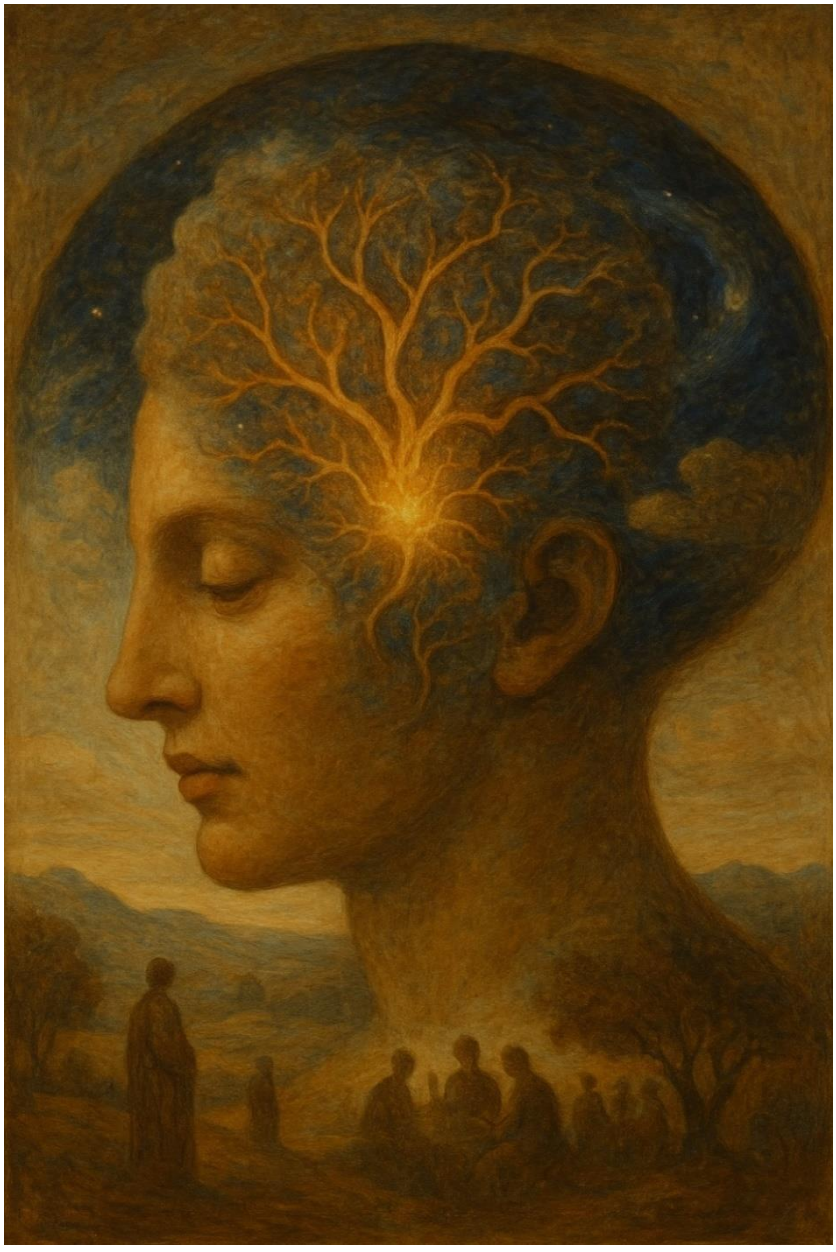
AUTORITÀ PER LE GARANZIE NELLE COMUNICAZIONI
AGCOM

Intelligenza Artificiale



II PARTE

Rapporto del Comitato sull'Intelligenza Artificiale, 2026

Sommario

Executive Summary

Il presente Rapporto costituisce il primo contributo organico del Comitato di Esperti sull'intelligenza artificiale istituito dall'Autorità per le Garanzie nelle Comunicazioni (AGCOM). Il Comitato è stato incaricato di analizzare le questioni giuridiche e regolatorie che l'Autorità dovrà affrontare nel breve e medio periodo, in un contesto segnato da una profonda trasformazione tecnologica e da un quadro normativo europeo ancora in piena evoluzione.

Il Rapporto si articola in sette capitoli, ciascuno affidato a uno o più esperti del Comitato, e si conclude con un capitolo di sintesi sulle prospettive di regolazione e governance dell'IA generativa. Di seguito si riassumono i principali contenuti e le raccomandazioni emerse dai singoli contributi.

L'introduzione, a cura di Andrea Renda, esamina la "fase di crisi" dell'AI Act europeo. Nato come strumento pionieristico per la regolazione dell'intelligenza artificiale con un approccio basato sul rischio, il Regolamento ha subito una trasformazione radicale a seguito dell'emergere dei sistemi di IA a uso generale (GPAI). L'entrata sul mercato di ChatGPT nel novembre 2022 ha costretto i co-legislatori europei a riaprire il negoziato, introducendo un corpus di regole specifiche per i GPAI e un nuovo sistema di governance incentrato sull'AI Office della Commissione europea. Il testo definitivo soffre tuttavia di lacune strutturali in materia di governance adattiva ed è sottoposto a crescenti pressioni verso la deregolamentazione — sia di origine interna (proposta di Digital Omnibus) sia internazionale (postura dell'amministrazione statunitense). Il capitolo identifica tre aree di debolezza fondamentali: l'assenza di un'architettura regolatoria sufficientemente flessibile, il carattere ibrido e non pienamente applicabile della normativa, e la carenza di un'interfaccia adeguata tra il regolatore europeo e le autorità di settore nazionali — tra cui l'AGCOM.

Il contributo di Giovanna de Minico (Capitolo 2) ricostruisce le competenze di AGCOM all'incrocio tra il Digital Services Act (DSA) e l'AI Act. Sebbene l'Autorità non sia stata espressamente designata come autorità di supervisione dell'IA, essa può esercitare i

propri poteri di enforcement — vigilanza, ordine, sanzione — ogniqualvolta i sistemi di IA, e in particolare gli LLM, vengano utilizzati come vettori di contenuti illeciti su piattaforme digitali di sua competenza. La competenza di AGCOM si fonda su un titolo indiretto, basato sulla sua qualità di Digital Services Coordinator (DSC) ai sensi del DSA, che le attribuisce poteri di intervento su fenomeni come la disinformazione, l'hate speech, la pubblicità occulta e la violazione del pluralismo dell'informazione. Il contributo esamina altresì le competenze in materia di pubblicità politica online (Reg. UE 2024/900), evidenziando le lacune di un regime limitato agli obblighi di trasparenza, che non introduce limiti quantitativi né vieta tecniche di manipolazione emotiva. Un ampio approfondimento teorico sul rapporto tra tecnica e diritto — inteso come "mosaico delle regole nell'ecosistema digitale" — completa il contributo, esaminando le condizioni di legittimità costituzionale dell'autoregolazione e il rischio che la tecnica si sostituisca alla politica come fonte del diritto.

Il capitolo 3, a cura di **Andrea Imperiali di Francavilla,** offre una mappatura sistematica e operativamente rilevante degli impatti dell'IA in tutti i settori di competenza di AGCOM: comunicazioni elettroniche, servizi digitali e media audiovisivi, servizi postali e tutela dei consumatori. Per ciascun settore sono identificati gli impatti effettivi e potenziali, le priorità di intervento e gli strumenti regolatori disponibili (ex ante, in itinere, ex post). Nelle comunicazioni elettroniche emergono come urgenti i rischi di discriminazione algoritmica del traffico dati (*throttling*, *zero rating* dinamico), la collusione algoritmica dei prezzi e la perdita di controllo umano nella gestione delle reti. Nel settore media si segnalano le minacce al pluralismo poste dagli algoritmi di raccomandazione, la proliferazione dei deepfake e la necessità di un regime di etichettatura obbligatoria dei contenuti generati dall'IA. Il capitolo include un benchmark comparativo dei principali regolatori europei (Ofcom, Arcom, BNetzA, BEREC, ERGA) e una ricostruzione del mosaico normativo applicabile (AI Act, DSA, GDPR, Digital Omnibus), con indicazioni operative per AGCOM su come costruire capacità tecnica interna, strumenti di testing indipendente e sistemi di monitoraggio permanente.

Nei capitoli 4 e 5, **Andrea Simoncini** e **Giovanni Boccia Artieri** analizzano la dimensione costituzionale della sfida posta dall'IA alla democrazia e alla libertà di informazione. Le piattaforme digitali, configurate come "formazioni sociali" che esercitano un potere normativo privato su scala globale, ridisegnano l'ecosistema dell'informazione e del discorso pubblico in modo difficilmente compatibile con i principi del costituzionalismo liberale. Il capitolo di Simoncini esamina il caso Romania come "laboratorio costituzionale" per le implicazioni dell'IA sulla sovranità democratica, e propugna un approccio di "costituzionalismo digitale europeo" capace di vincolare le piattaforme al rispetto dei diritti fondamentali e di riequilibrare il rapporto tra potere privato e sfera pubblica. Boccia Artieri esamina i rischi sistemici connessi all'uso dell'IA generativa come "infrastruttura cognitiva" del discorso pubblico. L'IA non è più un semplice strumento tecnico ma un ambiente epistemico che ridefinisce i criteri di credibilità, verità e autorevolezza. Il capitolo identifica tre vettori di rischio principali: la manipolazione informativa (deepfake, voice cloning, disinformazione sintetica scalabile), la degradazione discorsiva (hate speech amplificato algoritmicamente) e i meccanismi dell'economia dell'attenzione. Per ciascun rischio vengono proposti strumenti regolatori specifici — dai sistemi di provenance e labeling ai protocolli di risposta rapida ai contenuti lesivi — raccomandando ad AGCOM di posizionarsi come "infrastruttura della fiducia" nel panorama comunicativo italiano.

Nel capitolo 6, **Giuseppe Cassano** ricostruisce la problematica dell'autorialità e della protezione del diritto d'autore in un ecosistema dominato dall'IA generativa. Il capitolo esamina l'evoluzione del quadro normativo italiano, europeo e comparato in relazione alle opere create con l'ausilio dell'IA — distinguendo essenzialmente tra il contributo intellettuale umano e la generazione autonoma da parte della macchina. La legge italiana n. 132/2025 sull'intelligenza artificiale estende la protezione del diritto d'autore alle opere che riflettono il lavoro intellettuale dell'autore, anche se assistite dall'IA. Il capitolo analizza altresì le implicazioni del training dei modelli su dataset

protetti da copyright e individua nel Codice di Buone Pratiche per l'IA uno strumento per bilanciare i diritti dei creatori con le esigenze di sviluppo dei sistemi di IA.

Nel capitolo 7, **Mauro Giusto** affronta la dimensione educativa della sfida posta dall'IA, partendo dal "paradosso della generazione più connessa ma meno consapevole". I giovani sono esposti a rischi specifici — dall'abbandono educativo ai danni sull'identità, dalla dipendenza tecnologica all'esposizione ai deepfake — che richiedono una risposta sistemica che coinvolga scuola, famiglie, imprese e istituzioni. Il capitolo propone il concetto di "algoretica" — educazione al ragionamento critico nell'era degli algoritmi — come bussola per politiche pubbliche di alfabetizzazione digitale, e individua in AGCOM un attore fondamentale per la promozione di standard di protezione dei minori nell'ecosistema digitale.

Da ultimo, il capitolo 8 (a cura di **Giovanna de Minico**) approfondisce in chiave teorico-costituzionalistica il rapporto tra tecnica e diritto nell'era dell'intelligenza artificiale, proponendo una riflessione originale sul rischio che la tecnica si sostituisca alla politica come fonte autonoma del diritto. Partendo da una genealogia filosofica che va da Eschilo ad Aristotele fino ai mercati digitali contemporanei, il capitolo dimostra come la tecnica abbia progressivamente eroso la posizione di primazia della norma giuridica: da strumento servente della politica, essa si è trasformata in un attore normativo autonomo — e sovente opaco — che interviene sugli stessi oggetti del diritto con forza crescente e scarsa legittimazione democratica. Il capitolo analizza tre meccanismi attraverso cui la tecnica penetra nel sistema delle fonti: (i) il completamento discrezionale delle norme antitrust "aperte" attraverso parametri tecnici definiti caso per caso dall'Autorità, attribuendo all'Antitrust una funzione ibrida tra *iuris dictio* e indirizzo politico; (ii) la trasformazione dell'autoregolazione privata in fonte del diritto nell'ecosistema digitale, con le condizioni di legittimità costituzionale che tali codici-fonti devono rispettare per non degenerare in forme di potere privato incontrollato; (iii) il rischio che i sistemi di IA divengano direttamente *source of law*, emettendo statuizioni concrete e particolari che cancellano la certezza del diritto, perpetuano ingiustizie algoritmiche e sostituiscono la *rule of law* con la *rule of tech*.

L'analisi del DSA mostra come i codici volontari ivi previsti rischino di non soddisfare le condizioni di compatibilità costituzionale, mentre l'AI Act — pur nella sua pretesa di orientare la tecnica verso l'uomo — si affida eccessivamente ad autodichiarazioni dei fornitori, rinunciando a controlli pubblici imparziali. La proposta conclusiva riabilita il principio aristotelico del primato della politica sulla tecnica, declinato in chiave sovranazionale europea: il legislatore comunitario deve fare meno (eliminare regole superate e invasive) ma diversamente (riscrivere la disciplina sull'IA privilegiando controlli pubblici imparziali e introducendo rimedi strutturali adeguati contro la dominanza dei grandi operatori digitali). Solo a queste condizioni l'intreccio tra tecnica e politica potrà diventare un fattore di crescita dell'ordinamento democratico, invece che un vettore di sua erosione.

Sulla base dei contributi raccolti, il Comitato formula le seguenti raccomandazioni prioritarie per AGCOM:

- **Potenziamento della capacità tecnica interna:** Investire in un team dedicato di AI auditor, data scientist e giuristi specializzati, in grado di condurre ispezioni algoritmiche autonome e di sviluppare strumenti di testing indipendente sul modello delle esperienze di Ofcom e BNetzA.
- **Coordinamento istituzionale rafforzato:** Stringere protocolli di collaborazione formali con AgID, ACN, Garante privacy e AGCM, garantendo un enforcement coerente e proporzionato del quadro normativo europeo sull'IA nei settori di rispettiva competenza.
- **Istituzione di un Osservatorio permanente sull'IA:** Creare un sistema di monitoraggio continuo degli impatti dell'IA nei settori regolati, con *dashboard* pubbliche, report periodici e strumenti di rilevamento automatizzato di *bias*, discriminazioni e violazioni del pluralismo informativo.
- **Adeguamento del sistema sanzionatorio:** Attivare le sanzioni previste dal DSA (fino al 6% del fatturato globale) e dall'AI Act (fino al 7%), superando i limiti

strutturali della L. 249/1997 che rendono l'enforcement poco deterrente rispetto ai fatturati degli operatori.

- **Sviluppo di un framework regolatorio modulare:** Elaborare codici di condotta settoriali, linee guida e obblighi di trasparenza algoritmica armonizzati tra i diversi ambiti di competenza, fondati su principi comuni adattabili all'evoluzione tecnologica (trasparenza, *human oversight*, non-discriminazione, protezione dei vulnerabili).
- **Partecipazione attiva ai network europei:** Contribuire proattivamente ai gruppi di lavoro BEREC ed ERGA, co-costruendo strumenti comuni per l'audit dell'IA e l'armonizzazione delle pratiche di enforcement a livello europeo.
- **Sperimentazione e SupTech:** Avviare la sperimentazione di strumenti di *supervisory technology* (SupTech) per il monitoraggio automatizzato dei mercati regolati, aprendo la strada a forme di enforcement basate su dati in tempo reale invece che su ispezioni periodiche. L'AGCOM dovrebbe inoltre istituire una *regulatory sandbox* dedicata ai sistemi di IA che operano nei settori di propria competenza — comunicazioni elettroniche, media audiovisivi, servizi digitali e postali.

1 Introduzione. l'Intelligenza Artificiale come opportunità e responsabilità per le autorità indipendenti

Andrea Renda

Materia spinosa, in constante divenire, l'Intelligenza Artificiale (IA) continua a turbare il sonno di molti legislatori e regolatori nel mondo. Peraltro, mentre alcuni paesi appaiono sempre meno intenzionati a porre mano a strumenti legislativi per fronteggiare il rapido incedere dell'IA - che permea settori industriali, la ricerca e la scienza, nonché la vita pubblica e quella di governo – i rischi connessi all'incontrollata diffusione dei modelli di IA più potenti si fanno ogni giorno più evidenti. Non può sfuggire all'occhio attento l'attuale diatriba tra il governo statunitense (in particolare il Dipartimento della Difesa, ovvero della guerra, ai tempi di Trump) e imprese come Anthropic, preoccupate dell'uso indiscriminato che dei propri potentissimi modelli può fare l'esercito, o il Department of Homeland Security sempre più dedito alla sorveglianza di massa. Una *querelle* che, in assenza di una regolamentazione chiara ed efficace, non può che dar adito a casi di concorrenza al ribasso, come evidenziato nello specifico dall'opportunismo del principale rivale di Anthropic, OpenAI, pronta ad accogliere le richieste incondizionate dell'amministrazione Trump, in spregio ai principi etici e ai *guardrail* più basilari. Né sfugge all'occhio più attento la dinamica del consumo di dati, acqua ed energia dei modelli c.d. di frontiera, sui quali le imprese più avanzate spendono centinaia di miliardi di dollari l'anno, nell'attesa di una domanda e di applicazioni che non sembrano ancora prospettarsi all'orizzonte. Da ultimo, a proposito di *killer app*, non sfugge all'occhio attento (e critico) l'uso che dei c.d. Large Language Models si sta facendo in campo militare, come potenti strumenti di

disinformazione (si pensi all'ultima tendenza, quella del *LLM grooming*) e come formidabili alleati degli *hacker* negli attacchi cibernetici.

In Europa, queste e altre tendenze hanno generato una diffusa domanda di ripensamento, *à rebours*, rispetto a un percorso legislativo che, sin qui, si era distinto per la sua ambizione e un chiaro orientamento alla protezione dei diritti fondamentali, nonché della sicurezza dei prodotti e servizi che fanno uso di Intelligenza Artificiale. Un ripensamento che viene da lontano, e che include significative pressioni che le istituzioni comunitarie hanno subito negli ultimi mesi da parte del governo statunitense, come apparso già evidente nel discorso pronunciato dal vicepresidente J.D. Vance in occasione del Paris AI Action Summit del febbraio 2025. Un discorso tutto orientato alla deregolamentazione col fine di rimuovere ogni ostacolo allo sviluppo e alla diffusione dell'IA e accompagnato, in modo forse inconsapevole, dalla rinuncia della Commissione europea a portare avanti la proposta di direttiva sulla responsabilità civile al tempo dell'IA, pendente dal settembre 2022.[1] Un ripensamento che si sposa anche con il rinnovato imperativo che sovrasta l'agenda della nuova Commissione europea, quello della competitività e della semplificazione normativa, evocata dal Rapporto Draghi già nel 2024, e sfociata in una serie di proposte di legislazione c.d. *omnibus*, ivi inclusa una sul digitale che mira a semplificare, tra l'altro, l'AI Act e il regolamento sulla protezione dei dati personali.

Un ripensamento, sia chiaro, in parte dovuto, anche se non per i motivi sin qui esposti. Per capirne i motivi, è però necessario fare un passo indietro, tornando alla fase di gestazione dell'AI Act. Figlio di un Libro Bianco sull'intelligenza artificiale del 2020, che ambiva esplicitamente alla creazione di un ecosistema "della fiducia" e di uno "dell'eccellenza" in questo campo, l'AI Act – almeno nella versione originale proposta dalla Commissione europea – traduceva in termini normativi le linee guida sull'IA affidabile preposte dal Gruppo di esperti di alto livello nominato dalla Commissione europea alcuni anni addietro, nel 2018. Adottando un approccio basato sul rischio, l'AI

[1] Commissione europea – proposta di Direttiva COM(2022) 496: AI Liability Directive sul regime di responsabilità civile extracontrattuale applicato all'intelligenza artificiale.

Act ha introdotto una tassonomia su quattro livelli, che - rimanendo fermamente ancorata al principio di non regolazione della tecnologia in quanto tale - mirava a escludere dal mercato le applicazioni eccessivamente rischiose, e si concentrava sulla regolazione di un numero limitato di applicazioni considerate "ad alto rischio" di violazione dei diritti fondamentali, o in termini di sicurezza. In un approccio che, nonostante denunciato come eccessivamente interventista, appariva effettivamente minimalista, la proposta di regolamento si affidava come base giuridica alla normativa sulla sicurezza dei prodotti, e assoggettava le applicazioni di IA ad alto rischio al quadro (co-)regolamentare europeo che prevede l'apposizione del "marchio CE" come presupposto della circolazione nel mercato interno. Per giunta, nella più parte dei casi nell'*AI act* la verifica della conformità delle applicazioni ad alto rischio ai requisiti regolatori rimane soggetta a "controlli interni" da parte dei soggetti fornitori (*provider*), e solo in una minoranza di casi si prevedeva l'intervento di parti terze come previsto dalla normativa europea sul marchio CE.

1.1 La "fase di crisi" dell'AI Act

Anche se pionieristica e da subito controversa, in realtà la versione iniziale dell'aprile 2021 rappresentava un intervento piuttosto timido, minimalista nell'ambito di applicazione, cauto nell'imporre costi di *compliance*, e accompagnato da una governance alquanto esile. Spiccavano, nella proposta, una definizione di IA neutrale dal punto di vista tecnologico, e passibile di aggiornamento costante dato l'incessante divenire della materia; requisiti regolatori definiti in modo piuttosto vago, da interpretare caso per caso; e una possibilità – ancorché limitata, e per molti limitante – di rivedere la tassonomia delle applicazioni a rischio, poste in allegato alla proposta di regolamento per poterne consentire una più agile revisione; e l'introduzione di *regulatory sandboxes* governate a livello nazionale per facilitare l'accesso al mercato e la verifica di compatibilità regolatoria, soprattutto per le piccole e medie imprese. La proposta di AI Act, presentata in tempi di pandemia, ha generato un grande interesse a livello internazionale, con paesi come Corea del Sud, Brasile, Canada, Colombia inizialmente decisi a emularne l'approccio e i contenuti essenziali. Peraltro, non si

trattava di un approccio interamente orizzontale, nel senso di applicabile in via generale a tutti i settori e agli *use case* dell'IA. Al contrario: da un lato l'approccio basato sul rischio si concentrava su un numero di applicazioni che la Commissione stimava corrispondere a meno del 10% del mercato; dall'altro, alcuni problemi specifici dell'IA, come il caso della disinformazione e della moderazione dei contenuti sulle reti sociali, cadevano al di fuori dell'ambito di applicazione dell'Act, posta la quasi coeva introduzione del Digital Services Act. Quest'ultimo può essere a tutti gli effetti annoverato tra le misure legislative sull'intelligenza artificiale della Ue: eppure, l'introduzione ivi prevista di una analisi di rischio sistemico a opera di terze parti sui sistemi di IA utilizzati dalle grandi piattaforme e motori di ricerca attende ancora un'adeguata opera di riconciliazione con la procedura di valutazione di conformità prevista dall'AI Act, per non parlare dell'analisi del rischio prevista per i sistemi di IA di applicazione generale, di cui si dirà tra breve.

Erano, quelli del 2021-2022, giorni floridi per il dibattito sulla regolamentazione dell'IA. A parte le tracce alquanto evidenti di un incipiente *Brussels effect*, che inorgogliva i funzionari di Bruxelles, si assisteva gradualmente a un risveglio della cooperazione internazionale a partire dagli organismi di standardizzazione (ISO, IEC, IEEE) per arrivare a contesti come l'OCSE, la Global Partnership on AI (GPAI), il G-7 e il G-20, e gradualmente anche le Nazioni Unite grazie al c.d. Global Digital Compact e al Patto per il futuro.

Sennonché, in questa primavera della regolazione, due fattori cruciali hanno finito col "rompere il giocattolo" dell'AI Act. Il primo coincide con la graduale presa di coscienza della complessità e della continua trasformazione della materia oggetto di regolazione. Già nel settembre 2022, Engler e Renda (2022) avevano posto in evidenza che l'enfasi dell'AI Act sul c.d. "provider", l'entità che immette il sistema o modello di IA sul mercato, o lo mette in servizio con il proprio nome o marchio, mal si adattava a una struttura di mercato che andava concentrandosi nelle mani di imprese che sviluppavano sistemi versatili, offrendoli a fornitori *downstream* che avevano limitata contezza delle specifiche di design e sviluppo dei modello stessi. In altre e più banali parole, concentrarsi sul fornitore significava mancare il bersaglio, e porre sulle spalle del

soggetto meno informato - certo non il c.d. *cheapest cost avoider* – la responsabilità di assicurare la conformità dell'applicazione di IA ai requisiti regolatori dell'AI Act. Mentre il Parlamento europeo si affrettava a introdurre obblighi di cooperazione lungo la catena del valore dell'IA, tra soggetti *upstream* e fornitori, l'entrata sul mercato di ChatGPT nel novembre 2022 consacrò definitivamente la necessità di porre mano in modo ben più strutturale all'impianto concettuale e giuridico dell'AI Act. Di lì a poco, nel febbraio 2023, i co-legislatori decisero di riaprire la negoziazione sul testo del regolamento, dando la stura a un trilogo che ne avrebbe modificato l'impianto in modo sostanziale.

Ne è emerso un testo ben diverso dalla proposta originale, con un *corpus* di regole ritagliato sui c.d. *General Purpose AI Systems (GPAI)*, o sistemi di IA con finalità generali, per i quali si prevedeva una regolazione asimmetrica, simile a quella predisposta dal Digital Services Act, con obblighi più stringenti a carico dei fornitori di GPAI particolarmente potenti, dunque suscettibili di generare rischi sistemici. Tali regole si accompagnavano a un sistema di governance completamente nuovo e addizionale rispetto a quello previsto per le applicazioni ad alto rischio, e assoggettato alla competenza della Commissione europea, nella quale veniva creato uno nuovo servizio, l'AI Office. Si trattava di un *upgrade* per nulla scontato, frutto di una crescente presa di coscienza, quella dell'incompiutezza dell'impianto originale. Peraltro, il testo definitivo vide la luce nel silenzio del trilogo, quindi nella totale assenza di consultazione pubblica o di contributi da parte degli esperti. Il senso di urgenza che accompagnava la riforma radicale di un testo ancora non entrato in vigore ben rappresenta, in modo vivido, l'*impasse* del regolatore europeo, in preda a una materia fluida e quasi imprevedibile nella sua fenomenologia.

La necessità di accelerare l'iniziativa del regolatore comunitario si materializzò anche con un rinnovato attivismo sul fronte della *soft law*, con la Commissione intenta a perseguire un accordo di massima che potesse anticipare l'entrata in vigore delle regole essenziali dell'AI Act. Ne seguono l'AI "Pact", sottoscritto da dozzine di aziende come impegno preliminare allo sviluppo responsabile dell'IA; e il sofferto Codice di Condotta relativo ai GPAI, partorito dopo una lunga e travagliata gestazione da un

gruppo di esperti che, superato il migliaio di partecipanti, si è infine affidato alle (competenti) penne di una dozzina di esperti. In seguito, un profluvio di linee guida e atti delegati emessi dalla Commissione su vari aspetti dell'AI Act ha contribuito a chiarire, almeno in parte, quali siano gli obblighi di figure come i fornitori di GPAI e quelli di applicazioni *high risk*, come definire e individuare tali applicazioni e molti altri aspetti generali di non facile interpretazione nel testo del regolamento.

Tali sforzi sono divenuti ancor più eroici nel corso del 2024 e del 2025, per via di un secondo ordine di fattori che ha inciso in modo non irrilevante sulla complessità del progetto di regolamento europeo. Da un lato, la progressiva presa di coscienza della scarsa competitività europea nel campo dell'IA, a fronte di progressioni a dir poco vertiginose negli Stati Uniti e in Cina, ha convinto la Commissione europea – specialmente all'inizio del secondo mandato di Ursula von der Leyen – a perseguire un percorso di semplificazione normativa, a volte sfociata in tentativi di deregolamentazione, per rispondere all'accorato appello del "rapporto Draghi". Dei due ecosistemi originariamente contemplati dal Libro Bianco del 2020, quello di eccellenza appariva deficitario, mentre quello della fiducia sembrava ancora lontano dal suo pieno compimento. Di qui la decisione di abbandonare il progetto di direttiva sulla responsabilità civile, di cui si è già detto, e di introdurre una proposta di semplificazione – il c.d. *Digital Omnibus* – che modifica ulteriormente la portata e l'applicazione dell'AI Act, ancora una volta senza affidarsi a una vera e propria analisi di impatto normativo.

Dall'altro lato, a questa rinnovata vocazione alla deregolamentazione – assolutamente autoctona – si accompagnava la pressione internazionale esercitata dall'amministrazione statunitense, decisa a non supportare alcun tentativo di cooperazione internazionale in tema di *regulation* in questo campo, e pronta a fare della "semplificazione" della regolazione europea elemento negoziale a fronte di una politica inusitatamente aggressiva in materia di tariffe commerciali. Normative come l'AI Act e il DSA sono finite, così, sul banco degli imputati in un allineamento di intenti che ha finito col denervare in modo significativo sia la spinta a completare l'*opus magnum* dell'AI Act, sia processi di cooperazione internazionale in materia di AI, con la

sola eccezione del poco incisivo processo di Hiroshima ancora attivo, quasi senza forza, nell'ambito del G-7.

Da ultimo, è utile ricordare che il processo di applicazione dell'AI Act dipende in modo significativo dal completamento di un percorso parallelo, quello della standardizzazione, oggetto di un mandato specifico che la Commissione ha rivolto al CEN-CENELEC. Un processo che sin qui si è rivelato alquanto macchinoso e costellato di conflitti di interesse, se si pensa che imprese private sono chiamate a definire standard che una volta validati si applicheranno a loro stesse; e che più o meno la metà degli stati membri si è espressa, in seno al Consiglio Ue, per una significativa ridefinizione, al ribasso della portata dell'AI Act. Posta di fronte al problema, nel *Digital Omnibus* la Commissione europea ha finito col posporre ufficialmente l'applicazione delle regole relative alle applicazioni ad alto rischio di un anno, all'agosto 2027. E anche in quel caso, si è trovata a dover annunciare che, in assenza di standard sufficientemente dettagliati prodotti dal CEN-CENELEC, dovrà essere costretta a introdurre standard provvisori, la gestazione dei quali comporterebbe non poche difficoltà.

1.2 Un'opera destinata a rimanere incompiuta?

Di fronte a tante difficoltà, la traiettoria dell'AI Act appare oggi aver perso il suo slancio iniziale, per vari motivi che qui – senza pretesa di essere esaustivi – provo a riassumere. Primo, nonostante lo sforzo iniziale della Commissione, il testo manca di alcuni elementi fondamentali, che renderebbero l'AI Act più flessibile e in grado di reggere l'urto dell'incedere e del continuo divenire dell'IA. Se da un lato la definizione di IA è neutrale dal punto di vista tecnologico, le liste delle applicazioni inizialmente inserite nelle quattro categorie di rischio sono convenientemente poste in allegati al testo, e le *regulatory sandbox* (su cui si veda il capitolo 6 di questo Rapporto) promettono un'apertura all'innovazione specialmente per le PMI, dall'altro l'impianto regolatorio mancava di una vera e propria *governance* adattiva, nella forma di un'agenzia europea supportata da un gruppo di esperti in grado di guidarne l'azione. Sarebbe stato, da questo punto di vista, preferibile evitare di adottare un testo che supera le 250 pagine,

preferendo un intervento più agile, concentrato sui principi e gli obiettivi della misura regolatoria, e più concentrato sul demandare a un organismo di *governance* indipendente il costante aggiornamento dello stato dell'arte in materia di AI. Ciò avrebbe permesso un processo di approvazione ed entrata in vigore più snello, nonché un quadro regolatorio più rapido, flessibile e partecipativo. Inoltre, avrebbe permesso all'autorità indipendente europea di iniziare a cooperare da subito con i regolatori di settore, che – come vedremo – svolgono un ruolo essenziale nel rendere il regolamento efficace.

Secondo, nell'oscillare tra regolazione basata sui principi e normativa prescrittiva, l'AI Act finisce con l'atterrare a metà strada, troppo dettagliato per essere agile, e troppo poco prescrittivo per essere applicabile da subito. La causa di questa indeterminatezza sta nella natura stessa del problema – al tempo stesso transeunte e pervasivo – che costringe il regolatore europeo a metter mano costantemente a revisioni normative per correggerle ancor prima che abbiano effetto. Basta leggere il testo per rendersene conto: i requisiti regolatori associati alle applicazioni ad alto rischio sono necessariamente generici e poco applicabili dalle entità regolate. La ragione è semplice: non solo la lista delle applicazioni ad alto rischio comprende casi d'uso chiaramente riferiti a problemi di sicurezza e a rischio per i diritti fondamentali; in più, ciascun settore e *use case* presenta rischi di natura ed entità diversa, nonché misure di mitigazione del rischio molto diversi. Riferirsi, ad esempio, ad un'adeguata supervisione umana non basta a spiegare cosa il requisito implichi per un sistema di *triage* di un ospedale, o l'utilizzo di un sistema di AI in contesto giudiziario, o in un veicolo autonomo. Lo stesso può dirsi per ciascuno dei requisiti, che devono essere tradotti in indicazioni più concrete per ciascuna applicazione futura.

Di qui, come terzo fattore, emerge la necessità di predisporre un'adeguata "interfaccia" tra il regolatore europeo e le autorità competenti per materie specifiche, come i garanti per la protezione dei dati personali; e ancor più le autorità di regolazione di settore, sia che si tratti di enti governativi o autorità indipendenti. Queste istituzioni, spesso dimenticate nel dibattito europeo e internazionale, sono chiamate a riconciliare la normativa per le quali sono competenti con il disposto dell'AI Act, e più in generale con

la diffusione dell'IA nei prodotti e servizi per la supervisione dei quali tali autorità sono competenti. Grandi dimenticate della prima stagione dell'AI Act, i garanti e le autorità di settore potrebbero ora tornare alla ribalta come protagonisti di una seconda stagione, nella quale la rincorsa a perdifiato dietro il costante divenire dell'IA viene sostituita da un più ragionato approccio alle riforme regolatorie che si rendono necessarie per assicurare che la normativa specifica rimanga *fit for purpose* nell'era dell'IA più avanzata. In questo frangente, troveranno interlocutori diversi e non necessariamente allineati, se si pensa che per alcuni tipi di applicazioni essere dovranno far riferimento al contesto nazionale; e che però, soprattutto se la proposta di *Digital Omnibus* dovesse essere tradotta in legge, si vedranno affiancare dalla Commissione europea e dell'AI Office, che con l'omnibus annunciano di voler offrire più certezza del diritto riguardo alla sovrapposizione tra AI Act e altre norme europee, e propongono di centralizzare nell'AI Office la supervisione si sistemi di AI sviluppati a partire da GPAI, o utilizzati da grandi piattaforme o motori di ricerca (come definiti dal DSA).

Si tratta, come si è detto, di una proposta soggetta al vaglio di Parlamento e Consiglio. Se dovesse "passare", significherebbe un altro, repentino cambiamento per il sistema di *enforcement* dell'AI Act, nonché un cambio di scenario per le autorità di supervisione nazionali. Si arriverebbe quasi a una "normalizzazione" dell'AI Act, se si pensa che quasi tutti i paesi che hanno deciso di porre mano a una politica pubblica in tema di AI lo hanno fatto adottando un quadro normativo orizzontale molto esile, abbinato a un più marcato sforzo di regolazione settoriale.

1.3 Le autorità indipendenti nell'era dell'intelligenza artificiale: motivazione e struttura del lavoro

Di fronte a uno scenario tanto incerto quanto cangiante, le autorità di supervisione devono porsi alcune domande fondamentali. Quale AI Act si troveranno a dover implementare? Quali saranno gli attori principali, a livello europeo e nazionale, e con quali competenze? Come assicurarsi che la normativa di settore rifletta appieno, e in modo dinamico, l'evoluzione dell'AI? Quanto fare affidamento su autorità nazionali con

competenza generale, se esistenti (in molti paesi, non esiste un'autorità centrale preposta all'IA)? E ancora: quali strumenti di regolazione utilizzare, e quali standard socio-tecnici adottare a riferimento?

È in questo contesto che l'AGCOM ha deciso di creare un Comitato indipendente di esperti di IA, con forte competenza giuridica, che si occupa di individuare e analizzare le questioni giuridiche che AGCOM dovrà affrontare nel corso dei prossimi mesi e anni, nonché di individuare le opportunità che l'AI può offrire sia nei contesti di mercato di competenza di AGCOM, sia come ausilio al processo di regolazione stessa.

