che diviene persino diabolica in casi come quello esaminato, in cui la corrispondenza sia indirizzata a un numero rilevante di destinatari (ben 450 nella fattispecie oggetto della sentenza). V'è da dire, però, che nel caso trattato dalla Suprema Corte la prova circa il perfezionamento dell'evento diffamatorio risultava in concreto acquisita non mediante il ricorso a presunzioni, ma sulla scorta delle dichiarazioni testimoniali di due destinatari della corrispondenza elettronica, non contestate dal ricorrente, che in Palermo l'avevano ricevuta.

La decisione rileva anche sotto un diverso profilo, allorché tratta il tema, dedotto con un secondo motivo di ricorso, concernente la responsabilità dell'imputato e il riconoscimento dell'esimente dal medesimo invocata (esclusa dal Tribunale, quale giudice d'appello avverso la pronuncia del Giudice di Pace, per l'assenza di continenza delle espressioni utilizzate e persino per la mancanza di veridicità di alcuni dei fatti esposti nella corrispondenza).

La Corte ha ritenuto la doglianza processualmente inammissibile perché afferente a un mero vizio della motivazione, non deducibile in Cassazione rispetto a sentenze che siano, come nel caso di specie, relative a reati di competenza del Giudice di Pace.

Il principio di diritto espresso dalla Corte è aderente alla disciplina risultante dal combinato disposto degli artt. 606, c. 2-*bis,* c.p.p. e 39-*bis* d.lgs. 274/2000, che limitano l'ipotesi del gravame di legittimità ai soli casi contemplati nelle lett. a), b) e c) dell'art. 606 c.p.p.

L'insegnamento della Corte, persino scontato visto il tenore inequivoco delle norme, conferma che eventuali doglianze afferenti il "merito", in procedimenti di diffamazione che rimangano confinati nella competenza del Giudice di Pace (per l'assenza delle aggravanti previste dal c. 3 e 4 dell'art. 595 c.p.) potranno essere prospettate solo se sfocino in un'erronea applicazione della legge penale (o di altre norme giuridiche, di cui si deve tener conto nell'applicazione della legge penale), mentre ogni censura che attenga alla completezza, coerenza, logicità della motivazione, rimane preclusa.

Cronache

La tutela del pluralismo nell'ecosistema digitale

Augusto Preta

Sommario

1. Premessa. - 2. Introduzione. - 3. La televisione nel nuovo ecosistema digitale. - 4. Il mercato rilevante della tv. - 4.1. Il mercato rilevante nell'analisi antitrust. - 4.2. Il mercato rilevante della televisione nell'ecosistema digitale. - 4.3. Il potere di mercato. - 5. Il significativo potere di mercato e la tutela del pluralismo. - 5.1. Informazione e pluralismo nell'ecosistema digitale. - 5.2. Il pluralismo e la sua misurazione. - 5.3. La tutela del pluralismo nell'ecosistema digitale. - 6. Considerazioni conclusive

Keywords

pluralismo – ecosistema digitale – pluralismo informativo – televisione – mercati rilevanti

1. Premessa

L'AGCOM ha dato avvio al procedimento per l'adozione delle linee guida volte a definire la metodologia specifica per la verifica dell'esistenza di posizioni di significativo potere di mercato lesive del pluralismo di cui all'art. 51, c. 5, d.lgs. 208/2021.

Il presente studio mira a comprendere come l'evoluzione del mercato e le trasformazioni in atto possano modificare le tradizionali definizioni dei mercati e incidere conseguentemente nell'identificazione del significativo potere di mercato in tema di pluralismo.

Il lavoro si compone di tre capitoli.

Il primo capitolo analizza il contesto di riferimento e le profonde trasformazioni degli ultimi anni nel mondo televisivo. In particolare, il paper evidenzia come lo sviluppo dei servizi streaming nel nuovo contesto digitale abbia trasformato radicalmente il sistema televisivo, in particolare nell'ultimo biennio, divenendo parte di un ecosistema più ampio, e accentuando forme di ibridazione dei modelli di business. L'ingresso di nuovi operatori globali, in grado di competere sui diversi mercati nazionali e offrire una pluralità di servizi, a costo zero (pubblicità) o molto ridotto, ha profondamente modificato la struttura di mercato, accentuando gli elementi di pressione competitiva tra TV in chiaro e a pagamento e determinando le condizioni per il superamento di tale distinzione in termini di mercato rilevante.

Il secondo capitolo approfondisce come tali processi hanno reso più competitivo e contendibile il mercato televisivo, a partire dal settore della pay TV, dove la riduzione delle barriere all'ingresso e la crescita consistente delle offerte *on demand*, ha ampliato

la concorrenza e ridotto il livello di concentrazione.
Nel terzo capitolo viene approfondita la distinzione tra pluralismo e concorrenza e la necessità di definire strumenti specifici, idonei a misurare il livello di pluralismo, abbandonando la strumentazione antitrust, in linea con la nuova formulazione dell'art. 51 TUSMA, e i meccanismi basati unicamente sulla proprietà dei media e sulle quote di mercato legate ai fatturati. In coerenza con le risultanze degli studi di settore e l'evoluzione del quadro normativo e regolamentare a livello europeo, il focus si concentra verso modelli legati al mercato dell'attenzione, con particolare riferimento agli elementi di *audience* dei soli programmi/contenuti d'informazione. Il tutto in un quadro anch'esso di profonda trasformazione, in cui l'evoluzione dei servizi online ha determinato un vero e proprio cambio di paradigma, ponendo sfide significative, in particolare nel settore dell'informazione, quali il fenomeno delle *fake news* e la lotta alla disinformazione.

2. Introduzione

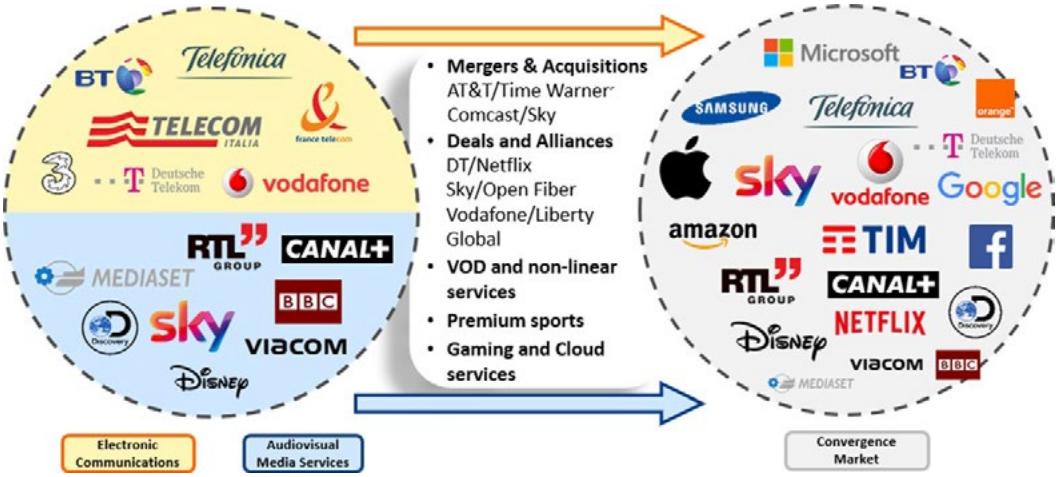
Il settore dei media è parte fondamentale dell'ecosistema digitale. Rappresenta una componente significativa dell'economia nazionale ed è uno dei pilastri delle nostre democrazie, fondate sulla libertà e il pluralismo dei mezzi d'informazione.
Questa duplice natura ha spinto in passato il legislatore a sovrapporre gli ambiti, estendendo la classica strumentazione antitrust, a tutela della concorrenza, a garanzia del pluralismo[1].
Il nuovo TUSMA, in particolare all'art. 51, muove da una diversa prospettiva, eliminando l'automatismo con cui il divieto di posizioni dominanti lesive del pluralismo viene applicato al superamento dei limiti anti-concentrativi basati sui ricavi, che diventa solo un indice sintomatico di potere di mercato e pone in capo all'Autorità l'adozione di linee guida, volte a definire la metodologia specifica per la «verifica dell'esistenza di posizioni di significativo potere di mercato lesive del pluralismo».
Al contempo la stessa norma elenca una lista di "indicatori" a cui l'AGCOM deve ispirarsi, che alla luce delle innovazioni dirompenti che hanno rimodellato l'industria dei media nell'era digitale (vedi cap.1), spinge a superare le ambiguità e le incertezze della precedente normativa, di tipo "adattativo", in cui la strumentazione antitrust rappresentava l'unica risorsa a disposizione del regolatore per misurare il significativo potere di mercato, e in ultima analisi, per tutelare il pluralismo (cap. 2).
La normativa attuale, che consente di intervenire in chiave diversa dal passato, offre l'opportunità di uscire da questo equivoco. La tutela del pluralismo, un bene non solo economico, che proprio nel contesto della trasformazione digitale assume connotazioni nuove, macroscopiche e ancor più degne di tutela – vedi la lotta alla disinfor-

[1] Il precedente TUSMAR ne era un evidente esempio, nel momento in cui prevedeva, in particolare all'art. 43, un'articolata disciplina sul divieto di costituzione (e mantenimento) di posizione dominante basata sul rispetto di precisi limiti anti concentrativi economici, commisurati ai ricavi e alle quote di mercato dei singoli operatori. Se rispetto alla pratica antitrust non richiedeva l'accertamento di comportamenti anti-competitivi che ne dimostrassero l'abuso, tutto il resto sostanzialmente ricalcava l'analisi dei mercati tipica delle autorità di concorrenza.

mazione- - richiede infatti di operare in una prospettiva diversa, con caratteristiche sempre più distintive rispetto alla tutela della concorrenza (cap. 3).
Il presente paper, nell'approfondire tutti questi aspetti, ha l'obiettivo di fornire una utile chiave di lettura in tal senso, che tenga conto della complessità del fenomeno e delle possibili ricadute a livello regolamentare, così da fornire un contributo concreto alla migliore definizione delle linee guida da parte di AGCOM.

3. La televisione nel nuovo ecosistema digitale

La televisione è da tempo al centro di un grande processo di trasformazione, a livello globale, legato all'esplosione del video streaming, ulteriormente accentuato dall'emergenza Coronavirus.
Lo sviluppo di internet e la conseguente rivoluzione digitale, hanno finalmente reso possibile quel processo, a lungo vagheggiato nell'ultima parte del secolo scorso, di convergenza tra infrastrutture/reti e servizi/contenuti: il primo tradizionalmente in capo al settore delle telecomunicazioni o comunicazioni elettroniche e il secondo a quello dei media.
Il nuovo contesto competitivo



Fonte: Pogorel/Preta per Fondazione Schumann, 2020

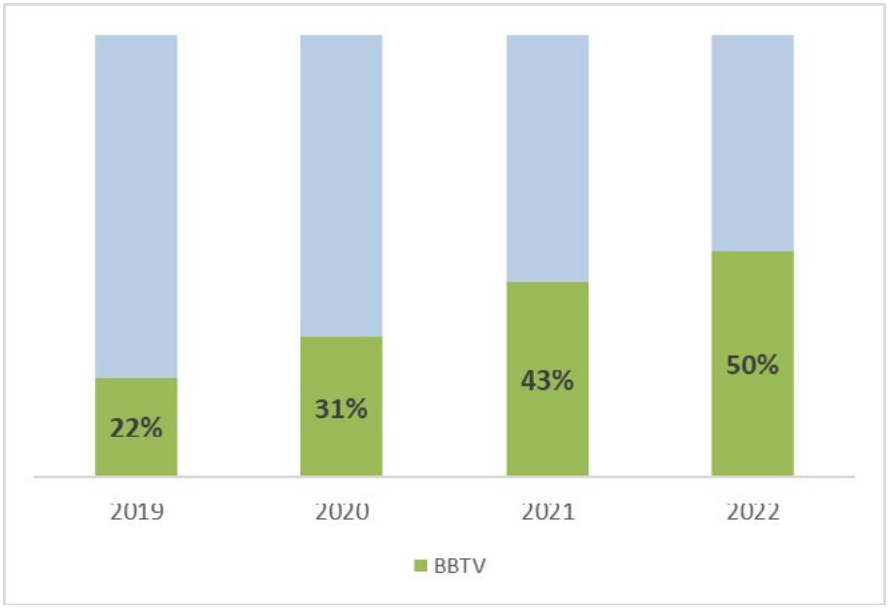
L'innovazione digitale, nel favorire tale processo, ha reso possibile la creazione di un più ampio ecosistema, come illustrato nella figura sopra, accrescendo la complessità e la reciproca interdipendenza tra quelli che un tempo erano mercati distinti e favorendo l'emergere di nuovi soggetti, le piattaforme internet (Google, Facebook, ecc.) e i fornitori di servizi di *video-streaming* (Netflix).
In questo senso, la pandemia ha rappresentato un punto di non ritorno anche in Italia, rimasta a lungo indietro nel processo di digitalizzazione, a seguito dei problemi infrastrutturali legati allo sviluppo della banda larga.
Infatti, nel periodo del *lockdown*, il tempo trascorso *online* è cresciuto costantemente, incoraggiando le persone che erano bloccate in casa a godersi sempre più l'intratteni-

mento in *streaming*.
Per categorie specifiche, meno inclini all'uso della tecnologia digitale, la proposta di prova gratuita durante la pandemia ha rappresentato un incentivo importante per familiarizzare e abituarsi per la prima volta ai servizi di *Video on demand* (VOD).
Come conseguenza di tutto ciò, il primo risultato evidente è stato la progressiva migrazione del pubblico dalla tv lineare verso i servizi di video streaming a banda larga.
In appena quattro anni, grazie soprattutto all'ingresso di nuovi operatori e alle offerte SVOD di servizi come Netflix, Disney+, Amazon Prime Video, Now TV, il settore della broadband TV e del video streaming ha assunto dimensioni rilevanti, essendo utilizzato non più soltanto da consumatori con specifiche caratteristiche demografiche (*millennials*), ma anche dal resto della popolazione, dando luogo a fenomeni di sostituzione (*cord-cutting*) e aumentando il numero degli attori e la competizione su questo terreno.
A ciò si è aggiunto, più di recente, un ulteriore vettore rappresentato dal passaggio dei diritti sportivi più pregiati dalle reti televisive broadcast ai servizi online, che ha spinto ad accelerare ancor più questo processo, favorendo la migrazione di parte degli abbonati della tv a pagamento lineare verso le offerte di video streaming.
In questo modo quello che non si era verificato in Italia nell'arco di decenni, è accaduto in pochi mesi.
Un cambiamento così radicale non è stato però soltanto il frutto della congiuntura macro-economica o dell'emergenza sanitaria, ma, sebbene favorito da questi fattori, è il risultato di una vera e propria trasformazione culturale nei comportamenti e nei consumi legati all'intrattenimento, che non è più solo generazionale, ma si è estesa ad altre fasce demografiche e ha coinvolto la maggior parte della popolazione.
Nei successivi due anni (2021-2022) il fenomeno è proseguito e lo streaming è diventato una presenza fissa e irrinunciabile nell'utilizzo del tempo libero nella maggioranza delle case degli italiani, facendo sì che ciò che appariva "straordinario" prima del COVID-19 sia diventato "normale" oggi.
Questa forte discontinuità è stata evidenziata anche dall'AGCOM nell'ultima sua Relazione Annuale, in cui afferma come «le piattaforme '*on demand*' dopo aver segnato un deciso balzo nel 2020 (passando da 11,2 a 14,3 milioni di utenti unici di siti e app nel mese medio dell'anno), presentano un trend ancora in crescita nel 2021, raggiungendo i 14,9 milioni di utenti unici»[2].
Più di recente, uno studio ITMedia Consulting mostra come si è più che raddoppiato, in appena 3 anni, il numero di famiglie abbonate a tali servizi, oltre 12,5 mln di abitazioni nel 2022, rispetto ai 5,9 mln del 2019[3].

