Ciberspazio e diritto, vol. 26, n. 79 (1 - 2025), pp. 11-27

La verità ai tempi dell'IA tra rischi e nuove sfide

Arianna Arini*

Sommario: 1. Premessa. – 2. L'impatto dell'intelligenza artificiale sulla disinformazione. – 2.1. Strumenti di intelligenza artificiale che generano disinformazione. – 2.2. Rischi e criticità. – 2.3. Strumenti giuridici in soccorso. – 3. Intelligenza artificiale e Verità: una possibile alleanza contro le fake news.

1. *Premessa*

È ormai parte dell'immaginario collettivo l'immagine diffusa sul Web ritraente Papa Francesco in abiti firmati, o di Donald Trump intento a baciare Kamala Harris o a consegnarsi all'ufficio del Procuratore distrettuale di Manhattan a New York per rispondere delle accuse penali contestate. Immagini incredibilmente realistiche, rappresentanti eventi che, tuttavia, non si sono mai verificati. Ciononostante, la viralità della loro circolazione nell'ambiente virtuale ha sicuramente gettato le basi per un'analisi di un fenomeno ormai ampiamente diffuso, ovvero quello delle immagini, notizie, informazioni fuorvianti, falsificate e ingannevoli, prodotte da un'intelligenza artificiale che è in grado oggi di manipolare e distorcere la nostra comune percezione della realtà, ponendo dubbi concreti sulla veridicità o meno dei contenuti di cui siamo quotidianamente fruitori nello spazio virtuale.

Il «New York Times», nel 2023, già sollevava preoccupazioni in merito alla possibilità di stabilire, nell'odierna società digitale, cosa sia effettivamente reale o meno, in un'epoca in cui immagini e informazioni manipolate, sebbene iper-realistiche, materializzano un timore ormai

* Avvocato in Novara. Assegnista di ricerca presso la Cattedra di Informatica giuridica dell'Università degli Studi di Milano, Dipartimento di Scienze giuridiche "Cesare Beccaria", Facoltà di Giurisprudenza come cultrice della materia, sui temi legati al diritto antidiscriminatorio, nuove tecnologie, tutela dei soggetti vulnerabili e minori.

doi 10.53148/CD202501001

diventato certezza[2]: le enormi potenzialità offerte dalle nuove tecnologie, e in particolare dall'avvento dell'intelligenza artificiale generativa, possono trasformarsi in un'arma potentissima di diffusione di fake news.

A tali preoccupazioni si aggiungano quelle relative a una considerazione provocatoria: sono e saranno sempre di più le macchine a discriminare il vero dal falso, a indirizzare la nostra percezione della realtà e, magari, a radicare in noi convinzioni, ideologie, gusti personali? Ci dimenticheremo presto come sviluppare e alimentare quotidianamente il nostro senso critico? Cosimo Accoto considera ormai questo scenario già insito in molti contesti digitali "tradizionali": dai server di posta elettronica che separano le e-mail utili dallo spam, ai data center bancari che distinguono gli accessi finanziari legittimi da quelli fraudolenti o criminali[3].

Se la "verità" può essere intesa in vario modo a seconda della cultura e del sistema di pensiero che la utilizza, esistono da sempre vari strumenti in grado di manipolarla e plasmarla, finendo per inquinare il dibattito pubblico e inficiare la qualità stessa della democrazia. Tra questi strumenti, la pervasività dei social media nella ricerca dell'informazione nell'esperienza quotidiana dell'onlife[4], ha aperto nuove possibilità di accesso alla diffusione di notizie false, dando spazio a contenuti totalmente o parzialmente falsi o errati, creati con diversi livelli di intenzionalità[5]. A questa considerazione, ormai pacifica tra gli studiosi del tema, si aggiunga come l'intelligenza artificiale sia oggi capace di amplificare il fenomeno, abbia

[2] Cfr. T. Hsu e S. Lee Myers, "Can We No longer Believe Anything We See?", «New York Times», 8 Aprile 2023.

[3] Cfr. C. Accoto, "Il (di)segno della verità nell'era della simulazione", «Harvard Business Review», 2024.

[4] Cfr. L. Floridi, *La quarta rivoluzione. Come l'infosfera sta trasformando il mondo*, Milano, Raffaello Cortina, 2017.

[5] Secondo la Treccani, per "disinformazione" si intende la «diffusione intenzionale di notizie o informazioni inesatte o distorte allo scopo di influenzare le azioni e le scelte di qualcuno» (https://www.treccani.it/vocabolario/disinformazione/?search=disinformazióne%2F.). La "misinformazione" fa riferimento a un insieme di «informazioni non accurate, inattendibili, i cui contenuti, diffusi frettolosamente, rischiano di essere accettati come veritieri perché difficili o impossibili da verificare» (https://www.treccani.it/vocabolario/neo-misinformazione_(Neologismi)/). La "cattiva informazione", traduzione del *malinformation* inglese, si riferisce, secondo il Collins Dictionary, alla pubblicazione di informazioni veritiere con l'intenzione di arrecare danno (https://www.collinsdictionary.com/dictionary/english/malinformation).

La verità ai tempi dell'IA tra rischi e nuove sfide

L'intelligenza artificiale generativa sta ridefinendo il concetto di verità nell'era digitale, facilitando la creazione di immagini, testi e contenuti iper-realistici che possono alimentare la disinformazione. Strumenti avanzati come deepfake e chatbot intelligenti hanno il potenziale, dunque, di manipolare l'opinione pubblica, influenzare le dinamiche politiche e alterare la percezione della realtà, rendendo sempre più difficile distinguere il vero dal falso.
L'articolo analizza i rischi e le criticità legate alla diffusione della disinformazione tramite l'intelligenza artificiale, così come considera anche la possibilità di un'alleanza tra i suoi potentissimi strumenti e la verità: così come algoritmi avanzati possono potenziare la quantità, la qualità e la personalizzazione delle fake news, essi possono altresì essere impiegati per il miglioramento della qualità informativa.

Truth in the time of AI between risks and new challenges

Generative artificial intelligence is redefining the concept of truth in the digital age, facilitating the creation of images, texts and hyper-realistic content that can feed disinformation. Advanced tools such as deepfakes and intelligent chatbots have the potential, therefore, to manipulate public opinion, influence political dynamics and alter the perception of reality, making it increasingly difficult to distinguish the true from the false.
The article analyses the risks and critical issues related to the spread of disinformation through artificial intelligence, as well as considers the possibility of an alliance between its powerful tools and the truth: just as advanced algorithms can enhance the quantity, quality and customisation of fake news, they can also be used to improve the quality of information.

Ciberspazio e diritto, vol. 26, n. 79 (1 - 2025), pp. 29-43

Il ruolo dell'AI nella manipolazione delle informazioni e strategie di intervento

Elena Enrica Bacigalupi*

Sommario: 1. L'importanza della qualità dei dati in input nei sistemi di AI. – 2. Lo sfruttamento dei sistemi di AI in un contesto tecnopolitico e tecnocratico. – 3. Impatti delle informazioni generate con sistemi di AI sugli individui. – 4. Scenari e strategie di intervento multidisciplinare per un'AI affidabile.

1. *L'importanza della qualità dei dati in input nei sistemi di AI*

L'Intelligenza Artificiale (di seguito anche "AI") ha il potenziale per migliorare la qualità della vita e trasformare positivamente la società, ma l'efficacia dei sistemi di AI dipende drasticamente dalla qualità dei dati utilizzati per addestrarli e farli funzionare. Come suggerito in un recente rapporto dell'UNESCO (2021)[1], la qualità dei dati in input è un elemento essenziale per garantire che l'AI operi in modo affidabile, il che si rifletterebbe positivamente anche in una migliore qualità di vita per gli esseri umani. Il concetto di affidabilità, infatti, fa riferimento sia ad aspetti tecnici che sociali, riflettendo caratteristiche di liceità, eticità e robustezza[2].

A tal proposito, dati incompleti, non rappresentativi o distorti possono introdurre *bias* significativi nei modelli di AI, portando a decisio-

* La Dott.ssa Elena Enrica Bacigalupi è consulente presso le Nazioni Unite alla United Nations International Computing Centre. È esperta in ambito compliance e mitigazione dei rischi per la data protection e la cybersecurity, ed è certificata Lead Auditor ISO/IEC 27001, Privacy Implementer ISO/IEC 27701 e Artificial Intelligence ISO/IEC 42001.

[1] Cfr. UNESCO - Organizzazione delle Nazioni Unite per l'Educazione, la Scienza e la Cultura, "Recommendation on the Ethics of Artificial Intelligence", 2021.

