

# Disinformazione digitale

**Modelli algoritmici  
e valori democratici**

a cura di  
Licia Califano  
Federica Fabrizi  
Giovanni Sartor

**FrancoAngeli** 

*Collana*

**di Diritto**

SAGGI E RICERCHE



# Disinformazione digitale

**Modelli algoritmici  
e valori democratici**

a cura di  
**Licia Califano  
Federica Fabrizzi  
Giovanni Sartor**

**FrancoAngeli** 

*Collana*

**di Diritto**

**SAGGI E RICERCHE**

# INDICE

*Introduzione* pag. 9

## **Sezione I Disinformazione e democrazia**

*Verità e disinformazione nella sfera pubblica digitale*  
Massimo Durante » 15

*La cura contro la disinformazione. Rimedi giuridici (e non)  
per contrastarne la diffusione*  
Federica Fabrizzi » 35

*La qualificabilità della disinformazione quale problema  
di cybersicurezza*  
Giovanni Barozzi Reggiani » 59

*Oltre la disinformazione: il problema della manipolazione nel quadro  
della regolazione algoritmica europea*  
Ludovica Durst » 75

*Il disordine informativo online: classificare le condotte per definirle  
e prevenirle giuridicamente*  
Andrea Ruffo » 101

*La crisi delle democrazie costituzionali al tempo  
della disinformazione: quale futuro per la liberal-democrazia?*  
Luca Maria Tonelli » 113

## Sezione II Informazione digitale e tutela dell'individuo

- Poteri pubblici, poteri privati e governo costituzionale dell'informazione digitale*  
Giulio Enea Vigevani pag. 131
- AI e nuovi modelli di garanzia dei diritti fondamentali: criticità e prospettive applicative*  
Giulia Vasino » 141
- Violenza digitale di genere in ambito politico: profili comparati tra Unione europea e Messico*  
Rosa Iannaccone » 165
- L'informazione nell'era dell'algoritmo: intelligenza artificiale e libertà di espressione nella governance europea*  
Giulia Napoli » 179
- Disclosure, Explainability e Human Oversight: nuovi diritti fondamentali nell'era dell'IA e dei Private Governors*  
Andrei-Mihai Pop » 193

## Sezione III Comunicazione elettorale e politica

- Comunicazione istituzionale e intelligenza artificiale. Appunti*  
Massimo Cavino » 207
- Il Regolamento europeo sulla pubblicità politica, ovvero del difficile tentativo di regolamentare la propaganda online (e non solo)*  
Anna Papa » 219
- La comunicazione politica nella digital era*  
Giuliaserena Stegher » 237

<i>Fighting Foreign Interference. Un'analisi comparata delle legislazioni contro le interferenze straniere nei processi elettorali in Australia, Canada, Irlanda, Regno Unito e Nuova Zelanda</i> Emanuele Gabriele	pag. 269
<i>Il contrasto alla disinformazione nei procedimenti elettorali: spunti comparatistici tra Italia e Spagna</i> Lorenzo De Carlo	» 283
<i>Comunicazione e nuove tecnologie: la propaganda politica online nella Costituzione</i> Allegra Dominici	» 297
<i>Il sistema di governance ed enforcement del Regolamento (UE) 2024/900 alla prova dell'effettività: un'analisi comparativa con il Digital Services Act</i> Emanuela Palomba	» 311

#### **Sezione IV** **Soluzioni regolatorie e tecniche** **nel mondo del digitale**

<i>The Governance of "Augmented Disinformation" between the Digital Services Act and the AI Act</i> Federico Galli, Giuseppe Contissa	» 325
<i>Fake news e Intelligenza Artificiale: sfide etiche e soluzioni tecnologiche</i> Andrea Ongarini, Federico Cerutti, Andrea Loreggia	» 357
<i>Misure europee e nazionali contro la disinformazione strategica online. Contributo a partire dal caso Storm-1516</i> Domenico Bruno	» 369
<i>Large Language Models: applicazioni e criticità nel dibattito pubblico-informativo. Spunti di riflessione informatico-giuridici</i> Casimiro Coniglione	» 385

*The state of the art legal and technical of automated fake news  
moderation systems*

Camilla Scarpellino

pag. 399

*Affrontare la disinformazione online:  
il sistema di mitigazione dei rischi nel Digital Services Act*

Sara Gallone

» 415

### **Riflessioni conclusive**

*L'attualità del costituzionalismo a garanzia dei diritti  
fondamentali nella società digitale*

Licia Califano

» 427

*Autrici e Autori*

» 439

# L'INFORMAZIONE NELL'ERA DELL'ALGORITMO: INTELLIGENZA ARTIFICIALE E LIBERTÀ DI ESPRESSIONE NELLA GOVERNANCE EUROPEA

Giulia Napoli\*

SOMMARIO: 1. La diffusione di notizie false nello spazio digitale: dal mito della rete alla crisi dell'informazione – 2. Intelligenza artificiale e manipolazione del consenso: la democrazia nella società algoritmica – 3. Dalla *soft law* alla *Constitutional governance*: il modello europeo di contrasto alla disinformazione *online* tra Dsa e AI Act – 4. *Pars construens*: l'IA al servizio del pluralismo informativo.

