Intelligenza Artificiale e *deepfakes*: le nuove frontiere della disinformazione e i possibili rimedi giuridici*.

Maria Esmeralda Bucalo

Abstract

Partendo da alcune considerazioni su tecnologia e diritti, il lavoro si occupa delle nuove frontiere della disinformazione determinate dal sorgere e dall'evolversi dell'Intelligenza Artificiale, la quale, oltre che comportare nuove sfide per il costituzionalismo, riacuisce formule già esistenti di discriminazione. Attraverso l'esame di alcuni casi concreti occorsi sia in Europa sia negli Stati Uniti, avrà modo di evidenziare, infatti, come i *deepfake* siano in effetti i "successori" tecnologicamente evoluti delle *fake news* e le diverse risposte che predispongono gli ordinamenti fra le due sponde dell'Atlantico per fronteggiarli, che sembrano ricalcare quelle già predisposte per le forme già conosciute di disinformazione nell'era digitale.

Starting from some considerations on the combination of technology and rights, the work deals with the new frontiers of disinformation determined by the rise and evolution of Artificial Intelligence, which, in addition to posing new challenges for constitutionalism, sharpens already existing formulas of discrimination. Through the examination of some concrete cases that occurred both in Europe and in the United States, it will be able to highlight, in fact, how deepfakes are in fact the technologically advanced "successors" of fake news and the different responses that predispose the orders between the two shores of the Atlantic to face them, which seem to follow those already prepared for the already known forms of disinformation in the digital age.

Sommario

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

italiano sull'Intelligenza Artificiale – 8. Considerazioni finali.

Keywords

Intelligenza Artificiale – *deepfake* – disinformazione – giurisprudenza UE – normativa europea

1. Premessa

Le riflessioni oggetto del presente lavoro devono partire dalla considerazione preliminare per la quale tecnologia e diritti fondamentali costituiscono un binomio la cui somma algebrica è rilevante in diritto e non può non essere tenuto in considerazione. Infatti, partendo dalla definizione matematica di binomio[1] ed estendendola poi agli studi giuridici, può constatarsi che, sebbene tecnologia e diritti fondamentali siano ontologicamente due monomi fra loro non assimilabili, il corso della storia ci ha insegnato che essi sono necessariamente congiunti: la prima, infatti, è stato il fattore che ha contribuito, da un lato, alla interpretazione evolutiva di quelli già positivizzati nelle Costituzioni, nelle Dichiarazioni e nelle Carte di livello internazionale e sovranazionale, consentendo il loro adeguamento alle nuove esigenze dei tempi e, dall'altro, alla codificazione dei cosiddetti "nuovi diritti" o "diritti di nuova generazione" [2].

Ciò è ancor più vero nel XXI secolo, nel quale la rivoluzione tecnologica digitale (c.d. "quarta rivoluzione industriale)[3] ha cambiato radicalmente il modo di vivere di ciascuno, permeando in modo capillare anche l'ambiente giuridico.

La diffusione di *Internet* ha però seguito un percorso diverso rispetto alla diffusione degli strumenti tecnologici del passato, intanto perché sono stati soggetti privati (e

[1] Secondo la definizione matematica il binomio è un polinomio, ridotto in forma normale e composto dalla somma o dalla differenza di due monomi non simili.

[2] Per quanto concerne la libertà di espressione e manifestazione del pensiero, intorno alla quale ruotano le considerazioni che verranno rese qui appresso, si pensi all'invenzione della stampa che, oltre a contribuire allo sviluppo di un nuovo settore industriale, «ruppe il monopolio delle informazioni in capo a pochi privilegiati e consentì [...] l'accesso alla cultura a gruppi sempre più ampi» (così L. Torchia, *Lo Stato digitale,* Bologna, 2023, 35).

Fu così che, insieme al progredire della tecnologia e allo sviluppo della radio e della televisione come ulteriori mezzi di comunicazione, alla visione strettamente "individualista" della libertà di espressione, che determinava la necessità che lo Stato ed ogni apparato pubblico si astenesse da ogni azione che con essa potesse interferire, si sommò anche quella "funzionale e sociale", che determinava la necessità che i pubblici apparati intervenissero e regolassero il settore rendendolo fruibile a tutti i consociati e garantendo il pluralismo. In tema si veda L. Torchia, *Stato digitale,* cit., 17.

[3] La fortunata definizione è dovuta a K. Schwab, *La quarta rivoluzione industriale*, Milano, 2016. Si ritiene convenzionalmente che la prima rivoluzione industriale, fra la fine del XVIII e i primi decenni del XIX secolo, sia stata determinata dall'invenzione del motore a scoppio ed il conseguente aumento della produzione; la seconda, nel secolo successivo, dallo sviluppo dell'elettricità, dalla conseguente diffusione delle catene di montaggio e, dunque, delle produzioni di massa, che non richiedevano più specifiche competenze dei lavoratori; infine, la terza c.d. "informatica", che comincia negli anni '50 del secolo scorso con l'invenzione dei calcolatori elettronici, dalla diffusione della tecnologia digitale e di *Internet*, come mezzo che rende le informazioni accessibili a tutti anche da dispositivi mobili sempre più potenti e sempre più economici.

non pubblici) a organizzare e rendere disponibile a quante più persone i nuovi servizi; in secondo luogo, perché la capacità di incisione di questi nuovi mezzi di informazione e comunicazione nella vita personale e sociale assume dimensioni straordinarie, non coprendo più soltanto l'area dell'informazione, ma anche quella della elaborazione e della trasmissione delle culture[4].

In questa nuova rivoluzione tecnologica accade dunque ciò che in passato non si era mai verificato, e cioè che l'innovazione sia «al tempo stesso una caratteristica intrinseca del potere pubblico e un fenomeno la cui regolazione è centrale per i rapporti economici e sociali complessivamente intesi»[5].

La locuzione usata per riassumere e definire questa novità è "Stato Digitale", che presenta due nuove caratteristiche rispetto al passato. La prima è che l'attività pubblica nel suo complesso viene trasformata, quanto a mezzi e modalità di svolgimento, dall'uso delle nuove tecnologie, che ne determinano una riarticolazione e una riorganizzazione di funzioni e strutture.

La seconda è determinata dalla capacità di incisione dello sviluppo tecnologico sui rapporti sociali ed economici in modo tale da rendere sempre più obsolete e inidonee le regole vigenti. Da qui la necessità di una nuova regolazione pubblica volta ad aggiornare le discipline esistenti, con l'introduzione di nuove regole che si adeguino alla realtà attuale[6].

Le considerazioni che seguono hanno ad oggetto le nuove forme di manifestazione del pensiero e di espressione generate per mezzo della tecnologia, e in specie dell'Intelligenza Artificiale, che si diffondono attraverso la Rete su scala globale. Esse appaiono a prima vista l'evoluzione di quelle già sviluppatesi in "età digitale", grazie allo sviluppo dei *social media* e, più in generale, di tutte le piattaforme di condivisione.

Se così è, conseguentemente anche la disinformazione ha subito una analoga evoluzione grazie all'avanzamento tecnologico. Oggi, infatti, la diffusione di *fake news* viene sopravanzata dalla diffusione di *fake* generati dalla IA (i cc.dd. *deepfake*), che, grazie al maggiore impatto delle immagini e dei video artificialmente creati, hanno una capacità nociva e dannosa potenziata rispetto alle prime, spesso integrando fattispecie di reato e riacuendo discriminazioni.

Che i *deepfake* siano l'evoluzione delle *fake news* sarà reso evidente anche dall'analisi delle risposte e dei rimedi predisposti da ordinamenti diversi, che, come si avrà modo

[4] U. De Siervo, *Informazione, comunicazione globale e privacy*. Secondo K. Schwab, *La quarta rivoluzione industriale,* cit., 23, il *punctum crucis* che distingue la "quarta rivoluzione industriale" da quelle precedenti, è che essa non riguarda soltanto la diffusione di nuovi strumenti di comunicazione e di produzione, ma il fatto che essa determini anche mutamenti sociali, cambiando drasticamente il mondo in cui viviamo. Ciò si rileva anche dal fatto che il dato rilevante ai fini della sua individuazione non sia tanto la diffusione di nuovi strumenti tecnologici (si pensi alla diffusione dei calcolatori elettronici considerata convenzionalmente l'elemento determinante il sorgere della "terza rivoluzione industriale"), bensì la velocità dell'avanzamento tecnologico, la portata e l'intensità delle innovazioni, oltre che l'impatto prodotto sui sistemi aziendali, di produzione, ma anche sociali in generale.

[5] L. Torchia, *Lo Stato digitale,* cit.

[6] L. Torchia, *Prefazione*, in V. Bontempi (a cura di), *Lo Stato digitale nel Piano di Ripresa e Resilienza*, Roma, 2022, 11 ss., ma anche in *Lo Stato digitale,* cit., 22, spec. 18-19 l'A. spiega come l'aggettivo "digitale" si sommi oggi a quelli che hanno accompagnato l'evoluzione dello Stato, il quale certamente continua a svolgere tutte le funzioni assunte in precedenza.

di approfondire, ricalcano quelli già apprestati per le seconde. Essi possono ascriversi tutti al "costituzionalismo digitale", inteso come "nuova stagione del costituzionalismo", che tende ad «ampliare, a completare e a rafforzare gli strumenti del costituzionalismo tradizionale […] effetto diretto delle trasformazioni che la scienza e la tecnica hanno determinato nella sfera fisica, psichica e relazionale della persona umana»[7].
Si tratterebbe dunque di un "costituzionalismo globale" espansione del paradigma "costituzionale tradizionale" che, pur non volendo rompere con la tradizione e con il passato, guardano al futuro, optando per una rifondazione della democrazia costituzionale mediante l'introduzione di adeguate tecniche e funzioni di garanzia[8].
La risposta del diritto alle "rotture costituzionali"[9] e alle alterazioni dell'"equilibrio costituzionale" prodotte dalle tecnologie digitali, determinate dalla evoluzione tecnologica, dovrebbe essere reperita in un "processo di costituzionalizzazione dell'ambiente digitale" che consenta di bilanciarle, identificando nell'insieme di valori del costituzionalismo tradizionale le necessarie "contromisure"[10].

2. Problemi definitori dell'Intelligenza Artificiale

Fra le tecnologie sviluppate l'Intelligenza Artificiale è certamente quella maggiormente rilevante, essendo divenuta la forma più progredita di raccolta ed elaborazione dei dati, grazie ai meccanismi di *machine learning*, in grado di incidere profondamente sui diritti di ciascuno.
Essa da sempre ha stimolato grandi opere letterarie e cinematografiche che hanno contribuito a costruirne progressivamente il mito[11]. Menzionarle in questa sede ha un senso, perché contribuisce a ricordarci che la cultura di partenza sul tema, che in qualche modo ha formato le convinzioni (e i pregiudizi) di ciascuno su questa tecnologia, la descrive essenzialmente come fonte di grandi pericoli per l'umanità, capace di ribellarsi in modo autonomo ai suoi inventori, i quali ne perdono il controllo e possono

[7] E. Cheli, *Conclusioni*, in *Osservatorio sulle fonti*, 2, 2021, 955-956.
Si veda anche P. Costanzo, *Il fattore tecnologico e le sue conseguenze*, relazione tenuta al convegno annuale della Associazione Italiana dei Costituzionalisti, svoltosi a Salerno 23-24 novembre 2012, 28, che parla di «costituzionalismo tecnologico»; G. Azzariti, *Internet e Costituzione*, in *costituzionalismo.it*, 2, 2011, e G. De Gregorio, *Digital Constitutionalism in Europe: Reframing Rights and Powers in the Algorythmic Society,* Cambridge, 2022.

[8] L. Ferrajoli, *Costituzionalismo oltre lo Stato*, Modena, 2017, 44 ss. L'A. più recentemente ha sviluppato le stesse tesi in *Il futuro del costituzionalismo*, in *costituzionalismo.it*, 2, 2022, 182 ss. Analogamente L. Antonini, *Globalizzazione e nuove sfide del costituzionalismo*, in *Diritto pubblico*, 2, 2019, 319 ss.

[9] F. Balaguer Callejon, *La Constitucion de l'algoritmo*, Zaragoza, 2023, 16-17, il quale afferma che la «Costituzione analogica» regola «un mondo che in parte non esiste più o è divenuto socialmente irrilevante» ed è necessario invece «analizzare la realtà digitale dal punto di vista delle rotture che essa sta generando e che hanno una dimensione costituzionale», al fine di «proporre soluzioni che permettano di mitigare tali rotture e facilitino una risposta costituzionale».

[10] E. Celeste, *Digital Constitutionalism: a new systematic theorization*, in *International Review of Law*, 1, 2019, 88, 93 e 99.

[11] Fra le prime va menzionata la produzione di Asimov datata anni 30 del secolo passato, mentre fra le seconde capolavori come *2001: Odissea nello spazio* (di Kubrick del 1968) *Blade Runner, Matrix* e più di recente *L'uomo bicentenario* e *Her.*

rimediarvi solo staccando la spina[12].

Come è stato rilevato, non si può non tenere in considerazione come tale portato culturale abbia influenzato (e influenzi a tutt'oggi) anche l'approccio giuridico, e scientifico in generale, a questa tecnologia[13], tanto che il suo sviluppo è stato negli anni molto ondivago, alternando a periodi di grandi interesse e investimenti, periodi di disaffezione per l'argomento (denominati comunemente *"AI Winter"*)[14].

L'indubbia complessità del tema deriva, non solo dalla difficile caratterizzazione delle diverse tipologie di funzionamento dell'IA, ma anche, in via preliminare, dal punto di vista definitorio[15].

Nel ricordare che il primo studio scientifico sulla IA fu firmato dal matematico Alan Mathison Turing e risale alla prima metà del XX secolo[16] e che ad esso fece seguito il "*Dartmouth Summer Project Research on Artificial Intelligence*"[17], che per la prima volta configurò l'Intelligenza Artificiale come autonoma disciplina scientifica[18], è bene rammentare anche il metodo, invero ancora attuale, ivi a tal fine adottato.

Turing, infatti, piuttosto che definire cosa fosse l'intelligenza, cosa invero assai difficile, preferiva confrontare i risultati di un processo: se il processo era qualificato intelligente quando svolto da un essere umano, allora il raggiungimento di uguali risultati attraverso un processo svolto da una macchina determina che anche quest'ultimo poteva essere definito "intelligente"[19].

---

[12] G. Finocchiaro, *Intelligenza Artificiale. Quali regole?,* Bologna, 2024, 16.

[13] G. Finocchiaro, ivi, 17 ritiene che tale portato culturale influenzi anche gli aspetti giuridici strettamente tecnici, poiché non si può dimenticare che «coloro che interpretano o scrivono le regole sono inevitabilmente condizionati dalla cultura di cui sono portatori».

[14] G.F. Italiano, *Intelligenza Artificiale: passato, presente, futuro,* in F. Pizzetti (a cura di), *Intelligenza artificiale, protezione dei dati personali e regolazione*, Torino, 2018, 209 ricostruisce un primo periodo (1974-1980), nel quale si registra la carenza di attenzione introno agli studi scientifici sull'argomento sono seguiti alla diffusione di due report (rispettivamente del 1966 dell'*Automatic Language Processing Advisory Committe* del governo statunitense e quello denominato *Lighthill* del 1973 del governo inglese), che ritennero improbabile nel breve periodo lo sviluppo di tale tecnologia, tagliando dunque i relativi finanziamenti; mentre il secondo *"AI Winter"* andrebbe dal 1987 al 1993. Sebbene ancora nei primi anni del XXI secolo il tema non godesse di grande attenzione, oggi le ricerche sull'IA vivono la loro *"AI Spring"* (o *"AI Boom")* registrando i più alti livelli di interesse e finanziamento nella storia.

[15] Sulla difficoltà definitoria associata all'IA, M.U. Scherer, *Regulating Artificial Intelligence Systems: risks, challanges competencies, and strategies,* in *Harvard Journal of Law & Technology,* 2, 2016, 359, nonché C. Casonato, *Intelligenza Artificiale e diritto costituzionale: prime considerazioni,* in *Diritto pubblico comparato ed europeo,* numero speciale, 2019, 102-103.

[16] Ci si riferisce in particolare agli studi di A.M. Turing, *Computing Machinery and Intelligence*, in *Mind*, 236, 1950, 433 ss.

[17] Proposta da J. Mccarthy – M.L. Minsky – N. Rochester – C.E. Shannon, *A proposal for the Dartmouth summer research project on artificial intelligence*, 31agosto 1955, 1.

[18] Alla conferenza parteciparono anche altri sei studiosi: Ray Solomonoff, Oliver Selfridge, Trenchard More, Arthur Samuel, Allen Newell e Herbert Simon. Fra questi in particolare, Newell e Simon presentarono il *Logic Theorist*, il primo programma esplicitamente progettato per imitare le capacità di *problem solving* degli esseri umani.

[19] A.M. Turing, *Computing Machinery and Intelligence,* cit., 433. Si tratta dell'*Imitation Game* di A.M. Turing, *Computing Machinery and Intelligence*, cit., 433, ideato per determinare se una macchina sia in grado di effettuare collegamenti, concatenare idee ed infine esprimerle. Con estrema esemplificazione può dirsi che secondo l'A. l'intelligenza artificiale è la scienza di far fare ai computer cose che richiedono intelligenza quando vengono compiute da esseri umani. Dunque, chiedersi se le macchine possono

A dimostrare l'importanza e l'interesse trasversale che suscita il tema, è bene notare che l'"approccio controfattuale", nato nell'area delle c.d. "scienze dure", è stato ripreso recentemente e sviluppato anche in ambito umanistico e filosofico da Luciano Floridi che ribadisce che l'IA non ha nulla a che fare con l'intelligenza, poiché «separa la capacità di risolvere un problema o di portare a termine un compito con successo dalla esigenza di essere intelligenti per farlo»[20].

