Freedom, Security & Justice: European Legal Studies

Rivista giuridica di classe A

2025, n. 2

EDITORIALE SCIENTIFICA

Rivista quadrimestrale on line "Freedom, Security & Justice: European Legal Studies" www.fsjeurostudies.eu
Editoriale Scientifica, Via San Biagio dei Librai, 39 - Napoli
CODICE ISSN 2532-2079 - Registrazione presso il Tribunale di Nocera Inferiore n° 3 del 3 marzo 2017

Indice-Sommario
2025, n. 2

Editoriale

Saggi, Articoli, Commenti e Note

FOCUS

Democracy and the Rule of Law: A New Push for European Values

Il Focus contiene contributi elaborati a seguito della riflessione realizzata nel Seminario conclusivo dello Jean Monnet Module Eu-Draw (2022-2025) "Democracy and the Rule of Law: A New Push for European Values", tenutosi presso l'Università degli Studi di Salerno (1 aprile 2025)

L'"APPROCCIO EUROPEO" AL CONTRASTO ALLA DISINFORMAZIONE DIGITALE E ALLA PROTEZIONE DEI VALORI DEMOCRATICI: QUALE CONTRIBUTO DELL'*AI ACT*?

Rossana Palladino*

SOMMARIO: 1. Introduzione. – 2. Valori democratici dell'UE e "approccio europeo" nel contrasto alla disinformazione digitale. – 3. L'approccio normativo *hard*: l'*AI Act*. – 3.1. La previsione dell'intelligenza artificiale manipolatoria tra le pratiche vietate. – 3.2. Le previsioni riguardanti i sistemi ad alto rischio. – 3.3. L'introduzione di obblighi di trasparenza per "altre pratiche manipolative". – 4. Considerazioni conclusive: contributo e limiti dell'*AI Act* nel contrasto alla disinformazione digitale e alla protezione dei valori democratici dell'UE.

1. Introduzione

L'intelligenza artificiale (IA), nell'attuale era digitale, sta profondamente trasformando le modalità di produzione e di diffusione delle informazioni, configurandosi al tempo stesso come un'opportunità e un elemento di sfida per la tenuta democratica delle nostre società. Da un lato, lo sviluppo complessivo della digitalizzazione e l'uso diffuso di internet rappresentano, come ha sottolineato la Corte europea dei diritti dell'uomo, "*an unprecedented platform for the exercise of freedom of expression*"[1]; dall'altro lato, la diffusione di disinformazione *online*, fenomeno potenziato e amplificato dall'impiego dell'intelligenza artificiale, può costituire fattore di crescente minaccia per l'integrità delle elezioni e dei sistemi democratici. Basti pensare al noto ed emblematico caso *Cambridge Analytica*[2], che ha portato

Double-blind peer reviewed article.
* Professoressa associata di Diritto dell'Unione europea, Università degli Studi di Salerno. Indirizzo e-mail: rpalladino@unisa.it.
[1] In particolare, v. Corte europea dei diritti dell'uomo, sentenza del 1° dicembre 2015, *Cengiz e altri* c. *Turchia*, ricorsi nn. 48226/10 e 14027/11, par. 52 e Corte europea dei diritti dell'uomo, Grande Camera, sentenza del 15 marzo 2023, *Sanchez* c. *Francia*, ricorso n. 45581/15, par. 159.
[2] È stato il giornalismo investigativo a rilevare e rendere pubbliche le massicce fughe di dati degli utenti Facebook in relazione all'accesso concesso da tale piattaforma ad applicazioni terze e il conseguente abuso di tali dati a fini di campagna elettorale nonché altre violazioni dei dati personali detenuti e raccolti dalle principali imprese nel settore dei media sociali che sono venute alla luce successivamente, come poi acclarato dal Parlamento europeo nella Risoluzione del 25 ottobre 2018, *sull'utilizzo dei dati degli utenti Facebook da parte di Cambridge Analytica e l'impatto sulla protezione dei dati* (2018/2855(RSP)).

all'attenzione pubblica i rischi connessi all'uso di *microtargeting* e della profilazione psicografica nel condizionamento delle preferenze elettorali. L'impiego dell'intelligenza artificiale in rete può agevolare siffatte pratiche, non soltanto tramite la profilazione degli utenti ma anche attraverso strumenti come i *deep fake*[3] e i contenuti generati automaticamente: strumenti avanzati come i *Large Language Models* (LLMs) e la generazione di immagini, audio e video sintetici possono amplificare il rischio di manipolazione dell'opinione pubblica e di disinformazione[4].

In ambito europeo, è oramai nota l'"operazione Doppelgänger", basata sullo sviluppo di cloni di siti web legittimi utilizzati per diffondere notizie false in Germania e Francia, con l'obiettivo di disorientare gli elettori. Inoltre, secondo il rapporto del Comitato europeo per i servizi digitali (EDMO) sulla disinformazione durante le elezioni del 2023 in Europa[5], la circolazione e diffusione di notizie false riguardanti frodi elettorali, interferenze straniere e pratiche sleali ha rappresentato una criticità diffusa in diversi paesi, sollecitando un'azione di "contrasto" più incisiva in vista delle elezioni del Parlamento europeo del 2024. Da ultimo, un caso particolarmente significativo è quello della Romania, dove le elezioni presidenziali del 2024 sono state annullate dalla Corte costituzionale, in considerazione dell'uso non trasparente dell'intelligenza artificiale nella manipolazione del voto e della conseguente alterazione delle pari opportunità tra i candidati[6].

Di fronte all'emersione di tali fenomeni[7], l'Unione europea (UE) ha progressivamente delineato un "approccio europeo" al contrasto della disinformazione *online*, riconoscendone i connotati di minaccia per la democrazia, valore indefettibile dell'UE stessa. Su tale approccio, definibile di *soft law* in assenza di una competenza normativa attribuita in *subiecta materia* all'UE, si innestano oggi le misure vincolanti

[3] Secondo la definizione ora contenuta nell'*AI Act*: un'immagine o un contenuto audio o video generato o manipolato dall'IA che assomiglia a persone, oggetti, luoghi, entità o eventi esistenti e che apparirebbe falsamente autentico o veritiero a una persona.

[4] Nel considerare l'accesso a una corretta informazione, quale *"precondition for an informed and genuine exercise of the right to vote*" v. R. MASTROIANNI, *Fake news, free speech and democracy: A (bad) lesson from Italy?*, in *Southwestern Journal of International Law*, 2019, pp. 42-74, spec. p. 45.

[5] Reperibile *online*: EDMO-TF-Elections-disinformation-narratives-2023.pdf. Si veda anche L'EUROPEAN UNION AGENCY FOR FUNDAMENTAL RIGHTS (FRA), *Fundamental Rights Report 2025*, Luxembourg, 2025, spec. p. 26 ss.

[6] Corte costituzionale romena, sentenza n. 32 del 6 dicembre 2024, par. 11, dove la Corte rileva, sulla base di una serie di "note informative" prese in considerazione, che "i principali aspetti contestati nel processo elettorale per l'elezione del Presidente della Romania del 2024 riguardano la manipolazione del voto degli elettori e la distorsione della parità di opportunità tra i concorrenti elettorali, attraverso l'uso non trasparente e in violazione della legislazione elettorale di tecnologie digitali e di intelligenza artificiale durante la campagna elettorale, nonché il finanziamento non dichiarato della campagna elettorale, anche online». Reperibile *online* nella traduzione in italiano: CorteCost.Romania-dec.6-12-24_n.32. Sul punto, M.G. LOSANO, *Le Corti costituzionali di Romania e Colombia sulle interferenze informatiche e sull'intelligenza artificiale*, in *Diritto pubblico comparato ed europeo*, 2025, n. 1, pp. 117-138; S. SASSI, A. STERPA, *La Corte costituzionale della Romania difende la democrazia liberale dalla disinformazione. Prime note sulla sentenza n. 32 del 6 dicembre 2024*, in *federalismi.it*, 2025, n. 4, pp. 162-181.

[7] In tema, evidenziando parallelamente l'utilizzo dell'intelligenza artificiale anche quale strumento a supporto del contrasto alla disinformazione, N. BONTRIDDER, Y. POULLET, *The role of artificial intelligence in disinformation*, in *Data & Policy*, 2021, n. 3, pp. 3-32.

adottate nella stagione regolatoria del mercato unico digitale[8], in particolare tramite il *Digital Services Act* (DSA)[9] che impone a grandi piattaforme – quali, ad esempio, *Facebook* e *TikTok* – l'obbligo di individuare e segnalare contenuti manipolati, inclusi i cd. *deep fake*. Tale strumento giuridico esulerà dall'oggetto della presente indagine[10], che intende, invero, concentrarsi sullo specifico strumento regolatorio dell'intelligenza artificiale adottato di recente dall'Unione europea, ossia il cd. *AI Act (Artificial Intelligence Act*)[11], con l'obiettivo di valutarne il contributo, in termini di "*hard law*", nel contrasto ai rischi legati alla disinformazione elettorale, tenendo conto che già nella fase preparatoria, numerose organizzazioni della società civile, autorità pubbliche, mondo accademico, avevano posto l'attenzione sull'uso dei sistemi di intelligenza artificiale nella manipolazione dei comportamenti umani, nella determinazione delle proprie opinioni e decisioni, evidenziando non solo un danno alla persona ma anche un detrimento dei processi democratici[12] e, in ultimo, della democrazia quale valore fondante dell'UE.

2. Valori democratici dell'UE e "approccio europeo" nel contrasto alla disinformazione digitale

Ai sensi dell'art. 2 del Trattato sull'Unione europea (TUE), la democrazia è ricompresa nel nucleo dei valori fondanti dell'ordinamento giuridico dell'Unione europea: valore comune a tutti gli Stati membri, unitamente al rispetto della dignità umana, alla libertà, all'uguaglianza, allo Stato di diritto e al rispetto dei diritti umani. Si tratta di valori che assumono una funzione identitaria per l'Unione europea, caratterizzandone il nucleo assiologico e avendo una duplice valenza: il loro rispetto è requisito imprescindibile per l'adesione di un nuovo Stato membro, nonché per la

[8] Sulla stagione europea del "costituzionalismo digitale", caratterizzata dalla "volontà di riappropriazione, da parte del legislatore, del ruolo di *law maker*" per lungo tempo esercitato, di fatto, dalla Corte di giustizia dell'Unione europea, v. O. POLLICINO, *Di cosa parliamo quando parliamo di costituzionalismo digitale?*, in *Quaderni costituzionali*, 2023, n. 3, pp. 569-594; nonché G. DI GREGORIO, *The Rise of Digital Constitutionalism in the European Union*, in *International Journal of Constitutional Law*, 2021, n. 1, pp. 41-70; G. PITRUZZELLA, *Il costituzionalismo digitale tra Stati Uniti e Europa*, in *Unione europea e Diritti*, 2025, n. 2, pp. 1-11.

[9] Regolamento (UE) 2022/2065 del Parlamento europeo e del Consiglio, del 19 ottobre 2022, *relativo a un mercato unico dei servizi digitali e che modifica la direttiva 2000/31/CE (regolamento sui servizi digitali)*, in GU L 277 del 27.10.2022.