[2] High-Level Expert Group on Artificial Intelligence set up by the European Commission, "Ethics Guidelines for Trustworthy AI", 2019, p. 5.

doi 10.53148/CD202501002

ni ingiuste o discriminatorie[3]. Ad esempio, un algoritmo di recruiting addestrato su dati storici che riflettono disparità di genere nelle assunzioni potrebbe perpetuare tali discriminazioni escludendo sistematicamente le candidate donne.

Oltre a introdurre pregiudizi, dati di scarsa qualità possono compromettere l'affidabilità delle decisioni prese dai sistemi di AI generando misinformazione, intesa come informazione inaccurata o inattendibile, rischiando che venga accettata come veritiera poiché difficile o impossibile da verificare. Ciò è particolarmente preoccupante in contesti ad alto rischio come la medicina, dove diagnosi errate basate su dati imprecisi potrebbero avere conseguenze letali[4].

Il processo di ingestione dei dati inizia con la fase di raccolta, che può avvenire attraverso molteplici canali e fonti. In particolare, i sistemi moderni si trovano a dover gestire un'enorme varietà di dati: da quelli strutturati provenienti da database relazionali e API, fino ai dati non strutturati come testi, immagini, audio e video. Tale varietà richiede approcci differenti nella loro gestione, ma anche un'attenzione costante alla preservazione della qualità e dell'integrità[5] delle informazioni durante tutto il processo di acquisizione, indipendentemente dalla tipologia dei dati utilizzati.

Una volta raccolti, è necessario procedere all'identificazione dei valori mancanti e alla normalizzazione, eliminando inconsistenze o duplicati e assicurando che i dati siano comparabili e utilizzabili dal sistema di AI. Un aspetto particolarmente delicato è quello, poi, dell'etichettatura dei dati (nota come "data labeling")[6], processo che può richiedere l'intervento di esperti del dominio per garantire la correttezza delle annotazioni, come per esempio giuristi che classificano documenti legali. L'impor-

[3] Cfr. N. Mehrabi, F. Morstatter, N. Saxena, K. Lerman, A. Galstyan, "A Survey on Bias and Fairness in Machine Learning", 2022, pp. 4-11.

[4] Cfr. M.A. Gianfrancesco, S. Tamang, J. Yazdany, et al. "Potential Biases in Machine Learning Algorithms Using Electronic Health Record Data", JAMA Internal Medicine, 178(11), 2018, pp. 1544-1547.

[5] Per "integrità" si intende «proprietà di accuratezza e completezza», riprendendo la definizione fornita dalla ISO/IEC 27000:2018 - Information Security Management Systems, 3.36.

[6] Processo attraverso il quale si associano a ciascun dato o *dataset* (insieme di dati) una o più *label* che identificano le caratteristiche o categorie rilevanti per il task di apprendimento. Concetto simile ma non alternativo a quello dei metadati, fondamentali, tra l'altro, per promuovere la rintracciabilità dei dati.

Il ruolo dell'AI nella manipolazione delle informazioni e strategie di intervento

L'Intelligenza Artificiale (AI) sta rapidamente ridefinendo il panorama tecnologico, economico e sociale del XXI secolo. Tuttavia, le informazioni mediate dall'AI possono intensificare l'impatto della misinformazione e della disinformazione su molteplici fronti. Nella prima sezione del presente contributo viene esaminata l'importanza di garantire dataset di alta qualità per prevenire *bias* algoritmici e generazione di output errati o imprecisi. La seconda sezione tratta, invece, il tema dello sfruttamento dell'AI per la diffusione di fake news in contesti tecnopolitici e tecnocratici, con il rischio di polarizzazione sociale e minaccia alla stabilità democratica. La terza parte è, poi, dedicata agli impatti cognitivi sugli individui, tra cui la dipendenza tecnologica e la potenziale perdita di capacità di pensiero critico. In conclusione, l'articolo ipotizza scenari futuri conseguenti alla manipolazione delle informazioni automatizzate e propone un approccio multidisciplinare, consolidando una proposta di soluzioni tecnologiche, normative ed educative per garantire un'AI affidabile al servizio del benessere collettivo e dei singoli individui.

The role of AI in information manipulation and intervention strategies

Artificial Intelligence (AI) is rapidly redefining the technological, economic, and social landscape of the 21st century. However, AI-mediated information can intensify the impact of misinformation and disinformation on multiple fronts. The first section examines the importance of ensuring high-quality datasets to prevent algorithmic biases and the generation of incorrect or inaccurate outputs. The second section addresses the exploitation of AI for spreading fake news in technopolitical and technocratic contexts, highlighting the risk of social polarization and threats to democratic stability. The third part is dedicated to the cognitive impacts on individuals, including technological dependency and the potential loss of critical thinking skills. In conclusion, the article hypothesizes future scenarios resulting from the automated information manipulation and proposes a multidisciplinary approach, consolidating a set of technological, regulatory, and educational solutions aimed at ensuring a trustworthy AI that serves the collective and individual well-being.

Ciberspazio e diritto, vol. 26, n. 79 (1 - 2025), pp. 45-59

Disinformazione e social network: la guerra alla verità dagli USA al DSA dell'Unione europea

Roberto Balzamo*

1. *Introduzione*

Persuadere qualcuno che qualcosa di falso sia la verità, o mantenere la verità nascosta a proprio vantaggio, è l'essenza di un fenomeno tanto diffuso quanto antico: la disinformazione. Questa è da intendersi, secondo il gruppo di esperti europeo noto come *High Level Expert Group on Fake news and Online Disinformation* (HLEG), come informazioni false o fuorvianti che vengono create, presentate e diffuse per ottenere un guadagno economico o per ingannare intenzionalmente il pubblico, e che possono arrecare danno alla collettività[1]; inoltre, essa ha un aspetto multiforme, rappresentando una delle maggiori sfide per gli ordinamenti giuridici in tutto il mondo. Da un punto di vista storico, l'esistenza della disinformazione è identificabile sin dagli albori dell'informazione; tuttavia, ciò che ne è mutato sono le modalità e gli strumenti utilizzati per disseminarla, dapprima con l'epocale invenzione della stampa nel quindicesimo secolo e successivamente con l'avvento dell'era digita-

* Laureato presso il corso di laurea triennale in "Scienze Politiche e delle Relazioni Internazionali" dell'Università degli Studi di Napoli "L'Orientale", Classe L-36, laureando presso il corso di laurea magistrale in "Law and Sustainable Development" dell'Università degli Studi di Milano, Classe LM/SC_GIUR.

[1] Cfr. S. Sassi, *L'Unione Europea e la lotta alla disinformazione online,* 2023, p. 185. Disponibile su: https://flore.unifi.it/retrieve/532d0890-3920-4f81-9fcc-885c7be8056f/Disinformazione%20-%20Federalismi.pdf.

doi 10.53148/CD202501003

le. Infatti, ben prima della nascita di Internet e dei moderni sistemi di informazione, nel 1923, la disinformazione veniva già classificata come una complessa tattica del tutto equiparabile a una "arma da guerra"[2]. Sebbene l'intenzione che si cela dietro la diffusione di contenuti volutamente deformati e imprecisi sia progressivamente sempre più chiara, tra cui influenzare processi democratici quali il momento elettorale e la modificazione dell'opinione pubblica col fine di frammentarla, da un punto di vista legale i suoi confini sono tutt'altro che omogeneamente definiti, così come dimostrerà l'analisi comparativa tra l'ordinamento giuridico statunitense e quello dell'Unione europea. In particolar modo, questo lavoro si propone di analizzare il ruolo delle piattaforme digitali come strumenti privilegiati per la diffusione di false informazioni, con un'attenzione specifica ai meccanismi algoritmici da essi impiegati, come i sistemi di raccomandazione, rei di innescare il fenomeno del c.d *filter bubbles*, contribuendo alla compartimentazione e alla polarizzazione del dibattito pubblico. Successivamente, si approfondirà il caso delle elezioni presidenziali statunitensi del 2016 e del 2024, emblema dell'impatto di tale fenomeno sui processi democratici. Di conseguenza, lo studio si concentrerà sul confronto tra le risposte normative adottate dall'Unione europea e dagli Stati Uniti, con particolare riferimento al *Digital Services Act* (DSA), per evidenziare come quest'ultimo tenti di regolamentare la diffusione di contenuti disinformativi attraverso obblighi specifici e approcci innovativi nel panorama normativo internazionale. Infine, si discuteranno le principali criticità e sfide aperte, tra cui l'ambiguità dell'assenza di una definizione di disinformazione nel DSA e i limiti pratici riscontrabili nell'attuazione delle normative in esame.