Lo stesso A. ritiene che di IA non esista una definizione univoca come per molte delle «cose importanti della vita […] (che) spesso non sono affatto definibili», ma che riconosciamo quando le vediamo. Di conseguenza è possibile ritenere che «IA non è un termine scientifico, […] ma un'espressione generica […] una scorciatoia, usata per riferirsi approssimativamente a diverse discipline, servizi, prodotti tecnoscientifici talora solo genericamente correlati»[21].

Riprendendo quanto sopra rilevato sul retaggio culturale intorno alla IA, è possibile allora ritenere, come lo stesso fa Floridi, che della IA esistano "due anime", una "ingegneristica" (altrimenti detta "riproduttiva") e l'altra "cognitiva" (detta anche "produttiva"). La prima interessata alla riproduzione di comportamenti umani definiti intelligenti (e che è quella di cui scientificamente ci si occupa), l'altra come settore della scienza cognitiva interessata alla produzione di intelligenza, che attualmente resta una idea fantascientifica[22].

La mancanza di una piena autonomia definitoria è evidente anche nell'ambito più strettamente giuridico, come dimostrato anche da recenti atti normativi (e para-normativi).

Nella Risoluzione del Parlamento Europeo del 16 febbraio 2017 recante raccomandazioni alla Commissione concernenti norme di diritto civile sulla robotica, l'IA sembra essere intesa quale elemento strumentale allo sviluppo di altre tecnologie, essendo maggiormente incentrata sul concetto di "robot autonomo intelligente" e non come fenomeno a sé stante e non necessariamente legato ad una componente *hardware*[23].

Più utile è invece quella resa dal Gruppo di Esperti sull'Intelligenza Artificiale nominato dalla Commissione Europea nel 2019, per la quale «I sistemi di Intelligenza Artificiale (AI) sono sistemi software (ed eventualmente anche hardware) progettati da esseri umani che, dato un obiettivo complesso, agiscono nella dimensione fisica o digitale percependo il loro ambiente attraverso l'acquisizione di dati, interpretando i dati strutturati o non strutturati raccolti, ragionando sulla conoscenza, o l'elaborazione

pensare era per Touring una domanda «troppo insensata per meritare di essere discussa». Sul punto si veda anche G. Finocchiaro, *Intelligenza Artificiale*, cit., 22.

[20] L. Floridi, *Etica dell'Intelligenza Artificiale. Sviluppi, opportunità, sfide*, Milano, 2022, 41, il quale ritiene anche che «l'IA non concerne la capacità di riprodurre l'intelligenza, ma in realtà la capacità di farne a meno».

[21] Ivi, 42.

[22] Ivi, 48-50.

[23] Risoluzione del Parlamento europeo del 16 febbraio 2017 recante raccomandazioni alla Commissione concernenti norme di diritto civile sulla robotica (2015/2013INL). Specificatamente si veda l'allegato a detta Risoluzione ove si rende una definizione di *robot* autonomo intelligente e non di IA in senso autonomo dalla robotica. Sottolineano le criticità di tale definizione C. Cath – S. Wachter – B. Mittelstadt – M. Taddeo – L. Floridi, *Artificial Intelligence and the "Good Society": the US, EU, and UK approach*, in *Science and Engineering Ethics*, 2, 2018, 514-515.

delle informazioni, derivata da questi dati e decidere le migliori azioni da intraprendere per raggiungere l'obiettivo prefissato. I sistemi di intelligenza artificiale possono utilizzare regole simboliche o apprendere un modello numerico e possono anche adattare il proprio comportamento analizzando il modo in cui l'ambiente è influenzato dalle loro azioni precedenti»[24].

La correttezza della definizione citata appare peraltro confermata dall'*AI Act* che al suo art. 3 definisce l'IA come «un sistema automatizzato progettato per funzionare con livelli di autonomia variabili e che può presentare adattabilità dopo la diffusione e che, per obiettivi espliciti o impliciti, deduce dall'*input* che riceve come generare *output* quali previsioni, contenuti, raccomandazioni o decisioni che possono influenzare ambienti fisici o virtuali»[25].

Entrambe le definizioni hanno il pregio di non antropomorfizzare l'IA ed evitano di riproporre il confronto uomo-macchina, che era presente negli studi sulla Intelligenza Artificiale fin dai suoi albori, non associando ai sistemi artificiali né il concetto di intelligenza umana né altre facoltà tipicamente legate alla sfera biologica[26].

Inoltre, la stessa consente anche di tracciare una distinzione fra ciò che ad oggi può essere inteso verosimilmente rientrare nel concetto di IA e ciò che, invece, rimane distante dalla realtà. In particolare, il Gruppo di Esperti fa una netta distinzione tra "IA generale" (o IA forte) e "IA ristretta" (o IA debole). Quest'ultima tipologia rimane circoscritta ad una forma di abilità funzionale per lo svolgimento di compiti specifici[27], mentre, l'IA generale descrive un sistema dotato di estrema versatilità e pertanto capace di svolgere qualsiasi attività eseguibile da parte di un essere umano.

Sarebbe, dunque, la "IA ristretta" a rappresentare la tipologia dei sistemi artificiali ad

---

[24] High-Level Expert Group on Artificial Intelligence, *Una definizione di IA: principali capacità e discipline*, Bruxelles, 8 aprile 2019, 6. La definizione consta anche di una seconda parte che considera la IA come una disciplina scientifica affermando che «In quanto disciplina scientifica, l'intelligenza artificiale comprende diversi approcci e tecniche, come l'apprendimento automatico (di cui il deep learning e l'apprendimento di rinforzo sono esempi specifici), il ragionamento automatico (che include pianificazione, programmazione, rappresentazione e ragionamento della conoscenza, ricerca e ottimizzazione) e la robotica (che comprende il controllo, la percezione, i sensori e gli attuatori, nonché l'integrazione di tutte le altre tecniche nei sistemi ciberfisici).» La definizione perfeziona quella precedentemente proposta dalla Commissione europea, *Comunicazione su L'Intelligenza artificiale per l'Europa*, Bruxelles, 25.4.2018, COM(2018) 237 final, 1 « L'intelligenza artificiale (AI) si riferisce a sistemi che mostrano un comportamento intelligente analizzando il loro ambiente e intraprendendo azioni – con un certo grado di autonomia – per raggiungere obiettivi specifici. I sistemi basati sull'intelligenza artificiale possono essere puramente basati su software, agendo nel mondo virtuale (ad esempio assistenti vocali, software di analisi delle immagini, motori di ricerca, sistemi di riconoscimento vocale e facciale) oppure l'intelligenza artificiale può essere incorporata in dispositivi hardware (ad esempio robot avanzati, automobili autonome, droni o applicazioni Internet of Things).»

[25] Sull'*AI Act* si veda *infra* § 7.2.

[26] Così R. Cucchiara, *Intelligenza Artificiale e Italia. Sfide e opportunità*, in 2, *Gnosis*, 2019, 48-49. Analogamente, L. Floridi, *What the near future of Artificial Intelligence cloud be*, in *Philosophy & Technology*, 32, 2019, 2-3, non fa riferimento al concetto di intelligenza quanto piuttosto a quello di *agency*, intesa come abilità d'azione, tanto da definire l'IA quale «reservoir of smart agency on tap». R. Cingolani, *Il corpo e la mente. Robot e uomini nel futuro dell'Intelligenza Artificiale*, in 2, *Gnosis*, 2019, 75-76 riassume efficacemente la distinzione fra uomo e IA affermando che «il robot segue le leggi dell'elettricità, il corpo umano quelle della biochimica». Sulla necessità di non antropomorfizzare l'IA anche G. Finocchiaro, *La regolazione dell'Intelligenza artificiale,* in *Rivista trimestrale di diritto pubblico,* 4, 2022, 1087.

[27] R. Cingolani, *L'altra specie. Otto domande su noi e loro*, Bologna, 2019, 105 ss.

oggi effettivamente implementati ed in uso, mentre la "IA generale" l'obiettivo, definito "utopico"[28], che suscita vivaci dibattiti, ma, in modo rassicurante per tutti, non trova applicazioni concrete[29].

Con molta semplificazione, può dirsi quindi che l'Intelligenza Artificiale non opera secondo un ragionamento logico, ma è un *software* che si avvale di un'impostazione *data-driven,* il quale, attraverso l'elaborazione dei dati, porta il sistema artificiale a sviluppare e seguire un proprio modello matematico per svolgere l'operazione che gli è stata assegnata.

Posta tale definizione, il Gruppo di Esperti ne propone una bipartizione che individua due "macroaree" distinte in base alla capacità di apprendere e ricalibrare la propria attività in relazione ai mutamenti intervenuti nell'ambiente in cui la macchina opera.

Una prima categoria di sistemi di IA si limita a riprodurre in maniera automatizzata il meccanismo in essi contenuto, risultando capaci di giungere all'obiettivo assegnatoli a partire dai dati raccolti in ingresso e seguendo stabilmente, volta dopo volta, il processo decisionale e di elaborazione-ragionamento descritto all'interno del loro codice (il c.d. *data-set* individuato dal programmatore e da lui inserito nel sistema di IA). In questi casi, dunque, la macchina esegue processi automatizzati lineari, caratterizzati da meccanismi di *input-output* e da una forte corrispondenza tra l'impulso iniziale e l'esito del processo.

Una seconda categoria più complessa costituita dai sistemi di IA che a quanto sopra descritto aggiungono una autonoma "capacità di apprendimento", cosicché, attraverso l'interazione con il contesto circostante e l'assimilazione di nuove informazioni, risultano in grado di cambiare la propria strategia comportamentale (c.d. sistemi *machine learning*).

Di conseguenza, questi sistemi non operano solo attraverso procedimenti lineari (*input-output*), ma, a prescindere dall'input immesso in avvio, sono in grado di tenere in considerazione sia l'ambiente circostante sia l' "esperienza pregressa" costituita dagli *output* precedentemente emessi. In tal modo acquisiscono una certa autonomia dall'intervento umano del programmatore e ricalibrano il proprio "comportamento" sulla base dei riscontri avuti dalle passate interazioni con il contesto di riferimento.

Fra queste si distinguono i modelli più semplici e quelle di *deep learning*. I modelli più semplici si organizzano su tre livelli: il livello di *input* dei dati, il livello nascosto (*hidden layer*) destinato all'elaborazione delle informazioni e, da ultimo, il livello di *output* che produce l'esito del processo algoritmico.

I secondi, invece, risultano più complessi e maggiormente articolati, e si caratterizzano per essere composti da più livelli, ciascuno dei quali riceve *input* dallo strato precedente e alimenta con il proprio *output* lo strato successivo. Sono dunque chiamati sistemi di *deep learning,* perché sono in grado di completare più sofisticati processi di addestramento o di analisi di dati, appunto con "maggiore profondità".

Se quanto appena rilevato è certamente utile ai fini della definizione e della compren-

[28] High-Level Expert Group on Artificial Intelligence, *Una definizione di IA: principali capacità e discipline*, cit., 5-6.

[29] Commission nationale informatique & libertes, *How can human keep the upper hand? The ethical mat-ters raised by algorithms and artificial intelligence*, December 2017, 19.

sione di ciò che è Intelligenza Artificiale, delle sue diverse tipologie e dei suoi meccanismi di funzionamento, ritornando all'ambito giuridico che ci compete, ciò che risulta ancora troppo poco indagato (e regolato) sono le modalità attraverso le quali i dati vengono elaborati e immessi in tali sistemi, nonché la ricostruzione del procedimento attraverso il quale gli stessi sistemi di IA raggiungono determinati risultati, soprattutto laddove si faccia riferimento ai citati sistemi di *deep learning*.

3. Intelligenza Artificiale e nuove forme di discriminazione e disinformazione: i *deepfake*

La crescente penetrazione dell'IA in ogni aspetto del nostro vivere quotidiano e l'oscurità delle regole sulla base delle quali questa nuova tecnologia elabora i dati pongono necessariamente degli interrogativi, ai quali il diritto costituzionale ha il compito di rispondere, in modo da trarne tutti i possibili vantaggi e benefici ed evitando che i diritti della persona subiscano una ingiustificata compressione.
La capacità di incisione di tale tecnologia sui diritti fondamentali è l'oggetto delle riflessioni che seguono e che si occuperanno in particolare del diritto di informazione e della libertà di manifestazione del pensiero, indagando anche i profondi pregiudizi che l'uso dell'Intelligenza Artificiale comporta per l'individuo e la sua personalità, ivi compresi gli aspetti più intimi e delicati della stessa, come quelli attinenti alla sessualità.
Su questo versante, sempre preliminarmente, si può fare riferimento, per esempio, ai *sex robots* definibili come «una entità artificiale con forma umanoide, comportamenti quasi umani, un certo grado di intelligenza, usati per scopi sessuali»[30]. Come è evidente l'uso di queste macchine pone questioni profonde e intime, che riguardano i caratteri connotativi dell'essere umano che, sul versante giuridico, coinvolgono il principio di eguaglianza di genere e di non discriminazione[31].
In tema di *sex robots,* anche al fine di indagare i risvolti disinformativi che può avere l'uso della IA nella generazione di immagini, va citato il caso della diffusione *on line* della notizia, accompagnata dalla relativa foto generata dalla IA, per la quale Elon Musk sarebbe stato sul punto di presentare la sua fidanzata *robots*. La notizia poi rivelatasi falsa[32] era esplosa sui *social networks*, senza però essere mai stata ufficialmente smentita dallo stesso Musk o dalle sue aziende, le quali invero lavorano già da tempo alla realizzazione di *robot* umanoidi progettati però per svolgere compiti fisici e non per

[30] R. Halwani nella recensione a J. Danaher – N. Macarthur (eds.), *Robot Sex: Social and Ethical Implications*, Cambridge, 2017 ospitata in *Bioethics*, 32, 2018, 639.

[31] C. Nardocci, *Intelligenza artificiale e discriminazione,* in *La Rivista Gruppo di Pisa,* 3, 2021, 9.
In riferimento alla questione di genere, inoltre, è stata osservata una azione definita come *robots gendering,* ossia quella volta ad avere un forte impatto sull'attività e il comportamento delle persone attraverso la manipolazione della voce e delle caratteristiche estetiche del *robot*. Gli studi dimostrano, infatti, come siano prevalentemente i maschi a considerare socialmente utili i *sexbots*, caratterizzati in senso per lo più femminile, e come evidentemente ciò conduca a stereotipi di genere già fortemente presenti in molte società. Si veda T. Nomura, *Robots and Gender*, in *Gender and the Genome*, 1, 2017, 18; M. Scheutz – T. Arnold, *Are we ready for sex robots?*, in *The Eleventh ACM/IEEE International Conference on human robot interaction*, 07 March 2016, 351.

[32] Si veda l'articolo pubblicato su *Open* il 30 maggio 2023.

"funzioni sociali" o "emotive"[33].

Se forse i *sexrobots* ci sembrano ancora una realtà eccessivamente lontana, basti pensare che le discriminazioni che il loro utilizzo può comportare hanno il loro prototipo negli "assistenti vocali" che utilizziamo quotidianamente (Siri o Alexa). Progettati per essere in ogni momento a completa disposizione del loro "padrone", difficilmente vengono realizzati con "sembianze" maschili, presentando, al contrario, il più delle volte, nomi e voci di donna, a cui viene affidato il compito di rivolgersi in maniera accondiscendente nei confronti dell'interlocutore e di esaudire qualsiasi richiesta gli venga fatta, anche laddove quest'ultima risulti sgradevole, offensiva o inopportuna.

Così una semplice scelta progettuale di questo tipo alla base dello strumento tecnologico perpetua e rafforza gli stereotipi discriminatori esistenti nei confronti del genere femminile, confinando la figura della donna ad un ruolo di subalternità e di assoggettamento, respingendo l'idea che la stessa possa occupare posizioni diverse[34].

A ciò possiamo aggiungere ulteriori utilizzi dell'IA che, oltre che riacuire la discriminazione fondata sul genere, possono produrne anche in ragione dell'origine etnica e del colore della pelle. Si fa riferimento, per esempio, all'implementazione sui dispositivi mobili e sui computer dei sistemi *facial recognition*, che soprattutto in occidente sono sviluppati per lo più sulla base di data set contenenti immagini di uomini e donne bianchi, con una conseguente sottorappresentazione di individui di diversa origine etnica e la possibilità di pervenire a risultai inesatti[35].

Le discriminazioni di genere ed etniche, già presenti nel passato, riemergono quindi in senso fortemente accentuato in quanto rinnovate dall'uso dello strumento tecnologico e dall'enorme quantitativo di dati, tanto che da più parti viene rilevato come in generale il diritto costituzionale e pubblico non risultino pienamente in grado di tutelare l'individuo dalle disparità di trattamento che possono derivare dall'utilizzo di sistemi di Intelligenza Artificiale[36].

La capacità discriminatoria della IA e la sua mancata neutralità[37] dipendono da una molteplicità di ragioni che attengono per lo più alla costruzione del *data-set*[38], cioè dall'introduzione di quell'insieme di dati di cui la macchina dispone per compiere le

[33] Si veda il video riportato dal *Corriere delle sera*, 11 ottobre 2024.

[34] Sul punto si veda N. Loideain – R. Adams, *From Alexa to Siri and the GDPR: The gendering of Virtual Personal Assis- tants and the role of Data Protection Impact Assessment*, in *Computer Law & Security Review*, 2020, 36.

[35] Si legga in tal senso European union agency for Fundamental rights, *Facial recognition technology: fundamental rights considerations in the context of law enforcement*, November 2019, 27.

[36] Si veda per esempio F.Z. Borgesius (eds.), *Discrimination, artificial intelligence, and algorithmic decision-making*, Council of Europe, Strasbourg, 2018, 18 ss. e L. Giacomelli, *Big brother is «gendering» you. Il diritto antidiscriminatorio alla prova dell'intelligenza artificiale: quale tutela per il corpo digitale*?, in *Biolaw Journal – Rivista di Biodiritto*, 2, 2019, 278 ss. Analogamente esprimeva la stessa opinione in tema di diritto del lavoro L. Peruzzi, *Il diritto antidiscriminatorio al test dell'Intelligenza Artificiale,* in *Labour & Law Issue,* 1, 2021, 50 ss.