[10] Trovando, invece, approfondimento nel contributo di A. FESTA, in questo numero della *Rivista*.

[11] Si tratta del Regolamento (UE) 2024/1689 del Parlamento europeo e del Consiglio, del 13 giugno 2024, *che stabilisce regole armonizzate sull'intelligenza artificiale e modifica i regolamenti (CE) n. 300/2008, (UE) n. 167/2013, (UE) n. 168/2013, (UE) 2018/858, (UE) 2018/1139 e (UE) 2019/2144 e le direttive 2014/90/UE, (UE) 2016/797 e (UE) 2020/1828 (regolamento sull'intelligenza artificiale)*, in GU L del 12 luglio 2024.

[12] European Commission, *Contributions to White Paper on Artificial Intelligence: Public Consultation Towards a European Approach for Excellence and Trust* (20 February-14 June 2020), disponibile *online*: White Paper on Artificial Intelligence: Public consultation towards a European approach for excellence and trust | Shaping Europe's digital future.

permanenza in tale organizzazione internazionale che se ne fa promotrice anche nelle relazioni con il resto del mondo (articolo 3, par. 5, del TUE).

Sebbene non sia facile definire in maniera univoca, sul piano normativo, il valore della democrazia ovvero ricondurlo a un "modello rigido" e tassativo[13], è ravvisabile un nucleo essenziale e indefettibile di tale principio nella derivazione dei poteri pubblici dalla volontà popolare. L'essenza del principio democratico si manifesta nella possibilità, per i cittadini, di esprimere liberamente le proprie opinioni, partecipare attivamente alla vita politica e istituzionale, eleggere i propri rappresentanti e incidere sul processo decisionale. Ciò implica la necessità di garantire uno spazio pubblico inclusivo, in cui sia consentita l'espressione di posizioni eterogenee, il diritto al dissenso e la possibilità di modificare gli assetti di governo mediante consultazioni elettorali libere e trasparenti, esenti da interferenze interne o esterne.

Nelle più attuali fasi dell'integrazione, progressiva centralità ha assunto l'Unione europea nel presidiare il rispetto del valore della democrazia, al pari degli altri valori comuni, in particolare lo Stato di diritto[14], che vanno condivisi e riconosciuti da tutti gli Stati membri[15] e che rappresentano l'essenza "dell'identità stessa dell'Unione"[16] quale ordinamento giuridico comune. Un "nuovo slancio" per la democrazia, che assicuri l'adesione a modelli pienamente democratici in tutti gli Stati membri dell'Unione europea, anche nel misurarsi con le questioni sollevate dall'evoluzione digitale, è

[13] Cfr. U. VILLANI, *Valori comuni e rilevanza delle identità nazionali e locali nel processo di integrazione europea*, Napoli, 2011. In generale, sui valori fondanti dell'Unione europea, ci si limita a richiamare M. MESSINA (a cura di), *I valori fondanti dell'Unione europea a 60 anni dai Trattati di Roma*, Napoli, 2017; B. NASCIMBENE, *Valori comuni dell'Unione europea*, in E. TRIGGIANI, F. CHERUBINI, I. INGRAVALLO, E. NALIN, R. VIRZO (a cura di), *Dialoghi con Ugo Villani*, tomo I, Bari, 2017, pp. 631-636. V. anche R. BARATTA, *Droits fondamentaux et "valeurs" dans le processus d'intégration européen*, in *Studi sull'integrazione europea*, 2019, n. 2, pp. 289-308. Con riferimento all'azione esterna dell'Unione europea, E. SCISO, R. BARATTA, C. MORVIDUCCI (a cura di), *I valori dell'Unione europea e l'azione esterna*, Torino, 2016. Per il profilo della "giustiziabilità" dei valori: L. DIMITRIOS SPIEKER, *EU Values Before the Court of Justice*, Oxford, 2024, nonché E. CANNIZZARO, *Il ruolo della Corte di giustizia nella tutela dei valori dell'Unione europea*, in A.A. V.V., Liber amicorum *Antonio Tizzano. De la Cour CECA à la Cour de l'Union: le long parcours de la justice européenne*, Torino, 2018, pp. 158-169; L.S. ROSSI, *Il valore giuridico dei valori. L'art. 2 TUE: relazioni con altre disposizioni del diritto primario dell'UE e rimedi giurisdizionali*, in *federalismi.it*, 2020, n. 20, pp. ix-xxvi; C. NOVI, *I valori dell'Unione europea e la loro tutela giurisdizionale*, in questa *Rivista*, 2025, n. 1, pp. 136-188.

[14] Ci si limita a rinviare a M. CARTA, *Unione europea e tutela dello Stato di diritto negli Stati membri*, Bari, 2020; A. FESTA, *Lo Stato di diritto nello spazio europeo. Il ruolo dell'Unione europea e delle altre organizzazioni internazionali*, Napoli, 2021; A. CIRCOLO, *Il valore dello Stato di diritto nell'Unione europea*, Napoli, 2023; nonché U. VILLANI, *Sul controllo dello Stato di diritto nell'Unione europea*, in questa *Rivista*, 2020, n. 1, pp. 10-27.

[15] La condivisione di siffatti valori rappresenta, infatti, la base su cui poggia la fiducia reciproca tra gli Stati membri e la creazione di uno spazio di (libertà, sicurezza e) giustizia. Con riferimento specifico all'indipendenza dei giudici, tanto è stato sottolineato dalla CGUE nella sentenza del 27 febbraio 2018, *Associação Sindical dos Juízes Portugueses* c. *Tribunal de Contas*, causa C-64/16, ECLI:EU:C:2018:117, punto 30. V. anche Parere della CGUE del 18 dicembre 2014, 2/13, ECLI:EU:C:2014:2454, punto 191, emesso ai sensi dell'art. 218, par. 11, del TFUE sull'adesione dell'Unione europea alla Convenzione europea per la salvaguardia dei diritti dell'uomo e delle libertà fondamentali (CEDU).

[16] Corte di giustizia, Seduta Plenaria, sentenze del 16 febbraio 2022, *Ungheria* c. *Parlamento europeo e Consiglio dell'Unione europea*, causa C-156/21, ECLI:EU:C:2022:97, punto 232 e *Polonia* c. *Parlamento europeo e Consiglio dell'Unione europea*, causa C-157/21, ECLI:EU:C:2022:98, punto 264.

obiettivo che si è prefissata la Commissione europea 2019-2023 nell'ambito del Piano d'azione dell'UE per la democrazia europea[17], cui si aggiungono un nuovo meccanismo europeo per lo Stato di diritto[18] e la nuova strategia per rafforzare l'applicazione della Carta dei diritti fondamentali[19].

Nel solco della "dimensione valoriale" dell'Unione europea, quest'ultima oramai da un decennio ha intrapreso l'impegno della lotta alla disinformazione, muovendosi su di un piano non strettamente di regolamentazione legislativa, in considerazione della ravvisabile insussistenza di una competenza attribuita in tal senso all'Unione europea, trattandosi di una materia che, stando a quanto ricavabile dagli artt. 2-6 del TUE, è da ritenersi appartenente alle competenze "proprie" degli Stati membri. Diversamente, il ruolo che l'Unione europea può ritagliarsi è di mero "sostegno"[20] agli Stati membri nell'ottica dell'adozione di azioni comuni volte a rafforzare il coordinamento, la comunicazione e l'adozione di buone pratiche, sospinti dall'ottica di dare concretezza alla dimensione valoriale dell'Unione e alla tutela dei diritti fondamentali contenuti nella Carta dei diritti fondamentali dell'UE, entrambi – ad ogni buon conto – non ammessi di per se stessi a fondare una competenza dell'UE in siffatta materia[21].

---

[17] Commissione europea, Comunicazione della Commissione al Parlamento europeo, al Consiglio, al Comitato economico e sociale europeo e al Comitato delle Regioni, *Sul piano d'azione per la democrazia europea*, Bruxelles, 3.12.2020, COM(2020) 790 final. Di rilievo è anche la Raccomandazione (UE) 2023/2829 della Commissione del 12 dicembre 2023, *relativa a processi elettorali inclusivi e resilienti nell'Unione e al rafforzamento della natura europea e dell'efficienza nello svolgimento delle elezioni del Parlamento europeo*, in GU L del 20.12.2023, ove si incoraggiano i partiti politici e gli organizzatori di campagne elettorali ad assumere impegni riguardo alle proprie campagne e ad adottare codici di condotta sull'integrità delle elezioni e sulla correttezza delle campagne, che prevedano anche la volontà di "astenersi da comportamenti manipolativi che minacciano i valori, le procedure e i processi politici, o che potrebbero incidere negativamente su di essi", in particolare "la produzione, l'utilizzo o la diffusione di dati o materiali falsificati, fabbricati, che rivelano informazioni personali (doxing) o rubati, compresi i deepfake generati da sistemi di intelligenza artificiale".

[18] Cfr. Comunicazione della Commissione al Parlamento europeo, al Consiglio europeo, al Consiglio, al Comitato economico e sociale europeo e al Comitato delle Regioni, *Rafforzare lo Stato di diritto nell'Unione. Programma d'azione*, Bruxelles, 17.7.2019, COM(2019) 343 final. Il nuovo Meccanismo europeo per lo Stato di diritto consiste in un ciclo di valutazione svolto, annualmente, dalla Commissione europea nei confronti degli Stati membri.

[19] Comunicazione della Commissione al Parlamento europeo, al Consiglio, al Comitato economico e sociale europeo e al Comitato delle regioni, *Strategia per rafforzare l'applicazione della Carta dei diritti fondamentali dell'Unione europea*, Bruxelles, 2.12.2020, COM(2020) 711 final.

[20] Seppur non espressamente fondato sull'art. 6 del TFUE, che individua i settori nei quali l'Unione ha competenza per svolgere azioni intese a sostenere, coordinare o completare l'azione degli Stati membri.

[21] Sull'impossibilità dell'art. 2 TUE di fondare una base giuridica, in particolare v. A. ADINOLFI, *L'intelligenza artificiale tra rischi di violazione dei diritti fondamentali e sostegno alla loro promozione: considerazioni sulla (difficile) costruzione di un quadro normativo dell'Unione*, in A. PAJNO, F. DONATI, A. PERRUCCI (a cura di), *Intelligenza artificiale e diritto: una rivoluzione? Diritti fondamentali, dati personali e regolazione*, vol. I, Bologna, 2022, pp. 127-164, al cui contributo si rinvia per approfondimenti sul punto. Con riferimento alla Carta dei diritti fondamentali dell'UE, è noto che sussistono una serie di "presidi" che non consentono di utilizzarla come base giuridica autonoma: l'art. 6 TUE precisa che "Le disposizioni della Carta non estendono in alcun modo le competenze dell'Unione definite nei trattati", aspetto che è ribadito anche dall'art. 51, par. 2, della Carta: "La presente Carta non estende l'ambito di applicazione del diritto dell'Unione al di là delle competenze dell'Unione, né introduce competenze nuove o compiti nuovi per l'Unione, né modifica le competenze e i compiti definiti nei trattati".