2. *Disinformazione e social media: gli strumenti per la diffusione di false informazioni*

Le maggiori piattaforme di social media si sono recentemente affermate quali protagoniste dell'informazione, per cui negli Stati Uniti circa il sessanta per cento della popolazione utilizza una piattaforma social come strumento privilegiato per l'accesso all'informazione, una quota in netto

[2] *Ibidem.*

Disinformazione e social network: la guerra alla verità dagli USA al DSA dell'Unione europea

La disinformazione rappresenta una delle principali sfide per le democrazie contemporanee, amplificata dalla diffusione capillare dei social media e dall'uso di sofisticati algoritmi di raccomandazione. Questo studio analizza il fenomeno partendo dal contesto statunitense, con particolare attenzione alle elezioni presidenziali del 2016 e del 2024, per poi confrontarlo con il quadro normativo europeo, incentrato sul "Digital Services Act" (DSA). Se da un lato gli Stati Uniti adottano un approccio liberale, basato sulla protezione della libertà di espressione e sulla limitata responsabilità delle piattaforme, l'Unione europea ha sviluppato una regolamentazione più strutturata, imponendo obblighi di trasparenza e misure di contrasto alla diffusione di contenuti falsi o dannosi. Tuttavia, permangono criticità sia nell'efficacia delle normative, sia nella loro implementazione, in particolare per l'assenza di una definizione giuridica univoca di disinformazione nel DSA. Lo studio evidenzia, inoltre, il ruolo dei social media nella polarizzazione del dibattito pubblico, aggravata dal fenomeno del cosiddetto *filter bubbles* e dall'uso di strumenti come *bot* e *deepfake*. In conclusione, si sottolinea la necessità di un approccio sinergico tra regolamentazione, responsabilità delle piattaforme e alfabetizzazione digitale, al fine di garantire un ecosistema informativo più trasparente e democratico.

Disinformation and social networks: the war on truth from the USA to the European Union's DSA

Disinformation represents one of the main challenges for contemporary democracies, amplified by the widespread diffusion of social media and the use of sophisticated recommendation algorithms. This study analyzes the phenomenon starting from the United States context, with particular attention to the 2016 and 2024 presidential elections, and subsequently compares it with the European regulatory framework, centered around the Digital Services Act (DSA). While the United States adopts a liberal approach, grounded in the protection of freedom of expression and the limited liability of platforms, the European Union has developed a more structured regulation, imposing transparency obligations and measures to counter the dissemination of false or harmful content. Nevertheless, critical issues persist regarding both the effectiveness of these regulations and their implementation, particularly due to the absence of a clear legal definition of disinformation within the DSA. The study also highlights the role of social media in the polarization of public discourse, exacerbated by the phe-

nomenon of so-called filter bubbles and the use of tools such as bots and deepfakes. In conclusion, the article underscores the need for a synergistic approach involving regulation, platform accountability, and digital literacy, in order to ensure a more transparent and democratic information ecosystem.

Ciberspazio e diritto, vol. 26, n. 79 (1 - 2025), pp. 61-80

Le dinamiche comunicative nell'era della post verità: riflessioni sociogiuridiche sul *framework* europeo di contrasto alla disinformazione online

Nicola Pierpaolo Barbuzzi*

1. *Le dinamiche della comunicazione nei social: dai modelli tradizionali ai feed algoritmici*

È ormai un dato incontrovertibile come il nostro tempo sia quello della comunicazione attraverso le immagini più che quello della parola. I social (media e network), ad oggi sempre meno considerati spazi neutri e democratici, hanno soppiantato il ruolo egemone della carta stampata, della radio e in ultimo della televisione nella formazione dell'informazione, proponendo uno codice narrativo composto da una miscela di contenuti facilmente comprensibili, immagini e video – molto spesso – fuori

* Nicola Pierpaolo Barbuzzi è docente a contratto di diritto privato presso l'Universitas Mercatorum. È avvocato nonché docente di diritto e processo penale presso la Scuola Allievi Finanzieri della Guardia di Finanza. Cultore della materia in diritto privato e PhD's presso l'Università Telematica Pegaso, già docente a contratto di diritto dei mezzi di comunicazione presso la medesima Università, è autore di numerose pubblicazioni su riviste scientifiche, contributi in volumi collettanei e monografie tra cui *Cyberbullismo, odio in rete e diffamazione nell'era digitale. Analisi giuridica e strategia di tutela*, Duepuntozero, 2024; *La filosofia punitiva nel Codice Rocco*, in S. Ricchitelli (a cura di) *La pena e la sua esecuzione in Italia*, Roma-Bari, Laterza, 2023; *Guida alla responsabilità dei genitori e dei precettori*, Molfetta, Duepuntozero, 2021; *Le intercettazioni di conversazioni e comunicazioni*, Molfetta, Duepuntozero, 2020.

contesto con accattivanti pubblicità che, di fatto, impediscono all'utente di concentrarsi su ciò che realmente passa sotto i propri occhi. L'ideologia californiana, sviluppatasi nella Silicon Valley negli anni Novanta, che accompagnava il lancio dei primi social network, accogliendo con entusiasmo il potere emancipativo delle nuove tecnologie dell'informazione, in grado di liberare il cittadino-utente dalle ormai obsolete strutture sociali ponendo al centro la libertà individuale[2], cede il passo a una visione pragmatica di un "non luogo" di vulnerabilità, a disposizione sì dell'utente, ma che ruota attorno agli interessi dell'intermediario fornitore/gestore in cui, inevitabilmente, rientrano anche le sue simpatie politiche, molto spesso cangianti e ondivaghe. La partecipazione dell'utente digitale ("prosumer"[3]) al processo comunicativo viene costantemente stimolata all'interno dei social network i quali, insistentemente, chiedono a cosa l'utente stia pensando, invitandolo continuamente a esprimere la sua opinione su qualsiasi argomento o a condividere, attraverso *stories*, frammenti del proprio vissuto.

La maggioranza dei contenuti che passano sui social si materializzano grazie ad algoritmi che massimizzano l'*engagement* (coinvolgimento emotivo), trattenendo gli utilizzatori sul *newsfeed* (🏠), la *home* del social per intenderci, attraverso titoli *clickbait*, in grado di catalizzare like e irretire una platea di utenti sempre più numerosa, incapace di distinguere la veridicità di una informazione, ma, paradossalmente, abile nel condividere quelle "verità" che a suo dire, nessuno avrà il coraggio di condividere.

L'utente ingenuo delle origini, allora bersagliato da un flusso ininterrotto e monodirezionale di informazioni, si è oggi trasformato in un utente proattivo, non più semplice fruitore di contenuti, ma partecipe, sino all'esasperazione (*pshycopathology of information overload*), del processo di formazione di un'informazione sempre più crossmediale. La continua e costante erosione dei collaudati standard comunicativi da parte dei social ha creato dei veri e propri canali di relazione decentralizzati[4],

[2] Cfr. M.A. Polesana, *Influencer e social media*, Milano, FrancoAngeli, 2023.

[3] Cfr. A. Toffler, *The third wave*, New York, William Morrow and Company, 1980.

[4] Cfr. M. De Salvia, V. Zagreblesky, *Diritti dell'uomo e libertà fondamentali. La giurisprudenza della Corte europea dei diritti dell'uomo e della Corte di giustizia delle Comunità europee*, Milano, Giuffrè, 2007; cfr. D.M. Orofino, *La libertà di espressione*

Le dinamiche comunicative nell'era della post verità: riflessioni sociogiuridiche sul framework *europeo di contrasto alla disinformazione online*

La moderazione dei contenuti è essenziale per contrastare la disinformazione online. Le *fake news*, amplificate dagli algoritmi, si diffondono rapidamente, influenzando società e mercato. Il Digital Services Act (DSA) dell'UE impone regole chiare alle grandi piattaforme, promuovendo un ambiente trasparente e neutrale, dove le informazioni fuorvianti siano limitate a tutela dell'infrastruttura sociale. Tuttavia, bilanciare controllo e libertà di espressione è cruciale. Politiche trasparenti e strumenti di *fact-checking* sono indispensabili per un ecosistema informativo sicuro e affidabile, in cui società civile e comunità scientifica svolgano un ruolo attivo.