[37] M.V. Craiut – I. Iancu, *Is technology gender neutral? A systematic literature review on gender stereoptypes attached to artificial intelligence,* in *Human Technology,* 2022, 18(3), 297-315 e M. Airoldi – D. Gambetta, *Sul mito della neutralità algoritmica*, in *The Lab's Quarterly*, 4, 2018, 25 ss. Analogamente M. D'Amico – C. Nardocci, *Intelligenza artificiale e discriminazione di genere: rischi e possibili soluzioni,* in G. Cerrina Feroni – C. Fontana – E.C. Raffiotta (a cura di), *AI Anthology,* cit., 251; M. D'Amico, *Una parità ambigua,* Milano, 314 e 319 e S. Barocas – A.D. Selbst, *Big Data Disparate Impact,* in *California law Review,* 2016, 671 ss.

[38] K. Crawford, *The Hidden Biases in Big Data,* in *Harvard Business review,* 2013.

scelte per le quali viene programmata[39]; dalla associazione tra i dati[40]; nonché da possibili casi di *proxy discrimination,* che si verificano quando un dato formalmente neutro viene elaborato dal sistema di apprendimento automatico in modo da realizzare una discriminazione in via mediata senza che l'utilizzatore ne sia consapevole[41]. La novità determinata dall'avvento di questa nuova tecnologia è legata alla circostanza che in questo caso il carattere *proxy* potrebbe essere individuato autonomamente dal sistema di apprendimento automatico, senza che l'utilizzatore ne sia consapevole.

In casi più gravi, è possibile che la macchina sia stata volutamente programmata per ottenere risultati discriminatori, poiché è proprio il programmatore umano a rendere la macchina uno strumento di discriminazione[42]. È evidente che queste ultime forme di discriminazione siano più facilmente e immediatamente individuabili, mentre tendono a sfuggire alle successive verifiche quelle inintenzionali, che originano dagli stessi algoritmi.

A queste discriminazioni di genere operate dall'Intelligenza Artificiale, per lo più dipendenti dall'uomo che agisce nella programmazione sulla base di pregiudizi legati alla struttura non paritaria della società, si deve aggiungere poi la marginalizzazione femminile che vede le donne estromesse da settori scientifici e disciplinari nevralgici per lo sviluppo dell'Intelligenza Artificiale che ancora sono "monopolio maschile"[43].

Se questi sono i dati di partenza, a rendere ancor più evidente la gravità del fenomeno dell'utilizzo non corretto della IA e come tale utilizzo non corretto possa incidere sulla libertà di espressione, anche in senso discriminatorio, sono i cosiddetti *deepfake.*

Il termine è un neologismo nato dall'incrocio tra la locuzione *deep learning* e *fake* e si riferisce a contenuti multimediali, quali immagini, video, audio e testo, falsi che sono generati o manipolati utilizzando algoritmi appunto di *deep learning*: l'intento è quello di indurre colui che li osserva a percepirli come una rappresentazione fedele della realtà[44].

---

[39] In pratica, se ad una macchina si forniscono fin dall'origine dati errati, falsati o incompleti, la macchina risponderà analogamente in modo errato, falsato o incompleto. In questo caso, i *data-set* possono anche essere costruiti sulla base di pregiudizi e discriminazioni implicite, presenti cioè nella mente e/o nella cultura del programmatore.

[40] I dati forniti possono essere corretti, ma, soprattutto per le *machine learning* che operano autonomamente, la loro associazione può essere errata e portare a risultati e scelte discriminatorie "in uscita", anche non previsti o non voluti in sede di programmazione.

[41] Si veda in tema A.E.R. Prince – D. Schwarcz, *Proxy Discrimination in the Age of Artificial Intelligence and Big Data,* in *Iowa Law revin,* 2020, 1257 ss.

Tali forme di discriminazione sono, invero, una pratica risalente cui si faceva ricorso ben prima della nascita dei sistemi di Intelligenza Artificiale. Si ricordi per esempio che, nella metà del 1900, per non concedere prestiti ai neri alcune banche utilizzarono l'indicazione dei codici postali o i confini dei quartieri al fine di escludere i prestiti a quartieri abitati per lo più da afroamericani, invece di operare una discriminazione diretta sulla base della razza. Ne parla F.Z. Borgesius, *Discrimination, Artificial Intelligence and Algorithmic Decision-Making,* Council of Europe, Strasbourg, 2018, 13-14.

[42] In tema si veda A. Venanzoni, *La valle del perturbante: il costituzionalismo alla prova delle intelligenze artificiali e della robotica*, in *Politica del diritto*, 2019, 2, 237-238, nonché P. Zuddas, *Intelligenza Artificiale,* in *Liber amicorum per Pasquale Costanzo*, 16 marzo 2020, 7 e G. Resta, *Governare l'innovazione tecnologica: decisioni algoritmiche, diritti digitali e principio di uguaglianza*, in *Politica del diritto*, 2, 2019, 208

[43] M. D'Amico – C. Nardocci, *Intelligenza artificiale,* cit., 256-258.

[44] Sulla nascita di tale fenomeno, sul suo intersecare la tematica della libertà di informazione, nonché per un esame della risposta normativa europea attraverso anche lo *Strengthened Code of Practice on Disinformation* si veda M. Cazzaniga, *Una nuova tecnica (anche) per veicolare disiformazione: le risposte europee ai*

La loro nascita è fissata convenzionalmente alla fine del 2017, quando un anonimo gruppo con lo pseudonimo *Deepfakes* pubblica i primi video falsi di natura pornografica su un *social* molto popolare chiamato *Reddit*, facendo uso dell'applicazione *FakeApp*: si trattava di immagini e video di star di Hollywood, realizzati a loro insaputa[45].
Dopo questo evento, nonostante la circolazione di *deepfake* pornografici sia stata proibita e bloccata tramite la rimozione degli stessi sulle piattaforme *social*, la creazione e divulgazione di *deepfake* è diventata inarrestabile[46].
L'ultimo caso eclatante che ha coinvolto un personaggio famoso risale peraltro solo a gennaio 2024, quando furono diffuse in modo virale sulla piattaforma *social X* immagini false, create con IA, della popstar Taylor Swift in atteggiamenti provocanti e atti sessuali[47]. Solo pochi giorni più tardi *X* metteva in atto una forma di censura inedita, rendendo inattive le ricerche col nome dell'artista. Il caso ha riaperto il dibattito in USA sulla necessità di disposizioni legislative per arginare i *fake*, sul modello del DSA europeo.
I *deepfake* vengono impiegati in diversi settori, come il cinema, la medicina, l'arte, le comunicazioni digitali, l'intrattenimento, per scopi commerciali (*e-commerce* e moda) e costituiscono la naturale evoluzione di altri metodi di contraffazione di dati sintetici, come la *Computer Graphica* (in grado di contraffare immagini e video digitali in 2D e 3D) e *Auto-tune* (primo *software* in grado di manipolare audio in modo autonomo).
Secondo il rapporto 2023 dell'Istituto Europeo per le norme delle Telecomunicazioni (ETSI) intitolato *"Securing Artificial Intelligence (SAI); Automated Manipulation of Multimedia Identity Representations*" «il rapido progresso della tecnologia informatica negli ultimi decenni ha reso sempre più semplice anche la manipolazione di foto, file audio e video, (perché) le tecniche di intelligenza artificiale consentono di automatizzare manipolazioni che in precedenza richiedevano una notevole quantità di lavoro manuale.»[48]
Dunque, a differenza del loro esordio, nel quale i *deepfake* ritraevano principalmente personaggi famosi, adesso invece chiunque può diventare vittima di tali contenuti. Ciò evidentemente determina «rischi sostanziali in vari contesti che vanno dalla diffamazione personale e alla apertura di conti bancari utilizzando false identità (attraverso attacchi alle procedure di autenticazione biometrica) fino alle campagne per influenzare l'opinione pubblica»[49].
Su quest'ultimo tema si ricordi da ultimo che a gennaio 2024, a poche ore dalle primarie statunitensi nel New Hampshire molti cittadini hanno ricevuto una *robocall* falsa, nella quale una voce del tutto uguale a quella del presidente Biden (evidentemente

deepfakes, in questa *Rivista*, 1, 2023, 170 ss.

45 Sul primo caso di deepfake, S. Maddocks, *'A Deepfake Porn Plot Intended to Silence Me': exploring continuities between pornographic and 'political' deepfakes*, in *Porn Studies*, 7, 2020, 415 ss.

46 Sul caso Taylor Swift, S. Ruiz Lichter, *Why the Taylor Swift AI Scandal is Pushing Lawmakers to Address Pornographic Deepfakes,* in *The National Law Review,* 22 aprile 2024. Sulla normativa europea *infra* § 7.

47 Le visualizzazioni di uno dei deepfake, di cui la cantante era vittima, sono state 45 milioni, con almeno 24 mila condivisioni e incalcolabili apprezzamenti degli utenti. "*Taylor Swift Ai*" è diventato presto un termine di ricerca in tendenza sulla piattaforma, con le immagini rimaste online per ore.

48 *Securing Artificial Intelligence (SAI); Automated Manipulation of Multimedia Identity Representations*, Istituto Europeo per le norme delle Telecomunicazioni (ETSI), rapporto 2023, 11.

49 *Securing Artificial Intelligence (SAI); Automated Manipulation of Multimedia Identity Representations*", cit., 10.

generata con IA), invitava a non votare e a conservare il proprio voto per le elezioni di novembre[50].

Ovviamente tutto ciò getta delle ombre oscure e preoccupanti sulla capacità manipolativa anche del consenso elettorale da parte di coloro che sono in grado di creare *deepfake*[51].

D'altro canto e in senso più specifico, i dati diffusi dall'Autorità Garante per protezione dei dati personali italiana, nonché filoni di ricerca recenti, evidenziano in modo preoccupante che oltre il 90% dei video *deepfake* si configurano come materiale a contenuto pornografico e che nella quasi totalità dei casi (98%) essi hanno ad oggetto donne[52].

Tale evoluzione deteriore dei *deepfake* è denominata "*deepfake pornography*", locuzione che fa riferimento all'utilizzo di tecniche generative per l'alterazione di immagini e video con contenuto pornografico.

Le vittime del contenuto in questione vengono "spogliate" artificialmente ed appaiono in atteggiamenti sessualmente espliciti, che in realtà sono a loro estranei. Poiché le immagini sono modificate artificialmente con la tecnica dello *face-swap* e, dunque, i protagonisti non sono effettivamente coinvolti nell'atto sessualmente esplicito, la creazione e la successiva divulgazione di tali contenuti comportano una lesione della loro dignità e della loro privacy, in quanto collocati forzatamente in un contesto a cui non appartengono, e conseguenzialmente l'insorgere di numerose fattispecie penalmente rilevanti, fra cui il furto di identità[53].

---

[50] Analogamente qualche giorno dopo, un video diffuso su *Facebook*, manipolato anch'esso attraverso l'IA, etichettava lo stesso presidente come "pedofilo schifo", mentre invece la ragazza baciata era la nipote. Si veda *Deepfake alle primarie USA. Clonata la voce del presidente Biden. Telefonata falsa finalizzata a invitare gli elettori a disertare il voto,* in *ansa.it,* 29 gennaio 2024.

[51] Si veda *Biden, un video manipolato sui social che lo accusa di pedofilia. Meta, un'etichetta contro i deepfake*, in *corriere.it,* 6 febbraio 2024.
Quanto alla formazione dell'opinione pubblica manipolata da *fake news* che circolano online soprattutto in periodi di elettorale si permetta di rinviare a M.E. Bucalo – F. Pacini *Informazione e formazione del consenso politico,* in M.E. Bucalo – M. Caporale – A. Sterpa (a cura di), *Diritto pubblico di Internet,* Napoli, 2024, 253-282, nonché a M.E. Bucalo, *I volti della libertà di manifestazione nell'era digitale, fra intermediari online, moderazione dei contenuti e regolazione,* Torino, 2023.

[52] Percentuale che scende al 77% nel caso di *deepfake* generici. I dati sono del Garante per la Protezione dei dati Personali, *Vademecum* 2020. Quanto alla dottrina si veda H. Kshetri, *The Economics of Deepfakes*, in *Computing's Economics*, 2023, 89 ss.; G. Macgregor, *Gun to your head: how deepfakes and other non-consensual synthetic media hold individual autonomy hostage*, in *UMKC Law Review*, 2, 2021, 431 ss.

[53] Così V. Azzali – N. Ellecosta, *La questione deepfake in Italia, una panoramica,* in questa *Rivista*, 3, 2023, 82, che individua tale fattispecie nella «artificiale "deprivazione" di un individuo del proprio volto, seguita dalla sovrapposizione dello stesso a quello di un'altra persona.».
Oltre allo stato di angoscia in cui permango le vittime di *deepfake,* determinato dalla paura che il video o le immagini falsi possano continuare a essere presenti sul *web* e ritenuti autentici, è bene rilevare anche che in un *report* della organizzazione *Home Security Heroes* pubblicato negli Stati Uniti nel 2024, è emerso che il 74% degli uomini che avevano consumato materiale pornografico *deepfake* non manifestasse sensi di colpa riguardo al proprio consumo. Tale dato è facilmente interpretabile nel senso di una pericolosa accettazione e normalizzazione per una parte significativa del pubblico di tali contenuti, come prodotti per l'intrattenimento per adulti.

4. Il moltiplicarsi dei casi di *deepfake* in Europa e negli Stati Uniti e i diversi modelli normativi predisposti a tutela delle vittime

La disponibilità crescente delle applicazioni, scaricabili gratuitamente su ogni dispositivo, attraverso le quali è possibile creare *deepnude* e la facilità del loro utilizzo, sono divenute evidenti in recentissimi casi di cronaca avvenuti in Spagna, in Italia e negli Stati Uniti.

Nel mese di settembre 2023 nella cittadina spagnola di Alendralejo una ventina di ragazzine minorenni hanno trovato in circolazione sul *web* video modificati con programmi di Intelligenza Artificiale che le ritraevano nude e coinvolte in atti sessualmente espliciti. I filmati, che mostrano il viso delle minori su corpi sconosciuti, sono stati condivisi in *chat Whatsapp* e su *Telegram*. La vicenda, resa nota dopo la denuncia delle madri delle vittime, che, oltre a rivolgersi alle autorità spagnole, si sono riversate sui *social* per condannare l'accaduto, ha portato alla luce le responsabilità della creazione e divulgazione dei video erano dei compagni di scuola delle vittime**,** anche loro minorenni.

In Italia un caso analogo si è verificato a marzo 2023, quando due ragazzi di una scuola superiore di Latina hanno "spogliato per scherzo" cinque compagne di classe e una docente per mezzo della *app BikiniOff*[54].

L'episodio appena citato è peraltro analogo a quello successo sempre a Latina, in un'altra scuola superiore, esattamente un anno dopo, quando tre ragazzi minorenni, oggi indagati, hanno analogamente usato delle foto di due compagne per "spogliarle" usando la medesima l'*app*. Anche in questo caso i fotomontaggi delle due studentesse nude sono stati poi condivisi attraverso *chat* e *social network* tanto che, in poche ore, sono diventate virali all'interno dell'istituto[55].

Come è chiaro dagli esempi sopra riportati, tali usi della l'Intelligenza Artificiale sono in grado di impattare con la libertà di espressione (specialmente laddove tali forme espressive vengano poi diffuse su piattaforme digitali e dunque abbiano una diffusione globale) e, laddove si tratti dei *deepfake pornografici,* ripropongono e amplificano

[54] *BikiniOff* è una applicazione 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. Il *deepnude* della docente risultò peraltro così convincente da comparire su due siti pornografici. Sul caso si vedano D. Barbera, *Tutti i rischi di usare Bikinioff, il chatbot che spoglia le donne*, in *wired.it*, 19 aprile 2023 e 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.

[55] In seguito ai sopracitati eventi verificatisi a Latina, la Procura dei Minori di Roma ha avviato due diverse inchieste e il Garante per la protezione dei dati personali ha avviato un'istruttoria nei confronti di *Telegram* e ha mantenuto alta la attenzione sul tema stilando anche un *vademecum* intitolato *Deepfake. Il falso che ti "ruba" la faccia (e la privacy)*, emettendo provvedimenti, documenti ufficiali e comunicati., affrontando il tema del diritto alla identità personale anche in ottica divulgativa, proprio per il costante emergere di nuove tecnologie.
Casi del genere si stanno peraltro velocemente moltiplicando in tutto il mondo. A maggio 2023, per esempio, una bufala girava sul *web*: si trattava della foto di una avvenente giovane donna con *decolté* importante esibito in una occasione pubblica e che veniva spacciata per il ministro giapponese della salute**.** In realtà il ministro giapponese della salute era un uomo di mezza età e l'immagine della donna non era reale ma creata dalla IA della piattaforma *ChatGPT*. Si veda K. Hao, *Deepfake porn is ruining women's lives. Now the law may finally ban it*, in *technologyreview.com*, 12 febbraio 2021; H. Laffier – A. Rehaman, *Deepfakes and Harm to Women*, in *Digital Life and Learning*, 1, 2023.

esponenzialmente il problema della discriminazione di genere, integrando peraltro ipotesi di reato.

Analogamente negli Stati Uniti, sta sollevando questioni legali e morali indubbiamente urgenti il caso di due adolescenti, che nello scorso anno sono stati arrestati in Florida in base ad una legge del 2022, che ha istituito il reato di diffusione di immagini sessualmente esplicite senza il consenso della vittima. I due, infatti, avevano creato e diffuso immagini *deepfake* di nudi, utilizzando l'IA per generare rappresentazioni esplicite dei loro compagni di classe minorenni.