Pertanto, attesa la competenza propria degli Stati membri in ordine alla protezione del processo elettorale, l'azione dell'Unione europea matura per fare fronte alla acquisita dimensione transfrontaliera del fenomeno della disinformazione in rete, al fine di assicurare interventi efficaci e coordinati che proteggano la democrazia europea nel suo complesso, considerato che la legittimità e la stessa ragione d'essere dell'UE si basa su fondamenta democratiche, che a loro volta dipendono da un elettorato informato, che esprime la propria volontà democratica tramite elezioni libere e regolari; di conseguenza, qualsiasi tentativo malevolo e intenzionale di diffondere sfiducia o manipolare l'opinione pubblica si traduce in una "grave minaccia" per l'Unione europea nel suo complesso.

È su tale presupposto che l'impegno dell'Unione europea nella lotta alla disinformazione può trovare un *dies a quo* nel 2015[22], allorquando il Consiglio europeo[23] ha invitato l'Alto Rappresentante per gli affari esteri e la politica di sicurezza a elaborare un piano d'azione in materia di comunicazione strategica, con l'obiettivo specifico di contrastare le campagne di disinformazione in corso da parte della Russia[24].

Sulla base dei lavori di un Gruppo di esperti ad alto livello istituito dalla Commissione europea, quest'ultima ha poi adottato un "approccio europeo" nel

[22] Per un quadro sulla lotta alla disinformazione nell'UE: J. BALTRIMAS, *Disinformation in the EU Law: Moral Theories and the Context*, in *Journal of the University of Latvia*, 2024, n. 17, pp. 273-290; J. BAYER, *The EU policy on disinformation: aims and legal basis*, in *Journal of Media Law*, 2024, n. 1, pp. 18-27; P. GABORIT, *A Sociopolitical Approach to Disinformation and AI: Concerns, Responses and Challenges*, in *Journal of Political Science and International Relations*, 2024, n. 4, pp. 75-88; S. SASSI, *L'Unione Europea e la lotta alla disinformazione* online, in *federalismi.it*, 2023, n. 3, pp. 183-201.

[23] Conclusioni della riunione del Consiglio europeo del 19 e 20 marzo 2015, punto 13, EUCO 11/15.

[24] Successivamente, al fine di contrastare le attività di disinformazione nei confronti dell'opinione pubblica degli Stati membri dell'UE, contro la Russia sono state adottate anche sanzioni individuali "atipiche", ossia non costituite dal blocco dei beni o da divieti di ingresso nel territorio degli Stati membri nei confronti di soggetti non statali, quanto piuttosto dal divieto di radiodiffusione nel territorio dell'UE. Con sentenza del 27 luglio 2022 (causa T-125/22, *RT France* c. *Consiglio dell'Unione europea*, ECLI:EU:T:2022:483) il Tribunale ha rigettato l'impugnazione evidenziando, tra l'altro, come le attività di propaganda e le campagne di disinformazione costituiscano una minaccia diretta all'ordine e alla sicurezza pubblica dell'Unione europea, minando le fondamenta di una società democratica e fanno parte dell'arsenale moderno di guerra (punti 53 e 56); sebbene poi le misure impugnate costituiscono un'ingerenza nel diritto alla libertà di informazione della ricorrente di natura temporanea, esse risultano giustificate in quanto sono previste dalla legge, rispettano il contenuto essenziale dei diritti di cui all'art. 11 della Carta dei diritti fondamentali dell'Unione europea, rispondono a un interesse generale riconosciuto come tale dall'UE e sono proporzionate. In argomento, A. MAFFEO, *La sottile linea di confine tra libertà di informazione e propaganda di guerra: il caso* RT France, in R. MASTROIANNI, F. FERRARO (a cura di), *Libertà di informazione e diritto dell'Unione europea. Le nuove sfide a tutela della democrazia e del pluralismo*, Napoli, 2022, pp. 195-216; S. POLI, *Le misure Ue di contrasto alle attività di disinformazione russe, alla prova della Carta europea dei diritti fondamentali*, in *Quaderni costituzionali*, 2022, n. 3, pp. 626-634; S. LATTANZI, *Su propaganda e ordine pubblico in Europa*, in *Politica del diritto*, 2023, n. 3, pp. 389-418; L. LONARDO, *Censorship in the EU as a result of the war in Ukraine: RT France v Council of the European Union*, in *European Law Review*, 2023, n. 6, pp. 707-719; J.-P. JACQUÉ, *Liberté d'information et mesures restrictives*, in *Revue trimestrielle des droits de l'homme*, 2024, n. 139, pp. 719-732. Più in generale, G. MORGESE, *Il contrasto alla disinformazione originata da ingerenze straniere nell'Unione Europea*, in M. MESSINA (a cura di), *Strengthening the European Union through the European Citizenship and the Rule of Law/Cittadinanza e Stato di diritto per un'Unione europea più forte*, Napoli, 2024, pp. 89-130.

contrasto alla disinformazione *online*[25] e un piano d'azione contro la disinformazione[26], definendo dieci azioni specifiche basate su quattro settori prioritari o "pilastri" indirizzate alla società nel suo complesso.

Ulteriormente, la politica di contrasto alla disinformazione *online* ha condotto all'adozione, nel 2018, di un *EU Code of Practice on Disinformation* ("rafforzato" nel 2022)[27] che ha impegnato – attraverso un approccio di tipo volontario basato sull'autoregolamentazione[28] – la Commissione europea, le piattaforme *online* e altre associazioni di categoria nell'adozione di buone pratiche. In particolare, attraverso tale *Code of practice*, le piattaforme *online* e le associazioni di categoria rappresentanti il settore pubblicitario si sono impegnate a trasmettere alla Commissione europea relazioni che delineino la situazione delle misure adottate per rispettare gli impegni assunti. Tali misure possono consistere nell'assicurare la trasparenza nei messaggi pubblicitari di natura politica, oppure nel chiudere profili falsi o, ancora, nell'evitare che i vettori di disinformazione ne traggano vantaggi economici.

Dalla serie di misure analizzate, è possibile anche delineare il concetto di "disinformazione" – in luogo di quello più comunemente invalso di *fake news* – da intendersi come "un'informazione rivelatasi falsa o fuorviante concepita, presentata e diffusa a scopo di lucro o per ingannare intenzionalmente il pubblico, e che può arrecare un pregiudizio pubblico".

Un concetto che, dunque, vale a ricomprendere tutte quelle informazioni che sostanziano in informazioni false o fuorvianti, che siano state create, presentate e diffuse al preciso scopo di trarne un profitto economico oppure per ingannare il pubblico in maniera intenzionale, determinando un pregiudizio qualificabile non a livello individuale ma a livello "pubblico". Nella fattispecie di "disinformazione" vengono in rilievo, pertanto, condotte di per se stesse non qualificate come illecite né dall'ordinamento europeo né dagli ordinamenti nazionali – diversamente da condotte che integrano fattispecie vietate come può essere, ad esempio, la diffamazione – laddove

[25] Comunicazione della Commissione al Parlamento europeo, al Consiglio, al Comitato economico e sociale europeo e al Comitato delle Regioni, *Contrastare la disinformazione online: un approccio europeo*, Bruxelles, 26.4.2018, COM(2018) 236 final. Si veda O. POLLICINO, *The European Approach to Disinformation: Comparing Supranational and National Measures*, in *Annuario di diritto comparato e studi legislativi*, 2020, pp. 175-212.

[26] Comunicazione congiunta al Parlamento europeo, al Consiglio europeo, al Consiglio, al Comitato economico e sociale europeo e al Comitato delle Regioni, *Piano d'azione contro la disinformazione*, Bruxelles, 5.12.2018, JOIN(2018) 36 final.

[27] Il cui testo è reperibile *online*: 2018_Code_of_Practice_on_Disinformation_l4DbpCSGHOu3e1vYe0Dbzq669k_87534 (1).pdf., così come il testo del Codice "rafforzato" del 2022: 2022_Strengthened_Code_of_Practice_Disinformation_TeAETn7bUPXR57PU2FsTqU8rMA_87585 (2).pdf.

[28] G. PAGANO, *Il Code of Practice on Disinformation. Note sulla natura giuridica di un atto misto di autoregolazione*, in *federalismi.it*, 2019, n. 11, p. 2; A. KUCZERAWY, *Fighting online disinformation: did the EU Code of Practice forget about freedom of expression?*, in G. TERZIS, D. KLOZA, E. KUŻELEWSKA, D. TROTTIER (eds.), *Disinformation and Digital Media as a Challenge for Democracy*, Cambridge, 2020, pp. 291-308; M. MONTI, *Lo 'strengthened Code of Practice on Disinformation': un'altra pietra della nuova fortezza digitale europea?*, in *MediaLaws*, 2022, n. 2, pp. 317-321.

l'elemento centrale è rappresentato dalla "possibilità" di arrecare un "pregiudizio pubblico", nel quale può considerarsi ricompresa la ingerenza nei processi democratici e la "minaccia" al corretto funzionamento del sistema democratico e dello Stato di diritto, in sostanza un detrimento a quei valori che fondano la stessa Unione europea.

Unitamente al potenziale "pregiudizio pubblico", l'altro elemento che connota la disinformazione è costituito dalla volontarietà dell'azione (concepimento, presentazione, diffusione di una informazione falsa o fuorviante) finalizzata all'ottenimento di un profitto oppure ad ingannare il pubblico, elemento che consente, peraltro, di segnare una distinzione rispetto ad altri fenomeni, quali, in particolare, la *misinformation* (o, altrimenti, disordine informazionale) e le *fake news*[29].

L'azione così delineata nella lotta alla disinformazione – diretta a contrastare contenuti disinformativi, suscettibili di produrre un pregiudizio pubblico, ma di per sé, come si è detto, non qualificabili come "illeciti" – pone evidentemente anche una sfida significativa nell'ottica della preservazione della libertà di opinione e di espressione, così come sancite nella Carta dei diritti fondamentali dell'Unione europea[30].

Viene in rilievo, all'uopo, l'articolo 11 della Carta dei diritti fondamentali dell'UE che garantisce il diritto di ogni persona alla libertà di espressione, inclusa la "libertà di opinione e la libertà di ricevere o di comunicare informazioni o idee senza che vi possa essere ingerenza da parte delle autorità pubbliche e senza limiti di frontiera", precisando anche che "la libertà dei media e il loro pluralismo sono rispettati".

Un diritto garantito nella stretta corrispondenza con l'articolo 10 della Convenzione europea dei diritti dell'uomo (CEDU), evidenziata anche dalla Corte di giustizia dell'UE allorquando ne ha enfatizzato il rilievo alla luce della particolare importanza che tale diritto riveste in qualsiasi "società democratica": un diritto fondamentale che

[29] La prima fattispecie (*misinformation*) si differenzia dalla disinformazione in ragione dell'elemento soggettivo, contraddistinguendosi per l'assenza di una precisa volontà nella diffusione di informazioni false e, dunque, nella diffusione inconsapevole di queste ultime. Evidenziano, nell'ecosistema di internet, la trasformazione di iniziali fattispecie di disinformazione in misinformazione, O. POLLICINO, P. DUNN, *Disinformazione e intelligenza artificiale nell'anno delle* global elections*: rischi (ed opportunità)*, in *federalismi.it*, 2024, n. 12, pp. iv-xxii, reperibile *online*. Il termine "*fake news*" ricopre, invece, forme di discorso considerate "a basso rischio", quali ad esempio parodie, discorsi politici di parte, screditamento di parti politiche, etc...