Communicative dynamics in the age of post-truth: socio-legal reflections on the European framework against online disinformation

Content moderation is essential in combating online misinformation. Fake news, amplified by algorithms, spreads rapidly, influencing society and markets. The European Union's Digital Services Act (DSA) establishes clear rules for large platforms, promoting a transparent and neutral environment where misleading information is limited to safeguard the social infrastructure. However, balancing control and freedom of expression is crucial. Transparent policies and fact-checking tools are indispensable for a secure and reliable information ecosystem, in which civil society and the scientific community play an active role.

Ciberspazio e diritto, vol. 26, n. 79 (1 - 2025), pp. 81-101

Falsi, disinformazione e complottismo come parte della società civile

Diego Giorio*

Sommario: 1. Premessa. – 2. Qualche esempio. – 3. Fake news e disinformazione. – 3.1. Le ragioni alla base di disinformazione e misinformazione. – 4. Il complottismo. – 4.1. Le responsabilità della scienza. – 5. La *cancel culture*. – 6. Come difendersi – 7. Il ruolo del Legislatore e della magistratura. – 8. Conclusioni.

1. *Premessa*

Da sempre l'umanità è alla ricerca della verità. La filosofia nata nella Magna Grecia è solamente un momento nel quale questa ricerca è stata in qualche modo organizzata, etichettata, ma, fin dagli albori della civiltà, espressioni religiose, riti iniziatici, osservazioni della realtà naturale hanno sempre cercato di scoprire le verità che governano il mondo. E, fin dalla notte dei tempi, vi è sempre stata una separazione, più o meno intenzionale, tra chi conosce e comprende e chi è privo di questa conoscenza. Le caste sacerdotali, i filosofi e gli scienziati – che creavano vere e proprie sette, come i pitagorici – erano nettamente separati dal popolo analfabeta: vivevano in modo distinto da quanti, per impossibilità di accesso all'istruzione o per insufficienti capacità intellettive, non riuscivano ad accedere al patrimonio culturale del momento. Al di là di altri aspetti, questo scollamento sociale ha inevitabilmente creato sospetti (conoscenze nascoste, segreti gruppi di potere, invenzioni secretate) e ha favorito la formazione e diffusione di convinzioni parallele (esseri immaginari, magie, filtri miracolosi), slegate dalla conoscenza scientifica e dai fatti dimostrabi-

* Servizi demografici del Comune di Villanova Canavese – Autore e membro del Consiglio di Redazione per SEPEL Editrice. Le opinioni nel presente articolo sono espresse a titolo puramente personale.

doi 10.53148/CD202501005

li, basate sulla percezione dei sensi (la terra piatta o il sole che ruota attorno alla terra) o su pseudoscienze fondate sull'autosuggestione o su trucchi da prestigiatore. Oggi, nonostante Internet consenta l'accesso a una fonte illimitata di sapere, non è cambiato molto. Ma cosa può fare il Legislatore, nazionale o sovranazionale, per evitare la disinformazione e contenerne gli effetti? Ben poco, a parere di chi scrive. Se la filosofia del diritto insegna che una norma non può prevedere e regolare ogni situazione reale che possa verificarsi, vale anche il limite opposto, ovvero che non tutti i comportamenti umani possono essere controllati da una norma.

2. *Qualche esempio*

Se il fenomeno della disinformazione è così diffuso nel tempo e nello spazio, diventa difficile scegliere il punto di partenza tra tanti spunti che si possono trovare. In questo caso si è deciso di iniziare dalla storia europea delle patate.

Il tubero, oggi così diffuso tanto nei fast food quanto nell'alta cucina, è arrivato in Europa dopo la scoperta dell'America, dov'era coltivato per uso alimentare da oltre 5000 anni e, poiché il misoneismo è diffuso quanto la disinformazione, sulle prime venne guardato con molto sospetto[1]. Oltretutto, venne inizialmente usato soprattutto come mangime per maiali e cavalli, il che accrebbe i sentimenti negativi verso il nuovo alimento: si diffusero notizie (naturalmente provenienti da fonti certe e autorevoli) che trasmettesse malattie, che facesse volare le streghe, che portasse malattie esantematiche. Senza contare che patate germogliate vennero date ai carcerati, causando realmente intossicazioni alimentari.

Quando il botanico Antoine Augustin Parmentier, che le aveva mangiate e studiate mentre era prigioniero in Prussia durante la guerra dei Sette anni, capì che in realtà si trattava di una grande opportunità per sfamare la popolazione e limitare le carestie, visto il valore nutritivo e la facile coltivazione in terreni molto diversi, ebbe un'idea geniale per diffon-

[1] Cfr. A.A. Parmentier, *Traité sur la culture et les usages des pommes de terre, de la patate e du topinambour*, Barrois, 1789; C.C. Mann, "How the Potato Changed the World", 2011, disponibile al seguente link: https://www.smithsonianmag.com/history/how-the-potato-changed-the-world-108470605/ (ultimo accesso in data 13 maggio 2025).

Falsi, disinformazione e complottismo come parte della società civile

Da sempre l'umanità è alla ricerca della verità. La filosofia nata nella Magna Grecia è solamente un momento nel quale questa ricerca è stata in qualche modo organizzata ed etichettata, ma, fin dagli albori della civiltà, espressioni religiose, riti iniziatici, osservazioni della realtà naturale hanno sempre cercato di scoprire le verità che governano il mondo. Tuttavia, molto spesso la verità non è raggiungibile e nascono quindi dicerie, fisime, paure. Oppure la verità non viene accettata, facendo nascere teorie complottiste. Il ruolo del Legislatore è, in questo caso, marginale, dato che non è possibile imporre una verità per legge, anche se questa verità fosse unica e certa. La filosofia del diritto insegna che una norma non può prevedere e regolare ogni situazione reale che possa verificarsi; vale però anche il limite opposto, ovvero non tutti i comportamenti possono essere controllati da una norma. Lo strumento migliore in mano alle democrazie moderne non è, quindi, la censura, ma l'istruzione, che consente ai cittadini di discernere le fonti, di analizzare i fatti, di formare un'opinione fondata sul pensiero critico.

Fake news, misinformation, and conspiracy theories as part of civil society

Since the dawn of civilization, religions, initiation rituals, observations of natural reality have always sought to discover the truths that govern the world. Philosophy, born in Magna Graecia, is just one way which that search has been organized and labeled, but there have been many others. The truth can be difficult to pinpoint, however. This gives rise to rumors, fixations, and fears. Or the facts are not accepted, giving rise to conspiracy theories. The role of the Legislator is marginal in such cases, since it is not possible to impose a truth by law, even if this truth were unique and certain. The philosophy of law teaches us that a norm cannot predict and regulate every real situation that may occur; however, the opposite limitation also applies, namely, not all behavior can be controlled by a norm. The best tool in the hands of modern democracies, therefore, is not censorship, but education, which enables citizens to discern sources, analyze facts, and form opinions based on critical thinking.

Ciberspazio e diritto, vol. 26, n. 79 (1 - 2025), pp. 103-120

L'interrelazione tra protezione dei dati personali e disinformazione in Unione europea, quali prospettive?

Elisabetta Stringhi[*]

Sommario: 1. Considerazioni introduttive: l'interrelazione tra protezione dati e regolazione della disinformazione a livello europeo. – 2. Il Regolamento (UE) 2024/900 e la disciplina della trasparenza e del *targeting* della pubblicità politica. – 3. L'intreccio tra protezione dati e contrasto alla disinformazione. – 4. Il ruolo dell'Autorità di protezione dei dati personali e il coordinamento con il Coordinatore dei servizi digitali. – 5. Conclusioni.

1. *Considerazioni introduttive: l'interrelazione tra protezione dati e regolazione della disinformazione a livello europeo*

L'ascesa dei sistemi algoritmici e di intelligenza artificiale ha alterato radicalmente il panorama della condivisione delle informazioni e del discorso politico, introducendo all'interno delle piattaforme online profilazione invasiva, pubblicità mirata e campagne di disinformazione che minacciano la fiducia del pubblico e l'integrità dei processi democratici[1].

[*] Elisabetta Stringhi lavora presso l'Autorità Garante per la Protezione dei Dati Personali ed è *research fellow* presso l'*Information Society Law Center* dell'Università degli Studi di Milano. In passato, ha prestato attività di consulenza legale in materia di protezione dei dati personali. È inoltre docente di numerosi corsi e *master* in diritto delle nuove tecnologie. Le posizioni espresse nel contributo sono strettamente personali e non vincolano l'Autorità di appartenenza né sono a essa in alcun modo attribuibili.