Questo Rapporto costituisce il primo contributo del Comitato, e si compone di nove capitoli. Nel primo contributo, Giovanna de Minico analizza l'interrelazione tra AI Act e DSA, evidenzia il possibile impatto delle due normative sui poteri dell'AGCOM e – nella seconda parte – approfondisce l'impatto del regolamento europeo sulla pubblicità politica. Nel Capitolo 3, Andrea Imperiali di Francavilla presenta un lavoro a dir poco monumentale di mappatura dell'impatto dell'intelligenza artificiale in tutti i settori regolati da AGCOM. Nel Capitolo 4, Andrea Simoncini analizza le implicazioni costituzionali dell'intelligenza artificiale per la democrazia e la libertà di informazione, proponendo un modello di costituzionalismo digitale europeo volto a garantire il rispetto dei diritti fondamentali e a riequilibrare il rapporto tra potere privato e sfera pubblica. Nel capitolo 5 Giovanni Boccia Artieri fa il punto su un tema oggi alquanto attuale e spinoso, quello della IA generativa e del suo impatto su disinformazione e *hate speech*. Il capitolo 6, a cura di Giuseppe Cassano, approfondisce il tema – divenuto cruciale soprattutto nell'era dell'IA generativa – della tutela del diritto d'autore. Nel Capitolo 7, Mauro Giusto offre una visione più sistemica, presentando le sue riflessioni in tema di IA ed educazione dei giovani. Il capitolo 8, anch'esso a firma di Giovanna de Minico, propone una più ampia riflessione di carattere costituzionale sul rapporto tra tecnologia e diritto, analizzando le condizioni nelle quali la tecnologia rischia di evolvere da strumento disciplinato dal diritto a fonte autonoma di potere normativo e individuando le garanzie necessarie a preservare il primato del processo decisionale democratico. Nel capitolo conclusivo, mi occupo di tirare le fila dei contributi precedenti, abbinandoli a una riflessione sugli attuali sviluppi dell'agenda comunitaria,

sulla prossima evoluzione dell'IA (a partire dall'*agentic AI* e dell'incertezza che sta generando anche per l'applicazione dell'AI Act), e sugli strumenti di regolazione e governance che le autorità indipendenti dovranno adottare nel corso dei prossimi anni, per equipaggiarsi a dovere in quello che si presenta come un cammino impervio a fronte di una materia regolata sempre più spinosa.

2 La regulation dell'AGCOM intorno all'ecosistema digitale

Giovanna de Minico

Sommario: Parte A): L'incontro tra DSA e *AI Act*: le competenze dell'AGCOM. – 1. Le domande del DSA e dell'*AI Act*. – 2. Quale rapporto corre tra i poteri dell'AGCOM. secondo il DSA e la struttura dell'*AI Act*? – 3. Una possibile chiave di lettura in risposta alla seconda domanda. – 4. Casistica futura.
Parte B): Il Regolamento europeo sulla pubblicità politica e la riserva di competenze all'AGCOM. – 1. Luci e ombre del Regolamento europeo sulla pubblicità politica. – 2. L'AGCOM. nel dialogo con le altre Autorità. – 3. Tabella di sintesi.

2.1 L'incontro tra DSA e l'AI Act: le competenze dell'AGCOM

2.1.1 Le domande del DSA e dell'AI Act

Richiamo l'attenzione del lettore su alcune *issue* che l'esame attento del quadro regolatorio, europeo e interno, ha posto, ma al tempo stesso ha lasciato aperte, affidandole così alla definizione futura dell'interprete che verrà.

1) Secondo il richiamato dato di diritto, l'AGCOM è titolare di competenze specifiche in materia di intelligenza artificiale (IA)?
2) Se a questa domanda rispondesse positivamente, quale sarebbe la natura di questa competenza e quali le sue materie di intervento?

Per rispondere agli interrogativi sollevati occorre partire dalla lettura combinata del *Digital Service Act* – (DSA)[2] e dell'*Artificial Intelligence Act* (AI Act)[3] con la legge nazionale del settembre 2025 sull'intelligenza artificiale[4]. Si tratta di tre fonti che vanno quindi interpretate l'una con le altre, senza farsi attrarre dalla tentazione di letture più semplici perché atomistiche, escluse peraltro dalla natura sistemica dei due livelli ordinamentali, interno ed europeo, che integrandosi a vicenda impongo che si tenga conto dell'uno alla luce dell'altro e viceversa (Betti).

2.1.2 Quale rapporto corre tra i poteri dell'AGCOM secondo il DSA e la struttura dell'AI Act?

Veniamo alla prima domanda.

Sappiamo che l'AGCOM. è stata indicata *Digital Services Coordinator* (DSC) nazionale (art. 49 DSA): pertanto, è responsabile del coordinamento, della vigilanza e dell'applicazione del DSA su tutto il territorio nazionale, nei confronti degli *hosting provider*, piattaforme *online*, motori di ricerca *online*[5] sottosoglia comunitaria.

Ai sensi del DSA, l'AGCOM. dispone, fra l'altro, del potere di richiedere informazioni, di inibire condotte *unlawful* (tramite ordini) ed emettere sanzioni (per esempio in caso di inottemperanza agli ordini) nei confronti dei fornitori di servizi digitali che operano

*Professoressa ordinaria di Diritto costituzionale e pubblico – Dipartimento di Giurisprudenza - Università degli Studi di Napoli Federico II.

[2] Regolamento (UE) 2022/2065 Del Parlamento Europeo e del Consiglio del 19 ottobre 2022 relativo a un mercato unico dei servizi digitali e che modifica la direttiva 2000/31/CE (regolamento sui servizi digitali)

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

[4] Legge 23 settembre 2025, n. 132, Disposizioni e deleghe al Governo in materia di intelligenza artificiale, in GU n. 223 del 25 settembre 2025.

[5] A tal fine è stata specificamente modificata la legge istitutiva dell'Autorità con l'inserimento all'art. 1, comma 6, lett. c), della legge n. 249/1997 del numero 14-ter), in base al quale l'Autorità *«esercita la funzione di Coordinatore dei Servizi Digitali e i relativi poteri previsti dal Regolamento (UE) 2022/2065 del Parlamento europeo e del Consiglio del 19 ottobre 2022 relativo a un mercato unico dei servizi digitali»*. (Art. 15 c.3, D.L. 123/2023, cd. "decreto Caivano").

sul territorio nazionale, qualora l'Autorità rilevi una "propagazione" di contenuti nocivi per la massa indefinita degli utenti[6].

L'AGCOM. è quindi l'ANR competente *ex lege* a adottare le azioni finalizzate a impedire la propagazione di «contenuti illeciti». Si noti che il DSA non fornisce una definizione, né riporta un *numerus clausus* per ciò che concerne i «contenuti illegali». L'art. 3, co. 1, lett. h) definisce «contenuto illegale» «qualsiasi informazione che, di per sé o in relazione a un'attività, tra cui la vendita di prodotti o la prestazione di servizi, non è conforme al diritto dell'Unione o di qualunque Stato membro conforme con il diritto dell'Unione, indipendentemente dalla natura o dall'oggetto specifico di tale diritto». Il Regolamento reca, quindi, un rinvio mobile al parametro normativo esterno (unionale o interno), lasciando alla discrezionalità del Regolatore l'individuazione della norma che si assume di volta in volta lesa e che conferisce l'attributo della illegalità al contenuto, nel senso della sua contrarietà al comando[7].

Inoltre, l'AGCOM (nei confronti delle piattaforme nazionali) e la Commissione (per i VLOP/VLOSE) sono i soggetti individuati dall'ordinamento unionale e interno per contrastare la diffusione dei contenuti illeciti, poiché la semplice circolazione e la

[6] Gli artt. 9 e 10 del DSA concernono, rispettivamente, l'ordine di rimozione dei contenuti illegali, e quello di fornire informazioni. Inoltre, ai sensi dell'art. 52, ove necessario per lo svolgimento dei loro compiti ai sensi del regolamento, i coordinatori dei servizi digitali dispongono dei seguenti poteri di *enforcement*, vale a dire di poteri "coercitivi" in senso lato (*i.e.:* di incidere/modificare la sfera giuridica del soggetto interessato, da esercitarsi nel rispetto delle indefettibili garanzie del contraddittorio) nei confronti dei fornitori di servizi intermediari che ricadono nella competenza del loro Stato membro. Per l'articola elencazione si rinvia al dato positivo

[7] Cfr. il Considerando 12: «Per conseguire l'obiettivo di garantire un ambiente online sicuro, prevedibile e affidabile, ai fini del presente regolamento il concetto di "contenuto illegale" dovrebbe rispecchiare ampiamente le norme vigenti nell'ambiente offline. In particolare, il concetto di "contenuto illegale" dovrebbe essere definito in senso lato per coprire anche le informazioni riguardanti i contenuti, i prodotti, i servizi e le attività illegali. Tale concetto dovrebbe, in particolare, intendersi riferito alle informazioni, indipendentemente dalla loro forma, che ai sensi del diritto applicabile sono di per sé illegali, quali l'illecito incitamento all'odio o i contenuti terroristici illegali e i contenuti discriminatori illegali, o che le norme applicabili rendono illegali in considerazione del fatto che riguardano attività illegali. Tra queste figurano, a titolo illustrativo, la condivisione di immagini che ritraggono abusi sessuali su minori, la condivisione non consensuale illegale di immagini private, il cyberstalking (pedinamento informatico), la vendita di prodotti non conformi o contraffatti, la vendita di prodotti o la prestazione di servizi in violazione della normativa sulla tutela dei consumatori, l'utilizzo non autorizzato di materiale protetto dal diritto d'autore, l'offerta illegale di servizi ricettivi o la vendita illegale di animali vivi. Per contro, un video di un testimone oculare di un potenziale reato non dovrebbe essere considerato un contenuto illegale per il solo motivo di mostrare un atto illecito quando la registrazione o la diffusione di tale video al pubblico non è illegale ai sensi del diritto nazionale o dell'Unione. A tale riguardo è irrilevante che l'illegalità delle informazioni o delle attività sia sancita dal diritto dell'Unione o dal diritto nazionale conforme al diritto dell'Unione e quale sia la natura esatta o l'oggetto preciso della legge in questione».

conseguente esposizione dell'utente è considerata un pericolo *ex se* –un evento temuto – anche qualora la sua diffusione non abbia prodotto un concreto ed effettivo effetto lesivo per l'utente (illecito da pericolo)[8].

Da qui l'assegnazione all'AGCOM. delle funzioni di ordine e di sanzione, nonché di limitati poteri regolamentari – prevalentemente di carattere organizzativo[9] – e, infine di coordinamento con le altre ANR (nazionali ed europee) e con la stessa Commissione europea.

Per quanto invece concerne invece il Regolamento europeo sull'IA, unitamente alla legge nazionale di recepimento, la sua interpretazione letterale porta a escludere l'assegnazione di poteri espressi all'Autorità in oggetto in materia di IA.

Infatti, il Regolamento IA prevede un approccio precauzionale volto a fronteggiare i cd. rischi sistemici derivanti dall'impiego dell'IA. La legge nazionale individua due Autorità specifiche, l'Agenzia per l'Italia Digitale (AgID) e l'Agenzia per la cybersicurezza nazionale (ACN)[10], senza conferire competenze specifiche all'AGCOM..

Ma è proprio la legge nazionale a fornire un elemento utile al nostro ragionamento. Infatti, una competenza per materia si potrebbe ricavare, con un'interpretazione deduttiva, lì dove la legge nazionale modifica la legge sul diritto d'autore, e specifica che sono protetti dal diritto d'autore anche contenuti prodotti mediante IA (Art. 25, comma 1: «Alla legge 22 aprile 1941, n. 633, sono apportate le seguenti modificazioni: a) all'articolo 1, primo comma, dopo le parole: "opere dell'ingegno" è inserita la seguente: "umano" e dopo le parole: "forma di espressione" sono aggiunte le seguenti:

[8] Mi sia consentito il richiamo a un mio lavoro, *Nuova tecnica per nuove diseguaglianze. Case law: Disciplina Telecomunicazioni, Digital Services Act e Neurodiritti*, in *Federalismi.it*, 6/2024, p. 17.

[9] Come nei fatti è avvenuto: cfr. il Regolamento in materia di certificazione degli organismi ADR (art. 21 DSA), di cui alla delibera n. 282/24/CONS; il Regolamento per la nomina dei «segnalatori attendibili» (art. 22 DSA), cioè gli enti qualificati per segnalare contenuti illegali alle piattaforme con priorità di trattamento (delibera n. 283/24/CONS); il Regolamento di procedura per la gestione dei reclami (art. 53 DSA), di cui alla delibera n. 25/25/CONS.

[10] L'AgID promuove e abilita lo sviluppo dell'IA, gestendo le attività di notifica, accreditamento e monitoraggio; l'ACN, di contro, vigila sulla sicurezza dei sistemi di IA e dispone di poteri di controllo e sanzionatori, promuovendo al contempo i profili di *cybersecurity* legati all'intelligenza artificiale.

", anche laddove create con l'ausilio di strumenti di intelligenza artificiale, purché costituenti risultato del lavoro intellettuale dell'autore»).

Ecco che dunque l'AGCOM., competente a tutelare in via diretta il diritto d'autore in rete, potrebbe rivendicare altresì una competenza indiretta nella tutela dei prodotti creati con l'ausilio dell'IA.

2.1.3 Una possibile chiave di lettura in risposta alla seconda domanda

L'AGCOM. non nasce come "Autorità dell'IA", ma va subito detto che la sua estraneità al perimetro del regolamento europeo deve fare i conti con la realtà dei *Large language models* (LLM). Questa innovativa forma di IA, tanto atipica da essere inizialmente ignorata dal Reg. eu., ha ricevuto un regime giudico in deroga a quello riservato alla fattispecie I.A. tradizionale, data la sua natura speciale rispetto al modello generale di base. Ebbene, i LLM entrano nel raggio d'azione dell'Autorità, pur nel silenzio assordante del Reg. e della legge italiana, quando il loro impiego finisce per intercettare competenze tipiche ed espresse dell'Autorità, perché comportano diffusione di contenuti *online*; servizi erogati dalla piattaforma, informazione, pubblicità, la tutela degli utenti (minori, pluralismo, disinformazione).

Possiamo escludere che l'Autorità abbia poteri sull'IA come oggetto *ex se*[11], questo rimane fuori della sua sfera d'azione, ma ciò non esclude che l'Autorità possa intervenire sugli effetti che derivano dal suo utilizzo, nel caso in cui i LLM contribuiscano a quella propagazione di contenuti «illegali», che, come tali, ricadono nella fattispecie aperta dell'art. 3 del DSA. Qualora, dunque, un LLM sia integrato in *social network*, in un motore di ricerca, in un *chatbot* informativo o in sistemi di raccomandazione, l'AGCOM. si deve considerare chiamata in causa sempre che rinvenga fattispecie di sua competenza, cioè disegnate nel DSA e nelle leggi nazionali, ad es. il TUSMA. Si pensi, a titolo esemplificativo, ai seguenti ambiti: la disinformazione; l'*hate speech*; la violazione della tutela dei minori; la mancata trasparenza algoritmica o la violazione del pluralismo dell'informazione.

[11] Ci sia consentito il rinvio a: *Too Many Rules or Zero Rules for the CHATGPT*, in G. De Minico, *Libertà Virtuali. Costituzione e Mercato*, Merita Edizioni, Torino, 2024, p. 246.

2.1.4 Casistica futura

Esporremo ancora un caso che può aiutare a comprendere l'interpretazione sistematica da noi proposta: nel contesto dei servizi digitali l'IA è usata per moderare contenuti (rimozione automatica di post o video), personalizzare pubblicità, rilevare comportamenti sospetti o illegali (es. deepfake, hate speech), gestire chatbot e assistenti virtuali[12]. Questo significa che un fornitore si avvale di un LLM per generare o amplificare contenuti illegali, AGCOM. può chiedere informazioni, valutare i rischi, imporre misure correttive, e pronunciare sanzionare per mancata trasparenza o moderazione inadeguata.

Nel comparto della pubblicità e comunicazioni commerciali, l'IA può essere impiegata per forme di *influencer marketing* automatizzato, e su questo terreno si possono aprire spazi di azione per l'AGCOM. lì dove si stiano realizzando forme di pubblicità occulta.

Quindi, l'Autorità non agisce in via immediata sull'IA, perché nessuna fonte le assegna questo compito, né tanto meno dispone di poteri regolamentari. Qui dobbiamo applicare in modo rigido il criterio di legalità formale, che ci porta a escludere una sua competenza perché l'AGCOM al pari di ogni altra pubblica autorità deve rispettare il principio di legalità, ciò significa che deve poter esibire in ogni momento un titolo ah hoc che giustifichi i poteri che intende esercitare. La legalità per il potere pubblico si esprime in positivo come esibizione della previa norma di legge, che in quanto norma di azione conferisce l'attribuzione che l'A. rivendica come propria, a differenza del settore governato dal diritto privato, dove vige la regola per cui è lecito per la persona, tutto ciò che il diritto non proibisce. Detto ciò, c'è un 'ma', perché un'analisi attenta rivela che l'AGCOM. si possa avvalere dei suoi poteri di *enforcement* - vigilanza e controllo, ordine, sanzione - quando l'IA diventa il mezzo, cioè la tecnologia di cui si servono le piattaforme per offrire servizi e contenuti illegali agli utenti finali. In questo contesto l'Autorità si avvarrà delle sole sue funzioni tipizzate, le quali nell'incidere sui contenuti ospitati dalla piattaforma, intercetterà anche in via mediata il prodotto

[12] ID., *Unione europea, mercato, tecnica*, Relazione al *Convegno annuale dell'Associazione Italiana dei Costituzionalisti "L'Unione europea a confronto con la Costituzione della Repubblica italiana" Università degli Studi di Torino,* 11/10/2025.

creato dal LLM, pur nel risetto sostanziale del principio di legalità formale. Qui, non siamo ricorsi alla teoria dei poteri impliciti – la cui compatibilità costituzionale con il principio di legalità formale corre sul filo del rasoio - perché non abbiamo desunto da un potere espresso e maggiore uno inespresso e minore, ma ci siamo limitati a dilatare il raggio del compasso che ha in mano l'Autorità sull'intero materiale ospitato dalla piattaforma, sia esso generato dagli utenti che lo postano o da una mente mecanica che lo genera in autonomia.

Veniamo dunque rapidamente alle conclusioni, ponendo l'accento su due ulteriori profili.

Appare chiaro che l'esercizio di tali funzioni potrebbe determinare delle sovrapposizioni di competenze, il che comporta, di conseguenza, la necessità per le varie ANR coinvolte di coordinarsi, secondo quanto il metodo europeo della leale collaborazione indica di fare.

Ad esempio, se tramite un canale che trasmette contenuti generati con l'IA vengono trasmessi contenuti nocivi per i minori, l'AGCOM. è competente a intervenire ai sensi del DSA (e anche della fonte nazionale, ad esempio del TUSMA e del Decreto Caivano); se poi quegli stessi contenuti violano anche la *privacy* del minore (per esempio mostrando il volto di un minorenne), sarà necessario anche l'intervento del Garante per la protezione dei dati personali. Se poi nell'ambito della medesima condotta si rinvengono anche potenziali rischi per la sicurezza complessiva della Rete, entreranno in gioco le attribuzioni anche di ACN. Lo strumento del coordinamento tra ANR tramite la stipula di Protocolli di collaborazione rappresenta dunque un passo indispensabile per il buon andamento dell'azione amministrativa, come alcuni di noi suggeriscono già da tempo.

Infine, tale potere di *enforcement* e la conseguente funzione sanzionatoria si intersecano con il grande tema della responsabilità delle piattaforme. La direttiva *e-commerce*, non modificata su questo punto dal DSA, prevede un regime di responsabilità condizionata per la piattaforma, se inconsapevole dei contenuti nocivi trasmessi, oppure se appena saputo si attiva per rimuoverli, la piattaforma è ritenuta irresponsabile. Tuttavia, il DSA fa salvo l'obbligo di diligenza a carico della piattaforma

per evitare il danno: quest'ultima, a sua volta, dovrà dotarsi di un sistema organizzativo interno idoneo a impedire la diffusione incontrollata di contenuti illeciti, anche tramite sistemi di IA (cfr. Cons. Stato, n. 4277/2024). Qualora la piattaforma ometta di agire in tal senso, o non dimostri di aver adottato tali sistemi, in sede di contenzioso non potrà ritenersi assolto l'onere della prova contraria.

Salutiamo con favore la giurisprudenza amministrativa recente che – per contrastare l'auto-esonero da responsabilità dei Giganti della rete basato sul mito della neutralità dell'*hosting provider* – ha elevato i sistemi di IA gestionali (diretti a prevenire il danno) a elemento organizzativo essenziale per dimostrare l'assolvimento dell'onere di diligenza da parte della piattaforma.

In questo contesto l'IA compie un salto di qualità: la tecnologia non è più soltanto il mezzo diffusivo di potenziali contenuti dannosi, ma diviene lo strumento privilegiato per arginare il danno temuto.

2.2 Il Regolamento europeo sulla pubblicità politica e le competenze dell'AGCOM

2.2.1 Luci e ombre del Regolamento europeo sulla pubblicità politica

Questo Regolamento nasce in risposta della *self-regulation* finora prevalsa per iniziativa dei motori di ricerca, *social* e piattaforme[13]. Il Regolamento segue la filosofia della *disclosure*, anche se non totale, poiché fondamentalmente impone un obbligo di trasparenza limitato a talune informazioni in capo allo sponsor della pubblicità politica Tale obbligo prevede l'apposizione di un'etichetta volta a chiarire che si tratti di messaggio pubblicitario e chi sia lo *sponsor*, cioè chi se ne avvantaggia. Ma nulla è detto su chi finanzia la campagna – punto su cui cade il silenzio – e manca, soprattutto, di imporre il divieto di addizionare spazi pubblicitari oltre un certo limite.

[13] Ci sia permesso il richiamo di quanto già consegnato in: *Le fonti del diritto: un argine all'intelligenza artificiale?* in *Rivista AIC*, 3/2025,78 ss.

In sintesi, la Commissione si è accontentata di rendere il cittadino edotto del fatto che sta per essere bombardato di messaggi di pubblicità politica a favore di chi può permetterselo, senza però tutelare il suo diritto a ottenere un'informazione pluralistica, cioè contribuita dalle tante famiglie politiche che articolano il pluralismo ideologico della società.

La domanda che ci poniamo, dunque, è se possa definirsi politicamente utile un regolamento sulla pubblicità che manca dei due elementi essenziali prima evidenziati.

A queste due gravi omissioni si aggiunge una questione di fondo: lo sfruttamento di tecniche emozionali per orientare l'opinione pubblica in una certa direzione[14]. Ebbene, anche questa *issue* non è stata affrontata, così come non lo è neppure nell'AI Act agli artt. 5 e 50.

2.2.2 L'AGCOM nel dialogo con le altre Autorità

Ai fini della presente analisi, per individuare se ci sia uno spazio di intervento per l'AGCOM. e quanto ampio esso possa essere è utile notare che il Garante privacy nazionale e quello europeo della protezione dei dati sono competenti per gli artt. 18 e 19 a trattare i dati per finalità di *targeting*.

Gli Stati membri designano, invece, le autorità deputate a controllare l'osservanza degli obblighi sulla pubblicità politica, imposti i prestatori di servizi intermediari (come definiti dal DSA). Tale autorità può coincidere con quella designata a norma del DSA; in Italia, ciò accaduto assegnando questo compito all'AGCOM.

Il Coordinatore dei Servizi Digitali (DSC), di cui all'articolo 49 del DSA, è infatti competente per il coordinamento a livello nazionale tra i due regolamenti (pubblicità politica e DSA). L'osservanza delle norme dagli articoli 7 a 21 è rimessa alla vigilanza delle stesse autorità designate dal DSA (AGCom), con la precisazione che il regolamento sui servizi digitali limita il governo delle autorità nazionali sui soli soggetti diversi dalle VLOP o dai VLOSE.

[14] Cfr. ID., *Nuova tecnica per nuove diseguaglianze*, cit., pp.19-21.

Di conseguenza, in virtù del rinvio operato dall'art. 22 al quadro del DSA, l'autorità competente per la pubblicità *online* è la Commissione europea qualora si tratti di una VLOP; laddove si tratti, invece, di realtà minori (come il "blog di quartiere"), la competenza resta in capo all'autorità nazionale. In caso di violazioni relative al trattamento dei dati, l'autorità di riferimento è il Garante privacy nazionale o europeo.

2.3 Tabella di sintesi

Poteri dell'AGCOM ai sensi del Regolamento (UE) 2024/900	
Regolamento (UE) 2024/900	**Commento**
Art. 22 ***Autorità competenti e punti di contatto*** 1. Le autorità di controllo di cui all'articolo 51 del regolamento (UE) 2016/679 o il Garante europeo della protezione dei dati di cui all'articolo 52 del regolamento (UE) 2018/1725 sono competenti a monitorare l'applicazione degli articoli 18 e 19 del presente regolamento nel settore di loro competenza. Si applicano, mutatis mutandis, l'articolo 58 del regolamento (UE) 2016/679 e l'articolo 58 del regolamento (UE) 2018/1725. Il capo VII del regolamento (UE) 2016/679 si applica alle attività contemplate agli articoli 18 e 19 del presente regolamento. 2. Il comitato europeo per la protezione dei dati di cui all'articolo 68 del regolamento (UE) 2016/679 elabora orientamenti, di propria iniziativa o su richiesta della Commissione, al fine di assistere le autorità di controllo di cui al regolamento (UE) 2016/679 nel valutare la conformità ai requisiti del presente regolamento.	1. L'AGCOM, quale coordinatore dei servizi digitali a norma del DSA, potrebbe essere designata a livello nazionale come autorità competente per controllare l'osservanza da parte dei prestatori di servizi intermediari *online* degli articoli da 7 a 17 e dell'art. 21, cioè degli obblighi di trasparenza e di quelli relativi al dovere di diligenza per i servizi di pubblicità politica, nonché delle disposizioni relative al rappresentante legale, ma al momento non possiamo dirlo perché attendiamo l'atto nazionale di designazione.

	2. Resta ferma la competenza del Garante privacy sul *targeting*, ex art. 22, par.
3. Gli Stati membri designano le autorità competenti a controllare l'osservanza, da parte dei prestatori di servizi intermediari ai sensi del regolamento (UE) 2022/2065, degli obblighi di cui agli articoli da 7 a 17 e all'articolo 21 del presente regolamento, ove applicabile. Le autorità competenti designate a norma del regolamento (UE) 2022/2065 possono essere anche le autorità competenti designate per controllare l'osservanza, da parte degli intermediari online, degli obblighi di cui agli articoli da 7 a 17 e all'articolo 21 del presente regolamento. Il coordinatore dei servizi digitali di cui all'articolo 49 del regolamento (UE) 2022/2065 di ogni Stato membro è competente per il coordinamento a livello nazionale nei confronti dei prestatori di «servizi intermediari» quali definiti dal regolamento (UE) 2022/2065. Alle materie connesse con l'applicazione del presente regolamento nei confronti dei prestatori di servizi intermediari si applicano l'articolo 49, l'articolo 58, paragrafi da 1 a 4, e l'articolo 60, paragrafo 1, del regolamento (UE) 2022/2065. L'articolo 51 del regolamento (UE) 2022/2065 si applica mutatis mutandis per quanto riguarda i poteri delle autorità competenti designate a norma del presente paragrafo.	3. Nelle materie residuali può essere individuata una terza autorità, anche diversa dalle precedenti.

4. Ogni Stato membro designa una o più autorità competenti incaricate dell'applicazione e dell'esecuzione degli aspetti del presente regolamento non contemplati ai paragrafi 1 e 3 del presente articolo. Tali autorità competenti possono essere diverse da quelle di cui ai paragrafi 1 e 3 del presente articolo e possono essere le stesse di quelle contemplate all'articolo 30 della direttiva 2010/13/UE. Ogni autorità competente designata ai sensi del presente paragrafo gode strutturalmente di piena indipendenza sia dal settore che da qualsiasi intervento esterno o pressione politica. Agendo in piena indipendenza monitora efficacemente e prende le misure necessarie e proporzionate per garantire la vigilanza, il rispetto e l'applicazione del presente regolamento. […] 8. I punti di contatto nazionali designati dagli Stati membri a norma del paragrafo 9, secondo comma, si riuniscono periodicamente a livello di Unione nella rete dei punti di contatto nazionali. La rete dei punti di contatto nazionali funge da piattaforma per lo scambio periodico di informazioni e prassi eccellenti, nonché per una cooperazione strutturata tra i punti di contatto nazionali e la Commissione in merito a tutti gli aspetti del presente regolamento. In particolare, la	4 Questa chiarezza nel riparto di competenze fra le autorità nazionali non si rinviene anche nel DSA, che assegna al coordinatore di servizi digitali il potere di richiedere informazioni sull'algoritmo e, quindi, anche sul *targeting* (ex art. 40 DSA e Considerando 110).

rete dei punti di contatto nazionali agevola la cooperazione a livello di Unione per quanto riguarda l'applicazione e l'esecuzione del presente regolamento e facilita l'elaborazione, in collaborazione con i pertinenti portatori di interessi, di orientamenti intesi ad aiutare gli sponsor e i prestatori di servizi di pubblicità politica a conformarsi ai requisiti del presente regolamento. La rete dei punti di contatto nazionali si riunisce almeno due volte all'anno e, se necessario, su richiesta debitamente motivata della Commissione o di uno Stato membro. Essa opera in stretta collaborazione con la rete europea di cooperazione in materia elettorale, il gruppo dei regolatori europei per i servizi di media audiovisivi e altre reti od organismi pertinenti, onde agevolare lo scambio rapido e sicuro di informazioni su questioni connesse al controllo e all'applicazione del presente regolamento. La Commissione partecipa alle riunioni della rete dei punti di contatto nazionali e presta assistenza amministrativa. 9. Lo Stato membro che designa più di un'autorità competente provvede a che i rispettivi compiti di tale autorità siano stabiliti chiaramente e che le autorità in questione cooperino strettamente ed efficacemente nello svolgimento dei loro compiti. Ciascuno Stato membro designa un'autorità competente come punto di contatto nazionale a livello	

dell'Unione ai fini di tutti gli aspetti del presente regolamento. I punti di contatto nazionali sostengono e agevolano la cooperazione efficace tra le autorità nazionali competenti e con i punti di contatto nazionali di altri Stati membri. Gli Stati membri mettono a disposizione del pubblico i dati di contatto dei loro punti di contatto nazionali. Gli Stati membri interessati comunicano, se del caso, il nome delle altre autorità competenti e i rispettivi compiti alla rete dei punti di contatto nazionali.	
Art. 24 ***Diritto di presentare un reclamo*** 1. Fatti salvi qualsiasi altra procedura amministrativa o mezzo di ricorso giudiziario, le autorità competenti trattano debitamente ogni notifica di un'eventuale violazione del presente regolamento e, su richiesta, informano la persona che ha effettuato la notifica del seguito dato. Durante l'ultimo mese che precede un'elezione o un referendum, qualsiasi notifica ricevuta in relazione a tale elezione o referendum è trattata senza indebito ritardo. 2. Le autorità competenti trasmettono senza indebito ritardo i reclami che rientrano tra le competenze di un'altra autorità competente in un altro Stato membro a tale autorità competente.	Ogni persona ha diritto di presentare un reclamo all'Autorità nazionale competente che dovrà trattare ogni notifica. Durante l'ultimo mese che precede un'elezione o un referendum, qualsiasi notifica deve essere esaminata senza indebito ritardo.

Articolo 25 ***Sanzioni*** 1. Gli Stati membri stabiliscono le norme relative alle sanzioni o alle altre misure necessarie applicabili agli sponsor o ai prestatori di servizi di pubblicità politica in caso di violazioni degli articoli da 5 a 17, 20 e 21 e adottano tutte le misure necessarie per garantirne la tempestiva attuazione. [...] 5. Le violazioni degli articoli 5, 7, 11, 12, 13, 15, 16 e 18 sono considerate particolarmente gravi quando riguardano messaggi di pubblicità politica pubblicati o diffusi durante l'ultimo mese che precede un'elezione o un referendum e rivolte a cittadini dello Stato membro in cui è organizzata l'elezione o il referendum. Gli Stati membri possono anche imporre sanzioni pecuniarie periodiche per costringere gli sponsor, i prestatori di servizi di pubblicità politica e gli editori di pubblicità politica a porre fine a una violazione grave e reiterata del presente regolamento.	Quanto alle sanzioni da applicare, spetta agli Stati membri stabilire le relative norme all'interno della cornice regolatoria fissata dal legislatore europeo. Alcune violazioni sono considerate automaticamente particolarmente gravi qualora riguardino messaggi di pubblicità politica pubblicati o diffusi durante l'ultimo mese che precede un'elezione o un referendum, in ragione della rilevanza delle suddette norme ai fini della protezione dei diritti dei destinatari.

3 Intelligenza artificiale nei settori regolati da AGCOM: mappatura degli impatti e framework operativo di intervento

Andrea Imperiali

3.1 Introduzione. L'intelligenza artificiale come paradigma trasversale

L'intelligenza artificiale non è una tecnologia riconducibile alle categorie di innovazione ordinaria. La sua portata trasformativa, la rapidità con cui penetra nei processi — visibili e invisibili — che governano la vita quotidiana e la fondamentale ambiguità del suo impatto la collocano in una posizione del tutto peculiare nel panorama regolatorio contemporaneo. L'AI può operare simultaneamente come abilitatore silenzioso di efficienza, come moltiplicatore di squilibri preesistenti e come strumento opaco di manipolazione. Soprattutto, essa possiede la capacità di attraversare — in diagonale — settori, competenze, diritti e linguaggi, ponendo sfide inedite a qualsiasi autorità di regolazione.

In questo scenario, un'autorità indipendente come l'AGCOM non può limitarsi a un'analisi settoriale tradizionale, né può attendere che il legislatore nazionale o europeo definisca un perimetro normativo definitivo che, nel caso dell'intelligenza artificiale, rischia strutturalmente di risultare in ritardo rispetto all'evoluzione tecnologica e dei mercati. Dal punto di vista giuridico-istituzionale, ciò implica la necessità di un approccio anticipatorio e adattivo, che combini la certezza del diritto con la flessibilità degli strumenti regolatori, in un contesto in cui la natura stessa dell'oggetto regolato — sistemi algoritmici autonomi, capaci di apprendere e adattarsi — sfida le tradizionali categorie giuridiche di responsabilità, trasparenza e controllo.

3.1.1 La scelta di partire dagli ambiti di competenza

L'approccio metodologico più diffuso in materia di AI articola l'analisi per settori industriali (sanità, istruzione, trasporti, comunicazioni) o per tipologie tecnologiche (modelli linguistici, sistemi predittivi, raccomandazioni personalizzate). Tale mappatura è utile, ma non sempre sufficiente per le autorità indipendenti, che operano non tanto lungo le linee verticali dei comparti produttivi, quanto lungo assi trasversali di garanzia: tutela dei diritti fondamentali, promozione della concorrenza, vigilanza sulla qualità dei servizi ed equilibrio informativo. Il Comitato ha pertanto scelto di articolare l'analisi in quattro blocchi fondamentali, ciascuno corrispondente a uno degli ambiti formali di competenza dell'Autorità: (i) comunicazioni elettroniche, (ii) servizi digitali, media audiovisivi e radiofonici, (iii) servizi postali, (iv) tutela del consumatore e degli utenti. Questa scelta, forse meno elegante dal punto di vista sistematico, risulta più aderente alla missione istituzionale dell'AGCOM e consente di ancorare l'analisi a ciò che l'Autorità è già chiamata a fare quotidianamente: regolare, monitorare, presidiare.

3.1.2 La distinzione tra impatti effettivi e potenziali

Per ciascun impatto individuato, il rapporto specifica se si tratta di un fenomeno già in atto (impatto effettivo) oppure ragionevolmente prevedibile nel breve-medio termine (impatto potenziale). In diversi casi, entrambe le qualificazioni coesistono. Questa distinzione non è meramente classificatoria, ma operativa e metodologicamente rilevante:

gli impatti **effettivi** richiedono risposte immediate e strumenti prontamente attivabili. Si tratta di fenomeni documentati, corroborati da evidenze empiriche, casi accertati o pronunce di autorità nazionali e internazionali;

gli impatti **potenziali** impongono un'azione anticipatoria, di studio, osservazione e preparazione degli strumenti regolatori. Si tratta di rischi ad alta probabilità di materializzazione nel breve-medio termine, la cui gestione richiede una postura regolatoria vigile e adattiva.

Ne deriva una duplice matrice di prioritizzazione: la prima classifica gli impatti per orizzonte temporale (immediato: 0-12 mesi; medio termine: 1-3 anni; prospettico: 3-5 anni); la seconda incrocia la gravità dell'impatto con il grado di preparazione normativa (gap

normativo), consentendo di classificare le priorità di intervento in quattro livelli: urgenza massima (impatto grave con gap normativo elevato), alta priorità (impatto medio-alto con gap significativo), monitoraggio (impatto gestibile con copertura parziale) e nessuna azione immediata (basso impatto o copertura adeguata).

3.1.3 La struttura analitica: ex ante, in itinere, ex post

Per ciascun impatto identificato, l'analisi è stata strutturata secondo cinque assi: (a) descrizione sintetica dell'impatto, con qualificazione della sua natura (effettivo/potenziale); (b) matrici di prioritizzazione (classificazione temporale e rapporto gravità/preparazione normativa); (c) attività ex ante (prevenzione e preparazione); (d) attività in itinere (vigilanza continua); (e) attività ex post (correzione e enforcement).