## 1. La diffusione di notizie false nello spazio digitale: dal mito della rete alla crisi dell'informazione

Ogni ordinamento democratico si costruisce attorno a un presupposto tanto semplice quanto decisivo: che il processo di formazione della volontà collettiva poggi su un'informazione libera, pluralista e attendibile<sup>1</sup>. Quando tale pre-

\* Dottoranda di ricerca in Diritto costituzionale e pubblico, Università degli Studi di Milano-Bicocca.

1. La formula riflette l'idea di democrazia deliberativa cara, tra gli altri, a J. Habermas (al quale si rimanda, in particolare in *Storia e critica dell'opinione pubblica*, alias *Strukturwandel der Öffentlichkeit*, 1962; Laterza, 2005), J. Rawls, R. Dahl e S. Rodotà, secondo i quali il consenso democratico non è autentico se non si fonda su un dibattito pubblico informato. Più di recente, questo assunto è stato sviluppato da M. Luciani, *Libertà di espressione e pluralismo politico nella giurisprudenza della Corte costituzionale*, Corte Costituzionale, 2025, ove si legge che «senza pluralismo non c'è democrazia e senza libertà di espressione non c'è pluralismo», perché «fra libertà di espressione, libertà di informazione, pluralismo e democrazia vi [è] una corrispondenza biunivoca». Si rimanda altresì a G. Sartori, *The Theory of Democracy Revisited*, Chatham House, 1987, ove egli scrive che «il potere elettorale è di per sé la garanzia meccanica della democrazia; ma la garanzia sostanziale è data dalle condizioni sotto le quali il cittadino acquisisce l'informazione». Sul rapporto tra pluralismo informativo e nuove tecnologie, si vedano anche, tra gli altri, G.E. Vigevani, *I media di servizio pubblico nell'età della rete. Verso un nuovo fondamento costituzionale, tra autonomia e pluralismo*, Giappichelli, 2018; A. Lamberti, *Libertà di informazione e democrazia ai tempi della società digitale: problemi e prospet-*

supposto si incrina, anche la sovranità popolare vacilla, poiché si altera la capacità dei cittadini di partecipare alla vita democratica consapevolmente. In questo delicato equilibrio tra libertà e verità, tra informazione e potere, si colloca il fenomeno della disinformazione: non una novità del nostro secolo o un portato della contemporaneità, ma una costante della storia e della comunicazione politica, che le tecnologie digitali e l'intelligenza artificiale hanno reso più pervasiva ed efficace, fino a trasformarla in una minaccia sistemica per l'integrità del dibattito pubblico e, di riflesso, per la salubrità dei processi democratici<sup>2</sup>.

Già Marc Bloch, nel celebre saggio *Réflexions d'un historien sur les fausses nouvelles de la guerre* (1921) osservava come le "fausses nouvelles" non costituissero soltanto l'espressione di meri difetti di percezione o di inaccurately comunicative, bensì un fenomeno politico e sociale, che rivela le credenze, le paure e le passioni di un'epoca e che risulta capace di incrinare il rapporto fiduciario tra cittadini e istituzioni<sup>3</sup>. Quell'intuizione, lungi dal risultare ormai datata, conserva del tutto intatta la propria attualità. Tuttavia, se nel passato la diffusione del falso era limitata dalla lentezza dei mezzi analogici di comunicazione, oggi la velocità e la capillarità della rete – alle quali si aggiunge la "deprofessionalizzazione" dell'informazione, favorita dalla facilità di accesso a strumenti di pubblicazione più immediati e dalla progressiva erosione del ruolo degli intermediari tradizionali – amplificano esponenzialmente il rischio di manipolazione dell'informazione. Il risultato è un ecosistema informativo frammentato, certamente polifono ma forse eccessivamente rumoroso, in cui la distinzione tra fonti autorevoli e contenuti amatoriali si assottiglia e l'attenzione dell'utente diventa una risorsa scarsa, contesa da logiche di mercato<sup>4</sup>.

In questo scenario si inseriscono le grandi piattaforme digitali – alcune delle quali molto note, quali Facebook, X, YouTube, TikTok –, attori privati divenuti in poco tempo parte integrante dell'infrastruttura pubblica della comunicazione e *tertium* nel tradizionale rapporto tra Stato e cittadino<sup>5</sup>. Per la loro

*tive*, Consulta Online, II, 2022; R. Mastroianni, O. Pollicino, M. Bassini (a cura di), *Il T.U. dei servizi di media audiovisivi*, Giuffrè, 2024.

2. Sul rapporto tra pluralismo, libertà di espressione e nuove tecnologie si veda M. Bassini, *Intelligenza Artificiale generativa e media*, in R. Mastroianni, O. Pollicino, M. Bassini (a cura di), *Il T.U. dei servizi di media audiovisivi*, Giuffrè, 2024, pp. 269-293.

3. M. Bloch, *Réflexions d'un historien sur les fausses nouvelles de la guerre*, 1921. Nella traduzione del saggio del 1997 edita Einaudi, dal titolo *Riflessioni d'uno storico sulle false notizie della guerra*, si parla proprio di "percezioni distorte", magari anche giuste al principio «ma mal interpretate – unanimemente stravolte per accordarsi agli ardenti desideri di tutti».

4. F. Nanni, *I rischi dell'utilizzo dell'IA sul diritto di informazione – La disinformazione*, in *Rivista di diritto dei media*, I, 2025.