[2] Relazione annuale AGCOM 2022

[3] ITMedia Consulting, *Il mercato televisivo in Italia 2022-2024: ibridi e connessi*, in *itmedia-consulting.com*, novembre 2022.

Evoluzione nella penetrazione della streaming TV (% su totale abitazioni TV)



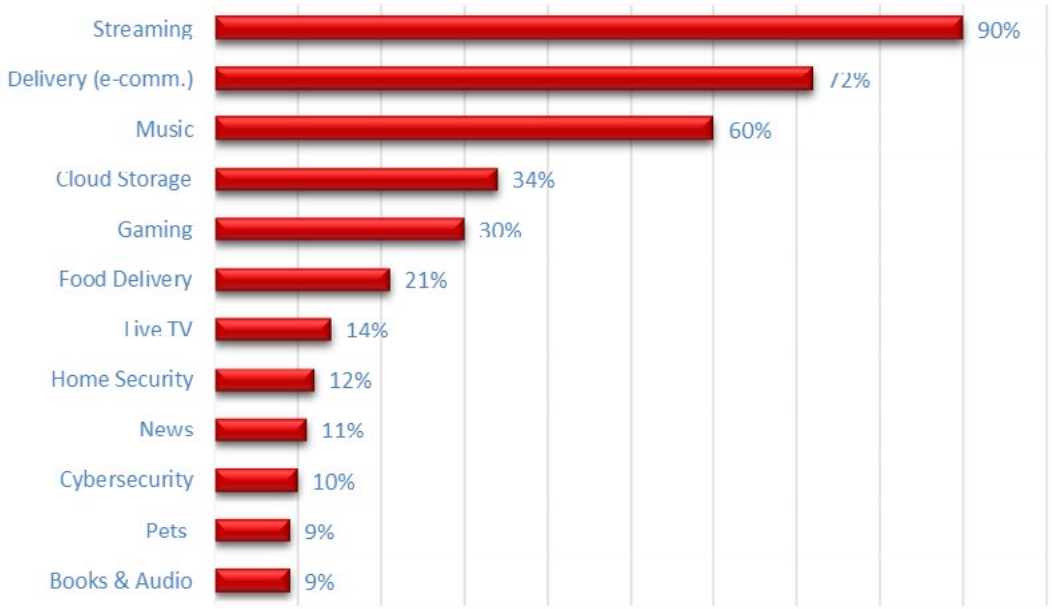
Fonte: ITMedia Consulting, 2022

Allo stesso tempo, però, tutto ciò che appare consolidato ora, viene immediatamente rimesso in discussione. Se dunque l'effetto sostituzione del video streaming si era concentrato fin qui soprattutto, ma non solo, sul comparto della TV pagamento (*pay-TV*), attualmente la difficile congiuntura economica e il generale clima d'incertezza fanno emergere i primi segnali di crisi dello streaming a pagamento (SVOD). In questo senso gli stessi operatori SVOD, a cominciare da Netflix, cercano di rispondere alla contrazione degli abbonamenti con la stretta sulla condivisione degli account, che però mostra al contempo anche una crescente fragilità del modello di business.
Ne consegue che il consumatore tende ad abbandonare (*churn*) l'abbonamento ad almeno una parte di questi operatori, riducendo così il fenomeno del *multihoming* sulla componente a pagamento, cioè del numero di servizi SVOD a disposizione di ogni abitazione, che ha caratterizzato fin qui il settore, per spostare parte di essi sui servizi online finanziati dalla pubblicità (AVOD), come quelli di altri operatori globali come Pluto, Warner Bros.Discovery o su servizi "ibridi", recentemente lanciati dalla stessa Netflix (e a seguire da Disney), dove l'accettazione della pubblicità comporta una riduzione del prezzo di abbonamento.
In questo modo, via servizi AVOD, si amplia la concorrenza diretta con la TV in chiaro, mentre sulla pubblicità online si trasferiscono nuove ulteriori risorse, che andranno a ridurre quelle della televisione lineare in chiaro, e dove già oggi nuovi operatori "non-televisivi" raccolgono la gran parte della torta pubblicitaria.
Ciò comporta rilevanti conseguenze, che meritano di essere sottolineate.
La prima riguarda l'evoluzione dello streaming e il passaggio dal modello *library* a utilità ripetuta (prevalentemente film e serie TV) a quello che ingloba anche gli eventi in diretta dal vivo (sport), fino a diventare un'alternativa a 360 gradi alla tradizionale radiodiffusione (broadcasting) televisiva[4].

[4] A. Preta, *Lo sport come game changer*, in *Agenda Digitale*, 2 luglio 2021.

La seconda, proprio nella logica di ecosistema precedentemente segnalata, è correlata a una crescente competizione per la conquista del pubblico, non solo tra una sempre più affollata e variegata schiera di *streamer*, e tra questi e i tradizionali *broadcaster,* ma anche con altre forme di intrattenimento.

Le categorie più popolari tra gli utenti internet



Fonte: Variety Intelligence Platform su Forbes Advisor, Dicembre 2022

Infatti, se attualmente le persone hanno a disposizione un maggior numero di opzioni a pagamento e gratuite per i servizi di streaming video, che rappresentano la modalità più popolare di accesso a internet (vedi grafico sopra), al contempo anche i videogiochi sono in crescita non più solo tra i millennials, come pure un numero crescente di utenti si rivolgono ai social media per l'intrattenimento e le notizie (YouTube, Tik-ToK, Twitter, Facebook, ecc.).
Indubbiamente, sia i servizi di social media e le piattaforme online che attraggono e fidelizzano i clienti grazie alla fornitura di feed video personalizzati, per lo più generati dagli utenti, con costi di produzione bassi o nulli, sia il videogioco, con il suo enorme fascino interattivo che coinvolge più attivamente l'utente, sono contenuti molto attraenti e in competizione diretta nella battaglia per l'attenzione e il tempo dei consumatori.[5]
Questa dimensione sistemica, plurale, interconnessa e strettamente interdipendente è

[5] «Oggi l'obiettivo è la fidelizzazione piuttosto che l'espansione. L'utilizzo della pubblicità è un modo per far avanzare il mercato, ma l'offerta si svilupperà in forme nuove in vari modi. È il caso della *gamification*, che unisce elementi di *game design* e principi di gioco nello streaming (gvod). L'obiettivo è consentire al pubblico di essere partecipativo e ottenere maggiore coinvolgimento. La competizione non è solo fra servizi simili ma sul tempo dei consumatori e sulla maggiore capacità di attrazione di tali servizi. Questo apre nuove opportunità per tutti i *player* nel mondo dell'intrattenimento, ma l'accresciuta competizione non avverrà più all'interno di mondi chiusi e separati come accaduto in passato». A. Secchi, *Streaming, crescita con gli spot,* in *Italia Oggi,* 23 maggio 2023

opportunamente riaffermata dalla stessa Commissione europea, che nel suo recente rapporto sull'industria dei media in Europa[6], definisce 3 diversi ambiti (sottosettori) - audiovisivo, videogiochi, informazione -, all'interno del più ampio contesto del sistema dei media.

In coerenza con quanto dalla stessa Commissione affermato, in definitiva lo scenario che emerge è di una continua, incessante e inarrestabile trasformazione, guidata dall'innovazione digitale (*Big data*, AI, *blockchain*, IoT, ecc.) che pone al centro il consumatore, con le grandi opportunità che questa propone, ma anche con i rischi e le preoccupazioni di natura etica e sociale, sui quali il ruolo della regolazione, a partire dalla tutela del pluralismo, diventa sempre più essenziale (come approfondito nel cap. 3).

Da questo grande processo di ricomposizione dell'intero settore televisivo, che abbiamo qui cercato di delineare, discendono alcune importanti conseguenze:

1. La prima riguarda la necessità di rivedere la definizione dei mercati alla luce delle profonde trasformazioni in atto. Se infatti la competizione sul tempo ha storicamente rappresentato, già nel mondo televisivo analogico, un fattore di pressione competitiva nella definizione del mercato rilevante, a maggior ragione questo vale nel nuovo ecosistema digitale.
2. La seconda, conseguenza della prima, è che anche la definizione di significativo potere di mercato va adeguata e misurata rispetto al nuovo e più ampio contesto competitivo, in cui un numero crescente di attori offrono un'ampia gamma di servizi in concorrenza tra loro, ampliando in questo modo le possibilità di scelta del consumatore e determinando al contempo una riduzione delle barriere all'ingresso.
3. La terza considerazione è relativa all'impatto che i fenomeni sopra descritti avranno sul pluralismo e sulla sua tutela, questione che assume una valenza fondamentale alla luce della mutata cornice normativa, rappresentata dal nuovo TUSMA.

Tutti questi aspetti vengono approfonditi nei capitoli che seguono.

4. Il mercato rilevante della TV

4.1 Il mercato rilevante nell'analisi antitrust

Come noto, un mercato rilevante è costituito da un prodotto (o un insieme di prodotti) e un'area geografica in cui è venduto, su cui è possibile esercitare un potere di mercato. Per definire dunque il mercato rilevante del prodotto – all'interno di un'area geografica omogenea – occorre considerare quell'insieme di prodotti che agli occhi del consumatore sono "intercambiabili" o "sostituibili". Di fatto è proprio la possibilità di sostituire il prodotto di un 'impresa con quello di un'altra impresa, che determina il livello di competitività del mercato e i vincoli concorrenziali che sono in grado di limitare il potere di mercato dell'impresa[7].

Storicamente, dall'ultimo decennio del secolo scorso, quando le *pay-TV* hanno iniziato ad affermarsi in Europa, la definizione di mercato rilevante è rimasta sostanzialmente

[6] Commissione europea, *The European Media Industry Report,* maggio 2023

[7] A. Preta, *Televisione e Mercati rilevanti*, Milano, 2012.

invariata, basata sulla distinzione tra due mercati, quello in chiaro, finanziato dalla pubblicità e quello a pagamento, finanziato direttamente dal consumatore finale tramite abbonamento (*pay TV*) e pagamento per singoli contenuti (PPV).
Ciò era una conseguenza del fatto che nonostante le autorità antitrust in Europa avessero riconosciuto l'esistenza di una pressione competitiva da parte delle tv in chiaro nella decisione del telespettatore di abbonarsi o meno ai canali a pagamento, questa non era tale da consentire la sua sostituibilità con questi ultimi.
In sostanza la presenza di contenuti *premium*, in esclusiva (film e calcio), destinati unicamente alle pay-TV, era l'elemento distintivo che determinava, sia lato domanda che lato offerta, l'esistenza di due mercati distinti
Successivamente, con il passaggio del *broadcasting* televisivo dall'analogico al digitale, in un panorama televisivo non più così rigidamente distinto, in virtù sia dello sviluppo di canali tematici sia in chiaro che a pagamento (*basic*) sia del maggiore impatto della pubblicità sul business delle offerte *pay* (*premium*), il tema della sostituibilità è emerso con più forza, attenuando da un lato i confini tra i due modelli, non più cosi rigidamente distinti, e dall'altro richiedendo un'analisi caso per caso per verificare se l'accresciuta pressione competitiva potesse configurarsi come vera e propria sostituibilità, all'interno di un unico mercato televisivo[8].
Di fatto però, anche nelle più recenti decisioni dell'Autorità Garante della Concorrenza e del Mercato (AGCM) la divisione è rimasta in vigore. Richiamandosi alla prassi nazionale ed europea[9], l'AGCM ricorda che:

> «il mercato della *pay-tv* sia tradizionalmente considerato distinto dal mercato della televisione in chiaro e, più specificamente, dal mercato della raccolta pubblicitaria su mezzo televisivo, benché tra questi intercorrano collegamenti derivanti da relazioni orizzontali. Ciò, in primo luogo, in quanto i programmi e i contenuti "*premium*" normalmente distribuiti attraverso la *pay-tv* non sono di norma sostituibili con i programmi e i contenuti trasmessi dai canali in chiaro. Inoltre, la principale fonte di ricavi degli operatori attivi nella pay-tv è rappresentata in misura quasi totalitaria dagli abbonamenti e dalla vendita di programmi in *pay-per-view*; invece, i ricavi degli operatori attivi nel mercato della televisione in chiaro derivano invece dalla pubblicità e/o dai contributi statali. Pertanto, le caratteristiche della domanda e dell'offerta sono profondamente diverse nei due mercati e risulta limitata anche la sostituibilità dal lato dell'offerta tra televisione *free-to-air* e *pay-tv*»[10].

In generale, sono diverse le obiezioni che possono essere mosse a questa lettura molto

[8] Ivi, 70 ss.

[9] AGCM, Provvedimento n. 27784, C12207 – *Sky Italia/R2*, in Boll. n. 21/2019, 20 maggio 2019; AGCM, Provvedimento n. 18932, A407 – *Conto TV/Sky Italia*, 2 ottobre 2008; AGCM, Provvedimento n. 15632, A362 – *Diritti calcistici*, 28 giugno 2006; AGCM, Provvedimento n. 24206, A418C – *Procedure selettive Lega calcio 2010/11 e 2011/12*, 6 febbraio 2013; Decisione della Commissione, COMP/M.5121 – *News Corp/Premiere*, 25 giugno 2008; Decisione della Commissione, COMP/M.2876 – *Newscorp/Telepiù*, 2 aprile 2003.

[10] AGCM, Provvedimento n. 30100, *C12207B - Sky Italia/r2 - revisione misure*, 12 aprile 2022.

conservativa.
La prima è la contraddittorietà tra l'affermare l'esistenza di due mercati distinti, uno a pagamento, finanziato direttamente dal consumatore, e l'altro finanziato dalla pubblicità, e poi non inserire tutta la pubblicità televisiva nello stesso mercato, come invece avviene ad esempio per la pubblicità online.
Per giustificare ciò si sostiene che la pubblicità dei canali a pagamento vada inserita nel mercato della tv a pagamento, risultandone una componente secondaria dei ricavi di quest'ultima. In questo modo però, lato domanda, ciò si potrebbe giustificare solo se l'investitore/inserzionista pubblicitario operasse a sua volta una distinzione tra le due modalità di offerta nella scelta dei programmi su cui investire.
Al contrario, la motivazione, salvo specifici prodotti che però costituiscono una quota marginale delle risorse pubblicitarie anche dell'emittente a pagamento, è in entrambi i casi la stessa: massimizzare l'audience relativamente al target sociodemografico di riferimento che si intende raggiungere, sia esso utente della tv in chiaro, sia di quella a pagamento. In altre parole, lo spot della Coca Cola o di TIM è lo stesso e tutte le emittenti, in chiaro o a pagamento, competono nella decisione dell'azienda di come e quanto investire in quel canale piuttosto che nell'altro, indipendentemente dal loro modello di business.
La seconda riguarda il tema degli *switching cost*. Se nel mondo analogico il passaggio da un'offerta in chiaro a una a pagamento imponeva di utilizzare una diversa rete distributiva (terrestre per la *free* tv e satellitare per la *pay* TV), con tutti gli ostacoli e le difficoltà, anche in termini di dotazioni tecnologiche, che questo comportava, il passaggio alla tecnologia digitale (a cominciare dal terrestre e poi anche dell'IP) ha ridotto i relativi costi, aumentando il numero dei canali in chiaro e a pagamento e anche il loro passaggio da un modello all'altro, ampliando in questo modo il livello della competizione e modificando, come meglio vedremo di seguito, lo stesso concetto di prezzo correlato alla distinzione tra i due modelli.
Tutto ciò, a parere di chi scrive, avrebbe dovuto già da tempo condurre, alla luce delle trasformazioni appena ricordate, a una diversa interpretazione in chiave evolutiva delle relazioni strategiche orizzontali tra i due modelli di business non solo in termini di pressioni competitive da tutti riconosciute, quanto di vera e propria sostituibilità.
In ogni caso tale esito non è più rinviabile, in relazione alle trasformazioni analizzate nel capitolo precedente e alla necessità di utilizzare una strumentazione più adeguata relativamente a quanto sopra ricordato.