Questa tripartizione temporale riflette il ciclo di vita di una politica regolatoria evoluta e moderna: prevenire, vigilare, correggere. Non si tratta di un esercizio teorico, ma della simulazione concreta di come l'AGCOM potrebbe intervenire di fronte a ciascun rischio o utilizzo dell'AI che incide sulle sue competenze. L'obiettivo naturalmente non è anticipare scelte regolatorie puntuali, ma fornire all'Autorità una bussola analitica per orientarsi nella complessità, individuando le aree di maggiore urgenza, i rischi sistemici emergenti e le opportunità di rafforzamento del ruolo istituzionale dell'AGCOM in un ecosistema digitale sempre più guidato da sistemi di AI.

Le attività ex ante sono state ulteriormente suddivise in quattro categorie: codici di condotta e co-regolamentazione; linee guida e regolamenti; eventuali sistemi di autorizzazione o notifica; attività di formazione e alfabetizzazione. Le attività in itinere includono: procedure e standard; strumenti specialistici (monitoraggio, audit, simulazioni); laboratori tecnici (*sandbox, testbed*, osservatori); attività in outsourcing o partnership tecnico-scientifiche. Le attività ex post comprendono: obblighi correttivi, misure di trasparenza verso l'utenza e obblighi di reporting o rendicontazione.

3.1.4 Nota per il lettore sul perimetro delle raccomandazioni

Il presente contributo è redatto in una prospettiva volutamente ad ampio spettro: pur muovendo dagli ambiti formali di competenza dell'AGCOM come asse organizzativo

dell'analisi, alcune delle implicazioni di policy individuate e delle raccomandazioni formulate trascendono il perimetro diretto delle attribuzioni dell'Autorità, investendo materie — quali la disciplina antitrust, la tutela della riservatezza, la governance istituzionale dell'AI Act — che ricadono nella competenza di altre autorità nazionali, del legislatore o della Commissione europea. Tale scelta è consapevole e metodologicamente giustificata: le autorità indipendenti operano non lungo gli assi verticali dei perimetri settoriali, ma lungo quelli orizzontali della tutela dei diritti, del pluralismo e dell'equilibrio dei mercati, e una mappatura analitica dell'AI non può che rispecchiare questa vocazione trasversale. Il Consiglio AGCOM, destinatario di questo Rapporto, valuterà, nella propria autonomia, quali raccomandazioni assumere come proprie e quali trasmettere alle sedi istituzionali competenti.

3.2 Comunicazioni elettroniche

3.2.1 Il quadro degli impatti

L'intelligenza artificiale sta trasformando profondamente il settore delle comunicazioni elettroniche, con impatti che investono la gestione delle infrastrutture di rete, la neutralità del traffico dati, la concorrenza tra operatori e la tutela degli utenti. Il Comitato ha identificato venti impatti specifici in questo ambito, di cui quattro classificati a urgenza massima.

Gestione dinamica delle reti. I sistemi AI sono già centrali nell'ottimizzazione del traffico, nell'allocazione automatizzata delle risorse di rete e nell'anticipazione dei guasti infrastrutturali. Tramite modelli predittivi e monitoraggio in tempo reale, l'AI migliora la stabilità e la resilienza dei servizi, con benefici tangibili quali la riduzione delle interruzioni e la manutenzione preventiva. Tuttavia, la progressiva delega di decisioni tecniche critiche agli algoritmi introduce rischi significativi che meritano attenzione regolatoria: errori non previsti che possono propagarsi su larga scala, vulnerabilità a sabotaggi informatici (*adversarial attacks* sui sistemi predittivi), e soprattutto, perdita progressiva del controllo umano nei processi critici. Il principio di *human oversight*, cardine dell'AI Act (art. 14), trova qui una delle sue applicazioni più rilevanti.

Neutralità della rete e discriminazione algoritmica del traffico dati. Uno degli impatti più rilevanti e urgenti riguarda il rischio che i sistemi AI per la gestione del traffico dati

operino discriminazioni non trasparenti tra diversi flussi. Nella pratica, ciò si declina in fenomeni specifici e documentati:

Zero rating dinamico: i sistemi AI identificano in tempo reale i servizi partner commerciali dell'operatore ed escludono il relativo traffico dal consumo dati, penalizzando i concorrenti. Il caso italiano più significativo ha riguardato un operatore mobile che offriva traffico illimitato per i servizi di streaming musicale partner, mentre il traffico verso concorrenti veniva conteggiato nel consumo.

Throttling selettivo: riduzione algoritmica della velocità per applicazioni specifiche come VoIP (WhatsApp calls, Skype) durante le ore di punta, senza trasparenza verso gli utenti, al fine di incentivare l'uso delle chiamate tradizionali a pagamento.

Prioritizzazione commerciale: allocazione preferenziale di banda a servizi "premium" (*cloud gaming*, videoconferenze *enterprise*) a scapito di servizi meno redditizi come siti informativi, piccoli e-commerce o servizi pubblici digitali. Test indipendenti come Wehe App, Northeastern University hanno evidenziato diversi casi di *traffic differentiation e throttling* su diverse reti mobili a livello globale, anche se non si registrano dati pubblici consolidati con percentuali specifiche per il mercato italiano.

Le evidenze raccolte dimostrano che queste pratiche sono già in atto: tra il 2018 e il 2023, AGCOM ha ricevuto diverse segnalazioni relative a sospette violazioni della neutralità della rete, aprendo procedimenti sanzionatori di cui alcuni conclusi con accertamento di violazione. Il Regolamento UE 2015/2120 sull'Open Internet vieta espressamente *blocking, throttling* e prioritizzazione commerciale (art. 3, par. 3), ma l'assenza di strumenti di testing automatizzato, l'adeguatezza delle sanzioni e l'onere della prova a carico dell'utente rendono l'enforcement strutturalmente debole.

Impatto infrastrutturale dei modelli AI di grandi dimensioni. La crescita dei servizi AI generativi sta aumentando significativamente il fabbisogno di capacità di rete a livello di interconnessioni tra data center e *peering* internazionale, con proiezioni che indicano che entro il 2030 una quota consistente del traffico di rete coinvolgerà elaborazione AI (Omdia/Frontier Economics, 2024); l'impatto sul traffico internet consumer rimane invece più contenuto per i servizi di tipo testuale, mentre diventa rilevante per quelli multimodali (audio, video generativo). Le implicazioni infrastrutturali sono molteplici: necessità di

anticipare gli investimenti di *capacity* di 2-3 anni rispetto ai piani originari; pressione sugli accordi di *peering* internazionali (traffico dati UE-USA per l'accesso ai modelli cloud); e difficoltà di pianificazione dovuta all'imprevedibilità dei pattern di utilizzo. Il dibattito europeo sul cosiddetto *fair share* — la richiesta dei grandi operatori TLC di far contribuire i principali generatori di traffico ai costi infrastrutturali — è stato rilanciato nel 2022-2023, ma la consultazione pubblica della Commissione (ottobre 2023) ne ha registrato il rigetto da parte della maggioranza dei rispondenti e del Parlamento europeo (BEREC, Internet Society, 2023).

***Edge computing* e sovranità digitale.** L'elaborazione dei dati sempre più vicina all'utente finale (*edge computing*) consente latenze ridotte e prestazioni migliorate, ma sposta il trattamento di informazioni personali su nodi distribuiti, spesso al di fuori del territorio nazionale. L'intelligenza artificiale distribuita su migliaia di *edge node* rende complessa la tracciabilità dei flussi di dati, la verifica della conformità al GDPR e l'esercizio dei diritti degli interessati. Il rischio è la perdita di controllo sulla localizzazione fisica dei dati, con implicazioni sulla sovranità digitale, sulla giurisdizione applicabile e sulla capacità di *enforcement*.

Concentrazione del potere di mercato. Gli operatori capaci di sviluppare algoritmi AI proprietari e di accedere a dataset massivi acquisiscono vantaggi competitivi difficili da colmare, creando nuove barriere all'ingresso per i concorrenti. Il fenomeno opera attraverso un circolo vizioso (*flywheel effect*): più dati generano algoritmi migliori, che producono servizi superiori, che attraggono più clienti, che generano ulteriori dati. I dati del settore sono eloquenti: in Italia, TIM, Vodafone (ora confluita in Fastweb+Vodafone) e WindTre controllano quasi l'80% del mercato mobile; gli investimenti AI dichiarati ammontano a diverse centinaia di milioni di euro per il triennio 2023-2025, contro poche decine di challenger come Iliad. A livello europeo, le acquisizioni *AI-centric* nel settore TLC sono cresciute significativamente tra il 2020 e il 2023, con la quasi totalità delle operazioni sotto le soglie di notifica obbligatoria — un dato che rivela l'inadeguatezza strutturale degli strumenti antitrust tradizionali per mercati dove il valore risiede in asset immateriali (dati, algoritmi, know-how).

***Pricing* dinamico e collusione algoritmica.** L'uso di algoritmi che monitorano in tempo reale le strategie di prezzo dei concorrenti e aggiornano automaticamente le tariffe può condurre a fenomeni di collusione tacita (*tacit algorithmic collusion*): convergenza dei prezzi verso livelli sovracompetitivi senza alcun accordo esplicito tra operatori. Il meccanismo opera attraverso quattro fasi: monitoraggio (*scraping* automatico dei *prezzi*), analisi (identificazione del *pricing* ottimale tramite *game theory* e *reinforcement learning*), risposta (aggiustamento automatico), ed equilibrio collusivo (tutti gli algoritmi convergono verso lo stesso *pricing* senza comunicazione). L'*Autorité de la concurrence* francese ha documentato casi di convergenza anomala dei prezzi tra operatori mobili, nonostante riduzioni dei costi infrastrutturali. Simulazioni accademiche (Calvano et al., 2020, American Economic Review) dimostrano che algoritmi di tipo Q-learning, non programmati per colludere, sviluppano autonomamente strategie collusive con prezzi sistematicamente sovracompetitivi. Il paradosso giuridico è che l'art. 101 TFUE richiede la prova di una "concertazione" che, per definizione, è assente nella collusione algoritmica.

Gestione dinamica dello spettro radioelettrico. L'AI permette l'allocazione dinamica e predittiva delle frequenze in tempo reale, ottimizzando l'uso dello spettro scarso. Tuttavia, algoritmi proprietari potrebbero favorire determinati operatori o servizi, generare interferenze non previste o allocare risorse in modo discriminatorio senza trasparenza. Il rischio è la perdita del controllo pubblico su una risorsa strategica, con possibili violazioni dei principi di equità nell'accesso allo spettro e distorsioni competitive.

3.2.2 Le priorità di intervento e le raccomandazioni operative

La matrice di prioritizzazione identifica sei impatti a urgenza massima nel settore delle comunicazioni elettroniche: la violazione della *net neutrality* per mezzo di algoritmi, la concentrazione del potere di mercato *AI-driven*, il *pricing* dinamico collusivo, le fusioni *AI-centric*, la sorveglianza e privacy, e il *traffic management* discriminatorio. Per ciascuno di essi, il gap normativo è stato valutato come elevato, richiedendo interventi regolatori immediati.

Tra gli interventi più urgenti, peraltro non tutti di diretta competenza di AGCOM, si segnalano: l'adozione di strumenti automatizzati di testing della *net neutrality* (sul modello del *Net Neutrality Reference Measurement System* sviluppato da BEREC e già adottato da

BNetzA in Germania nell'ambito del servizio breitbandmessung.de, disponibile in open source su GitHub); l'introduzione di obblighi di notifica preventiva per i sistemi AI utilizzati nella gestione del traffico; valutare l'introduzione di strumenti che riequilibrino l'onere probatorio in caso di segnalazioni di discriminazione (l'operatore deve dimostrare la non-discriminazione, non l'utente dimostrare la violazione); la definizione di criteri di valutazione del potere di mercato che includano il possesso di dataset strategici, le capacità AI proprietarie e l'integrazione verticale algoritmi-infrastruttura-servizi; la creazione di un sistema di *monitoring* automatizzato del parallelismo dei prezzi con alert in caso di convergenze anomale; e l'istituzione di un registro volontario delle acquisizioni *AI-centric* con incentivi alla notifica (*fast-track pre-clearance*) e disincentivi all'omissione (investigazione approfondita in caso di scoperta post-facto).

3.3 Servizi digitali, media audiovisivi e radiofonici

3.3.1 Il quadro degli impatti

Il settore dei servizi digitali e dei media audiovisivi è quello in cui l'impatto dell'intelligenza artificiale assume le dimensioni più allarmanti, tanto per la pervasività dei fenomeni quanto per la loro rilevanza in termini di diritti fondamentali — libertà di espressione, pluralismo informativo, dignità della persona, tutela dei minori, diritto d'autore. Il Comitato ha identificato trenta impatti in questo ambito, di cui quattro classificati a urgenza massima.

Deepfake ed ecosistema della disinformazione. La generazione automatica di contenuti audiovisivi — *deepfake video*, sintesi vocale, *synthetic media* — ha raggiunto livelli di sofisticazione tali da rendere i contenuti falsi indistinguibili da quelli autentici per la grande maggioranza degli utenti. Le evidenze empiriche sono inequivocabili: nelle elezioni parlamentari slovacche del settembre 2023, un audio deepfake di un leader politico ha sollevato preoccupazioni circa un potenziale impatto sulla perdita stimata nei sondaggi exit poll e persino di qualche seggio in Parlamento — e il fact-checking ha impiegato 18 ore per la smentita, un tempo incompatibile con la velocità della diffusione virale; nelle primarie statunitensi del 2024, *robocall* con voce clonata del Presidente Biden hanno invitato gli elettori a non votare; video deepfake di candidati in situazioni compromettenti sono stati

diffusi su X e TikTok. In Italia, in occasione delle elezioni europee del 2024, AGCOM ha ricevuto diverse segnalazioni di possibili contenuti deepfake. L'AI Act (art. 50, par. 4) prevede l'obbligo di etichettatura dei contenuti generati o manipolati da AI «in modo chiaro e distinguibile», ma l'ambiguità di tale formulazione, l'assenza di standard tecnici attuativi (formato del label, posizionamento, linguaggio), la limitata capacità di *detection* da parte delle autorità e l'assenza di meccanismi di risposta rapida rendono il *framework* normativo attualmente insufficiente.

Motori di ricerca *AI-powered* e sostenibilità del giornalismo. L'introduzione di funzionalità *AI-powered* nei motori di ricerca (Google AI Overview, Bing AI Mode, Perplexity) modifica radicalmente l'interazione utente-informazione: invece di fornire link a fonti esterne, questi sistemi generano risposte sintetiche dirette basate su contenuti estratti dal web. Lo *shift* produce conseguenze multiple: una drastica riduzione del traffico verso siti editoriali (con cali documentati), con impatto diretto sulla sostenibilità economica del giornalismo e in particolare del giornalismo locale; una concentrazione informativa mediante la presentazione di una "risposta unica" invece di una pluralità di fonti, compromettendo così l'esposizione degli utenti a prospettive diverse (anche quando i suddetti motori di ricerca indicano in calce le fonti); il riutilizzo di contenuti giornalistici senza adeguato compenso, configurando potenziale violazione dell'art. 15 della Direttiva Copyright; opacità su fonti e metodologia di sintesi, con rischio di disinformazione se l'AI interpreta erroneamente o estrapola contenuti fuori contesto; e fenomeni di *self-preferencing* da parte di piattaforme in posizione dominante che trattengono gli utenti invece di rimandarli al web aperto, con profili rilevanti sotto il Digital Markets Act.

Algoritmi di raccomandazione e pluralismo. Gli algoritmi di raccomandazione ottimizzati tramite AI modellano le preferenze degli utenti creando «bolle informative» che limitano l'esposizione a prospettive diverse. I dati empirici sono significativi: analisi indipendenti (es. Mozilla Foundation) hanno evidenziato che il 71% dei contenuti ha ricevuto raccomandazioni "regrettabili" (violenza, disinformazione, teorie complottiste) e che molte delle raccomandazioni problematiche riguardavano contenuti mai cercati dall'utente, proposti proattivamente dall'algoritmo. Il DSA (artt. 34 e 38) impone alle VLOP di valutare i rischi sistemici, inclusi quelli derivanti dagli algoritmi di raccomandazione, e di offrire agli

utenti almeno un'opzione non basata sulla profilazione, ma non definisce in modo puntuale metriche standardizzate e audit indipendenti verificabili.

Diritto d'autore e addestramento dei modelli. L'utilizzo massivo di opere audiovisive protette per l'addestramento dei modelli generativi solleva questioni critiche di equo compenso, consenso degli autori e tracciabilità dell'origine dei contenuti. È ampiamente riconosciuto che i dataset utilizzati per l'addestramento dei modelli linguistici di grandi dimensioni includono, in misura significativa, contenuti protetti da diritto d'autore, spesso acquisiti attraverso tecniche di raccolta massiva di dati online. La questione del possibile utilizzo di opere protette senza adeguati meccanismi di licenza o remunerazione è attualmente oggetto di contenziosi giudiziari e di un intenso dibattito regolatorio a livello europeo e internazionale, anche in relazione alla definizione di modelli sostenibili di compensazione per gli aventi diritto. In Italia, soggetti rappresentativi degli autori e degli editori, tra cui SIAE e AIE, hanno preso posizione nel dibattito sull'utilizzo delle opere protette da parte dei sistemi di intelligenza artificiale, sollecitando maggiore trasparenza, meccanismi di licensing e forme di remunerazione per gli aventi diritto. L'eccezione per il *Text and Data Mining* prevista dall'art. 4 della Direttiva (UE) 2019/790 consente l'uso di contenuti anche per finalità commerciali, salvo *opt-out* espresso dai titolari dei diritti. Tuttavia, nel contesto dei modelli di IA generativa, sono emerse criticità applicative legate alla difficoltà di verificare in modo sistematico tali *opt-out* nei processi di raccolta massiva dei dati. L'AI Act introduce obblighi specifici per i modelli di uso generale (GPAI), tra cui la pubblicazione di una sintesi dei dati di addestramento e misure di conformità al diritto d'autore. Tuttavia, il regolamento non affronta direttamente il tema della remunerazione degli aventi diritto, lasciando aperta la necessità di ulteriori sviluppi normativi in materia di licensing e compensazione.

Pubblicità *AI-driven*, *chatbot* pubblicitari e pratiche commerciali. L'integrazione di sistemi AI conversazionali nel marketing digitale assume forme sempre più sofisticate e giuridicamente problematiche. Da un lato, chatbot che fungono da brand influencer sviluppano relazioni parasociali per promuovere prodotti; dall'altro, OpenAI ha annunciato per il 2025 l'inserimento di contenuti sponsorizzati nelle conversazioni con ChatGPT. Questi strumenti sfruttano personalizzazione conversazionale, analisi emotiva in tempo reale e

capacità di adattamento comportamentale per orientare scelte d'acquisto attraverso interazioni apparentemente neutre o assistive. Tra i rischi emergenti: pubblicità occulta mascherata da raccomandazioni personalizzate, manipolazione comportamentale mediante sfruttamento di vulnerabilità psicologiche, profilazione massiva senza consenso informato adeguato, e targeting di minori e soggetti vulnerabili. La commistione tra funzione assistiva e finalità commerciale rende poco trasparente la natura pubblicitaria dell'interazione, eludendo i requisiti di *disclosure* previsti dalla normativa su trasparenza pubblicitaria e pratiche commerciali sleali, e potenzialmente configurando forme di manipolazione cognitiva subliminale vietate dall'art. 5 dell'AI Act.

Manipolazione elettorale, par condicio e micro-targeting politico. Gli algoritmi di raccomandazione e ranking delle piattaforme digitali possono alterare la visibilità di candidati e partiti in modo non trasparente, compromettendo la par condicio. La logica con cui le piattaforme mostrano determinati contenuti politici è spesso imperscrutabile: candidati, forze politiche o idee minoritarie possono essere sistematicamente penalizzati senza possibilità di verifica esterna. Il micro-targeting politico tramite AI predittiva consente campagne mirate su base individuale, sfruttando profilazione psicografica e vulnerabilità degli utenti — come dimostrato dal precedente Cambridge Analytica e dalle sue evoluzioni tecnologiche. La produzione automatica di contenuti politici sintetici (*fake interview, synthetic news, video deepfake*) rappresenta una minaccia concreta alla integrità dei processi democratici. La Legge 28/2000 (par condicio), concepita in un contesto mediale prevalentemente radiotelevisivo, non contempla le dinamiche proprie dell'ecosistema digitale, né la dimensione algoritmica della distribuzione e visibilità dei contenuti politici.

Concentrazione della filiera nelle mani di piattaforme AI-owned. Le principali piattaforme tecnologiche integrano in misura crescente diverse componenti della filiera digitale, includendo lo sviluppo di sistemi di intelligenza artificiale, la gestione delle infrastrutture e il controllo dei canali di distribuzione dei contenuti. Tale integrazione rafforza dinamiche di concentrazione del potere di mercato, già riconosciute a livello europeo attraverso la disciplina dei "*gatekeeper*", e può favorire pratiche di *self-preferencing*, incidendo sulle condizioni di accesso al mercato per operatori terzi e sulla distribuzione del valore economico tra piattaforme e creatori di contenuti. Inoltre, il ruolo centrale dei sistemi

di raccomandazione algoritmica consente alle piattaforme di influenzare in modo significativo la visibilità e la circolazione delle informazioni nel dibattito pubblico. In questo contesto, le tradizionali misure antitrust, concepite per mercati industriali, mostrano limiti nell'applicazione a modelli di business fondati su dati, algoritmi e dinamiche di piattaforma, come evidenziato dall'evoluzione del quadro regolatorio europeo.

3.3.2 Le priorità di intervento e le raccomandazioni operative

Il settore registra quattro impatti a urgenza massima e sei prioritari: deepfake e disinformazione, manipolazione elettorale AI, violazione del diritto d'autore per addestramento, microtargeting politico, opacità del ranking dei contenuti politici e concentrazione proprietaria delle piattaforme AI-owned. Le raccomandazioni operative, anche qui non tutte da riferirsi in modo diretto ed esclusivo ad AGCOM, includono: l'espansione e rafforzamento della task force sulla *par condicio* elettorale, con capacità operativa H24 durante le campagne (composta da AGCOM, Polizia Postale e rappresentanti delle piattaforme); la definizione di protocolli con le principali piattaforme per la rimozione rapida di deepfake politici (entro 2 ore dalla segnalazione); la definizione di standard attuativi per l'etichettatura dei contenuti *AI-generated* ai sensi dell'art. 50 dell'AI Act (formato, posizionamento, linguaggio); l'acquisizione o sviluppo di capacità forense per la *detection* dei deepfake (partnership con Politecnico di Milano, FBK Trento, o software commerciale come *Sensity* e *Reality Defender*); l'avvio di un tavolo con SIAE, Ministero della Cultura e Garante Privacy per un framework italiano di licensing ed equo compenso; la promozione di standard di *watermarking* crittografico (C2PA, *Content Authenticity Initiative*) per l'autenticità dei contenuti; e una proposta legislativa per valutare l'opportunità di introdurre fattispecie specifiche per l'introduzione di un reato di *deepfake* elettorale con pene aggravate e inversione dell'onere della prova in sede civile.

3.4 Servizi postali

3.4.1 Il quadro degli impatti

Sebbene il settore postale presenti una complessità moderata rispetto ai precedenti (due priorità alte su otto impatti identificati), l'intelligenza artificiale vi produce impatti significativi, con implicazioni rilevanti in termini di coesione territoriale, inclusione sociale e inclusione finanziaria.

Ottimizzazione logistica e divari territoriali. I sistemi AI per il *sorting*, il *routing* e la previsione delle consegne migliorano l'efficienza operativa, ma tendono strutturalmente a privilegiare la redditività economica (costo/consegna, densità spedizioni/km^2), penalizzando sistematicamente le aree a bassa densità abitativa, i comuni con popolazione anziana e le isole minori. I dati di Poste Italiane documentano riduzioni significative dei punti di accesso fisico e degli orari di servizio nelle aree rurali, montane e del Mezzogiorno. Negli ultimi anni, le attività di monitoraggio e le segnalazioni pervenute all'Autorità e ad altri soggetti istituzionali evidenziano criticità nella qualità e nella continuità del servizio universale postale, in particolare in alcune aree territoriali. In questo contesto, l'introduzione di strumenti avanzati di ottimizzazione operativa, inclusi sistemi algoritmici, pone una questione regolatoria rilevante relativa al bilanciamento tra obiettivi di efficienza e obblighi di servizio universale. Tale profilo richiama direttamente i principi fondanti della regolazione postale, come definiti dalla Direttiva 97/67/CE e dal D.lgs. 261/1999.

Servizi finanziari postali e discriminazione algoritmica. Gli operatori postali offrono servizi finanziari (BancoPosta, PostePay, prestiti, investimenti) che integrano AI per *credit scoring*, consulenza automatizzata (*robo-advisory*) e rilevamento frodi. Il rischio è la discriminazione algoritmica: modelli addestrati su dati storici che riflettono *bias* socioeconomici, geografici e anagrafici possono penalizzare sistematicamente donne, giovani, residenti in aree svantaggiate e lavoratori precari. La mancanza di trasparenza nei criteri decisionali impedisce di contestare dinieghi e verificare l'equità delle decisioni, in violazione dei principi di pari trattamento e tutela del consumatore. L'AI Act classifica i sistemi di valutazione del merito creditizio (credit scoring) tra i sistemi ad alto rischio (art. 6 e Allegato III, punto 5(b)), imponendo obblighi in materia di gestione del rischio, qualità

dei dati e mitigazione dei *bias*, trasparenza, supervisione umana e robustezza dei sistemi. Il sistema di *enforcement* è tuttavia in fase di progressiva implementazione e richiederà lo sviluppo di standard tecnici, nonché il rafforzamento delle capacità di vigilanza e del coordinamento tra le autorità competenti.

***Green logistics* e *greenwashing* algoritmico.** I sistemi AI per l'ottimizzazione dei percorsi di consegna possono ridurre emissioni e sprechi, ma il rischio emergente è il greenwashing algoritmico: operatori che dichiarano riduzioni di emissioni basate su modelli AI non verificabili, con poca trasparenza su dati, metodologie e risultati reali. Senza standard di misurazione indipendenti, audit e certificazioni credibili, le promesse di sostenibilità *AI-based* rischiano di tradursi in marketing ingannevole.

Tracciabilità *blockchain+AI*. L'integrazione di *blockchain* e AI promette tracciabilità *end-to-end* delle spedizioni e contrasto alla contraffazione, ma solleva questioni di sorveglianza commerciale: ogni movimento di merci, e indirettamente di persone, è monitorato, profilato e analizzato per finalità commerciali, con rischi di profilazione delle abitudini di acquisto e possibili discriminazioni di fornitori o clienti.

3.4.2 Le priorità di intervento

La mappatura suggerisce: l'introduzione di vincoli espliciti negli algoritmi di ottimizzazione logistica che garantiscano standard minimi del servizio universale indipendentemente dalla redditività; la definizione di KPI di accessibilità territoriale non derogabili, con audit periodici; l'adozione di obblighi di *explainability* e *audit* antidiscriminazione per i sistemi di credit scoring postale, con diritto dell'utente a una motivazione dettagliata dei dinieghi e a una revisione umana; e la definizione di standard di reporting ambientale per le dichiarazioni di *green logistics AI-driven*, con certificazioni indipendenti e sanzioni per *greenwashing* accertato.

3.5 Tutela del consumatore e degli utenti

3.5.1 Il quadro degli impatti

La tutela del consumatore e degli utenti rappresenta il settore con la criticità massima: la mappatura identifica ben quattro impatti a urgenza massima su quattordici impatti complessivi, con rischi che investono direttamente i diritti fondamentali della persona.

Decisioni automatizzate opache. Un utente può ricevere una decisione — accettazione o rifiuto di un rimborso, modifica contrattuale, applicazione di penali, offerta personalizzata — senza alcuna possibilità di comprenderne la logica sottostante o di contestarne l'esito. Il diritto a non essere sottoposti a decisioni basate unicamente su trattamento automatizzato con effetti giuridici o analogamente significativi, sancito dall'art. 22 del GDPR, presenta rilevanti criticità applicative nella prassi. In particolare, gli operatori possono fare ricorso alle eccezioni previste dal regolamento, tra cui quelle legate alla necessità contrattuale, mentre le modalità di esercizio dei diritti e di contestazione risultano in molti casi complesse o poco trasparenti per gli utenti. Inoltre, l'utente medio può incontrare difficoltà nell'identificare e comprendere decisioni automatizzate e nell'attivare efficacemente i meccanismi di tutela previsti. L'AI Act introduce obblighi complementari, in particolare in materia di trasparenza e documentazione tecnica dei sistemi ad alto rischio, ma permangono sfide rilevanti in termini di enforcement e capacità di vigilanza.

Sorveglianza biometrica. I sistemi AI per l'identificazione biometrica — riconoscimento facciale, vocale, analisi dell'andatura, biometria comportamentale — sono già diffusamente impiegati da operatori TLC, *service provider* e piattaforme per sicurezza di rete, autenticazione clienti e cooperazione con l*aw enforcement*. Il caso Clearview AI (2022), conclusosi con una sanzione di 20 milioni di euro da parte del Garante per la protezione dei dati personali per l'utilizzo di sistemi di riconoscimento facciale basati su *scraping* massivo di immagini online, rappresenta un esempio emblematico dei rischi connessi all'uso di dati biometrici. Ulteriore profilo di rischio è rappresentato dal cosiddetto *function creep*, ossia il riutilizzo di dati biometrici raccolti per finalità specifiche (es. autenticazione) per scopi ulteriori, quali marketing o profilazione, in assenza di una base giuridica adeguata. L'AI Act vieta alcune pratiche, tra cui il *social scoring* (art. 5, par. 1, lett. c) e la categorizzazione

biometrica volta a inferire caratteristiche sensibili (art. 5, par. 1, lett. d), nonché l'identificazione biometrica remota in tempo reale nei luoghi pubblici, salvo specifiche eccezioni per finalità di contrasto a reati gravi. L'ampiezza e la complessità delle condizioni previste per tali eccezioni hanno sollevato interrogativi circa il rischio di applicazioni estensive, mentre l'effettività dell'enforcement dei divieti previsti richiederà un significativo rafforzamento delle capacità di vigilanza.

Manipolazione cognitiva dei soggetti vulnerabili. Interfacce adattive e sistemi predittivi che sfruttano vulnerabilità psicologiche possono indurre comportamenti non consapevoli, specialmente tra minori e soggetti fragili: dipendenza da social media (*addiction by design*), trappole d'acquisto, polarizzazione estrema di opinioni. L'AI Act vieta la manipolazione cognitiva subliminale che causa danni significativi (art. 5, par. 1, lett. a), ma i cosiddetti *dark patterns AI-powered* — design dell'interfaccia che sfrutta vulnerabilità cognitive tramite AI per orientare le scelte dell'utente — occupano una zona grigia normativa. In Italia, circa il 45–46% della popolazione tra i 16 e i 74 anni possiede almeno competenze digitali di base (fonte: Eurostat). Permangono inoltre divari significativi legati all'età, al livello di istruzione e al contesto territoriale, particolarmente evidenti nel Mezzogiorno e nelle aree interne.

AI nei sistemi ADR e *legaltech*. L'integrazione di AI nei sistemi di risoluzione alternativa delle controversie (ODR) per analizzare reclami, proporre soluzioni e mediare tra le parti presenta rischi di *bias* pro-azienda (i modelli sono addestrati su storici decisionali che favoriscono l'impresa), mancanza di personalizzazione per casi complessi, e compressione delle garanzie procedurali. Parallelamente, la diffusione di strumenti di assistenza legale automatizzata (legaltech), inclusi chatbot giuridici, può favorire l'accesso a informazioni legali di base, contribuendo a ridurre barriere economiche e informative. Tuttavia, tali strumenti possono comportare rischi di informazioni giuridiche errate o incomplete, difficoltà nell'adattare le risposte alle specificità del caso concreto e profili di incertezza in merito alla responsabilità per le indicazioni fornite. In questo contesto, assume particolare rilievo la distinzione tra informazione legale generica, generalmente ammissibile, e consulenza personalizzata, riservata ai professionisti abilitati, ponendo l'esigenza di chiarire anche sul piano regolatorio i confini tra le due fattispecie e i relativi regimi di responsabilità.

Discriminazione algoritmica in assicurazioni e credito. Il settore assicurativo e creditizio utilizza massivamente AI per valutare rischi (underwriting), calcolare premi personalizzati (pricing) e gestire sinistri (claims processing). La letteratura accademica e istituzionale documenta come i sistemi algoritmici possano produrre effetti discriminatori, anche indiretti, a danno di specifici gruppi, in relazione a caratteristiche quali genere, origine etnica, condizioni socio-economiche o stato di salute. Tali effetti possono emergere anche in assenza di un utilizzo esplicito di dati sensibili, attraverso l'impiego di variabili apparentemente neutrali (proxy), come la localizzazione geografica, il comportamento digitale o altri indicatori indiretti, che risultano statisticamente correlati a caratteristiche protette.
Questo fenomeno pone rilevanti criticità sotto il profilo della non discriminazione e della tutela degli utenti, richiedendo adeguati strumenti di valutazione, trasparenza e controllo dei sistemi algoritmici. La profilazione predittiva rischia di trasformare probabilità statistiche in destino individuale, cristallizzando e amplificando le disuguaglianze esistenti.

3.5.2 Le priorità di intervento

I quattro impatti a urgenza massima — decisioni automatizzate opache, sorveglianza biometrica, manipolazione cognitiva dei vulnerabili e deepfake con danni individuali — richiedono interventi immediati e strutturali. Si raccomandano: l'adozione di obblighi di *explainability* per tutte le decisioni automatizzate; il diritto sempre garantito alla revisione umana su domanda; l'istituzione di uno sportello centralizzato per le segnalazioni dei consumatori relative a decisioni automatizzate, integrato con i meccanismi esistenti di tutela degli utenti e gestione dei reclami; una moratoria sull'uso del riconoscimento facciale per finalità diverse dall'autenticazione 1-a-1; protocolli formali di coordinamento con il Garante Privacy, l'AGCM, Banca d'Italia e IVASS per audit congiunti; il crescente utilizzo di tecniche di generazione sintetica dei contenuti ha alimentato il dibattito sulla possibile introduzione di fattispecie specifiche relative ai deepfake non consensuali, anche con riferimento al rafforzamento degli strumenti di tutela delle vittime. Parallelamente, il quadro normativo europeo – in particolare il Digital Services Act – introduce divieti in materia di dark patterns e prevede specifiche tutele per i minori. L'evoluzione dei sistemi di intelligenza artificiale, inclusi quelli in grado di personalizzare e ottimizzare le interazioni con gli utenti, rende

tuttavia necessario un ulteriore sviluppo di strumenti regolatori volti a prevenire pratiche manipolative.

3.6 Conclusioni: verso una strategia operativa integrata

3.6.1 Sintesi trasversale delle priorità

Dall'analisi trasversale dei quattro ambiti di competenza emergono diversi impatti a urgenza massima, con rilevanza cross-settoriale. L'elenco che segue non è una mera catalogazione, ma una gerarchia operativa: ogni voce riflette la combinazione di gravità dell'impatto, ampiezza del gap normativo, trasversalità rispetto agli ambiti di competenza AGCOM e urgenza temporale dell'intervento.

Manipolazione elettorale AI (Media, Tutela consumatori) — richiede linee guida per la par condicio algoritmica in ambiente digitale e *labeling* obbligatorio dei contenuti politici *AI-generated*. L'urgenza è massima in vista delle prossime elezioni politiche italiane (2027): senza un protocollo operativo attivabile con le piattaforme, il rischio di un «caso Slovacchia italiano» è concreto.

Deepfake e disinformazione (Media, Tutela consumatori) — richiede un framework integrato di *detection*, trasparenza sull'origine dei contenuti e risposta rapida. Il tempo che intercorre tra la diffusione di un *deepfake* e la sua smentita (attualmente stimato in 18 ore nel migliore dei casi) è incompatibile con i cicli dell'informazione digitale.

Violazione della *net neutrality* algoritmica (Telecomunicazioni) — l'evoluzione dei sistemi di gestione del traffico e delle reti, sempre più basati su logiche automatizzate e algoritmiche, pone nuovi interrogativi in merito all'applicazione del principio di neutralità della rete. In questo contesto, può emergere l'esigenza di estendere gli strumenti di monitoraggio e verifica anche alle modalità algoritmiche di gestione del traffico, attraverso lo sviluppo di standard tecnici adeguati e il rafforzamento dei meccanismi di *enforcement*. Strumenti tecnici sviluppati in ambito BEREC e da alcune autorità nazionali consentono già il monitoraggio di specifici aspetti della neutralità della rete e potrebbero costituire una base per ulteriori sviluppi in chiave evolutiva.

Decisioni automatizzate non trasparenti (Telecomunicazioni, Postali, Tutela consumatori) — richiede *explainability* obbligatoria, diritto di ricorso umano e un sistema di segnalazione accessibile. È l'impatto con la maggiore diffusione trasversale: dal *customer care* TLC al *credit scoring* postale, dalla moderazione dei contenuti alle procedure ADR.