5. Il riferimento imprescindibile qui è J. Balkin, *Free speech is a triangle*, in *Columbia Law Review*, n. 118/2011.

capacità di connettere, selezionare e diffondere contenuti su scala globale, esse si configurano di fatto come la principale *agorà* del discorso pubblico, lo spazio in cui si formano le opinioni, si consolidano le credenze e si plasmano gli orientamenti politici.

Ma proprio in virtù di tale indubbia centralità, che le ha rese le vere e proprie *gatekeepers* dell'informazione<sup>6</sup>, tali piattaforme meritano oggi un'attenzione particolare. In effetti, le *digital platforms* non si limitano a svolgere – o, quantomeno, non più – una funzione di intermediazione neutrale dei contenuti immessi dagli utenti: il principio di neutralità della rete, che in origine le voleva meri vettori del flusso comunicativo, scembrare da interventi di moderazione e immuni da qualsivoglia responsabilità, ha ceduto il passo a una realtà assai diversa, nella quale gli operatori digitali assumono un ruolo attivo di selezione, filtraggio e orientamento dei messaggi<sup>7</sup>. Tale potere – talvolta esercitato attraverso sistemi algoritmici automatizzati di *content moderation* sulla base dei *community standards* – comporta, nei fatti, la possibilità di sopprimere, oscurare o declassare contenuti e profili sulla base di criteri non sempre trasparenti. In questo senso, le piattaforme si trovano a esercitare un potere di incidenza sulla libertà di espressione che, pur muovendosi formalmente nel perimetro del diritto privato, assume effetti sostanzialmente pubblici: esse decidono, infatti, quali discorsi restano visibili, quali vengono rimossi e, in ultima analisi, quali idee possono circolare nello spazio digitale.

Quel che ne deriva è un rischio strutturale di “censura privata”, ossia di limitazione del discorso pubblico non imposta da un'autorità pubblica, ma da regole contrattuali e algoritmiche – spesso opache e comunque governate da logiche di mero profitto – che sfuggono a ogni forma di controllo democratico e che, in concreto, determinano in misura sempre crescente la forma e la qualità del dibattito pubblico<sup>8</sup>.

Eppure, per struttura e per logica di funzionamento, la rete avrebbe dovuto ampliare e rendere più immediata – e in parte effettivamente lo ha fatto – la fruizione da parte dei cittadini dei diritti e delle libertà fondamentali, in particolare la libertà di manifestazione del pensiero e il correlato diritto all'informazione: l'accesso diretto a una quantità potenzialmente illimitata di fonti,

6. Questa fortunata espressione è stata istituzionalizzata dal Digital Markets Act del 2022: sul punto, si veda C. Caruso, *Il tempo delle istituzioni di libertà. Piattaforme digitali, disinformazione e discorso pubblico europeo*, in *Quaderni costituzionali*, n. 3/2023, pp. 558-559.

7. Si veda G.E. Vigevani, *Piattaforme digitali private, potere pubblico e libertà di espressione*, in *Diritti di libertà e nuove tecnologie*, a cura di A. Morelli, *Diritto Costituzionale*, 1, VI, 2023.

8. Sul tema della cosiddetta “privatizzazione della censura” si rimanda a M. Monti, *Privatizzazione della censura e Internet Platforms: la libertà d'espressione e i nuovi censori dell'agorà digitale*, in *Rivista italiana di informatica e diritto*, n. 1/2019.

la possibilità di produrre e diffondere contenuti in tempo reale e la riduzione delle barriere all'ingresso nel dibattito pubblico sembravano tradurre in pratica l'ideale che aveva caratterizzato Internet ai suoi albori. Quest'ultimo, infatti, incarnava l'immagine di un'arena orizzontale, capace di emancipare la parola dai vincoli della censura e dai condizionamenti del potere politico ed economico, divenendo simbolo di una libertà finalmente svincolata da mediazioni istituzionali. Sintetizza perfettamente questa immagine la *Declaration of the Independence of Cyberspace* del poeta statunitense John Perry Barlow, nella quale egli scriveva «Governments of the Industrial World, you weary giants of flesh and steel [...] You have no sovereignty where we gather», cristallizzando quella fiducia libertaria nel *cyberspazio* come nuova frontiera dell'autodeterminazione collettiva e individuale<sup>9</sup>.

Tuttavia, come si è detto poc'anzi, l'esperienza ha poi rivelato il rovescio di quell'originaria promessa, mostrando il definitivo tramonto del mito della rete come spazio neutrale e incontaminato. Alla visione di un'arena comunicativa aperta e orizzontale si è sostituita la realtà di un ecosistema informativo governato da pochi attori privati, titolari di un potere inedito di selezione e orientamento dei contenuti. Le stesse infrastrutture che avrebbero dovuto "liberare" la parola si sono trasformate in dispositivi di filtraggio e profilazione, nei quali la visibilità dei messaggi è determinata da logiche di mercato più che da criteri di rilevanza o veridicità delle informazioni. A ciò si aggiunge un ulteriore elemento di criticità: la rete, che in astratto avrebbe dovuto ampliare lo spazio della libertà comunicativa, si è rivelata anche un potente veicolo di contenuti che, pur non necessariamente illeciti, si dimostrano parimenti nocivi o manipolatori. La combinazione tra l'accesso illimitato, l'assenza di "filtri" professionali e la presenza di algoritmi di amplificazione ha reso più agevole la diffusione di messaggi falsi, distorti o deliberatamente ingannevoli, capaci di incidere non solo sull'informazione individuale, ma sull'integrità stessa del dibattito democratico.