4.2. Il mercato rilevante della televisione nell'ecosistema digitale

La pandemia, come analizzato in precedenza, ha costituito un punto di non ritorno nella trasformazione del sistema televisivo in Italia. Attualmente, seguendo le tendenze prevalenti nel resto d'Europa, si assiste, come abbiamo visto, allo sviluppo di servizi in chiaro (AVOD) e di servizi "ibridi", che si affiancano a quelli "tradizionali" in abbonamento (SVOD).

In questo modo, indipendentemente dal modello di business, si assiste a una crescente competizione per la conquista del pubblico non solo tra l'affollata e variegata schiera di streamer, e tra questi e i tradizionali broadcaster, ma anche con altre forme di intrattenimento, quali videogiochi e social media. Il tutto nel più ampio contesto dell'ecosistema digitale.

La presenza di forti effetti di rete in tale ecosistema (internet) si unisce in questo caso alla particolare natura delle piattaforme che offrono tali servizi come mercati multi-versante.

Il fatto che due gruppi siano attratti l'uno dall'altro non è un fenomeno recente e nella teoria economica è conosciuto come effetto di rete[11]. Tuttavia, nei mercati a più versanti questi effetti esterni di rete sono internalizzati da un intermediario – la piattaforma – che unisce due o più gruppi di consumatori e le relazioni che instaura con ciascuno di loro influenza anche il comportamento degli altri gruppi[12] .

In tal senso, il ruolo della piattaforma è essenziale poiché al gruppo di utenti che genera il più alto livello di effetti di rete sarà applicato un prezzo relativamente più basso, un prezzo inferiore al costo marginale, o addirittura sotto lo zero (versante sussidiato). Il contrario accade agli utenti dell'altro versante (versante *profit-making*). In questo modo, la piattaforma provvede a rendere possibile la relazione tra i gruppi, i cui costi di transazione o i cui benefici economici non sarebbero altrimenti sostenibili o realizzabili[13].

Questa struttura di mercato determina dunque in teoria benefici per tutti i soggetti interessati, poiché la piattaforma, internalizzando le esternalità positive tra i diversi gruppi di utenti (inserzionista, editore e utente) determina utilità per questi soggetti che non avrebbero massimizzato se la piattaforma non li avesse messi in contatto, non essendo in relazione diretta tra loro, ovvero relazionandosi in maniera meno efficace senza di essa.

È parimenti importante sottolineare che in questo nuovo contesto i costi di switching sono praticamente nulli e la scelta del modello di business risulta "indifferente" ai contenuti proposti.

Infatti, l'ibridizzazione delle offerte, dei formati e dei contenuti fa venir meno una delle principali motivazioni che hanno giustificato fin qui, agli occhi delle autorità antitrust, la distinzione in due mercati distinti, legati alla presenza di contenuti premium, in esclusiva (film e calcio), destinati unicamente alle *pay TV*.

---

[11] L'esempio classico è quello della telefonia, dove maggiore è il numero di utenti di una rete, maggiore è l'utilità che ne ricava ogni singolo utente, rappresentando ciò un incentivo anche nei confronti di coloro che non ne fanno ancora parte. In questo caso si parla di effetti positivi, ma ci sono anche quelli negativi, come nel caso delle autostrade, dove maggiore è il numero degli automobilisti tanto più probabile è il rischio di congestione del traffico. Per ulteriori approfondimenti, si veda: N. Economides, *The Economics of Networks*, in *International Journal of Industrial Organization*, 14, 1996, 6.

[12] La prima analisi dei mercati multiversante risale ad uno studio di William Baxter, capo della divisione antitrust del Department of Justice, sul mercato delle carte di pagamento. Il concetto di mercato a più versanti è stato poi ripreso in un successivo studio di J. C. Rochet - J. Tirole, *Platform Competition in Two-Sided Markets*, in *Journal of the European Economic Association,* 4, 2003, 990 ss. Tale documento rivelò che molti business in diversi settori potevano essere analizzati attraverso l'approccio del mercato a più versanti: a titolo esemplificativo basti pensare ai sistemi operativi dei computer, alle console per video games, ai giornali, ai centri commerciali, alle carte di credito, ai locali di incontri, ecc.

[13] C. Shapiro - H.R. Varian, *Information rules: A strategic guide to the network economy,* Leeds, 1999.

Più in generale ne discende che oggi più che mai, non è tanto la disponibilità a spendere e dunque il surplus del consumatore a determinate le tipologie d'offerta e i comportamenti dei consumatori, quanto la crescente competizione tra operatori sulla cattura del tempo del consumatore e sulla sua fidelizzazione, in un contesto sempre più ampio e convergente, quale quello rappresentato dal mercato dell'attenzione[14].
Nel definire il mercato rilevante, l'analisi *antitrust* dovrà quindi adeguare la sua strumentazione a questo cambiamento di paradigma, analizzando e valutando caso per caso le relazioni strategiche orizzontali tra le diverse componenti dell'ecosistema, a seconda dello stato di evoluzione del mercato, risolvendolo, all'interno di ciascuna componente, in termini di pressione competitiva o di vera e propria sostituibilità

4.3. Il potere di mercato

Tuttavia, nonostante le evidenze qui mostrate, come abbiamo ricordato, ancora nella sua ultima decisione dell'aprile dello scorso anno, l'AGCM ha confermato la distinzione tra i due mercati.
Di conseguenza ciò ha condotto ad analizzare la fase immediatamente successiva che è quella relativa all'esistenza del potere di mercato. In questo senso, prima di entrare nel merito della decisione, vanno sottolineate alcune contraddizioni che emergono dai dati di mercato utilizzati dell'Autorità.
Mettendo a confronto i dati pubblicati da AGCM e provenienti da una fonte terza (Statista) con gli stessi dati di autorevoli società e istituti di ricerca, si evince che il valore di mercato utilizzato dall'Antitrust italiana è meno attendibile.
Spesa dei consumatori in Italia per i servizi televisivi a pagamento on-demand (€ mln)

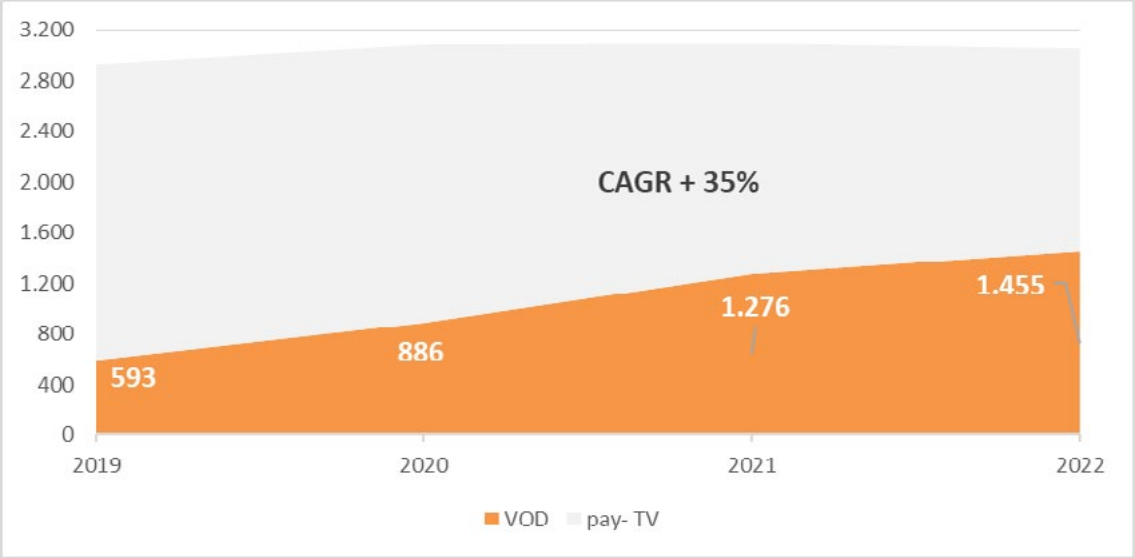
	AGCM (su Statista)	Media 3 Istituti	PWC	Mediobanca	ITMedia Consulting
2013	48,80				
2014	48,00				
2015	50,20	74,5	77		72
2016	52,70	128,5	116		141
2017	67,20	261	331		191
2018	78,00	349	413		285
2019	87,90	533	534	471	593
2020	93,70	805	772	756	886
2021		1.172	1.245	995	1.276

Fonte: elaborazione dell'Autore su varie

[14] «Tutto ciò richiede però strumenti sempre più sofisticati, servono numeri e dati per poterlo misurare e capire meglio, inseguendo e abbracciando le nuove sfide che l'ecosistema digitale pone. Un modello che abbia davvero al centro il consumatore, *super partes*, univoco e in tempo reale, che consenta di superare le incertezze e i dubbi delle attuali rilevazioni e che abbia come obiettivo finale di migliorare l'esperienza di fruizione per evitare qualsiasi fenomeno di rifiuto e di ostilità». Per maggiori informazioni vedasi F. Setti, *Misura per Misura, Medium per Medium*, in *group*.it, 22 gennaio 2022.

Oltre a ciò, la tabella sopra riportata rafforza anche quanto in precedenza sottolineato, sulla rilevanza crescente dello streaming nel settore della *pay TV*.

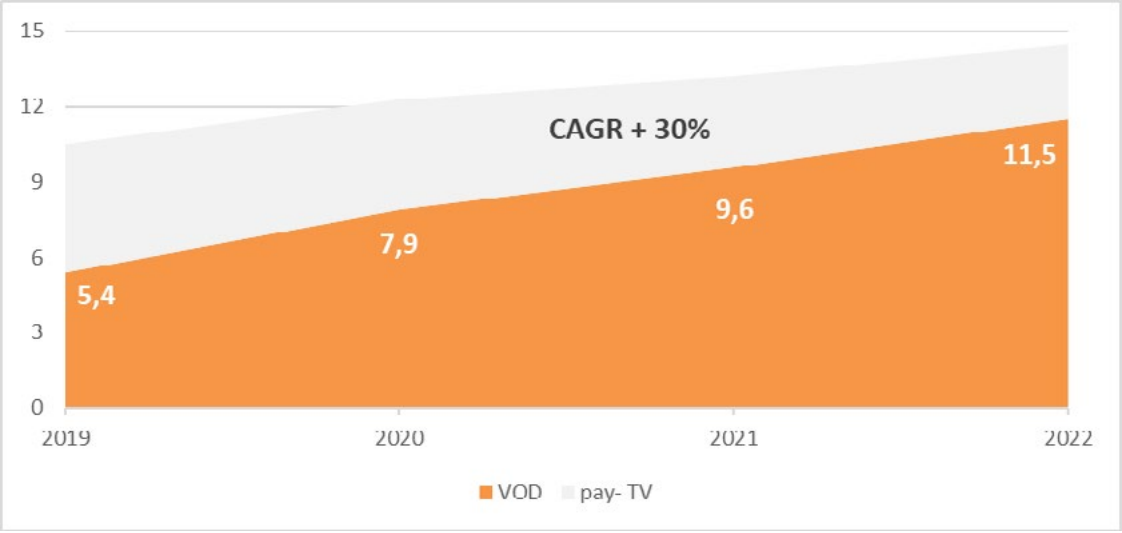
Ricavi della Pay-TV (€ mln)



Fonte: ITMedia Consulting, 2022

Questo vale anche in termini di quote di mercato, laddove il VOD recupera in pochi anni il ritardo nei confronti del *broadcast*, con una crescita media annua del 35% e si appresta entro la fine 2023 a colmare il *gap*, sorpassando la *broadcast pay TV*, che nel frattempo continua a diminuire in valore.

Abbonati alla Pay-TV (Mln abitazioni TV)



Fonte: ITMedia Consulting, 2022

Effetti ancora più evidenti riguardano il numero di abbonati, laddove il VOD domina abbondantemente, con oltre l'80% dei sottoscrittori totali (grafico sotto)[15].
Peraltro, le dimensioni di questi fenomeni (dinamiche delle quote di mercato degli operatori e dei sottoscrittori dei relativi servizi, peraltro rivalutati nella Sezione VII del provvedimento in maniera maggiormente coerente rispetto ai nostri dati di mercato), hanno indotto l'AGCM a revocare anticipatamente le misure imposte nel 2019.
In tal senso, quanto alla posizione detenuta da Sky, l'Autorità ha osservato che: «dal 2019 al novembre 2021 si è registrata una decisiva riduzione della quota di mercato, la quale era superiore all'80% nel 2019 ed è scesa al di sotto del 55% nel periodo luglio – novembre 2021 (successivamente all'avvio del nuovo Campionato di Serie A). Quanto all'indice di concentrazione, l'indice HHI nel 2019 è pari a oltre 8.900 punti e si è ridotto a meno di 4.000 punti nel periodo luglio – novembre 2021. Analogo andamento si è osservato per l'indice di concentrazione CR2»[16].
Nonostante, dunque, la scelta di AGCM di operare in continuità con la decisione del 2019 sugli aspetti preliminari del procedimento (definizione dei mercati, etc.) si è infine giunti ad una revoca delle misure.
La riduzione del livello di concentrazione e la tendenza duratura a un minor peso dell'operatore storico sono, infatti, chiari indicatori di una riduzione delle barriere all'ingresso nel mondo dell'online e di un forte livello di concorrenza nel settore, tale da ampliare in maniera significativa le possibilità di scelta del consumatore.
Questo aspetto è particolarmente rilevante nell'applicazione del significativo potere di mercato a tutela del pluralismo, perché la sua applicazione in passato, a differenza di quanto avveniva in ambito antitrust, era indipendente dal suo abuso, ma scattava automaticamente al raggiungimento di determinate soglie di mercato.
Anche qualora criteri che si richiamano maggiormente al diritto della concorrenza vengano adottati dall'AGCOM come indicatori del significativo potere di mercato, una volta che la medesima soglia quantitativa venga superata, è evidente come indicatori relativi alla presenza o meno di barriere all'ingresso, insieme ad altri indicatori "concorrenziali" diventano essenziali per definire concretamente il significativo potere di mercato.
Quello che qui si vuol far notare, tuttavia, è come nella *pay TV*, se da un lato esiste un operatore storico che potrebbe superare i limiti del significativo potere di mercato (> 50% del totale risorse), chi al contrario pone in essere i comportamenti tipicamente e potenzialmente anticoncorrenziali sono proprio quei soggetti che in questo "mercato" hanno una quota di ricavi inferiore, come ad esempio Amazon, in termini di "sinergie derivanti dalle attività svolte in mercati differenti ma contigui", oppure la stessa Amazon insieme ad altri (Netflix, Apple) relativamente al tema della "disponibilità e controllo dei dati" e di "efficienza economica ed effetti di rete".
Questo dovrebbe certamente condurre a esiti diversi, ma che qui non vengono approfonditi, perché da un lato sarebbero in contraddizione con l'assunto sviluppato al

[15] Va notato infatti come la differenza tra la quota di abbonamenti e la quota di ricavi è determinata dal diverso ARPU, cioè dal costo del singolo abbonamento che nel caso della tradizionale pay TV è ben più alto.