Concentrazione del potere di mercato *AI-driven* (Telecomunicazioni, Media) — la crescente concentrazione del potere di mercato nei settori delle telecomunicazioni e dei media digitali è ulteriormente rafforzata dall'utilizzo intensivo di dati e sistemi algoritmici. In questo contesto, emerge l'esigenza di evolvere gli strumenti di analisi antitrust, anche al fine di considerare il ruolo dei dati e degli algoritmi nella determinazione del potere di mercato, nonché di rafforzare meccanismi quali la portabilità dei dati per ridurre le barriere al cambio di piattaforma per gli utenti. Inoltre, l'inadeguatezza delle tradizionali soglie di notifica delle operazioni di concentrazione, basate principalmente sul fatturato, rappresenta una delle principali criticità evidenziate nel dibattito europeo sui mercati digitali.

Sorveglianza biometrica (Telecomunicazioni, Tutela consumatori) — richiede moratoria sul riconoscimento facciale in spazi pubblici commerciali, DPIA obbligatoria per sistemi biometrici e notifica formale alle autorità. L'intersezione tra DPI (*Deep Packet Inspection*) e *AI analytics* rappresenta la forma più insidiosa di potenziale sorveglianza occulta.

Pricing collusivo algoritmico (Telecomunicazioni) — richiede tool di detection, sandbox regolatorio per algoritmi di pricing e revisione degli strumenti probatori antitrust. L'impossibilità strutturale di provare un "accordo" nella collusione algoritmica rende necessaria l'introduzione di nuovi approcci nelle presunzioni legali.

Violazione del diritto d'autore per addestramento (Media) — l'utilizzo di contenuti protetti da diritto d'autore nei processi di addestramento dei modelli di intelligenza artificiale solleva rilevanti questioni giuridiche, attualmente oggetto di contenzioso e dibattito a livello europeo e internazionale.

In questo contesto, sono state avanzate proposte relative all'introduzione di strumenti quali registri dei dataset di addestramento, modelli di licensing collettivo e meccanismi di compensazione per gli aventi diritto. L'eccezione per il *Text and Data Mining* prevista dall'art. 4 della Direttiva (UE) 2019/790, che consente l'uso di contenuti salvo *opt-out* dei titolari dei diritti, è oggetto di interpretazioni controverse, anche alla luce delle modalità di raccolta

massiva dei dati nei sistemi di AI generativa, e potrebbe richiedere ulteriori chiarimenti normativi e applicativi, a livello europeo e nazionale.

Manipolazione cognitiva dei vulnerabili (Media, Tutela consumatori) — i sistemi di intelligenza artificiale possono essere utilizzati per influenzare il comportamento degli utenti, con rischi particolarmente rilevanti per soggetti vulnerabili, tra cui i minori. In questo contesto, emerge l'esigenza di rafforzare le misure di contrasto alle pratiche manipolative, inclusi i cosiddetti *dark patterns*, anche alla luce dell'utilizzo crescente di tecniche algoritmiche avanzate, nonché di sviluppare strumenti efficaci di protezione dei minori, quali meccanismi di verifica dell'età e limitazioni all'uso di tecniche persuasive. L'AI Act (art. 5) introduce divieti relativi a pratiche manipolative e allo sfruttamento delle vulnerabilità, ma la loro applicazione pratica solleva questioni interpretative, in particolare con riferimento alla definizione e all'individuazione delle tecniche di manipolazione subliminale.

***Bias* discriminatori sistemici** (tutti gli ambiti) — richiede audit algoritmico e approccio *fairness by design* come requisito di compliance. È l'unico impatto che attraversa tutti e quattro gli ambiti di competenza AGCOM.

3.6.2 Le attività ex ante: prevenire, preparare, abilitare

Sul fronte della prevenzione, l'analisi restituisce un messaggio chiaro: la fase ex ante non è un mero esercizio tecnico-preparatorio, ma un atto di posizionamento istituzionale che presuppone scelte robuste e visione strategica.

Codici di condotta e co-regolamentazione rappresentano lo strumento più ricorrente nelle raccomandazioni operative, e quello a più alta flessibilità. Garantiscono adattività (possono essere aggiornati con frequenza superiore a quella legislativa), *ownership* da parte dei soggetti regolati (che partecipano alla loro definizione), e legittimità istituzionale (vengono adottati o approvati dall'Autorità). Tuttavia, devono essere progettati con una visione cross-settoriale: moltiplicare codici per ciascun comparto (uno per il customer care, uno per i media, uno per la pubblicità) potrebbe essere controproducente. Il rapporto pone la questione — lasciata aperta alla riflessione — dell'opportunità di un unico Codice AGCOM per l'uso responsabile dell'intelligenza artificiale nei settori regolati, con sezioni specifiche

modulari ma principi comuni (trasparenza, *human oversight*, non-discriminazione, protezione dei vulnerabili).

Linee guida e regolamenti tecnici devono essere adottati con urgenza nei settori più esposti: linee guida e regolamenti tecnici appaiono particolarmente opportuni nei settori più esposti. In materia di raccomandazione algoritmica, il quadro europeo già impone, per le piattaforme di grandi dimensioni, almeno un'opzione non basata sulla profilazione, ma lascia aperta la definizione di metriche più puntuali in tema di pluralismo e diversità dei contenuti. In materia di contenuti *AI-generated*, l'evoluzione applicativa del quadro europeo potrà richiedere standard più dettagliati su *labeling* e tracciabilità. Sul pricing dinamico, il diritto UE già impone trasparenza in caso di personalizzazione basata su decisioni automatizzate, mentre restano aperti possibili sviluppi su disclosure più granulari. Analogamente, nella trasparenza pubblicitaria e nella protezione dei minori, il DSA offre già una base rilevante, ma l'uso crescente di *chatbot* e assistenti AI può richiedere ulteriori specificazioni applicative.

La possibilità di prevedere autorizzazioni e notifiche emerge come strumento da adottare in futuro in casi specifici ad alto rischio: AI in servizi postali decisionali (credit scoring, gestione reclami), AI in contenuti politici (durante periodi elettorali), AI nella gestione discriminatoria del traffico dati (traffic management, zero rating), AI nella sorveglianza biometrica (riconoscimento facciale, analisi comportamentale). Il framework notificatorio deve essere snello — per non paralizzare l'innovazione o costituire un deterrente agli investimenti — ma sufficientemente chiaro sui confini non valicabili e sulle conseguenze dell'omessa notifica.

Formazione e alfabetizzazione costituiscono un presidio fondamentale, spesso trascurato, che dovrebbe operare su tre livelli distinti: (a) risorse interne dell'AGCOM, con programmi strutturati di formazione su AI Act, *risk assessment*, *audit* algoritmico, *explainability* e *detection* dei *deepfake*, rivolti sia al personale tecnico sia a quello giuridico-amministrativo; (b) operatori e stakeholder, con workshop, linee guida pratiche e FAQ su obblighi di compliance, uso responsabile dell'AI e best practices; (c) utenti finali e cittadini, con campagne di alfabetizzazione mediatica e digitale sul riconoscimento dei contenuti *AI-generated*, sui diritti digitali e sulle modalità di segnalazione di abusi o discriminazioni.

3.6.3 Le attività in itinere: vigilare mentre tutto evolve

L'AI è dinamica per definizione: i modelli apprendono, i dataset cambiano, le applicazioni evolvono. La vigilanza continua non può essere affidata a logiche esclusivamente ex post (intervento dopo la segnalazione o il danno) e richiede strumenti specifici, diversi da quelli utilizzati per la vigilanza sui servizi tradizionali.

Procedure e standard di vigilanza. È fondamentale definire uno standard minimo comune per l'audit dei sistemi AI, applicabile trasversalmente e anche in collaborazione con altre autorità. Le procedure devono abilitare non solo la registrazione delle informazioni, ma anche azioni rapide di intervento. Il rapporto suggerisce l'adozione di un protocollo di audit AI strutturato in quattro fasi: (1) notifica e raccolta informazioni; (2) analisi tecnica del sistema (documentazione, testing, simulazioni); (3) valutazione di conformità normativa (AI Act, DSA, EECC, GDPR); (4) disposizioni operative (misure correttive, obblighi di adeguamento).

Tool specialistici. Questo ambito rappresenta uno degli investimenti strategici più rilevanti per l'evoluzione delle capacità operative dell'Autorità. In particolare, l'Autorità dovrebbe dotarsi di strumenti tecnici avanzati per: il monitoraggio dei sistemi di raccomandazione algoritmica (analisi di diversità, bias e polarizzazione); il rilevamento forense di contenuti sintetici (deepfake audio, video e immagini); l'analisi di bias e discriminazioni nei sistemi automatizzati (es. scoring, pricing, moderazione dei contenuti); il tracciamento e la verifica della trasparenza dei contenuti pubblicitari, inclusi quelli generati o veicolati tramite sistemi di intelligenza artificiale; l'analisi di dinamiche di mercato emergenti, inclusi scenari di pricing algoritmico e possibili comportamenti collusivi; il testing della neutralità della rete (simulazioni di traffico, analisi QoS, rilevamento di pratiche discriminatorie). Strumenti sviluppati in ambito europeo (es. BEREC, ERGA) possono costituire una base di riferimento metodologica, mentre soluzioni già adottate da altre autorità nazionali, tra cui la Bundesnetzagentur, potrebbero essere oggetto di valutazione ai fini di un eventuale adattamento. Per ambiti specifici, quali la rilevazione dei contenuti sintetici, appaiono percorribili nel breve termine sia lo sviluppo di partnership con università e centri di ricerca

nazionali (es. Politecnico di Milano, FBK, CNR), sia l'adozione di soluzioni tecnologiche disponibili sul mercato.

Laboratori tecnici e ambienti di test. Alla luce della crescente complessità dei sistemi di intelligenza artificiale, emerge l'esigenza di rafforzare le capacità tecniche interne dell'Autorità attraverso la creazione di ambienti strutturati di analisi e sperimentazione. In tale prospettiva, appare opportuno sviluppare capacità articolate lungo due direttrici principali: un ambito dedicato ai profili AI & Media, orientato al testing di sistemi di raccomandazione, alla valutazione del pluralismo algoritmico, al rilevamento di contenuti sintetici (deepfake) e all'analisi dei sistemi di moderazione automatizzata; un ambito dedicato a AI & Regolazione, finalizzato alla simulazione di scenari di impatto, allo sviluppo di strumenti predittivi, all'analisi dei comportamenti di mercato e alla valutazione di sistemi automatizzati di pricing. Tali capacità dovrebbero essere sviluppate in coordinamento con università, enti di ricerca, altre autorità e centri di eccellenza, a livello nazionale e internazionale, in una logica di condivisione delle conoscenze e rafforzamento delle competenze.

Outsourcing tecnico-scientifico. Alcune funzioni possono essere delegate in modo controllato: audit tecnici specifici, valutazioni d'impatto su sistemi complessi, simulazioni su larga scala, stress test algoritmici. Il rapporto suggerisce la creazione di un albo accreditato di soggetti terzi, con criteri chiari di trasparenza, indipendenza (assenza di conflitti di interesse con i soggetti regolati), competenza tecnica certificata e accountability.

3.6.4 Le attività ex post: intervenire, correggere, rendere visibile

Misure correttive. Le misure correttive dovrebbero essere valutate caso per caso, secondo un approccio proporzionato, modulare e progressivo, orientato prioritariamente alla mitigazione dei rischi e all'adeguamento dei sistemi. In tale contesto, gli interventi possono includere, a seconda dei casi: la modifica dei sistemi algoritmici, il rafforzamento degli obblighi di trasparenza e comprensibilità dei criteri di funzionamento, la reintroduzione o il potenziamento della supervisione umana, nonché l'imposizione di obblighi di adeguamento entro termini definiti. Nei casi di rischio grave e imminente per gli utenti o per l'interesse

pubblico, possono essere adottate anche misure più incisive, fino alla sospensione del servizio.

Trasparenza verso l'utenza. In molti ambiti — pubblicità, customer care, raccomandazione, moderazione dei contenuti — la trasparenza è la prima e più efficace forma di tutela. Il rapporto identifica un set minimo di obblighi necessari e immediatamente implementabili: disclosure automatizzata ("stai interagendo con un sistema di intelligenza artificiale"), *label* visibile e comprensibile sui contenuti *AI-generated* ("questo contenuto è stato generato/modificato con intelligenza artificiale"), avvisi chiari quando l'interazione è con un sistema automatizzato e non con un operatore umano, e informazione proattiva sui diritti dell'utente (diritto di richiedere revisione umana, diritto di conoscere la logica della decisione, diritto di reclamo).

Reportistica pubblica. La reportistica pubblica può essere valorizzata come strumento di accountability e benchmarking, contribuendo a rafforzare la trasparenza e l'efficacia dell'azione regolatoria. In tale prospettiva, si propone l'istituzione di un Rapporto annuale AGCOM sull'intelligenza artificiale nei settori regolati, che integri dati quantitativi (es. numero di segnalazioni, audit condotti, violazioni accertate, sanzioni irrogate) e analisi qualitative (trend emergenti, casi studio, best practices). Il rapporto potrebbe inoltre includere una valutazione delle misure correttive adottate e dei relativi esiti, nonché formulare raccomandazioni sistemiche rivolte al legislatore e al Governo.

Sanzioni. L'adeguatezza del sistema sanzionatorio rappresenta un elemento essenziale per l'efficacia complessiva del framework regolatorio, in quanto incide direttamente sulla capacità di garantire un adeguato livello di deterrenza. In questo contesto, i regimi sanzionatori previsti dalla normativa nazionale di riferimento possono presentare, in alcuni casi, limiti in termini di entità delle sanzioni, soprattutto in relazione a operatori di grandi dimensioni e a modelli di business ad alta intensità di dati. Il quadro europeo introduce livelli sanzionatori significativamente più elevati: il Digital Services Act prevede sanzioni fino al 6% del fatturato globale annuo (art. 52), mentre l'AI Act stabilisce sanzioni fino al 7% del fatturato globale per le violazioni più gravi (art. 99), e fino al 3% per altre tipologie di violazione. In tale prospettiva, appare rilevante garantire un efficace coordinamento tra i diversi regimi sanzionatori, assicurando proporzionalità, coerenza e capacità deterrente.

3.6.5 Pochi strumenti, ben progettati

Dalla mappatura trasversale emerge che molti strumenti ricorrono in più ambiti. Non occorre moltiplicarli: occorre progettarli una volta sola, con cura, e renderli adattabili a tutti gli utilizzi. In particolare, tre assi appaiono imprescindibili:

1. Un **framework regolatorio modulare**, con codici di condotta, linee guida e obblighi di trasparenza armonizzati tra i settori, basato su principi comuni (trasparenza, *human oversight*, non-discriminazione, protezione dei vulnerabili) declinati in obblighi operativi specifici per ciascun contesto;
2. Una **dotazione tecnico-operativa moderna**, con tool *AI-based*, laboratori interni, partnership scientifiche e collaborazioni esterne accreditate, capace di evolvere con la tecnologia;
3. Una **capacità di reazione regolatoria integrata**, che tenga insieme trasparenza verso gli utenti, obblighi correttivi modulari, sanzioni proporzionate e deterrenti, e cooperazione istituzionale con le altre autorità nazionali ed europee.

3.7 Benchmark regolatori europei

L'AGCOM opera all'interno di un ecosistema regolatorio europeo caratterizzato da un crescente livello di coordinamento tra autorità nazionali e organismi sovranazionali. In questo contesto, l'analisi comparata delle esperienze di altri regolatori — tra cui Ofcom (Regno Unito), Arcom (Francia), Bundesnetzagentur (Germania), nonché delle reti europee BEREC ed ERGA — evidenzia alcune direttrici evolutive comuni nell'approccio alla regolazione dei sistemi di intelligenza artificiale nei settori delle comunicazioni, dei media e dei servizi digitali. In particolare, tali esperienze mostrano: un progressivo rafforzamento delle capacità tecniche interne delle autorità; lo sviluppo di ambienti di analisi e sperimentazione (lab, sandbox, unità data science); una crescente attenzione ai sistemi di raccomandazione algoritmica, ai contenuti sintetici (deepfake) e ai rischi per il pluralismo informativo; il ricorso a forme strutturate di cooperazione con università, centri di ricerca e altre autorità; l'integrazione tra strumenti regolatori tradizionali e approcci data-driven e risk-based. Tali elementi costituiscono un riferimento utile per l'evoluzione delle capacità operative dell'Autorità.

3.7.1 Ofcom (Regno Unito)

Ofcom rappresenta uno dei casi più avanzati in Europa nell'integrazione di competenze tecnologiche all'interno della funzione regolatoria. L'autorità ha progressivamente rafforzato le proprie capacità interne in ambito digitale e di intelligenza artificiale, anche attraverso la creazione di team specializzati e di strutture dedicate all'analisi delle tecnologie emergenti (Ofcom, *Approach to AI*, 2025; Ofcom Annual Plan). In particolare, Ofcom: ha sviluppato una strategia esplicita sull'uso e sulla regolazione dell'intelligenza artificiale (Ofcom, *Supporting and harnessing AI innovation safely*, 2025); dispone di competenze interne avanzate in ambito data science e AI (Ofcom Annual Report; Ofcom statements on in-house technology expertise); ha istituito ambienti di test e analisi tecnologica, tra cui un Technology Lab nell'ambito del regime di Online Safety (Ofcom, Online Safety regime implementation); collabora con istituzioni accademiche e centri di ricerca, tra cui l'Alan Turing Institute (Ofcom in collaboration with the Alan Turing Institute on safety technology taxonomy). L'approccio adottato si caratterizza per: una combinazione di promozione dell'innovazione e gestione dei rischi (Ofcom, *Approach to AI*, 2025); un progressivo rafforzamento degli strumenti di trasparenza e accountability, in linea con il Digital Services Act (Reg. UE 2022/2065); un investimento significativo nelle capacità tecniche dell'autorità. Spunti per AGCOM: l'approccio risk-based richiede un forte investimento in competenze tecniche interne; la capacità di analizzare sistemi complessi è una condizione necessaria per un enforcement efficace.

3.7.2 Arcom (Francia)

Arcom, istituita nel 2022 dalla fusione tra CSA e HADOPI, rappresenta un modello di integrazione tra competenze audiovisive, digitali e di tutela dei contenuti (Ordonnance n° 2021-580; Arcom Rapport Annuel 2024). L'autorità ha progressivamente sviluppato: attività di monitoraggio dei servizi digitali e delle piattaforme online, anche in attuazione del Digital Services Act (Arcom, reports on online platforms and intermediary services); strumenti di analisi dei contenuti e dei fenomeni informativi (Arcom, studies on media pluralism and information integrity); iniziative di cooperazione con soggetti istituzionali e operatori privati, in particolare in ambito elettorale e di contrasto alla disinformazione (Arcom,

electoral recommendations and guidance to platforms). Il modello francese si caratterizza per: una forte attenzione al pluralismo informativo e alla diversità culturale (Arcom Rapport Annuel; legacy of CSA audiovisual regulation); un approccio orientato alla trasparenza delle piattaforme, in linea con il DSA; l'utilizzo di strumenti sia regolatori sia cooperativi (dialogo con piattaforme, raccomandazioni, linee guida). Spunti per AGCOM: il monitoraggio continuo e strutturato dei fenomeni digitali è più efficace di interventi sporadici; le forme di cooperazione con le piattaforme possono essere utili, ma richiedono un solido quadro regolatorio di riferimento.

3.7.3 BNetzA (Germania)

La Bundesnetzagentur (BNetzA) ha sviluppato un approccio particolarmente avanzato nel campo della regolazione tecnica delle infrastrutture di comunicazione, con particolare riferimento alla neutralità della rete e alla tutela degli utenti (Bundesnetzagentur, net neutrality framework; Reg. UE 2015/2120). In particolare: partecipa attivamente allo sviluppo delle metodologie europee in ambito BEREC (BEREC Guidelines on Net Neutrality); utilizza strumenti tecnici per il monitoraggio della qualità del servizio e della neutralità del traffico (Bundesnetzagentur broadband measurement tools; BEREC net neutrality measurement methodologies); ha sviluppato sistemi di raccolta delle segnalazioni da parte degli utenti (Bundesnetzagentur consumer complaint and reporting systems). Il modello tedesco evidenzia l'importanza di: strumenti di misurazione indipendenti delle prestazioni e del comportamento degli operatori; integrazione tra analisi tecnica e attività di vigilanza; utilizzo di dati e segnalazioni per identificare criticità sistemiche. Spunti per AGCOM: il testing tecnico indipendente rappresenta un elemento essenziale della regolazione; le dichiarazioni degli operatori devono essere integrate da strumenti di verifica autonoma.

3.7.4 BEREC ed ERGA

A livello europeo, le reti di regolatori svolgono un ruolo crescente nello sviluppo di approcci comuni. Il BEREC ha analizzato l'impatto dell'intelligenza artificiale nel settore delle telecomunicazioni (BEREC Report BoR (23) 93), evidenziando: la crescente diffusione di soluzioni AI nelle reti e nei servizi; le implicazioni per la regolazione, in particolare in materia

di trasparenza e neutralità della rete; la necessità di sviluppare strumenti e competenze adeguate per l'audit dei sistemi. L'ERGA ha invece focalizzato l'attenzione su: impatto dei sistemi algoritmici sul pluralismo informativo; diffusione di contenuti sintetici (deepfake); tutela dei minori e degli utenti vulnerabili (ERGA reports and statements on disinformation, media pluralism and platform regulation). Entrambe le reti stanno contribuendo allo sviluppo di: metodologie condivise; linee guida e raccomandazioni; spazi di coordinamento tra autorità nazionali (BEREC Work Programme; ERGA annual reports). Spunti per AGCOM: la partecipazione attiva ai network europei è essenziale per contribuire alla definizione degli standard futuri e per rafforzare la coerenza dell'azione regolatoria a livello UE.

3.7.5 Spunti operativi per AGCOM

Dal confronto comparato emergono alcuni spunti operativi rilevanti:

Il rafforzamento delle capacità tecniche interne rappresenta un fattore abilitante essenziale: la disponibilità di competenze specialistiche (data science, ingegneria dei sistemi AI, auditing algoritmico) consente di ridurre l'asimmetria informativa rispetto ai soggetti regolati.

Il testing indipendente e automatizzato costituisce un elemento fondamentale dell'attività di vigilanza: l'integrazione tra strumenti tecnici e analisi regolatoria rafforza l'efficacia dell'enforcement.

Il monitoraggio continuativo dei sistemi digitali può risultare particolarmente efficace rispetto a interventi basati esclusivamente su verifiche periodiche, soprattutto in contesti caratterizzati da elevata dinamicità e adattività dei sistemi di intelligenza artificiale.

Le forme di cooperazione con le piattaforme risultano efficaci se inserite in un quadro regolatorio chiaro e supportate da strumenti di enforcement credibili, in grado di garantire il rispetto degli impegni assunti.

La cooperazione internazionale rappresenta un elemento imprescindibile, in considerazione della natura transfrontaliera dei fenomeni legati all'intelligenza artificiale, tra cui la diffusione di contenuti sintetici, i rischi di discriminazione algoritmica e le dinamiche di mercato digitali.

3.8 Raccordo normativo: il mosaico giuridico europeo

3.8.1 L'architrave dell'AI Act

Il Regolamento (UE) 2024/1689 (AI Act), adottato il 13 giugno 2024 e applicabile in via generale dal 2 agosto 2026, con applicazione progressiva di alcune disposizioni fino al 2 agosto 2027, costituisce il principale riferimento del framework giuridico europeo in materia di intelligenza artificiale. La sua architettura *risk-based* distingue tra pratiche vietate (art. 5), sistemi ad alto rischio (art. 6 e Allegato III), sistemi soggetti a specifici obblighi di trasparenza (art. 50) e sistemi non soggetti, in linea generale, a obblighi specifici ulteriori. Per AGCOM, la rilevanza del regolamento è potenzialmente elevata e trasversale, fermo restando che la designazione delle autorità competenti nazionali è rimessa agli Stati membri ai sensi dell'art. 70. Nel settore delle comunicazioni elettroniche, alcuni sistemi AI impiegati nella gestione di infrastrutture digitali critiche potrebbero rientrare tra i sistemi ad alto rischio, in particolare ove assumano la funzione di *safety components* ai sensi dell'Allegato III, punto 2. In tal caso troverebbero applicazione gli obblighi di *risk management*, data governance, documentazione tecnica, *record-keeping*, trasparenza e *human oversight* previsti dal Titolo III del regolamento. Per tali sistemi, il regime di registrazione previsto dall'art. 49 va letto tenendo conto delle specificità dell'Allegato III, incluso il fatto che i sistemi di cui al punto 2 sono registrati a livello nazionale. L'intersezione con il Codice europeo delle comunicazioni elettroniche è operativa: le disposizioni sulla sicurezza delle reti e dei servizi, sulla disponibilità del servizio e sull'accesso ai servizi di emergenza (artt. 108–109 EECC) pongono vincoli che devono restare compatibili con l'impiego di sistemi AI nella gestione delle reti. Analogamente, il principio di neutralità della rete interagisce con i requisiti di trasparenza e controllo dei sistemi automatizzati.

3.8.2 L'intersezione DSA-AI Act

L'AGCOM, quale Coordinatore dei servizi digitali (DSC) per l'Italia, esercita già funzioni che presentano rilevanti punti di contatto con il quadro delineato dall'AI Act. Il DSA impone alle piattaforme online molto grandi (VLOP) e ai motori di ricerca online molto grandi (VLOSE) obblighi di valutazione dei rischi sistemici (art. 34), che includono anche il funzionamento

dei sistemi algoritmici, in particolare dei sistemi di raccomandazione, e i loro possibili effetti su disinformazione, dibattito civico, processi elettorali, salute pubblica e tutela dei minori; obblighi di audit indipendenti annuali (art. 37); nonché l'obbligo di offrire almeno un'opzione di raccomandazione non basata sulla profilazione (art. 38). In questo contesto emerge un'area significativa di coordinamento tra i due regimi, utile a evitare duplicazioni e a favorire una vigilanza integrata. Tale convergenza appare particolarmente evidente su almeno tre fronti: il DSA impone la valutazione dei rischi sistemici derivanti anche dai sistemi algoritmici, mentre l'AI Act introduce obblighi di trasparenza per specifiche tipologie di contenuti sintetici, inclusi i deepfake; il DSA richiede trasparenza sui principali parametri dei sistemi di raccomandazione per gli utenti, mentre l'AI Act impone obblighi di documentazione e trasparenza nei confronti delle autorità competenti; il DSA vieta i dark patterns (art. 25), mentre l'AI Act vieta specifiche pratiche subliminali o manipolative che incidono in modo significativo sull'autonomia decisionale delle persone (art. 5, par. 1, lett. a).

3.8.3 DMA, Direttiva Copyright ed ePrivacy

l Digital Markets Act (Regolamento UE 2022/1925), pur non rientrando tra le competenze dirette dell'AGCOM, incide significativamente sull'ecosistema digitale in cui l'Autorità opera. I gatekeeper designati dalla Commissione controllano una pluralità di servizi di piattaforma di base – tra cui motori di ricerca, app store, servizi di messaggistica e piattaforme social – con effetti rilevanti sulle dinamiche concorrenziali e sulla distribuzione dei contenuti digitali. In particolare, gli obblighi di portabilità dei dati (art. 6, par. 9) e di interoperabilità dei servizi di comunicazione interpersonale indipendenti dal numero (art. 7) costituiscono riferimenti importanti per l'evoluzione del quadro regolatorio europeo in materia di apertura dei mercati digitali.

La Direttiva Copyright (UE) 2019/790, e in particolare l'art. 4 sul *text and data mining*, solleva questioni ancora aperte rispetto all'addestramento commerciale dei modelli di intelligenza artificiale su opere protette. La norma consente il *text and data mining* anche per finalità commerciali, salvo che i titolari dei diritti abbiano riservato espressamente tale utilizzo in modo appropriato, anche mediante strumenti *machine-readable* per i contenuti resi disponibili online. In questo contesto, permangono rilevanti criticità applicative, anche in

relazione alla verifica degli *opt-out* nei processi di raccolta massiva dei dati, mentre il quadro vigente non prevede un meccanismo generale di compensazione per gli aventi diritto. Ciò alimenta il dibattito sull'opportunità di ulteriori sviluppi normativi o contrattuali, anche in materia di *licensing* collettivo e trasparenza sui *dataset* di addestramento.

La Direttiva ePrivacy (2002/58/CE) resta inoltre rilevante con riferimento a tecniche di analisi del traffico e delle comunicazioni, incluso il *deep packet inspection*. L'art. 5 tutela la riservatezza delle comunicazioni e dei dati di traffico e vieta, in linea di principio, forme di intercettazione o sorveglianza da parte di soggetti diversi dagli utenti, salvo consenso o specifica base legale. Ne consegue che pratiche di *deep packet inspection* combinate con *analytics* o profilazione possono porre rilevanti criticità sotto il profilo della riservatezza delle comunicazioni, salvo i casi in cui siano strettamente necessarie alla trasmissione del servizio o altrimenti giustificate dal quadro normativo applicabile.

3.8.4 Un framework operativo per navigare la complessità

Il rapporto propone un framework operativo articolato su quattro livelli per navigare la complessità e il mosaico normativo:

- **AGCOM come naturale snodo** tra DSA, AI Act, EECC e TUSMA, valorizzando la concentrazione di competenze già presenti nell'Autorità e favorendo un enforcement coerente e non duplicativo;
- **protocolli di coordinamento con altre autorità e istituzioni competenti**, tra cui il Garante per la protezione dei dati personali (GDPR, ePrivacy, data governance), l'AGCM (concorrenza, DMA, collusione algoritmica), il Ministero della Cultura e gli organismi rappresentativi dei titolari dei diritti (copyright e licensing dei dati di addestramento), nonché Banca d'Italia e IVASS (credit scoring e assicurazioni AI);
- **tavoli tecnici intersettoriali**, ad esempio su AI & Telecomunicazioni (AGCOM, ENISA, operatori), AI & Media (AGCOM, *broadcaster*, piattaforme, *fact-checkers*, EDMO), AI & Consumatori (AGCOM, associazioni dei consumatori, Garante, AGCM);
- ***capacity building***, attraverso il rafforzamento delle competenze tecniche specialistiche interne, programmi di formazione diffusa sull'AI per gli uffici dell'Autorità, convenzioni

con università e centri di ricerca per il supporto tecnico all'audit dei sistemi e forme di collaborazione strutturata con altri regolatori europei.

3.9 Implicazioni per il quadro giuridico italiano

3.9.1 Le cinque urgenze per giuristi e policymaker

L'analisi condotta dal Comitato evidenzia un divario significativo tra la rapidità di adozione dei sistemi di intelligenza artificiale nei settori regolati e la capacità del quadro normativo vigente di presidiare i rischi emergenti. Mentre l'AI Act e il Digital Services Act forniscono un'architettura europea di riferimento, la loro efficacia dipenderà in larga misura dalla qualità dell'implementazione nazionale e dalla capacità delle autorità competenti di dotarsi di strumenti tecnici, organizzativi e di competenza adeguati a garantire un enforcement effettivo. In tale contesto, emergono cinque aree di particolare rilevanza.

Prima. La piena operatività delle autorità competenti ai sensi dell'art. 70 dell'AI Act richiede chiarezza istituzionale — in termini di attribuzione di competenze, poteri e risorse — e adeguata dotazione operativa. In questo quadro, se AgiD a ACN saranno coinvolte prevalentemente per aspetti relativi al *design* delle soluzioni, in base al DSA sarà AGCOM ad essere coinvolta per i profili di competenza sui temi relativi ai servizi.

Seconda. Il coordinamento tra i regimi sanzionatori nazionali ed europei rappresenta una questione centrale. Le sanzioni previste dal quadro nazionale possono risultare limitate rispetto alla dimensione economica degli operatori digitali, mentre i regimi europei (DSA e AI Act) introducono livelli sanzionatori significativamente più elevati. Ciò richiede meccanismi applicativi chiari e un possibile intervento di coordinamento normativo.

Terza. L'emergere di fenomeni di collusione algoritmica pone nuove sfide agli strumenti antitrust tradizionali. Il quadro attuale, fondato sui concetti di accordo e pratica concordata, potrebbe richiedere adattamenti interpretativi o evoluzioni normative per affrontare dinamiche di coordinamento tacito mediate da sistemi algoritmici.

Quarta. La tutela del diritto d'autore nell'era dell'AI generativa solleva rilevanti criticità applicative, in particolare rispetto all'uso di opere protette nei processi di addestramento. Il quadro vigente, basato sull'eccezione TDM con *opt-out*, alimenta il dibattito sull'opportunità

di ulteriori sviluppi normativi o soluzioni di mercato, anche in materia di *licensing* e trasparenza sui *dataset*.

Quinta. La protezione dei soggetti vulnerabili — tra cui minori, anziani e persone con bassa alfabetizzazione digitale — richiede un rafforzamento degli strumenti esistenti e una valutazione dell'adeguatezza delle misure attuali, anche con riferimento a pratiche manipolative, trasparenza delle interfacce e alfabetizzazione digitale.

3.9.2 La sfida della proporzionalità e dell'innovazione

Il rapporto è consapevole del rischio di una regolazione eccessivamente gravosa, che possa rallentare l'innovazione senza produrre tutele effettive. L'approccio proposto si fonda su tre principi:

– **modularità**, intesa come adozione di strumenti adattabili a diversi contesti e livelli di rischio, evitando approcci uniformi;

– **proporzionalità**, con interventi commisurati alla gravità dell'impatto e al gap normativo, e caratterizzati da gradualità nell'escalation;

– **trasversalità**, attraverso lo sviluppo di strumenti progettati in modo riutilizzabile e coerente tra ambiti diversi, al fine di evitare duplicazioni e frammentazione.

In questo quadro, la co-regolamentazione e i codici di condotta rappresentano strumenti particolarmente rilevanti, in quanto in grado di garantire flessibilità e adattamento alle specificità settoriali, purché supportati da meccanismi di enforcement credibili e da processi di revisione periodica che assicurino l'aggiornamento continuo del framework alla luce dell'evoluzione tecnologica.

4 Democrazia costituzionale e libertà dell'informazione al tempo dell'IA

Andrea Simoncini

4.1 La dimensione costituzionale della libertà di conoscenza, informazione e comunicazione

La Costituzione italiana riconosce la libertà di informazione come diritto fondamentale e inviolabile all'art. 21, laddove prevede che "tutti hanno diritto di manifestare liberamente il proprio pensiero con la parola, lo scritto e con ogni altro mezzo di diffusione" (1).

Per comprendere a pieno la portata di tale libertà, occorre considerarla in una prospettiva *tridimensionale*: la libertà di *conoscere*, intesa come la possibilità per ogni individuo di accedere alle informazioni sul mondo, sugli eventi e sulle opinioni altrui (art. 21 e 33 Cost.), la libertà di *informare*, ossia la facoltà di esprimere i propri pensieri, opinioni e conoscenze al pubblico (art. 21 Cost.) e la libertà di *comunicare*, come la possibilità di interagire e dialogare con altri soggetti senza interferenze altrui (art. 15 Cost.) (2).

Così intesa, la libertà di conoscenza, informazione e comunicazione costituisce il presupposto necessario per la realizzazione di molti altri diritti fondamentali: l'esistenza di un spazio effettivo di libertà conoscitiva, informativa e comunicativa è una condizione imprescindibile per il funzionamento dello stesso sistema democratico che, fondandosi sul metodo del libero concorso tra diverse opinioni politiche, presuppone il pluralismo dell'offerta politica e l'esistenza di cittadini realmente liberi nelle loro decisioni, perché informati e criticamente consapevoli.

Pertanto, la stessa realizzazione della libertà *morale* (art. 13 Cost.) è connessa alla libertà di manifestazione del pensiero, senza la quale l'individuo non può esercitare *pienamente* la propria libertà personale. Allo stesso modo, l'art. 21 si pone come condizione essenziale per

l'educazione e lo sviluppo del senso critico, così come per la stessa ricerca scientifica (artt. 33 e 34 Cost.).

Come si legge nell'art. 21, l'unico limite all'esercizio di tale libertà è rappresentato dal *buon costume*, che sulla base della giurisprudenza costituzionale (sin dalla Sent. 9/1965 della Corte costituzionale) è stato ricondotto alla sfera del pudore sessuale e, nel tempo, sempre piu' avendo particolare attenzione alla tutela dello sviluppo e della personalità dei minori, abbandonando l'iniziale riferimento alla cd. "morale comune" o all'etica prevalente.

Ovviamente, oltre questo limite espresso, la libertà di manifestazione del pensiero ne incontra altri, legati ad interessi costituzionalmente rilevanti (diritto alla riservatezza, all'onore e alla reputazione, alla sicurezza dello Stato, ecc.) ulteriormente individuati dal legislatore - e poi valutati dal giudice costituzionale - sulla base di un giusto bilanciamento tra interessi di rango costituzionale.

L'esistenza di limiti costituzionali alla libertà di manifestazione del pensiero differenzia il nostro modello costituzionale da quello americano, nel quale, come noto, la dottrina costituzionale sul primo emendamento e sulla *freedom of speech* non ammette, praticamente, limitazioni efficaci a tale libertà.

Un'ultima osservazione - utile a comprendere come il nostro sistema democratico costituzionale considera la libertà di informazione, soprattutto in prospettiva evolutiva - muove dalla formulazione della parte finale del primo comma dell'art. 21.

In esso si legge che "tutti hanno diritto di manifestare liberamente il proprio pensiero con la parola, lo scritto *e con ogni altro mezzo di diffusione"* (nostro il corsivo).