Per qualche anno, le istituzioni hanno osservato questo processo con una prudenziale inerzia, ritenendo, tra l'altro, che la rapidità dell'innovazione tecnologica rendesse inadeguato un intervento normativo, che sarebbe repentinamente tramutato in un atto obsoleto. Di qui il ricorso a forme di autoregolamentazione, con le quali le piattaforme si sono impegnate – su base volontaria – a contrastare la disinformazione e i discorsi d'odio attraverso codici di condotta e *policy* interne.

Questa strategia, tuttavia, ha presto mostrato i propri limiti: l'assen-

9. J. Perry Barlow, *A Declaration of the Independence of Cyberspace*, in *Stanford Law Review*, 1996.

za di controlli esterni e di criteri uniformi ha reso infatti evidente che la *self-regulation* sia una pratica meramente discrezionale e, come tale, suscettibile di decisioni arbitrarie e spesso guidate da interessi economici piuttosto che da principi posti a tutela dei diritti fondamentali. Quel che ne deriva, come anticipato poc' anzi, è una forma di privatizzazione della censura, nella quale attori privati esercitano poteri pubblici senza il corrispettivo di adeguate garanzie di trasparenza o di ricorso.

È proprio per superare tale squilibrio che l'Unione europea ha progressivamente elaborato un modello alternativo, fondato su una logica di coregolamentazione e, più di recente, sull'adozione di strumenti di *hard law*<sup>10</sup>. Diversamente dall'approccio statunitense – ancorato a una concezione negativa della libertà di espressione – e da quello cinese, centrato sul controllo politico del discorso, l'Europa ha perseguito un paradigma *rights-based*, nel quale libertà e responsabilità coesistono come poli inscindibili della cittadinanza digitale. Il percorso che conduce dal Codice di condotta del 2018 fino al Digital Services Act<sup>11</sup> e all'Artificial Intelligence Act<sup>12</sup> riflette questa ambizione: riportare la *governance* della comunicazione *online* entro il perimetro del diritto, garantendo la responsabilità algoritmica, la trasparenza dei processi decisionali automatizzati e il bilanciamento tra libertà di espressione, pluralismo e protezione degli utenti.

Da questa tensione prende le mosse la domanda che orienta il presente contributo: le misure oggi vigenti – in particolare il Dsa e l'AI Act – sono davvero idonee a garantire un'informazione pluralista e affidabile nell'ecosistema digitale o rischiano invece di tradursi in un sacrificio eccessivo della libertà di manifestazione del pensiero?

## 2. Intelligenza artificiale e manipolazione del consenso: la democrazia nella società algoritmica

Come anticipato nel paragrafo che precede, una dialettica democratica salubre presuppone un circuito informativo aperto, pluralista e quanto più possibile attendibile. Quando tale circuito viene saturato dall'immissione di con-

10. A dire il vero, prima dell'adozione del Digital Services Act non vigevo soltanto un regime di *self-regulation*, essendo già stata emanata la Direttiva n. 31 del 2000 nota come Direttiva "E-commerce". Pur rimanendo una forma di regolamentazione piuttosto debole in materia, infatti poi superata con estrema velocità dall'evoluzione tecnologica, essa rappresentava senz'altro un intervento di *hard law* degno di menzione.

11. Regolamento UE n. 2065 del 2022.

12. Regolamento UE n. 1689 del 2024.

tenuti falsi o fuorvianti, il diritto dei cittadini a essere informati subisce un *vulnus* che incide sulla libera formazione della volontà popolare, momento essenziale per l'esercizio della relativa sovranità. In quest'ottica, l'informazione, tradizionalmente intesa come "cane da guardia della democrazia" o "quarto potere" dello Stato<sup>13</sup>, svolge un ruolo imprescindibile in ogni ordinamento costituzionale, perché «soltanto un'opinione pubblica consapevole, che si nutre di un processo costante di diffusione delle idee, può dare effettività al diritto dei cittadini di concorrere alla determinazione della politica nazionale», fornendo loro gli strumenti – e dunque le informazioni necessarie – per esercitare consapevolmente i propri diritti politici<sup>14</sup>.

In questa prospettiva, la disinformazione non si esaurisce soltanto in un problema di accuratezza delle notizie riportate o di veridicità dei fatti esposti, ma rappresenta un fenomeno delicato, capace di incidere direttamente sulla legittimazione del potere democratico.

A rendere ulteriormente complesso il già precario ecosistema informativo è intervenuto lo sviluppo dei sistemi di intelligenza artificiale, che ha favorito la produzione e la successiva proliferazione bulimica di contenuti sintetici, amplificando i rischi di una cattiva informazione. Queste tecnologie – oggi impiegate tanto nella forma generativa, capace di creare autonomamente contenuti testuali, visivi o sonori di elevata verosimiglianza, quanto in quella moderativa, utilizzata dalle piattaforme per determinare la visibilità, l'ordine e la rilevanza nella modalità di diffusione dei contenuti condivisi dagli utenti, secondo criteri di mercato e di massimizzazione dell'*engagement* – esercitano un'influenza crescente sul modo in cui le informazioni vengono generate, veicolate e, soprattutto, recepite.