Nonostante ciò, il fenomeno dei nudi e delle immagini esplicite generate dall'IA da parte di minori sta diventando un problema sempre più comune nei distretti scolastici degli Stati Uniti, tanto che anche altri Stati si sono dotati di leggi analoghe a quelle della Florida.

In Virginia, per esempio, una legge del 2014 puniva la diffusione di foto o video con l'intento di costringere, molestare o intimidire un'altra persona. Nel 2019, tale legge è stata emendata estendendo la fattispecie e includendovi video o immagini statiche false e prevendendo una pena fino a un anno e una multa fino a 2500 dollari.

Analogamente in California nel 2019 sono state approvate due leggi in tema di *deepfake*: una prima che punisce chi pubblica video o immagini manipolate dei politici con l'intento di screditarli nei 60 giorni che precedono un'elezione che li vede coinvolti e una seconda che, invece, permette a chi si ritrova suo malgrado protagonista di un video *hard*, pur non avendone mai girato uno, di fare causa all'autore del *deepfake*.

Pur non di meno, negli Stati Uniti non esiste una legge di livello federale che affronti specificatamente il tema dei nudi *deepfake* non consensuali, lasciando dunque ai singoli Stati il compito di gestire, autonomamente, l'impatto dell'IA generativa su questioni delicate come il materiale di abuso sessuale infantile, il *revenge porn* e la formazione del consenso a fini politici.

A tal fine, il 30 ottobre 2023, il presidente Biden aveva emesso un ordine esecutivo[56], incaricando il Dipartimento del Commercio di sviluppare linee guida sul *watermarking* dei contenuti generati dall'IA, al fine di segnalare quando un contenuto multimediale è stato creato artificialmente e come è stato successivamente modificato[57].

A tal fine, l'ordine delegava ai Dipartimenti e alle Agenzie il ruolo essenziale di definire le linee guida, nonché eventuali standard di sicurezza generali, puntando sulla formazione, sullo sviluppo e sulla collaborazione volontaria delle imprese per raggiungere la trasparenza e l'affidabilità dei sistemi.

Quanto al sistema di marchiatura dei contenuti, esso oggi assolve essenzialmente a due funzioni. In primo luogo, esso può essere utilizzato per segnalare al fruitore di un'immagine, di un video o di un documento quale sia il suo vero autore, fungendo così da equivalente grafico del *copyright*. In questo caso, è usato come marchio per definire la proprietà di un prodotto audiovisivo.

---

[56] Executive Order 14110 of October 30, 2023 *Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence*.

[57] Per *watermarking* si intende la "marchiatura" dei contenuti volta a identificarne gli autori e la autenticità. La pratica è tutt'altro che recente: suoi precursori *ante litteram* sono per esempio la filigrana nella carta, nelle marche da bollo e nelle banconote dal XIII secolo in avanti.

Accanto a tale funzione tradizionale, ne esiste un'altra che sta acquisendo sempre più rilevanza, visti i recenti sviluppi dell'Intelligenza Artificiale e dei suoi possibili impieghi. L'applicazione di questa firma digitale, infatti, può costituire un valido strumento, per aiutare l'utente a distinguere un'immagine reale da una artificiale, potendo anche essere invisibile e inserita senza che l'immagine subisca alcun tipo di cambiamento percepibile, mantenendo inalterata la loro qualità e la fruibilità venga alterata[58].

La spinta americana sulla tecnologia *watermarking* evidenziava dunque un approccio molto diverso al problema da parte degli Stati Uniti, rispetto a quanto non stesse accadendo nell'Unione europea con l'approvazione del *AI Act*. Esso, infatti, appariva orientato a favorire le imprese e lo sviluppo dei sistemi di IA, che hanno da sempre un impatto politico decisivo, godendo di un forte sostegno da parte del Governo, con il quale sono interconnesse.

Tale approccio, da un lato, incentivava la frammentazione legislativa, spingendo i singoli Stati ad approvare per sé leggi spesso molto diverse, dall'altro la scelta di non dotarsi di una stringente regolazione del fenomeno era volta ad evitare di porre limiti allo sviluppo della tecnologia IA e alla competitività delle imprese americane impegnate nel settore. Analogamente succede per la manifestazione del pensiero sulle piattaforme digitale che negli Stati Uniti non è regolata in modo uniforme a livello federale[59], manifestandosi anche in quella sede un approccio maggiormente *business friendly* in favore delle aziende di quanto non accada, come vedremo, in Europa.

Le linee guida fissate dal citato ordine esecutivo sono state oggi abrogate dal presidente Trump il 23 gennaio 2025 con un nuovo ordine esecutivo[60], che evidentemente manifesta la volontà della nuova amministrazione di sviluppare un approccio del tutto libero e non regolamentato della evoluzione di tale tecnologia.

Ciò, da un lato, certamente accelera l'innovazione e attrae investimenti nel settore da parte di quelle aziende che si occupano di sviluppare l'IA; dall'altro, però, numerosi sono i rischi che tale politica determina quanto ai diritti degli utenti, soprattutto in relazione alla possibile proliferazione di contenuti falsi e disinformativi e non facilmente individuabili come tali, cui si aggiungono quelli relativi alle possibili discriminazioni algoritmiche e alle violazioni della privacy, che potrebbero ingenerare la sfiducia del consumatore verso questa tecnologia.

Per questo la sfida, cui la nuova deregolamentazione statunitense dovrà far fronte, sarà quella di trovare un equilibrio fra il sostegno allo sviluppo tecnologico, la minimizzazione dei rischi e la tutela dei diritti degli utenti.

Deve comunque segnalarsi che, dopo il caso occorso alla popstar Taylor Swift e in considerazione del fatto che cominciano a moltiplicarsi le azioni giudiziarie delle vittime del fenomeno soprattutto negli Stati che non si sono ancora dotati di una pro-

[58] Nonostante la sua utilità, un *report* intitolato *Detecting AI fingerprints: A guide to watermarking and beyond* del centro di ricerca americano *Brooking Institute* del gennaio 2024 sull'uso del *watermarking* in relazione all'IA sottolinea però come tale strumento di firma non sia privo di margini di errore, perché una volta inserito non è difficile da rimuovere per chi possieda le competenze per farlo.

[59] Anche qui si permetta di rinviare a M.E. Bucalo, *I volti della libertà di manifestazione del pensiero,* cit.

[60] Executive Order 14179 of January 23, 2025 *Removing Barriers to American Leadership in Artificial Intelligence*.

pria legislazione[61], il 30 gennaio 2024 al Congresso statunitense è stata presentata una proposta di legge detta *Defiance Act* (*Disrupt Explicit Forged Images And Non-Consensual Edits)*[62], che mira a fornire una disciplina unitaria del fenomeno, consentendo alle vittime di *deepfake* di richiedere un risarcimento a coloro che producono o possiedono intenzionalmente tali immagini con l'intento di propagarle.

Il modello di rimedi predisposti in Europa appare *ictu oculi* molto diverso da quello statunitense[63].

Volendo analizzare il modello italiano, esso, a differenza di quello statunitense, predispone già nel codice penale fattispecie generiche applicabili ai casi di diffusione di *deepfake* (anche pornografici): si tratta della sostituzione di persona di cui all'art. 494 c.p. e della frode informatica *ex* art. 640 *ter*, c. 3, c.p.

Quanto alla prima disposizione, che sanziona con la reclusione fino a un anno «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», la giurisprudenza della Corte di cassazione si è premurata di ricomprendervi anche condotte poste in essere mediante le nuove tecnologie, ed in particolare la creazione di un *account* sui *social network* utilizzando abusivamente l'immagine di una persona inconsapevole, associata ad un *nickname* di fantasia ed a caratteristiche personali negative[64].

La frode informatica sanziona, invece, con la reclusione da sei mesi a tre anni la condotta di colui il quale procura a sé o ad altri un ingiusto profitto con altrui danno «alterando in qualsiasi modo il funzionamento di un sistema informatico o telematico o intervenendo senza diritto con qualsiasi modalità su dati, informazioni o programmi contenuti in un sistema informatico o telematico». Ad essa l'art. 9, c. 1, lett. a), del d.l. 14 agosto 2013, n. 93, convertito dalla l. 15 ottobre 2013, n. 119, ha poi aggiunto un c. 3, che stabilisce una pena maggiore, per chiunque ponga in essere la condotta con furto o indebito utilizzo dell'identità digitale in danno di uno o più soggetti. Disciplinando dunque il furto d'identità digitale, la norma citata è quella che nell'ordinamento italiano più si avvicina alla qualificazione del *deepfake* come illecito.

A tali fattispecie generiche, si somma il reato di "Diffusione illecita di immagini o video sessualmente espliciti", comunemente detto *revenge porn*[65], di cui all'art. 612 ter c.p., introdotto

[61] Nel New Jersey, che non ha una propria legislazione in tema, per esempio, i casi di *deepfake* porno sono proliferati nei licei, tanto che un adolescente ha citato in giudizio un compagno di classe per aver condiviso falsi nudi realizzati con l'IA. Su questi casi si veda l'articolo pubblicato su CBS news del 2 novembre 2023.

[62] S. 3696 – 118th Congress (2023-2024). Il testo risulta fermo presso il Senato dalla fine di luglio 2024.

[63] Le ragioni profonde che determinano tali diversità negli approcci fra le due sponde dell'Atlantico, sia a livello giurisprudenziale sia a livello regolatorio, sono note e possono essere brevemente riassunte nella tutela costituzionale, che potremmo definire "rafforzata", che il primo emendamento della Costituzione degli Stati Uniti assicura alla libertà di manifestazione del pensiero, a fronte del quale il vecchio continente oppone una visione che bilancia caso per caso la medesima libertà con gli altri diritti che con essa eventualmente entrassero in conflitto.

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

[65] Sulla correttezza della intitolazione informale del suddetto reato quale "*revenge porn*" è in corso il dibattito fra chi, soprattutto in ambienti giornalistici, lo ritiene un valido abbreviativo per individuare

dalla l. n. 69 del 2019, proprio per cercare di fronteggiare il fenomeno della diffusione di immagini e video sessuali senza il consenso della persona in questione, a seguito di riproduzione o sottrazione fraudolenta, e che prevede a tal fine la reclusione da uno a sei anni e la multa da 5.000 a 15.00 euro.

La previsione specifica di tale reato è presente solo in pochissimi ordinamenti al mondo ed in particolare, oltre che in Italia anche in Australia, Canada, Filippine, Giappone, Israele, Malta, Regno Unito e in alcuni Stati degli Stati Uniti.

Il tratto distintivo della norma risiede nella molteplicità dei destinatari della stessa. Infatti, l'articolo non mira a sanzionare soltanto coloro i quali abbiano diffuso i contenuti dopo averli personalmente realizzati o sottratti, ma anche tutti quei soggetti che, pur non avendo materialmente contribuito alla produzione o al furto degli stessi, abbiano contribuito a farli circolare dopo averli semplicemente ricevuti.

Ovviamente, condizione essenziale per l'applicazione è l'assenza del consenso della persona rappresentata, da intendersi cioè come manifestazione di volontà positiva ed esplicita. Infine, ai commi 3 e 4, l'articolo 612 ter c.p. prevede degli inasprimenti delle sanzioni nel caso di integrazione della fattispecie in situazioni particolarmente gravi, consistenti cioè in una relazione affettiva tra il reo e la vittima, o una condizione di inferiorità fisica o psichica della stessa.

La particolare attenzione rivolta in Italia a fattispecie come quelle sopra trattate è evidenziata anche dalla adozione di un altro strumento di "natura stragiudiziale", che imprime una notevole accelerazione alla loro risoluzione, evitando o comunque prevenendo la lunghezza dei tempi della giustizia. Si tratta dell'accordo concluso nel 2021 fra *Facebook* e Garante per la Privacy, in tema di pornografia *on line* non consensuale, che consente alle possibili vittime di iniziare un procedimento di segnalazione urgente, con un accesso privilegiato dal sito istituzionale del Garante stesso[66].

A tal fine, l'utente deve segnalare i *link* e i *post* da rimuovere e allegare le foto. Il sistema di collegamento fra Garante e *Facebook* consente quindi a quest'ultimo di individuare le foto allegate con velocità, cifrarle, rendendole così irriconoscibili e distruggerle, nonché attraverso una tecnologia di comparazione bloccarne le possibili condivisioni e nuove pubblicazioni anche su altre piattaforme

5. I *deepfake* in quanto "successori" delle *fake news*

La oggettiva gravità dei fatti trattati nel precedente paragrafo, rimarcata dalla potenza del mezzo informatico capace di diffonderne la lesività a livello globale, pone in luce anche il ruolo svolto dalle piattaforme *online*, veri signori delle porte di accesso alla rete (*Gatekeepers*), i quali forniscono gli strumenti e i servizi, che agevolano la diffusione di *deepfake*.

A fronte di tutto ciò, bisogna però rilevare la crescente pressione dell'opinione pubbli-

immediatamente la fattispecie e chi invece sostiene che essa, dal punto di vista giuridico e sociale, sia fuorviante e possa prestarsi ad interpretazioni pericolose e a derive che potrebbero in qualche modo tentare di giustificare questi atti. La vendetta, per quanto possa essere in astratto discutibile, presuppone infatti il fatto che esista alla base un torto o uno sgarbo per cui vendicarsi, cosa che in effetti non è. Si veda in tema *ex multis* G.M. Caletti, *"Revenge porn" e tutela penale*, in *Diritto Penale Contemporaneo Rivista Trimestrale,* 3, 2018, 63 ss.; F. Florio, *Non chiamatelo "revenge porn",* Milano, 2022; E. Strighi, *Revenge porn: lettura di genere di una fattispecie (incompresa),* in *Sociologia del diritto,* 1, 2021, 33 ss.

[66] Si veda la apposita pagina dedicata sul sito del Garante della protezione dei dati personali.

ca, volta a costringere le società di informazione a limitare la diffusione di tali contenuti e ad assumersene le relative responsabilità. A tale pressione, in prima battuta, pare aver dato risposta la Corte di Giustizia dell'Unione europea, che, in tema di *fake news*, ha oramai assunto come principio generale quello della responsabilità del *provider*.

A ben vedere tale principio, oramai consolidato nella giurisprudenza unionale, risulta applicabile anche ai *deepfake*, che a buon diritto possono essere considerati i "successori" tecnologicamente evoluti delle *fake news*.

In anni recenti, infatti, il fenomeno della disinformazione si è sviluppato in modo significativo non solo a causa dell'espansione di *Internet* e delle strategie di comunicazione digitale, ma anche grazie allo sviluppo e alla diffusione dei sistemi di IA[67].

Fra *fake news* e *deepfake* esistono infatti molti punti di contatto e molte analogie, a cominciare per esempio dalle difficoltà definitorie, che riguardano le prime e che sono analoghe per i secondi[68].

Vero è che le prime trovano il loro archetipo nelle bufale di età digitale, che esistono da sempre[69], ma è altrettanto vero che la cifra distintiva fra queste ultime e le attuali *fake news* è che nell'ecosistema digitale la diffusione di questi fenomeni assume una velocità mai avuta in età analogica. Per questo non si può che concordare con chi sostiene che «il maggior discrimine fra le *fake news* del passato e quelle attuali sta proprio nella "cinetica" con la quale si propagano al punto da costituire un elemento ontologico dirimente»[70].

Analogamente ciò può dirsi mettendo a raffronto *fake news* e *deepfake,* poiché analoghe sono le ragioni che determinano tale velocità di diffusione e che possono brevemente riassumersi nell'assetto oligopolistico degli *Internet Service Providers*, nel decentramento

[67] Sul punto O. Pollicino – P. Dunn, *Disinformazione e intelligenza artificiale nell'anno delle global election: rischi (ed opportunità),* in *federalismi.it,* 12, 2024, ix, i quali si soffermano sulla capacità dei sistemi di IA di produrre disinformazione, facendo riferimento, per il primo profilo, all'emersione di innumerevoli sistemi, legati in particolare alla cosiddetta "IA generativa", ai modelli fondativi e ai *large language models* (LLM), capaci di creare immagini, video e testi sintetici altamente realistici.

[68] Laddove si volesse tentare di definire il fenomeno delle *fake news*, secondo H. Allcott – M. Gentzkkow, *Social Media and Fake news in the 2016 elections*, in *Journal of Economic Perspectives,* 2, 2017, 211 ss., potrebbero intendersi quelle notizie che sono intenzionalmente e verificabilmente false e potrebbero trarre in inganno chi vi si imbatte. Verrebbero così escluse tutte le informazioni che, pur vicine al concetto di *fake*, non lo sono, collocandosi nel territorio del pensiero manifestato liberamente, ma si inserirebbero a buon diritto tutte le notizie false, costruite ad arte da gruppi di potere con l'obiettivo di modificare l'agenda pubblica, manipolando l'informazione e la formazione dell'opinione pubblica anche tramite tecnologie sofisticate, nonché tutte le notizie false che ledono interessi individuali o collettivi.

[69] Si ricordi per esempio che H. Von Kleist scrisse nel 1809 *Manuale dell'informazione francese*, una satira in risposta alla propaganda di guerra di Napoleone, nel quale descriveva come i giornali francesi montavano e diffondevano bufale, finalizzate solo ad esaltare l'imperatore. Leggendo R. Dale, *'Napoleon is Dead': Lord Cochrane and the Great Stock Exchange Scandal*, Londra, 2007 si scopre che la bufala della morte di Napoleone fu in grado di dirottare grandi capitali in borsa e fare decollare i titoli di Stato. Anche negli Stati Uniti le bufale esistevano già nel XIX sec., una fake news divenuta famosa fu quella che è stata chiamata la "Great Moon Hoax". Nel 1835 il New York Times pubblicava una serie di articoli che parlavano della scoperta della vita sulla Luna, falsamente attribuite a sir John Herschel, il più noto astronomo del tempo. L'idea dell'anonimo autore degli articoli, poi rivelatosi Richard Adams Locke, era presumibilmente solo quella di fare satira, che tuttavia fu creduta vera, sebbene le notizie suscitarono notevole scalpore e furono tradotti in diverse lingue nel mondo.