[16] AGCM, Provvedimento n. 30100, *C12207B - Sky Italia/r2 - revisione misure*, 12 aprile 2022, 15.

punto precedente, dell'esistenza di un unico mercato all'interno dell'ecosistema digitale, e dall'altro perché, come approfondiamo di seguito, siamo fortemente convinti che la strumentazione adottata per la tutela della concorrenza, solo perché già consolidata e quindi di più facile applicazione, non sia trasferibile e non possa più essere utilizzata secondo le logiche tradizionali per le verifiche di cui all'art. 51 del TUSMA – stante il mutato quadro di riferimento - nell'ottica della tutela del pluralismo.
Infatti, noi riteniamo che l'analisi *antitrust* non sia in grado di cogliere le profonde differenze che esistono tra i due ambiti, e che richiedono, come previsto anche nella normativa, in ambiente digitale per la tutela del pluralismo, una strumentazione ad hoc, che sia in grado di intervenire e regolare fenomeni che hanno profonde ripercussioni, non tanto e non solo sul mercato, quanto sulla vita delle persone e sulla formazione delle idee, in ultima analisi sulla stessa democrazia.
Come ciò potrebbe essere realizzato, viene approfondito nel capitolo che segue.

5. Il significativo potere di mercato e la tutela del pluralismo

L'art. 51 del TUSMA, in particolare al c. 5, modifica dunque il precedente meccanismo del SIC legato al solo superamento della soglia antitrust, cioè di un sistema basato su ricavi e quote di mercato per identificare le posizioni dominanti in materia di pluralismo, sostituendolo con quello attuale, che collega a tale sistema solo una presunzione di dominanza, introducendo e integrandolo con ulteriori criteri (indici) e lasciando all'AGCOM il compito di valutarne la rilevanza, al fine di accertare la presenza di posizioni di significativo potere di mercato lesive del pluralismo.
Come già ricordato, anche in virtù delle sollecitazioni provenienti dalla Corte di Giustizia europea[17] il legislatore riconosce che pluralismo e concorrenza siano principi non coincidenti, suscettibili di diversa tutela.
Gli indicatori che si richiamano alla tradizionale strumentazione antitrust di analisi dei mercati e definizione delle posizioni dominanti volte a tutelare la concorrenza, devono essere quindi interpretati ed applicati in senso evolutivo, avendo in mente la diversa natura di bene tutelato rappresentato dal pluralismo e non dalla concorrenza[18].
Questo vale anche rispetto alla definizione dei mercati precedenti, a partire da quella relativa al mercato della TV in chiaro, dove confluirebbero sia le risorse da pubblicità che quelle da canone, giustificata quest'ultima in passato, diversamente da quanto avviene nei mercati rilevanti antitrust, proprio in ragione del pluralismo.

[17] Alcune di queste sono disponibili sul sito *curia.europa.eu.*

[18] L'Autorità deve tener conto dei seguenti "indicatori": «oltre che dei 1) ricavi, 2) del livello di concorrenza statica e dinamica all'interno del sistema, 3) delle barriere all'ingresso nello stesso, 4) della convergenza fra i settori e mercati, 5) delle sinergie derivanti dalle attività svolte in mercati differenti ma contigui, 6) della integrazione verticale e conglomerale delle società, 7) della disponibilità e del controllo di dati, 8) del controllo diretto o indiretto di risorse scarse necessarie, quali le frequenze trasmissive, 9) delle dimensioni di efficienza economica dell'impresa, anche in termini di economie di scala, gamma e rete, nonché 10) degli indici quantitativi di diffusione dei programmi radiotelevisivi, anche con riferimento ai programmi di informazione, delle opere cinematografiche, dei prodotti e servizi editoriali e online». (numerazione aggiunta).

Poiché trattasi di risorsa del tutto estranea al mercato, non contendibile e come tale anche distorsiva del corretto andamento del mercato, il canone va a finanziare proprio quei servizi pubblici che tutelano il cosiddetto pluralismo interno e che proprio per questo godono di un'eccezione riconosciuta a livello europeo dal trattato di Amsterdam[19].

Il SIC in tal senso opera una evidente forzatura che non ha tanto lo scopo di porre il servizio pubblico sotto la lente dello scrutinio ai fini della identificazione del significativo potere di mercato, dal momento che un servizio pubblico in ragione della sua missione non può che tutelare il pluralismo, che ne costituisce il suo elemento costitutivo e la sua ragion d'essere, quanto quello di ampliare il terreno di gioco impedendo a chi eventualmente detiene una posizione dominante sull'unico mercato in cui opera la TV in chiaro, che è quello finanziato dalla pubblicità televisiva, di poter essere riconosciuto come tale, con le conseguenze che questo comporterebbe. È la stessa ragione in fondo per cui, come abbiamo visto in precedenza, la stessa pubblicità televisiva viene ripartita tra *pay* e *free*, laddove, come abbiamo cercato di dimostrare, tale distinzione non ha alcuna ragion d'essere in termini di struttura dei mercati e di comportamenti dal lato della domanda (l'inserzionista)[20].

5.1. Informazione e pluralismo nell'ecosistema digitale

Tutto ciò premesso, appare in ogni caso opportuno preliminarmente affrontare il tema del pluralismo nel più ampio contesto del diritto all'informazione, tutelato dall'art. 21 della Costituzione.

Informazione e pluralismo sono infatti due concetti strettamente correlati e inscindibili. L'informazione è un bene pubblico, fondamentale, non escludibile e non rivale[21], e rappresenta lo strumento a disposizione della collettività per permettere ai cittadini di operare le proprie scelte e consentire il realizzarsi nella società delle dinamiche de-

[19] Commissione europea, *Comunicazione relativa all'applicazione delle norme sugli aiuti di Stato al servizio pubblico di emittenza radiotelevisiva*, 2009/C 257/01.

[20] Anche l'ipotesi di prevedere una dominanza congiunta appare scarsamente praticabile, sia per le perplessità e le difficoltà di dimostrazione di questo tipo di intese nel diritto antitrust, sia per le ragioni esposte sopra, relativamente all'eccezione relativa al servizio pubblico radiotelevisivo.

[21] Il bene pubblico presenta due caratteristiche che lo distinguono dal bene privato: la non rivalità (o indivisibilità) e la non escludibilità. Un bene non rivale è un bene il cui consumo da parte di una persona non diminuisce la quantità disponibile per gli altri. L'esempio classico è quello delle candele: quando una seconda candela si accende utilizzando la fiamma della prima, non c'è nella prima alcuna perdita di luce o di calore. Rientrano in questa categoria la difesa nazionale, le strade, l'informazione. Considerando la nullità dei costi marginali, il bene non rivale o indivisibile garantisce il massimo benessere sociale quando il prezzo è pari a zero. Se così non avviene, coloro che non sono disposti a pagare sarebbero esclusi, laddove il loro consumo non determinerebbe alcun costo. La non escludibilità comporta invece l'impossibilità di escludere qualcuno dal consumo del bene, anche se non paga alcun prezzo per lo stesso. Ciò deriva da considerazioni tecniche o etiche, a differenza dell'indivisibilità che è una proprietà specifica del bene. Va peraltro sottolineato come la loro collocazione non sia rigida, dal momento che alcuni di questi beni hanno perduto (o ridotto) nel tempo tali caratteristiche. Ad esempio, le strade sono un bene non rivale, ma possono essere escludibile quando venga richiesto un pedaggio per entrarvi. Per ulteriori approfondimenti: K. J. Arrow, *Economic Welfare and the Allocation of Resources for Innovation*, in R. R. Nelson (a cura di), *The rate and direction of inventive activity*, Princeton, 1962.

mocratiche, tutelate appunto dall'art. 21 della Costituzione. In tal senso, il pluralismo (dell'informazione) è una condizione essenziale affinché il processo democratico possa realizzarsi e rafforzarsi nel tempo[22].

Ciò comporta dunque prioritariamente che anche nella definizione del "mercato rilevante del prodotto", occorra operare una chiara distinzione tra concorrenza e pluralismo, poiché non tutti i contenuti che hanno valore nel mercato rilevante della prima, lo hanno nel secondo.

Questo evidentemente perché i media offrono non solo contenuti di svago e d'intrattenimento, ma contenuti importanti per il pubblico o che hanno benefici pubblici più ampi, come l'informazione e il giornalismo. Ne discende che sebbene a priori non si possa escludere che anche altri generi e contenuti possano contribuire al ruolo di "formazione dell'opinione pubblica" che costituisce il fine ultimo della tutela del pluralismo, si può altresì affermare che solo l'informazione ne costituisca parte integrante. Richiamando infatti la stessa ripartizione operata dalla Commissione europea e sopra ricordata[23], appare difficile immaginare che tutto ciò che è puro intrattenimento (sport, eventi musicali, ecc.) possa assolvere a tale funzione, come pure considerare altri generi (film, serie, ecc.) come contenuti dotati della stessa dignità e rilevanza dell'informazione[24].

Nell'ecosistema digitale questa distinzione è se possibile ancora più chiara, perché, per le dinamiche in precedenza ricordate e per la natura stessa di bene pubblico oggetto dell'attività d'impresa (editoriale), l'accesso tende a essere universale e i ricavi si orientano vieppiù verso un determinato gruppo di utenti: gli inserzionisti.

Tutto questo comporta alcune conseguenze:

1. L'analisi del mercato non può essere settoriale, ma richiede di identificare le posizioni di significativo potere di mercato in una dimensione di sistema (o ancor meglio di ecosistema) dell'informazione.
2. La valutazione non può essere fatta in termini di semplice accesso alle fonti di informazione, e dunque sui ricavi, ma in concreto in ragione delle dinamiche della domanda e dei consumi reali d'informazione, e dunque sulle audience.
3. Gli indicatori individuati debbono essere coerenti con gli obiettivi di efficienza economica, ma anche con altre preoccupazioni, non solo concorrenziali, in materia di tutela del pluralismo

[22] G. Pitruzzella, *La libertà di informazione nell'era di Internet*, in questa *Rivista*, 1, 2018, 19 ss.

[23] Commissione europea, *The European Media Industry Report,* cit.

[24] Per i contenuti diversi dall'informazione appare controversa, e a nostro avviso impraticabile, l'idea di definire e misurare indici di pluralismo collegati ad altri generi, come ad esempio le opere cinematografiche, così come ipotizzato all'art. 51, c. 5, nel nuovo TUSMA. Questi contenuti, oltre ad avere una finalità d'intrattenimento, avrebbe l'ulteriore problema di richiedere la definizione di parametri qualitativi attualmente inesistenti, dal momento che richiederebbero un'analisi del contenuto per ogni singola opera, a differenza degli indici quantitativi (generi, ore di programmazione, indici di ascolto) che in chiave di pluralismo non possono certo applicarsi indifferentemente a tutti i contenuti filmici offerti dall'emittente/fornitore del servizio.

5.2. Il pluralismo e la sua misurazione

Per le ragioni ricordate in precedenza, proprio per la particolare rilevanza dei contenuti in oggetto, la necessità di definire il mercato e di valutarne il significativo potere sotto il profilo del pluralismo non può essere considerato, a nostro avviso, dal punto di vista anti-concentrativo, delle quote di mercato sul sistema integrato delle comunicazioni (elemento che ne costituisce peraltro il presupposto), quanto piuttosto concentrarsi nel valutare se il superamento di tali limiti possa incidere nel diritto dei cittadini ad essere informati, nella protezione della dignità delle persone, in un quadro peraltro di pieno rispetto della libertà d'espressione e della libertà editoriale.

Da questo punto di vista alcuni fenomeni meritano di essere sottolineati.

Il primo è che i consumatori hanno potuto beneficiare, grazie al nuovo ecosistema digitale, come approfondito di seguito, dell'accesso a una varietà sempre più ampia di fonti d'informazione. La televisione, sebbene perda terreno soprattutto nei confronti di internet, continua ad avere un ruolo rilevante nel sistema dell'informazione, come ricordato nell'ultima relazione annuale dell'AGCOM:

> «Lo *share* delle principali edizioni dei telegiornali si mantiene su valori non distanti da quelli del 2020, registrando in diversi casi un incremento rispetto agli anni antecedenti alla pandemia. Rimane consistente la quota (48% nel complesso) degli spettatori dei telegiornali delle 20, Rimane consistente la quota (48% nel complesso) degli spettatori dei telegiornali delle 20 (Tg1, Tg5 e Tg La 7), con il Tg1 che conferma il proprio primato (5,1 milioni di fruitori in media nell'edizione serale), seguito dal Tg5 (circa 4 milioni di spettatori)»[25].

Si tratta di dati allineati per molti versi alle altre esperienze europee (vedi Regno Unito sotto), in base alle quali ancora oggi la televisione, pur se in diminuzione rispetto al passato, continua ad essere mediamente la principale fonte di informazione e di accesso alle notizie[26].

D'altro canto, però Internet continua a crescere, avvicinando sempre più la televisione e superandola nettamente in alcune fasce della popolazione. In questo contesto, al suo interno, i social media conquistano sempre maggiori popolarità, rappresentando il modo più comune di accedere all'informazione per i giovani e in generale la popolazione fino ai 34 anni.

[25] AGCOM, *Relazione annuale 2022 sull'attività svolta e sui programmi di lavoro,* Luglio 2022, 45.

[26] Ofcom, *The future of media plurality in the UK,* in o*fcom.org.uk,* 17 novembre 2021, 46. Si veda anche Ofcom, *News Consumption in the UK 2022*, in *ofcom.org.uk,* luglio 2022.

Use of main platforms for news nowadays 2021 – by age

Platform	Total	16-24	25-34	35-44	45-54	55-64	65-74	75+
Television	79%	61%	67%	77%	82%	91%	92%	95%
Internet*	73%	89%	85%	80%	72%	66%	54%	43%
Social media	49%	79%	67%	58%	44%	33%	19%	16%
Radio	46%	29%	40%	49%	51%	57%	51%	49%
Newspapers (print)	32%	16%	22%	28%	34%	37%	49%	53%

**Internet figures include use of social media, podcasts and all other websites/apps accessed via any device.*

Fonte: Ofcom, *The future of media plurality in the UK 2021*

Il secondo è che, per operare in maniera efficace, il sistema di misurazione debba evolversi, utilizzando strumenti di rilevazione più adeguati al mutato contesto competitivo.

In tal senso, anche alla luce di quanto emerge a livello internazionale, il vero criterio distintivo per misurare il pluralismo dovrebbe essere costituito proprio da un sistema non più incentrato sui ricavi ma sugli ascolti.

In questa chiave, la costituzione di posizioni di significativo potere di mercato lesive del pluralismo dovrebbe essere valutata in funzione delle quote di ascolto dei programmi e dei contenuti d'informazione da parte di ciascun operatore.

Se dunque tuttora la televisione rimane il riferimento per l'informazione nella numerosità della platea che vi accede per finalità informative[27], il significativo potere di mercato andrà misurato sull'indice quantitativo di diffusione dei programmi radiotelevisivi più convenzionale, lo share medio (Auditel nel nostro caso), misurato sul totale dei programmi d'informazione (giorno e prime time).

Ma è chiaro altresì che questo sistema dovrà evolvere in fretta, come già sta facendo, verso nuove metriche e nuovi sistemi di rilevazione, come quelli basati sulla *total audience*, gli unici in grado di cogliere le trasformazioni dei consumi nel mutato contesto digitale.

I sistemi tradizionali, come Auditel, infatti, sono ancora organizzati verticalmente, per mezzo, in alcuni casi per *device*, producendo metriche tra loro non comparabili che tendono alla frammentazione.