Inserendo l'inciso "con ogni altro mezzo di diffusione", i costituenti si sono dimostrati particolarmente *anteveggenti*, creando una norma costituzionale proiettata verso il futuro ("*future proof"*, si direbbe oggi); capace, cioè, di includere strumenti di comunicazione non ancora esistenti al momento della sua stesura: dalle trasmissioni televisive, ai social media e, come si vedrà, alle piattaforme digitali.

Questo spiega come un articolo scritto quando i mezzi di comunicazione di massa erano solo il giornale e la radio, abbia potuto continuare ad ispirare e guidare l'evoluzione legislativa e giurisprudenziale anche quando e' nata la televisione e poi nel passaggio dal monopolio pubblico al sistema radiotelevisivo integrato pubbico-privato.

Si pensi, solo per citare alcuni esempi rilevanti, alle sentt. 225 e 226 del 1974 che hanno fissato i canoni costituzionali cui ancorare la disciplina del monopolio pubblico nel contesto radiotelevisivo; e poi, alla sent. 202 del 1976 con la quale si è proceduto alla liberalizzazione della radiotelevisione a livello locale; ai moniti enunciati con la sent. 826 del 1988, cui e' seguito il varo della l. 223/1990, sulla disciplina del sistema radiotelevisivo "misto" pubblico e privato. Successivamente, le altre iniziative del legislatore, con la cd. terza legge di sistema (l. 112/2004) sino ad arrivare all'attuale Testo unico dei servizi di media audiovisivi (d.lgs. 208/2021) (3).

Analogamente, con riferimento alla libertà della comunicazione (art. 15), sempre grazie all'opera della Corte costituzionale, si è estesa tale libertà a tutti i mezzi tecnologici disponibili, includendo i canali digitali, le e-mail, le chat, le piattaforme social e, più recentemente, i servizi di messaggistica (sent. 170/2023).

In conclusione, il nostro sistema costituzionale si e' mostrato resiliente dinanzi ai cambiamenti tecnologici che in questi anni hanno completamente trasformato i mezzi della comunicazione e dell'informazione.

4.2 Lo "stato" della democrazia costituzionale nel contesto globale

Questa considerazione che possiamo derivare dal quadro costituzionale nazionale deve, però, collocarsi nello scenario globale in cui queste trasformazioni stanno avvenendo. Ed il panorama oggi e' decisamente cambiato rispetto alla fase "crescente" delle democrazie successiva alla seconda guerra mondiale: oggi la democrazia non è più la forma di organizzazione politica prevalente nel mondo.

Secondo il *Democracy Index* elaborato da Economist Intelligence Unit, nel 2024 soltanto il 45% della popolazione mondiale vive in regimi qualificabili come democratici, mentre una quota significativa — pari a circa il 39% — è sottoposta a regimi autoritari e il restante 15% si colloca in sistemi "ibridi".

Il contesto internazionale è caratterizzato da una crescente competizione tra modelli alternativi, alcuni dei quali apertamente mettono in discussione l'universalità dei diritti fondamentali e la centralità della persona: di conseguenza, anche la convinzione secondo cui

debbono esistere diritti fondamentali inviolabili da garantire contro ogni forma di potere, appare sempre meno condivisa e molto spesso contestata.

In tale contesto, la libertà di informazione assume un ruolo cruciale: da un lato, essa costituisce uno degli indicatori fondamentali per misurare la "qualità democratica" di un ordinamento; dall'altro, è essa stessa oggetto di tensioni, compressioni e ri-definizioni, anche all'interno degli ordinamenti formalmente democratici.

Per meglio comprendere la natura delle sfide contemporanee, è necessario dunque porre questa analisi in prospettiva storica. Difatti, la tutela giuridica delle libertà fondamentali – quello che chiamiamo comunemente "costituzionalismo" - non è mai stata statica, bensì si è sempre evoluta in funzione delle forme assunte dai poteri pubblici e privati capaci di interferire con tali liberta'.

Nel costituzionalismo rivoluzionario tra XVIII e XIX secolo, il potere da limitare era quello assoluto del sovrano. La risposta fu l'elaborazione del principio di legalità e dello Stato di diritto; la legge, espressione della volontà generale, divenne lo strumento attraverso cui garantire l'uguaglianza e la libertà.

Tuttavia, l'esperienza dei totalitarismi del XX secolo ha mostrato come la stessa legge potesse trasformarsi in strumento di oppressione. Il potere pubblico, anziché essere vincolato dal diritto, poteva appropriarsene e utilizzarlo per fini autoritari.

La reazione a tale crisi ha dato origine al costituzionalismo del secondo dopoguerra, caratterizzato dalla rigidità delle Costituzioni, la centralità dei diritti fondamentali, l'istituzione di Corti costituzionali e, quindi, dal controllo di costituzionalità delle leggi approvate dal Parlamento sovrano.

In quel paradigma, il potere da limitare non era più soltanto quello esecutivo, ma anche quello legislativo e la Costituzione si affermava come norma suprema, capace di vincolare tutti i poteri dello Stato.

Se proseguiamo il senso di questa linea storica, dobbiamo renderci conto che oggi stiamo assistendo *all'emersione di una nuova forma di potere*, *che non si identifica né col tradizionale potere pubblico né con quello economico in senso classico.*

La trasformazione digitale della società ha prodotto un'infrastruttura informativa globale all'interno della quale si svolgono una pluralità di attività essenziali, come la comunicazione,

l'informazione, la socializzazione, il lavoro e il dibattito civico dei cittadini. Questa infrastruttura e' oggi di fatto necessaria per il pieno dispiegarsi delle attività umane e, di conseguenza, finisce per condizionare inevitabilmente qualsiasi forma espressiva della libertà.

Le piattaforme digitali, sviluppate e gestite da grandi imprese private, esercitano così un potere che può definirsi *infrastrutturale*, in quanto condiziona l'accesso a tantissimi servizi fondamentali, e primo fra tutti: all'informazione.

E' una nuova forma di potere *cibernetico* (3-bis), che influenza la formazione della conoscenza e, dunque, anche *relazionale*, in quanto struttura le interazioni sociali.

Questo potere di dimensioni globali incide direttamente sull'esercizio dei diritti fondamentali e presenta forti tratti di opacità, ponendo sfide del tutto inedite ai sistemi democratico-costituzionali così come li conosciamo.

4.3 Le piattaforme digitali come "formazioni sociali". Libertà fondamentali e potere privato

Un nodo teorico di particolare rilevanza riguarda la qualificazione delle piattaforme digitali nell'ambito del diritto costituzionale.

L'articolo 2 della Costituzione italiana, com'è noto, riconosce e garantisce i diritti inviolabili dell'uomo, sia come individuo sia nelle "formazioni sociali ove si svolge la sua personalità". Tradizionalmente, tali formazioni sono state individuate in contesti quali la famiglia, gli enti locali, l'ambiente di lavoro, le associazioni, i partiti politici e così via.

Nell'attuale contesto, occorre domandarsi se oggi le piattaforme digitali possono essere individuate come nuove forme di organizzazione sociale, nelle quali si sviluppano relazioni/connessioni decisive per la costruzione dell'identità personale.

A questo riguardo, va ricordata la duplice dimensione delle formazioni sociali nell'articolo 2 della Costituzione, laddove esso afferma che "la Repubblica riconosce e garantisce i diritti inviolabili dell'uomo, sia come individuo sia *nelle* formazioni sociali ove si svolge la sua personalità" (nostro il corsivo).

Il fatto che, i diritti debbano essere riconosciuti e garantiti "*nelle*" formazioni, esprime questa dualità delle formazioni sociali che sono, da un lato, spazio e condizione di esercizio delle

libertà e, dall'altro, potenziale fonte di compressione e restrizione dei diritti di chi vi partecipa.

Il diritto costituzionale si è storicamente sviluppato in prevalenza all'interno di confini territoriali ben definiti, nei quali lo Stato ovvero le organizzazioni politiche sovranazionali rappresentavano (e rappresentano) l'ambito entro cui si esercita la sovranità.

La "rivoluzione digitale" ha profondamente alterato questo assetto: le grandi piattaforme esercitano un'influenza che, per ampiezza e intensità, è paragonabile - se non superiore - a quella degli Stati, ma, tuttavia, senza sottostare agli stessi vincoli costituzionali cui sono assoggettati questi ultimi.

Tale carattere induce a riflettere sulla ridefinizione del rapporto tra diritto e tecnologia, sulla necessità di adattare gli strumenti di tutela attualmente esistenti e le correlative categorie giuridiche. Ad emergere è, soprattutto, l'esigenza di sviluppare forme di regolazione internazionale e sovranazionale, sulle quali si soffermano gli altri contributi in questo Rapporto.

4.4 L'evoluzione dell'ecosistema dell'informazione e della comunicazione

Per comprendere in maniera più puntuale come sta cambiando il quadro costituzionale entro cui si svolge la libertà di informazione, occorrerà spostare la nostra attenzione sulla trasformazione subita dal cosiddetto ecosistema della comunicazione negli anni più recenti.

Se fino a metà degli anni Novanta il pluralismo informativo era principalmente garantito dall'equilibrio tra stampa cartacea, radio e televisione, oggi, le piattaforme digitali, i *social network*, i motori di ricerca online e da ultimo le chatbot di IA generativa, costituiscono i principali "mediatori" tra i cittadini e l'informazione.

In tal senso, i dati AGCOM evidenziano come l'informazione online abbia ormai sorpassato i media tradizionali come fonte principale di notizie, con profonde implicazioni sociali, culturali e politiche.

Secondo i dati più recenti dell'*Osservatorio annuale sul sistema dell'informazione dell'Autorità per le Garanzie nelle Comunicazioni* (AGCOM, 2024) un italiano su due si informa prevalentemente online e il 50,5% degli utenti social dichiara di ricevere notizie dalle

piattaforme digitali *prima* delle fonti ufficiali. Dall'Osservatorio AGCOM emerge anche come 9 italiani su 10 posseggano uno smartphone e che ben l'87% dei bambini tra 6 e 13 anni ne possieda uno.

Inoltre, la principale attività svolta sui dispositivi mobili dagli italiani riguarda la ricerca di informazioni (48,8%) e la comunicazione e social networking (36%). Tra i minori, invece, l'ordine si inverte, coi *social network* a dominare la dieta informativa, facendo emergere nuove forme di socializzazione e di esposizione precoce a contenuti digitali.

Dal "*Rapporto: I fabbisogni di alfabetizzazione mediatica e digitale*", del luglio 2025, di AGCOM emerge poi una forte diffusione (non solo) di smartphone, ma anche di tablet e di smart TV tra tutte le fasce di età, ovviamente, con significative differenze tra utenti giovani e anziani nella fruizione dei contenuti digitali, coi primi maggiormente orientati ai social e ai contenuti interattivi.

Questa segmentazione generazionale implica però la necessità di predisporre strumenti di educazione digitale differenziata, in grado di rendere i cittadini consapevoli dei flussi informativi cui sono esposti e dei meccanismi di funzionamento degli algoritmi.

Anche l'esperienza pandemica (COVID-19) ha notevolmente accelerato il passaggio alla "socializzazione online" e modificato radicalmente i comportamenti informativi, con numerose attività quotidiane che si sono spostate nel mondo digitale.

Le piattaforme, pertanto, non sono più strumenti "opzionali", ma infrastrutture non rinunciabili per chi voglia avere una vita sociale ordinaria.

Anche su tali aspetti, pertanto, le piattaforme digitali - social network, motori di ricerca e applicazioni di messaggistica - hanno assunto un ruolo di *gatekeeper* dell'informazione, mediando e selezionando i contenuti accessibili agli utenti.

Il potere informativo si concentra oggi nelle mani di grandi piattaforme private (Meta, Google, TikTok, solo per citarne alcuni) e gli algoritmi determinano i contenuti diffusi per ogni utente, configurando un fenomeno di profilazione e "filtraggio informativo" personalizzato per ogni utente (cd. di *filter bubble*). (4)

Tale fenomeno comporta evidenti rischi per il pluralismo, in quanto la diversità di opinioni è subordinata a logiche commerciali e algoritmiche. A ciò, si aggiungano anche i possibili rischi legati alla diffusione di contenuti manipolati (v. par. successivo).

Inoltre, con l'avvento delle chatbot basate su IA generativa (ChatGPT, novembre 2022) e i Large Language Models (LLM) si è aperta una nuova e ulteriore fase: gli utenti non si limitano più a cercare contenuti sui motori di ricerca, per poi essere indirizzati ai websites relativi, ma la piattaforma di IA generativa sulla base dei dati sui cui e' addestrata direttamente offre la risposta in linguaggio naturale, eliminando potenzialmente qualsiasi riferimento ad un eventuale website da consultare.

E' evidente che la generazione di contenuti informativi attraverso l'IA non solo solleva problemi di verifica e di *accountability*, ma rappresenta un attacco ulteriore e formidabile alla sostenibilità economica del giornalismo tradizionale.

A tal proposito, alcune recenti azioni legali (es. *New York Times vs. Perplexity AI*, 5 dicembre 2025) hanno provato a contestare nel concreto il conflitto tra proprietà intellettuale e generazione automatica di contenuti informativi.

Il New York Times ha affermato nella sua denuncia che Perplexity ha realizzato la "copia e distribuzione illegale su larga scala" di milioni dei suoi articoli per costruire il suo "motore di risposte" basato sull'intelligenza artificiale. La denuncia sosteneva che i "*prodotti di Perplexity sostituiscono direttamente i contenuti del giornale, compromettendone così l'attività e svalutandone il giornalismo"* e che la condotta di tale IA *"minaccia questa eredità e impedisce alla stampa libera di continuare a svolgere il suo ruolo nel sostenere una cittadinanza informata e una democrazia sana".*

Oggi, dunque, lo spazio informativo digitale non è più composto da una infinità di siti web semplicemente intermediati dalle piattaforme - *search engines* o *social media* - ma si stia rapidissimamente contraendo e concentrando attorno ai *chatbot* e ai LLM che rispondono direttamente alle esigenze informative, facendo scomparire i siti web di origine.

4.5 La disinformazione come problema costituzionale e il caso Romania come "laboratorio costituzionale"

Come accennato (v. par. precedente), uno dei problemi legati alla divulgazione di informazioni tramite strumenti di IA e di chatbot è quello relativo alla verifica e alla attendibilità delle informazioni divulgate.

La disinformazione, nella sua configurazione contemporanea, non può più essere interpretata come un fenomeno marginale o patologico dell'ecosistema informativo, essa costituisce un elemento strutturale del funzionamento delle piattaforme digitali e, conseguentemente, diviene un problema di natura costituzionale.

La libertà di informazione necessita della presenza di una pluralità di fonti, della piena accessibilità delle informazioni e della capacità critica dei destinatari.

Nel contesto digitale, tali condizioni risultano spesso profondamente alterate. La concentrazione delle fonti informative, la personalizzazione dei contenuti e l'asimmetria informativa tra piattaforme e utenti compromettono il funzionamento stesso della democrazia.

Inoltre, secondo uno studio dell'*Observatory on information and democracy* EUI (5), del dicembre 2024, le piattaforme digitali, diversamente dai media tradizionali, possono trarre un rilevante profitto dalla disinformazione.

Difatti, nel 2021, i siti web che pubblicavano disinformazione hanno generato ben 2,6 miliardi di dollari di entrate pubblicitarie a livello globale, con Meta che ha guadagnato almeno 30,3 milioni di dollari.

La monetizzazione della disinformazione solleva questioni esiziali per la libertà di informazione, in quanto trasforma un diritto fondamentale in un bene scambiato per profitto, con effetti sistemici sulla formazione dell'opinione pubblica e sulla qualità della democrazia. Si configura così un vero e proprio mercato della disinformazione, in cui la verità perde valore economico.

In tale contesto si colloca una importante decisione adottata dalla Corte costituzionale rumena nel dicembre del 2024, che costituisce uno degli esempi più rilevanti di conflitto tra tecnologie digitali, libertà di informazione e processi elettorali nel contesto costituzionale europeo (6). I giudici costituzionali di Bucarest hanno annullato l'intero procedimento per l'elezione del Presidente della Repubblica avviato il 24 novembre 2024, ordinando al Governo di fissare una nuova data per le elezioni, a causa di gravi irregolarità "per tutta la durata e in tutte le fasi" del voto, riconducibili all'uso non trasparente delle tecnologie digitali, all'impiego di intelligenza artificiale nella campagna elettorale e al finanziamento elettorale non dichiarato da uno dei candidati.

Sulla base di *Note informative* rese pubbliche dal Governo, richiamate in motivazione ma senza riportare analiticamente i fatti descritti, la Corte ha ritenuto "*che il processo elettorale per l'elezione del Presidente della Romania sia stato inficiato per tutta la sua durata e in tutte le fasi da molteplici irregolarità e violazioni della legislazione elettorale che hanno distorto la natura libera e corretta del voto espresso dai cittadini e le pari opportunità dei concorrenti elettorali, hanno influito sulla natura trasparente ed equa della campagna elettorale e non hanno rispettato le norme legali sul suo finanziamento. Tutto ciò ha avuto l'effetto convergente di non rispettare i principi essenziali delle elezioni democratiche* (…) *attraverso l'uso non trasparente e in violazione delle leggi elettorali delle tecnologie digitali e dell'intelligenza artificiale nello svolgimento della campagna elettorale, nonché attraverso il finanziamento della campagna elettorale da fonti non dichiarate, anche online".*

La Corte rumena ha agito sulla base di una sua funzione "generale" di garante della Costituzione e ha deciso così di interpretare in maniera molto estensiva la sua competenza "speciale" di giudice elettorale.

Al di là dei dubbi sul piano procedurale e competenziale, la decisione è rilevante perché muove dalla constatazione che l'esercizio delle libertà costituzionali oggi avviene nello "spazio digitale", tant' è vero che le violazioni lamentate sono tutte perpetrate per mezzo di nuove tecnologie cibernetiche. Le campagne elettorali, in particolare, si svolgono (prevalentemente) attraverso nuovi *media* digitali, rispetto ai quali occorre ricalibrare completamente i mezzi della tutela giuridica.

Con riferimento a questa vicenda, al di là delle ombre dell'intervento della Corte rumena, questa azione dimostra come, nell'epoca delle *fake news*, sia sempre più decisivo prevedere forme trasparenti di controllo *indipendente* dei processi elettorali - in particolare, nelle elezioni dirette - e che tale controllo dev'essere, laddove possibile, *preventivo*, dato che i giudizi *ex post* potrebbero alimentare fortissime tensioni sul piano interno.

La Corte rumena ha deciso di auto-assegnarsi il giudizio sul procedimento elettorale, "integrando", quanto meno, la normativa esistente ed esercitando un potere che, proprio perché non regolato, rischia di mettere in moto una serie di reazioni a catena sul piano politico difficilmente arginabili.

Va però segnalato che un sostegno fondamentale alla iniziativa dei giudici di Bucarest, è venuto dalla stessa Commissione europea, la quale il giorno prima della decisione - il 5 dicembre 2024 - ha emesso nei confronti della piattaforma digitale TikTok (ritenuta dal governo rumeno la principale responsabile della disinformazione) un ordine di conservazione dei dati relativi ai processi elettorali in corso, tra il 24 novembre 2024 e il 31 marzo 2025.

Di fatto, la Commissione con tale azione ha fortemente corroborato l'ipotesi dell'interferenza nelle elezioni rumene da parte della piattaforma cinese (manovrata, in ipotesi, da mandanti russi).

Il rilievo di questa decisione, oltre il riscontro mediatico, sta nel fatto che si sta sempre più consolidando a livello europeo quello che oggi è definibile come un vero e proprio "*digital acquis*" (7), di fatto, un corpus normativo unico al mondo relativo all'impatto della trasformazione digitale sugli strumenti di informazione, fino alla disciplina della comunicazione politica (8).

4.6 Conclusioni: verso un costituzionalismo digitale europeo

La libertà di informazione, se guardata in prospettiva diacronica, si trova oggi al centro di una tensione che investe l'intero edificio del costituzionalismo contemporaneo.

Non si tratta semplicemente di adattare categorie esistenti a nuove tecnologie, ma di affrontare una vera e propria mutazione genetica del rapporto tra potere, conoscenza e democrazia.

Il primo elemento che emerge con chiarezza è che la libertà di informazione non può più essere garantita soltanto come assenza di interferenze da parte del potere pubblico. Tale concezione, adeguata nel contesto del costituzionalismo liberale e post-bellico, risulta insufficiente dinanzi a piattaforme private di dimensione globale. In questo contesto, la minaccia alle libertà fondamentali non deriva più soltanto dai rischi legati alla censura o alla repressione, ma anche da quelli legati alla manipolazione, selezione e all'orientamento dei flussi informativi.

In secondo luogo, l'analisi dell'ecosistema digitale evidenzia come il potere informativo si sia progressivamente concentrato in capo a pochi attori globali. Le piattaforme digitali e i sistemi

di intelligenza artificiale non si limitano a facilitare l'accesso all'informazione, ma ne determinano la struttura. Questo fenomeno pone una questione cruciale, relativa alla qualificazione di tali soggetti come meri intermediari, oppure se debbano essere considerati titolari di responsabilità analoghe - seppur non identiche - a quelle tradizionalmente attribuite agli editori.

La difficoltà di rispondere a tale interrogativo riflette una tensione più profonda tra due esigenze contrapposte: da un lato, la tutela della libertà di espressione e dell'innovazione tecnologica; dall'altro, la necessità di garantire la qualità e l'affidabilità dell'informazione. In questo equilibrio instabile, il rischio è che una normativa maggiormente stringente potrebbe comprimere la libertà, mentre una regolazione insufficiente potrebbe compromettere le condizioni stesse della democrazia.

Un terzo profilo riguarda la disinformazione, che emerge come uno dei principali fattori di destabilizzazione dei sistemi democratici. Come si è visto, essa non è un fenomeno marginale e, anzi, la sua propagazione incide sulla capacità dei cittadini di formarsi opinioni autonome, libere e consapevoli.

In questo senso, la decisione della Corte costituzionale rumena del 2024 rappresenta un segnale paradigmatico: per la prima volta, la disinformazione digitale viene considerata idonea a compromettere la validità di un processo elettorale. Al di là delle indubbie criticità della pronuncia, essa evidenzia una presa d'atto cruciale, per cui la democrazia non può essere ridotta a una procedura formale, ma richiede condizioni sostanziali relative ad una corretta informazione per chi esercita diritti fondamentali.

Un elemento finale di riflessione riguarda il ruolo crescente del diritto europeo, che verrà esaminato nel dettaglio più avanti.

Gli interventi normativi dell'Unione — dal Digital Services Act al Regolamento sull'Intelligenza Artificiale — delineano un modello di regolazione che mira a coniugare innovazione e tutela dei diritti fondamentali. Tali strumenti non si limitano a disciplinare aspetti tecnici, ma introducono principi di carattere costituzionale, quali la trasparenza, la responsabilità e la gestione dei rischi sistemici.

Di fatto essi si trovano a svolgere una vera e propria supplenza costituzionale dinanzi ai fenomeni inediti e rapidissimi che abbiamo provato a delineare

Si può, dunque, parlare, con sempre maggiore consapevolezza, di un nascente *costituzionalismo digitale europeo*, inteso come insieme di norme e principi volti a governare il potere tecnologico in chiave garantista. Tuttavia, questo processo è ancora in fase di consolidamento e presenta numerose criticità; soprattutto il rischio di una eccessiva frammentazione e iperproduzione normativa che pone rilevantissime difficoltà di applicazione effettiva delle regole (la recente proposta del *Digital Omnibus Package* è segno di questa consapevolezza)

La possibilità di accedere a informazioni illimitate, la personalizzazione dei contenuti e l'interazione continua con sistemi automatizzati richiedono anche politiche attive di alfabetizzazione mediatica e digitale, come evidenziato anche dai rapporti di AGCOM di cui si è dato conto sopra.

Anche con riferimento alla tutela dei soggetti vulnerabili, in particolare dei minori, i rischi legati alla pervasività dei contenuti e alla difficoltà di controllo assumono una particolare rilevanza. La protezione dei minori rappresenta, infatti, il primo banco di prova del nascente costituzionalismo digitale europeo.

In prospettiva, la sfida principale consiste nel ripensare la forma della democrazia costituzionale alla luce delle nuove forme di potere digitale. Ciò implica, da un lato, l'elaborazione di nuove categorie giuridiche e, dall'altro, il rafforzamento degli strumenti esistenti.

In particolare, è necessario richiedere maggiore trasparenza algoritmica, maggiore responsabilità delle piattaforme, un migliore sviluppo di strumenti di educazione digitale e una maggiore cooperazione internazionale, necessaria per affrontare fenomeni globali.

In definitiva, la libertà di informazione richiede un approccio integrato, capace di tenere insieme dimensione giuridica, tecnologica ed economica. Il rischio, altrimenti, è quello di una progressiva erosione dell'autenticità della stessa democrazia, nell'era dell'intelligenza artificiale, in cui cittadini, formalmente liberi, rischiano di essere pesantemente condizionati nella formazione delle proprie opinioni dai sistemi tecnologici che utilizzano.

Note

(1) A. Pace - E. Manetti, *Art. 21. Rapporti civili. La libertà di manifestazione dle pensiero,* in *Commentario della Costituzione,* fondato da G. Branca e continuato da A. Pizzorusso, Bologna-Roma, 2006;

(2) P. Barile - E. Cheli, voce *Corrispondenza (libertà di),* in *Enc. Dir.,* vol. X, Milano, Giuffrè, 1962, pp. 743 ss.;

(3) P. Caretti – A. Cardone, *Il diritto dell'informazione e della comunicazione nell'era dell'intelligenza artificiale. Seconda edizione*, Milano, Il Mulino, 2024, pp. 120 ss.;

(3-bis) A. Simoncini *Potere cibernetico e futuro del diritto costituzionale*, in *Quaderni costituzionali* n. 1/2026

(4) E. Longo, *Dai big data alle «bolle filtro»: nuovi rischi per i sistemi democratici,* in *Percorsi* Costituzionali, 12/2019, pp. 29 ss.;

(5) Observatory on information and democracy EUI, Information Ecosystems And Troubled Democracy, A Global Synthesis Of The State Of Knowledge On News Media, AI And Data Governance, in www. observatory.informationdemocracy.org, 3 dicembre 2024;

(6) Corte costituzionale, sent. n. 32 del 6 dicembre 2024; per una prima analisi della decisione cfr. A. Simoncini, *L'annullamento delle elezioni presidenziali in Romania. Luci ed ombre di una divisiva decisione costituzionale*, in *Quaderno costituzionali*, 1/2025, pp. 234 ss.;

(7) A. Bogucki – A. Engler – C. Perarnaud e A. Renda, *The AI Act and emerging EU digital acquis. Overlaps, gaps and inconsistencies*, in www.ceps.eu, 14 settembre 2022.

(8) Da un lato, si pensi al Regolamento sulla trasparenza e sul targeting nella pubblicità politica (2024/900/CE) ed alla Dichiarazione interpretativa della Commissione di Venezia, dall'altro, si pensi al richiamato intervento della Commissione europea nei confronti di TikTok, attraverso i poteri ad essa assegnati dal Digital Services Act (2022/2065/CE), che - come detto - ha legittimato "esternamente" l'azione della Corte rumena.

5 IA generativa, disinformazione e *hate speech*: rischi sistemici e leve regolatorie per AGCOM

Giovanni Boccia Artieri

5.1 Introduzione: l'IA come ambiente discorsivo e infrastruttura cognitiva

L'intelligenza artificiale (IA) non può essere considerata un semplice insieme di strumenti tecnici esterni, ma deve essere compresa come un ambiente informazionale pervasivo che interviene sulle condizioni stesse del pensare. Opera, piuttosto, come una vera e propria infrastruttura cognitiva e un ambiente simbolico (Suchman, 2007; Floridi, 2002) capace di riconfigurare le pratiche quotidiane, le grammatiche dell'attenzione e le forme dell'immaginazione sociale secondo una logica di cognizione distribuita. In questo contesto, l'IA generativa segna una svolta decisiva che è rappresentata dalla capacità di automatizzare e scalare porzioni rilevanti del lavoro cognitivo attraverso la produzione di testi, immagini e simulazioni, intervenendo direttamente sulle condizioni della conoscenza, ridefinendo ciò che appare credibile, autorevole e rilevante.

Questi sistemi tendono a funzionare come un vero e proprio dispositivo epistemico (Crawford, 2021) che contribuisce a strutturare le condizioni entro cui il pensiero prende forma, viene riconosciuto e validato. Attraverso processi di modellizzazione statistica, i sistemi generativi trasformano archivi estesi di dati in superfici probabilistiche di significato (Bender et al., 2021), offrendo versioni normalizzate del sapere che tendono a occultare i bias e le lacune degli archivi da cui apprendono. Tale processo rende sempre più porosa la soglia tra l'esperienza vissuta e la sua rappresentazione statistica, naturalizzando un sapere di tipo correlazionale e previsionale che rischia di ridurre la complessità e la conflittualità del conoscere. Allineandosi alle regolarità statistiche dei discorsi pregressi, l'IA può agire come una "maschera epistemica", marginalizzando le voci dissonanti e le forme di sapere

situato o contro-egemonico (Boccia Artieri, 2025). Infatti, poiché i sistemi generativi operano in assenza di incarnazione ed esperienza vissuta, essi offrono una comprensione per proxy basata su archivi pregressi che riduce la complessità epistemica a favore di versioni del sapere normalizzate e apparentemente neutre, che tendono a occultare i propri bias e le proprie lacune materiali.

In questo scenario, emergono rischi sistemici che minacciano l'integrità dell'ecosistema comunicativo. Da un lato, la manipolazione informativa si manifesta attraverso contenuti sintetici come *deepfake* e *voice cloning*; dall'altro, si assiste a una degradazione discorsiva alimentata da discorsi d'odio (*hate speech*) e tossicità online. Tali fenomeni sono strettamente legati a meccanismi di amplificazione algoritmica che connettono la produzione automatizzata di contenuti con le economie dell'attenzione e le pratiche di micro-targeting. L'abitudine a ricevere risposte immediate e ben formate favorisce una relazione al sapere centrata sulla "soluzione" piuttosto che sulla problematizzazione e la verifica, trasformando il pensiero critico in una mera competenza di controllo dei risultati prodotti dalla macchina. Per l'Autorità per le Garanzie nelle Comunicazioni, l'IA rappresenta un fattore trasversale di trasformazione che attraversa diritti fondamentali e meccanismi di tutela. Il presente contributo analizza queste criticità collocandole all'incrocio tra i principali quadri regolatori europei, in particolare il Digital Services Act (DSA) e l'AI Act. In qualità di Digital Services Coordinator (DSC), AGCOM è chiamata a vigilare sulla gestione di questi rischi, superando una visione di mera supervisione tecnica per esercitare un giudizio umano situato che mantenga attivi il senso dei problemi e la pluralità del senso. L'obiettivo del capitolo è fornire una bussola operativa per rafforzare l'intervento dell'Autorità attraverso leve regolatorie – misure ex-ante, strumenti in itinere e azioni ex-post – capaci di garantire che l'innovazione tecnologica non comprima la qualità del dibattito democratico.

5.2 Dimensioni critiche: manipolazione informativa e degradazione discorsiva

L'integrazione massiva dell'IA generativa nei più diversi processi comunicativi ha amplificato minacce sistemiche preesistenti, determinando un salto qualitativo nella produzione e diffusione di contenuti problematici. Questi sistemi, infatti, non si limitano a veicolare

informazioni, ma operano come dispositivi epistemici che ristrutturano i criteri di verità e affidabilità (Boccia Artieri, 2025).

5.2.1 La manipolazione informativa e l'era del sintetico

La manipolazione informativa contemporanea si avvale di contenuti sintetici avanzati quali *deepfake*, *face swapping*, *lip-syncing* e *voice cloning* che rendono instabili i criteri di autenticità, facendo diventare più difficile distinguere tra evento e sua ricostruzione artificiale, e tra documento e simulazione verosimile. Se l'ecosistema della disinformazione era già stato descritto come un intreccio di *mis-information*, *dis-information* e *mal-information* (Wardle & Derakhshan, 2017), l'avvento della generazione automatica ad alta qualità produce un salto di scala per cui la falsificazione non è più un'eccezione di tipo "artigianale", ma un processo industrializzabile, replicabile e adattabile.

L'IA generativa permette infatti di produrre contenuti falsi estremamente verosimili a costo marginale quasi nullo, riducendo barriere tecniche e tempi di realizzazione e rendendo plausibile una produzione continua di falsi in molte varianti, calibrate su pubblici differenti. La conseguenza non è soltanto l'aumento quantitativo del materiale ingannevole, ma una trasformazione qualitativa del disordine informativo. Si consolida così un regime di *sinteticità diffusa*, nel quale immagini, audio e testi non devono più essere manipolazioni di materiali preesistenti, ma possono essere interamente generati come artefatti sintetici dotati di sufficiente coerenza visiva, narrativa o sonora da risultare socialmente plausibili. In altre parole, la manipolazione non agisce solo su ciò che circola, ma su ciò che appare probabile, e quindi credibile, nel flusso.

Un rischio intrinseco, inoltre, è rappresentato dalle cosiddette "allucinazioni", ovvero la generazione di informazioni inesatte o inventate fornite con un tono autorevole. La letteratura scientifica ha chiarito che le allucinazioni non sono un "bug" occasionale, ma un effetto strutturale dei modelli generativi, che ottimizzano la probabilità di sequenze linguistiche coerenti più che l'aderenza a un referente verificabile. I modelli, infatti, non accedono alla "verità", ma inferiscono output plausibili sulla base di correlazioni apprese nei dati (Huang et al., 2025). Questo quadro si amplia se si considera che le allucinazioni attraversano ormai più modalità (testo, immagine, audio, video) e tendono a presentarsi in

forme miste, dove micro-errori vengono inglobati in narrazioni fluenti, rendendo più difficile l'individuazione del falso da parte di utenti non specialisti. In un panorama in cui, in Italia, le fonti online sono diventate il principale mezzo di informazione (con il sorpasso della televisione nella dieta informativa, secondo l'Osservatorio AGCOM) (AGCOM, 2025), questa capacità di "scrivere come noi" e "parlare come noi" dei cosiddetti pappagalli stocastici (Bender et al., 2021) aumenta la fragilità del dibattito pubblico, non tanto perché ogni output sia falso, quanto perché la distinzione tra affidabile e non affidabile diventa un lavoro cognitivo permanente.

L'era del sintetico introduce così un doppio vettore di vulnerabilità. Il primo è *deceptive*, cioè la possibilità di fabbricare prove audiovisive convincenti (video di un candidato o di una candidata che pronuncia frasi mai dette, audio spacciati come "rubati" ma in realtà prodotti sinteticamente) e di farle circolare in tempi compatibili con la velocità dell'attenzione. Il secondo è epistemico, e ha a che fare con la crescita di un clima di indeterminatezza che erode la fiducia e normalizza l'idea che "non si sappia più cosa sia vero". Su questo punto, la ricerca in comunicazione politica ha mostrato che l'impatto dei *deepfake* non va pensato solo in termini di persuasione diretta, ma anche come produzione di incertezza e corrosione della fiducia nel *news environment*: la sintesi realistica contribuisce a un "ambiente di dubbio" che peggiora la qualità della cultura civica online e indebolisce la capacità di attribuire verità e responsabilità (Vaccari & Chadwick, 2020). Qui si innesta un effetto politico cruciale e cioè che la disponibilità sociale dell'idea "potrebbe essere un deepfake" genera il cosiddetto *liar's dividend*, cioè il vantaggio di chi può squalificare contenuti autentici etichettandoli come artificiali, sottraendosi a accountability e sanzione reputazionale (Chesney & Citron, 2019; Schiff et al., 2025).

A questa dinamica si aggiunge la specificità del *voice cloning*, che abbassa ulteriormente la soglia di accesso alla manipolazione perché l'audio è spesso percepito come più "intimo" e immediato. Infatti la voce, nella comunicazione quotidiana, funziona come marcatore identitario e come indice di presenza. Per questo l'audio sintetico ha un rendimento elevato in scenari di disinformazione e frode, e pone un problema tecnico ulteriore, quello della rilevazione. Le rassegne recenti mostrano che il riconoscimento automatico di audio *deepfake* è un campo in rapida evoluzione, ma anche intrinsecamente conflittuale, perché i

sistemi di detection sono vulnerabili a strategie di elusione e a vere e proprie forme di attacco avversario (Zhang, 2025). Sul piano sistemico, questo significa che l'"antidoto" non può essere pensato come soluzione definitiva: detection e generazione co-evolvono, e la sicurezza informativa diventa una corsa tra capacità di sintesi e capacità di verifica (Khan et al., 2025; Pei et al. 2024).

Infine, la manipolazione informativa nell'era del sintetico non riguarda solo la produzione di falsi, ma anche l'infrastruttura della fiducia (Boccia Artieri, 2025). La risposta più promettente, secondo molte linee di policy e ricerca applicata, combina strumenti di detection con strategie di *provenance*, cioè non chiedere solo "è falso?", ma "da dove viene?" e "quali trasformazioni ha subito?". In questo senso, gli standard di tracciamento e attestazione – come quelli sviluppati dalla Coalition for Content Provenance and Authenticity (C2PA) – mirano a fornire metadati verificabili sull'origine e sulle modifiche dei contenuti (C2PA Specification, v2.2). L'obiettivo è spostare il baricentro dalla caccia al falso alla costruzione di catene di autenticità leggibili e controllabili, pur sapendo che tali soluzioni non eliminano il problema poiché possono essere aggirate, non coprono tutto l'ecosistema, richiedono adozione industriale e alfabetizzazione.