[70] Per questo non si può che concordare con A. Sciortino, *Fake news e post-verità nella società dell'algoritmo*, in *dirittifondamentali.it,* 2, 2021, 426.

della produzione della informazione, che non è più soggetta ai controlli legalmente imposti agli editori[71] e nella progressiva perdita di fiducia in tv e carta stampata come strumenti di informazione, utili per il confronto fra opinioni diverse, ma oramai tacciate di parzialità e commistione con gli apparati pubblici o con i grandi poteri economici[72].
Epperò i *deepfake* possono considerarsi una evoluzione, in senso peggiorativo, delle *fake news,* considerato che essi associano alla "cinetica" della loro propagazione attraverso *Internet,* anche un "aggravamento" della falsità in quanto l'immagine (o il video) ha un impatto certamente maggiore sull'utente della Rete, di quanto non lo abbiano le opinioni o gli scritti[73]. A tutto ciò, che evidentemente accelera ulteriormente la diffusione, si somma la preoccupante considerazione che l'uso della IA rende ancor più difficilmente distinguibile ciò che è veritiero visivamente da ciò che non lo è.
Ancora, deve segnalarsi come per le *fake news*, la peculiare organizzazione dei contenuti in *Internet*, che si fonda su sistemi di raccomandazione (*recommender* o *recommendation systems*) i quali, partendo dai dati e dalle informazioni raccolte sulle preferenze del singolo utente, sono in grado di predirne l'indice di gradimento con riferimento a nuovi contenuti ed elementi[74].
Tale sistema di diffusione delle informazioni *online* presenta certamente il pregio di poter essere usato proattivamente al fine di garantire una diffusione dei contenuti quanto più plurale possibile, ma anche un effetto collaterale. Poiché infatti a muovere l'*engagement* degli utenti da parte dei gestori delle piattaforme è sostanzialmente il loro interesse economico e poiché i contenuti altamente divisivi e polarizzanti, come si diceva, tendono ad attrarre maggiormente l'attenzione del pubblico, esiste il concreto rischio che «l'algoritmo, pur di suscitare l'interesse degli utenti, sia disincentivato a ridurre la diffusione di disinformazione»[75].
I rischi sono poi ulteriormente accentuati dal meccanismo di filtraggio che viene operato dalle piattaforme e dai *social network* nella pubblicazione e nella diffusione dei contenuti. La semplice navigazione degli utenti, infatti, determina da parte dei gestori la raccolta dei loro dati, profilati poi per mezzo di operazioni algoritmiche. In tal modo essi sono in grado, per così dire, di "predire" in futuro i comportamenti individuali e collettivi degli utenti e conseguentemente di influenzarne le decisioni.

---

[71] G.E. Vigevani, *L'informazione e i suoi limiti: il diritto di cronaca*, in G.E. Vigevani – O. Pollicino – C. Melzi D'Eril – M. Cuniberti – M. Bassini, *Diritto dell'informazione e dei media*, Torino, 2019, 25 ss.

[72] Si veda in tema G. Pitruzzella, *Libertà di manifestazione del pensiero nell'era di Internet*, in G. Pitruzzella – O. Pollicino – S. Quintarelli, *Parole e potere. Libertà di espressione*, hate speech *e* fake news, Milano, 2017, 70 ss.

[73] O. Pollicino – P. Dunn, *Disinformazione e intelligenza artificiale,* cit., ix-x. e xii, affermano che «quanto alla "disseminazione" di *deepfake online* va rilevato che, da un lato l'IA viene utilizzata dai produttori o da soggetti comunque interessati alla diffusione di materiali falsi in rete al fine precipuo di aumentarne l'impatto» e ritengono particolarmente diffusa «la pratica di ricorrere a *social bot*, ovverosia ad *account* falsi gestiti in modo automatico o semiautomatico (in quest'ultimo caso si parla di "*cyborg*", cioè di profili gestiti in parte da persone vere e in parte dall'IA) con il preciso obiettivo di contribuire alla diffusione di materiali "inquinanti"».

[74] Secondo O. Pollicino – P. Dunn, *Disinformazione e intelligenza artificiale,* cit.,xiii, tali sistemi svolgono un ruolo centrale nella diffusione delle informazioni e, pertanto, nella formazione della stessa coscienza pubblica, avendo la capacità di influenzare e strutturare le stesse preferenze degli utenti e di guidarne le scelte sia a livello individuale sia a livello sociale e collettivo.

[75] Ivi, xiii.

Da un lato le attività della vita degli utenti sono registrate ogni volta che si connettono a *Internet*, perché i dati che costoro lasciano nella Rete restituiscono un loro "profilo" utile poi alle piattaforme per indirizzarli nelle ricerche successive. Dall'altro però, il procedimento seguito per aggregare gli "indizi" relativi alla loro personalità (i dati appunto), che "prediranno" i comportamenti futuri, è del tutto oscuro (*black box society*)[76].
La profilazione dei contenuti determina come immediata conseguenza anche la personalizzazione delle informazioni, operata per ciascun utente dagli stessi motori di ricerca e *social network*, i quali rendono disponibili le informazioni ricercate secondo un *ranking,* il cui ordine è stabilito per mezzo delle stesse operazioni algoritmiche che hanno contribuito alla sua profilazione.
Così, se le notizie, i contenuti o i risultati delle ricerche sono profilati in modo da "predire" la personalità dell'utente, allora quest'ultimo visualizzerà sempre e solo quelle informazioni conformi al suo pensiero e si convincerà che nella realtà esistono soltanto persone, che esprimono le sue stesse idee. La possibilità di accedere a fonti di informazione o pareri che discordino da queste convinzioni sarà evidentemente molto limitata, come di conseguenza lo sarà anche la formazione di una opinione genuinamente consapevole.
Si tratta di un vero e proprio processo di «inscatolamento del nostro mondo informativo»[77] e della conseguente costruzione di mondi che sono solo a immagine e somiglianza di colui che naviga in Rete.
Volendo usare una nota metafora, ciascuno sembra chiuso in una *filter bubble*[78] (o anche *echo chamber*), che però «amplifica le divisioni e le polarizzazioni, tradendo una delle missioni più profonde della libertà di espressione e del confronto: la tolleranza reciproca fra opinioni differenti»[79].

6. I rimedi predisposti dalla giurisprudenza dell'Unione europea

Analizzando adesso i rimedi predisposti a fronte della diffusione *online* dei *deepfake* e rivolgendo l'attenzione in prima battuta alla giurisprudenza dell'Unione europea, è noto che il modello decisorio della Corte di Giustizia risente, oltre che della assenza di una disciplina specifica (sanata solo nel 2023 con l'approvazione del *Digital Service Package*), anche della genesi del suo ordinamento, improntata ad una visione prettamen-

[76] Sul punto si veda F. Pasquale, *The Black Box Society. The Secret Algorithms That Control Money and Information*, Harvard, 2015, 9 ss. Sulla "doppia natura" conoscitiva e predittiva dell'algoritmo, nonché sulla necessità di regole che ne assicurino la trasparenza e il rispetto dei diritti degli utenti della Rete, prima fra tutte la privacy si veda anche L. Torchia, *Stato digitale*, cit., 24-25. Analogamente, A. Koltay, *New Media and Freedom of Expression: Rethinking the Constitutional Foundation of Public Sphere*, Oxford, 2021, 86.

[77] M. Calise – F. Musella, *Il Principe digitale*, Roma-Bari, 2019, 11.

[78] E. Parisier, *Filter Bubble: How the New Personalized Web Is Changing What We Read and how We Think*, New York, 2011.

[79] C. Bologna, *Libera di espressione e "riservatezza" nella Rete? Alcune osservazioni sul mercato delle idee nell'agorà digitale*, in *Rivista del Gruppo di Pisa*, fascicolo speciale, 3, 2021, 71.

te mercantilistica dei diritti fondamentali e della loro tutela[80]. Ciò ha determinato una giurisprudenza, che si occupa del tema in modo, per così dire, indiretto, improntata al bilanciamento fra il diritto di espressione e gli altri diritti di estrazione economica, spesso ritenuti prevalenti o comunque dotati di una tutela rafforzata.

Fra questi il primo che va certamente menzionato è il diritto alla *privacy*[81], la cui tutela nell'ambito della società dell'informazione aveva stimolato il ruolo di supplenza della Corte di Giustizia, in considerazione della perdurante inerzia del legislatore europeo nell'opera di necessario aggiornamento della disciplina previgente.

L'archetipo di questo filone giurisprudenziale è la sentenza *Lindqvist*[82], con la quale per la prima volta la Corte di Giustizia definì l'ambito di applicazione della Direttiva 95/46 in tema di privacy (previgente rispetto al GDPR) nel quadro del nuovo scenario tecnologico, nonché indagare il rapporto fra protezione dei dati personali e libertà di espressione[83].

Volendo assicurare alla tutela dei dati personali una protezione a tutto tondo, la Corte affermava che il bilanciamento fra tutela della *privacy* e le altre libertà, andava effettuato caso per caso e che era compito delle autorità nazionali e dei giudici non solo interpretare il diritto nazionale in conformità con la direttiva 65/46, ma anche «provvedere a non fondarsi su un'interpretazione di quest'ultima che entri in conflitto con i diritti fondamentali tutelati nell'ordinamento giuridico comunitario»[84].

L'altra pronuncia "archetipo" [85] da menzionare necessariamente in questa sede è quella che determinò l'annullamento della direttiva 2006/24/CE (c.d. "*Data Retention*") relativa alla conservazione dei dati di traffico per violazione degli artt. 7 e 8 della Carta europea dei diritti fondamentali[86].

---

[80] M. Bassini, *Internet e libertà di espressione. Prospettive costituzionali nazionali e sovranazionali*, Roma, 2019, 319

[81] Il quale oggi gode, oltre che di una consolidata giurisprudenza, anche di una forte accelerazione impressagli dal GDPR del 2016, che ha sostituito la oramai risalente disciplina della direttiva 95/46/CE.

[82] CGUE, C-101/01, *Lindqvist* (2003). Sul punto si vedano i commenti di A. Palmieri – R. Pardolesi, *Il codice in materia di dati personali e l'intangibilità della* privacy *comunitaria*, in *Foro italiano*, 2, 2004, 59 ss.; T.M. Ubertazzi*, Il caso Lindqvist: i limiti della privacy*, in *Danno e responsabilità*, 24, 2004, 382 ss.

[83] Il caso riguardava la pubblicazione fatta da una cittadina svedese sul suo sito internet dei dati di alcune persone che lavoravano con lei come volontari in una parrocchia, senza averne ricevuto il consenso. Dopo la condanna in primo grado, la Corte di Appello aveva deciso di sollevare avanti la Corte di Giustizia sette questioni pregiudiziali riguardanti l'ambito di applicazione della direttiva sulla *privacy*, una delle quali espressamente chiedeva se la disciplina della direttiva potesse ritenersi compatibile con i principi generali in materia di libertà di espressione.

[84] CGUE, C-101/01, *Lindqvist* (2003), § 87.

[85] La pronuncia è infatti il fondamento della successiva giurisprudenza unionale in tema di tutela dei dati personali e da ultimo è stata recentemente ripresa *ex multis* da CGUE, C-746/18, *H.K. c. Prokuratuur* (2021) e CGUE, C-140/20, *G.D. contro The Commissioner of the Garda Síochána e a.* (2022).

[86] CGUE, C-293/12 e C-594/12, *Digital Rights Ireland Ltd e Kärntner Landesregierung* (2014). La sentenza è stata seguita dalla pronuncia sul caso analogo CGUE, C-203/15 e C-698/15, *Tele2Sverige* (2016). Moltissimi i commenti a questa sentenza. *Ex multis* R. Flor, *Dalla "data retention" al diritto all'oblio. Dalle paure orwelliane alla recente giurisprudenza della Corte di Giustizia. Quali effetti per il sistema di giustizia penale e quali prospettive* "de jure condendo"?, in *Il diritto dell'informazione e dell'informatica*, 4-5, 2014, 775 ss.; L. Trucco, "Data retention*": la Corte di Giustizia si appella alla Carta UE dei diritti fondamentali*, in *Giurisprudenza italiana*, 8-9, 2014, 1850 ss.; G. Tiberi*, La Corte di Giustizia sulla conservazione dei dati: la protezione dei diritti fondamentali nel "dopo-Lisbona"*, in *Quaderni costituzionali*, 3, 2014, 719 ss.; A. Arena, *La Corte di Giustizia*

La direttiva richiedeva agli Stati membri di conservare in modo generalizzato i dati di traffico connessi a qualsiasi mezzo comunicativo di tutti gli utenti, senza alcuna distinzione, limitazione o eccezione rispetto all'obiettivo di contrasto alla criminalità. Peraltro, l'archiviazione aveva ad oggetto dati di persone che, nemmeno indirettamente, si trovavano nella situazione di dare adito a procedimenti penali o di essere collegate, anche solo in modo remoto, a reati gravi.

La sentenza si rivelava di importanza fondamentale per la rilevanza che in essa assume il canone della proporzionalità delle misure limitative degli indicati diritti e che, nel caso di specie, non risultava rispettato, a causa della raccolta e conservazione generalizzata dei dati di tutti i soggetti, disposta dalla direttiva e che dunque risultava eccessivamente intrusiva rispetto ai diritti sanciti dalla Carta di Nizza.

È però la celebre sentenza *Google Spain*[87] a dimostrare in modo chiaro l'influenza della tecnologia digitale nel rapporto fra libertà di espressione e diritto alla protezione dei dati personali e ad individuare un primo possibile argine alla progressiva espansione del potere delle piattaforme nel cyberspazio, imponendo la loro responsabilità nella tutela del diritto alla privacy e degli altri diritti sanciti nella Carta di Nizza[88].

sulla conservazione dei dati: quali conseguenze per le misure nazionali di recepimento, in *Quaderni costituzionali*, 3, 2014, 7872 ss.

[87] CGUE, C-131/12, *Google Spain* (2014). Moltissimi i commenti alla sentenza, se ne citano qui solo alcuni *ex plurimis* F. Pizzetti, *La decisione della Corte di Giustizia sul caso* Google Spain: *più problemi che soluzioni, in federalismi.it*, 12, 2014; T.E. Frosini*, Diritto all'oblio e Internet*, ivi; tutti i contributi nel volume G. Resta – V. Zeno-Zencovich (a cura di*), Il diritto all'oblio dopo la sentenza* Google Spain, Roma, 2015 fra i quali O. Pollicino, *Un* digital right to privacy *preso (troppo) sul serio*, ivi, 17 ss.; G. Finocchiaro*, Il diritto all'oblio nel quadro dei diritti della personalità,* ivi*,* 29 ss.; T.E. Frosini, *Google e il diritto all'oblio preso sul serio,* ivi*,* 1 ss.; O. Pollicino – M. Bassini*, Reconciling right to be forgotten and freedom of information in the digital age. Past and future of personal data protection in the European Union*, in *DPCE*, 2, 2014, 641 ss.; M. Bassini, Google *davanti alla Corte di Giustizia: il diritto all'oblio*, in *Quaderni costituzionali*, 3, 2014, 730 ss.; M.C. D'Arienzo, *I nuovi scenari della tutela della* privacy *nell'era della digitalizzazione alla luce delle recenti pronunce sul diritto all'oblio*, in *federalismi.it*, 2, 2015; L. De Grazia, *La libertà di stampa e il diritto all'oblio nei casi di diffusione di articoli attraverso Internet: argomenti comparativi*, in *Rivista AIC*, 4, 2013; S. Leucci, *Diritto all'oblio, verità, design tecnologico: una prospettiva di ricerca,* in questa *Rivista*, 1, 2017, 116 ss.; A. Palmieri – R. Pardolesi, *Dal diritto all'oblio all'occultamento in rete: traversie dell'informazione ai tempi di* Google, in *Nuovi quaderni del foro italiano*, 1, 2014; R. Pastena, Internet *e* privacy: *una relazione complicata (A margine della sentenza della Corte di Giustizia del 13 maggio 2014)*, in *Osservatorio Aic*, 2, 2014; R.C. Post*, Data Privacy and Dignitary Privacy: Google Spain and the right to be forgotten, and the construction of public sphere*, in *Duke*, 2018, 981 ss.; E. Frantziou, *Further Developments in the Right to be Forgotten: The European Court of Justice's Judgment in Case C-131/12, Google Spain, SL, Google Inc v Agencia Espanola de Proteccion de Datos*, in *Human Rights Law Review*, 14, 2014, 76 ss. ; O. Spataro*, Il diritto all'oblio tra definizione sostanziale e rimedi di tutela. Riflessioni alla luce della giurisprudenza più recente della Corte di Cassazione e della Corte di Giustizia dell'Unione Europea in materia di deindicizzazione*, in *Diritto costituzionale*, 1, 2023, 133 ss.

[88] Per comprendere appieno la portata della pronuncia è necessario partire dalla fattispecie che la ha originata e che vedeva opposti un cittadino spagnolo e Google avanti l'Autorità spagnola per la protezione dei dati personali (*AEPD*). In particolare, il ricorrente sig. Costeja Gonzalez aveva chiesto che fosse ordinato al motore di ricerca di cancellare dalla pagina della ricerca alcuni link, che rimandavano ad una vicenda giudiziaria, in materia di pignoramenti immobiliari per debiti previdenziali, che lo aveva riguardato molti anni prima e che si era completamente conclusa. Analogamente lo stesso ricorrente chiedeva che fosse ordinato al giornale, che per ordine dell'autorità giudiziaria doveva dare massima pubblicità alla vicenda, di cancellare permanentemente le relative pagine o che ivi fossero occultati i suoi dati personali. Il Garante spagnolo aveva accolto il ricorso nella parte in cui chiedeva a *Google* di deindicizzare i link che riportavano al ricorrente, ma non quanto alla richiesta rivolta al quotidiano. Avverso questa decisione *Google* ricorreva avanti *l'Audiencia Nacional*, che sollevava alcune questioni pregiudiziali.