Nel mutato, convergente scenario, abbiamo visto invece come sia necessario operare in una logica cross-mediale, ancora più necessaria in un settore come quello dell'informazione, caratterizzato dalla pluralità delle modalità d'accesso.

A livello europeo, la forma prevalente dei soggetti che realizzano le indagini sulle rilevazioni degli ascolti al fine di garantire una effettiva rappresentatività dell'intero settore di riferimento è quella del JIC – *Joint Industry Committee* – e, seppure con livelli di avanzamento differenti tra loro, gli sviluppi dei differenti sistemi di rilevazione sembrano orientarsi verso un'analisi integrata di tutti i media[28].

[27] Ofcom, ivi, 45.

[28] Oltre allo storico modello di Médiamétrie in Francia, sistemi di misurazione convergenti sono rinvenibili nell'esperienza olandese (NMO) - che ha avviato un sistema *cross* piattaforma e *cross device* per la misurazione di tutti i mezzi, superando così i precedenti JIC – e anche in Canada – dove è stato replicato un modello analogo prevedendo l'utilizzazione anche dei dati di profilazione in forma aggregata delle

Così, anche in Italia, si è posta la necessità di effettuare analisi cross-mediali che diano conto del consumo dei contenuti fruiti sui mezzi tradizionali anche sui nuovi supporti connessi alla rete internet, nell'ottica di una graduale convergenza delle metriche in logica *cross-piattaforma*, *cross-device* e *cross-mediale* di tipo *consumer-centric*, in grado di fornire informazioni circa il coinvolgimento del consumatore rispetto ai contenuti e alla pubblicità diffusa sulle diverse piattaforme di comunicazione[29].

Anche Ofcom, occupandosi di pluralismo, ha riscontrato come i consumatori abbiano potuto beneficiare dell'accesso a una varietà di nuove fonti, in particolare nell'informazione[30].

L'esigenza di migliorare il quadro regolamentare impone, secondo l'Ofcom, di garantire maggiore trasparenza e accesso ai dati per dare priorità alle notizie; algoritmi di raccomandazione che promuovono i valori pubblici attraverso la diversità di esposizione, piuttosto che servire semplicemente contenuti per ottimizzare i "click" (*media literacy by design*); nuovi "meccanismi" che consentirebbero all'Ofcom di esaminare potenziali problemi di pluralità dei media al di fuori del contesto di una specifica fusione, come pure nuovi approcci volti a misurare il consumo di notizie e comprendere meglio i comportamenti e le abitudini dei consumatori nei confronti dei contenuti giornalistici.

Parallelamente, l'AGCOM ha auspicato la revisione delle infrastrutture di misurazione e delle metodologie adottate dai JIC, al fine di giungere ad un progressivo processo di coordinamento e convergenza degli attuali sistemi di rilevazione, che conduca a metriche univoche, all'integrabilità delle tecnologie e delle metodologie di rilevazione e alla condivisione degli asset di misurazione, in una logica di sistema e nella prospettiva del mercato, rispondendo a principi di equità, parità di trattamento e non discriminazione verso tutti i soggetti coinvolti, indipendentemente dal mezzo o dalla piattaforma di cui viene rilevato il consumo[31].

Ribadendo che devono essere compiuti tutti gli sforzi necessari per individuare metriche univoche tali da rendere comparabili i dati di consumo riferiti ai vari mezzi, in un'ottica *consumer-centric*, l'Autorità auspica di disporre di un dato complessivo certificato di "*total audience*" che consenta di misurare, su tutte le piattaforme e tutti i device, in maniera univoca e al netto delle duplicazioni, la fruizione dei contenuti e che presenti

Telco operanti sul territorio canadese, al fine di accrescere ulteriormente l'accuratezza delle misurazioni, in ragione della preminente funzione pubblica e di sistema che il governo ha riconosciuto al JIC. Nella stessa direzione si muove il JIC svizzero Mediapulse.

[29] In questa direzione rileva l'iniziativa assunta dalla federazione mondiale degli utenti della pubblicità (World Federation Advertisers – WFA - cui aderisce anche UPA in Italia) con il manifesto denominato WFA C*ross Media Initiative* con il quale i principali investitori su scala mondiale hanno chiesto l'unificazione degli attuali sistemi di rilevazione indicando i requisiti metodologici e tecnici che tale unificazione dovrebbe prevedere. Il manifesto è basato su principi condivisi e postula la flessibilità nell'approccio al fine di potersi adeguare alle specificità locali, ma individua quattro aree principali sulle quali dovrebbe innestarsi l'invocato percorso: *governance*, standard e metriche, privacy e sicurezza del processo, e un'infrastruttura tecnologica (*Pipework*) in grado di consentire la deduplicazione dei dati e, così come per la metodologia, sia calibrata e convalidata mediante l'uso di un *panel single source* indipendente. Un approccio ibrido (panel più censuario) è indicato come il modo migliore per acquisire i dati del pubblico in un ecosistema digitale così frammentato.

[30] Ofcom, *The future of media plurality in the UK,* cit., 9.

[31] AGCOM, Delibera n. 194/21/CONS, "Indirizzi in materia di sistemi di rilevazione degli indici di ascolto nel nuovo ecosistema digitale", giugno 2021.

le caratteristiche di affidabilità e granularità necessarie per il mercato e che risponda ai principi contenuti nell'atto di indirizzo[32].

Infine, nella recente proposta di regolamento dell'European Media Freedom Act (EMFA)[33], incentrata sulle tematiche dell'informazione, si ribadisce come la misurazione dell'audience – in particolare nel settore audiovisivo – abbia un impatto diretto sull'allocazione e sui prezzi della pubblicità, che rappresenta una fonte di reddito fondamentale per il settore poiché consente di valutare le prestazioni dei contenuti mediatici e comprendere le preferenze del pubblico per pianificare la futura produzione di contenuti.

La proposta evidenzia come alcuni nuovi operatori online emersi nell'ecosistema dei media forniscono i propri servizi di auto misurazione – che si sviluppano come "valute alternative" in concorrenza con quelle concordate dal mercato e che, in alcuni casi, portano a risultati di misurazione effettivi diversi nella pratica[34] – senza rendere disponibili informazioni sulle loro metodologie, facendo leva sulla loro integrazione verticale e sulla posizione di significativo potere di mercato di cui godono nel settore della pubblicità online.

L'opacità dei diversi sistemi di misurazione dell'audience è causa di asimmetrie informative e potenziali distorsioni tra gli operatori del mercato dei media, poiché incide negativamente sulle aziende del settore e svantaggia i concorrenti che forniscono servizi di misurazione dell'audience che rispettano gli standard concordati.

Dal momento che alcuni Stati membri hanno limitazioni alla proprietà basate sulle quote di mercato legate alle audience, altri hanno limitazioni basate sulle quote di mercato basate sui ricavi, altri ancora sulle restrizioni al controllo dei capitali o alla proprietà *cross-mediale*, le norme antitrust non possono affrontare in modo strutturato e omogeneo la metodologia poco trasparente per la misurazione dell'audience online.

Tali norme, infatti, non affrontano direttamente l'impatto che le concentrazioni di mercato potrebbero avere sul pluralismo e sull'indipendenza dei media, e le norme sugli aiuti di Stato, che sono applicate caso per caso (spesso *ex post*), non affrontano sufficientemente i problemi creati dall'iniqua allocazione di risorse statali ai fornitori di servizi di media.

Gli operatori del mercato dei media, in particolare i fornitori di servizi e gli inserzionisti, dovrebbero poter contare su dati di audience oggettivi, derivanti da soluzioni di misurazione dell'audience trasparenti, imparziali e verificabili, volte a ridurre le distorsioni del mercato, rafforzando ulteriormente la parità di condizioni tra i fornitori di servizi di media e gli operatori *online* a beneficio in particolare dei servizi di media audiovisivi e della stampa online, nonché degli inserzionisti *online*.

---

[32] AGCOM, Delibera n. 262/22/CONS, "Avvio di una consultazione pubblica finalizzata alla predisposizione di una relazione sullo stato di implementazione dell'atto di indirizzo di cui alla delibera n. 194/21/CONS", luglio 2022.

[33] European Commission, "Proposal for a regulation of the European Parliament and of the Council establishing a common framework for media services in the internal market (European Media Freedom Act) and amending Directive 2010/13/EU", Brussels, 16.9.2022 COM (2022) 457 final.

[34] Si cita a titolo di esempio la società Dazn in Italia. L'AGCOM ha riscontrato che i dati di audience raccolti dal provider di streaming sulla base di un sistema di auto-misurazione era superiore del 50% rispetto ai dati di audience misurati da Auditel.

Pertanto, la Sezione 6 dell'EMFA stabilisce che i fornitori di sistemi di misurazione dell'audience proprietari dovranno fornire, senza indebiti ritardi e gratuitamente, ai fornitori di servizi di media e agli inserzionisti, nonché a terzi autorizzati dai fornitori di servizi media e dagli inserzionisti, informazioni accurate, dettagliate, complete, intelligibili e aggiornate sulla metodologia utilizzata dai loro sistemi di misurazione dell'audience.
È importante, pertanto, che l'AGCOM continui a svolgere un ruolo da protagonista in questo ambito, nel far sì che questi nuovi strumenti di misurazione dell'audience possano rapidamente svilupparsi, dando la possibilità di applicare al meglio le norme poste a tutela del pluralismo informativo, identificando nel nuovo ecosistema digitale le posizioni di significativo potere di mercato lesive del pluralismo.
Auspicando un rapido passaggio all'utilizzo di queste nuove metriche e dei sistemi di rilevazione cross-mediale di *total audience*, nel periodo transitorio converrà comunque continuare ad avvalersi di Auditel e, per la parte non coperta dei sistemi tradizionali, delle *survey*.

5.3. La tutela del pluralismo nell'ecosistema digitale

Per ultimo, come abbiamo sottolineato, il raggiungimento della maggiore efficienza economica, legato alla personalizzazione dei servizi, se da un lato riduce alcune preoccupazioni concorrenziali, dall'altro le accresce in tema di pluralismo.
L'informazione online è, caratterizzata infatti dall'abbondanza e dalla personalizzazione dei contenuti[35]. Nel contesto tradizionale, infatti, un giornale o una televisione offrono un pacchetto (*bundle*) di articoli o servizi (il giornale o il telegiornale appunto) che il consumatore acquista e/o consuma congiuntamente.
Internet, come ha fatto prima con l'industria musicale e successivamente con quella audiovisiva, procede invece spacchettando l'offerta informativa tradizionale e riaggregandola attraverso nuovi soggetti che operano direttamente nell'informazione (es. Drudge Report e Huffington Post), ovvero che offrono servizi web orizzontali (motori di ricerca, portali, *social network*).
Anche questi ultimi differiscono tra loro poiché alcuni prevedono una valutazione dei contenuti informativi prima della loro pubblicazione e altri, invece, no.
La diffusione delle nuove tecnologie non necessariamente però si traduce in effettiva pluralità informativa, qualora priva di un concreto pluralismo di idee e di opinioni.
Le strategie organizzate di disinformazione, che sono cosa diversa dalle banali bufale o fake news, sono una esternalità negativa che «inquina il mondo dell'informazione e che agisce da selezione tossica dei contenuti che riceviamo. La disinformazione aumenta il costo di cercare e ricevere informazioni corrette»[36].
Gli intermediari, in virtù degli effetti di rete e delle esternalità positive legati alla natura multi-versante della loro attività, più volte ricordati, assumono dunque un ruolo centrale nel nuovo sistema dell'informazione, sia nel favorire o combattere questa tendenza,

[35] Per approfondimenti, si rimanda a G. Pitruzzella, *La libertà di informazione,* cit.,19 ss.

[36] A. Nicita, *Il mercato delle verità. Come la disinformazione minaccia la democrazia,* Bologna, 2021.

sia nell'accrescere il valore delle risorse che affluiscono al sistema (pubblicità online), regolando l'accesso e il consumo d'informazione da parte degli individui/consumatori. La disponibilità di dati (personali e non) consente alle piattaforme, in quanto possessori di maggiori informazioni sugli utenti, di avere una posizione di vantaggio competitivo, ove l'uso del dato è di vitale importanza nell'offerta di prodotti/servizi ("uso primario del dato"), consentendo, altresì, di sviluppare e migliorare i prodotti/servizi offerti (c.d. "uso secondario del dato")[37].

Il valore elevato di queste informazioni per configurare specifici profili di abitudini di consumo spinge le piattaforme online a fare in modo di catturare quanta più attenzione possibile, anche attraverso la promozione e la proposizione di contenuti graditi all'utente, che viene monetizzata attraverso la vendita di appositi spazi per l'online advertising, raggiungibili dagli utenti tramite terminali fissi e mobili (con sistemi cd. di *programmatic* e *reservation advertising*).

Questa struttura di mercato determina dunque in teoria benefici economici per i soggetti interessati, poiché la piattaforma, internalizzando le esternalità positive tra i diversi gruppi di utenti (inserzionista, editore e lettore), determina utilità per questi soggetti che non avrebbero massimizzato se la piattaforma non li avesse messi in contatto, non essendo in relazione diretta tra loro, ovvero relazionandosi in maniera meno efficace senza di essa.

Al contempo però fa emergere dubbi e preoccupazioni proprio in relazione alla tutela del pluralismo.

Come ampiamente illustrato, le nuove tecnologie hanno profondamente cambiato il modo in cui i contenuti vengono prodotti, condivisi e diffusi. Uno dei cambiamenti più recenti (e preoccupanti) è il fenomeno delle *fake news*, soprattutto perché la disinformazione e il travisamento intenzionale delle informazioni reali hanno iniziato a influenzare il processo decisionale individuale nella sfera politica.

È un fenomeno molto preoccupante perché la diffusione di *fake news* può sfidare i valori democratici e minare la sicurezza nazionale.

È senz'altro vero che la nascita del web ha favorito inizialmente l'incremento del pluralismo poiché, accanto alle agenzie di stampa e agli editori tradizionali, che diffondono le notizie in forma strutturata con propri siti di informazione, in rete operano, come abbiamo visto (par. 2.2.3), numerosi altri soggetti che forniscono servizi assai differenziati tra loro.

Ciò nonostante, negli ultimi anni, sono emerse tendenze poco rassicuranti, la cui motivazione deriva da una più approfondita analisi della natura stessa dell'informazione, definito come "bene esperienza", ossia un bene il cui valore può essere giudicato solo attraverso il consumo e tale consumo non è facilmente determinabile a priori, in quanto la comprensione dei comportamenti di consumo è legato al funzionamento della mente umana in tutte le situazioni in cui l'individuo è chiamato ad assumere delle decisioni.[38]

Di conseguenza, la comunità scientifica ha posto grande attenzione allo studio dei

[37] AGCM – AGCOM - GPDP, *Indagine conoscitiva sui Big Data,* febbraio 2020.

[38] AGCOM, *Percezione o disinformazione: molto razionali o troppo pigri?,* Servizio economico statistico, Febbraio 2020.

processi di scelta degli individui, che sono divenuti un tema di crescente interesse e attualità da parte di diversi ambiti disciplinari: l'economia, la sociologia, la psicologia, la filosofia, le neuroscienze, tanto da assurgere a dignità di dominio scientifico specifico, che comunemente va sotto il nome di scienze cognitive.

Le preoccupazioni maggiori connesse al fenomeno della personalizzazione dei servizi risiedono proprio in relazione alla tutela del pluralismo, in contrasto alla disinformazione.