E proprio l'interazione tra IA generativa e IA moderativa incide profondamente sulla struttura del discorso pubblico, perché ciò che deriva dalla loro combinazione è un ambiente comunicativo polarizzato, nel quale i contenuti più divisivi e sensazionalistici – spesso falsi o fuorvianti – prevalgono su quelli più complessi o analitici, alimentando un processo di radicalizzazione delle opinioni e di impoverimento del confronto pubblico. La disinformazione non appare più come una deviazione patologica, ma come un effetto fisiologico di un ecosistema informativo governato da logiche di attenzione e profitto.

13. Come noto, si tratta di principi ormai radicati nella giurisprudenza costituzionale italiana e, soprattutto, della Corte di Strasburgo, ove la stampa e i giornalisti in particolare sono definiti i "*public watchdog*" della democrazia.

14. G.E. Vigevari, *Informazione e potere*, in *Enciclopedia del diritto*, a cura di M. Cartabia e M. Ruotolo, Giuffrè, V, 2023, p. 230.

I rischi di manipolazione si accentuano nei momenti di competizione elettorale, quando la sovrapposizione tra tecniche di *microtargeting* politico, *profiling* comportamentale e uso coordinato di *social bot* o *cyborg accounts* consente di orientare in modo mirato la comunicazione verso specifici segmenti specifici di elettorato. Le elezioni presidenziali statunitensi del 2016, il referendum sulla Brexit e, più di recente, le elezioni rumene del 2024 mostrano quanto fragile sia la linea di confine tra libertà di informazione e manipolazione del consenso: anche in assenza di prove definitive in merito a esiti elettorali concretamente alterati dall'impiego di sistemi di IA, il rischio di interferenze mirate appare ormai sempre più tangibile<sup>15</sup>.

In questo orizzonte, assume rilievo la scelta del legislatore europeo di riconoscere, all'interno del Regolamento europeo sull'intelligenza artificiale, i processi democratici come bene giuridico meritevole di protezione. Pur non rappresentando una normativa "sulla disinformazione", il regolamento individua nelle classificazioni per rischio come particolarmente sensibili i sistemi di IA idonei a influenzare l'esito delle elezioni o a orientare il comportamento di voto, di fatto riconoscendo che la tutela della democrazia non si esaurisce nel diritto formale al suffragio, ma include anche le condizioni cognitive e comunicative che ne garantiscono l'effettività.

Tale impostazione comporta due importanti implicazioni costituzionali: in primo luogo, questa anticipa l'intervento della tutela. La protezione dei diritti politici, infatti, non si limita più al momento dell'espressione del voto, ma si estende alla qualità del circuito informativo che lo precede e lo rende possibile. In secondo luogo, essa solleva il problema della misurabilità del pregiudizio: il modello di regolazione del rischio, centrato su nessi causali e soglie di danno individuale, fatica a cogliere gli effetti diffusi e cumulativi che erodono la qualità della deliberazione pubblica senza incidere immediatamente sul singolo individuo.

È in virtù di questa tensione – tra tutela preventiva e garanzia sostanziale, tra rischio individuale e rischio sistemico – che ci si interroga sulle modalità di preservazione del diritto dei cittadini a essere informati e, conseguentemente, l'integrità dei processi democratici in un ecosistema informativo "mediato" dall'intelligenza artificiale, senza con ciò sacrificare la libertà di espressione e il pluralismo informativo.

E il punto di raccordo tra queste istanze, allo stato, pare doversi ricercare nel quadro regolatorio ora vigente e, in particolare, nel rapporto tra Digital Services Act e AI Act.

15. G.E. Vigevani, *Informazione e potere*, cit., p. 230.

### 3. Dalla *soft law* alla *Constitutional governance*: il modello europeo di contrasto alla disinformazione *online* tra Dsa e AI Act

Una volta delineato il perimetro del problema, chiarito perché la disinformazione costituisca oggi una minaccia più grave rispetto al passato ed evidenziate le sue ricadute sui fondamenti costituzionali della democrazia rappresentativa, occorre interrogarsi sulle risposte che l'Unione europea ha progressivamente elaborato per fronteggiare il fenomeno e sulle peculiarità che distinguono l'approccio europeo rispetto ad altri contesti ordinamentali.

Svanito, infatti, l'iniziale entusiasmo che aveva accompagnato la nascita del *cyberspazio*<sup>16</sup> – un fervore che, come accade nelle svolte epocali, si intrecciava a timori distopici e ad allarmismi fisiologici – si è gradualmente fatta strada la consapevolezza dell'esigenza di intervenire sul "lato oscuro della rete"<sup>17</sup>: la diffusione di contenuti nocivi, l'incitamento all'odio, la radicalizzazione dei discorsi pubblici e, più di recente, la manipolazione informativa integrata mediante i sistemi di intelligenza artificiale hanno progressivamente rivelato la necessità di un intervento regolatorio volto a limitare i rischi generati da simili fenomeni<sup>18</sup>.

Nei primi anni del dibattito europeo sul tema, la scelta – lo si è anticipato – è stata quella del *laissez-faire*, giustificata da una pluralità di ragioni: innanzitutto, le tendenze anarcoidi che avevano connotato l'internet degli albori; in secondo luogo, la fiducia nella capacità autocorrettiva del libero mercato delle idee; ancora, la percezione che la rigidità del diritto positivo fosse intrinsecamente incompatibile con la mutevolezza e la rapidità evolutiva delle tecnologie digitali, a fronte di operatori privati dotati di risorse tecniche e capacità di intervento più immediate; infine, la convinzione che una normativa vincolante potesse compromettere la libertà d'impresa di tali attori e, in ultima analisi, incidere sulla stessa libertà di espressione.