Le questioni sollevate dal remittente avevano ad oggetto, oltre che l'ambito territoriale di applicazione della allora vigente direttiva sulla *privacy*, l'identificazione della attività della piattaforma come trattamento dei dati personali, la portata dei diritti di cancellazione e opposizione al trattamento dei dati previsti dall'art. 12 lett. b) della direttiva stessa[89] e l'obbligo della stessa di rispettare i diritti umani determinandone la responsabilità.

Dopo aver affermato che le disposizioni in questione si applicano anche al soggetto o alla società che, pur non avendo sede nell'Unione, vi siano comunque stabiliti, determinando così la vigenza extraterritoriale e potenzialmente globale del diritto dell'Unione[90], la pronuncia si dedicava alla individuazione del ruolo svolto dalle piattaforme in ambiente digitale, ne identifica il potere e conseguentemente, vista l'estensione dello stesso, la piena responsabilità.

A tal fine la Corte identifica preliminarmente nei motori di ricerca i titolari del trattamento dei dati, perché li diffondono in modo globale, rendendoli accessibili a tutti gli utenti di *Internet*, li aggregano e li organizzano, in modo che l'utente che fa la ricerca possa averne una visione complessiva, e infine li conservano nei propri *server*[91].

Questa attività, che incide «in modo significativo […] sui diritti fondamentali di cui agli artt. 7 e 8 della Carta di Nizza», determina in capo al motore di ricerca l'obbligo di svolgerla in modo che le garanzie previste dalla direttiva 95/46 possano sviluppare pienamente i loro effetti[92].

Conseguenzialmente nel caso di mancato rispetto delle norme della direttiva stessa, ne consegue che il gestore del motore di ricerca ha l'obbligo di «sopprimere, dall'elenco di risultati [...] i link verso pagine web pubblicate da terzi e contenenti informazioni relative all'utente, anche nel caso in cui tale nome o tali informazioni non vengano previamente o simultaneamente cancellati dalle pagine» originarie, ancorché tali pubblicazioni siano lecite[93].

A tale responsabilità del *provider*, i giudici di Lussemburgo riconoscono allo stesso l'ulteriore obbligo di valutare in ordine alla sussistenza dei presupposti per l'esercizio del diritto e alla compatibilità con la libertà di informazione, che devono valutare caso per caso rilevando l'eventuale sussistenza di ragioni particolari, che determinerebbero il prevalere dell'interesse pubblico ad avere accesso all'informazione o meno[94].

La pronuncia però, come rilevato dalla dottrina maggioritaria, determinava il rischio di responsabilizzare eccessivamente il *provider,* che veniva indotto a cancellare il più

[89] Il quale disponeva che «Gli Stati membri garantiscono a qualsiasi persona interessata il diritto di ottenere dal responsabile del trattamento: [...] a seconda dei casi, la rettifica, la cancellazione o il congelamento dei dati il cui trattamento non è conforme alle disposizioni della presente direttiva, in particolare a causa del carattere incompleto o inesatto dei dati».

[90] Sul punto si veda G. Sartor – M. Viola De Azevedo Cunha, *Il caso Google e i rapporti regolatori USA/EU*, in G. Resta – V. Zeno-Zencovich (a cura di), *Il diritto all'oblio*, cit., 99 ss.

[91] CGUE, C-131/12, *Google Spain* (2014), §§ 36-38 e 83.

[92] Ivi, § 38.

[93] Ivi, §§ 70, 76 e 87-88.

[94] Ivi, § 97. Secondo M. Bassini, *Internet e libertà di espressione*, cit., 328, la Corte così sembra volere costruire intorno ai provider una responsabilità piena e a trecentosessanta gradi, rendendoli arbitri del conflitto fra i due diritti e affidando loro compiti difficilmente gestibili.

possibile le notizie, assegnandogli il compito di valutare ciò che può essere pubblicato e indicizzato perché di pubblico interesse e ciò che invece non lo è, facendo emergere il contrasto lampante con il principio costituzionale della riserva di giurisdizione nei casi di possibili restrizioni dei diritti fondamentali «che caratterizza il nucleo duro di qualsiasi ordinamento che si fondi sulla *rule of law*»[95].

D'altro canto, dalla sentenza citata anche la posizione degli utenti sembrava tutelata solo a metà, poiché ad essi era pienamente riconosciuto il diritto alla deindicizzazione, ma al contrario risultava gravemente leso il loro diritto all'informazione[96].

Alle ortodossie dimostrate dalla sentenza *Google Spain* sono stati posti dei correttivi con la sentenza *Google c. Commission Nationale de l'Informatique e des Libertes (CNIL – Autorità francese per la protezione dei dati personali)*[97], esito di un rinvio pregiudiziale promosso dal *Conseil d'Etat,* davanti al quale pendeva un ricorso avverso una decisione dell'Autorità del 2015, che aveva irrogato una sanzione nei confronti di *Google* a causa del diniego di operare una deindicizzazione su tutte le estensioni del nome a dominio del suo motore di ricerca e non già solo sulle declinazioni nazionali (.fr).

In questo caso la Corte di giustizia corregge parzialmente il tiro rispetto alla pronuncia *Google Spain*, affermando che l'obbligo di deindicizzazione gravasse sul *provider* e nella sua versione nazionale e nelle sue versioni relative ad altri Stati membri, ma certamente non potesse avere carattere globale, poiché, non esistendo un diritto alla deindicizzazione globale negli ordinamenti degli Stati, l'equilibrio fra diritto all'oblio e libertà di informazione degli utenti in Internet può «variare notevolmente nel mondo». Prova ne è il fatto che i singoli Stati membri possono prevedere discipline differenziate «in particolare per il trattamento a fini esclusivamente giornalistici o di espressione artistica o letteraria e per le esenzioni e le deroghe necessarie per conciliare tali diritti con la libertà di informazione»[98].

A questo ridimensionamento territoriale, non segue però alcun ridimensionamento degli obblighi e delle responsabilità del *provider*, al quale è comunque imposto di adottare le misure efficaci, affinché venga soddisfatto il bilanciamento fra i diritti fonda-

---

[95] O. Pollicino, *Un digital right to privacy preso (troppo) sul serio*, in G. Resta – V. Zeno-Zencovich (a cura di), *Il diritto all'oblio*, cit., 18.

[96] M. Bassini, *Internet e libertà di espressione*, cit., 339.

[97] CGUE, C-517/17, *Google v. CNIL* (2019). Fra i molti commenti F. Balducci Romano, *La Corte di giustizia "resetta" il diritto all'oblio*, in *federalismi.it*, 3, 2020, 3 ss.; A. Iannotti Della Valle, *Il diritto all'oblio "preso meno sul serio",* in *Rivista AIC*, 2, 2020, 495 ss.; G. Bellomo, *"Diritto all'oblio" e portata territoriale del "diritto alla deindicizzazione": la Corte ridisegna i confini applicativi*, in *DPCE online*, 4, 2019, 2987 ss.; G. Bevilacqua, *La dimensione territoriale dell'oblio in uno spazio globale e universale*, in *federalismi.it*, 23, 2019, 1 ss.; A. Correra, *La tutela dei dati personali e la portata territoriale dell'obbligo di deindicizzazione dei contenuti online*, in *Eurojus*, 3, 2020, 35 ss.; M. Orefice*, Diritto alla deindicizzazione: dimensione digitale e sovranità territoriale*, in *Rivista AIC*, 1, 2020, 653 ss.; F. Giovanella, *From the "right to delisting" to the "right to relisting"*, in questa *Rivista*, 2, 2022, 124 ss.; nonché e O. Spataro, *Il diritto all'oblio,* spec. 134 ss. Per la dottrina straniera, B. Martìn, *Google v. CNIL and the Right to Be Forgotten: A Judgment of Solomon*, in *Global Privacy Law Review*, 1, 2020, 61 ss.; M. Zalnieriute, Google LLC v. *Commission Nationale de l'Informatique et des Libertés (CNIL),* in *American Journal of International Law*, 2, 2020, 261 ss.; Y. Miadzvetskaya – G. Van Calster, *Google at the Kirchberg Dock. On Delisting Requests, and on the Territorial Reach of the EU's GDPR (C-136/17 GC and Others v CNIL, C□ 507/17 Google Inc v CNIL)*, in *European Data Protection Law Review*, 1, 2020, 143 ss.; O.J. Gstrein, *Right to be Forgotten: EU-ropean Data Imperialism, National Privilege, or Universal Human Right?,* in *Review of European Administrative Law (REALaw)*, 1, 2020, 125 ss.

[98] CGUE, C-517/17, *Google v. CNIL* (2019, § 67.

mentali e si impedisca agli utenti «di avere accesso ai link in questione a partire da una ricerca effettuata sulla base del nome»[99], conferendogli nuovamente una responsabilità *ultra vires*, che comporta il rischio della radicalizzazione della deindicizzazione e quello dell'attribuzione agli stessi di poteri gestori, che invece dovrebbero competere allo Stato e alla giurisdizione.

L'evoluzione e la razionalizzazione delle scelte giurisprudenziali europee in materia pare essersi compiuta solo con la sentenza *TU e RE c. Google*[100], che peraltro contiene assunti utili anche ai fini dell'individuazione delle responsabilità del *provider* in caso di *deepfake*.

Essa giunge all'esito di un rinvio pregiudiziale sollevato dalla Corte federale di Giustizia tedesca, avanti la quale avevano fatto ricorso due persone che avevano chiesto a *Google* di deindicizzare dall'elenco dei risultati di ricerca i *link* relativi ad alcuni articoli, apparsi su un sito che esponevano valutazioni critiche sul modello di investimento attuato da una società del loro gruppo, in quanto contenevano affermazioni inesatte e opinioni diffamatorie.

La questione posta, e che interessa particolarmente ai fini di questo studio, aveva riguardo al diritto alla deindicizzazione e chiedeva se la relativa richiesta potesse fondarsi sul fatto che le allegazioni contenute nel *link* fossero contestate nella loro veridicità dal ricorrente o se dovesse essere necessario un previo provvedimento giudiziario, che risolvesse la questione della attendibilità del contenuto visualizzato.

La Corte non pare discostarsi dalla sua precedente giurisprudenza e infatti ribadisce che l'attività del motore di ricerca deve essere qualificata "trattamento dei dati personali" ai sensi del GDPR e che dunque il gestore deve essere qualificato "responsabile del trattamento"[101].

Il dato in più, che costituisce la cifra distintiva della sentenza ora in analisi rispetto alle precedenti, è che la Corte sembra aggiungere nuovi criteri, a quelli già individuati, affinché il motore di ricerca proceda alla deindicizzazione. In particolare, essi sono individuabili nella falsità dell'informazione e nella sua palese inesattezza, le quali costituiscono «un elemento pertinente nell'ambito della valutazione delle condizioni di applicazione previste all'articolo 17, paragrafo 3, lettera a), del GDPR, al fine di valutare se il diritto all'informazione degli utenti di *Internet* e la libertà di espressione del fornitore di contenuti possano prevalere sui diritti del richiedente la deindicizzazione»[102].

L'idea che qui la Corte fa propria è quella che la falsità palese, l'inesattezza evidente non possono essere ricomprese nella libertà di informazione, perché «tale diritto, nella sua duplice valenza, attiva e passiva, se riferito ad un'informazione falsa, non può comunque essere posto sullo stesso piano dei diritti fondamentali al rispetto della vita

[99] Ivi, § 70.

[100] CGUE, C-460/20, *Tu. e Re. v. Google* (2022). Si veda G. Napoli, *Diritto alla deindicizzazione e notizie false: la Corte di giustizia precisa i confini tra oblio e libertà di espressione*, in questa *Rivista*, 1, 2023; F. Paolucci, I (Don't) remember my name: *il diritto all'oblio nella recente pronuncia C-460/2020 della Corte di Giustizia dell'Unione Europea,* in *Diritti Comparati*, 19 gennaio 2023.

[101] CGUE, C-460/20, *Tu. e Re. v. Google,* (2022), § 44.

[102] Ivi, § 64.

privata e alla tutela dei dati personali»[103]. Per far questo la prevalenza è assegnata alla dignità umana, in quanto valore fondamentale dell'Unione, rendendo di fatto inesistente il conflitto fra il diritto alla riservatezza e il diritto di espressione. In questo caso la richiesta di deindicizzazione potrà essere posta direttamente dall'interessato al motore di ricerca, poiché dal punto di vista probatorio, non è necessario che costui la accompagni con un previo provvedimento giurisdizionale (o amministrativo) che accerti l'inesattezza medesima o la falsità[104].

Si instaura così un rapporto diretto fra piattaforma e utente, cui segue anche una sorta di alleggerimento della posizione del *provider*, che evita anche i menzionati rischi di "deindicizzazione di massa", perché nel caso in cui il soggetto che ha presentato una siffatta richiesta, apportando elementi di prova pertinenti e sufficienti a dimostrare inesattezza o la falsità delle informazioni, il gestore del motore di ricerca sarà tenuto ad accogliere detta richiesta di deindicizzazione. Solo nel caso in cui tale inesattezza non risulti manifesta, avanti il rifiuto di deindicizzazione del gestore della piattaforma, l'utente dovrà adire l'autorità giudiziaria.

Anche la seconda questione posta nel medesimo rinvio interessa particolarmente la presente analisi, perché era relativa alla possibilità che la deindicizzazione potesse avere ad oggetto anche le foto di persone fisiche che, nell'ambito di una ricerca nominativa, fossero visualizzate come miniature (*"thumbnails"*) e dovesse tener conto in modo determinante del contesto della pubblicazione originaria, anche quando il motore di ricerca, visualizzando la miniatura, in effetti rimanda al sito originario, ma senza menzionarlo concretamente.

La Corte, dunque, imprime una marcia in più alla tutela della persona eventualmente lesa dalla pubblicazione delle immagini, perché esse rispetto alla comunicazione verbale hanno un impatto più forte sugli utenti di Internet, dunque, nella valutazione della richiesta di tale deindicizzazione deve essere attribuito loro un valore informativo superiore che prescinde dal contesto della loro pubblicazione nelle pagine *web* originarie[105].

---

[103] Così le conclusioni dell'Avvocato Generale Pitruzzella in C-460/20, § 30.

[104] Infatti, per evitare di far gravare sulla persona un eccessivo onere accertativo, la Corte afferma che essa sia tenuta unicamente a fornire elementi di prova, dei quali può ragionevolmente essere in possesso, atti a dimostrare l'inesattezza manifesta. CGUE, C-460/20, *Tu. e Re. v. Google* (2022), § 68.
Deve segnalarsi che il canone della falsità e dell'inesattezza come criteri identificativi della responsabilità delle piattaforme *online* e del loro obbligo di rimozione immediata dei contenuti pubblicati è posto fondamento anche della pronuncia del Tribunale di Milano, sez. I civ., 15 febbraio 2023, n. 1208, emessa in un procedimento per risarcimento del danno promosso dalla società *Snaitech* (nota concessionaria per la gestione dei giochi legali e autorizzati in Italia) contro *Facebook*.
In essa il *social network*, anche in considerazione dell'«ampia capacità diffusiva dei contenuti che ospitano le piattaforme c.d. social» (12-13), viene condannato al risarcimento del danno per non aver rimosso le pagine recanti post palesemente falsi, sulla base della costatazione per la quale il «diritto di critica, il quale costituisce notoriamente espressione della libertà di manifestazione del pensiero di matrice costituzionale (art. 21 Cost.)» non è configurabile nel caso in cui il contraddittore aggredisca «con accuse di perpetrazione di veri e propri delitti o comunque di condotte infamanti in rapporto alla dimensione personale, sociale o professionale del destinatario» (11).

[105] CGUE, C-460/20, *Tu e Re. v. Google* (2022), § 85.

7. I rimedi predisposti dal diritto positivo dell'Unione europea

7.1 Il Digital Market Package

Quanto invece ai rimedi diritto positivo avverso la diffusione di *deepfake* e *deep porn,* deve segnalarsi come l'Unione europea sia stata la prima nel panorama globale a dotarsi di una normazione in materia di mercati e servizi digitali resi dalle società di informazione, garantendo peraltro alle vittime di *deepfake* strumenti di tutela ulteriori rispetto alla possibilità di adire la autorità giudiziaria, adeguati alla velocità con la quale si diffondono *online* tali contenuti dannosi[106].

Ci si riferisce al *Digital Markets Package*, oggi integrato dal regolamento *Artificial Intelligence Act.*

Il primo è un pacchetto di due regolamenti*,* in vigore dal mese di maggio 2023 (*Digital Market Act*[107] e *Digital Service Act*[108]), che delinea un modello di intervento regolativo che mira a contemperare, da un lato, le esigenze dello sviluppo economico del settore digitale e la capacità di innovazione del tessuto imprenditoriale europeo, e, dall'altro, gli interessi economici e politici dell'Unione; esso si caratterizza anche per la precisa scelta normativa di prediligere norme di carattere prevalentemente procedurale o procedimentale, che procedono di pari passo con un più generale quadro assiologico, fondato sui valori dell'Unione europea[109].

La scelta dello strumento normativo, il regolamento e non la direttiva, latore di norme armonizzate e direttamente applicabili, manifesta la volontà del legislatore europeo, volta a riaccentrare la disciplina di materie così complesse nelle sue mani[110] e ad evitare la frammentazione normativa a livello dei singoli Stati membri[111], pervenendo all'uniformazione delle condizioni alle quali gli operatori digitali dovranno soggiacere per poter prestare i propri servizi nel mercato interno dell'UE[112].