Le strategie organizzate di disinformazione, che sono cosa ben diversa dalle banali bufale o *fake news*, sono una esternalità negativa che inquina il mondo dell'informazione e che agisce da selezione tossica dei contenuti che riceviamo.

In questo contesto, la valorizzazione dell'informazione risulta uno snodo cruciale per comprendere il futuro di Internet e delle democrazie moderne, che passa dunque per la lotta alla disinformazione, obiettivo primario anche dell'azione dell'AGCOM. Naturalmente, in tale contesto, l'AGCOM è protagonista.

Tra le numerose attività, ha, tra l'altro, svolto indagini sull'evoluzione dell'offerta di informazione tramite le piattaforme online,[39] avviato tavoli tecnici,[40] promosso iniziative condivise con gli operatori stessi, come il costante monitoraggio e l'attività di vigilanza dei fenomeni di disinformazione,[41] la creazione di una piattaforma di coordinamento delle attività autonome di *fact-checking* e la definizione di standard giornalistici in materia di trasparenza, etica e qualità della struttura editoriale, come pure la predisposizione di linee guida volte a fornire ai cittadini strumenti per un uso consapevole e critico delle piattaforme.

Ancora, l'Autorità è stata chiamata a coordinare e a indirizzare l'attività di auto-regolazione nonché la predisposizione di codici di condotta e di ulteriori forme di co-regolazione con le piattaforme.[42], fino ad arrivare al recente regolamento in materia di *copyright*[43].

È dunque importante che i criteri e gli indicatori per misurare il pluralismo non entrino in conflitto con i risultati già raggiunti dall'AGCOM e anzi costituiscano una natura-

[39] A titolo di esempio, si ricorda l'indagine conoscitiva su "Piattaforme digitali e sistema dell'informazione", da cui sono scaturiti i Rapporti "News vs. fake nel sistema dell'informazione" e "Percezioni e disinformazione. Molto razionali o troppo pigri?". Per approfondimenti, si veda AGCOM, *Indagine conoscitiva su Piattaforme digitali e sistema dell'informazione*. Ancora, l'Autorità ha avviato forme di co-regolazione, ad esempio, nell'ambito della trasposizione della nuova direttiva europea sui servizi media audiovisivi rispetto alle piattaforme di condivisione di video. AGCOM, *AGCOM: approvato il regolamento per il contrasto alle espressioni d'odio (hatespeech)*, Comunicato stampa, 23 maggio 2019.

[40] Si rammenta il Tavolo Tecnico per la garanzia del pluralismo e la correttezza dell'informazione sulle piattaforme online, che ha portato alla redazione del primo Rapporto tecnico sulle "strategie di disinformazione online e la filiera dei contenuti fake". Per approfondimenti, si veda il sito *agcom.it*.

[41] Attraverso l'Osservatorio sulle piattaforme online, l'Osservatorio sulla disinformazione online e il Sistema integrato delle comunicazioni. Per approfondimenti, si veda *agcom.it*.

[42] A titolo esemplificativo, l'AGCOM svolge monitoraggio delle iniziative di auto-regolamentazione delle piattaforme online volte a contrastare la disinformazione sulle tematiche Covid-19. AGCOM, *Coronavirus, Whatsapp avvia fact-checking delle Informazioni al tavolo di autoregolamentazione su "piattaforme Digitali e big data*", Comunicato Stampa, 2 aprile 2020.

[43] AGCOM, Delibera n. 3/23/CONS, "Regolamento in materia di individuazione dei criteri di riferimento per la determinazione dell'equo compenso per l'utilizzo online di pubblicazioni di carattere giornalistico di cui all'articolo 43-bis della legge 22 aprile 1944, n. 633".

le evoluzione di quanto già acquisito anche attraverso la sua attività di regolazione e co-regolazione.
In definitiva, occorre far sì che gli indicatori individuati, oltre a tutelare il corretto funzionamento del mercato, non incentivino al contempo pratiche di disinformazione. Questo, come abbiamo visto, non solo a tutela del mercato e del consumatore, che trarrebbe nocumento dai comportamenti opportunistici appena ricordati, ma più in generale a tutela di un ulteriore bene anch'esso garantito dalla Costituzione, nella fattispecie dall'art. 21, rappresentato dal diritto all'informazione, di cui il pluralismo costituisce uno degli elementi caratterizzanti e la stessa ragion d'essere dell'AGCOM.

6. Considerazioni conclusive

In definitiva, ciò che emerge dal quadro qui delineato, è, che alla luce del nuovo contesto normativo, ai fini della più efficace tutela del pluralismo, particolare rilevanza vada data all'informazione e al relativo mercato, dando prevalenza agli "indici quantitativi di diffusione dei programmi", misurati in chiave cross-mediale, rispetto al sistema dei media nel suo complesso.
A fronte di un contesto che cambia radicalmente, l'adeguamento della strumentazione antitrust, con poche eccezioni, non sembra in grado di cogliere gli elementi di novità collegati alla trasformazione digitale, che sia in tema di concorrenza, ma ancor più nell'ambito del pluralismo, presenta forti elementi di discontinuità, soprattutto nel momento in cui all'efficienza economica tipica del diritto della concorrenza, si contrappongono fenomeni del tutto diversi e che destano forti preoccupazioni in chiave di pluralismo come ad esempio la disinformazione.
Da questo punto di vista la trasformazione digitale porta con sé anche una più ampia divaricazione, come mai in passato, tra concorrenza e pluralismo.
L'AGCOM è dunque chiamata a intervenire all'interno di un percorso tracciato dalla legge, nelle sue linee generali, ma che l'Autorità è chiamata ad attuare secondo approcci meno tradizionali, nella consapevolezza che la tutela di un bene ancor più fondamentale oggi, in questa particolare fase storica, sia primario rispetto a ogni altro interesse.
In questa chiave, cercando di fornire una utile chiave di lettura che possa contribuire alla migliore definizione delle linee guida, ripercorriamo le principali evidenze emerse in questo nostro lavoro:

- Il profondo processo di trasformazione digitale ha reso possibile la convergenza tra mercati in precedenza separati all'interno di un più ampio ecosistema, accrescendo la complessità e la reciproca interdipendenza tra i diversi settori.
- Nel nuovo allargato campo di gioco, favorito dall'esplosione dello streaming, il primo risultato è stato la progressiva migrazione del pubblico dalla tv lineare, broadcast, verso i servizi online di video streaming. Questo fenomeno è destinato a durare, nonostante gli incessanti continui cambiamenti nei comportamenti di consumo e nei modelli di business.
- Ciò impone una revisione nella definizione dei mercati, in primo luogo in ambito antitrust, rendendo ormai obsoleto e non più giustificabile, in termini di mercati

rilevanti, la distinzione tra TV in chiaro e TV a pagamento

- Un'ulteriore conseguenza è la necessità di allargare l'analisi anche al di fuori dell'ambito televisivo, in particolare valutando caso per caso le relazioni strategiche orizzontali tra le diverse componenti dell'ecosistema, a seconda dello stato di evoluzione del mercato, risolvendolo, all'interno di ciascuna componente, in termini di pressione competitiva o di vera e propria sostituibilità.
- In termini di potere di mercato, si evidenzia la contendibilità, indipendentemente dalla quota di mercato in termini di ricavi, del settore della pay TV, dove si assiste a un sensibile aumento del numero di operatori e dei loro ricavi all'interno di una torta che rimane sostanzialmente stabile, e dove i comportamenti potenzialmente anti-competitivi vengono messi in atto soprattutto dai nuovi entranti
- Più nello specifico, analizzando la questione del significativo potere di mercato in tema di pluralismo, diventa centrale il tema dell'informazione (e della disinformazione) dove questo unico genere, e non altri che poco o nulla hanno a che fare con il pluralismo, merita di essere analizzato e misurato in una prospettiva più ampia e complessa, multimediale e cross-piattaforma e non settoriale, legata al mercato dell'attenzione e con sistemi di rilevazione e metriche più sofisticate, a partire dalla total audience.
- La posta in gioco è molto elevata e va ben aldilà della sola questione concorrenziale. L'AGCOM, facendo tesoro delle tante esperienze in questo ambito, approfondendo e sviluppando la componente più specifica collegata agli indici di pluralismo, potrà certamente trovare le opportune soluzioni in grado di garantire il rispetto delle regole, affinché il libero gioco democratico non venga messo in discussione.

Transumanesimo: tra privacy e *AI*

Giulio Lombardi - Antonio Lombardi - Antongiulio Lombardi

Sommario

1. Cenni alle esigenze sottese alla disciplina della protezione dei dati personali ed a quella della regolamentazione dell'*AI*. - 2. Transumanesimo. - 3. Transumanesimo, tutela dei dati e *AI*. - 4. Conclusioni.

Keywords

privacy – intelligenza artificiale – transumanesimo – dati – AI Act

1. Cenni alle esigenze sottese alla disciplina della protezione dei dati personali ed a quella della regolamentazione dell'*AI*

La disciplina della tutela dei dati personali[1] e quella relativa alle applicazioni di *AI*[2] si stanno rivelando in concreto estremamente interconnesse. Le due discipline nate in momenti storici diversi ed in luoghi diversi, l'una alla fine del 1800 negli Stati Uniti d'America[3] e l'altra all'inizio del XXI secolo nell'Unione Europea[4], rispondono ad esigenze solo in apparenza profondamente diverse: la prima quella della tutela dell'individuo nella circolazione dei suoi dati e la garanzia della libera circolazione degli stessi[5], la seconda quella di assicurare un sano sviluppo dei sistemi di *AI* limitando i rischi per gli individui [6] quando si servono di detti sistemi. Le due finalità sono in realtà coincidenti pur nella diversità dell'oggetto; astraendo dal contesto potremmo definir-

[1] Regolamento (UE) 2016/679 del Parlamento europeo e del Consiglio, del 27 aprile 2016, relativo alla protezione delle persone fisiche con riguardo al trattamento dei dati personali, nonché alla libera circolazione di tali dati e che abroga la direttiva 95/46/CE (regolamento generale sulla protezione dei dati).

[2] Proposta di regolamento del Parlamento europeo e del Consiglio che stabilisce regole armonizzate sull'intelligenza artificiale (legge sull'intelligenza artificiale) e modifica alcuni atti legislativi dell'Unione.

[3] S. D. Warren - L. D. Brandeis, *The Right to Privacy*, in *Harvard Law Review*, 4(5), 1890, 193 ss.

[4] *Emendamenti del Parlamento europeo, approvati il 14 giugno 2023, alla proposta di regolamento del Parlamento europeo e del Consiglio che stabilisce regole armonizzate sull'intelligenza artificiale (legge sull'intelligenza artificiale) e modifica alcuni atti legislativi dell'Unione*, in *europarl.europa.eu*, 14 giugno 2023.

[5] Art. 1 - Oggetto e finalità del GDPR.

[6] *Amendment 140 Proposal for a regulation, Article 1 – paragraph 1 (new) «1. The purpose of this Regulation is to promote the uptake of human-centric and trustworthy artificial intelligence and to ensure a high level of protection of health, safety, fundamental rights, democracy and the rule of law, and the environment from harmful effects of artificial intelligence systems in the Union while supporting innovation».*

le come: i) esigenza di diffusione e circolazione e ii) necessità di tutela dell'interesse dell'individuo.

Va sottolineato che se la disciplina a tutela dei dati personali è stata soggetta ad un lungo ed articolato processo di creazione ed evoluzione durato più di un secolo, al contrario, la disciplina degli utilizzi di sistemi di *AI* si è venuta definendo negli ultimi dieci anni ed ancora non risulta cristallizzata in un atto normativo generale né nell'Unione Europea né negli Stati Uniti d'America mentre solo nella Repubblica Popolare Cinese[7] sono state recentemente approvare disposizioni generali vincolanti limitatamente *AI* sistemi di *AI* generativa; inoltre, l'evoluzione dei contenuti della proposta di AI Act, che solo con gli emendamenti del dicembre 2022[8] ha offerto una prima disciplina dei *Deep Fake*[9] e dell'*AI* generativa[10], dimostra come sia estremamente difficile la definizione di regole utili a rispondere alle esigenze della costante evoluzione dell'*AI*.

Di particolare rilevanza è anche la considerazione che la tutela dei dati personali si è evoluta insieme al crescere degli utilizzi e del numero dei dati scambiati senza che ci fosse mai un momento di vuoto normativo, di assenza di regole da applicare, lo stesso non può sostenersi per i sistemi di *AI* che vedono oggi miliardi di utilizzi a livello mondiale in assenza di regole specifiche quando solo per una porzione limitata di detti utilizzi è presente negli individui la consapevolezza della loro esistenza.

In tale ottica, in via generale, possiamo sostenere che, se il consenso[11] dell'interessato, libero e consapevole è uno strumento utile a risolvere gran parte delle possibili problematiche connesse con la tutela dei dati personali, detto istituto risulta non idoneo al medesimo fine in relazione alla disciplina dei sistemi di *AI* che per la pervasività nelle loro applicazioni travalicano necessariamente gli interessi del singolo, rendendo non dirimente il consenso dello stesso: tale elemento risulta particolarmente rilevante nell'analisi delle implicazioni del transumanesimo, che ad oggi risulta totalmente non regolato in modo organico e consapevole.

2. Transumanesimo

Il transumanesimo studia i possibili utilizzi di innovazioni tecnologiche volti a migliorare la funzionalità fisica o mentale dell'uomo, analizzando le implicazioni tecniche, mediche e giuridiche.[12] Gli occhiali da vista, la sedia a rotelle, le protesi sostitutive delle

[7] *Measures for the management of generative Artificial intelligence service Interim Measures for the management of generative Artificial intelligence service*, in *cac.gov.cn*, luglio 2023.

[8] *Amendments adopted by the European Parliament on 14 June 2023 on the proposal for a regulation of the European Parliament and of the Council on laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts*, in *europarl.europa.eu*, 14 giugno 2023.

[9] Emendamenti 101, 102, 399.

[10] Emendamenti 203, 486, 487.

[11] Art. 4, n. 11, GDPR.

[12] Il transumanesimo è un movimento che interessa la scienza medica, il diritto e la filosofia. Le sue declinazioni sono estremamente variegate e vanno dalla rigorosa valutazione scientifica sino a progetti che allo stato attuale della scienza e della tecnica rasentano la fantascienza. Mentre le applicazioni più concrete che comportano un miglioramento attuale del livello di vita reale dei cittadini necessitano di

gambe in kevlar, il sintetizzatore vocale che trasforma in suono il testo "digitato" con gli occhi, sono alcuni dei tanti esempi che ci aiutano a capire che il fenomeno non è né nuovo né rivoluzionario in sé ma è solo l'ultimo tassello di una evoluzione iniziata molto tempo fa. Indubbiamente sono nuove le sue implicazioni future, ove il transumanesimo sia applicato in modo massivo per migliorare funzionalità già esistenti modificando la struttura fisica e mentale di un individuo. Il massimo sviluppo del Transumanesimo dovrebbe portare secondo i suoi fautori ad una liberazione dell'umanità dal vincolo dell'evoluzione tramite un suo controllo preventivo.

Per la prima volta fu teorizzato da Julian Huxley[13] nel 1957 - negli stessi anni in cui nasceva il concetto di intelligenza artificiale - nel libro *"New Bottles for New Wine"*[14].