[30] Cfr. J. BAYER, I. KATSIREA, O. BATURA, B. HOLZNAGEL, S. HARTMANN, K. LUBIANIEC, *The fight against disinformation and the right to freedom of expression*, 2021, Policy Department for Citizens' Rights and Constitutional Affairs, reperibile *online*. In generale, sulla libertà di espressione nel diritto dell'Unione europea, si vedano M. CASTELLANETA, *La libertà di stampa nel diritto internazionale ed europeo*, Bari, 2012, spec. p. 10 ss.; V. SALVATORE, *La libertà di espressione, una prospettiva di diritto comparato. Unione europea*, studio EPRS, Servizio Ricerca del Parlamento europeo, 2019, reperibile *online*; F. FERRARO, *Brevi note sulla libertà di espressione e di informazione nel diritto dell'Unione*, in R. MASTROIANNI, F. FERRARO (a cura di), *Libertà di informazione e diritto dell'Unione europea*, *op. cit.*, pp. 1-20; C. MAUBERNARD, S. PLATON, R. TINIÈRE (cur.), *Les mutations de la liberté d'expression dans l'Union européenne*, Bruxelles, 2025.

"costituisce uno dei fondamenti essenziali di una società democratica e pluralista, facente parte dei valori sui quali, a norma dell'articolo 2 TUE, l'Unione è fondata"[31].

Si tratta, pertanto, di una libertà intrinsecamente connessa alla democrazia – che, a sua volta, rappresenta l'unico sistema politico in grado di garantire la protezione dei diritti umani – essenziale a siffatto sistema tanto da trovare applicazione anche in caso di informazioni "*that offend, shock or disturb the State or any sector of the population*", essendo ciò richiesto da quel pluralismo, tolleranza e spirito di apertura in assenza dei quali una "società democratica" non esisterebbe affatto[32].

In linea con l'art. 10 della CEDU – nonché con l'art. 19 del Patto Internazionale sui diritti civili e politici – anche l'art. 11 della Carta dei diritti fondamentali dell'UE ricomprende una "dimensione passiva" del diritto alla libertà di espressione, consistente nel diritto di cercare e ricevere informazioni, ossia come il diritto di "formarsi un'opinione libera e indipendente" che dà concretezza a quella che è definibile come "libertà di opinione", la cui possibile compressione è considerata sostanzialmente a rischio grazie al progressivo utilizzo di tecniche manipolative che le piattaforme digitali, gli attori statali e gli altri soggetti *online* possono porre in essere per influenzare le persone.

Con specifico riferimento allo "spazio digitale", il diritto alla libertà di espressione e di informazione è ribadito anche nell'art. 12 della Dichiarazione sui diritti e principi digitali[33], ove allo stesso tempo si considera il ruolo delle piattaforme *online* – specie quelle di dimensioni molto grandi – nel "sostenere il libero dibattito democratico *online*" e, considerato il ruolo svolto dai loro servizi "nel plasmare l'opinione pubblica e il dibattito pubblico", per primo è individuato un obbligo di attenuazione dei rischi derivanti dal funzionamento e dall'uso dei loro servizi, "anche in relazione alle campagne di disinformazione e cattiva informazione, e tutelare la libertà di espressione".

---

[31] *Ex pluribus*, Corte di giustizia dell'UE, Grande Sezione, sentenza del 21 dicembre 2016, cause riunite C-203/15 e C-698/15, *Tele2 Sverige AB* c. *Post- och telestyrelsen e Secretary of State for the Home Department contro Tom Watson e a*., ECLI:EU:C:2016:970, punto 93.

[32] Mettendo, dunque, in luce la duplice valenza della libertà di espressione, riconosciuta come essenziale sia per lo sviluppo personale di qualsiasi essere umano, sia per il progresso della società nella sua collettività, cfr. Corte europea dei diritti dell'uomo, *Handyside* c. *Regno Unito*, sentenza del 7 dicembre 1976, ricorso n. 5493/72, par. 49. Va anche sottolineato che, allorquando si è trovata a confrontarsi con casi di cd. "*hate speech*", la Corte EDU abbia evidenziato la necessità di tenere conto del contesto specifico entro cui le espressioni offensive trovano formulazione; così, in particolare, cfr. Corte EDU, *Gunduz c. Turchia*, sentenza del 4 dicembre 2003, ricorso n. 35071/97, parr. 40-41. In tema, M. NINO, *The Freedom of Expression and Hate Speech in Cyberspace*, in *La Comunità internazionale*, 2023, n. 1, pp. 33-57 e ID., *Il rapporto tra libertà di espressione e discorsi d'odio nell'era digitale alla luce della normativa internazionale ed europea*, in A. ORIOLO, A.R. CASTALDO, A. DI STASI, M. NINO (a cura di), *Criminalità transnazionale e Unione europea*, Napoli, 2024, pp. 767-790.

[33] Dichiarazioni comuni, Parlamento europeo, Consiglio, Commissione europea, *Dichiarazione europea sui diritti e i principi digitali per il decennio digitale*, 2023/C 23/01. Per un'analisi di tale Dichiarazione, anche in relazione al suo valore giuridico, v. L. CIANCI, *Dichiarazione europea sui diritti e i principi digitali:* quid pluris*?*, in *Diritto pubblico comparato ed europeo*, 2022, pp. 381-390; P. DE PASQUALE, *Verso una Carta dei diritti digitali (fondamentali) dell'Unione europea?*, in *Il Diritto dell'Unione europea*, 2022, p. 163 ss.

3. L'approccio normativo *hard*: l'*AI Act*

Sul piano legislativo, come si traduce l'impegno dell'UE nel contrasto alla disinformazione *online*? Si è già osservato che, in carenza di una competenza in tal senso dell'Unione europea e di una specifica base giuridica, non si può immaginare l'adozione di un quadro normativo *hard* incentrato sul contrasto alla disinformazione; piuttosto si può evidenziare che siffatto obiettivo è perseguito, in via mediata, nell'ambito degli strumenti legislativi che danno concretezza al "decennio digitale europeo" e alla "sovranità digitale dell'UE", complessivamente improntati al rispetto dei valori propri dell'Unione europea[34].

In primis, è il già citato *Digital Services Act* a tradurre sul piano legislativo l'obbligo, cui sono soggetti i fornitori di piattaforme *online* e i motori di ricerca di dimensioni molto grandi, di effettuare una valutazione periodica dei rischi sistemici che i loro servizi presentano per la società, compresa la libertà di espressione, o il rischio che i loro servizi siano usati come strumento per campagne di disinformazione, soprattutto al fine di tutelare i processi elettorali. Quale soluzione fondamentale per attenuare tali rischi[35], essi sono invitati a partecipare all'elaborazione di codici di condotta e protocolli di crisi volontari, di cui ne costituisce principale esempio il citato Codice "rafforzato" di buone pratiche sulla disinformazione del 2022, recentemente riconosciuto come "Codice di condotta" *ex* art. 45 del DSA ai fini dell'attenuazione dei rischi sistemici derivanti dalla disinformazione[36].

---

[34] Individua il susseguirsi di 3 stagioni regolatorie nell'approccio europeo al contrasto alla disinformazione (in termini di *auto-regulation*, *self-regulation* e regolamentazione "dall'alto") O. POLLICINO, *Disinformazione e intelligenza artificiale: "A Groovy Kind of Love?"*, in *Annuario di diritto comparato e di studi legislativi*, Napoli, 2025, pp. 239-276. In termini di "autoregolazione" e "approccio regolatorio leggero" v. S. PLATON, *La lutte contre la désinformation en Droit de l'Union européenne*, in C. MAUBERNARD, S. PLATON, R. TINIÈRE (cur.), *Les mutations de la liberté d'expression dans l'Union européenne*, *op. cit*., pp. 193-220; in tema v. anche E. SHATTOCK, *Lies, Liability, and Lawful Content: Citiquing the Approaches to Online Disinformation in the EU*, in *Common Market Law Review*, 2023, pp. 1313-1348. Si iscrivono nell'ultima "stagione regolatoria" anche il Regolamento (UE) 2024/900 del Parlamento europeo e del Consiglio, del 13 marzo 2024, *relativo alla trasparenza e al targeting della pubblicità politica*, in GU L 2024/900 del 20.3.2024 e il Regolamento (UE) 2024/1083 del Parlamento europeo e del Consiglio, dell'11 aprile 2024, *che istituisce un quadro comune per i servizi di media nell'ambito del mercato interno e che modifica la direttiva 2010/13/UE (regolamento europeo sulla libertà dei media)*, in GU L 2024/1083 del 17.4.2024.

[35] Sottolinea l'imposizione di obblighi di tipo "procedurale" piuttosto che di tipo "sostanziale", M. HUSOVEC, *The Digital Services Act's red line: what the Commission can and cannot do about disinformation*, in *Journal of Media Law*, 2024, n. 1, pp. 47-56; sul contrasto alla disinformazione *online* nel *Digital Services Act*, v. anche G. CAGGIANO, *Il contrasto alla disinformazione tra nuovi obblighi delle piattaforme* online *e tutela dei diritti fondamentali nel quadro del* Digital Service Act *e della co-regolamentazione*, in *Papers di diritto europeo*, 2021, n. 1, pp. 45-71; A. STROWEL, J. DE MEYERE, *The Digital Services Act: transparency as an efficient tool to curb the spread of disinformation on online platforms?*, in *Journal of Intellectual Property, Information Technology and Electronic Commerce Law*, 2023, n. 1, pp. 66-83; L. NANNINI, E. BONEL, D. BASSI, M.J. MAGGINI, *Beyond phase-in: assessing impacts on disinformation of the EU Digital Services Act*, in *AI and Ethics*, 2025, n. 5, pp. 1241-1269.

[36] Adottato il 15 febbraio 2025: Code_of_Conduct_on_Disinformation_f9bhfVbrSm6IEbiMmtGRVsLHZKA_112678.pdf. Infatti, l'art. 45 del DSA prevede la possibilità di elaborare, generalmente su invito della Commissione, codici di

È da osservarsi che già all'interno di tale Codice di buone pratiche è posta particolare attenzione sui rischi legati ai *deep fake* e ad altre tecniche manipolative, come *bot* e profili falsi e, di conseguenza, richiamata l'esigenza di trasparenza e di conformità alle regole fissate in un apposito atto legislativo disciplinante l'intelligenza artificiale, che è appunto in questo scritto preso specificatamente in considerazione[37].