[1] Esiste ormai un'ampia letteratura sull'argomento. Sono numerosi i casi documentati di utilizzi scorretti dei dati personali che hanno influenzato lo svolgimento o, addirittura, l'esito di processi elettorali. Particolarmente preoccupante il caso delle elezioni brasiliane del 2022, in cui l'uso combinato di tecniche di *micro-targeting* e l'utilizzo illecito dei dati personali dei cittadini per lanciare campagne di disinformazione su Whatsapp e Telegram ha fortemente contribuito a esacerbare le divisioni interne, fino a influenzare l'opinione degli elettori e minare l'ordine pubblico interno. Per un approfondimento, cfr. R. Cazzamatta, A. Santos, G. Albuquerque, "Unveiling Disinfor-

Lo svolgimento di campagne elettorali *data-driven* e l'uso (anche illecito)[2] di tecniche di *targeting* e *micro-targeting* per la consegna mirata del messaggio politico al potenziale elettore hanno sollevato a più riprese a livello europeo la pressante necessità di contrastare il dilagante fenomeno della disinformazione e salvaguardare il corretto svolgimento delle operazioni elettorali attraverso, da un lato, lo strumento della protezione dei dati personali[3] e, dall'altro, con una regolazione giuridica della pubblicità politica[4].

Secondo l'intuizione profonda del legislatore europeo, la protezione dei dati personali è un elemento centrale per la salvaguardia delle libertà d'opinione e politiche, perché consente una formazione dell'identità politica libera da indebiti condizionamenti e manipolazioni, protegge gli elettori da forme invasive di profilazione e contribuisce a ridurre fenomeni di polarizzazione, innalzando la qualità del dibattito pubblico.

In questa prospettiva, la protezione dei dati personali assume un ruolo di primo piano per la tutela dell'ordinamento democratico, introdu-

mation: Mapping Attacks on Brazil's Electoral System and the Response of the Superior Electoral Court", «International Journal of Communication», vol. 18, 2024.

[2] Pur essendo state valutate come libere dagli osservatori internazionali, le elezioni ungheresi del 2022 hanno sollevato notevoli preoccupazioni per l'uso scorretto dei dati personali degli elettori per *advertising* politico. È stato, infatti, documentato che i dati personali raccolti dal governo per finalità connesse all'esercizio di poteri pubblici o allo svolgimento di compiti pubblici sono stati ri-utilizzati impropriamente per veicolare i messaggi della campagna elettorale di Fidesz, evidenziando una profonda erosione dell'ordinamento democratico interno. Anche questo caso è dimostrativo della profonda connessione tra protezione dati, uso lecito dei dati personali degli elettori per finalità di *political advertising* e preservazione della democrazia. Cfr. D. Brown, *Trapped in a web: The exploitation of personal data in Hungary's 2022 elections*.

[3] Cfr. conss. 51 e 57 del Regolamento (UE) 2016/679.

[4] Cfr. cons. 4 del Regolamento (UE) 2024/900: «La pubblicità politica può essere un vettore di disinformazione, specie se non ne è esplicitata la natura politica, se proviene da *sponsor* esterni all'Unione o se è oggetto di tecniche di *targeting* o tecniche di consegna dei messaggi pubblicitari. È necessario un livello elevato di trasparenza anche per sostenere un dibattito politico e campagne politiche equi e aperti, come pure elezioni o referendum liberi e regolari, e per combattere la manipolazione dell'informazione e le interferenze, nonché le interferenze illecite anche da paesi terzi. Una pubblicità politica trasparente aiuta l'elettore e gli individui in generale a capire meglio quando è in presenza di un messaggio di pubblicità politica, per conto di chi è fatta quella pubblicità nonché come e perché è diventato il *target* di un prestatore di servizi pubblicitari, ponendolo così in condizioni migliori per una scelta informata».

L'interrelazione tra protezione dei dati personali e disinformazione in Unione europea, quali prospettive?

Il contributo approfondisce l'interrelazione tra protezione dei dati personali e regolazione della disinformazione a livello europeo, analizzando la disciplina della trasparenza e del *targeting* della pubblicità politica online del Regolamento (UE) 2024/900. Tale Regolamento affronta il problema della disinformazione, contrastandola come forma di annuncio pubblicitario o di messaggio politico amplificato verso il pagamento di un corrispettivo. In particolare, il Regolamento introduce diversi obblighi di trasparenza e di *due diligence* per la pubblicità politica. Inoltre, esso disciplina l'uso di tecniche di *targeting* e di consegna del messaggio pubblicitario politico, introducendo norme imperative che vietano direttamente determinati trattamenti di specifiche categorie di dati personali, apposite condizioni di trattamento e ulteriori obblighi di trasparenza nel contesto del *marketing* politico. Si evidenzia che l'intreccio tra tutela dei dati personali e regolazione della pubblicità politica richiede all'interprete di individuare i ruoli e le responsabilità privacy degli attori del *marketing* politico. Infine, si sottolinea che la complessa cornice di *governance* delineata dal Regolamento solleva importanti sfide di coordinamento tra le Autorità competenti, specialmente nell'ottica di valorizzare il ruolo di vigilanza delle Autorità di protezione dati sul *targeting* politico.

The interrelationship between personal data protection and disinformation in the European Union: what are the prospects?

The paper explores the interrelationship between personal data protection and the regulation of disinformation at the European level, analyzing the regulation of transparency and targeting of online political advertising in EU Regulation 2024/900. This Regulation addresses the problem of disinformation by counteracting it as a form of advertisement or amplified political message towards payment of a fee. Specifically, the Regulation introduces several transparency and due diligence requirements for political advertising. In addition, it regulates the use of targeting and delivery techniques of political advertising message, introducing mandatory rules directly prohibiting certain processing of specific categories of personal data, special processing conditions, and additional transparency obligations in the context of political marketing. It is pointed out that the intertwining of personal data protection and political advertising regulation requires the interpreter to identify the privacy roles and responsibilities of polit-

ical marketing actors. Finally, it is emphasized that the complex governance framework outlined by the Regulation raises important coordination challenges among the relevant authorities, especially with a view to enhancing the supervisory role of Data Protection Authorities on political targeting.

Ciberspazio e diritto, vol. 26, n. 79 (1 - 2025), pp. 121-137

Il "riposo del vero". La dissimulazione artificiale per la tutela del diritto d'autore

Alessia Palladino[*]

Sommario: 1. Il demiurgo artificiale. Il caso dell'IA generativa nella dialettica "vero – falso". – 2. *Offendicula 4.0*: i casi Glaze e Nightshade come strumenti di autotutela tecnica del diritto d'autore. – 3. La dissimulazione artificiale come "riposo tecnico" del vero. – 4. Considerazioni conclusive. La dissimulazione come ulteriore forma di degenerazione etico – giuridica.

1. *Il demiurgo artificiale. Il caso dell'IA generativa nella dialettica "vero – falso"*

Le più recenti frontiere dell'Intelligenza Artificiale ("IA")[1] segnano il progressivo ingresso nella "Quarta rivoluzione industriale"[2], ove l'uso massiccio di macchine[3], sempre più pensanti, sugella nuove e più stringenti forme di integrazione del mondo fisico, digitale e biologico.

[*] Alessia Palladino è Assegnista di Ricerca in Informatica Giuridica e Cultore della Materia in "Computer Law" [SSD GIUR-17/A], presso Università degli Studi di Cagliari, Corso di Laurea in Computer Engineering, Cybersecurity and Artificial Intelligence.

[1] Cfr. M.A. Boden, *L'intelligenza Artificiale*, Bologna, Il Mulino, 2019. Cfr. L. Corso, "Intelligenza collettiva, intelligenza artificiale e principio democratico", in R. Giordano, A. Panzarola, A. Police, S. Preziosi, M. Proto (a cura di), *Il diritto nell'era digitale. Persona, mercato, amministrazione, giustizia*, Milano, Giuffrè, 2022, p. 443-459. Cfr. A. Punzi, "La persona del futuro. Il dialogo delle intelligenze tra umanesimo e tecnoscienze", «Persona e Mercato», 2023, pp. 161-168.

[2] Questo concetto è stato formalizzato nel 2016 da Klaus Schwab, fondatore del World Economic Forum, descrivendo tale rivoluzione come un punto di convergenza tra tecnologie avanzate come l'intelligenza artificiale (IA), la robotica, l'Internet delle Cose (IoT), la stampa 3D, le biotecnologie e i computer quantistici. Cfr. J. Barrat, *La nostra invenzione finale. L'intelligenza artificiale e la fine dell'età dell'uomo*, D. Pezzella, M. Pezzella (trad. it.), Roma, Nutrimenti, 2019, p. 9.