La finalità è ridefinire i limiti del corpo ed in prospettiva della mente dell'Uomo venendo o a porre rimedio a patologie o a potenziare situazioni nelle quali è totalmente assente ogni deficienza funzionale, con limiti e valutazioni etiche ben più complesse in questo ultimo caso[15]. L'aumento delle capacità fisiche è oggi una realtà e gli esperimenti per il miglioramento della capacità mentali sono molto avanzati: possibilità che oggi possono sembrare al limite della fantascienza saranno diffuse e nella disponibilità di molti nel medio periodo ed è per questo che una riflessione sulle implicazioni in diritto delle loro applicazioni è doverosa in particolare in relazione alla tutela dei dati personali dell'utilizzatore non tanto nei confronti del soggetto che gestisce il potenziamento il cui ruolo ed attività risulta giustificato dalla necessità di dare corso al contratto quanto piuttosto verso i terzi che vengono in relazione con l'individuo potenziato sia che il potenziamento abbia sanato gli effetti di una situazione patologica sia che si sia migliorata una situazione fisiologica.

Per completezza si fa presente che non vengono qui analizzate le implicazioni degli utilizzi militari del potenziamento umano: come sappiamo detto ambito risulta escluso da ogni regolamentazione in vigore ed in via di approvazione in materia di tutela dei

una impellente disciplina, l'applicazione più estrema che prevede un ancora teorico download della coscienza si scontra con la semplice constatazione che ogni download di dati contenuti nella nostra mente sta alla stessa come una nostra fotografia sta a noi. Per tale ragione detto argomento e ogni valutazione filosofica è stata attentamente evitata nel presente testo. Tra i tanti testi si citano: J. M. MacFerlane, *Transhumanism as a New Social Movement. The Techno-Centred Imagination,* Londra, 2020; S. Lilley, *Transhumanism and Society: The Social Debate over Human Enhancement*, Berlino-Heidelberg-Dordrecht-New York, 2012; N. Lee, *The Transhumanism Handbook*, Berlino-Heidelberg-Dordrecht-New York, 2019.

[13] Fratello di Haldous autore de *Il mondo nuovo*. Il nonno era Thomas Henry che sostenne le teorie di Darwin.

[14] Julian Huxley, *New Bottles for New Wine*, Londra, 1957 «*I believe in transhumanism: once there are enough people who can truly say that, the human species will be on the threshold of a new kind of existence, as different from ours as ours is from that of Peking man. It will at last be consciously fulfilling its real destiny*».

[15] U. Ruffolo - A. Amidei distinguono tra modifiche "curative" e modifiche "migliorative" in *Intelligenza artificiale, human enhancement e diritti della persona*, in *Intelligenza artificiale il diritto, i diritti, l'etica*, Milano, 2020. Altri distinguono tra 1) miglioramento fisico 2) miglioramento cognitivo 3) miglioramento emotivo e morale. Si veda: S. Holm - M. McNamee, *Physical enhancement: What baseline, whose judgment?*, in J. Savulescu - R. ter Meulen - G. Kahane (a cura di), Enhancing human capacities, Hoboken, 2011, 291 ss.; T. Kjærsgaard, *Enhancing motivation by use of prescription stimulants: The Ethics of Motivation Enhancement*, in *Ajob Neuroscience*, 6(1), 2015, 4 ss.; S. K. Nagel, *Shaping children: Ethical and social questions that arise when enhancing the young*, Berlino-Heidelberg-Dordrecht-New York, 2019.

dati e di *AI*[16].

Il transumanesimo è quindi un fenomeno già in essere le cui conseguenze estreme possono oggi inquietare alcuni, in modo non diverso di quanto un trapianto o una trasfusione avrebbe potuto spaventare un uomo medio della metà del 1800, ma che, tuttavia, necessita di una attenta e preventiva analisi giuridica delle implicazioni, specialmente in tema di tutela dei dati personali.

È certo che le applicazioni di *AI* utilizzate per sistemi di potenziamento umano che implicano le maggiori cautele etiche, che sono quelle connesse con il potenziamento delle capacità cognitive, sono oggi ancora in una fase del tutto sperimentale, a differenza di quelle che consentono un potenziamento fisico: l'utilizzo di impianti per consentire l'interlavoro diretto tra uomo e macchina o il caricamento di dati ha evidentemente un livello di potenziale lesività dei diritti connessi con i dati personali che non può essere semplicemente disciplinato facendo riferimento a categorie esistenti: né all'epoca della prima proposta di quello che sarebbe diventato il GDPR né all'epoca della sua pubblicazione era infatti neanche presente nel legislatore nazionale o dell'Unione Europea la consapevolezza dell'esigenza di disciplinare dette fattispecie[17].

I dati relativi *AI* sistemi di potenziamento umano sono di per sé dati medici e per questo dati particolari ai sensi del GDPR ma sono anche dati indispensabili all'esecuzione del rapporto contrattuale con il soggetto che esegue, monitora e riallinea, se necessario, il potenziamento: in tale ottica è necessario provvedere ad una analisi del rapporto tra consenso dell'interessato e diritto del terzo in relazione ai dati personali utilizzati delineando il fondamento giuridico *in primis* del potenziamento e poi del trattamento dei dati personali conseguenti allo stesso. In tale ottica l'art. 32 della Costituzione[18] è per alcuni[19] una possibile fonte giuridica per disporre da parte dell'interessato le "mo-

[16] Gli utilizzi militari sono oggi ben più diffusi e avanzati di quanto sia pubblico. Evidentemente tali utilizzi sono parte dell'attività che il militare deve compiere per il suo ufficio e i dati relativi sono nella totale disponibilità dei singoli eserciti. In internet è disponibile un'ampia descrizione di detti utilizzi che vanno dal potenziamento di sensi (vista e udito) ad impianti neurali ricercando articolo scientifici con le parole chiave *military human enhancement*.

[17] Ciò è un chiaro esempio della non tempestività dell'aggiornamento della normativa alle esigenze presentate dalla evoluzione tecnologica che risulta uno dei principali temi da affrontare per ogni branca del diritto che abbia una interrelazione tecnologica.

[18] Vanno rammentati anche i contenuti dell'articolo 3 della Carta dei diritti fondamentali dell'Unione Europea, Diritto all'integrità della persona: «1. Ogni individuo ha diritto alla propria identità fisica e psichica.
2. Nell'ambito della biologia e della medicina devono essere in particolare rispettati il consenso libero e informato della persona interessata, secondo le modalità stabilite dalla legge;
– il divieto delle pratiche eugenetiche, in particolare di quelle aventi come scopo la selezione delle persone;
– il divieto di fare del corpo umano e delle sue parti in quanto tali una fonte di lucro;
– il divieto della clonazione riproduttiva degli esseri umani».

[19] U. Ruffolo – A. Amidei, *Intelligenza Artificiale e diritti della persona: le frontiere del transumanesimo*, in *Giurisprudenza italiana*, 7, 2019, 1658 ss., Capitolo 4 «Libertà di autodeterminazione e "diritto a potenziarsi". L'art. 5 c.c. e gli artt. 32 e 2 (e 13) Cost.»: «[..]la dialettica tra gli artt. 2 e 32 Cost., la reciproca integrazione dei due precetti costituzionali ha ispirato una interpretazione dinamica della nozione di "salute" ex art. 32 Cost., intesa non quale mera assenza di malattia, bensì quale generale ed armonica "condizione [...] di benessere fisico e psichico della persona", come la definisce anche il Codice di deontologia medica; condizione perseguita anche dall'esercizio delle libertà del singolo, dominus del proprio corpo, nella realizzazione della propria personalità in sintonia con la percezione

difiche" fisiche che ritenga necessarie sia a carattere curativo che migliorativo, purché non disposte a vantaggio di terzi; il riferimento all'art. 5 del Codice civile[20] appare del tutto inconferente in quanto le modifiche sono disposte dall'interessato per sua scelta e volontà e senza che ci sia un vantaggio per terzi. Già più di dieci anni fa ben prima che le attuali potenzialità tecniche si affacciassero alla soglia del possibile, Rodotà aveva teorizzato la modifica del corpo per attribuire nuove capacità o per restituirne di perdute[21]. Ciò che è certo è che in ogni caso le applicazioni che consentono il potenziamento umano, anche note come *Human Enhancement Technologies* (*HET*), devono sempre rispettare i diritti fondamentali e la dignità umana.

Per completezza si rappresenta che i sistemi *HET* devono rispettare a livello internazionale: 1) le previsioni della Carta dei diritti fondamentali dell'Unione Europea in particolare con riferimento alla dignità umana di cui all'art. 1[22] ed il diritto alla integrità

che ha di sé (anche in relazione al diritto alla "identità personale")». Si pensi alla disposizione del proprio corpo che ha luogo nel processo di modifica del sesso.

[20] «Gli atti di disposizione del proprio corpo sono vietati quando cagionino una diminuzione permanente della integrità fisica, o quando siano altrimenti contrari alla legge, all'ordine pubblico o al buon costume».

[21] S. Rodotà, *Costruzione del Corpo*, in *Treccani.it*, 2009, «Così, in maniera sempre più marcata, il corpo si presenta come perennemente 'incompiuto', disponibile per una ininterrotta e sempre più incisiva attività di costruzione/modificazione. Su di esso è possibile intervenire per reintegrarne funzioni perdute o mai possedute (amputazioni, cecità, sordità) o proiettarlo al di là della sua antropologica normalità, rafforzandone le funzioni o aggiungendone di nuove, sempre in nome del benessere della persona, o della sua competitività sociale (incremento delle attitudini sportive, "protesi" per l'intelligenza). Siamo di fronte a "*repairing and capacity enhancing technologies*", a una moltiplicazione delle tecnologie *body-friendly*, che dilatano e modificano la nozione di cura del corpo e annunciano l'avvento dei *cyborgs*, del corpo postumano. «Nelle nostre società il corpo tende a divenire una materia prima modellabile secondo l'ambiente del momento». Si allargano così le possibilità di intervento individuale, ma crescono anche le opportunità di interventi politici di controllo del corpo attraverso le tecnologie»; S. Rodotà, *Il diritto di avere diritti*, Roma-Bari, 2013. Degna di menzione è la categorizzazione proposta nel 2005 dal Gruppo europeo per l'etica delle scienze e delle nuove tecnologie: *The European group on ethics in science and new technologies*, *Ethical* aspects of ICT implants in the human body, in *ec.europa.eu*, 3 marzo 2009, «Dispositivi ICT: dispositivi che si avvalgono delle tecnologie dell'informazione e della comunicazione, solitamente basati sulla tecnologia dei chip di silicio. Dispositivo medico attivo: qualsiasi dispositivo medico il cui funzionamento si basa su una fonte di energia elettrica interna e indipendente, ovvero su una fonte di energia diversa da quella generata direttamente dal corpo umano o dalla gravità. Dispositivo medico attivo impiantabile: qualsiasi dispositivo medico attivo destinato a essere impiantato interamente o parzialmente mediante intervento chirurgico nel corpo umano, o mediante intervento medico in un orifizio naturale, e destinato a rimanervi dopo l'intervento.

Dispositivi ICT passivi impiantabili: dispositivi ICT impiantabili nel corpo umano che utilizzano per il funzionamento un campo elettromagnetico esterno (Sezione 3.1.1. relativa al 'Verichip'). Dispositivi ICT impiantabili online: dispositivi ITC impiantabili che utilizzano per il funzionamento una connessione ('online') con un computer esterno, o che sono interrogabili ('online') da un computer esterno (Sezione 3.1.2. relativa *AI* biosensori). Dispositivi ICT impiantabili offline: dispositivi ICT impiantabili il cui funzionamento non dipende da dispositivi ICT esterni (eventualmente dopo un'operazione iniziale di configurazione, come nel caso della stimolazione cerebrale profonda)».

[22] Art. 1 – Dignità umana. La dignità umana è inviolabile. Essa deve essere rispettata e tutelata.

della persona di cui all'art. 3[23] ma anche l'art. 7[24] relativo al diritto alla vita privata e familiare; 2) le previsioni della Convenzione Europea dei diritti dell'Uomo;[25] 3) le previsioni della Convenzione di Oveido[26] ed in particolare l'art. 1 e 2; 4) nel caso di utilizzo di apparati elettronici situazione che ha luogo sempre quando i sistemi *HET* utilizzano *AI*, le previsioni del regolamento sulle apparecchiature mediche 2017/45; 5) La direttiva NIS 216/1148 deve sicuramente essere rispettata in caso di connessione in rete dell'apparato di *AI* a supporto del *HET*.

3. Transumanesimo, tutela dei dati e *AI*

Di tutti i possibili utilizzi dell'*AI*, la sua applicazione a sistemi che rendono possibile il miglioramento delle funzionalità umane è sicuramente tra quelli che pongono maggiore necessità di analisi in relazione alla tutela dei dati degli utilizzatori ma anche delle persone che si trovano vicino agli stessi. Se infatti l'interessato ha di norma la piena disponibilità delle proprie scelte in tema di diffusione o limitazione della circolazione dei propri dati, i sistemi di *AI* che aumentano le capacità dell'individuo rendono indispensabile non solo il consenso dell'interessato ma la puntuale consapevolezza dei terzi di trovarsi di fronte ad un utilizzo di un sistema di *AI*[27].
La prospettiva è quindi rovesciata nell'*AI Act* rispetto a quella del GDPR e tale impostazione sarà applicabile anche agli utilizzi di sistemi di *AI* utilizzati in strumenti di

[23] Art. 3 – Diritto all'integrità della persona. 1. Ogni individuo ha diritto alla propria integrità fisica e psichica.
2. Nell'ambito della medicina e della biologia devono essere in particolare rispettati: il consenso libero e informato della persona interessata, secondo le modalità definite dalla legge, il divieto delle pratiche eugenetiche, in particolare di quelle aventi come scopo la selezione delle persone, il divieto di fare del corpo umano e delle sue parti in quanto tali una fonte di lucro, il divieto della clonazione riproduttiva degli esseri umani.»

[24] Art. 7 – Rispetto della vita privata e della vita familiare. Ogni individuo ha diritto al rispetto della propria vita privata e familiare, del proprio domicilio e delle sue comunicazioni.

[25] Art. 8 – Diritto al rispetto della vita privata e familiare. 1. Ogni persona ha diritto al rispetto della propria vita privata e familiare, del proprio domicilio e della propria corrispondenza. 2. Non può esservi ingerenza di una autorità pubblica nell'esercizio di tale diritto a meno che tale ingerenza sia prevista dalla legge e costituisca una misura che, in una società democratica, è necessaria alla sicurezza nazionale, alla pubblica sicurezza, al benessere economico del paese, alla difesa dell'ordine e alla prevenzione dei reati, alla protezione della salute o della morale, o alla protezione dei diritti e delle libertà altrui.

[26] Convenzione per la protezione dei Diritti dell'Uomo e della dignità dell'essere umano nei confronti dell'applicazioni della biologia e della medicina: Convenzione sui Diritti dell'Uomo e la biomedicina, Oviedo, 4 aprile 1997.
Art. 1 – Oggetto e finalità. Le Parti di cui alla presente Convenzione proteggono l'essere umano nella sua dignità e nella sua identità e garantiscono ad ogni persona, senza discriminazione, il rispetto della sua integrità e dei suoi altri diritti e libertà fondamentali riguardo alle applicazioni della biologia e della medicina. Ogni Parte prende nel suo diritto interno le misure necessarie per rendere effettive le disposizioni della presente Convenzione.
Art. 2 – Primato dell'essere umano. L'interesse e il bene dell'essere umano debbono prevalere sul solo interesse della società o della scienza.

[27] Come previsto dallo schema di Regolamento *AI Act* che impone la consapevolezza nell'umano di interagire con l'*AI*. *Recital* 16 del testo approvato il 14 giugno 2023: l'*AI Act* è pervaso dalla esigenza di consapevolezza dell'umano di interagire con un sistema di AI.

transumanesimo.