L'*AI Act*[38] è, anch'esso, uno strumento regolatorio proprio dell'Unione europea[39] con cui si pone un tassello nella realizzazione di uno dei cinque "macro-obiettivi" stabiliti nell'articolo 3 del TUE laddove include la promozione del "progresso scientifico e tecnologico" tra le misure da promuovere al fine di realizzare un mercato interno libero, inclusivo e concorrenziale.

Occorre precisare che l'adozione di un regolamento in materia di intelligenza artificiale è finalizzata precipuamente a regolare il buon funzionamento del mercato interno dell'intelligenza artificiale attraverso regole armonizzate inerenti allo sviluppo, all'immissione in mercato e all'utilizzo di prodotti e servizi che ricorrono all'intelligenza artificiale, così come di sistemi di IA "*stand-alone*"[40]. In altre parole, si tratta di uno degli atti di diritto derivato adottato per il ravvicinamento delle legislazioni, il completamento e il buon funzionamento nel mercato interno, nella sua declinazione

condotta che prevedano, tra l'altro, l'adozione di misure specifiche di attenuazione del rischio e un quadro di comunicazione periodica sulle specifiche misure adottate e i relativi risultati.

[37] Individua quattro azioni dell'UE in termini di "*prohibition*", "*transparency*", "*risk management*", "*education*" nella regolazione dell'intelligenza artificiale, J. CUPAĆ, M. SIENKNECHT, *Regulate against the machine: how the EU mitigates AI harm to democracy*, in *Democratization*, 2024, n. 1, pp. 1067-1090.

[38] Per un'analisi esaustiva dell'impianto del regolamento 2024/1689, si vedano C. MUÑOZ GARCÍA, *Regulación de la inteligencia artificial en Europa*, Valencia, 2023; N. TH. NIKOLINAKOS, *EU Policy and Legal Framework for Artificial Intelligence, Robotics and Related Technologies-The AI Act*, Cham, 2023; nonché F. FERRI (a cura di), *L'Unione europea e la nuova disciplina sull'intelligenza artificiale: questioni e prospettive*, in *Rivista Quaderni AISDUE*, fasc. spec. 2/2024, reperibile *online*, in particolare M. INGLESE, *Il regolamento sull'intelligenza artificiale come atto per il completamento e il buon funzionamento del mercato interno?*, pp. 71-90; C. NECATI PEHLIVAN, *The EU Artificial Intelligence (AI) Act: An Introduction*, in *Global Privacy Law Review*, 2024, n. 1, pp. 31-42; R. PETRUSO, G. SMORTO, *Il Regolamento europeo sull'intelligenza artificiale: una prima lettura*, in *La Nuova Giurisprudenza civile commentata*, 2024, n. 1, pp. 989-1004.

[39] L'Unione europea ha anche firmato la Convenzione quadro del Consiglio d'Europa sull'intelligenza artificiale (Council of Europe Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law, Vilnius, 5.IX.2024). In merito, si rinvia a A. IERMANO, in questo numero della *Rivista*; nonché a F.P. LEVATINO, F. PAOLUCCI, *Advancing the Protection of Fundamental Rights Through AI Regulation: How the EU and the Council of Europe are Shaping the Future*, New York, 2024; G. ZACCARONI, *Of Artificial Intelligence and Fundamental Rights Charters: How AI Could Bridge the EU and the Council of Europe to Strengthen Fundamental Rights*, in J.G. WITSCHHOFF (ed.), *Europe's Fundation and its Future: The EU Charter in Focus*, Berlin, 2024, pp. 99-108; F. SEATZU, *Assessing the Council of Europe's AI Convention: Challenges and Prospects for Protecting Human Rights and Democracy in the Age of AI*, in *Studi sull'integrazione europea*, 2025, n. 1, pp. 9-22.

[40] Rileva la definizione di "sistema di IA" di cui all'art. 3 del regolamento 2024/1689, quale "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". Sull'eccessiva vaghezza di una nozione "onnicomprensiva" e sulle ricadute applicative, ci si limita a rinviare a quanto rilevato da G. CONTISSA, F. GALLI, *AI Act e diritti fondamentali: presupposti tecnologici e ricadute normative*, in *Quaderni costituzionali*, 2024, n. 3, pp. 738-741.

di "mercato unico digitale"[41], come suggerisce l'identificazione della base giuridica nell'articolo 114 del Trattato sul Funzionamento dell'Unione europea (TFUE)[42], centrale rispetto alla seconda base giuridica, l'articolo 16 TFUE, in relazione ai profili concernenti il trattamento e la tutela dei dati personali. La disciplina prevista è, pertanto, di stampo prevalentemente "economico", ossia finalizzata alla costituzione di un mercato unico per lo scambio e l'utilizzo di sistemi di intelligenza artificiale affidabili, il cui impiego, cionondimeno, è da realizzarsi nel rispetto dei valori e dei diritti fondamentali dell'Unione europea.

L'"approccio normativo orizzontale"[43], che connota il regolamento 2024/1689, recepisce l'esigenza, da una parte, di "accelerare il processo di sviluppo e immissione sul mercato dei sistemi di IA" e, dall'altra parte, di consentire che il cambiamento sia accettato dai singoli e dal settore imprenditoriale: sia i cittadini sia le imprese "devono poter avere fiducia nella tecnologia con cui interagiscono, disporre di un contesto normativo prevedibile e contare su efficaci misure di salvaguardia che proteggano i loro diritti e le loro libertà fondamentali"[44].

---

[41] Ossia quel mercato "[…] in cui è garantita la libera circolazione delle merci, delle persone, dei servizi e dei capitali e in cui, quale che sia la loro cittadinanza o nazionalità o il luogo di residenza, persone e imprese non incontrano ostacoli all'accesso e all'esercizio delle attività online in condizioni di concorrenza leale e potendo contare su un livello elevato di protezione dei consumatori e dei dati personali". In questi termini si è espressa la Commissione europea, nella sua Comunicazione al Parlamento europeo, al Consiglio, al Comitato economico e sociale europeo e al Comitato delle regioni, *Strategia per il mercato unico digitale in Europa*, del 06.05.2015, COM(2015) 192 final, p. 3. Nella revisione intermedia del 2017 emerge, per la prima volta, l'espressa necessità per l'UE di intervenire anche nel settore dell'IA e, nel 2018, la Commissione europea ha adottato una prima comunicazione sull'IA (Comunicazione della Commissione al Parlamento europeo, al Consiglio, al Comitato economico e sociale europeo e al Comitato delle Regioni, *L'intelligenza artificiale per l'Europa*, COM (2018) 237final, del 25 aprile 2018.) e poi il Libro bianco sull'intelligenza artificiale – Un approccio europeo all'eccellenza e alla fiducia, COM (2020) 65final, del 19 febbraio 2020. Da ultimo, con la proposta di regolamento è stata adottata la Comunicazione della Commissione al Parlamento europeo, al Consiglio, al Comitato economico e sociale europeo e al Comitato delle Regioni, *Promuovere un approccio europeo all'intelligenza artificiale*, COM (2021) 205final, del 21 aprile 2021. In generale, v. F. FERRI, *Il bilanciamento dei diritti fondamentali nel mercato unico digitale*, Torino, 2022.

[42] Sull'art. 114 TFUE come base giuridica si richiama lo studio di T.M. MOSCHETTA, *Il ravvicinamento delle normative nazionali per il mercato interno. Riflessioni sul sistema delle fonti alla luce dell'art. 114 TFUE*, Bari, 2018; nonché S. POLI, *Il rafforzamento della sovranità tecnologica europea e il problema delle basi giuridiche*, in *I* Post *di AISDUE*, 20 dicembre 2021, pp. 69-84.

[43] Ossia la regolamentazione dell'intelligenza artificiale nel suo complesso piuttosto che attraverso il ricorso alla regolazione delle applicazioni dell'IA in specifici settori o singole materie; cfr. G. FINOCCHIARO, *La regolazione dell'intelligenza artificiale*, in *Rivista trimestrale di diritto pubblico*, 2022, n. 4, pp. 1085-1099.

[44] Come recita la citata Comunicazione, *L'intelligenza artificiale per l'Europa*, p. 16. Sulla proposta di regolamento, ci si limita a rinviare all'analisi approfondita di A. ADINOLFI, *L'Unione europea dinnanzi allo sviluppo dell'intelligenza artificiale: la costruzione di uno schema di regolamentazione europeo tra mercato unico digitale e tutela dei diritti fondamentali*, in S. DORIGO (a cura di), *Il ragionamento giuridico nell'era dell'intelligenza artificiale*, Pisa, 2020, pp. 13-36; G. CONTALDI, *La proposta di regolamento sull'intelligenza artificiale e la protezione di dati personali*, in G. CAGGIANO, G. CONTALDI, P. MANZINI (a cura di), *Verso una legislazione europea su mercati e servizi digitali*, Bari, 2022, p. 205 ss.; F. DONATI, *Diritti fondamentali e algoritmi nella proposta di regolamento sull'intelligenza artificiale*, in *Il Diritto dell'Unione europea*, 2021, n. 3-4, pp. 453-466; A. ODDENNINO, *Intelligenza artificiale e tutela dei diritti fondamentali: alcune notazioni critiche sulla recente Proposta di Regolamento della UE, con particolare riferimento all'approccio basato sul rischio e al pericolo di discriminazione algoritmica*, in A. PAJNO, F.

È in questa prospettiva che l'obiettivo dichiarato del legislatore europeo si concretizza nella progettazione, sviluppo e distribuzione dell'intelligenza artificiale secondo "le proprie modalità e i propri valori"[45]: un'IA "sicura, affidabile ed etica"[46], incentrata sul rispetto dei valori fondanti. Siffatto tentativo di contemperamento delle istanze di mercato e delle istanze valoriali trova, poi, specifico riflesso nella tecnica legislativa adottata che, oltrepassando la mera logica del mercato, amplifica il rilievo per la protezione dei diritti fondamentali, nonché dei processi democratici e dello Stato di diritto.

Assecondando tale prospettiva, il regolamento 2024/1689 introduce un sistema di classificazione proporzionale del rischio (*risk-based approach*)[47] su cui graduare, con diverse intensità, una pluralità di obblighi tecnico-giuridici incombenti sia sugli "operatori" dei sistemi di intelligenza artificiale, sia sugli "utilizzatori" (siano pubbliche amministrazioni o persone fisiche e giuridiche di natura privata), nonché sugli Stati membri. Si tratta, dunque, di un approccio regolatorio basato sul livello di rischio dei sistemi, concernente sia sistemi basati su algoritmi deterministici sia ML (sistemi con diversi gradi di autonomia) che opera una distinzione tra categorie di applicazioni vietate (rischio inaccettabile), ad alto rischio (soggette a verifica di conformità secondo i requisiti fissati dal regolamento) e a rischio limitato o minimo (per le quali sono previsti soltanto oneri di informazione e l'adesione volontaria a codici di condotta), cui si aggiunge una quarta categoria basata sul "*trasparency risk*".