[3] Per un maggior approfondimento si rinvia a A.C. Amato Mangiameli, "Algoritmi e big data. Dalla carta alla robotica", «Rivista di Filosofia del diritto», n. 1, 2019,

Specialmente nel corso dell'ultimo decennio, l'intelligenza artificiale ha acquisito rinnovato slancio sia a livello internazionale che nazionale[4], proponendosi come un fenomeno poliedrico[5] e uno strumento sempre più pervasivo nella vita quotidiana[6] – tanto nel settore privato[7] quanto pubblico[8] – capace di innovare tanto il *modus vivendi*[9], quanto l'*ars pensandi*[10].

p. 108; G. Pasceri, *Intelligenza artificiale, algoritmo e machine learning*, Milano, Giuffrè, 2021, p. 11. L'Autore rileva che «l'intelligenza artificiale è figlia del naturale sviluppo dell'innovazione tecnologica come conseguenza ordinaria della crescita scientifica, tecnica e culturale dell'uomo. L'errore, in cui spesso si incorre, è quello di identificarla, diversamente, come un processo tecnologico moderno frutto della capacità di calcolo e dell'informatizzazione dei processi».

[4] Le tecnologie dell'informazione e della comunicazione (ICT) hanno guidato l'aumento della produttività europea dal 1995. Cfr. Comunicazione della Commissione al Parlamento europeo, al Consiglio, al Comitato economico e sociale europeo e al Comitato delle regioni, *Strategia per il mercato unico digitale per l'Europa*, COM/2015/0192 final del 6.5.2015. Disponibile al seguente link: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A52015DC0192. Cfr. anche Comunicazione della Commissione al Parlamento europeo, al Consiglio, al Comitato economico e sociale e al Comitato delle Regioni, *Creare fiducia nell'intelligenza artificiale antropocentrica,* COM(2019) 168 final del 08.04.2019. Disponibile al seguente link: https://eur-lex.europa.eu/legal-content/IT/TXT/HTML/?uri=CELEX:52019DC0168&from=EN.

[5] Sugli stadi evolutivi dell'intelligenza artificiale si veda M. Farina, "Brevi riflessioni sullo status delle 'persone elettroniche'", «L'Ircocervo», n. 2, 2021, pp. 106-126.

[6] Cfr. M.U. Scherer, "Regulating Artificial Intelligence Systems: Risks, Challenges, Competencies, And Strategies", «Harvard Journal of Law & Technology», vol. 29, n. 2, Berlino, Spring, 2016; M. Craglia, *et al.*, *Artificial Intelligence - A European perspective*, European Commission, Joint Research Centre, Artificial intelligence: European perspective, Publications Office, 2019; L. Floridi, *The Fourth Revolution, How the Infosphere is Reshaping Human Reality*, Oxford, Oxford University Press, 2014.

[7] Per una panoramica, si v. M.L. Montagnani, "Governance societaria e governance dell'intelligenza artificiale", «Mercato, concorrenza, regole», n. 2, 2022, pp. 271-290.

[8] Cfr. I. Martin Delgado, "Automazione, intelligenza artificiale e pubblica amministrazione: vecchie categorie concettuali per nuovi problemi?", «Ist. Federalismo», n. 3, 2019, p. 643.

[9] Cfr. A. Gehlen, *L'uomo nell'era della tecnica. Problemi socio-psicologici della civiltà industriale*, Milano, SugarCo, 1967, p. 12.

[10] Cfr. H. Jonas, *Dalla fede antica all'uomo tecnologico* (1974), trad. it., Bologna, Il Mulino, 1991, p. 9; cfr. A. Longo, G. Scorza, *Intelligenza artificiale. L'impatto sulle nostre vite, diritti e libertà*, Milano, Mondadori, 2020, pp. 57-58.

Il "riposo del vero". La dissimulazione artificiale per la tutela del diritto d'autore

L'IA generativa dimostra una latente vocazione dato centrica, che ritrova nel *web scraping* una massa critica pressoché inesauribile di dati per alimentare i *datasets* e le attività di addestramento.
Tali attività, specialmente nella variante *text-image*, tuttavia, rischiano di minare il diritto d'autore, rivelando tutta l'inadeguatezza della mera tutela regolatoria, incapace di ammantare di effettività la tutela stessa di tali diritti.
Per tali ragioni, verranno esaminate le più recenti iniziative, volte alla tutela "tecnica" del diritto d'autore, per esprimerne vantaggi e criticità.
Tali contromisure, infatti, se da un lato contribuiscono a neutralizzare l'utilizzo improprio di tali immagini, al contempo alimentano un circuito "avvelenato", che rischia, su larga scala, di contribuire al dilagare di nuove forme di degenerazione dell'IA.
Per tali ragioni, si rifletterà sul rapporto tra vero e falso, nonché tra dissimulazione e degenerazione.

The "rest of truth". Artificial dissimulation for the protection of copyright

Generative AI demonstrates a latent data-centric vocation, which finds in web scraping an almost inexhaustible critical mass of data to feed datasets and training activities.
Such activities, especially in the text-image variant, however, risk undermining copyright, revealing all the inadequacy of mere regulatory protection, incapable of cloaking the protection of these rights itself with effectiveness.
For these reasons, the most recent initiatives aimed at the "technical" protection of copyright will be examined to express their advantages and critical issues.
In fact, these countermeasures, if on the one hand contribute to neutralizing the improper use of these images, at the same time fuel a "poisoned" circuit, which risks, on a large scale, contributing to the spread of new forms of AI degeneration.
For these reasons, we will reflect on the relationship between true and false, as well as between dissimulation and degeneration.

Ciberspazio e diritto, vol. 26, n. 79 (1 - 2025), pp. 139-158

Governance dell'IA nell'epoca dei deepfake: note comparatistiche sulle esperienze di Cina e Giappone

Sonia Sforza[*], Davide Luigi Totaro[**]

1. *Verso la* Trustworthy AI*: le sfide del deepfake*

Il tema e le problematiche legate alla diffusione di fake news[1], riemersi prepotentemente sulla scena internazionale nell'ultimo decennio[2], primeggiano tra le sfide cogenti per il mondo del diritto. Tale affermazione è tutto fuorché iperbolica, specie in considerazione del fatto che dette notizie sono passibili non solo di pregiudicare primari valori quali l'integrità dell'informazione, il diritto all'onore e all'immagine nonché la leale concorrenza, ma anche di minare la stabilità sociale e l'ordine pubblico, manipolando condotte, amplificando divisioni, alimentando tensioni politiche fino a compromettere la stessa fiducia nelle istituzioni[3]. Tutto ciò è a fortiori problematico nell'era dei social media[4].

[*] Dottoranda di ricerca in Diritto Privato Comparato presso Università degli studi di Milano, è autrice dei paragrafi 1 e 2.

[**] Docente presso Hitotsubashi Institute for Advanced Study dell'Hitotsubashi University di Tokyo e Professore aggregato presso Rikkyo University College of Law and Politics di Tokyo, è autore dei paragrafi 3 e 4.

[1] Cfr. D.M.J. Lazer, *et al.*, "The science of fake news", «Science», vol. 359, n. 6380, 2018, p. 1094.

[2] Cfr. G. Pennycook, D.G. Rand, "The Psychology of Fake News", «Trends in Cognitive Sciences», vol. 25, n. 5, 2021, p. 388.

[3] Cfr. G. Di Domenico *et al.*, "Fake news, social media and marketing: A systematic review", «Journal of Business Research», vol. 123, 2021, pp. 335-336.

[4] Cfr. E. Aïmeur, S. Amri, G. Brassard, "Fake news, disinformation and misinformation in social media: a review", «Social Network Analysis and Mining», vol. 13, n. 30, 2024, p. 30.

In questo contesto, le rapide e trasformative innovazioni tecnologiche legate all'intelligenza artificiale (IA) che hanno caratterizzato l'ultimo quinquennio, hanno riportato la questione sotto i riflettori mediatici e dei regolatori, a seguito della nascita di una nuova generazione di contenuti digitali comunemente noti come deepfake. Questi output, generati attraverso tecniche avanzate di deep learning, esacerbano la succitata problematica sia da un punto di vista quantitativo, in considerazione della potenza generativa dei moderni sistemi di IA, ma anche e soprattutto da un punto di vista qualitativo, in quanto tramite i deepfake è possibile replicare volti, voci e comportamenti umani con un realismo senza precedenti, rendendo particolarmente difficoltoso ogni procedimento di riconoscimento e verifica[5]; in ultimo, andando a rendere il confine tra realtà e finzione sempre più incerto e sfumato agli occhi di utenti consumatori e professionali.