Nell'ambito delle applicazioni di transumanesimo partendo dalle protesi intelligenti che potenziano per l'esecuzione di lavori pesanti o da quelle che consentono di camminare creando una connessione elettrica che supera le conseguenze di una lesione fisica, per arrivare ad utilizzi ancora futuribili che potrebbero consentire un collegamento diretto tra mente umana e *AI*, appare imprescindibile una revisione normativa che consenta una lettura unitaria e ragionata delle norme del GDPR e *AI Act* ma anche la determinazione di nuovi principi ove quelli dei due regolamenti non fossero in grado di rispondere all'esigenza concreta in modo adeguato.

Ad esempio, riteniamo che un sistematico utilizzo di *DPIA*[28] e di *AI*[29], auspicabilmente redatte congiuntamente, sia indispensabile per evitare una erronea valutazione dei rischi connessi alle singole applicazioni e sistemi di potenziamento dell'uomo e, al contrario, per assicurare un utilizzo ottimale dei sistemi che combinano *AI* e interfaccia dirette con l'uomo. Il comma 6 dell'articolo 29 dell'ultima versione del *draft AI Act* prefigura tale impostazione.[30]

Inoltre, dovrà essere identificata una modalità sicura e non invasiva di comunicazione a tutti i terzi della presenza di un sistema di *AI* che supporta una applicazione di super umanesimo, ciò senza ledere il diritto dell'utilizzatore; una soluzione potrebbe essere l'obbligo per il sistema di *AI* di inviare un messaggio che avverta della presenza del sistema di *AI* senza indicare il mittente, che avrà sempre la possibilità per sua scelta di avvertire della propria identità.

Grande attenzione deve poi essere posta nel contemperare le esigenze non perfettamente allineate di GDPR e *AI Act* in relazione alla qualità e natura dei dati richiesti dall'*AI Act* e dal GDPR. Per l'art. 10 dell'*AI Act* i set di dati di addestramento e, ove applicabile, i set di dati di convalida e prova devono essere pertinenti, sufficientemente rappresentativi, adeguatamente verificati in termini di errori e il più possibile completi alla luce della finalità prevista[31]. Per il GDPR i dati devono essere minimizzati,

---

[28] Art. 35 del Regolamento (UE) 679/2016, "Valutazione d'impatto sulla protezione dei dati".

[29] Art. 29 della proposta di regolamento del Parlamento europeo e del consiglio che stabilisce regole armonizzate sull'intelligenza artificiale (legge sull'intelligenza artificiale) e modifica alcuni atti legislativi dell'Unione. ECNL - Data&Society, *Recommendations for Assessing AI Impacts to Human Rights, Democracy, and the Rule of Law*, in *ecnl.org*, 23 novembre 2021; Vanja Skoric, *Critical Criteria for AI Impact Assessment: An Aggregated View*, in *papers.ssrn.com*, 2 giugno 2023. B. C. Stahl - J. Antoniou - N. Bhalla - L. Brooks *et. al.*, *A systematic review of artificial intelligence impact assessments,* in *Artificial Intelligence Review,* 56, 11, 2023, 12799 ss.

[30] «*Users of high-risk AI systems shall use the information provided under Article 13 to comply with their obligation to carry out a data protection impact assessment under Article 35 of Regulation (EU) 2016/679 or Article 27 of Directive (EU) 2016/680, where applicable*».

[31] L'evoluzione dell'art. 10 è traccia evidente del lavoro svolto su questo tema. *Text proposed by the Commission Amendment 3. «Training, validation and testing data sets shall be relevant, representative, free of errors and complete. They shall have the appropriate statistical properties, including, where applicable, as regards the persons or groups of persons on which the high-risk AI system is intended to be used. These characteristics of the data sets may be met at the level of individual data sets or a combination thereof». Amendment 3. «Training datasets, and where they are used, validation and testing datasets, including the labels, shall be relevant, sufficiently representative, appropriately vetted for errors and be as complete as possible in view of the intended purpose. They shall have the appropriate statistical properties, including, where applicable, as regards the persons or groups of persons in relation to whom the high-risk AI system is intended to be used. These characteristics of the datasets shall be met at the level of individual datasets or a combination thereof».*

necessari e proporzionati[32]. In tale ottica appare indispensabile, ancor più che in altri utilizzi, nei sistemi che utilizzando l'*AI* vengano a potenziare le capacità umane, che il programmatore riesca ad assicurare che tutti i criteri citati siano rispettati. Detta attività è stata resa indubbiamente più agevole dalla introduzione nell'ultima versione della proposta di *AI Act* di numerose norme di coordinamento generale tra disposizioni contenute nell'*AI Act* e previsioni del GDPR. Ricordiamo tra le altre l'emendamento 7 al Considerando 2-*ter*[33], l'emendamento 53 al Considerando 26-*quinquies*[34] ma soprattutto l'emendamento 80 al Considerando 45-*bis* [35] che conferma la rilevanza del principio della minimizzazione dei dati e fornisce un criterio interpretativo generale per la corretta lettura dell'art. 10 dell'*AI Act* ed in particolare del concetto di dati "sufficientemente rappresentativi".

Altro aspetto che necessita di una attenta analisi è la possibilità di "leggere" stati mentali che in soggetti che siano stati sottoposti ad opportuni *enhancements* risulta strutturalmente applicabile. Tale situazione ha fatto teorizzare il diritto alla *privacy* mentale (in inglese il "*right to mental privacy")* che dovrebbe evitare utilizzi indebiti delle informazioni acquisite con le tecniche che consentano l'accesso diretto alla mente[36], tanto che qualcuno ha chiesto l'introduzione di un *Mental Data Protection Impact Assesment*[37].

---

[32] Art. 5, c. 1, lett. *c).*

[33] Considerando 2-*ter*. Il diritto fondamentale alla protezione dei dati personali è garantito in particolare dai regolamenti (UE) 2016/679 e (UE) 2018/1725 e dalla direttiva (UE) 2016/680. La direttiva 2002/58/CE tutela, inoltre, la vita privata e la riservatezza delle comunicazioni, in particolare stabilendo le condizioni per l'archiviazione di dati personali e non personali e l'accesso AI dati archiviati in apparecchi terminali. Tali atti giuridici costituiscono la base di un trattamento dei dati sostenibile e responsabile, anche laddove i set di dati includano una combinazione di dati personali e non personali. Il presente regolamento non mira a pregiudicare l'applicazione del diritto dell'Unione esistente che disciplina il trattamento dei dati personali, inclusi i compiti e i poteri delle autorità di controllo indipendenti competenti a monitorare la conformità con tali strumenti. Inoltre, esso non pregiudica il diritto fondamentale alla vita privata e alla protezione dei dati personali previsto dal diritto dell'Unione in materia di protezione dei dati e della vita privata e sancito dalla Carta dei diritti fondamentali dell'Unione europea (la "Carta").

[34] Considerando 26-*quinquies*. Il presente regolamento non dovrebbe incidere sulle pratiche vietate dalla legislazione dell'Unione, ivi incluso dalla normativa in materia di protezione dei dati, non discriminazione, protezione dei consumatori e concorrenza.

[35] Considerando 45-*bis*. Il diritto alla vita privata e alla protezione dei dati personali deve essere garantito durante l'intero ciclo di vita del sistema di AI. A tale riguardo, i principi della minimizzazione dei dati e della protezione dei dati fin dalla progettazione e per impostazione predefinita, sanciti dal diritto dell'Unione in materia di protezione dei dati, sono essenziali quando il trattamento dei dati comporta rischi significativi per i diritti fondamentali delle persone fisiche. I fornitori e gli utenti dei sistemi di AI dovrebbero attuare misure tecniche e organizzative all'avanguardia al fine di tutelare tali diritti. Tali misure dovrebbero includere non solo l'anonimizzazione e la cifratura, ma anche l'uso di tecnologie sempre più disponibili che consentano di inserire algoritmi nei dati e di ricavare informazioni preziose senza la trasmissione tra le parti o un'inutile copia degli stessi dati grezzi o strutturati.

[36] S. Ligthart - T. Douglas - C. Bublitz - T. Kooijmans *et al.*, *Forensic Brain-Reading and Mental Privacy in European Human Rights Law: Foundations and challenges*, in *Neuroethics*, 14(2), 2020, 191 ss.; A. W. Paz, *Is your neural data part of your mind? Exploring the conceptual basis of mental privacy*, in *Minds and Machines*, 32(2), 2021, 395 ss.

[37] Si intende per *digital mind technology* una tecnologia volta all'esplorazione, all'analisi, all'influenza dei dati mentali.
Per dato mentale si intende ogni dato che può essere organizzato e processato per dedurre gli stati mentali della persona includendo gli stati cognitivi, affettivi e conativi. M. Ienca - G. Malgieri, *Mental data protection and the GDPR*, in *Journal of Law and the Biosciences*, 9(1), 2022.

4. Conclusioni

Dall'analisi di quanto accaduto a livello tecnologico e normativo negli ultimi anni appare evidente che il super umanesimo avrà uno sviluppo sempre più pervasivo e che una disciplina unitaria dello stesso appare necessaria per tutelare in modo adeguato gli utilizzatori ed i terzi, non risultando del tutto idonee le disposizioni esistenti in altri settori ad indirizzare tutte le problematiche che il super umanesimo pone. Alla luce delle diverse problematiche sollevate, la soluzione più opportuna è sicuramente quella di prevedere una disciplina ex ante che sia in grado di indirizzare in modo completo le esigenze legali, tecniche e mediche ponendo dei chiari limiti espressi e preventivi ad utilizzi tecnicamente possibili ma che possano venire a violare i principi applicabili alla base del GDPR e dell'attuale testo dell'*AI Act*. In particolare, possiamo sin d'ora identificare alcune esigenze che sono tratte sia dalle disposizioni del GDPR che dell'*AI Act* che i sistemi di *HET* che utilizzano *AI* dovranno rispettare: 1) l'esigenza che l'umano scelga sempre in modo consapevole ed informato di sottoporsi ad un trattamento di *HET* ed il connesso tema della tracciabilità e spiegabilità dei sistemi di *AI* asserviti a sistemi di *HET*; 2) la presenza costante della possibilità per l'umano di orientare sempre ciò che viene attuato dal sistema di *AI* asservito al sistema di *HET*; 3) il rispetto da parte di sistemi di *AI* che offrono *HET* del basilare principio di non discriminazione; 4) l'esigenza diversa di trasparenza verso i terzi che interagiscano con umani che utilizzino sistemi di *HET* che a loro volta beneficiano di sistemi di *AI*; 5) l'esigenza che i dati siano accurati, robusti e sicuri e che i livelli di accuratezza dei dati e le metriche di accuratezza siano dichiarate espressamente nelle istruzioni per l'uso; 6) l'esigenza che il sistema *HET* sia sviluppato ed offerto da un soggetto che abbia le necessarie conoscenze e qualifiche tecniche e che sia in grado con le necessarie assicurazioni di far fronte a richieste di risarcimento; 7) l'esigenza primaria che ogni sistema HET sia applicato senza porre a rischio la vita dell'umano considerando quanto ragionevolmente sicuro secondo lo stato delle conoscenze tecniche e mediche; 8) l'esigenza di non promuovere i sistemi di *HET* con campagne pubblicitarie che ne sviliscano la portata e i rischi.

In conclusione, considerando non sufficiente la regolamentazione di altri settori come quello medico o della responsabilità da prodotto, pure applicabili nelle more di future evoluzioni, è necessario prevedere un nuovo sistema di regole ed autorizzazioni specifiche che consentano l'offerta di sistemi *HET* solo da soggetti qualificati e solo a persone che siano state adeguatamente informate delle peculiarità ed implicazioni dell'impiego dei suddetti sistemi.

Elenco autori

Veronica Azzali
avvocato in Bolzano; docente di diritto dei media, Libera Università di Bolzano
Raffaele Bifulco
professore ordinario di diritto costituzionale, LUISS Guido Carli
Daniele Butturini
professore associato di diritto costituzionale, Università degli Studi di Verona
Nicol Ellecosta
dottoranda di ricerca in Scienze educative e sociali, Libera Università di Bolzano
Giuseppe Gallo
dottorando di ricerca in Diritti e tutele nei mercati globalizzati, Università degli Studi di Bari
Antongiulio Lombardi
direttore affari regolamentari
Antonio Lombardi
avvocato praticante in Roma
Giulio Lombardi
avvocato in Roma
Giulia Olivato
dottoranda in Studi giuridici comparati ed europei, Università degli Studi di Trento
Andrea Palumbo
avvocato in Milano
Maurizio Pedrazza Gorlero
professore emerito di diritto costituzionale, Università degli Studi di Verona
Jacopo Piemonte
avvocato in Milano
Augusto Preta
presidente dell'International Institute of Communications – Italian Chapter
Federico Riboldi
avvocato in Milano
Andrea Ranghino
avvocato in Milano
Federico Serini
dottorando di ricerca in Diritto pubblico, comparato e internazionale, La Sapienza – Università di Roma
Alessandro Tedeschi Toschi
avvocato praticante in Milano

CODICE ETICO

La **Rivista di diritto dei media** intende garantire la qualità dei contributi scientifici ivi pubblicati. A questo scopo, la direzione, il Comitato degli esperti per la valutazione e gli autori devono agire nel rispetto degli standard internazionali editoriali di carattere etico.

Autori: in sede di invio di un contributo, gli autori sono tenuti a fornire ogni informazione richiesta in base alla policy relativa alle submissions. Fornire informazioni fraudolente o dolosamente false o inesatte costituisce un comportamento contrario a etica. Gli autori garantiscono che i contributi costituiscono interamente opere originali, dando adeguatamente conto dei casi in cui il lavoro o i lavori di terzi sia/siano stati utilizzati. Qualsiasi forma di plagio deve ritenersi inaccettabile. Costituisce parimenti una condotta contraria a etica, oltre che una violazione della policy relativa alle submission, l'invio concomitante dello stesso manoscritto ad altre riviste. Eventuali co-autori devono essere al corrente della submission e approvare la versione finale del contributo prima della sua pubblicazione. Le rassegne di dottrina e giurisprudenza devono dare esaustivamente e accuratamente conto dello stato dell'arte.

Direzione: la direzione (ivi compresi direttori e vice-direttori) si impegna a effettuare la selezione dei contributi esclusivamente in base al relativo valore scientifico. I membri della direzione (ivi compresi direttori e vice-direttori) non potranno fare uso di alcuna delle informazioni acquisite per effetto del loro ruolo in assenza di un'esplicita autorizzazione da parte dell'autore o degli autori. La direzione è tenuta ad attivarsi prontamente nel caso qualsiasi questione etica sia portata alla sua attenzione o emerga in relazione a un contributo inviato per la valutazione ovvero pubblicato.

Comitato degli esperti della valutazione: i contributi sottoposti a valutazione costituiscono documentazione a carattere confidenziale per l'intera durata del processo. Le informazioni o idee acquisite confidenzialmente dai valutatori per effetto del processo di revisione non possono pertanto essere utilizzate per conseguire un vantaggio personale. Le valutazioni devono essere effettuate con profondità di analisi, fornendo commenti e suggerimenti che consentano agli autori di migliorare la qualità delle loro ricerche e dei rispettivi contributi. I revisori dovranno astenersi dal prendere in carico la valutazione di contributi relativi ad argomenti o questioni con i quali sono privi di familiarità e dovranno rispettare la tempistica del processo di valutazione. I revisori dovranno informare la direzione ed evitare di procedere alla valutazione nel caso di conflitto di interessi, derivante per esempio dall'esistenza di perduranti rapporti professionali con l'autore o la relativa istituzione accademica di affiliazione.

mediaLAWS

ANNO 2023 - CODICE ISSN 2532-9146