3.1. La previsione dell'intelligenza artificiale manipolatoria tra le pratiche vietate

Alla luce della logica di classificazione del "rischio", in considerazione della netta contrarietà con i valori dell'Unione europea – quali il rispetto della dignità umana, la libertà, l'uguaglianza, la democrazia e lo Stato di diritto – e con i diritti sanciti dalla Carta dei diritti fondamentali, compresi il diritto alla non discriminazione, alla protezione dei dati e alla vita privata e i diritti dei minori, l'*AI Act* individua una serie di

DONATI, A. PERRUCCI (a cura di), *Intelligenza artificiale e diritto: una rivoluzione? Diritti fondamentali, dati personali e regolazione*, vol. I, Bologna, 2022, pp. 135-171.

[45] COM(2018) 237 final, cit., p. 21.

[46] Cfr. considerando 8.

[47] Come espressamente dichiara il considerando 26 del regolamento 2024/1689. Il "rischio" è definito nell'art. 3, par. 2, del regolamento come "la combinazione della probabilità del verificarsi di un danno e la gravità del danno stesso". Sul *risk-based regulatory approach* si veda G. DE GREGORIO, P. DUNN, *The European risk-based approaches: Connecting constitutional dots in the digital age*, in *Common Market Law Review*, 2022, n. 2, pp. 473-500; J. CHAMBERLAIN, A. KOTSIOS, *Defining risk and Promoting Trust in AI Systems*, in M. BERGSTRÖM, V. MITSILEGAS (eds.), *EU Law in the Digital Age. Swedish Studies in European Law*, Oxford, 2025, pp. 105-122; in una prospettiva critica, v. anche C. NOVELLI, F. CASOLARI, A. ROTOLO, M. TADDEO, L. FLORIDI, *AI Risk Assessment: A Scenario-Based, Proportional Methodology for the AI Act*, in *Digital Society*, 2024, pp. 1-29.

pratiche considerate "vietate" [48].

In particolare, l'art. 5 del regolamento 2024/1689 fissa un "numero chiuso" di pratiche in cui il divieto è strutturato in termini di proibizione assoluta, ossia in maniera indipendente da una qualsiasi valutazione in concreto del rischio che possa essere effettuata da parte dei fornitori o degli utilizzatori di tali sistemi (definiti «*deployer*» dall'art. 3, par. 1, n. 4), secondo un apprezzamento da cui non è consentito discostarsi nemmeno con il consenso degli stessi destinatari del risultato finale elaborato dal sistema (il c.d. *output*)[49].

Orbene, al fine della presente indagine, volta a individuare il contributo dell'*AI Act* nel perseguire gli obiettivi di contrasto alla disinformazione digitale, può venire in rilievo l'inserimento tra le pratiche vietate di tre tipi di tecniche manipolative, atteso che sono vietate, ai sensi dell'articolo 5, lett. a) del regolamento 2024/1689, l'immissione sul mercato, la messa in servizio o l'uso di sistemi di intelligenza artificiale che utilizzano tecniche subliminali che agiscono senza che una persona ne sia consapevole o tecniche volutamente manipolative o ingannevoli aventi lo scopo o l'effetto di distorcere materialmente il comportamento di una persona o di un gruppo di persone, pregiudicando in modo considerevole la loro capacità di prendere una decisione informata, inducendole pertanto a prendere una decisione che non avrebbero altrimenti preso, in un modo che provochi o possa ragionevolmente provocare a tale persona, a un'altra persona o a un gruppo di persone un danno significativo.

Tra le diverse condizioni cumulative che devono essere soddisfatte, l'elemento, *in primis*, di difficile definizione è la qualificazione stessa di "tecniche subliminali", "tecniche manipolative intenzionali" e "tecniche ingannevoli", in assenza di precise indicazioni all'interno dell'*AI Act*. Per quanto concerne le "tecniche subliminali"[50] può prendersi, invero, in considerazione il considerando n. 29 del regolamento 2024/1689, e il riferimento ivi contenuto all'utilizzo di "stimoli audio, grafici e video che le persone non sono in grado di percepire poiché tali stimoli vanno al di là della percezione umana e che hanno la capacità di sovvertire o pregiudicare l'autonomia, il processo decisionale o la libera scelta di una persona senza che sia consapevole di tali tecniche". Non dovrebbe, tuttavia, trattarsi di un numero chiuso di ipotesi, arricchite, peraltro, dalle indicazioni fornite nelle Linee guida in vigore dal 2 febbraio 2025 adottate sulla base dell'art. 96, par. 1, lett. b) del regolamento *de quo* (che tuttavia non presentano carattere

[48] Specificatamente sulle pratiche vietate, con riferimento alla proposta di regolamento, R.J. NEUWIRTH, *Prohibited Artificial Intelligence Practices in the Proposed EU Artificial Intelligence Act*, in *Computer Law & Security Review*, 2023.

[49] In ipotesi di pratiche vietate, neanche la Commissione europea ha le competenze normative e di esecuzione volte a rendere flessibile tale elenco rivedendo i divieti in parola, diversamente da quanto previsto con riferimento ai sistemi di intelligenza artificiale definiti ad "alto rischio". La non conformità al divieto delle pratiche di IA di cui all'articolo 5 è soggetta a sanzioni amministrative pecuniarie (fino a € 35.000.000 o, se l'autore del reato è un'impresa, fino al 7 % del fatturato mondiale totale annuo dell'esercizio precedente, se superiore) così come previsto dall'art. 99 dell'*AI Act*.

[50] Sulla difficoltà di definire le "tecniche subliminali" v. anche R.J. NEUWIRTH, *The EU Artificial Intelligence Act. Regulating Subliminal AI System*, Abingdon, 2023.

vincolante)[51]; tali Linee guida tengono conto della velocità con cui si sviluppano questi sistemi, anche di tipo "interfacce cervello-computer" o di "realtà virtuale", che consentono un livello più elevato di controllo degli stimoli presentati alle persone, nella misura in cui sono in grado di determinarne una distorsione materiale del comportamento in modo significativamente nocivo.

Si potrebbe, in conclusione, considerare che, allorquando si faccia riferimento a "tecniche subliminali" che agiscono senza che una persona ne sia consapevole, si voglia stigmatizzare tutto ciò che, nel contenuto, è in grado di agire sulle scelte delle persone in maniera inconsapevole, laddove nell'ipotesi di "tecniche volutamente manipolative o ingannevoli" si presti attenzione non soltanto al contenuto o all'effetto della tecnica ma anche al mero "intento" ossia a un tentativo di determinare una manipolazione o un inganno.

Orbene, l'elemento dell'intento non è un elemento necessario, considerato che l'art. 5, par. 1, lett. a) dell'*AI Act* copre anche quelle pratiche che abbiano il solo "effetto" di causare una distorsione. Tuttavia, è necessaria la sussistenza di un impatto sostanziale sul comportamento della persona e sulla sua autonomia di libera scelta che sia minata, non potendosi considerare come sufficiente una semplice "influenza" sul comportamento; in altre parole, è necessario un nesso causale plausibile o ragionevolmente probabile tra la potenziale distorsione materiale del comportamento e la tecnica subliminale, intenzionalmente manipolativa o ingannevole impiegata dal sistema di intelligenza artificiale.

Inoltre, la previsione normativa fa riferimento alla causazione – o ragionevole causazione – di un "danno significativo": all'uopo i principali tipi di danni rilevanti ai sensi dell'articolo 5, par. 1, lett. a) paiono ricomprendere i danni fisici, psicologici, finanziari ed economici, alla luce del considerando n. 29 del regolamento 2024/1689 che a essi fa espressamente riferimento; laddove incorrono nel divieto quelle tecniche che implichino evidentemente impatti negativi "significativi" sulla salute fisica, psicologica o sugli interessi finanziari di persone o gruppi di persone.

La presenza di siffatti elementi determina, allora, un ridimensionamento della formulazione in termini assoluti della previsione di cui all'art. 5 dell'*AI Act*, atteso il venire in rilievo di specifiche caratteristiche delle pratiche di intelligenza artificiale che incorrono nel divieto. I divieti, infatti, sono costruiti su una serie di limiti che ampliano le maglie dei sistemi che si sottraggono al divieto stesso e il cui utilizzo è consentito da parte di attori sia pubblici sia privati, palesando che l'inaccettabilità non è connessa al sistema di IA in quanto tale, quanto piuttosto al suo utilizzo o all'impiego eccessivamente rischioso a cui esso può prestarsi.

[51] Communication to the Commission, *Approval of the Content of the draft Communication from the Commission – Commission Guidelines on prohibited artificial intelligence practices established by Regulation (EU) 2024/1689 (AI Act)*, Brussels, 4.2.2025, C(2025) 884 final.

3.2. Le previsioni riguardanti i sistemi ad alto rischio

Diversamente dai sistemi di intelligenza artificiale che determinano rischi inaccettabili, la qualificazione di un sistema di IA come "ad alto rischio" non comporta un divieto, quanto piuttosto il sorgere dell'obbligo di ricondurre i rischi generati da questi sistemi entro livelli accettabili così da trovare un equilibrio tra gli interessi in gioco, anche in ragione dei costi che la regolamentazione comporta.

Nell'ambito della categoria dei sistemi di intelligenza artificiale ad alto rischio, il regolamento 2024/1689 pone una differenziazione, in relazione alla disciplina applicabile, tra due sottogruppi, distinguibili a seconda che l'intelligenza artificiale sia incorporata, anche come componente di sicurezza, in uno dei prodotti la cui fabbricazione è già regolamentata dalla normativa di armonizzazione (ad esempio, giocattoli, ascensori, imbarcazioni da diporto, dispositivi medici o di protezione individuale) oppure che sia progettata per essere utilizzata come «elemento indipendente» (c.d. sistemi "*stand-alone*").

Con riferimento a questa seconda categoria, rileva la considerazione quali "ad alto rischio" di quei sistemi di intelligenza artificiale "destinati a essere utilizzati per influenzare l'esito di un'elezione o di un referendum o il comportamento di voto delle persone fisiche nell'esercizio del loro voto alle elezioni o ai referendum"[52]. Rispetto a tali sistemi di IA trova, pertanto, applicazione il sistema di "gestione dei rischi" strutturato dall'*AI Act*[53], nonché la previsione di obblighi in capo ai fornitori e ai *deployer* degli stessi sistemi; in particolare, rileva l'imposizione, in capo ai (soli) utilizzatori di sistemi di intelligenza artificiale – siano essi organismi di diritto pubblico o enti privati che forniscono servizi pubblici – di effettuare una valutazione dell'impatto che tali tecnologie possono avere sui diritti fondamentali.

Tale "valutazione d'impatto sui diritti fondamentali"[54] costituisce un elemento centrale nell'analisi dei rischi connessi all'uso dell'intelligenza artificiale che potrebbe

[52] Punto 8, lett. b) dell'Allegato III all'*AI Act*, che individua i *Sistemi di IA ad alto rischio di cui all'articolo 6, paragrafo 2*; restano espressamente esclusi i sistemi di IA ai cui *output* le persone fisiche non sono direttamente esposte, come gli strumenti utilizzati per organizzare, ottimizzare e strutturare le campagne politiche da un punto di vista amministrativo e logistico, come precisato nello stesso Allegato III.