Nel complesso, l'era del sintetico non produce soltanto "più disinformazione", ma riconfigura le condizioni della credibilità pubblica. Da un lato, rende scalabile la falsificazione; dall'altro, rende scalabile il dubbio, e quindi la possibilità di disinnescare prove, responsabilità e conflitti interpretativi. In un contesto in cui il consumo informativo online è strutturalmente dominante, il rischio è che la sfera pubblica scivoli verso una condizione di post-affidabilità (Boccia Artieri 2025) in cui la verifica diventa eccezione, la plausibilità sostituisce la prova e l'attenzione – già fragile – viene catturata da oggetti comunicativi ottimizzati per sembrare veri, indipendentemente dal loro statuto.

5.2.2 Degradazione discorsiva: hate speech e tossicità automatizzata

La seconda dimensione critica riguarda la degradazione discorsiva, alimentata da discorsi d'odio, inciviltà e forme di *e-bile* (Jane, 2014), espressione con cui si indicano quelle pratiche comunicative online intrise di livore, invettiva, disprezzo e aggressività verbale, che contribuiscono ad avvelenare il clima del confronto pubblico e a normalizzare registri ostili

nella comunicazione digitale. In questa prospettiva, l'odio online non va considerato come un "rumore" collaterale della comunicazione digitale, ma va pensato come una grammatica di interazione che combina denigrazione, minaccia e umiliazione, colpendo in modo sproporzionato soggetti e gruppi esposti nello spazio pubblico. La letteratura ha mostrato come l'architettura delle piattaforme intensifichi tali dinamiche attraverso specifiche affordance – anonimato relativo, asincronia, visibilità intermittente, assenza di contatto faccia-a-faccia – che favoriscono un effetto di disinibizione e abbassano il costo sociale dell'aggressione verbale (Suler, 2004). Ne deriva un ambiente in cui la conflittualità non si esprime soltanto come agonismo democratico, ma può degenerare in inciviltà sistemica e intimidazione, con effetti misurabili sulla qualità del confronto politico e sulla disponibilità a partecipare al dibattito (Papacharissi, 2004).

Gli strumenti digitali facilitano inoltre comportamenti di *swarming*, in cui utenti ordinari si coalizzano attorno a eventi o persone, rinforzando pratiche discriminatorie e rendendo l'attacco più efficace perché distribuito, reiterato e difficilmente imputabile a singoli attori (Gagliardone, 2019). In queste dinamiche l'odio assume spesso la forma di *harassment assemblage*, cioè una molteplicità di micro-atti (commenti, meme, segnalazioni, *doxxing*, citazioni ostili) che, sommati, producono un effetto di pressione e silenziamento. L'inciviltà, inoltre, non è solo un problema morale, infatti studi sperimentali mostrano che l'esposizione a discussioni incivili incide sulla valutazione della razionalità degli argomenti e può irrigidire la percezione dell'outgroup, contribuendo a polarizzare ulteriormente le identificazioni (Popan et al., 2019). In questa cornice, l'hate speech può essere inteso – nella definizione ampia adottata da molte rassegne – come comunicazione che disprezza o attacca persone o gruppi sulla base di caratteristiche come etnia, religione, nazionalità, genere o orientamento sessuale, con capacità di produrre danni sociali e politici che eccedono il singolo episodio (Castaño-Pulgarín et al., 2021).

In questo quadro, l'IA generativa non si limita ad amplificare la circolazione dei contenuti, ma interviene direttamente nella produzione di discorso e nella formazione delle opinioni attraverso forme di iperpersuasione (*hyperpersuasion*), fondate su meccanismi di allineamento adattivo e ottimizzazione empatica: i modelli tendono ad allinearsi alle preferenze dell'utente, producendo risposte che ne confermano le aspettative e ne rafforzano

i bias, fino a configurare una persuasione "invisibile" che opera per micro-aggiustamenti successivi (AGCOM, 2025). Ricerche recenti mostrano inoltre che, in contesti di dibattito online, i sistemi conversazionali possono risultare più efficaci del confronto umano nel modificare opinioni e orientamenti, soprattutto quando dispongono di informazioni personali minime sull'interlocutore: la personalizzazione aumenta in modo statisticamente significativo la capacità persuasiva in conversazioni strutturate (Salvi et al., 2025).

Il passaggio decisivo sta nel fatto che questa efficacia non è generica poiché l'argomentazione viene modellata sui tratti demografici e psicologici dell'interlocutore, producendo un'iper-personalizzazione che aggrega consenso sfruttando vulnerabilità individuali e asimmetrie cognitive, e insieme riduce l'esposizione a punti di vista dissonanti. In un ecosistema già segnato da selezione algoritmica, abitudini di consumo omofile e diete mediali polarizzate, questo tipo di interazione "su misura" può contribuire a rafforzare bolle cognitive ed *echo chambers*, riducendo l'attrito con l'alterità e rendendo meno probabile l'incontro con argomenti controintuitivi o minoritari (Fletcher & Nielsen, 2018). Ne derivano rischi sistemici per l'integrità dei processi democratici, perché la persuasione adattiva può simulare falso consenso, erodere la legittimità del dissenso deliberativo e, in modo più sottile, innescare una depoliticizzazione strisciante. L'apparente neutralità delle risposte può, infatti, sostituire il conflitto democratico con una delega cognitiva, spostando l'autonomia dell'elettore verso la semplice verifica di output generati da macchine. Qui l'IA generativa si inserisce nella degradazione discorsiva con un effetto paradossale poiché può attenuare il conflitto "in superficie" (tono più pacato, risposte educate) e, insieme, indebolire la dimensione pubblica del disaccordo come pratica democratica, normalizzando un'interazione consulenziale che non richiede confronto tra posizioni, ma ottimizzazione dell'adesione.

A tutto ciò si aggiunge una seconda linea di rischio, quella della tossicità automatizzata come produzione e riproduzione su larga scala di linguaggi ostili. Le piattaforme hanno già conosciuto l'automazione della propaganda e dell'amplificazione artificiale; oggi, con l'uso combinato di bot e modelli generativi, la produzione di commenti, insulti, insinuazioni e *dog whistles* può essere industrializzata, resa "variabile" (per eludere filtri) e adattata ai contesti locali. La nozione di *computational propaganda* descrive proprio l'assemblaggio di

piattaforme, agenti automatici, algoritmi e dati orientati alla manipolazione dell'opinione pubblica (Woolley & Howard, 2016). In questo senso, la *swarming* non è solo un problema di quantità ma diventa una strategia percettiva che trasforma la visibilità in credibilità, facendo apparire diffuso ciò che è orchestrato e spontaneo ciò che è automatizzato.

Infine, l'uso di LLM per moderazione e contrasto all'odio apre un'ulteriore tensione. Questi modelli possono essere utili nel rilevamento di linguaggio ostile, ma la ricerca evidenzia anche vulnerabilità, opacità e limiti contestuali (ad esempio nella comprensione di riferimenti impliciti, ironia, contesto comunitario), con il rischio di falsi positivi sui discorsi di rivendicazione e falsi negativi sui discorsi codificati (Roy et al., 2023; Guo et al., 2023). In altre parole, l'automazione non riguarda soltanto la produzione dell'odio, ma anche la sua "gestione" istituzionale, poiché la governance discorsiva tende a spostarsi verso classificazioni automatiche che, se non verificabili e contestualizzate, possono redistribuire asimmetricamente visibilità e silenziamento, con effetti politici non intenzionali.

Questo ecosistema resta, inoltre, alimentato da modelli di business che monetizzano l'attenzione, trattando i messaggi tossici come "esternalità negativa" della raccolta pubblicitaria (Gruppo di lavoro Odio Online, 2021), e può essere ulteriormente destabilizzato da pratiche di propaganda computazionale e *swarming*, in cui bot e sistemi automatizzati amplificano artificialmente voci marginali, trasformandone la visibilità in percezione di rilevanza politica e orientando il clima d'opinione. In sintesi, la degradazione discorsiva nell'era dell'IA non coincide solo con "più odio", ma con una riconfigurazione delle condizioni di produzione, circolazione e legittimazione del discorso pubblico in cui la tossicità diventa più scalabile, più adattiva, più difficile da attribuire, mentre la persuasione si fa più personalizzata e meno osservabile, spostando la vulnerabilità democratica dal contenuto isolato all'ecologia complessiva dell'interazione.

5.2.3 L'effetto "black box" e la sfida dell'accountability

Ad aggravare le dimensioni descritte interviene l'opacità intrinseca di una parte rilevante dei modelli contemporanei di *deep learning*. L'IA opera spesso come una "scatola nera" (*black box*), e non perché manchi qualunque informazione sul suo funzionamento, ma perché la catena che lega dati, addestramento, architettura e output resta difficilmente ricostruibile in

modo comprensibile e verificabile, soprattutto quando i sistemi sono complessi, multi-componente e continuamente aggiornati. Le rassegne recenti sull'interpretabilità mostrano con chiarezza come l'opacità sia un tratto strutturale di molti modelli deep e come le tecniche di *explainabili*ty (ŞAHiN et al., 2025) forniscano spesso spiegazioni parziali, locali e dipendenti dal contesto, utili ma non equivalenti a una vera intelligibilità del processo decisionale.

In termini di responsabilità pubblica, ciò significa che l'accountability non può essere ridotta a una spiegazione ex post del singolo output. La letteratura sull'interpretabilità insiste su un punto chiave, che non esiste un consenso univoco su cosa significhi "interpretabilità" e su come misurarla rigorosamente; il bisogno di spiegazione varia per dominio, rischi e destinatari (sviluppatori, utenti, regolatori, soggetti colpiti), e richiede quindi criteri e protocolli diversi. Parallelamente, una parte della ricerca critica avverte che affidarsi a spiegazioni post-hoc di modelli opachi può legittimare cattive pratiche nei contesti ad alta posta in gioco, cioè quando sono in gioco diritti, accesso a servizi o integrità democratica, la soluzione non è sempre "spiegare meglio", ma anche progettare sistemi più controllabili, documentati e, dove possibile, intrinsecamente interpretabili.

È qui che l'effetto *black box* diventa un problema politico e regolatorio. Se, come osserva la riflessione critica sull'IA, il potere dei sistemi algoritmici non risiede soltanto nei loro risultati, ma nelle infrastrutture materiali e nelle asimmetrie di controllo che li rendono operativi su larga scala, l'opacità diventa una forma di governo "per delega", che sposta la possibilità di comprendere e contestare le decisioni lontano dagli spazi pubblici e dai soggetti esposti alle conseguenze. In questa prospettiva, la mancanza di trasparenza rende più difficile sia tracciare l'origine delle manipolazioni (catene di produzione e diffusione di contenuti, strategie di amplificazione, strumenti di targeting) sia spiegare output discriminatori o distorsivi, perché ciò che serve non è solo un motivo "narrabile", ma evidenza verificabile su dati, metriche, obiettivi e vincoli del sistema. Per questo la ricerca propone, accanto (e talvolta prima) delle tecniche di spiegazione, un insieme di dispositivi di responsabilizzazione documentale quali *model cards* (Mitchell et al. 2019), per rendere espliciti scopi, limiti, prestazioni e rischi dei modelli, e *datasheets for datasets* (Gebru et al, 2021), per chiarire provenienza, composizione e condizioni d'uso dei dati di addestramento.

Questi strumenti non "aprono" completamente la scatola nera, ma costruiscono una base minima di tracciabilità e responsabilità che consente confronto pubblico, valutazione comparativa e contestazione informata. A livello giuridico-tecnologico, un'altra linea di lavoro fondamentale (Kroll, 2017) insiste sul fatto che l'accountability non coincide con la sola trasparenza, ma servono meccanismi che rendano possibile dimostrare conformità a standard di equità e correttezza e permettano forme di controllo indipendente anche quando parte del sistema resta proprietaria.

Dentro questo quadro si colloca la lettura sugli *AI incidents*, che includono disinformazione automatizzata e danni informativi, come questione regolatoria cruciale. In qualità di Digital Services Coordinator per l'Italia, l'Autorità è chiamata a vigilare su rischi che non sono soltanto tecnici, ma investono diritti fondamentali e integrità dei processi democratici. La difficoltà, però, è proprio la natura sistemica e opaca dei processi, poiché non basta individuare "contenuti problematici", perché il danno spesso emerge dall'insieme di scelte di progettazione (raccomandazione, ranking, frizioni, moderazione, pubblicità, pattern di engagement) che determinano che cosa diventa visibile, per chi e con quale velocità. Il Digital Services Act, del resto, spinge esplicitamente in questa direzione quando sostiene che le piattaforme di dimensione molto grande devono valutare e mitigare i rischi sistemici, tenendo conto del ruolo della progettazione dei sistemi algoritmici, inclusi i *recommender* e i sistemi di moderazione. In parallelo, la Commissione europea ha avviato un ciclo annuale di analisi del *risk landscape* che integra le valutazioni di rischio delle piattaforme, i risultati degli audit indipendenti e la reportistica di trasparenza prevista dal DSA (European Board for Digital Services, 2025; Reg. UE 2022/2065, artt. 34 e 37) che rappresenta un segnale del fatto che la governance sta tentando di spostarsi dalla reazione al singolo episodio verso una supervisione più strutturale delle architetture. La sfida, allora, è quella di superare la mera supervisione ex post per intervenire sulle architetture algoritmiche che governano la visibilità dei contenuti. Ciò implica almeno tre spostamenti: 1. dalla prova del singolo falso alla tracciabilità dei processi che lo rendono efficace; 2. dall'output "spiegato" alla valutazione documentata di dati, performance e impatti; 3. dalla trasparenza dichiarativa a forme di controllo sottoponibili a scrutinio e verifiche indipendenti, capaci di misurare l'efficacia reale delle mitigazioni e di rendere contestabili decisioni che, altrimenti, restano

schermate dietro l'opacità tecnica e proprietaria. In questa traiettoria, l'accountability non è un accessorio etico, ma una condizione di possibilità per governare il rischio informativo nell'ecosistema digitale contemporaneo.

5.3 Meccanismi di amplificazione algoritmica ed economia dell'attenzione

I rischi sistemici descritti non operano isolatamente, ma traggono la loro efficacia da un'architettura tecnologica ed economica che lega la produzione automatizzata di contenuti alle dinamiche di micro-targeting e alla gestione algoritmica dell'attenzione. In questo contesto, l'IA agisce come un moltiplicatore di scala e di precisione, trasformando il disordine informativo in un processo industriale. Una parte della letteratura ha mostrato che le piattaforme siano infrastrutture che governano la visibilità e la rilevanza attraverso sistemi opachi di ranking e raccomandazione, producendo effetti pubblici difficili da imputare e contestare (Tufekci, 2015). L'economia dell'attenzione, inoltre, incentiva la selezione di contenuti ad alto rendimento emotivo e interazionale, rendendo strutturalmente "competitivi" messaggi polarizzanti, scandalistici o tossici (Zuboff, 2023).

5.3.1 Produzione scalabile e micro-targeting psicografico

L'IA generativa permette di automatizzare porzioni rilevanti del lavoro cognitivo, rendendo la creazione di contenuti manipolatori un processo replicabile a costo marginale quasi nullo. Questa capacità produttiva si salda con le pratiche di micro-targeting, dove l'IA non si limita a selezionare il pubblico, ma interviene nella costruzione stessa del messaggio per colpire specifiche vulnerabilità psicologiche o tratti demografici dell'interlocutore. L'efficacia di questo meccanismo è visibile nell'ottimizzazione empatica. Ricerche recenti mostrano che, in contesti di dibattito online, un LLM può risultare significativamente più persuasivo del confronto umano e che la persuasione aumenta quando il sistema dispone di informazioni sociodemografiche sull'interlocutore, abilitando una forma di personalizzazione argomentativa (Salvi et al., 2025).

In parallelo, sul versante del micro-targeting psicografico, la ricerca ha evidenziato che messaggi politicamente "congruenti" con tratti di personalità possono produrre effetti

misurabili su atteggiamenti e intenzioni di voto, indicando che la personalizzazione non opera soltanto su interessi e segmenti demografici, ma può agire sulla compatibilità tra frame comunicativo e predisposizioni individuali (Matz et al., 2017). Il punto decisivo sta nel fatto che questa efficacia non è generica: l'argomentazione viene modellata sui tratti dell'interlocutore, producendo un'iper-personalizzazione che può aggregare consenso sfruttando asimmetrie cognitive e riducendo l'esposizione a punti di vista dissonanti.

Ne consegue una trasformazione politica del micro-targeting. È una tecnica di distribuzione, e insieme una tecnologia di adattamento del messaggio, nella quale la persuasione tende a diventare meno osservabile e più difficile da sottoporre a scrutinio pubblico, perché frammentata in molteplici varianti individualizzate (Bennett & Livingston, 2018). Questa è una delle criticità evidenziate anche dalla letteratura giuridico-politica sul micro-targeting, che mette a fuoco come la personalizzazione possa sostenere strategie comunicative differenti per pubblici differenti, erodendo la dimensione comune e verificabile del discorso politico (Zuiderveen Borgesius et al., 2018)

5.3.2 Sistemi di raccomandazione e feedback loops algoritmici

Il ruolo dei sistemi di raccomandazione è cruciale nel determinare la visibilità e la circolazione dei contenuti. Questi operano come veri e propri filtri cognitivi che condizionano la percezione della realtà, ristrutturando l'esperienza informativa in sequenze "assistite" di esposizione e selezione (Tufekci, 2015). Attraverso circuiti di retroazione positiva (*positive feedback loops*), ogni interazione dell'utente alimenta i meccanismi di apprendimento del sistema: più dati vengono assorbiti, maggiore è la capacità della piattaforma di trattenere l'attenzione, rafforzando progressivamente il valore economico dell'infrastruttura. La letteratura sui sistemi di raccomandazione ha mostrato che i feedback loop possono produrre effetti cumulativi tra cui amplificazione di bias di popolarità, concentrazione della visibilità, omogeneizzazione progressiva dell'offerta e difficoltà di correggere effetti sistemici con valutazioni locali o di breve periodo (Chaney et al., 2018). In questa dinamica, la personalizzazione non è soltanto una "comodità" d'uso, ma diventa un'infrastruttura che seleziona ciò che appare saliente, credibile e socialmente disponibile nel tempo dell'attenzione.

Tuttavia, questo processo può generare rischi di esclusione epistemica. La personalizzazione algoritmica può ridurre l'esposizione incidentale a contenuti dissonanti e rendere più probabile la normalizzazione di versioni parziali del sapere che circolano, con un impatto sulla qualità della dieta informativa. Inoltre, la ricerca sull'"algorithmic confounding" e sulle dinamiche di esposizione mostra che l'interazione tra preferenze, design della piattaforma e sistemi di raccomandazione rende difficile isolare causalmente gli effetti e, proprio per questo, rende la governance più complessa. Il monitoraggio di questi fenomeni è reso complesso dall'opacità delle architetture e il danno sistemico spesso non deriva da singoli contenuti illeciti, ma dall'insieme di scelte progettuali (ranking, pattern di engagement, ecc.) che determinano cosa diventa visibile e con quale velocità (AGCOM, 2025).

5.3.3 Consapevolezza algoritmica e asimmetrie di tutela

I dati dell'Autorità evidenziano una significativa asimmetria nella consapevolezza di questi meccanismi. Più della metà della popolazione italiana dai 14 anni in su (58,9%) dichiara di essere a conoscenza del ruolo degli algoritmi di raccomandazione, mentre tra gli anziani la quota scende al 35,9%; inoltre, circa il 48% è consapevole della possibilità di personalizzare la propria esperienza di fruizione tramite strumenti di curation o segnalazione (AGCOM, 2025). Questa lacuna di *algorithmic literacy* è particolarmente critica perché molti utenti sperimentano gli algoritmi come invisibili o "naturali" finché non vengono messi a confronto con alternative non filtrate. Studi qualitativi hanno mostrato che tale invisibilità riduce la capacità di attribuire causalità agli algoritmi e di attivare pratiche di controllo, generando un deficit di agency informativa (Eslami et al., 2015).

L'asimmetria di tutela cresce ulteriormente perché l'economia dell'attenzione tende a monetizzare ogni forma di coinvolgimento. In questa cornice, contenuti tossici o polarizzanti possono essere tollerati come esternalità negative quando sostengono engagement e tempo di permanenza, e ciò rende la protezione dipendente dalla capacità individuale di riconoscere e gestire i meccanismi di selezione. La reazione più comune degli utenti di fronte ai rischi, cioè evitare il canale o la piattaforma, segnala una difficoltà nell'uso di strumenti proattivi (segnalazione, verifica delle fonti, gestione delle preferenze), che restano più frequenti tra le fasce con maggiore capitale digitale.

L'intervento regolatorio deve quindi superare la logica della reazione al singolo episodio per concentrarsi sulla supervisione strutturale delle architetture. In questa direzione, il Digital Services Act prevede obblighi specifici per la gestione dei rischi sistemici, inclusi risk assessment e audit indipendenti per le piattaforme molto grandi (Reg. UE 2022/2065, artt. 34 e 37). L'obiettivo è evitare che l'ambiente digitale si trasformi in un surrogato di autorità epistemica non contestabile, riportando visibilità, verificabilità e controllo pubblico al livello in cui oggi si produce una parte crescente della realtà sociale, cioè nelle infrastrutture algoritmiche della distribuzione.

5.4 Il quadro regolatorio: intersezioni tra AI Act e Digital Services Act

Il panorama regolatorio europeo sta effettivamente compiendo un passaggio di fase: dalla reazione al singolo contenuto illecito verso una vigilanza sistemica e di architettura che prende di mira *come* le piattaforme producono visibilità, amplificazione e asimmetrie di tutela. Questa direzione è inscritta nella grammatica del DSA, che lega i rischi sistemici non solo ai contenuti, ma alle scelte di design: *recommender systems*, pubblicità, moderazione, termini d'uso e pratiche sui dati devono essere considerati nelle valutazioni di rischio. L'AI Act, in parallelo, entra in scena come regolazione "a monte" della filiera. Non disciplina in prima battuta la circolazione del discorso pubblico, ma le condizioni di progettazione, immissione sul mercato e impiego di sistemi di IA, includendo obblighi di trasparenza specifici per i contenuti sintetici e per l'interazione uomo-macchina (Reg. UE 2024/1689).

La formula della complementarità ("AI Act a monte, DSA a valle") funziona come mappa preliminare, ma tende a semplificare eccessivamente la realtà regolatoria. Nella pratica, i due dispositivi si intersecano proprio dove si gioca l'accountability, cioè lungo la catena di responsabilità tra chi sviluppa i modelli, chi li integra in servizi e chi governa l'ambiente di circolazione (provider/deployer/piattaforma). Qui non c'è una divisione netta dei compiti, perché gli stessi effetti di rischio emergono dall'assemblaggio: un modello generativo "conforme" può produrre danni sistemici quando viene innestato in un sistema di raccomandazione, targeting e moderazione; al tempo stesso una piattaforma può rispettare formalmente obblighi DSA sui contenuti e restare opaca sulle scelte di design e sulle dipendenze tecnologiche che alimentano tali rischi.

L'intersezione tra regolamenti rischia di produrre zone grigie di imputazione e una forma di responsabilità distribuita che, senza strumenti di verifica robusti, diventa facilmente "scaricabile" lungo la filiera. Ciascun attore può rivendicare conformità al proprio perimetro, mentre l'effetto complessivo resta poco governabile. In questo quadro, l'Unione europea tenta di costruire accountability soprattutto tramite *obblighi di processo* (risk assessment, audit, reporting) e *obblighi di trasparenza* (disclosure, marking/labeling). Il punto critico è che questi dispositivi operano dentro un ecosistema in cui le architetture sono in larga parte proprietarie, la misurazione dei rischi dipende spesso da dati e metriche controllati dalle piattaforme, e gli incentivi economici restano orientati alla massimizzazione dell'engagement. Ne deriva un rischio specifico: che l'accountability si traduca in una conformità prevalentemente documentale – una governance per report e procedure – che aumenta la leggibilità formale del sistema senza necessariamente incidere sulle sue logiche operative, cioè sulle condizioni tecniche ed economiche che rendono certi contenuti più visibili, più persistenti e più persuasivi.

Un indicatore concreto di questo spostamento "architettonico" è l'avvio di un **c**iclo annuale sul *risk landscape* previsto dal DSA: Commissione e Board dei Digital Services Coordinator pubblicano report ricorrenti sui rischi sistemici più prominenti e sulle misure di mitigazione adottate, nel tentativo di uscire dalla logica episodica del "caso" e costruire un orizzonte comparativo e cumulativo di sorveglianza (European Board for Digital Services, 2025). Il passaggio è rilevante perché segnala un cambio di baricentro che sottolinea come non basti più intervenire sul singolo contenuto, conta la capacità strutturale delle piattaforme di riconoscere, misurare e ridurre i rischi che esse stesse producono attraverso i propri sistemi di raccomandazione, targeting, moderazione e monetizzazione. Qui si rende però necessaria un'osservazione critica: proprio perché l'accountability viene costruita in larga parte attraverso obblighi di processo e obblighi di trasparenza, questi strumenti possono diventare dispositivi di apprendimento istituzionale e di standardizzazione progressiva delle aspettative regolatorie, ma allo stesso tempo rischiano di scivolare verso un adempimento documentale "a bassa densità" se non sono accompagnati da condizioni di verificabilità effettiva.

- La prima riguarda l'accesso ai dati e alle informazioni necessarie a rendere controllabili le affermazioni delle piattaforme. Senza accesso, il risk assessment tende a restare auto-descrittivo e a riflettere il perimetro cognitivo della piattaforma più che un'analisi indipendente del rischio.
- La seconda riguarda la qualità e l'indipendenza delle verifiche. Se l'audit non è in grado di misurare con indicatori robusti l'efficacia delle mitigazioni, può ridursi a procedura rituale, formalmente corretta e sostanzialmente debole.
- La terza condizione è la costruzione di standard condivisi (metodologici e metrici) che rendano comparabili nel tempo e tra piattaforme le valutazioni e i risultati, riducendo l'arbitrarietà delle definizioni operative di rischio e dei criteri di successo.

È proprio su questo crinale che le analisi critiche dei primi round di risk assessment hanno già evidenziato una tensione strutturale: senza un'infrastruttura stabile di accesso ai dati e senza criteri comuni, la governance rischia di produrre un'"accountability per report", ossia una rendicontazione formalmente ricca ma politicamente debole perché non incide sulle scelte di design e sugli incentivi che alimentano l'amplificazione (Knight-Georgetown Institute, 2025).

5.4.1 AGCOM come Digital Services Coordinator e la gestione dei rischi sistemici

In qualità di Digital Services Coordinator, AGCOM diventa il perno nazionale dell'implementazione del DSA, con un ruolo di raccordo tra il livello europeo della Commissione e quello della cooperazione tra coordinatori, e con responsabilità dirette nella costruzione di procedure, strumenti e capacità di enforcement. La sua funzione, in questo quadro, non coincide con una semplice attività di rimozione o sanzione su contenuti puntuali. Il compito principale è contribuire a rendere operativa la logica del DSA come regolazione dei rischi sistemici, cioè di quelle condizioni strutturali che derivano dal modo in cui le piattaforme progettano e ottimizzano ranking, raccomandazioni, pubblicità e moderazione.
Questa impostazione sposta l'attenzione dal "che cosa circola" al "come circola". Per AGCOM ciò significa valutare se le grandi piattaforme – e in particolare le VLOPs – siano in grado di identificare e mitigare rischi che riguardano il pluralismo informativo, l'integrità del dibattito

pubblico e, più in generale, l'esercizio di diritti fondamentali. Il punto sensibile non è solo la presenza di contenuti nocivi o manipolatori, ma la combinazione tra meccanismi di amplificazione, design dell'engagement e micro-targeting che può trasformare segnali marginali in percezioni di consenso, irrigidire polarizzazioni e rendere più difficile la formazione di un'opinione informata. In questo senso, la vigilanza riguarda direttamente le architetture che distribuiscono attenzione e visibilità, perché è lì che si determinano esposizione, salienza e velocità di propagazione.

L'azione del Coordinatore, inoltre, si misura con un vincolo strutturale: la capacità di controllo dipende dalla disponibilità di evidenze verificabili sulle pratiche delle piattaforme. Per questo il ruolo di AGCOM, più che sostituirsi alle piattaforme nel "governo quotidiano" del feed, consiste nel costruire condizioni di contestabilità: accesso a dati e documentazione, capacità di valutare risk assessment e mitigazioni, possibilità di verifiche indipendenti e confronto con standard metodologici condivisi. È qui che l'intersezione tra governance nazionale e governance europea diventa decisiva. La tenuta del modello DSA non dipende solo dalle norme, ma dalla capacità istituzionale di trasformare la compliance procedurale in un effettivo cambiamento delle pratiche di progettazione.

In parallelo, AGCOM si colloca in una rete di cooperazione multilivello – europea e internazionale – che riguarda sia la definizione delle priorità di rischio, sia la produzione di strumenti comuni, sia la promozione di pratiche di trasparenza e alfabetizzazione. Questa dimensione di network è essenziale perché i rischi che il DSA intende presidiare non sono confinabili su base nazionale: operano per scalabilità, replicabilità e sincronizzazione transfrontaliera. La gestione dei rischi sistemici, di conseguenza, richiede un coordinamento che tenga insieme enforcement, tutela del pluralismo e capacità di risposta rapida, evitando che l'intervento pubblico resti confinato alla dimensione reattiva e intermittente dei "casi" e diventi invece una forma stabile di supervisione sulle infrastrutture algoritmiche che organizzano lo spazio pubblico digitale.

5.4.2 Obblighi di trasparenza e accountability degli algoritmi

L'integrazione tra gli obblighi di trasparenza introdotti dall'AI Act e il regime di mitigazione dei rischi sistemici del DSA rappresenta una leva centrale per l'azione regolatoria. Nello

scenario di *sinteticità* diffusa, la manipolazione informativa tende a diventare un processo industrializzato, capace di indebolire la distinzione tra evento e sua ricostruzione sintetica, e di destabilizzare i criteri ordinari di credibilità pubblica. In questo contesto, la trasparenza non è un obiettivo "morale" o comunicativo, ma una condizione funzionale per ridurre due rischi convergenti: da un lato la proliferazione di contenuti sintetici difficili da distinguere; dall'altro l'aumento del dubbio strategico e del *liar's dividend*, cioè la possibilità di squalificare contenuti autentici come "artificiali" per sottrarsi a responsabilità e sanzione reputazionale (Schiff, Schiff & Bueno, 2025). In questa cornice, l'AI Act introduce obblighi che incidono direttamente sul problema dell'attribuzione e della riconoscibilità: i fornitori di sistemi generativi devono predisporre misure affinché i contenuti siano riconoscibili come generati o manipolati e, nei casi più critici (deepfake e simulazioni verosimili di persone, oggetti o eventi reali), sia garantita un'informazione chiara agli utenti sul carattere artificiale del materiale (Reg. UE 2024/1689).

L'orientamento è spostare una parte della tutela "a monte", riducendo l'asimmetria informativa tra chi produce e chi fruisce, e rendendo più rapida l'identificazione di contenuti sintetici in contesti ad alta esposizione (campagne elettorali, crisi, conflitti, eventi di cronaca). Affinché questi obblighi siano realmente utili all'attuazione del DSA, tuttavia, l'etichettatura non può essere intesa come semplice segnaletica visiva. Serve un'infrastruttura tecnica che renda possibile il tracciamento in ambienti di circolazione rapida e multilivello. Due famiglie di soluzioni sono particolarmente rilevanti:

a) watermarking e marcature integrate nei contenuti – con problemi noti di robustezza e persistenza nelle catene di condivisione)
b) standard di *provenance*, che allegano metadati verificabili capaci di ricostruire origine e trasformazioni del contenuto. In questa direzione, lo standard C2PA rappresenta oggi un riferimento avanzato per una catena di autenticità basata su metadati firmati e verificabili, utile soprattutto a documentare contenuti "buoni" e a costruire procedure di autenticazione nei contesti istituzionali e giornalistici.

La criticità è che *marking/labeling* e *provenance* diventano leve di governo del rischio solo se sono "azionabili" nelle architetture di visibilità. La questione non è soltanto *se* un contenuto sia etichettato, ma *che cosa produce* l'etichettatura nel sistema, e cioè se attiva frizioni, avvisi

contestuali, priorità di verifica, de-amplificazione, limiti di targeting o escalation di moderazione; e, soprattutto, se consente di anticipare le cascate di circolazione invece di inseguirle a posteriori. In assenza di questo innesto, la trasparenza rischia di ridursi a un adempimento formale che segnala l'artificialità senza ridurre l'impatto sistemico dell'amplificazione.

Il secondo asse del problema riguarda l'accountability dei modelli e dei sistemi che li incorporano. L'effetto *black box* dei sistemi di *deep learning* rende spesso impossibile ricostruire in modo pienamente intelligibile la catena che porta dall'addestramento all'output, soprattutto in contesti di aggiornamento continuo e integrazione multi-componente. In questo scenario, l'accountability deve configurarsi come responsabilizzazione documentale. Non occorrerebbe quindi "aprire" completamente il sistema, ma produrre evidenze controllabili su scopi, limiti, prestazioni e rischi, e sulle condizioni d'uso che rendono l'output accettabile o problematico. È qui che strumenti come le *model cards* (Mitchell et al., 2019) e i *datasheets for datasets* (Gebru et al., 2021) assumono una funzione strategica: riducono la distanza tra dichiarazioni di conformità e verificabilità effettiva, e rendono possibili audit e contestazioni informate da parte di autorità e ricerca indipendente, anche quando parte dell'infrastruttura resta proprietaria.

In prospettiva operativa, l'obiettivo non è accumulare documentazione, ma trasformare la trasparenza in una condizione di contestabilità. Per l'Autorità ciò implica orientare enforcement e interlocuzione regolatoria verso alcune verifiche ricorrenti chiedendosi quali meccanismi di marking e disclosure sono effettivamente implementati (e con quale copertura sulle diverse "superfici" della piattaforma), quanto tali marcature resistano a download, re-upload, compressione e ricondivisione, e se siano integrate in procedure che incidono sul ranking e sui recommender. In parallelo, occorre verificare se e come vengono adottati standard interoperabili di *provenance* (come C2PA), con quali strumenti di verifica accessibili e con quali percorsi di contestazione quando label e provenance risultino assenti o incoerenti. Sul piano dell'accountability dei modelli, diventa essenziale accertare la disponibilità di documentazione anche quando il modello è di terza parte e integrato come servizio, e valutare le metriche rilevanti per il rischio pubblico – tassi di errore, falsi positivi e falsi negativi nelle procedure di etichettatura o detection, gestione dei casi limite (satira,

parodia, giornalismo, pubblico interesse). Solo attraverso questa infrastruttura di verificabilità, la trasparenza smette di essere un segnale simbolico e diventa una leva per governare i meccanismi di amplificazione che organizzano lo spazio pubblico digitale, riducendo l'asimmetria tra compliance formale e controllo sostanziale.

5.4.3 Tutela degli utenti vulnerabili e la protezione dei minori

Un focus prioritario dell'azione regolatoria riguarda la protezione dei minori e dei soggetti vulnerabili dall'esposizione a contenuti manipolativi, inappropriati e da architetture di ingaggio che possono produrre danni cumulativi. Il punto non è soltanto "rimuovere il contenuto nocivo", ma presidiare le condizioni tecniche che lo rendono **r**aggiungibile, persistente e ripetibile – catene di raccomandazione, profilazione, frizioni assenti, dinamiche di ricompensa e interazione che trasformano singole occorrenze in abitudini di consumo. È precisamente questo il senso dell'Art. 28 del DSA, che impone alle piattaforme accessibili alle minori misure adeguate per garantire un elevato livello di privacy, sicurezza e protezione nello sviluppo psico-fisico, e che la Commissione ha ulteriormente tradotto in linee guida operative per l'enforcement (Reg. UE 2022/2065, art. 28).

Dentro questa cornice si colloca anche la questione – ancora in definizione, ma già rilevante sul piano preventivo – del cosiddetto "debito cognitivo". L'ipotesi che l'uso intensivo e non mediato di assistenti basati su IA generativa possa incoraggiare forme di *cognitive offloading* e ridurre l'ingaggio cognitivo in compiti di apprendimento, con effetti su memoria e consolidamento. Alcune evidenze sperimentali – finora raccolte soprattutto su adulti in contesti controllati – suggeriscono che l'impiego di assistenti generativi in compiti di produzione testuale possa incidere su *come* si svolge il lavoro cognitivo mentre si scrive. In particolare, in questi studi l'assistenza del modello tende a spostare una parte dell'attività dall'elaborazione interna (pianificazione, formulazione, revisione) verso una modalità più centrata su selezione, rifinitura e validazione di output proposti dalla macchina. Questo cambiamento si accompagna a differenze nei pattern di attivazione misurati durante il compito e, sul piano del prodotto, a una maggiore similarità tra testi: quando la generazione automatica fornisce strutture, formulazioni e soluzioni "pronte", gli elaborati finiscono più facilmente per convergere su lessico, ritmo argomentativo e scelte retoriche ricorrenti,

riducendo la variabilità idiosincratica che deriva dal percorso di scrittura individuale (Kosmyna et al., 2025). Proprio perché si tratta di risultati preliminari e non centrati specificamente sui minori, è opportuno trattarli come segnale di rischio, ma la loro importanza sta nell'indicare una direzione: la dipendenza da strumenti "frictionless" può rendere più frequente la delega dell'elaborazione, un effetto coerente con una letteratura più consolidata sul *cognitive offloading* e sulla memoria transattiva nell'ecosistema digitale (Risko & Gilbert, 2016).