Non stupisce, dunque, osservare come tale recente evoluzione sia percepita come un'addizionale minaccia per l'ordine sociale – lo stato di diritto[6] – nonché la fiducia del mercato e degli investitori nei confronti proprio delle nuove tecnologie. Pertanto, lo sviluppo di una *trustworthy AI* rappresenta oggi una priorità condivisa dai regolatori di tutto il mondo[7]. A ciò si affianca anche la connessa, ma non coincidente, istanza comune di «realizzare sistemi di IA antropocentrici»[8] e al servizio dell'uomo, obiettivo che, a ben vedere, non può prescindere dalla crea-

[5] Cfr. K. Kertysova, "Artificial intelligence and disinformation", «Security and Human Rights», vol. 29, nn. 1-4, 2018, pp. 66-68.

[6] Cfr. C. Vaccari, A. Chadwick, "Deepfakes and Disinformation: Exploring the Impact of Synthetic Political Video on Deception, Uncertainty, and Trust in News", «Social Media + Society», vol. 6, n. 1, 2020, p. 3.

[7] A titolo esemplificativo i G20 AI Principles; le Ethics guidelines for trustworthy AI, elaborate dall'Unione Europea; i Social Principles of Human-Centric AI elaborati dal governo giapponese, tutti nel 2019, nonché il White Paper on Trustworthy Artificial Intelligence elaborato dalla China Academy of Information and Communications Technology del 2021. In generale, alla base dell'idea della *trustworthy AI* si individua il convincimento per cui individui, organizzazioni e società nel suo complesso potranno realizzare appieno il potenziale dell'IA solo se sarà possibile instaurare un rapporto di fiducia nel suo sviluppo, nella sua implementazione e nel suo utilizzo. Cfr. S. Thiebes, S. Lins, A. Sunyaev, "Trustworthy artificial intelligence", «Electronic Markets», vol. 31, 2021, pp. 447-448.

[8] Cfr. Ethics guidelines for trustworthy AI, 2019.

Governance dell'IA nell'epoca dei deepfake: note comparatistiche sulle esperienze di Cina e Giappone

L'avvento dei deepfake si presenta come una cruciale sfida per il diritto nel nuovo millennio, incidendo sull'integrità dell'informazione, sui diritti individuali e sulla stabilità sociale. Le trasformative capacità dei sistemi di intelligenza artificiale e il sempre più elevato grado di realismo, che essi consentono ai contenuti generati, impongono un'accurata analisi degli strumenti giuridici a disposizione per la governance di tale fenomeno. Forte di questa consapevolezza, il presente contributo propone un'analisi comparata degli approcci regolamentari preposti alla gestione dei deepfake e delle problematiche a essi associate in sistemi tecnologicamente all'avanguardia quali Cina e Giappone. La prospettata analisi si propone di osservare gli strumenti di hard law e di soft law ivi adottati e applicabili al tema in esame, di evidenziare affinità e divergenze e di valutare l'affermazione di nuovi modelli nel panorama internazionale dell'AI governance.

AI governance in the age of deepfakes: comparative notes on the experiences of China and Japan

The advent of deepfakes represents a crucial challenge for law in the new millennium, affecting the integrity of information, individual rights, and social stability. The transformative capabilities of artificial intelligence systems and the increasingly high degree of realism they enable in generated content necessitate a thorough analysis of the legal instruments available to govern this phenomenon. With this awareness, the present contribution offers a comparative analysis of the regulatory approaches adopted for the governance of deepfakes and the related issues in technologically advanced systems such as China and Japan. The proposed analysis aims to examine the hard law and soft law instruments implemented and applicable to the subject, highlight similarities and differences, and assess the emergence of new models in the AI governance international landscape.

Protezione dei dati e sicurezza

Ciberspazio e diritto, vol. 26, n. 79 (1 - 2025), pp. 161-170

L'istituto giuridico delle *regulatory sandboxes* tra diritto, protezione dei dati e intelligenza artificiale

Francesca Santoro[*]

Sommario: 1. L'istituto giuridico delle *regulatory sandboxes* nel panorama normativo europeo e italiano. – 2. La *ratio* e l'approccio basato sul rischio e sulla supervisione regolamentare in tempo reale. – 3. Le *AI regulatory sandboxes* nella recente disciplina del Regolamento UE 2024/1689 (AI Act). 4. Condizioni per il trattamento dei dati personali all'interno delle *regulatory sandboxes* per lo sviluppo di sistemi di AI nell'interesse pubblico.

1. *L'istituto giuridico delle* regulatory sandboxes *nel panorama normativo europeo e italiano*

Con il termine *regulatory sandbox* o spazio di sperimentazione si intende un recente istituto giuridico che ha l'obiettivo primario di garantire che gli ordinamenti nazionali e la regolamentazione vadano di pari passo alla rapida evoluzione tecnologica, senza creare una barriera già dalle prime fasi di sperimentazione dell'innovazione. In particolare, tale istituto giuridico risponde all'esigenza di non incorrere nel rischio che considerevoli tempi di normazione abbiano un impatto negativo sul progresso tecnologico molto rapido, nonché di introdurre un modello più flessibile che risponda meglio alle esigenze di disciplina della società odierna. In particolare, occorre pensare che il progresso tecnologico sta rendendo conveniente ciò che prima era dispendioso[1].

[*] Francesca Santoro è giurista esperta in informatica giuridica e diritto delle nuove tecnologie. Attualmente ricopre il ruolo di Director presso Deloitte in Italia, dove guida i servizi di compliance in materia di Privacy e Digital Regulation. Si occupa da anni di progetti di trasformazione digitale, con focus su data risk management, compliance regolamentare e cybersecurity, operando in contesti nazionali e internazionali. È in possesso di diverse certificazioni professionali, collabora con università e associazioni per attività formative e ha contribuito a diverse pubblicazioni scientifiche.

[1] Crf. A. Agrawal, J. Gans, A. Goldfarb, "Macchine predittive", «Harvard Business Review Press», ristampa FrancoAngeli, 2018, pp. 21-25.

Difatti, proprio al fine di testare la massima efficacia degli output derivanti, ad esempio, dall'intelligenza artificiale, mitigando al contempo i potenziali rischi sotto diversi profili, molti ordinamenti giuridici hanno iniziato ad adottare lo strumento delle regulatory sandboxes[2].

Questo nuovo modello fa riferimento alla temporanea deroga di alcune barriere normative al fine di consentire di eseguire test e sperimentazioni in libertà, ancorché sotto la vigilanza delle autorità.

Volendo tracciare le origini di tale istituto, da ultimo disciplinato nel Regolamento (UE) 2024/1689 sull'intelligenza artificiale (AI Act), esse possono identificarsi nel 2015 quando nel Regno Unito emerse uno dei primi modelli nel contesto del settore Fintech. Da lì in avanti, ulteriori Paesi hanno progressivamente adottato questo modello di sperimentazione controllata delle tecnologie più innovative.

Con riferimento all'Italia, a partire dal 2020, è entrata in vigore nel nostro ordinamento una disciplina generale delle regulatory sandboxes riflessa nell'art. 36 del d.l. n. 76/2020[3], convertito con modificazioni dalla legge n. 120/2020, relativo alle "Misure urgenti per la semplificazione e l'innovazione digitale".

Tale d.l., da ultimo aggiornato all'atto pubblicato il 27 dicembre 2024, nel Capo IV "Misure per l'innovazione" all'articolo 36 prevede: «al fine di favorire la trasformazione digitale della pubblica amministrazione, nonché lo sviluppo, la diffusione e l'impiego delle tecnologie emergenti e di iniziative ad alto valore tecnologico, le imprese, le Università, gli enti di ricerca pubblici e privati e le società con caratteristiche di spin off o di start up universitari […] che intendono sperimentare iniziative attinenti all'innovazione tecnologica e alla digitalizzazione, possono presentare alla struttura della Presidenza del Consiglio dei ministri competente per la trasformazione digitale i relativi progetti, con contestuale domanda di temporanea deroga alle norme dello Stato […] che impediscono la sperimentazione».

Si noti che, se da un lato viene recepita a livello di disciplina una sostanziale, seppur temporanea, deroga alle norme dello Stato per fini di sperimentazione innovativa, dall'altro lato rimangono esclusi da que-

[2] Cfr. P. Dal Checco, "Sandbox", «Dizionario Legal tech», a cura di G. Ziccardi e P. Perri, Milano, Giuffrè Francis Lefebvre, 2020, pp. 845-846.

[3] Cfr. parere del Consiglio di Stato, sez. atti normativi, 29 gennaio 2021, n. 109.