A ciò si aggiungeva la difficoltà di mettere a fuoco e di inquadrare giuridicamente quella «zona grigia» che connota in misura peculiare i contenuti disinformativi<sup>19</sup>: messaggi ambigui o "*borderline*", formalmente leciti – perché non

16. Sul punto, si veda Christopher S. Yoo, *Free Speech and the Myth of the Internet as an Unintermediated Experience*, in *George Washington Law Review*, vol. 78, n. 4, 2010; M. Bassini, *Fundamental rights and private enforcement in the digital age*, in *European Law Journal*, n. 25, 2019, pp. 184-187.

17. Per un approfondimento sul tema, si veda, tra gli altri, A. Missiroli, *The Dark Side of the Web: Cyber as a Threat*, in *European Foreign Affairs Review*, n. 24, 2019, pp. 135-152.

18. D. Keller, *Lawful but Awful? Control over Legal Speech by Platforms, Governments, and Internet Users*, in *The University of Chicago Law Review*, 2022.

19. G.E. Vigevani, *Piattaforme digitali private, potere pubblico e libertà di espressione*, in *Diritti di libertà e nuove tecnologie*, cit., p. 44.

lesivi di altri diritti o libertà costituzionalmente rilevanti – ma potenzialmente idonei a compromettere la qualità del dibattito pubblico e ad alterare il regolare svolgimento dei processi democratici. Tutti questi elementi indussero le istituzioni europee a ritenere, in un primo momento, «che l'autoregolamentazione dei titolari delle piattaforme bastasse a contrastare la disinformazione»<sup>20</sup>, evitando il rischio che un intervento normativo eccessivamente prescrittivo si traducesse in un ostacolo alla libertà di manifestazione del pensiero.

Su tali premesse si fondava il Codice di condotta sulla disinformazione del 2018<sup>21</sup>, espressione di quella regolamentazione volontaria che impegnava le principali piattaforme digitali a introdurre misure interne di contrasto alla diffusione di notizie false, a migliorare la trasparenza nella pubblicità politica e a collaborare con i verificatori di fatti indipendenti.

Questo primo tentativo, tuttavia, rivelò ben presto i propri limiti strutturali: l'assenza di obblighi giuridicamente vincolanti, la mancanza di meccanismi sanzionatori, l'opacità delle procedure di attuazione e la sostanziale discrezionalità nella definizione degli standard ridussero l'efficacia del codice a quella di un mero strumento reputazionale.

Nel frattempo, la centralità assunta da questi attori rese evidente l'ampiezza del potere da essi esercitato. È oggi innegabile che il concreto esercizio della libertà di espressione e il godimento effettivo del diritto all'informazione dipendano, in larga misura, dalla presenza dell'individuo all'interno delle piattaforme, rivelandosi così condizionato dalle regole di accesso e di visibilità imposte da tali “nuovi governatori” dello spazio digitale<sup>22</sup>.

Pertanto, la consapevolezza dell'insufficienza di tale modello ha condotto, in una seconda fase<sup>23</sup>, all'elaborazione di un paradigma di *coregolamentazione*, fondato sulla cooperazione tra poteri pubblici e attori privati. L'obiettivo dichiarato era quello di responsabilizzare le piattaforme senza esautorare la loro capacità tecnica e innovativa, mantenendo un equilibrio tra la flessibilità operativa e la salvaguardia dei valori costituzionali dell'ordinamento europeo.

In tale prospettiva, le istituzioni europee hanno perseguito una logica di

20. C. Pinelli, U. Ruffolo, *I diritti nelle piattaforme*, Giappichelli, 2023, p. 13.

21. Istituito nel 2018, come noto, il Codice di buone pratiche è stato poi notevolmente rafforzato prima nel 2022, con l'obiettivo di essere riconosciuto come Codice di condotta ai sensi della legge sui servizi digitali, e poi il 13 febbraio 2025, quando la Commissione e il Comitato europeo per i servizi digitali hanno approvato l'integrazione del Codice di buone pratiche sulla disinformazione del 2022 come Codice di condotta sulla disinformazione nel quadro della legge sui servizi digitali.

22. K. Klonick, *The New Governors: The People, Rules, and Processes Governing Online Speech*, in *Harvard Law Review*, n. 131, 2018.

23. Sulle fasi di intervento europeo di contrasto alla disinformazione, si veda O. Pollicino, P. Dunn, *Intelligenza artificiale e democrazia*, cit., pp. 127-134.

*shared responsibility*: lo Stato definisce gli obiettivi di interesse pubblico (trasparenza, sicurezza, tutela dei diritti fondamentali), mentre le piattaforme conservano la libertà di individuare gli strumenti più idonei al loro raggiungimento.

È in questo quadro di progressiva maturazione che si inserisce, nel 2022, l'adozione del Digital Services Act, che ha costituito un vero e proprio punto di svolta nella regolazione dello spazio digitale europeo. Il Dsa rappresenta, infatti, il primo tentativo di istituzionalizzare la cooperazione tra poteri pubblici e privati, attraverso l'introduzione di obblighi giuridici vincolanti per le *very large online platforms* (Vlop) e i *very large online search engines* (Vlose), chiamati a valutare annualmente i rischi sistemici derivanti dai propri servizi, tra i quali rientrano espressamente la disinformazione, la manipolazione elettorale e le interferenze informative.