A questa dimensione si aggiunge un altro possibile effetto cumulativo, con implicazioni specifiche per adolescenti e giovani, e cioè la tendenza alla standardizzazione linguistica e alla convergenza di stili e formule discorsive quando i modelli diventano un medium di massa. Anche qui, le evidenze disponibili sono in parte preliminari e vanno lette criticamente. Tuttavia, la ricerca recente discute la possibilità di un feedback culturale fra output dei modelli e produzione umana, con ricadute sulla varietà linguistica e sui registri di espressione (Yakura et al., 2024/2025). L'idea di "feedback culturale" è che i modelli generativi non si limitano a riprodurre linguaggi esistenti, ma possono contribuire a stabilizzare e diffondere certe forme espressive, perché entrano in modo ripetuto e quotidiano nei processi di scrittura, riscrittura e conversazione. Se milioni di utenti usano gli stessi strumenti per formulare mail, compiti, post, caption, commenti o discorsi, gli output del modello diventano una sorta di "linguaggio ausiliario" che offre soluzioni standard quali strutture argomentative, formule di cortesia, pattern narrativi, lessico ad alta probabilità. Nel tempo, questi pattern possono rientrare nel circuito sociale come modelli imitabili, creando un anello di retroazione. In pratica le persone assorbono stilistiche e scelte linguistiche suggerite dall'IA, le reimmettono nelle piattaforme e nelle pratiche comunicative, e questo materiale – direttamente o indirettamente – alimenta ulteriori addestramenti, fine-tuning o preferenze di mercato, rafforzando ulteriormente le stesse forme.

Le ricadute possibili riguardano due aspetti. Il primo è la varietà linguistica, cioè quando prevalgono soluzioni "medie" e altamente compatibili, può ridursi la presenza di idioletti, deviazioni creative, registri minoritari o locali, e in generale della differenza stilistica prodotta da vincoli situati (contesto sociale, appartenenza, oralità, sperimentazione). Il secondo è il regime dei registri di espressione, e in tal senso i modelli tendono a privilegiare

toni e cornici discorsive "sicure" (neutre, concilianti, razionalizzate, spesso con una retorica di equilibrio), e ciò può spingere parte della produzione umana verso un linguaggio più uniforme, meno marcato e più prevedibile. In questa prospettiva, il rischio non è una uniformazione totale, ma una pressione statistica verso forme espressive dominanti che, a parità di condizioni, diventano più facili, più disponibili e più spesso riutilizzate, con un effetto cumulativo sul paesaggio linguistico. Per la tutela dei minori, questi segnali contano non perché "l'IA rende tutti uguali" in senso deterministico, ma perché mostrano come l'ambiente conversazionale possa diventare una *grammatica pronta all'uso* che abbassa la soglia di problematizzazione e spinge verso forme di delega epistemica.

In questo quadro, è importante evitare la retorica dei minori come "nativi digitali" automaticamente competenti. La letteratura sulla diseguaglianza delle competenze online mostra che, anche tra giovani con accesso e familiarità elevati, le abilità variano in modo sistematico e sono correlate a capitale culturale e condizioni sociali. Questo punto è decisivo perché l'opacità algoritmica e la persuasione adattiva colpiscono con maggiore forza chi non dispone di strumenti per riconoscerle.

Sul versante dei contenuti e dell'amplificazione, il rischio minori è strettamente legato ai *rabbit holes*, cioè percorsi di raccomandazione che, a partire da segnali minimi di interesse o vulnerabilità, possono aumentare l'esposizione a contenuti dannosi (disturbi alimentari, autolesionismo, violenza, pornografia, estremismi). Su alcuni di questi ambiti – ad esempio *disordered eating* e contenuti pro-ED – la letteratura ha discusso in modo esplicito la "logica del rabbit hole" come effetto dell'interazione fra affordance, contenuti aspirazionali e raccomandazione (Harriger et al., 2022; e, in chiave di rischio e policy, anche Amnesty International, 2023). In parallelo, lavori interdisciplinari su algoritmi e salute mentale adolescenziale hanno messo in relazione la spinta all'engagement, l'esposizione ripetuta e i danni potenziali, sollecitando misure di policy che incidano sul design e non solo sulla moderation "a valle".

Alcuni interventi recenti su contenuti legati a "sfide social" pericolose – tra cui il caso della cosiddetta "cicatrice francese" – mostrano che, in situazioni di urgenza, è possibile ridurre rapidamente esposizione e circolazione. Il nodo, nel quadro DSA/AI Act, è rendere questa capacità meno episodica e più strutturale, quindi non limitarsi alla rimozione emergenziale,

ma imporre obblighi verificabili su design e distribuzione, perché il danno nasce soprattutto dalla ripetizione e dall'amplificazione. In assenza di evidenze misurabili sull'efficacia delle mitigazioni, l'intervento rischia di restare una gestione dell'eccezione, mentre l'architettura che produce il rischio continua a operare invariata.

Un ulteriore vettore di vulnerabilità riguarda l'iperpersuasione, cioè il fatto che sistemi conversazionali siano capaci di adattare tono e argomentazione a tratti dell'utente, con potenziale superiore al confronto umano in contesti controllati. Per i minori, la criticità non è tanto l'"opinione politica" in sé, quanto l'assuefazione alla risposta fluida e rassicurante come surrogato di competenza, e la normalizzazione di una relazione in cui la macchina diventa interlocutore autorevole senza le frizioni del confronto educativo. La tutela passa quindi anche da una misura di principio: preservare le condizioni del giudizio autonomo, evitando che l'ambiente conversazionale sostituisca le pratiche di apprendimento con una delega epistemica sistematica.

Da qui discende una linea di prevenzione che integra alfabetizzazione, mediazione educativa e vigilanza istituzionale. La prevenzione del rischio di "debito cognitivo" nei minori – inteso come indebolimento progressivo delle facoltà di elaborazione, memoria e attenzione quando l'assistenza generativa diventa una scorciatoia sistematica – non può essere affidata a un'unica misura. Richiede, piuttosto, un insieme di leve che agiscano sulle competenze, sulle pratiche quotidiane e sulle condizioni ambientali della fruizione.

1. **Prompt culture come pratica riflessiva.** Il prompting andrebbe insegnato non come tecnica di efficienza, ma come chiarificazione del pensiero: esplicitare il problema, definire contesto e criteri di qualità, mantenendo attiva la distinzione fra fluidità linguistica e affidabilità del contenuto.
2. **Alfabetizzazione algoritmica e curation attiva.** Occorre rendere intelligibili i filtri: come *recommender* e *ranking* orientano la percezione della realtà e possono innescare *rabbit holes*. A questa consapevolezza va affiancata una competenza pratica nell'uso degli strumenti di personalizzazione e segnalazione, per passare da fruizione passiva a gestione consapevole del feed.

3. **Mediazione genitoriale abilitante.** La sola restrizione (tempo/divieti) è spesso insufficiente; le evidenze comparate indicano l'efficacia di strategie di accompagnamento e *co-using*, che trasformano l'interazione in un'occasione di sviluppo di criteri critici e di competenze.
4. **Alleanza tra scuola e istituzioni.** Servono curricoli verticali di cittadinanza digitale e alfabetizzazione algoritmica come infrastruttura pubblica stabile; in parallelo, la vigilanza istituzionale deve assicurare che le misure previste dal DSA per i minori siano effettive e verificabili, soprattutto quando il rischio deriva dall'architettura di raccomandazione e ingaggio.

In sintesi, proteggere minori e vulnerabili significa evitare che il lavoro interpretativo venga delegato integralmente alla macchina, preservando il giudizio situato e la capacità di problematizzare le risposte, dentro un ambiente progettato per ridurre esposizione, reiterazione e vulnerabilità persuasiva.

5.5 Proposta operativa: leve d'intervento per AGCOM

Questa sezione traduce il quadro analitico in una bussola operativa per l'Autorità nel ruolo di Digital Services Coordinator. L'obiettivo non è inseguire singoli episodi, ma rendere misurabili e contestabili le architetture che producono rischio: sistemi di raccomandazione, ranking, advertising, moderazione e design dell'engagement. In questa prospettiva, AI Act e DSA devono essere trattati come un unico dispositivo di governo. Laddove il primo rende più stringenti gli obblighi di trasparenza e tracciabilità "a monte", il secondo fornisce la cornice per valutare e mitigare i rischi "a valle" con strumenti di supervisione sistemica.

5.5.1 Misure ex-ante: costruire l'infrastruttura della fiducia

Le misure preventive servono a ridurre l'asimmetria informativa tra piattaforme, utenti e autorità e a impedire che la "sinteticità diffusa" diventi un moltiplicatore automatico di disordine e sfiducia.

a) **Trasparenza algoritmica "azionabile" (non dichiarativa).** Definire requisiti minimi per rendere intelligibili e verificabili i criteri di ranking e raccomandazione: non tanto "spiegazioni" generiche, quanto informazioni operative su obiettivi ottimizzati, segnali d'ingresso (features), frizioni, policy di de-amplificazione, e condizioni di escalation. La trasparenza deve essere concepita come condizione di contestabilità, perciò ciò che conta è poter verificare, non solo leggere.
b) **Responsabilizzazione documentale su modelli e dati.** Richiedere e standardizzare, nei perimetri applicabili, pratiche di documentazione (model cards e datasheets), con particolare attenzione a: limiti, bias noti, performance in condizioni avverse, casi limite (satira, pubblico interesse, minori), procedure di aggiornamento. Il punto è costruire un set di evidenze comparabili che renda praticabile audit e controllo indipendente.
c) **Labelling e *provenance* dei contenuti sintetici come infrastruttura, non come "bollino".** Promuovere l'implementazione di *marking/labeling* per deepfake e contenuti sintetici con criteri di robustezza (persistenza in re-upload/compressione), e valorizzare standard interoperabili di provenance (es. C2PA) per ricostruire catene di autenticità. La condizione di efficacia è l'innesto nella distribuzione: l'etichetta deve attivare misure (frizioni, priorità di verifica, limitazioni del targeting, de-amplificazione), altrimenti resta segnaletica priva di impatto.
d) **AI literacy come mindset critico e prevenzione della delega cognitiva.** Sostenere programmi di alfabetizzazione che non riducano l'IA a "competenza d'uso", ma insegnino criteri di valutazione: distinzione tra fluidità e affidabilità, verifica delle fonti, riconoscimento di pattern persuasivi, comprensione dei filtri algoritmici e pratiche di curation attiva. La posta in gioco è evitare che l'interazione con l'IA produca delega epistemica e vulnerabilità persuasiva.

Per evitare che la trasparenza resti dichiarativa, occorre definire quali evidenze minime sono considerate sufficienti: schema standard di reporting su obiettivi ottimizzati, segnali/feature, frizioni, policy di de-amplificazione ed escalation, con formato comparabile nel tempo e tra piattaforme (*minimum evidence + standardized reporting*). Per labeling/provenance, la prova non è l'adozione formale dello standard, ma la sua efficacia lungo la filiera: label retention

(sopravvivenza dell'etichetta dopo download, re-upload, compressione, ricondivisione) e impatti osservabili sulla distribuzione (frizioni attivate, priorità di verifica, limitazioni del targeting, de-amplificazione).

5.5.2 Strumenti in itinere: vigilanza, dati e cooperazione sistemica

La vigilanza continua serve a misurare se le mitigazioni funzionano nel tempo e a intercettare mutazioni rapide quali nuovi formati sintetici, nuove tattiche di amplificazione, spostamenti di pubblico, ecc.

a) **Monitoraggio dei rischi sistemici e audit sulle architetture di visibilità.** Rafforzare capacità di analisi su recommender e ranking: non solo enforcement sul contenuto, ma verifica di metriche di esposizione, reach e velocità di propagazione per classi di rischio (disinformazione sintetica, tossicità, "rabbit holes" su minori). Gli audit devono essere orientati agli esiti: cosa cambia in termini di distribuzione dopo una misura di mitigazione.
b) **Data access e ricerca indipendente (DSA art. 40).** Rendere operativo l'accesso ai dati per ricercatori qualificati e laboratori indipendenti, con protocolli chiari (privacy-by-design, minimizzazione, ambienti sicuri). Senza accesso, i risk assessment restano auto-descrittivi; con accesso, diventano verificabili e comparabili.
c) **Cooperazione inter-Autorità e coordinamento europeo.** Costruire un circuito stabile con le altre autorità rilevanti (privacy, concorrenza, tutela consumatori) e con i network europei dei DSC: i rischi sono transfrontalieri e spesso "multi-regime" (dati + advertising + design + contenuti). La cooperazione deve produrre anche standard comuni di valutazione.

Gli audit devono essere orientati agli esiti, non ai soli processi: le metriche chiave includono exposure (quota di utenti esposti), reach e velocity (velocità di propagazione), reiteration (ripubblicazioni/re-upload) e, dove pertinente, profondità dei percorsi di raccomandazione (*rabbit-hole depth*). Le mitigazioni vanno giudicate su variazioni misurabili di questi indicatori prima/dopo l'intervento.

5.5.3 Azioni ex-post: accountability, rimedi e protocolli di risposta

Le misure a valle devono evitare che la violazione diventi "costo di esercizio" e, soprattutto, devono produrre apprendimento regolatorio.

a) **Enforcement orientato ai modelli organizzativi.** In caso di violazioni sistematiche (trasparenza insufficiente, mitigazioni inefficaci, fallimento persistente nella protezione dei minori), attivare misure correttive e sanzionatorie che incidano su processi e design: requisiti di frizione, restrizioni del targeting, modifiche dei recommender, obblighi di reporting rafforzato. L'obiettivo è cambiare le condizioni che producono rischio, non inseguire il singolo contenuto.
b) **Protocollo incident response e osservatorio sugli AI incidents.** Istituire un sistema di monitoraggio e classificazione degli incidenti (disinformazione sintetica ad alta viralità, errori di labeling, allucinazioni in contesti sensibili, escalation tossiche automatizzate), con procedure di risposta rapida e canali di smentita/controbilanciamento. L'archivio degli incidenti deve alimentare l'aggiornamento periodico delle linee guida ex-ante: chiudere il ciclo della governance adattiva.

Per incident response, la prova di efficacia è duplice: (i) **tempi di risposta** (detection–mitigation–comunicazione) e (ii) riduzione documentata di **reach/propagazione** dell'evento (riduzione della curva di diffusione, abbattimento della velocità, contenimento della reiterazione). Gli incidenti devono alimentare un aggiornamento ciclico delle misure ex-ante, chiudendo il circuito della governance adattiva.

Fase	Leva	Cosa richiede (in pratica)	Evidenza/output atteso
Ex-ante (prevenzione)	**Trasparenza "azionabile"**	Requisiti minimi verificabili su obiettivi ottimizzati, segnali/feature, frizioni, policy di de-amplificazione, criteri di escalation	Schema standard di trasparenza + reporting comparabile
	Documentazione obbligatoria	Model cards/datasheets su limiti, bias, performance in condizioni avverse, casi limite (minori, pubblico interesse)	Documentazione verificabile per audit e contestazione

	Contenuti sintetici	Labeling/marking + disclosure (AI Act) con robustezza (re-upload/compressione) e integrazione nella logica di distribuzione	Etichette persistenti + attivazione di frizioni/limitazioni/de-amplificazione
	Provenance interoperabile (C2PA)	Adozione di standard di provenance e metadati verificabili lungo la filiera	Catena di autenticità consultabile e verificabile
	AI literacy	Distinzione fluidità/affidabilità; pratiche di verifica; consapevolezza dei filtri; curation attiva	Programmi e materiali per competenze critiche e riduzione delega cognitiva

In itinere (vigilanza continua)	**Audit orientati agli esiti**	Misura di exposure/reach/ velocity; profondità rabbit-hole; reiterazione; impatto delle mitigazioni	KPI e report di efficacia delle mitigazioni (non solo compliance)
	Data access (DSA art. 40)	Protocollo nazionale per accesso alla ricerca indipendente (privacy-by-design, ambienti sicuri)	Accesso dati tracciabile + studi riproducibili/verifiche esterne
	Coordinamento multi-autorità + rete DSC	Standard comuni, scambio informativo, enforcement armonizzato	Linee condivise e azioni coordinate su rischi transfrontalieri

Ex-post (rimedi)	**Enforcement "design-based"**	Rimedi su frizioni, limiti di targeting, de-amplificazione, modifiche recommender; non solo rimozione	Misure correttive cogenti che incidono sull'architettura di distribuzione
	Osservatorio AI incidents + risposta rapida	Tassonomia incidenti (deepfake virali, errori labeling, escalation tossiche) + protocolli di risposta + aggiornamento ciclico linee guida	Incident response operativo + apprendimento regolatorio (update ex-ante)

Tab. 1 Matrice operativa delle leve di intervento: ex-ante / in itinere / ex-post

5.6 Conclusioni: infrastruttura della fiducia e governance adattiva

L'IA generativa non aggiunge semplicemente nuovi contenuti al flusso informativo, ma riconfigura le condizioni di produzione, circolazione e credibilità del discorso pubblico, intrecciando automazione della scrittura e dell'audiovisivo, micro-targeting e architetture di raccomandazione che trasformano l'attenzione in valore. In questo scenario, la vulnerabilità democratica non coincide solo con l'aumento di falsi, tossicità o inciviltà, ma con il modo in cui tali fenomeni vengono resi scalabili, persistenti e adattivi, fino a incidere su pluralismo,

attribuzione di autenticità, legittimità del dissenso e tutela dei soggetti vulnerabili. Il punto chiave emerso lungo l'analisi è che il rischio è spesso sistemico: nasce dall'assemblaggio tra modelli generativi e infrastrutture di distribuzione, più che dalla singola violazione.

Da questa diagnosi discende un principio di metodo. L'intervento regolatorio efficace non può limitarsi alla rimozione ex post o alla trasparenza dichiarativa, ma deve orientarsi alla contestabilità delle architetture. L'intersezione tra AI Act e DSA rende disponibile un set di leve complementari quali obblighi di marking/labeling e tracciabilità per i contenuti sintetici; valutazioni e mitigazioni dei rischi sistemici; audit e accesso ai dati; rimedi capaci di incidere su ranking, raccomandazione, targeting e frizioni. La sfida è far sì che questi strumenti producano effetti misurabili sulla distribuzione, riducendo esposizione, reiterazione e velocità di propagazione dei contenuti a rischio, e impedendo che l'accountability scivoli in una compliance meramente documentale. In questa prospettiva, la "bussola operativa" proposta mira a consolidare un circuito di governance adattiva: misure ex-ante che costruiscano un'infrastruttura della fiducia, vigilanza in itinere che renda verificabili rischi e mitigazioni, e azioni ex post che trasformino gli incidenti in apprendimento istituzionale. Il risultato atteso non è una sfera pubblica "sterilizzata", ma un ambiente digitale in cui la libertà di espressione e il pluralismo siano compatibili con la protezione dei minori e degli utenti vulnerabili e con l'integrità del dibattito democratico. La posta in gioco, in ultima analisi, è mantenere governabile l'ecosistema comunicativo in un'epoca in cui la persuasione può diventare invisibile e la credibilità può essere prodotta, e non solo verificata, dalle infrastrutture che organizzano la visibilità.

Bibliografia

AGCOM (2025). *Intelligenza artificiale. Rapporto 2025* (documento interno / bozza di lavoro). Roma: Autorità per le Garanzie nelle Comunicazioni.

Amnesty International (2023). *Dragged into the Rabbit Hole: New Evidence of TikTok's Risks to Children's Mental Health* (report). London: Amnesty International.

https://www.amnestyusa.org/reports/dragged-into-the-rabbit-hole-new-evidence-of-tiktoks-risks-to-childrens-mental-health/

Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). On the dangers of stochastic parrots: Can language models be too big? In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency (FAccT '21)* (pp. 610–623). ACM.

Bennett, W. L., & Livingston, S. (2018). The disinformation order: Disruptive communication and the decline of democratic institutions. *European journal of communication*, *33*(2), 122-139.

Boccia Artieri, G. (2025). *Sfiduciati. Democrazia e disordine comunicativo nella società esposta*. Milano: Feltrinelli.

C2PA (Coalition for Content Provenance and Authenticity) (2025). *C2PA Specifications Overview: Latest Version 2.2 (released May 2025).* https://c2pa.wiki/specifications/

Castaño-Pulgarín, S. A., Suárez-Betancur, N., Vega, L. M. T., & López, H. M. H. (2021). Internet, social media and online hate speech: Systematic review. *Aggression and Violent Behavior, 58*, 101608.

Chaney, A. J. B., Stewart, B. M., & Engelhardt, B. E. (2018). How algorithmic confounding in recommendation systems increases homogeneity and decreases utility. In *Proceedings of the 12th ACM Conference on Recommender Systems (RecSys '18)* (pp. 224–232). ACM.

Chesney, R., & Citron, D. K. (2019). Deep fakes: A looming challenge for privacy, democracy, and national security. *California Law Review, 107*, 1753–1819.

Commissione europea (2025). *Digital Services Act report lays out landscape of systemic risks online*. Shaping Europe's digital future. https://digital-strategy.ec.europa.eu/en/news/digital-services-act-report-lays-out-landscape-systemic-risks-online

Crawford, K. (2021). *The Atlas of AI: Power, Politics, and the Planetary Costs of Artificial Intelligence*. New Haven, CT: Yale University Press.

Eslami, M., Rickman, A., Vaccaro, K., Aleyasen, A., Vuong, A., Karahalios, K., Hamilton, K., & Sandvig, C. (2015). "I always assumed that I wasn't really that close to [her]": Reasoning about invisible algorithms in news feeds. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (CHI '15)* (pp. 153–162). ACM.

European Board for Digital Services (2025). *First report of the European Board for Digital Services in cooperation with the Commission pursuant to Article 35(2) DSA on the most prominent and recurrent systemic risks as well as mitigation measures* (18 November 2025). Bruxelles: European Board for Digital Services / Commissione europea. https://cdn.table.media/assets/europe/first_article_352_dsa_report_on_systemic_risks_and_mitigations_final_ddxkzxhwga8vftj3unr0mgkwqvk_121707.pdf

Fletcher, R., & Nielsen, R. K. (2018). Are people incidentally exposed to news on social media? A comparative analysis. *New Media & Society, 20*(7), 2450–2468.

Floridi, L. (2002). What is the philosophy of information? *Metaphilosophy, 33*(1–2), 123–145.

Gagliardone, I. (2019). Defining online hate and its "public lives": What is the place for "extreme speech"? *International Journal of Communication, 13*, 20.

Gebru, T., Morgenstern, J., Vecchione, B., Vaughan, J. W., Wallach, H., Daumé III, H., & Crawford, K. (2021). Datasheets for datasets. *Communications of the ACM, 64*(12), 86–92.

Guo, K., Hu, A., Mu, J., Shi, Z., Zhao, Z., Vishwamitra, N., & Hu, H. (2023). An investigation of large language models for real-world hate speech detection. In *2023 International Conference on Machine Learning and Applications (ICMLA)* (pp. 1568–1573). IEEE.

Gruppo di lavoro sull'odio online (2021). *Odio online. Rapporto finale* (5 febbraio 2021).

Harriger, J. A., Evans, J. A., Thompson, J. K., & Tylka, T. L. (2022). The dangers of the rabbit hole: Reflections on social media as a portal into a distorted world of edited bodies and eating disorder risk and the role of algorithms. *Body Image, 41*, 292–297.

Hargittai, E. (2010). Digital Na(t)ives? Variation in Internet skills and uses among members of the "Net Generation". *Sociological Inquiry, 80*(1), 92–113.

Huang, L., Yu, W., Ma, W., Zhong, W., Feng, Z., Wang, H., & Liu, T. (2025). A survey on hallucination in large language models: Principles, taxonomy, challenges, and open questions. *ACM Transactions on Information Systems, 43*(2), 1–55.

Jane, E. A. (2014). "You're a ugly, whorish, slut": Understanding e-bile. *Feminist Media Studies, 14*(4), 531–546.

Khan, A. A., Laghari, A. A., Inam, S. A., Ullah, S., Shahzad, M., & Syed, D. (2025). A survey on multimedia-enabled deepfake detection: State-of-the-art tools and techniques, emerging trends, current challenges & limitations, and future directions. *Discover Computing, 28*(1), 48.

Kosmyna, N., Hauptmann, E., Yuan, Y. T., Situ, J., Liao, X.-H., Beresnitzky, A. V., & Maes, P. (2025). *Your brain on ChatGPT: Accumulation of cognitive debt when using an AI assistant for essay writing task.* arXiv:2506.08872.

Knight-Georgetown Institute (2025). *Systemic Risk Assessment under the Digital Services Act* (Commentary, 15 May 2025). https://kgi.georgetown.edu/research-and-commentary/systemic-risk-assessment-under-the-digital-services-act/

Kroll, J. A., Huey, J., Barocas, S., Felten, E. W., Reidenberg, J. R., Robinson, D. G., & Yu, H. (2017). Accountable algorithms. *University of Pennsylvania Law Review, 165*(3), 633–705.

Matz, S. C., Kosinski, M., Nave, G., & Stillwell, D. J. (2017). Psychological targeting as an effective approach to digital mass persuasion. *Proceedings of the National Academy of Sciences, 114*(48), 12714–12719.

Mitchell, M., Wu, S., Zaldivar, A., Barnes, P., Vasserman, L., Hutchinson, B., Spitzer, E., Raji, I. D., & Gebru, T. (2019). Model cards for model reporting. In *Proceedings of the Conference on Fairness, Accountability, and Transparency (FAT*)* (pp. 220–229). ACM.

Papacharissi, Z. (2004). Democracy online: Civility, politeness, and the democratic potential of online political discussion groups. *New Media & Society, 6*(2), 259–283.

Pei, G., Zhang, J., Hu, M., Zhang, Z., Wang, C., Wu, Y., & Tao, D. (2024). *Deepfake generation and detection: A benchmark and survey.* arXiv:2403.17881.

Popan, J. R., Coursey, L., Acosta, J., & Kenworthy, J. (2019). Testing the effects of incivility during internet political discussion on perceptions of rational argument and evaluations of a political outgroup. *Computers in Human Behavior, 96*, 123–132.

Risko, E. F., & Gilbert, S. J. (2016). Cognitive offloading. *Trends in Cognitive Sciences, 20*(9), 676–688.

Roy, S., Harshvardhan, A., Mukherjee, A., & Saha, P. (2023). Probing LLMs for hate speech detection: Strengths and vulnerabilities. In *Findings of the Association for Computational Linguistics: EMNLP 2023* (pp. 6116–6128).

Şahin, E., Arslan, N. N., & Özdemir, D. (2025). Unlocking the black box: An in-depth review on interpretability, explainability, and reliability in deep learning. *Neural Computing and Applications, 37*(2), 859–965.

Salvi, F., Horta Ribeiro, M., Gallotti, R., & West, R. (2025). On the conversational persuasiveness of GPT-4. *Nature Human Behaviour, 9*(8), 1645–1653.

Schiff, K. J., Schiff, D. S., & Bueno, N. S. (2025). The liar's dividend: Can politicians claim misinformation to evade accountability? *American Political Science Review, 119*(1), 71–90.

Suler, J. (2004). The online disinhibition effect. *CyberPsychology & Behavior, 7*(3), 321–326.

Suchman, L. A. (2007). *Human–Machine Reconfigurations: Plans and Situated Actions*. Cambridge: Cambridge University Press.

Tufekci, Z. (2015). Algorithmic harms beyond Facebook and Google: Emergent challenges of computational agency. *Colorado Technology Law Journal, 13*, 203–218.

Vaccari, C., & Chadwick, A. (2020). Deepfakes and disinformation: Exploring the impact of synthetic political video on deception, uncertainty, and trust in news. *Social Media + Society, 6*(1).

Wardle, C., & Derakhshan, H. (2017). *Information disorder: Toward an interdisciplinary framework for research and policy making.* Strasbourg: Council of Europe.

Woolley, S. C., & Howard, P. N. (2016). Political communication, computational propaganda, and autonomous agents: Introduction. *International Journal of Communication, 10*.

Yakura, H., Lopez-Lopez, E., Brinkmann, L., Serna, I., Gupta, P., Soraperra, I., & Rahwan, I. (2024). *Empirical evidence of Large Language Model's influence on human spoken communication.* arXiv:2409.01754.

Zhang, B., Cui, H., Nguyen, V., & Whitty, M. (2025). Audio deepfake detection: What has been achieved and what lies ahead. *Sensors, 25*(7), 1989.

Zuboff, S. (2023). *Il capitalismo della sorveglianza. Il futuro dell'umanità nell'era dei nuovi poteri.* Roma: Luiss University Press.

Zuiderveen Borgesius, F., Möller, J., Kruikemeier, S., Ó Fathaigh, R., Irion, K., Dobber, T., & de Vreese, C. H. (2018). Online political microtargeting: Promises and threats for democracy. *Utrecht Law Review, 14*(1), 82–96.

6 La tutela del diritto d'autore nell'era di ChatGPT – quale ruolo per AGCOM?

Giuseppe Cassano

Sommario: 1. Introduzione. - 2. Evoluzione della tecnica e diritto d'autore. - 3. Opere generate dalla IA e diritto d'autore. - 4. IA e informazione visiva. - 5. L'autore artificiale, fu vero autore?. - 6. Esperienze normative a confronto. - 7. L'oggetto della tutela del diritto d'autore al tempo dell'IA . - 8. L'IA e le modifiche introdotte dal Legislatore. - 9. Addestramento dei sistemi di IA e tutela autoriale. - 10. Codice di Buone Pratiche per l'IA

6.1 Introduzione

All'AGCOM compete un ruolo di primo ordine nella materia della tutela del diritto d'autore essendosi in tal senso più volte espresso lo stesso Legislatore nel corso degli anni (v. L. 22 aprile 1941, n. 633; D.Lgs. 9 aprile 2003, n. 70; D.Lgs. 15 marzo 2017, n. 35; D.Lgs. 8 novembre 2021, n. 208).

Da ultimo si registrano:

- il D.Lgs. 8 novembre 2021, n. 177 (Attuazione della direttiva (UE) 2019/790 del Parlamento europeo e del Consiglio, del 17 aprile 2019, sul diritto d'autore e sui diritti connessi nel mercato unico digitale) secondo cui "decorsi due anni dalla data di entrata in vigore del presente decreto, l'Autorità per le garanzie nelle comunicazioni trasmette alle Camere una relazione, integrata da verifica d'impatto della regolazione, sull'applicazione di propria competenza delle disposizioni che vi sono contenute, con particolare riferimento ai criteri e alle modalità di determinazione dell'equo compenso per gli editori di pubblicazioni giornalistiche, di cui all'articolo 43-bis della legge 22 aprile 1941, n. 633, e alla procedura di determinazione dei compensi per gli autori, artisti, interpreti e esecutori, previo confronto con gli organismi di gestione collettiva o entità di gestione indipendenti di cui al decreto legislativo 15 marzo 2017 n. 35" (art. 2);

- la Legge 14 luglio 2023, n. 93 (Disposizioni per la prevenzione e la repressione della diffusione illecita di contenuti tutelati dal diritto d'autore mediante le reti di comunicazione elettronica) secondo cui, tra l'altro, l'Agcom "con proprio provvedimento, ordina ai prestatori di servizi, compresi i prestatori di accesso alla rete, di disabilitare l'accesso a contenuti diffusi abusivamente mediante il blocco della risoluzione DNS dei nomi di dominio e il blocco dell'instradamento del traffico di rete verso gli indirizzi IP prevalentemente destinati ad attività illecite" (comma I). E ancora "Con il provvedimento di cui al comma 1, l'Autorità ordina anche il blocco di ogni altro futuro nome di dominio, sottodominio, o indirizzo IP, a chiunque riconducibili, comprese le variazioni del nome o della semplice declinazione o estensione (cosiddetto top level domain), che consenta l'accesso ai medesimi contenuti diffusi abusivamente e a contenuti della stessa natura" (comma II).

A chiusura di questa breve sintesi del quadro normativo di riferimento si ritiene utile un richiamo al Regolamento del Parlamento Europeo relativo a un mercato unico dei servizi digitali - regolamento sui servizi digitali, o "Digital Services Act" (DSA) – secondo cui "gli Stati membri designano una o più autorità competenti incaricate della vigilanza dei fornitori di servizi intermediari e dell'esecuzione del presente regolamento («autorità competenti»)" (art. 49).

Orbene, l'Autorità per le garanzie nelle comunicazioni è stata designata quale Coordinatore dei Servizi Digitali (art. 15 D.L. 15 settembre 2023, n. 123 - Misure urgenti di contrasto al disagio giovanile, alla povertà educativa e alla criminalità minorile, nonché per la sicurezza dei minori in ambito digitale - convertito in legge, con modificazioni, dall'art. 1, I, L. 13 novembre 2023, n. 159).

6.2 Evoluzione della tecnica e diritto d'autore

La diffusione incessante, inarrestabile e quasi autoritaria, dell'Intelligenza Artificiale (AI) nella quotidianità di singoli, imprese e soggetti pubblici tutti, pone nuove frontiere di dibattito in tema di tutela del diritto d'autore, né potrebbe essere diversamente giacchè ad ogni evoluzione tecnica di rilievo ha fatto sempre da contraltare la necessità della rimodulazione della tutela qui in esame (Adelaide Rossi).

L'intreccio tra evoluzione della tecnica e diritto d'autore non è dunque un tema esclusivo appannaggio di questi nostri giorni, né tipicamente proprio della intelligenza artificiale generativa (qui indicata anche come GenAI).

Per "intelligenza artificiale generativa" <<si intende il campo dell'intelligenza artificiale che si concentra sulla creazione di contenuti nuovi e originali rispetto ai dati di input in risposta alle richieste (prompt) dell'utente, attraverso l'utilizzo di algoritmi prevalentemente di tipo neurale.

Per "rete neurale" si intende un modello computazionale standard applicabile nei contesti più diversificati che permette il riconoscimento di oggetti, forme o pattern all'interno di un dato o un insieme di dati (ad esempio, un volto umano in una fotografia).

Gli algoritmi di intelligenza artificiale generativa sono impiegati in una vasta gamma di applicazioni, tra cui il riconoscimento e la generazione di immagini, di tracce vocali o musicali, di testi e di video>> (Garante per la protezione dei dati personali - Provvedimento del 10 aprile 2025, n. 232).

6.3 Opere generate dalla IA e diritto d'autore

Già nel tempo si è dibattuta la tutela del diritto d'autore in merito alla fotografia, alle banche dati (CGUE, sez. III, 18 ottobre 2012, n. 173/11), ai software, ai progetti tecnici (da ultimo T.a.r. Emilia-Romagna, Bologna, sez. I, 22 ottobre 2025, n. 1180) etc. giungendo a soluzioni che possono dirsi ancora utili anche in riferimento alla IA.

In tal senso, quanto alla fotografia, è certamente valido l'insegnamento della CGUE (sez. III, 1 dicembre 2011, n. 145) secondo cui un ritratto fotografico può essere protetto dal diritto d'autore alla condizione, che spetta al Giudice nazionale verificare in ogni singolo caso di specie, che costituisca una creazione intellettuale dell'autore che ne rifletta la personalità e si manifesti attraverso scelte libere e creative di quest'ultimo nella realizzazione di tale ritratto.

Proprio con riferimento ad una fotografia si è addivenuti in Italia ad una primissima pronuncia della Suprema Corte di Cassazione in materia di software e tutela del diritto d'autore per un'opera generata dalla IA.

Tutto ha inizio con la richiesta della creatrice dell'opera grafica "The Scent of the Night" dell'accertamento della violazione dei propri diritti d'autore, nonché della condanna della convenuta (RAI) al risarcimento dei danni per l'utilizzo, non autorizzato, della stessa come scenografia fissa del Festival di Sanremo 2016.

Il Giudice di prime cure (Tribunale di Genova, sentenza 6 giugno 2018, n. 1640) ha accolto tale richiesta con pronuncia che ha poi superato indenne il vaglio critico del Giudice di Appello (Appello di Genova, sez. spec. in materia di imprese, sentenza 11 novembre 2020, n. 1066).

All'esito la Corte di Cassazione, con ordinanza del 16 gennaio 2023, n. 1107, pur pronunciandosi solo incidentalmente sul tema, ha comunque dato semaforo verde alla possibilità di attribuire carattere creativo ad una fotografia realizzata dal suo autore tramite un software di IA.

Secondo gli Ermellini "l'ammissione …di aver utilizzato un software per generare l'immagine" è "circostanza … pur sempre compatibile con l'elaborazione di un'opera dell'ingegno con un tasso di creatività che andrebbe solo scrutinato con maggior rigore …".