I due regolamenti appaiono formalmente distinti, occupandosi di materie diverse. Spe-

---

106 In Italia è nota la vicenda che qualche anno fa coinvolse una ragazza, che si era tolta la vita per la vergogna provata per il fatto che per alcuni mesi, senza il suo consenso, erano circolati su *social network* foto e video di intimità sessuali pubblicati su Facebook dal fidanzato. La madre della ragazza aveva proposto ricorso d'urgenza contro la piattaforma sociale, che, nonostante le richieste, non aveva rimosso i post, ottenendo il risarcimento del danno solo tre anni dopo e solo dopo una lunga vicenda giudiziaria (Tribunale di Napoli, sez. II civ., 3 novembre 2016).

107 Regolamento (UE) 2022/1925 del Parlamento europeo e del Consiglio del 14 settembre 2022 relativo a mercati equi e contendibili nel settore digitale e che modi- fica le direttive (UE) 2019/1937 e (UE) 2020/1828 (*Digital Market Act*).

108 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 diretti- va 2000/31/CE (*Digital Service Act).*

109 Così anche O. Pollicino, *Regolazione e innovazione tecnologica nell'"ordinamento della Rete",* relazione tenuta al Convegno Nazionale della Associazione Italiana dei Costituzionalisti, *"La libertà di manifestazione del pensiero",* svoltosi a Salerno 15-16 novembre 2024, 40.

110 M.R. Allegri, *Il futuro digitale dell'Unione Europea: nuove categorie di intermediari digitali, nuove forme di responsabilità*, in *Rivista italiana di informatica e diritto*, 1, 2021, 11 e 12.

111 L. Torchia, *Stato digitale*, cit., 77.

112 M.R. Allegri, *Il futuro digitale dell'Unione Europea,* cit., 10.

cificatamente, il DMA si occupa dei rapporti fra le piattaforme e i fornitori dei servizi, mentre il *Digital Service Act* dei rapporti fra questi ultimi e gli utenti.
Tale distinzione è utile meramente a fini esplicativi, poiché è evidente che nel mercato digitale i momenti di sovrapposizione fra i due tipi di regolazione sono inevitabili, tanto più che le piattaforme sono solo degli intermediari fra i fornitori di servizi e coloro che invece li richiedono.
In realtà essi vanno letti insieme, poiché prevedono ampi e penetranti poteri di vigilanza, controllo e sanzione in capo alla Commissione sulle *Over the Top Companies* (a tal fine indentificate preliminarmente come *Gatekeepers*), cui poi si aggiungono anche quelli degli Stati membri.
Le ragioni che hanno convinto i legislatori europei alla approvazione del DMA sono quelle determinate dalla consapevolezza della insufficienza delle regole *antitrust*, a soccorrere alle nuove esigenze del mercato digitale, poiché sostanzialmente fondate su controlli *ex post* e, quindi, che ad esse debbano necessariamente affiancarsi nuove regole che dispongano anche un quadro di obblighi *ex ante* da imporre a queste imprese, indipendentemente dalla individuazione del mercato rilevante.
Per questi fini il Regolamento stabilisce anche una serie di precisi indicatori sulla base dei quali individuare il prestatore di servizi qualificabile come *Gatekeeper* (art. 3).
In questo modo viene istituito un sistema di presunzione *ex ante* della qualificazione, con la contestuale attribuzione alla Commissione di un ruolo centrale, la quale peraltro vigila sui numerosi e articolati obblighi imposti ai *Gatekeepers* (artt. 5-17), con ampi poteri di indagine, monitoraggio ed esecuzione delle norme del DMA, oltre che di quelli sanzionatori e la capacità di imporre rimedi comportamentali o strutturali.
La disciplina dettata dal *Digital Service Act* (*DSA*), invece, reca regole a protezione degli utenti *online*, per garantire la loro libertà di espressione, ma anche la libertà di iniziativa economica delle piattaforme[113].
La normativa ruota intorno a tre cardini fondamentali: l'individuazione della responsabilità dei *provider*, gli obblighi di diligenza e la cooperazione con le autorità.
Quanto al primo punto, la disciplina, abrogando con l'art. 71 la previgente relativa ai servizi delle società di informazione 2000/31, determina una ridefinizione generale della responsabilità degli intermediari (tra questi ovviamente anche le piattaforme e i *social networks*), che prescinde dalla tipologia del servizio che svolgono, imponendo che la eventuale irresponsabilità del *provider* debba essere provata di volta in volta, divenendo, peraltro, sempre più ardua via via che aumenta la complessità del servizio offerto.
Si determina così una palese rivoluzione copernicana, rispetto al regime precedente, nel quale invece vigeva al contrario la presunzione di neutralità dell'intermediario[114].

[113] Così anche O. Pollicino, *Regolazione e innovazione tecnologica nell'"ordinamento della Rete",* cit., 40, che afferma che scopo della disciplina è «"addomesticare" i giganti del mercato digitale e, quindi, andare direttamente a intervenire su uno degli aspetti caratterizzanti la società algoritmica […] combinando insieme, da un lato, novità concernenti gli obblighi dei *provider* alla tutela di un ambiente digitale trasparente e sicuro e, dall'altro lato, nuove regole relative alla promozione della concorrenza».

[114] Gli artt. 14 e 15 del DSA individuano nella procedura di *notice and take down* il principale strumento di cooperazione fra intermediari e utenti, finalizzato alla rimozione dei contenuti, imponendo alle prime che l'accesso alle notifiche da parte dei secondi sia facile e a che la notifica sia formulata con precisione, in modo da consentire una effettiva conoscenza dell'illecito e all'intermediario di adottare le decisioni conseguenti (art. 14). La eventuale decisione di rimozione del contenuto dovrà essere poi notificata al

Ciò che non viene definito dal DSA è però l'esatta nozione di contenuto illecito, dedicandosi più che altro a dettare norme di carattere procedurale volte alla moderazione e alla eventuale rimozione dello stesso e lasciando tale individuazione alle autorità giudiziarie nazionali (considerando 29 del DSA)[115].

In questo modo, secondo una condivisibile opinione della dottrina, i *providers* per discernere e valutare quali contenuti dovranno necessariamente essere moderati, avranno a disposizione un *corpus* normativo vastissimo sia di livello europeo sia di livello nazionale, che potrebbe facilmente determinare l'insorgere di controversie fra utenti, autorità nazionali e operatori digitali[116].

Un ulteriore aspetto innovativo del DSA è che esso modula e diversifica gli obblighi di cooperazione e vigilanza a seconda delle dimensioni degli intermediari e alla complessità dei servizi da loro offerti, distinguendoli in quattro categorie: *intermediary services, hosting, online platform e le very large online platforms*[117].

A queste ultime è richiesto, come obbligo aggiuntivo di cooperazione, quello di istituire ulteriori meccanismi di gestione dei reclami, che dovranno essere gestiti in modo tempestivo, diligente e obiettivo, come anche le decisioni sui contenuti.

Gli obblighi di moderazione dei contenuti, come anche quelli di controllo, imposti alle *Very Large Platforms* sono quindi ancora più severi. Esse, infatti, sono onerate anche

destinatario con adeguata motivazione, in modo da consentirgli di reclamare la decisione attraverso i meccanismi interni, o di risoluzione extragiudiziale delle controversie o ancora per via giudiziaria (art. 15).

[115] Il considerando 29 dispone infatti che «a secondo dell'ordinamento giuridico di ciascuno Stato membro di ciascuno Stato membro e del settore del diritto in questione, le autorità giudiziarie o amministrative nazionali possono ordinare ai prestatori di servizi intermediari di contrastare determinati contenuti illegali specifici o di fornire determinate informazioni specifiche»

Per discernere ciò che costituisca contenuto illegale, l'interprete deve infatti rifarsi, intanto, al considerato n. 12 che ne fornisce una ampia interpretazione, introducendovi «informazioni, indipendente- mente dalla loro forma, che ai sensi del diritto applicabile sono di per sé illegali [...] A tale riguardo è irrilevante che l'illegalità delle informazioni o delle attività sia sancita dal diritto dell'Unione o dal diritto nazionale conforme al diritto dell'Unione e quale sia la natura esatta o l'oggetto preciso della legge in questione».

A questo si deve anche aggiungere, l'art. 2 lett. g) del DSA, che qualifica "contenuto illegale" «qualsiasi informazione che, di per sé o in relazione ad un'attività, tra cui la vendita di prodotti o la prestazione di servizi, non è conforme alle disposizioni normative dell'Unione o di uno Stato membro, indipendentemente dalla natura o dall'oggetto specifico di tali disposizioni».

[116] M.R. Allegri, *Il futuro digitale dell'Unione Europea,* cit., 15.

D'altro canto, non può sorprendere la decisione del legislatore europeo di non disporre in modo sostanziale della definizione di "contenuto illegale", per diversi ordini di ragioni: i citati problemi definitori di ciò che sia *fake news* (e oggi *deepfake*) e disinformazione; il rischio che la cristallizzazione in disposizioni positive di tali definizioni non sarebbe in grado di seguire l'evoluzione tecnologica sottesa alla manifestazione delle opinioni online e renderebbe precocemente obsolete le norme stesse; le limitazioni alla libertà di espressione che da tale definizioni deriverebbero.

[117] Queste ultime sono le piattaforme digitali che hanno in media oltre 45 milioni di utenti attivi mensili nell'Unione europea, pari al 10% della popolazione europea.

A imporre la distinzione sono i considerando da 53 a 63 del DSA. In particolare nel considerando 53 si legge che «Data l'importanza che le piattaforme online di dimensioni molto grandi, per via del loro raggio d'azione, espresso in particolare come numero di destinatari del servizio, rivestono nel facilitare il dibattito pubblico, le operazioni economiche e la diffusione di informazioni, opinioni e idee e nell'influenzare il modo in cui i destinatari ottengono e comunicano informazioni *online*, è necessario imporre a tali piattaforme obblighi specifici, in aggiunta agli obblighi applicabili a tutte le piattaforme online. Tali obblighi supplementari per le piattaforme online di dimensioni molto grandi sono necessari per affrontare tali preoccupazioni di interesse pubblico, in quanto non esistono misure alternative e meno restrittive che consentano di conseguire efficacemente lo stesso risultato».

di valutare i rischi sistemici connessi al funzionamento e all'uso dei loro servizi e ai possibili abusi da parte dei destinatari, con il conseguente obbligo di adottare anche le misure per attenuarli.

La distinzione fra piattaforme e *Very Large Platforms* si coglie anche quanto ai poteri di controllo e di vigilanza. Infatti, per le prime è previsto che in ogni Stato membro si costituiscano appositi organismi di risoluzione extragiudiziale delle controversie fra intermediari e utenti, cui questi ultimi potranno rivolgersi se insoddisfatti dalle scelte delle piattaforme in esito ai reclami o in alternativa rispetto ai reclami stessi.

Inoltre, è previsto che sia individuato presso gli Stati membri anche un coordinatore dei servizi digitali[118], che certifica la sussistenza a livello nazionale di questi organismi e contribuisce alla applicazione coerente del Regolamento. I coordinatori sono inoltre tenuti a cooperare fra loro, con la Commissione e con il comitato europeo per i servizi digitali (art. 18).

Si crea, dunque, quanto al controllo sulle attività delle piattaforme, un sistema reticolare che vede la cooperazione fra i singoli Stati, raccordati in un organismo di livello sovranazionale, e fra questi e la Commissione[119].

La disciplina dei controlli e della vigilanza sulle *Very Large Platforms* è invece del tutto diversa da quella appena descritta per le piattaforme per così dire "ordinarie" e assume una struttura fortemente accentrata.

Esse sono anche soggette, ai sensi degli artt. 50 e ss., ad una procedura di vigilanza rafforzata, relativa alla conformità delle loro attività rispetto alle norme del Regolamento, che coinvolge la Commissione europea. Il ruolo di quest'ultima, infatti, sovrasta quello degli altri organismi di controllo, potendo essa intervenire direttamente di sua iniziativa nel caso di persistenza delle violazioni, svolgendo autonomamente indagini e audizioni, ispezioni in loco, adottando misure provvisorie, rendendo vincolanti impegni e finanche irrogare sanzioni pecuniarie per le violazioni del regolamento (artt. 50-66).

Ciò evidentemente avvicina il modello di controllo delle *Very Large Platforms* a quello *antitrust* e presenta degli importanti punti di collegamento alla disciplina di controllo di cui al *Digital Market Act*, tanto che parte della dottrina lo ha definito "controllo gemello" rispetto a quest'ultimo[120].

---

[118] I poteri dei quali sono disciplinati agli artt. 38, 39 e 41 del DSA.

[119] L. Torchia, *I poteri di vigilanza, controllo e sanzionatori nella regolazione europea della trasformazione digitale,* in *Rivista trimestrale di diritto pubblico*, 2022, 1110.
Al sistema di controllo reticolare appena descritto si aggiunge il sistema dei "segnalatori attendibili", di cui all'art. 19 DSA. Si tratta in particolare di enti accreditati dagli Stati membri che rappresentano interessi collettivi e sono indipendenti dagli intermediari digitali, le cui segnalazioni vanno trattate con priorità. Come con analoga priorità le piattaforme potranno reagire sospendendo i servizi nei confronti di quegli utenti che con frequenza diffondono contenuti manifestamente illegali (art. 20, par. 1).

[120] M.R. Allegri, *Il futuro digitale dell'Unione Europea,* cit., 17.

7.2. (*segue*) L'*AI Act* e il disegno di legge italiano sull'Intelligenza Artificiale

Alla normativa dettata dal *Digital Market Package* oggi si aggiunge l'*AI Act*[121] entrato in vigore il 2 agosto 2024, è caratterizzato da termini progressivi di applicazione delle sue disposizioni, che saranno applicabili nella loro totalità a 36 mesi dalla entrata in vigore stessa. Il rischio, a fronte di questa applicazione graduale, è quello del precoce invecchiamento di norme dettate molti mesi prima in una materia che affronta l'inarrestabile evoluzione tecnologica, la quale per sua natura procede ad un ritmo e ad una velocità non contenibile entro confini di natura temporale.

Esso si pone come pietra miliare della disciplina europea in tema di Intelligenza Artificiale, in una società che grazie a questa tecnologia si sta velocemente evolvendo da società digitale a *cybersociety*[122], modificando l'approccio normativo dalla «automazione fondata sull'"algoritmo" a una prospettiva sempre più fondata sull'"intelligenza artificiale"». Laddove, «mentre l'algoritmo si sostanzia in una sequenza di istruzioni ben definite, non ambigue e, dunque, applicate in modo meccanico dalla macchina, l'intelligenza artificiale, fondandosi per lo più su sistemi di *machine learning*, si caratterizza per il fatto di essere in grado di elaborare autonomamente regole di inferenza a partire dai dati usati per l'allenamento»[123].

Fra gli obiettivi principali del Regolamento quello della protezione degli utenti e dei loro diritti fondamentali, imponendo alle piattaforme gestorie obblighi di informazione laddove i singoli interagiscano con tali sistemi (anche e soprattutto con immagini, contenuti audio o video artificiali o manipolati, come nel caso dei *deepfake*) e implementando i controlli sul trattamento e la gestione dei dati personali.

In primo luogo, esaminando il regolamento, va sottolineato come l'ambito di applicazione (art. 2) si estenda ai fornitori che immettono sul mercato o mettono in servizio sistemi di Intelligenza Artificiale nell'Unione, indipendentemente dal fatto che siano stabiliti nell'Unione o in un Paese terzo, nonché agli utenti dei sistemi di IA situati nell'Unione e ai fornitori e agli utenti di sistemi di IA situati in un Paese terzo, ove l'*output* prodotto dal sistema sia utilizzato nell'Unione.

La strategia dell'Unione europea è, dunque, quella di porsi come *leader* nella produzione normativa anche in questo campo, facendo sì che il modello europeo divenga un riferimento globale e possa essere adottato nelle altre regioni del mondo, per esempio con

[121] Regolamento (UE) 2024/1689 del Parlamento europeo e del Consiglio del 13 giugno 2024, che stabilisce regole armonizzate sull'intelligenza artificiale.

[122] Così si esprime O. Pollicino, *Regolazione e innovazione,* cit., 44-45 che cita anche L. Violante, *Diritto e potere nell'era digitale. Cybersociety, cybercommunity, cyberstate, cyberspace: tredici tesi*, in *Biolaw Journal – Rivista di Biodiritto,* 1, 2022, 145 ss. e A. Simoncini - S. Suweis, *Il cambio di paradigma nell'intelligenza artificiale e il suo impatto sul diritto costituzionale*, in *Rivista di filosofia del diritto*, 1, 2019, 87–106.

[123] O. Pollicino, *Regolazione e innovazione*, cit*.,* 45. Si veda anche A. Simoncini, *Il linguaggio dell'Intelligenza Artificiale e la tutela costituzionale dei diritti*, in *Rivista AIC*, 2, 2023, 1–39.

gli Stati Uniti[124] (il cosiddetto «effetto Bruxelles»)[125].

Leggendo l'*AI Act* in combinato disposto col GDPR, nonché con il DMA e il DSA, si ha quindi l'idea di un disegno normativo generale nell'Unione europea volto a salvaguardare non solo i diritti di tutti coloro che entrino in contatto con la tecnologia digitale, ma anche i "valori" europei in senso culturale. Infatti, il termine "valori" è menzionato più volte nell'*AI Act*, in modo da «sottolineare che il modello elaborato non è solo normativo, ma culturale. Si vuole rendere evidente che non si tratta soltanto di regole giuridiche, ma anche della cultura che quelle regole esprimono»[126].

L'approccio adottato dal legislatore europeo per la regolazione della Intelligenza Artificiale è stato definito "orizzontale"[127], con norme anche in questo caso estremamente generali di carattere procedurale volte «non a risolvere specifici problemi o a colmare determinate lacune dell'ordinamento, ma applicabili a qualunque settore, […] per delineare un quadro complessivo, un contesto di riferimento nel quale opereranno i sistemi di intelligenza artificiale, anche quelli ancora da venire»[128].