Pertanto, la logica sottesa a detto Regolamento si dimostra essenzialmente prevenzionistica e orizzontale, poiché non mira alla rimozione selettiva dei contenuti, bensì alla gestione del rischio attraverso strumenti di trasparenza, *risk assessment* e *audit* indipendenti sugli algoritmi. Si tratta, in altri termini, di una regolazione che, più che incidere sul singolo contenuto, interviene sulle condizioni strutturali della comunicazione digitale, riconducendo la responsabilità delle piattaforme a un piano sistemico.

Nel solco di tale approccio, e in perfetta sintonia con la logica della “società del rischio” che permea il diritto europeo contemporaneo<sup>24</sup>, l'Unione ha ulteriormente arricchito il proprio quadro normativo con l'adozione nel 2024 del Regolamento europeo sull'Intelligenza Artificiale, adottato al fine di mettere a punto «un assetto giuridico adeguato alla complessità e poliedricità della tecnologia» in questione<sup>25</sup>. Il risultato è una disciplina che pare confermare il modello di regolazione europeo orientato su una logica *rights-based*, ove la tutela dei diritti fondamentali non è considerata un limite alla libertà ma, al contrario, la condizione del suo esercizio pieno e responsabile.

Già nel 2018 la Commissione europea aveva adottato il Piano coordinato sull'intelligenza artificiale, ove aveva tentato di individuare «le coordinate attorno alle quali dovesse ruotare l'intelligenza artificiale»<sup>26</sup>, evidenziando che soltanto la sussistenza di un «quadro di riferimento chiaro e completo» avrebbe consentito ai cittadini di avere fiducia nello sviluppo tecnologico<sup>27</sup>.

24. O. Caramaschi, *Il costituzionalismo al cospetto dell'intelligenza artificiale: nuove sfide, quali soluzioni?*, in *Rivista italiana di informatica e diritto*, n. 1/2025, p. 30.

25. M. Bassini, *Intelligenza Artificiale generativa: alcune questioni problematiche*, in *Rivista di diritto dei media*, n. 2/2023, p. 392.

26. D. Martire, *Intelligenza artificiale e Stato costituzionale*, in *Diritto Pubblico*, n. 2/2022, p. 404.

27. *Ibidem*.

In tal senso, l'AI Act rappresenta senz'altro un tentativo degno di plauso, al quale si deve riconoscere il merito di aver messo a punto una normativa di carattere settoriale che considera l'IA come un insieme di tecnologie rischiose. Si tratta, nello specifico, di tecnologie a rischi graduati, classificati secondo il potenziale impatto sui diritti e sulla sicurezza degli individui e della collettività.

Pur non affrontando in modo diretto il tema della disinformazione, il Regolamento lo richiama indirettamente all'art. 50, imponendo ai fornitori di sistemi di IA capaci di generare contenuti sintetici – testi, immagini, audio, video – obblighi di *disclosure* e tracciabilità, volti a evitare che tali prodotti possano essere confusi con informazioni autentiche.

Tuttavia, come si anticipava nel paragrafo che precede, la vera novità concettuale pare risiedere nella qualificazione dei processi democratici come ambiti “ad alto rischio”. L'AI Act, infatti, include espressamente tra i sistemi sottoposti a obblighi rafforzati quelli destinati a influenzare l'esito delle elezioni o delle consultazioni referendarie, nonché a incidere sul comportamento di voto degli elettori.

In tal modo, il legislatore europeo riconosce in modo esplicito che la tutela della democrazia nell'era dell'algoritmo non si esaurisce nel momento formale del voto, ma comprende anche le condizioni cognitive e informative che ne determinano la genuinità.

Si tratta di un passaggio concettuale di grande rilievo, perché anche in assenza di un espresso riferimento alla disinformazione l'AI Act ne presuppone l'incidenza potenziale, collocando la protezione dei processi deliberativi e dell'opinione pubblica tra le finalità stesse della regolazione.

Resta, tuttavia, un nodo irrisolto: il rapporto tra Dsa e AI Act. Se il primo regola la circolazione dei contenuti *online*, il secondo, invece, disciplina i sistemi che li generano o li elaborano. Entrambi perseguono la tutela dei diritti fondamentali nello spazio digitale, ma si muovono lungo assi regolatori differenti: *systemic risk management* nel Dsa, *risk-based approach* tecnologico nell'AI Act. Ci si chiede, dunque, se – e in quale misura – le garanzie poste dal Dsa possano o debbano estendersi anche ai *deployer* e ai *provider* di sistemi di intelligenza artificiale, e se un simile ampliamento possa realmente rafforzare la tutela del diritto all'informazione, senza introdurre forme di controllo preventivo incompatibili con la libertà di espressione<sup>28</sup>. La risposta non sembra risiedere in un'estensione formale dell'applicabilità, bensì in una integrazione sistematica e funzionale tra i due regimi: un dialogo normativo capace di preservare il nu-

28. Sul punto, si veda M. Bassini, *Speech without a Speaker: Constitutional Coverage for Generative AI Output?*, in *European Constitutional Law Review*, 2025.

cleo duro della libertà di manifestazione del pensiero e, al contempo, di garantire la responsabilità algoritmica degli attori privati che operano nell'ecosistema informativo digitale.