[53] Inteso, ai sensi dell'art. 9, par. 2, del regolamento 2024/1689, come "un processo iterativo continuo pianificato ed eseguito nel corso dell'intero ciclo di vita di un sistema di IA ad alto rischio, che richiede un riesame e un aggiornamento costanti e sistematici", comprendente 4 fasi: a) identificazione e analisi dei rischi noti e ragionevolmente prevedibili che il sistema di IA ad alto rischio può porre per la salute, la sicurezza e i diritti fondamentali quando il sistema di IA ad alto rischio è utilizzato conformemente alla sua finalità prevista; b) stima e valutazione dei rischi che possono emergere quando il sistema di IA ad alto rischio è usato conformemente alla sua finalità prevista e in condizioni di uso improprio ragionevolmente prevedibile; c) valutazione di altri eventuali rischi derivanti dall'analisi dei dati raccolti dal sistema di monitoraggio successivo all'immissione sul mercato di cui all'articolo 72; d) adozione di misure di gestione dei rischi opportune e mirate intese ad affrontare i rischi individuati ai sensi della lettera a).

[54] Prevista dall'art. 27 del regolamento *de quo*, la valutazione d'impatto sui diritti fondamentali deve comprendere: una descrizione dei processi del *deployer* in cui il sistema di IA ad alto rischio sarà utilizzato in linea con la sua finalità prevista; una descrizione del periodo di tempo entro il quale ciascun sistema di

essere impiegata per la diffusione di disinformazione *online*. Sebbene siffatto obbligo non trovi applicazione rispetto agli attori privati – comprese le piattaforme *online*, che, ad ogni buon conto, restano assoggettate agli obblighi di valutazione del rischio previsti dal *Digital Services Act* – esso rappresenta una nitida espressione dell'approccio europeo volto a una progressiva responsabilizzazione degli operatori che impiegano sistemi di intelligenza artificiale[55].

Tuttavia, occorre precisare che i sistemi di intelligenza artificiale "destinati a essere utilizzati per influenzare l'esito di un'elezione o di un referendum o il comportamento di voto delle persone fisiche nell'esercizio del loro voto alle elezioni o ai referendum" di cui si tratta, rientrano nell'art. 6, par. 2, del regolamento 2024/1689, ai quali si applica la precisazione contenuta al par. 3 che, in deroga al par. 2, non considera un sistema di IA di cui all'allegato III "ad alto rischio" se non presenti un "rischio significativo di danno" per la salute, la sicurezza o i diritti fondamentali delle persone fisiche. In particolare, poi, tale deroga riguarda i casi in cui il sistema non influenzi "materialmente il risultato del processo decisionale".

Ciò implica la previsione di una valutazione di "significatività" del danno che si muove nel senso di mitigare la predeterminazione del rischio, attraverso l'introduzione di un concetto di cui resta vago il criterio di misurazione che, ad ogni buon conto, pare debba essere precipuamente rapportato alla salute, alla sicurezza o ai diritti fondamentali delle persone fisiche e non già a un "danno sociale" o "collettivo" quale può essere la disinformazione a detrimento del processo democratico.

3.3. L'introduzione di obblighi di trasparenza per "altre pratiche manipolative"

In ultimo, a prescindere dalla qualificazione del sistema come ad alto rischio e programmato per finalità specifiche o per finalità generali, rilevano i requisiti di trasparenza che riguardano quei sistemi di IA che potrebbero comportare rischi di impersonificazione o inganno, in considerazione del fatto che siffatti sistemi potrebbero arrecare danno ai consumatori o minacciare i processi democratici, diffondendo su vasta scala informazioni false o manipolate[56].

IA ad alto rischio è destinato a essere utilizzato e con che frequenza; le categorie di persone fisiche e gruppi verosimilmente interessati dal suo uso nel contesto specifico; i rischi specifici di danno che possono incidere sulle categorie di persone fisiche o sui gruppi di persone; una descrizione dell'attuazione delle misure di sorveglianza umana, secondo le istruzioni per l'uso; le misure da adottare qualora tali rischi si concretizzino, comprese le disposizioni relative alla governance interna e ai meccanismi di reclamo.

[55] Mette in risalto l'elevato rischio di "burocratizzazione" della "complessa impalcatura finalizzata alla gestione del rischio ed alla valutazione dell'impatto in materia di diritti fondamentali", una volta che la Commissione e le autorità di vigilanza avranno dato attuazione alle parti dell'*AI Act* che necessitano di atti di diritto derivato, G. ZACCARONI, *Intelligenza artificiale e principio democratico: riflessioni a margine dell'emersione di un quadro normativo europeo*, in *Quaderni AISDUE*, fasc. spec. n. 2/2024, pp. 19-50, spec. p. 44.

[56] Cfr. considerando 133 dell'*AI Act*.

Al fine di contrastare tali rischi, l'*AI Act* stabilisce, in primo luogo, in caso di sistemi destinati a interagire direttamente con le persone fisiche (come per i *chatbot*) l'obbligo in capo ai fornitori di progettare e sviluppare siffatto sistema in modo tale che i singoli siano informati del fatto di stare interagendo con un sistema di intelligenza artificiale, salvo che ciò risulti evidente "dal punto di vista di una persona fisica ragionevolmente informata, attenta e avveduta, tenendo conto delle circostanze e del contesto di utilizzo"[57].

In tal modo, trova previsione una "clausola di esenzione" che, seppure possa considerarsi coerente con il principio di proporzionalità, può prestarsi a interpretazioni restrittive in grado di indebolire l'efficacia degli obblighi informativi previsti dal regolamento 2024/1689. Invero, la codificazione di siffatti obblighi avrebbe potuto essere improntata ad un approccio più stringente anche abbracciando i principi emersi dalla giurisprudenza della Corte di giustizia dell'UE che ha evidenziato l'importanza di notifiche chiare, "sufficientemente precise e dimostrate"[58], specie per consentire alle piattaforme *online* di reagire tempestivamente alla diffusione di contenuti illeciti, senza dover condurre un esame giuridico approfondito per ogni singola segnalazione, rafforzando l'obbligo di trasparenza nei confronti degli utenti, in particolare in caso di contenuti ingannevoli, come i *deep fake* utilizzati a fini di disinformazione elettorale. La previsione di requisiti di notifica più stringenti e standardizzati avrebbe avuto il pregio di promuovere una maggiore responsabilizzazione delle piattaforme e di offrire ai singoli strumenti di tutela più efficaci, garantendo comunque un equilibrio tra il contrasto alla diffusione di contenuti disinformativi e la tutela della libertà di espressione.

In secondo luogo, i fornitori sono tenuti a mitigare i rischi di impersonificazione o di inganno discendenti dalla generazione, da parte dei rispettivi sistemi di intelligenza artificiale, di contenuti audio, immagini, video o testuali sintetici: tali operatori sono obbligati a marchiare l'*output* prodotto dal sistema in un formato leggibile meccanicamente, attraverso soluzioni tecniche efficaci, interoperabili, solide ed efficaci e a rilevarne l'origine artificiale[59].

Infine, obblighi di trasparenza sono previsti anche per i *deployer*: rileva, in particolare, l'obbligo per gli utilizzatori di un sistema di IA capace di generare o

[57] Come recita l'art. 50, par. 1 del regolamento 2024/1689. In tema di regolamentazione dei *deep fake*, si veda A. ORLANDO, *La regolamentazione del deepfake in Europa, Stati Uniti e Cina*, in *Media Laws*, 2025, reperibile *online*; M. CAZZANIGA, *Una nuova tecnica (anche) per veicolare disinformazione: le risposte europee ai* deepfakes, in *Media Laws*, 2023, pp. 170-187; F. ROMERO MORENO, *Generative AI and deepfakes: a human rights approach to tackling harmful content*, in *International Review of Law, Computers & Technology*, 2024, n. 3, pp. 297-326.

[58] Corte di giustizia dell'UE, Grande Sezione, sentenza del 22 giugno 2021, *Frank Peterson* c. *Google LLC e a*. e *Elsevier Inc*. c. *Cyando AG*, cause riunite C-682/18 e C-683/19, ECLI:EU:C:2021:503, punti 115-116, in relazione all'interpretazione dell'art. 14, par. 1, della direttiva sul commercio elettronico (direttiva 2000/31/CE del Parlamento europeo e del Consiglio, dell'8 giugno 2000, *relativa a taluni aspetti giuridici dei servizi della società dell'informazione, in particolare il commercio elettronico, nel mercato interno*, in GU 2000, L 178, p. 1).

[59] Ai sensi dell'art. 50, par. 2 del regolamento 2024/1689.

manipolare immagini o contenuti audio o video che assomigliano notevolmente a persone, oggetti, luoghi, entità o eventi esistenti e che potrebbero apparire falsamente autentici o veritieri (ossia, i cd. *deep fake*) di rendere noto in modo chiaro e distinto che il contenuto è stato generato o manipolato artificialmente, etichettando di conseguenza i relativi *output* e rivelandone l'origine artificiale[60]. Allo stesso modo, i *deployer* che utilizzano un sistema di intelligenza artificiale per generare o manipolare testi allo scopo di informare su questioni di interesse pubblico sono soggetti ad un medesimo obbligo di trasparenza, ad eccezione dell'ipotesi in cui il contenuto generato dall'intelligenza artificiale sia sottoposto a un processo di revisione umana o a un controllo editoriale e una persona fisica o giuridica sia responsabile della sua pubblicazione.

Scopo della norma nel suo complesso, come si è analizzato, è precipuamente quello di minimizzare l'impatto che l'intelligenza artificiale può avere a livello informativo sui cittadini dell'Unione, atteso che il legislatore europeo mira proprio a ridurre "i nuovi rischi di cattiva informazione e manipolazione su vasta scala", come espressamente recita il considerando 133 dell'*AI Act*. Tuttavia, pur prevedendo obblighi di trasparenza vincolanti in materia di etichettatura e di rilevamento dei *deep fake*, ad essi si associa – sulla base del considerando 135 e dell'articolo 50, par. 7, del regolamento 2024/1689 – la necessità di un supporto operativo (che, peraltro, ricalca il modello di co-regolamentazione sposato dal *Digital Services Act*) che, pur ponendosi l'obiettivo di conformazione da parte dei destinatari, non è accompagnato da forme di coercizione diretta ed è suscettibile di ingenerare preoccupazioni in termini di trasparenza e chiarezza giuridica[61].