Seguendo, dunque, lo stesso modello adottato per il *Digital Markets Package*, il regolamento muove dalla classificazione delle tecnologie di IA fondata su quattro categorie, in ragione del rischio che presentano: sistemi a rischio inaccettabile, ad alto rischio e a rischio basso o minimo.

I primi, identificati dall'art. 5, sono assolutamente vietati. Per i sistemi di IA a basso rischio, invece, sono previsti alcuni obblighi di trasparenza e si incoraggia l'adozione di codici di condotta (art. 93)[129]. Lo stesso regolamento dispone che per i *deepfake*, considerati anche essi sistemi a basso rischio, a meno che non comportino la commissione di reati, è previsto che gli utenti rendano noto che il contenuto è stato generato o manipolato artificialmente.

Infine, la gran parte del regolamento è dedicata a prevedere in dettaglio gli obblighi per l'adozione di sistemi di IA ad alto rischio[130], che saranno soggetti a determinati vincoli prima di poter essere utilizzati tra i quali l'obbligo di fornire adeguata documentazione contenente tutte le informazioni necessarie sullo scopo del sistema, affinché le autorità possano valutarne la conformità[131], la predisposizione di una valutazione dei rischi, la

---

[124] G. Finocchiaro, *La regolazione,* cit., 1091-1092, afferma che la strategia normativa dell'UE ha un evidente obiettivo geopolitico, volendo contrastare in questo modo la *leadership* tecnologica cinese e statunitense. In particolare, come si è già accennato al § 5, l'approccio statunitense sembra sviluppare un «modello auto-regolatorio basato sull'*antitrust*».

[125] A. Bradford, *The Brussels Effect: How the European Union Rules the World*, New York, 2020.

[126] G. Finocchiaro, *La regolazione,* cit., 1091.

[127] Ivi, 1089.

[128] Ivi*,* 1093.

[129] Ad esempio, per i sistemi di IA destinati a interagire con le persone fisiche, è richiesto che esse siano informate del fatto che stanno interagendo con un sistema di IA; per i sistemi di riconoscimento delle emozioni o di categorizzazione biometrica, è prescritto agli utenti di informare delle loro modalità di funzionamento le persone fisiche che vi siano esposte.

[130] Si tratta di tutte le applicazioni che comprendono sistemi in grado di arrecare danni significativi alla salute, alla sicurezza, ai diritti fondamentali o all'ambiente delle persone, includendovi quelli utilizzati per l'influenzare gli elettori e i risultati delle elezioni ed i sistemi di raccomandazione utilizzati dalle piattaforme dei *social media* con una base utenti superiore a 45 milioni.

[131] Si tratta di una procedura di valutazione della conformità *ex ante*, la quale si conclude con l'apposizione

garanzia sulla tracciabilità dei risultati.

Sebbene l'*AI Act* sia il primo atto normativo che ambisce a regolare per intero questo settore, esso presenta delle indubbie criticità[132].

In primo luogo, la classificazione dei sistemi di IA sulla base del rischio cristallizza oggi una tecnologia in continuo divenire e dunque anche queste tipologie saranno necessariamente soggette a revisione, poiché certamente verranno sviluppati nuovi sistemi e nuovi metodi per implementare quanto già esistente, modificando il livello di rischio.

In secondo luogo, le medesime soluzioni, anche in termini di *accountability*, sono adottate indiscriminatamente per soggetti e ambiti assai diversi fra loro e a prescindere dalle dimensioni delle imprese, profilandosi dunque forti rischi per le imprese di piccole dimensioni e per le *start-up*.

Dal punto di vista sostanziale, infine, pur ponendosi a tutela dei valori europei, il regolamento si limita a vietare i sistemi di intelligenza artificiale che comportano un rischio inaccettabile, rinviando poi, in modo implicito o esplicito, ai principi generali che sono ormai al cuore del diritto europeo (come la dignità, la trasparenza, la protezione dei dati personali), senza prevedere delle specifiche modalità di applicazione degli stessi ai sistemi di intelligenza artificiale, né forme nuove e più efficaci di tutela dell'individuo[133].

Che quello della regolazione sia il modello prescelto in Europa, lo si comprende anche dal fatto che alla luce dell'*AI Act* alcuni Stati stanno promuovendo l'approvazione di proprie normative nazionale sul tema. In Italia, per esempio, presso l'Ottava Commissione (congiunta con la Decima) del Senato della Repubblica si sta svolgendo l'esame in sede referente del disegno di legge n. 1146 di iniziativa governativa intitolato "Disposizioni e delega al Governo in tema di Intelligenza Artificiale"[134].

La proposta di legge mira a definire, nel rispetto dell'*AI Act*, per la cui attuazione viene data delega di adozione di uno o più decreti legislativi, un quadro normativo domestico in relazione ad alcuni aspetti cruciali connessi all'utilizzo dei sistemi di IA, con particolare riferimento a quei settori nei quali tale utilizzo potrebbe avere un impatto significativo a livello sociale ed economico e con l'ambizione di fornire una risposta ad alcune delle preoccupazioni manifestate a proposito dell'utilizzo di IA in settori come la sanità, la Pubblica Amministrazione, la giustizia e le professioni.

L'intento del Governo proponente sembra quello di voler accelerare l'introduzione di alcuni dei principi previsti dall'*AI ACT*, di meglio disciplinare alcune aree potenzialmente critiche interessate dall'utilizzo dei sistemi di IA, consentendo all'Italia di avere una presenza strategica nel contesto europeo.

Per questo motivo, il testo contiene richiami espliciti ai diritti fondamentali ed alle libertà previste dalla Costituzione italiana e dal diritto dell'UE, nonché ai principi di

della marcatura CE (art. 48).

[132] Si veda L. Floridi, *The European Legislation on AI: A Brief Analysis of its Philosophical Approach*, in *Philosophy and Technology*, 2021,, 216 e G. Finocchiaro, *La regolazione,* cit., 1094 ss.

[133] Come, infatti, rilevato da G. Finocchiaro, *La regolazione,* cit., 1098, il regolamento «definisce una cornice di natura amministrativa per l'immissione nel mercato dei prodotti di Intelligenza Artificiale. Il quadro generale dovrà però essere completato dalle norme tecniche e dagli standard, che rivestiranno un'importanza fondamentale, e dovrà essere continuamente aggiornato».

[134] Il disegno di legge di iniziativa del Governo è stato approvato dal Consiglio dei ministri il 23 aprile 2024.

trasparenza, proporzionalità, sicurezza, valorizzazione e protezione dei dati personali, accessibilità e non discriminazione, a presidio dell'autonomia e dell'autodeterminazione umana. Analogamente la definizione di "sistema di intelligenza artificiale" corrisponde esattamente a quella contenuta nel regolamento europeo.

Quanto alla privacy, il testo ribadisce il principio fissato dall'art. 22 del GDPR secondo cui ciascuno ha il diritto di non essere sottoposto a una decisione basata unicamente sul trattamento automatizzato, compresi i trattamenti svolti dai sistemi di intelligenza artificiale, sebbene non vengano fornite indicazioni specifiche su come debba essere gestito il trattamento dei dati personali, limitandosi a rimandare alla legislazione corrente.

Limitatamente al tema oggetto di questo lavoro, l'art. 4 del citato disegno di legge nel disporre che «L'utilizzo di sistemi di intelligenza artificiale nell'informazione avviene senza pregiudizio alla libertà e al pluralismo dei mezzi di comunicazione, alla libertà di espressione e […] dell'informazione», garantisce anche «il *trattamento lecito, corretto e trasparente* dei dati personali […] in conformità con il diritto dell'Unione europea in materia di dati personali e di tutela della riservatezza»[135], evidentemente rinviando all'*AI Act*, e, inoltre, rafforza la tutela dei minori disponendo che il loro accesso a tale tecnologia avvenga «consenso di chi esercita la responsabilità genitoriale»[136].

Considerazioni finali

L'analisi dei rimedi posti dal diritto dell'Unione europea (e proposti nell'ordinamento italiano), volti a disciplinare i servizi resi dagli intermediari *online* (i.e. il *Digital Markets Package*) anche per mezzo dell'Intelligenza Artificiale (i.e. *AI Act*) e a porre un argine sul fronte del diritto penale ai reati commessi per mezzo di questa nuova tecnologia, sollecita delle brevissime riflessioni conclusive.

In prima battuta si è avuto modo di rilevare che in assenza di una disciplina positiva era stata la giurisprudenza (in specie si è analizzata quella della Corte di Giustizia dell'Unione europea, ma analoghi rimedi giurisprudenziali sono stati posti al livello dei singoli Stat e della Corte EDU) a porre un argine a tutte quelle situazioni giuridiche lesive dei diritti dei singoli e che sorgono a causa del diffondersi della tecnologia digitale e dei sistemi di IA.

In particolare, si è avuto modo di constatare come nell'Unione europea è facilmente estendibile alle fattispecie di *deepfakes* illegali la giurisprudenza sulla disinformazione e come negli Stati Uniti i giudici dei singoli Stati stiano apprestando analoghe tutele alle vittime, laddove però esistano leggi statali specifiche sul tema.

Va ribadito che in USA la mancanza di una legge federale determina la grave frammentazione dei possibili argini al diffondersi di reati di tal fatta e lascia ai singoli Stati il compito di gestire, autonomamente, l'impatto dell'IA generativa in questioni così delicate[137].

[135] Art. 4, c. I e II.

[136] Art. 4, c. IV, il quale ulteriormente dispone che: «Il minore degli anni diciotto, che abbia compiuto quattordici anni, può esprimere il proprio consenso per il trattamento dei dati personali connessi all'utilizzo di sistemi di intelligenza artificiale, purché le informazioni e le comunicazioni di cui al comma 3 siano facilmente accessibili e comprensibili».

[137] Per completezza si deve segnalare che tuttavia, alcune leggi federali esistenti riguardano l'AI, sebbene con applicazioni limitate. Un esempio è il *National AI Initiative Act* del 2020 (aggiornato, da ultimo, nel

In questo è evidente come l'approccio europeo al tema sia molto diverso da quello statunitense, così come invero accade anche per la tutela della libertà di espressione che si esprime attraverso modelli essenzialmente diversi fra le due sponde dell'Atlantico, avvicinando anche sotto questo aspetto le fattispecie di *fake news* a quelle di *deepfake*, i quali sembrano i "successori" delle prime, progrediti grazie all'evoluzione della tecnologia.
Quello statunitense manifesta oggi una visione del tutto *business friendly* volta a favorire le imprese e a tutelarne lo sviluppo, per mezzo delle sole disposizioni di natura *antitrust* e di una totale deregolazione dello sviluppo delle tecnologie di IA, mirato a garantire in questo modo la *leadership* americana del settore.
L'Unione europea con il DMA, il DSA e l'*AI Act* mostra invece la sua visione "umanocentrica", che pone invece al centro i diritti dei singoli e i valori dell'Unione.
Entrambe le soluzioni però lasciano irrisolto un nodo che accompagna tutte riflessioni rese nel presente lavoro, restando sempre sullo sfondo: la necessità di garantire alle vittime di *deepfake* una tutela effettiva, ma anche rapida, che non aspetti i tempi lunghi della giustizia e che si dimostri adeguata alla velocità con la quale la disinformazione si diffonde nel *web*.
A tal proposito soccorre in aiuto un'idea sviluppata anni fa da certa parte della dottrina italiana[138] in tema di manifestazione del pensiero, che potrebbe validamente essere applicata anche in questi casi, che, come già sostenuto, possono essere considerati una evoluzione del problema delle *fake news*, determinata dallo sviluppo della tecnologia.
La tesi sosteneva che sarebbe stato possibile costruire una rete di controllo di Autorità nazionali, collegate alla Commissione europea, sul modello delle Autorità *antitrust*, in modo che velocemente i *fake* (*fake news*, ma oggi potrebbe dirsi analogamente per i *deepfake*) possano essere valutati come tali e altrettanto velocemente rimossi.
L'idea è stata molto criticata in origine[139], ma oggi ritrova la sua validità se aggiornata e rivalutata in base al nuovo contesto normativo vigente nell'Unione europea (specificatamente il *Digital Market Package* e l'*Artificial Intelligence Act*).
Il primo, infatti, impone alle piattaforme di cooperare con le autorità europee nell'identificare i *fake,* adottando il principio della responsabilità dell'intermediario *online,* e individua nella Commissione europea l'organo deputato al controllo ultimo e centralizzato.
L'*AI Act* poi all'art. 70 espressamente dispone l'istituzione o l'individuazione di Autorità Indipendenti presso i singoli Stati membri, che svolgano l'attività di controllo, notifica e vigilanza in stretta comunicazione con la Commissione[140].

2023), che ha l'obiettivo di ampliare la ricerca e lo sviluppo nel campo dell'AI e ha istituito il *National Artificial Intelligence Initiative Office*, responsabile della supervisione e dell'implementazione della strategia nazionale statunitense sull'AI.

[138] La tesi di G. Pitruzzella resa nota dapprima con due interviste concesse al *Financial Times* il 30 dicembre 2016 *Italy antitrust chief urges EU to help beat fake news* e al *Corriere della Sera* il 2 gennaio 2017 *Quel filtro necessario per le notizie false sul web*, poi ulteriormente sviluppata in Id., *La libertà di informazione*, 82, 93-95.

[139] C.A. Carnevale Maffé, *Neppure l'Autorità della Veridicità può fermare il mercato delle bufale*, in *Il Foglio*, 7 gennaio 2017 e analogamente si vedano C. Melzi D'Eril – G.E. Vigevani, *Difesa giuridica dal sociachiacchiericcio*, in *Il Sole 24 Ore*, 2 aprile 2017; M. Bassini, *Fake news: perché non è un lavoro da spazzini (del web)*, in *medialaws.eu*, 16 marzo 2017; N. Zanon, *Fake News e diffusione dei social media: abbiamo bisogno di un'"Autorità Pubblica della Verità"?*, in questa *Rivista*, 1, 2018, 17.

[140] Art. 70, par. 1, dell'*AI Act* dispone infatti che: «Ciascuno Stato membro istituisce o designa come

Analogamente la proposta di legge italiana sulla IA, in attuazione delle disposizioni del regolamento europeo, pur mantenendo impregiudicate le attribuzioni dell'Autorità Garante per la Protezione dei Dati Personali, individua nell'Agenzia Italiana per il Digitale (AgID) e nell'Agenzia per la Cybersicurezza Nazionale (ACN) le Autorità Nazionali per l'Intelligenza Artificiale, distinguendone i compiti e le funzioni (art. 18).

In particolare, la prima sarà responsabile di promuovere l'innovazione e lo sviluppo dell'Intelligenza Artificiale e di provvedere a definire le procedure e ad esercitare le funzioni e i compiti in materia di notifica, valutazione, accreditamento e monitoraggio dei soggetti incaricati di verificare la conformità dei sistemi di intelligenza artificiale, secondo quanto previsto dalla normativa nazionale e dell'Unione europea. L'ACN, anche ai fini di assicurare la tutela della cybersicurezza, sarà responsabile per la vigilanza, ivi incluse le attività ispettive e sanzionatorie, dei sistemi di intelligenza artificiale, secondo quanto previsto dalla normativa nazionale e dell'Unione europea. Entrambe, invece, per quanto di rispettiva competenza, assicurano l'istituzione e la gestione congiunta di spazi di sperimentazione finalizzati alla realizzazione di sistemi di intelligenza artificiale conformi alla normativa nazionale e dell'Unione europea.

Da ultimo una considerazione finale, i temi trattati in questo lavoro si pongono al confine fra diritto pubblico e diritto privato, in un momento in cui il potere pubblico sembra riappropriarsi della sua sovranità normativa, che, fino a qualche tempo fa a causa del vuoto normativo persistente in materia, sembrava persa in favore delle grandi piattaforme sovrane del mondo digitale.

Si tratta di un terreno accidentato e per lo più sconosciuto, che pone il quesito di «come edificare il nuovo diritto costituzionale dell'Intelligenza Artificiale tra dimensione pubblica e privata»[141].

A tale quesito si può forse rispondere ritenendo che le prospettive aperte dalla nuova regolazione europea dell'IA devono essere accompagnare necessariamente dalla valorizzazione degli strumenti giuridici esistenti, come la tutela rafforzata dei diritti fondamentali e la garanzia di adeguati e omogenei livelli qualitativi dei servizi indispensabili al soddisfacimento dei diritti fondamentali, che imporrebbero ai proprietari privati dei sistemi di IA un dovere di protezione, divenendo titolari di obblighi tipici dei soggetti pubblici, come la trasparenza e l'imparzialità.

In questo modo il potere (espresso anche da soggetti privati che gestiscono le piattaforme *online*) sarebbe nuovamente catturato e limitato e il costituzionalismo riscoprirebbe la propria primigenia missione[142].

---

autorità nazionali competenti ai fini del presente regolamento almeno un'autorità di notifica e almeno un'autorità di vigilanza del mercato. Tali autorità nazionali competenti esercitano i loro poteri in modo indipendente, imparziale e senza pregiudizi, in modo da salvaguardare i principi di obiettività delle loro attività e dei loro compiti e garantire l'applicazione e l'attuazione del presente regolamento. I membri di tali autorità si astengono da qualsiasi atto incompatibile con le loro funzioni. A condizione che siano rispettati detti principi, tali compiti e attività possono essere svolti da una o più autorità designate, conformemente alle esigenze organizzative dello Stato membro».

141 A. Simoncini, *La dimensione costituzionale dell'Intelligenza Artificiale,* in G. Cerrina Feroni – C. Fontana – E.C. Raffiotta (a cura di), *AI Anthology,* cit., 150.

142 A. Simoncini, *La dimensione costituzionale,* cit., 154 riprende anche l'idea di M. Luciani, *Costituzionalismo irenico e costituzionalismo polemico*, in *Giurisprudenza costituzionale,* 4, 2006, 1668.