L'istituto giuridico delle regulatory sandaboxes *tra diritto, protezione dei dati e intelligenza artificiale*

Un nuovo modello di regolazione delle tecnologie innovative, in particolare dell'intelligenza artificiale, denominato regulatory sandbox, sta emergendo nell'ambito degli ordinamenti giuridici. Tale modello fa riferimento alla temporanea deroga di alcune barriere normative, al fine di consentire di testare tali tecnologie in libertà, ancorché sotto la vigilanza delle Autorità. L'istituzione di tali c.d. spazi di sperimentazione potrà contribuire in concreto a migliorare la certezza del diritto e l'apprendimento normativo basato su dati concreti, nonché a sostenere la condivisione delle migliori pratiche attraverso la cooperazione con le Autorità. Obiettivo di questo Articolo è di offrire una panoramica dell'istituto giuridico delle regulatory sandboxes, con focus sulla recente disciplina introdotta dal Regolamento (UE) 2024/1689 sull'intelligenza artificiale, analizzando le principali questioni giuridiche connesse e gli aspetti procedimentali.

The legal institution of regulatory sandboxes between law, data protection and artificial intelligence

A new model for regulating innovative technologies, particularly artificial intelligence, known as regulatory sandbox, is emerging within legal systems. This model refers to the temporary derogation of certain regulatory barriers in order to allow the testing of the most innovative technologies freely, under the supervision of Authorities. The establishment of such testing spaces can concretely contribute to improve legal certainty and support regulatory learning based on real data, as well as promote the sharing of best practices through cooperation with Authorities. The aim of this Article is to provide an overview of the legal concept of regulatory sandboxes, with a focus on the provisions regulated by Regulation (EU) 2024/1689 on artificial intelligence recently, analyzing the key legal issues involved and the main procedural aspect.

Ciberspazio e diritto, vol. 26, n. 79 (1 - 2025), pp. 171-191

Il cyber scudo europeo e l'architettura di cybersicurezza italiana

Alessio Vergnano*

Sommario: 1. Introduzione. – 2. L'attuale legislazione europea. – 3. Il *Cyber Solidarity Act*. – 4. La disciplina italiana di fronte alle nuove cyber frontiere comuni. – 5. La cybersecurity europea nel contesto digitale globale. – 6. Conclusioni.

1. *Introduzione*

La nuova era digitale ha portato rivoluzioni significative in molti settori, tra cui il lavoro, l'educazione, la finanza e la sicurezza. La pandemia di COVID-19 ha accelerato la trasformazione del settore sanitario, mentre l'ascesa delle criptovalute ha rivoluzionato il mercato finanziario, evidenziando l'importanza della sicurezza dei dati personali. Questi cambiamenti radicali presentano nuove sfide per la società, coinvolgendo enti statali, istituzioni europee, cittadini e aziende di ogni dimensione.

Per questi soggetti la sicurezza informatica, un tempo considerata un aspetto marginale, è ora al centro dell'attenzione a causa dei crescenti attacchi informatici che minacciano di generare caos e sfruttare le vulnerabilità tecniche e legislative. È essenziale analizzare come la cybersicurezza stia diventando una componente cruciale nella difesa militare e in molti altri settori: «la drammaticità della situazione non emerge soltanto dagli enormi danni economici per le infrastrutture strategiche statali derivanti da questa situazione di Far West digitale – danni che per natura, gravità, sistematicità e dimensione travalicano costantemente i confini delle tecnologie dell'informazione e della stessa cybersicurezza – ma anche e soprattutto per l'aumento esponenziale di attacchi diffusi, molti dei quali di natura para-militare, frutto delle tensioni internazionali tra superpotenze e del conflitto ad alta intensità combattuto ai confini dell'Europa»[1].

* Studente di Giurisprudenza presso l'Università degli Studi di Milano-Bicocca, Segretario dell'Associazione studentesca Legal Hackers Bicocca, Vice-Chair del Cyber-

Degne di nota sono le evoluzioni verificatesi nel panorama degli attacchi informatici, che mirano a generare caos, esercitare controllo e sfruttare, sia direttamente sia indirettamente, le lacune presenti in termini di competenze tecniche, strumenti, conoscenze e coesione legislativa. Risulta quindi evidente che «in un mondo dove ogni comportamento può essere tradotto in dati digitali, la protezione di questi dati è fondamentale non solo per comprendere i comportamenti passati, ma anche per predire quelli futuri»[2].

Il fenomeno è senza dubbio in crescita sia a livello nazionale[3], sia a livello europeo e globale[4].

2. *L'attuale legislazione europea*

Il campo della sicurezza informatica in Europa è nel pieno del suo sviluppo normativo. Negli ultimi anni, l'Unione europea ha emanato una serie di Direttive e Regolamenti per rafforzare la protezione contro le crescenti minacce informatiche, riconoscendo l'importanza cruciale della cybersicurezza in vari contesti.

«Il quadro giuridico sulla *cybersecurity* europea si va articolando in maniera sempre più complessa ed emerge un rafforzamento non solo delle competenze tecniche e operative dei rispettivi organismi coinvolti ma anche della collaborazione tra gli stessi: si pensi ad esempio alla cooperazione tra ENISA ed Europol ai fini di contrasto della cybercriminalità»[5].

security Strategic Committee presso Steppo Jean Monnet European Union Centre of Excellence dell'Università di Milano Bicocca.

[1] Cfr. A. Gatti, M. Giannelli, "Presupposti per la configurazione e la dichiarazione di guerra cibernetica", «Convegno DPCE Pescara 2023 in DPCE online», Sp-1, 2024, pp. 456-457.

[2] Cfr. A. Di Corinto, "Data commons: privacy e cybersecurity sono diritti umani fondamentali", «Rivista italiana di Informatica e Diritto», Fascicolo n. 1, 2022, p. 34.

[3] Cfr. Agenzia per la Cybersicurezza Nazionale, "Relazione annuale al Parlamento 2024", 2025, disponibile al link https://www.acn.gov.it/portale/relazione-annuale/2024 (ultimo accesso in data 20/05/2025).

[4] Cfr. Clusit, "Rapporto 2025 sulla Cybersicurty in Italia e nel mondo", 2025, disponibile al link https://clusit.it/rapporto-clusit/ (ultimo accesso in data 20/05/2025).

[5] Cfr. D. Marrani, "Il coordinamento delle politiche per la cybersecurity dell'UE nello spazio di libertà, sicurezza e giustizia", «Freedom, Security & Justice: European Legal Studies - Rivista quadrimestrale on line sullo Spazio europeo di libertà, sicurezza e giustizia», n. 1, 2021, p. 88.

Il cyber scudo europeo e l'architettura di cybersicurezza italiana

Il presente contributo, dopo una breve analisi del contesto storico e delle minacce informatiche presenti attualmente, procederà con una visione sullo stato dell'arte della legislazione europea di cybersicurezza.
Entrando nel merito della ricerca, *in primis* verrà esposto il nuovo Regolamento *Cyber Solidarity Act*, indicando capo per capo le istituzioni e gli enti su cui ricadono gli oneri espressi dal Regolamento e che formano il nuovo cyber scudo europeo.
Successivamente, attraverso un'analisi sistematica dell'Agenzia per la Cybersicurezza Nazionale, verrà delineata la struttura di quest'ultima e i nuovi organismi che si svilupperanno per cooperare alla cybersecurity nazionale in un'ottica di maggior capillarizzazione sul territorio nazionale e di cooperazione europea.
Infine, l'articolo si concluderà con una valutazione della cybersecurity come priorità per l'Unione europea, non solo a causa della sua fondamentale importanza per la difesa, ma soprattutto in virtù del suo potenziale per far avanzare l'Unione europea verso una maggiore coesione, anche in materie ancora fortemente sottoposte a interessi nazionali.

The European cyber shield and Italy's cybersecurity structure

After a brief analysis of the historical context and current cyber threats, this paper will proceed with an overview of the state of the art of European cybersecurity legislation.
Going into the merits of the research, firstly, the new Cyber Solidarity Act Regulation will be presented, indicating, chapter by chapter, the institutions and bodies on which the obligations expressed in the Regulation fall and which form the new European cyber shield.
Subsequently, through a systematic analysis of the National Cybersecurity Agency, the structure of the latter will be outlined, along with the new bodies that will be developed to cooperate on national cybersecurity with a view to greater coverage across the country and European cooperation.
The article will conclude with an assessment of cybersecurity as a priority for the European Union, not only because of its fundamental importance for defense, but above all because of its potential to advance the European Union towards greater cohesion, even in areas still strongly subject to national interests.