Si intende fare riferimento a quanto disposto dal citato articolo 50, par. 7, dell'*AI Act*, che pone in capo all'Ufficio per l'IA il compito di incoraggiare e agevolare l'elaborazione di codici di buone pratiche a livello dell'Unione: tali codici hanno proprio lo scopo di semplificare l'efficace attuazione degli obblighi relativi alla rilevazione e all'etichettatura dei contenuti generati o manipolati artificialmente, da parte sia dei fornitori sia degli utilizzatori. In secondo luogo, la Commissione può adottare atti di esecuzione per approvare tali codici di buone pratiche secondo la procedura di cui all'articolo 56, par. 6, dell'*AI Act*, mantenendo la facoltà di introdurre atti di esecuzione vincolanti qualora i codici risultassero inadeguati, secondo le procedure specifiche previste dallo stesso *AI Act*. Inoltre, ai sensi dell'articolo 96, par. 1, lett. d) dell'*AI Act*, la Commissione europea è tenuta a elaborare orientamenti sull'attuazione pratica**,**

[60] Ai sensi dell'art. 50, par. 4 del regolamento. Un obbligo di trasparenza "attenuato" è, tuttavia, previsto con riferimento ai *deep fake* che facciano parte di «un'opera o di un programma manifestamente creativo, satirico, artistico o fittizio»: in tali casi, si legge al considerando 134, «l'obbligo di trasparenza per i deep fake di cui al presente regolamento si limita alla rivelazione dell'esistenza di tali contenuti generati o manipolati in modo adeguato che non ostacoli l'esposizione o il godimento dell'opera, compresi il suo normale sfruttamento e utilizzo, mantenendo nel contempo l'utilità e la qualità dell'opera».

[61] Si vedano anche le preoccupazioni espresse dal Parlamento europeo, *Better regulation and the improvement of EU regulatory environment. Institutional and legal implications of the use of "soft law" instruments*, March 2007.

specificamente volti a supportare l'attuazione degli obblighi di trasparenza sui *deep fake* previsti dall'articolo 50, a beneficio di fornitori e utilizzatori.

4. Considerazioni conclusive: contributo e limiti dell'*AI Act* nel contrasto alla disinformazione digitale e alla protezione dei valori democratici dell'UE

Nell'*AI Act* adottato dall'Unione europea è possibile rinvenire, in conclusione, un iniziale tentativo di regolamentazione dei sempre più sofisticati strumenti che alimentano fenomeni di disinformazione e manipolazione dell'opinione pubblica, primi tra tutti i *deep fake*. Come si è analizzato in questo contributo, i sistemi di creazione di questi ultimi non rientrano nella classificazione né di sistemi proibiti né di sistemi ad alto rischio; piuttosto, il regolamento 2024/1689 introduce obblighi di trasparenza, imponendo la marcatura dei contenuti sintetici in un formato leggibile da macchine e il riconoscimento della loro natura artificiale, attraverso previsioni che pure lasciano emergere alcune lacune e criticità che qui si intendono evidenziare.

In primo luogo, dalla strutturazione dell'*AI Act* emerge un regime di responsabilità ricadente principalmente in capo ai *deployer* (ossia, come si è detto, gli utilizzatori dei sistemi), atteso che i *provider* (gli sviluppatori) risultano soggetti soltanto a obblighi di etichettatura non direttamente visibili al pubblico: approccio che rischia di risultare inefficace nel contrastare attori intenzionati malevolmente a creare e a diffondere *deep fake* con fini manipolatori. Per di più, la lacuna evidenziata può trovare accentuazione in considerazione della previsione della cd. "*household exemption*"[62] che implica l'esclusione dall'ambito di applicazione del regolamento 2024/1689 dell'utilizzo di natura non professionale e che, dunque, non tiene debitamente conto della capacità dei *deep fake* – pur se originariamente creati per scopi "privati" – di diffondersi in maniera virale ed esponenziale. Infine, sotto il profilo dell'*enforcement*, può considerarsi che l'*AI Act* risulta incentrato – come analizzato *supra* al par. 3.3. – sulla previsione di obblighi di trasparenza quale principale strumento di mitigazione del rischio, senza aprirsi alla previsione di misure più incisive, quale può essere l'obbligo di rimozione dei contenuti manipolati o l'introduzione di sanzioni da comminare in capo a coloro che svolgono attività di diffusione dei *deep fake* con intenti manipolatori.

Eventualmente, parte dei contenuti manipolati dall'intelligenza artificiale aventi incidenza sul processo democratico potranno venire in rilievo come sistemi "ad alto rischio", atteso che il regolamento 2024/1689 contempla in tale categoria quei sistemi di intelligenza artificiale "destinati a essere utilizzati per influenzare l'esito di un'elezione o di un referendum o il comportamento di voto delle persone fisiche

[62] Il regolamento non si applica, infatti, agli obblighi dei *deployer* che sono persone fisiche che utilizzano sistemi di IA nel corso di un'attività non professionale puramente personale (art. 10, par. 2, dell'*AI Act*). Per una valutazione "insufficiente" della disciplina europea in materia di *deep fake*, si veda anche M.J. BLOCK, *A Critical Evaluation of Deepfake Regulation through the AI Act in the European Union*, in *Journal of European Consumer and Market Law*, 2024, n. 4, pp. 184-192.

nell'esercizio del loro voto alle elezioni o ai referendum". Tuttavia, nel considerare il contributo dell'*AI Act* nel contrasto a fenomeni come la manipolazione e la disinformazione digitale, è da valutare il limite derivante dall'adozione di una logica di valutazione del rischio – mutuata dalla sicurezza dei prodotti[63] – che, seppure coerente con la logica del mercato interno e della protezione dei diritti fondamentali della persona, non appare pienamente adeguata a fronteggiare minacce "sistemiche" che fenomeni come la disinformazione rappresentano per la democrazia in quanto tale.

Il principale limite che, in questo contesto, si ritiene ascrivibile all'*AI Act*, è legato alla centralità riconosciuta all'emersione di un "danno significativo", riconnesso alla produzione di effetti – fisici, psicologici, economici – misurabili, quale criterio principe per l'applicazione dei divieti e degli obblighi normativi previsti dal regolamento 2024/1689, piuttosto che al riconoscimento di danni "diffusi" o "sociali", quali potrebbero essere considerati la polarizzazione politica o l'erosione della fiducia nelle istituzioni, che, dal canto loro, tuttavia, risultano essere di difficile identificazione, misurazione e quantificazione.

Siffatto approccio, si osservi, connota non soltanto i sistemi di intelligenza artificiale "ad alto rischio" ma anche il divieto di intelligenza artificiale manipolatoria (fissato dall'art. 5 dell'*AI Act*, come si è analizzato *supra* al par. 3.1.) e determina il rischio di escludere dalla regolamentazione diverse forme di manipolazione dell'informazione, in specie quelle che non trovano traduzione in una lesione diretta della persona, e che, tuttavia, possono avere conseguenze più ampie sul funzionamento democratico. La disinformazione, infatti, non colpisce solo gli individui, ma opera su scala collettiva, alterando il dibattito pubblico e interferendo nei processi elettorali, come emerge dai recenti episodi di campagne di disinformazione *online* in Europa. È tenendo conto di tale fattore, nonché della definizione di "disinformazione *online*" quale fenomeno in grado di determinare un "pregiudizio pubblico" (come pure si è specificato *supra* al par. 2) che il Comitato economico e sociale europeo ha – nella fase di negoziazione del regolamento – suggerito di ampliarne l'art. 5, par. 1, lett. a) nel senso di includere pratiche manipolative che causano "pregiudizio ai diritti fondamentali, compresi quelli alla salute e alla sicurezza fisiche o psicologiche, di un'altra persona o di un gruppo di persone, o alla democrazia e allo Stato di diritto"[64]. Suggerimento che, invero, non ha trovato accoglimento all'interno di un testo normativo che, per quanto si è analizzato, risulta orientato alla protezione dei diritti "individuali" della persona più che sulla salvaguardia diretta della democrazia come "bene collettivo", nonché improntato alla strategia di contrasto alla disinformazione propria del *Digital Services Act*, fondata, più che su logiche "repressive", sul tentativo di bilanciare i diritti e le libertà "costituzionali"

[63] Approfondiscono questa impostazione M. ALMADA, N. PETIT, *The EU AI Act: Between the rock of product safety and the hard place of fundamental rights*, in *Common Market Law Review*, 2025, pp. 85-120.

[64] Cfr. Parere del Comitato economico e sociale europeo sulla proposta di regolamento del Parlamento europeo e del Consiglio che stabilisce regole armonizzate sull'intelligenza artificiale (legge sull'intelligenza artificiale) e modifica alcuni atti legislativi dell'Unione, COM(2021) 206 final – 2021/0106 (COD).

potenzialmente in conflitto nello spazio digitale, facendo principalmente leva su meccanismi di trasparenza, co-regolamentazione e responsabilizzazione degli attori coinvolti. E ciò tenendo conto che, anche quando le tecnologie di intelligenza artificiale sono coinvolte nella produzione e diffusione di contenuti artificiali, resta il nodo centrale del ruolo che la libertà di espressione occupa nelle società democratiche: una libertà che, peraltro, non gode dello stesso livello di protezione in tutti gli ordinamenti degli Stati membri dell'UE[65]. In tale contesto, l'eventuale introduzione di divieti più ampi e indeterminati, ad esempio riferiti alla semplice diffusione di *deep fake*, se non ancorati a un danno concreto e identificabile rispetto al bene giuridico tutelato rischierebbe di non essere compatibile con gli standard europei e internazionali di protezione della libertà di espressione[66].

ABSTRACT: Dopo aver esaminato l'evoluzione dell'"approccio europeo" al contrasto alla disinformazione digitale, il contributo si concentra sulla disciplina introdotta dal regolamento 2024/1689 – come il divieto di pratiche manipolative, la regolazione dei sistemi ad alto rischio e gli obblighi di trasparenza – analizzando il contributo (e i limiti) dell'*AI Act* nel contrasto alla disinformazione digitale e nella tutela dei valori democratici dell'Unione europea.

KEYWORDS: Intelligenza artificiale – disinformazione – democrazia – valori europei – libertà di espressione.

THE "EUROPEAN APPROACH" TO COUNTERING DIGITAL DISINFORMATION AND SAFEGUARDING DEMOCRATIC VALUES: WHAT CONTRIBUTION FROM THE AI ACT?

ABSTRACT: Building on an analysis of the evolution of the "European approach" to countering digital disinformation, the paper focuses on the regulatory framework introduced by Regulation 2024/1689 – including the prohibition of manipulative

[65] In tal senso, cfr. quanto emerge da Corte di giustizia, Grande Sezione, sentenza del 24 settembre 2019, causa C-507/17, *Google LLC*, succeduta alla *Google Inc*. c. *Commission nationale de l'informatique et des libertés*, ECLI:EU:C:2019:772, spec. punto 67.

[66] Con riferimento agli *standard* internazionali, si veda quanto espresso da The United Nations (UN) Special Rapporteur on Freedom of Opinion and Expression, the Organization for Security and Co-operation in Europe (OSCE) Representative on Freedom of the Media, the Organization of American States (OAS) Special Rapporteur on Freedom of Expression and the African Commission on Human and Peoples' Rights (ACHPR) Special Rapporteur on Freedom of Expression and Access to Information, nella *Joint Declaration on freedom of expression and "fake news", disinformation and propaganda*, 2017, disponibile *online*: https://www.osce.org/files/f/documents/6/8/302796.pdf.

practices, the regulation of high-risk systems, and transparency obligations – analyzing the contribution (and the limits) of the AI Act in addressing digital disinformation and safeguarding the democratic values of the European Union.

KEYWORDS: Artificial Intelligence – Disinformation – Democracy – European Values – Freedom of Expression.