#### 4. *Pars construens*: l'IA al servizio del pluralismo informativo

Le considerazioni sin qui svolte mostrano come la disinformazione, amplificata dall'intermediazione algoritmica e dalla pervasività delle piattaforme digitali, rappresenti oggi una delle principali sfide per la tenuta della democrazia.

Resta tuttavia una domanda di fondo, che travalica il piano strettamente normativo: è possibile che la stessa tecnologia che alimenta la crisi del discorso pubblico diventi anche parte della sua soluzione?

L'intelligenza artificiale, in effetti, si presenta come un fenomeno intrinsecamente ambivalente, rappresentando certamente una fonte di rischio per la diffusione di contenuti disinformativi, ma al contempo anche una potenziale risorsa per contrastarne gli effetti e promuovere la produzione di informazione pluralistica e di qualità<sup>29</sup>. Se da un lato, infatti, i sistemi generativi e predittivi possono amplificare la circolazione di contenuti falsi, polarizzanti o sintetici, dall'altro le medesime architetture possono essere ripensate in funzione garantista, per promuovere la diversità delle fonti, la trasparenza delle scelte algoritmiche e l'inclusione di prospettive eterogenee nello spazio informativo.

È in questo orizzonte che si ritiene di poter inserire una riflessione su quello che la dottrina ha definito "*diversity by design*", concetto elaborato nell'ambito dell'ingegneria dei sistemi informativi per indicare la possibilità di intervenire, sin dalla fase di progettazione degli algoritmi, prevedendo dei meccanismi che favoriscano l'esposizione degli utenti a contenuti diversificati e che contrastino la formazione di bolle informative<sup>30</sup>.

Trasposto nel dibattito giuridico e politologico, tale principio può divenire una chiave di lettura per le politiche europee sulla *media diversity*: non più solo come obiettivo di *policy* da perseguire *ex post* mediante i più canonici strumenti regolatori, ma come criterio intrinseco di progettazione degli ambienti digitali. In altre parole, il *diversity by design* rappresenta la traduzione in termi-

29. Sul tema, si rimanda al saggio di G. Contissa, F. Galli, *La governance della "disinformazione aumentata" tra Digital Services Act e AI Act*, in *federalismi.it*, n. 15, 2025, pp. 129-131. Per una puntuale ricostruzione delle numerose modalità di impiego "positivo" dei sistemi di intelligenza artificiale si rimanda a O. Pollicino, P. Dunn, *Intelligenza artificiale e democrazia*, cit., pp. 150-153.

30. N. Helberger, *Diversity by design*, in *Journal of Information Policy*, n. 1/2011, pp. 441-469.

ni “tecnologici” del principio del pluralismo informativo: volto a incorporare nei codici delle piattaforme quei valori di apertura e molteplicità che l’art. 21 Cost. e l’art. 11 Cdfue pongono a presidio della sfera pubblica democratica, sostanzialmente orientando l’algoritmo a promuovere l’esposizione degli utenti a informazioni diversificate, invece di consolidare meccanismi di chiusura cognitiva come le *filter bubbles* e le *echo chambers*. In questi termini, se orientata secondo i riconosciuti principi di trasparenza, *accountability* e “diversità”, l’intelligenza artificiale potrebbe così svestire gli abiti del vettore della crisi per divenire alleata nella garanzia della qualità del discorso pubblico.

Tuttavia, anche tale prospettiva non è esente da criticità. Se l’IA viene istruita privilegiando soltanto fonti considerate “autorevoli” si corre il rischio di un appiattimento informativo che soffoca l’eterogeneità dei punti di vista e riduce il pluralismo a mera omologazione. Se, al contrario, il sistema viene nutrito di dati viziati, incompleti o polarizzati, il rischio è opposto: la diversità non sarebbe più una misura posta a sostegno del pluralismo, bensì una forma sofisticata di relativismo informativo.

In entrambi i casi pare riaffiorare una domanda antica, che nessuna tecnologia riesce a eludere: chi decide che cosa è vero e secondo quali criteri? L’illusione di poter delegare a un algoritmo il discernimento tra il vero e il falso porterebbe infatti a riprodurre, in forma automatizzata, le stesse asimmetrie di potere che il diritto costituzionale si propone di evitare.

Da questa prospettiva, pertanto, la sfida non è quella di integrare al *diversity by design* un modello di “verità programmata”, ma nel costruire un equilibrio dinamico tra apertura e affidabilità: un sistema che preservi la pluralità delle opinioni senza rinunciare all’attendibilità delle informazioni; un sistema in cui la tecnologia non pretenda di determinare la verità, ma si limiti a rendere le condizioni della sua ricerca più accessibili e trasparenti.

In quest’ottica, la tecnica e il diritto non si porrebbero più in contrapposizione, ma come strumenti complementari di una stessa architettura regolatoria: la prima come strumento di attuazione del secondo, in un’ottica di co-evoluzione tra garanzie costituzionali e innovazione tecnologica. Forse, raggiunto un simile equilibrio, l’intelligenza artificiale potrà realmente divenire alleata nella tutela del discorso democratico, contribuendo a preservare un flusso informativo aperto, affidabile e plurale.