6.4 IA e informazione visiva

La GenAi di immagini rappresenta una frontiera verso cui correre governando i temi controversi che pone ma, al tempo stesso, profittando delle numerose sue applicazioni utili per singoli e collettività.

D'altronde, non è seriamente contestabile che «sapere e comprendere … è fondamentale, prima di decidere se e come servirsi dello strumento.

(…)sebbene di recente siano tutti intenti a parlare di Intelligenza Artificiale - dividendosi più o meno equamente fra coloro che si approcciano all'IA con utopico ottimismo (formulando affascinanti ipotesi sulle sue innumerevoli e mirabolanti applicazioni possibili alla vita quotidiana, al lavoro, ma anche all'attività dello Stato e della Pubblica Amministrazione nelle sue molteplici attività) e coloro che, invece, vi si approcciano con maggiore paura e diffidenza - pochi sanno di che cosa stanno davvero parlando.

Anche gli studiosi di area umanistica che si avvicinano al tema IA - e i giuristi rientrano ovviamente all'interno di questa categoria - si trovano sovente in una situazione di vero e

proprio "deficit cognitivo": nel senso che discutono di un oggetto per loro largamente sconosciuto, una sorta di UFO, potremmo dire» (Diana-Urania Galetta).

Una applicazione della GenAi di immagini certamente utile per la collettività può rinvenirsi nella visione artificiale applicata alla ricostruzione di artefatti visivi (quali antichi affreschi, superfici pittoriche, dipinti frammentari o deteriorati) per riconoscere pattern cromatici, texture, tracce materiali, porzioni mancanti e stratificazioni delle singole opere attenzionate. Le tecniche di computer vision permettono di ricostruire parti perdute, suggerire integrazioni plausibili, identificare stesure pittoriche successive, classificare pigmenti o individuare interventi di restauro pregressi. L'obiettivo diventa, in altre parole, la ricostruzione dell'immagine, la comprensione della sua struttura visiva e la restituzione quanto più fedele possibile dell'aspetto originario dell'artefatto.

Si contribuisce così alla valorizzazione e allo studio del patrimonio culturale (cd. intelligenza artificiale applicata ai beni culturali).

L'IA è qui chiamata ad operare esclusivamente sull'informazione visiva: analizzare forme, colori, texture e pattern per identificare parti danneggiate, migliorare la leggibilità delle superfici o proporre ricostruzioni plausibili delle zone mancanti. Si tratta quindi di un lavoro centrato sull'elaborazione dell'immagine, che richiede competenze di tipo tecnico-informatico (computer vision, modelli generativi, segmentazione visiva) (T.a.r. Campania, Napoli, sez. V, 5 marzo 2026, n. 1525).

6.5 L'autore artificiale, fu vero autore?

Orbene, le numerose questioni che, oggi, gli operatori del diritto sono chiamati ad affrontare quanto al rapporto tra IA generativa e diritto d'autore impongono la ricerca di soluzioni di ampio respiro che si collochino oltre il solco tracciato dalle esperienze normative nazionali come modellate, nel corso degli anni, dall'apporto della giurisprudenza; con la precisazione che la GenAI implica il ripensamento delle coordinate che descrivono i soggetti e l'oggetto del diritto.

Quanto al profilo soggettivo della tutela l'interrogativo che pone la GenAI è la possibilità, o meno, di configurare un sorta di autore "artificiale".

Sul punto è opportuna una considerazione preliminare. Il dato normativo di riferimento – sia nazionale, che euro-unionale che, infine, di matrice internazionale – non è dirimente al fine di risolvere l'interrogativo in quanto, a volte, fa riferimento all'autore quale persona fisica, altre volte ammette l'autore "persona giuridica", senza mai spingersi fino ad accordare tutela anche all'autore artificiale.

È vero, al tempo stesso, che diversi fattori lasciano propendere per la limitazione alla sola persona fisica (o entro certi limiti giuridica) così neutralizzando, azzerandola, la questione dell'autore artificiale.

È con riferimento alla persona umana, invero, che il Legislatore accorda, in caso di accertate violazioni, la tutela anche del diritto morale, così come è solo per essa che può parlarsi di evento di "morte" nel settore qui in esame, come noto, rilevante ai fini del conteggio della durata nel tempo del sistema di protezione.

Ed infine, ma non da ultimo, l'autore pone in essere scelte consapevoli quando realizza la propria opera, non potendosi accordare alcuna consapevolezza ai sistema di IA.

Né può omettersi un richiamo all'art. 27 della Dichiarazione Universale dei Diritti Umani secondo cui "ogni individuo ha diritto di prendere parte liberamente alla vita culturale della comunità, di godere delle arti e di partecipare al progresso scientifico ed ai suoi benefici" (I comma) e, ancora, "ogni individuo ha diritto alla protezione degli interessi morali e materiali derivanti da ogni produzione scientifica, letteraria e artistica di cui egli sia autore".

Orbene, è fuori da ogni dubbio come non si possa parlare di diritti umani nei confronti di una macchina che, per restare al tema in esame, e per approntare una prima conclusione non è consapevole, non soffre, né prova emozioni (cui possa accedere una tutela in termini di diritti non patrimoniali), né potrà mai morire, potendo al più cessare di funzionare.

6.6 Esperienze normative a confronto

La questione appena tratteggiata impone alcune riflessioni alla luce delle normative vigenti in Paesi anche lontani dal nostro e che, tuttavia, possono offrire spunti di riflessione utili e da accogliere con interesse. D'altronde, come evidenziato sin dalle prime battute, il tema qui in esame necessita di soluzioni quanto più possibile omogenee che superino i confini nazionali.

Il primo rinvio che si vuol operare è all'art. 11 della legge cinese sul diritto d'autore il quale è chiaro nell'indicare come autore di un'opera intellettuale la sola persona fisica suo "creatore", così mettendo fuori gioco ogni possibile interpretazione atta a ricondurre in tale concetto l'autore artificiale.

Precisamente, secondo tale disposizione: "An author is a natural person who creates a work. The author is the natural person who creates the work. For works hosted by a corporate or unincorporated organization, created on behalf of the will of the corporate or unincorporated organization, and for which the corporate or unincorporated organization assumes responsibility, the corporate or unincorporated organization shall be regarded as the author". Emerge così il ruolo di mero ausilio che deve svolgere il sistema di GenAI. Per un'applicazione giurisprudenziale di tale normativa si rinvia alla pronuncia resa dal il Tribunale di Internet di Pechino (Beijing Internet Court o BIC) 27 novembre 2023, n. 11279 di condanna per l'illecito utilizzo di un'immagine generata da Stable Diffusion (Virgilio D'Antonio - Ciro Maria Ruocco).

Nell'esperienza normativa degli USA tanto il Copyright Act, quanto, e in termini ancora più netti, il Compendium of U.S. Copyright Office Practices negano tutela alle opere frutto di soli processi meccanici.

Guardando ancora all'esperienza degli USA è utile richiamare le linee guida (del 16 marzo 2023) dell'U.S. Copyright Office che hanno chiarito come il copyright è uno strumento di tutela delle solo opere create dagli uomini, osservando che "public guidance is needed on the registration of works containing AI-generated content. This statement of policy describes how the Office applies copyright law's human authorship requirement to applications to register such works and provides guidance to applicants".

Con la precisazione, al tempo stesso, che le conclusioni sono nel senso che "this policy statement sets out the Office's approach to registration of works containing material generated by AI technology. The Office continues to monitor new factual and legal developments involving AI and copyright and may issue additional guidance in the future related to registration or the other copyright issues implicated by this technology".

Quanto al sistema messicano, il riferimento è al Federal Law of Author's Right (Ley Federal del Derecho de Autor - LFDA) che pone la persona fisica al centro della tutela.

Di recente la decisione della Suprema Corte di Giustizia della Nazione (SCJN) del Messico, Segunda Sala; sentencia 2 de julio de 2025; Amparo Directo 6/2025, ha sottolineato la dimensione antropocentrica del diritto d'autore, escludendo ogni possibile riconoscimento di status autoriale ai sistemi di intelligenza artificiale (Virgilio D'Antonio - Ciro Maria Ruocco).

Approdando ora nel vecchio continente è da rilevare come la normativa britannica si discosti, e non di poco, da quella continentale (Ciro Maria Ruocco).

Il § 178 Copyright, Patents and Designs Act (CDPA -1988) prevede, tra l'altro, nelle minor definitions, <<"computer-generated", in relation to a work, means that the work is generated by computer in circumstances such that there is no human author of the work>>; è chiaro qui il riferimento ad un'opera generata da un computer in circostanze tali che non vi è alcun "autore umano dell'opera".

Si tratta, a ben vedere, di una norma che si riferisce ai semplici pc, e non alle macchine intelligenti, ma il cui raggio di azione è esteso, oggi, anche a queste ultime.

Sul punto della paternità dell'opera creata da tali macchine viene in aiuto (pur non risolvendo del tutto i molti nodi interpretativi) il § 9 (3) secondo cui <<in the case of a literary, dramatic, musical or artistic work which is computer-generated, the author shall be taken to be the person by whom the arrangements necessary for the creation of the work are undertaken>>; dunque, l'autore è la persona che si occupa di predisporre le misure necessarie alla creazione dell'opera.

Non può mancarsi poi un cenno all'art. 15 della Convenzione di Berna per la protezione delle opere letterarie ed artistiche (cui lo Stato italiano ha dato ratifica ed esecuzione con la L. 20 giugno 1978, n. 399) il quale prevede, espressamente, che "affinché gli autori di opere letterarie ed artistiche protette dalla presente Convenzione siano fino a prova contraria ritenuti tali, ed ammessi in conseguenza ad agire contro i contraffattori davanti ai tribunali dei Paesi dell'Unione, è sufficiente che il nome dell'autore sia indicato sull'opera nei modi d'uso. Il presente alinea è applicabile anche se il nome sia uno pseudonimo, purché questo non lasci dubbi sull'identità dell'autore" (I comma); tale disposizione dà per presupposto - scontato, quanto necessario risalendo la sua firma al 9 settembre del 1886 - che l'autore sia una persona fisica.

Infine, dell'esperienza eurounitaria, si considerino:

- la Direttiva 96/9/CE del Parlamento europeo e del Consiglio, dell'11 marzo 1996, relativa alla tutela giuridica delle banche di dati (cd. "Direttiva Database") secondo cui "l'autore di una banca di dati è la persona fisica o il gruppo di persone fisiche che l'ha creata o, qualora la legislazione dello Stato membro interessato lo consenta, la persona giuridica individuata da tale legislazione come titolare del diritto" (art. 4, I);
- la Direttiva 2009/24/CE del Parlamento europeo e del Consiglio del 23 aprile 2009 relativa alla tutela giuridica dei programmi per elaboratore (cd. "Direttiva Software") secondo cui "la tutela è riconosciuta a tutte le persone fisiche o giuridiche aventi i requisiti previsti dalla legislazione nazionale sul diritto d'autore applicata alle opere letterarie" (art. 3).

Emerge come, secondo il Legislatore Europeo, autore del database o del software possa essere una persona fisica, o un gruppo di persone fisiche e, quando la legislazione dello Stato nazionale lo permetta, anche la persona giuridica che ha creato il programma.

6.7 L'oggetto della tutela del diritto d'autore al tempo dell'IA

Soffermandoci ora sul profilo oggettivo della tutela del diritto d'autore al tempo dell'IA, non può non dedicarsi qualche riferimento agli elementi costitutivi di tale diritto.

Secondo un consolidato approdo della giurisprudenza la protezione del diritto d'autore postula i requisiti della originalità e della creatività che devono dirsi sussistere anche quando l'opera sia composta da idee e nozioni semplici, comprese nel patrimonio intellettuale di persone aventi esperienza nella materia propria dell'opera stessa, purché formulate ed organizzate in modo personale ed autonomo rispetto alle precedenti (Cass. civ., 13 giugno 2014, n. 13524; Cass. civ., 28 novembre 2011, n. 25173; Cass. civ., 12 marzo 2004, n. 5089; Cass. civ., 2 dicembre 1993, n. 11953).

In particolare, la creatività consiste non già nell'idea che è alla base della sua realizzazione, ma nella forma della sua espressione, ovvero nella sua soggettività, di modo che la stessa idea può essere alla base di diverse opere d'autore, come è ovvio nelle opere degli artisti, le quali tuttavia sono (o possono essere) diverse per la creatività soggettiva che ciascuno degli autori spende, e che in quanto tale rileva per l'ottenimento della protezione (Cass. civ., 29 maggio 2020, n. 10300).

Questi principi di diritto, enunciati dalla giurisprudenza nazionale (v. da ultimo Cass. civ., sez. I, ord., 10 febbraio 2025, n. 3393), sono conformi ai principi elaborati, in materia, dalla giurisprudenza della CGUE.

Secondo quest'ultima, muovendo dal testo della Dir. 2001/29/CE del Parlamento Europeo e del Consiglio del 22 maggio 2001 (sull'armonizzazione di taluni aspetti del diritto d'autore e dei diritti connessi nella società dell'informazione), la nozione di "opera" implica che esista un oggetto originale, nel senso che detto oggetto rappresenta una creazione intellettuale propria del suo autore, anche se la qualifica di opera è riservata agli elementi che sono espressione di tale creazione (CGUE 12 settembre 2019, Cofemel, C-683/17, punto 29; CGUE 13 novembre 2018, Levola Hengelo, C-310/17, punti 33 e da 35 a 37).

Perché un oggetto possa essere considerato originale, è necessario e sufficiente che rifletta la personalità del suo autore, manifestando le scelte libere e creative di quest'ultimo (CGUE 12 settembre 2019, Cofemel, cit., punto 30; CGUE 7 agosto 2018, Renckhoff, C-161/17, punto 14; CGUE 10 dicembre 2011, Painer, C 145/10, punti 88, 89 e 94); la nozione di "opera", poi, importa necessariamente l'esistenza di un oggetto identificabile con sufficiente precisione e oggettività (CG UE 12 settembre 2019, Cofemel, cit., punto 32; CGUE 13 novembre 2018, Levola Hengelo, cit., punto 40).

Ed ancora, <<Gli articoli da 2 a 5 della direttiva 2001/29 del Parlamento europeo e del Consiglio, del 22 maggio 2001, sull'armonizzazione di taluni aspetti del diritto d'autore e dei diritti connessi nella società dell'informazione, devono essere interpretati nel senso che la protezione a titolo del diritto d'autore da essi prevista si applica a un prodotto la cui forma è, quantomeno in parte, necessaria per ottenere un risultato tecnico, qualora tale prodotto costituisca un'opera originale risultante da una creazione intellettuale in quanto, mediante tale forma, il suo autore esprime la propria capacità creativa in maniera originale effettuando scelte libere e creative, di modo che detta forma riflette la sua personalità, circostanza che spetta al giudice del rinvio verificare tenendo conto di tutti gli elementi pertinenti della controversia principale>> (CGUE, sez. V, 11 giugno 2020, n. 833/18)

Con la precisazione, infine, che nella materia che qui ci occupa si ha violazione dell'esclusiva non solo quando un'opera è copiata integralmente, cioè quando vi sia riproduzione abusiva,

ma anche nel caso della contraffazione che ricorre quando i tratti essenziali dell'opera anteriore si ripetono in quella successiva.

Tali tratti essenziali dell'opera (che dunque non sono passibili di replicazione) vanno individuati muovendo da quanto è frutto della creatività del suo autore; quel che rileva, dunque, è la forma espressiva intesa nella sua soggettività, vale a dire la scelta dell'autore stesso quanto alla rappresentazione dell'idea, non l'idea in quanto tale (Cass. civ., sez. I, ord., 29 luglio 2025, n. 21851).

6.8 L'IA e le modifiche introdotte dal Legislatore

Orbene alla luce di quanto fin qui detto deve ora passarsi ad una analisi della normativa dopo gli interventi del Legislatore finalizzati a disciplinare il binomio diritto d'autore, da un lato, e GenAI, dall'altro lato.

In primis va richiamato l'art. 1 L. n. 633/1941, nel testo come risultante dopo l'intervento della L. 23 settembre 2025, n. 132 (Disposizioni e deleghe al Governo in materia di intelligenza artificiale) secondo cui "sono protette ai sensi di questa legge le opere dell'ingegno umano di carattere creativo che appartengono alla letteratura, alla musica, alle arti figurative, all'architettura, al teatro ed alla cinematografia, qualunque ne sia il modo o la forma di espressione, anche laddove create con l'ausilio di strumenti di intelligenza artificiale, purché costituenti risultato del lavoro intellettuale dell'autore" (I comma)

Il Legislatore del 2025 – che, curiosamente, non interviene a modificare anche l'art. 2575 c.c. – accorda così tutela anche alle opere create con l'ausilio di strumenti di IA, con la precisazione che punto fermo è, e deve restare, il "risultato del lavoro intellettuale dell'autore".

La prospettiva è dunque quella antropocentrica, né avrebbe potuto essere diversamente giacchè tutto il sistema della L n. 132 cit. è improntato ad una "dimensione antropocentrica" dell'intelligenza artificiale (art. 1)

Dalla lettura dell'art. 1 L. n. 633 emerge dunque come, con riferimento al tema qui trattato (binomio: diritto d'autore - GenAI) non sorgano questioni nuove per le opere frutto dell'esclusivo apporto umano (ovviamente, per il motivo che esse non si confrontano con i sistema di GenAI) e per quelle opere frutto dell'esclusivo apporto della AI giacchè, queste

ultime, sono messe fuori gioco dall'impianto normativo (gli strumenti di intelligenza artificiale, come visto, dovendo essere solo di ausilio).

Di diritto d'autore può, e deve, invece parlarsi quanto alle opere frutto dell'apporto umano che si avvale della GenAI, diritto che impone, a livello (almeno) europeo, soluzioni condivise e omogenee.

Vediamo più da vicino i nodi da sciogliere.

Precisamente il tema è quello della rilevanza giuridica sotto il profilo della tutela del diritto d'autore quanto all'input (ossia ai dati adoperati per l'addestramento del sistema di IA).

Quanto all'output (cioè all'opera generata dal sistema di IA), come detto, se trattasi di opera integralmente frutto dell'autore artificiale si pone fuori dallo schema della tutela legale; se trattasi opera realizzata dall'autore umano con l'ausilio di sistema di IA la tutela seguirà le vie ordinarie.

6.9 Addestramento dei sistemi di IA e tutela autoriale

Come noto, l'IA generativa di immagini con carattere di originalità – facilmente accessibile grazie anche ai suoi costi spesso piuttosto contenuti – impone agli operatori di diritto l'analisi della rilevanza giuridica dell'immissione, da parte di un utente, di un prompt (cioè di un testo) in un sistema di apprendimento profondo (Deep Learning) per la creazione di una immagine nuova e inedita.

L'IA generativa necessita di attingere a molte informazioni (nell'ordine di miliardi) per il suo apprendimento e spesso attinge anche ad opere coperte dal diritto d'autore ponendo la questione della tutela di quest'ultimo che vede, a livello normativo, soluzione tra loro differenti.

Così, mentre negli Stati Uniti si ricorre al principio del "fair use" (disciplinato dalla U.S. Copyright Law -Chapter 1, Section 107, che legittima l'incorporazione di materiale protetto da copyright in opere aventi scopo di critica, insegnamento, ricerca e giornalismo), l'esperienza euro-unionale si fonda sull'eccezione per il cd. "Text and Data Mining" (TDM).

In Italia, in particolare, il comma 1 bis dell'art. 70 L. n. 633/1941 espressamente consente "la libera pubblicazione attraverso la rete internet, a titolo gratuito, di immagini e musiche a bassa risoluzione o degradate, per uso didattico o scientifico e solo nel caso in cui tale utilizzo

non sia a scopo di lucro. Con decreto del Ministro per i beni e le attività culturali, sentiti il Ministro della pubblica istruzione e il Ministro dell'università e della ricerca, previo parere delle Commissioni parlamentari competenti, sono definiti i limiti all'uso didattico o scientifico di cui al presente comma".

A sua volta il successivo art. 70 ter consente le riproduzioni compiute da organismi di ricerca, e da istituti di tutela del patrimonio culturale, per scopi di ricerca scientifica, ai fini dell'estrazione di testo e di dati da opere o da altri materiali disponibili in reti o banche di dati cui essi hanno lecitamente accesso, nonché la comunicazione al pubblico degli esiti della ricerca ove espressi in nuove opere originali.

Infine, a norma dell'art. 70 quater: "Fermo restando quanto previsto dall'articolo 70-ter, sono consentite le riproduzioni e le estrazioni da opere o da altri materiali contenuti in reti o in banche di dati cui si ha legittimamente accesso ai fini dell'estrazione di testo e di dati. L'estrazione di testo e di dati è consentita quando l'utilizzo delle opere e degli altri materiali non è stato espressamente riservato dai titolari del diritto d'autore e dei diritti connessi nonché dai titolari delle banche dati" (I comma).

Per concludere con i richiami normativi, l'art. 70 septies stabilisce: "Fermo restando quanto previsto dalla Convenzione di Berna per la protezione delle opere letterarie ed artistiche, ratificata e resa esecutiva ai sensi della legge 20 giugno 1978, n. 399, le riproduzioni e le estrazioni da opere o da altri materiali contenuti in rete o in banche di dati a cui si ha legittimamente accesso, ai fini dell'estrazione di testo e di dati attraverso modelli e sistemi di intelligenza artificiale, anche generativa, sono consentite in conformità alle disposizioni di cui agli articoli 70-ter e 70-quater."

Orbene, quest'ultima norma, inserita dall'art. 25, I, lett. b), L. n. 132/2025, espressamente include la disciplina del text and data mining nella L. n. 633/1941 dimostrando la volontà del Legislatore di tutelare le esigenze di addestramento dei sistemi di intelligenza artificiale.

Essa trova il suo referente più immediato nel Regolamento del Parlamento europeo e del Consiglio (che stabilisce regole armonizzate sull'intelligenza artificiale) n. 2024/1689/UE, e precisamente nell'art. 53 sugli "Obblighi dei fornitori di modelli di IA per finalità generali" per quanto, dal raffronto dei due testi emerga come, l'art. 53 non faccia riferimento, a

differenza del testo italiano, al text and data mining (che in verità si trova al considerando 105).

Ne consegue che "il training di sistemi e modelli di IA su materiale tutelato dal diritto d'autore è lecito, grazie all'eccezione per le estrazioni di testo e dati, a condizione che tale materiale sia legittimamente accessibile, e che l'avente diritto non abbia esercitato l'opt-out. Ovviamente, restano da precisare le modalità di esercizio di tale opzione, così come la questione della responsabilità in caso di mancato rispetto della medesima" (Bruno Tassone – Sara Manni). La questione è oggettivamente più complesa e può essere meglio compresa attraverso i temi affrontati dal Tribunale di Monaco. Più in particolare l'11 novembre 2025 il Tribunale di Monaco I (Landgericht München I) ha condannato OpenAI nella causa promossa dalla società "collecting" tedesca GEMA.

Il tribunale ha in particolare sancito che l'utilizzo di testi di canzoni da parte della piattaforma ha violato il diritto d'autore:

- Ø memorizzando testi (di canzoni)
- Ø visualizzando parti dei testi nell'output.

Il Tribunale ha dunque stabilito che non si applica la limitazione di text-and-data mining ai sensi dell'art. 4 della direttiva 790/2019, in quanto ha ritenuto che tali testi, essendo contenuti nei parametri del modello e potendo essere restituiti come output, non siano semplicemente "utilizzati" a fini analitici, ma "memorizzati" in senso proprio, cioè incorporati in modo stabile e riutilizzabile.

Tale fenomeno di "memorizzazione" segna il punto di discrimine rispetto alla mera estrazione di testo e dati che caratterizza il text and data mining.

La memorizzazione si verifica quando, durante l'addestramento, i modelli linguistici non solo estraggono informazioni dal set di dati di addestramento, ma mostrano anche una completa incorporazione dei dati di addestramento nei parametri specificati dopo l'addestramento, tanto che i testi delle canzoni sono stati anche riprodotti negli output

Il Tribunale ha distinto tra:

- Riproduzioni meramente preparatorie, necessarie per l'analisi dei dati
- Riproduzioni che danno luogo a un'incorporazione stabile dell'opera nel modello.

Nel primo caso, la legge ammette la riproduzione temporanea in quanto funzionale alla successiva analisi.

Nel secondo, invece, la riproduzione permanente inciderebbe sui diritti di sfruttamento economico dell'autore.
Pertanto, il Landesgericht Monaco, con la Sentenza 11.11.2025 - 42 O 14139/24, ha condannato le convenute OpenAI USA e OpenAl Ireland a cessare:

- di riprodurre, in tutto o in parte, senza il consenso dell'attrice GEMA, i testi delle canzoni allegati nel fascicolo K1

all'interno di modelli linguistici (Large Language Models) e/o di far eseguire tali atti da terzi;

- di rendere pubblicamente accessibili e/o riprodurre, in tutto o in parte e/o sotto forma di elaborazioni, senza il consenso dell'attrice, i testi delle canzoni allegati nel fascicolo K1 negli output ("Outputs") di un chatbot, e/o di far eseguire tali atti da terzi (fascicolo K2) tali atti da terzi (fascicolo K2).

Le convenute sono state, altresì, condannate a fornire all'attrice le informazioni sull'estensione delle condotte descritte e quali introiti esse ne abbiano ricavato, indicando:

- il numero e l'entità delle condotte;
- i ricavi per ciascuna delle convenute.

Si accerta che le convenute sono obbligate a risarcire all'attrice ogni danno che le è già derivato e/o le deriverà dalle condotte descritte al punto 1, secondo quanto risulterà dalle informazioni rese ai sensi del punto 2.
Dunque, siamo in presenza di Nuove sfide e un criterio di riferimento già ben noto Il Three-step test (art. 9(2) Convenzione di Berna; art. 5 Dir. 2001/29/CE) , che prevedono come il ricorso alle eccezioni e limitazioni al diritto d'autore è ammesso alle seguenti condizioni:

- Caso speciale
  (L'uso rientra in un'eccezione chiaramente definita e di natura circoscritta).
- Assenza di conflitto con il normale sfruttamento dell'opera
  (L'eccezione non deve compromettere il mercato attuale o potenziale dell'opera).
- Nessun pregiudizio ingiustificato agli interessi legittimi del titolare
  (L'impatto sui diritti dell'autore deve rimanere proporzionato e non irragionevole).

Il tutto senza dimenticar che in ogni caso «Le eccezioni e le limitazioni non possono essere applicate in modo da arrecare pregiudizio agli interessi legittimi dei titolari dei diritti o da essere in contrasto con la normale utilizzazione economica delle loro opere o materiali protetti … È pertanto possibile che la portata di alcune eccezioni o limitazioni debba essere ulteriormente limitata nel caso di taluni nuovi utilizzi di opere e materiali protetti (considerando 44)

6.10 Codice di Buone Pratiche per l'IA

I fornitori di modelli di intelligenza artificiale generativa dovranno infine attenersi al Codice di Buone Pratiche per l'IA (CPAI) secondo cui, tra l'altro, ed espressamente:

<<In order to help ensure that Signatories will identify and comply with, including through state-of-the-art technologies, machine-readable reservations of rights expressed pursuant to Article 4(3) of Directive (EU) 2019/790 if they use web-crawlers or have such web-crawlers used on their behalf to scrape or otherwise compile data for the purpose of text and data mining as defined in Article 2(2) of Directive (EU) 2019/790 and the training of their general-purpose AI models, Signatories commit:

a) to employ web-crawlers that read and follow instructions expressed in accordance with the Robot Exclusion Protocol (robots.txt), as specified in the Internet Engineering Task Force (IETF) Request for Comments No. 9309, and any subsequent version of this Protocol for which the IETF demonstrates that it is technically feasible and implementable by AI providers and content providers, including rightsholders, and

b) to identify and comply with other appropriate machine-readable protocols to express rights reservations pursuant to Article 4(3) of Directive (EU) 2019/790, for example through asset-based or location-based metadata, that have either have been adopted by international or European standardisation organisations, or are state-of-the-art, including technically implementable, and widely adopted by rightsholders, considering different cultural sectors, and generally agreed through an inclusive process based on bona fide discussions to be facilitated at EU level with the involvement of rightsholders, AI providers and other relevant stakeholders as a more immediate solution, while anticipating the development of standards>>.

Quindi, in sintesi, qualora i fornitori (firmatari del Codice) utilizzino web crawler, o ne facciano utilizzare per loro conto, per estrarre o altrimenti compilare dati ai fini del text and data mining, per l'addestramento dei loro modelli di intelligenza artificiale generici, si devono impegnare a:

a) impiegare web crawler che leggano e seguano le istruzioni espresse in conformità al Robot Exclusion Protocol (robots.txt),

b) identificare e rispettare altri protocolli leggibili da una macchina appropriati per esprimere diritti / riserve ai sensi dell'art. 4(3) della direttiva (UE) 2019/790.

Con la precisazione, infine, che si applica una sanzione penale per chi "a-ter) riproduce o estrae testo o dati da opere o altri materiali disponibili in rete o in banche di dati in violazione degli articoli 70-ter e 70-quater, anche attraverso sistemi di intelligenza artificiale" (art. 171 L. n. 633/1941 come modificato dalla L. n. 132/2025).

Per una recente applicazione pratica di quanto fin qui detto si rinvia alla pronuncia del Tribunale di Amburgo (Landgericht Hamburg, Zivilkammer 10, Az. 310 O 227/23, Verkündet am 27 settembre 2024), che affrontando proprio il tema della liceità, o meno, della raccolta di opere protette dal diritto d'autore per creare dataset per l'addestramento di sistemi di IA ha giudicato legittima siffatta attività posta in essere, nel caso di specie, da una organizzazione no profit (la decisione si fonda sul § 60d UrhG che recepisce i principi ex art. 3 Direttiva 790/2019) (Ciro Maria Ruocco - Vera Iuzzolino).

Ed ancora si veda la già citata pronuncia resa dal Tribunale di Monaco di Baviera (Landgericht München I, 42. Zivilkammer, Urteil vom 11. November 2025, Az. 42 O 14139/24) che tratta dell'impiego di testi di canzoni protetti dal diritto d'autore ai fini dell'addestramento e per il funzionamento di modelli di GenAI negando la possibilità di ricorrere, nella specie, all'eccezione sul text and data mining.

7 L'IA e l'educazione dei giovani

Mauro Giusto

7.1 Il paradosso della generazione più connessa ma meno consapevole

I giovani di oggi (dai 10 / 25 anni) sono i primi a non ricordare un mondo senza algoritmi, senza assistenti vocali, senza feed personalizzati che selezionano per loro la realtà. In Italia e in Europa, l'uso dell'IA generativa tra bambini e adolescenti sfiora il 90% nella fascia 9-16 anni. Eppure, sono anche la generazione che meno comprende i meccanismi che la governano. Usano gli strumenti ma non sanno chi li usa attraverso di loro.

Tale inconsapevolezza non è un semplice problema tecnico: è una questione antropologica, politica e soprattutto etica fondamentale. E' esattamente il punto da cui parte la riflessione di Padre Paolo Benanti, francescano, teologo e docente di Etica delle tecnologie alla Pontificia Università Gregoriana, presidente della Commissione IA per l'Informazione del Governo italiano, unico nostro connazionale nel Comitato ONU sull'Intelligenza Artificiale, consigliere degli ultimi due Papi su questi temi. La sua posizione è molto chiara e al tempo stesso dirompente: non è la macchina in sé. Il problema vero è lasciare i ragazzi soli di fronte a quello strumento e alle sue dinamiche.

Se li immergiamo nel digitale fuori da relazioni umane significative, abbiamo già deciso implicitamente che non ci interessa educarli davvero.

Attorno a questa convinzione si è costruita, negli ultimi anni, una costellazione di pensatori, istituzioni ed esperienze educative che convergono su un punto centrale: l'intelligenza artificiale non può prescindere dall' etica, e i giovani non possono essere lasciati senza mezzi per capirla e per abitarla in modo libero e responsabile.

7.2 Non è solo tecnologia: è una trasformazione dell'umano

Il primo errore da evitare quando si parla di IA e giovani è di ridurre la questione a un argomento tecnico — di competenze, di alfabetizzazione informatica e di coding. La

rivoluzione dell'intelligenza artificiale non è soltanto tecnologica quindi è una trasformazione antropologica, culturale e sociologica. Non investe solo ciò che i ragazzi fanno ma cambia ciò che essi sono e come si percepiscono. Come costruiscono la propria identità, come si relazionano agli altri, come imparano, come si emozionano e come immaginano il futuro, proprio e del mondo.

Per comprendere la portata di questo cambiamento, toniamo al pensiero di padre Benanti. Esiste quindi una categoria filosofica nuova quanto definita: la specie tecno-umana. L'essere umano non ha mai affrontato il mondo a mani nude. Da decine di migliaia di anni lo modifica attraverso le idee che sa realizzare — il fuoco, la scrittura, la stampa, il vapore, l'elettricità, il digitale — e questi artefatti, a loro volta, trasformano l'essere umano stesso. La tecnologia non può essere neutra ma porta con sé una visione dell'universo, valori, un insieme di poteri e di rischi. Ogni nuovo strumento muta i confini di ciò che è possibile e, facendo ciò, cambia anche chi lo sa usare.

Questo pensiero, che ha profonde radici nella tradizione filosofica della tecnica, a partire da Aristotele, trova nella figura dell'algoritmo la sua incarnazione attuale più radicale. Un algoritmo non è semplicemente un codice informatico, è una disposizione di potere e una forma d'ordine. Ogni sistema di IA che entra in un contesto sociale ridistribuisce visibilità, accessi, gerarchie, opportunità. La visibilità di un contenuto, di un'opinione, di un candidato, di un prodotto non dipende più solo dalla sua qualità o dal suo merito, dipende dalla posizione che l'algoritmo gli attribuisce. Questo è un potere che agisce, nella maggior parte dei casi, in modo invisibile ai suoi stessi utenti ed è quindi particolarmente insidioso nei confronti degli umani più giovani, che non hanno ancora gli strumenti per riconoscerlo.

I giovani sono immersi in questo potere fin dall'infanzia e il sistema formativo non svela loro i rischi oltre alle potenzialità.

7.3 È necessaria una bussola etica per navigare il mondo degli algoritmi

C'è un neologismo ormai entrato nel vocabolario internazionale: algoretica (Benanti). La parola è nata dalla fusione di algoritmo ed etica, ed è riconosciuta dall'Accademia della Crusca e adottata dall'ONU nel proprio lessico. Tale parola nasce da una constatazione tanto

semplice quanto allarmante: se una macchina può decidere se concedere o negare un mutuo, se un algoritmo può orientare la pena in un processo penale, se un sistema automatizzato può selezionare le informazioni che un ragazzo riceve ogni giorno e plasmare silenziosamente la sua visione del mondo, allora quella macchina non esegue semplicemente un codice. In questo caso essa agisce su valori etici prodotti dall'uomo ma questi valori devono essere resi computabili in modo consapevole e responsabile, non occultati.

L'algoretica non attribuisce alla macchina una coscienza morale: la macchina non è un soggetto etico. Essa riconosce che i sistemi algoritmici incorporano valori, criteri e gerarchie che sono stati scelti — più o meno consapevolmente — da esseri umani. Ogni qualvolta che un algoritmo sceglie cosa mostrare a un ragazzo di quattordici anni, sta operando una scelta di valore: cosa è rilevante, cosa è desiderabile, cosa è normale. E quella scelta è stata fatta da qualcuno che molto spesso, non è l'adolescente, né la sua famiglia né la sua scuola, ma una corporate tecnologica che risponde anzitutto alle proprie logiche di profitto.

Per questo, l'algoretica, ormai per molti, non è solo una questione per ingegneri ed esperti: è una competenza di cittadinanza che deve essere insegnata fin dalla scuola con l'obiettivo preciso di evitare che ogni individuo — e in particolare ogni giovane — rinunci alla propria coscienza critica lasciando alla macchina di agire per suo conto, perdendo così il senso di ciò che fa e di ciò che è.

7.4 L'abbandono educativo

Non è quindi la macchina il nemico. Il vero problema è l'abbandono educativo: lasciare cioè i ragazzi soli di fronte a strumenti di inusitata potenza, in assenza di adulti capaci di affiancarli in questo confronto.

Una macchina che dà sempre ragione, che risponde sempre, che non si stanca mai, e non litiga e non si arrabbia, è una tentazione infinita per un adolescente che sta costruendo la propria identità e cerca conferme. Ma questa tentazione diventa una trappola quando non c'è nessuno — genitore o educatore — che aiuti il ragazzo a capire cosa stia cercando realmente, o perché lo stia cercando in un sistema tecnologico invece che in una relazione umana.

La vera domanda da porsi è perché un ragazzo non ha qualcuno con cui confrontarsi. Il tema è che quel minore nel pieno della crescita trova nelle risposte della AI l'unica fonte

disponibile. Chiaramente questo è un tema che non riguarda la tecnologia ma il ruolo e le responsabilità degli adulti.

I problemi di relazione e di crescita non si risolvono con le norme. Sono uno spazio esclusivamente umano che gli adulti devono interamente condividere con i più giovani. Ciò richiede fatica, dedizione e impegno. Quando ciò non accade i ragazzi rischiano di pagare a caro prezzo e in solitudine, quella presenza che gli adulti non riescono a garantire.

7.5 La macchina risponde sempre: rischi per l'identità

L' AI generativa entra nella vita dei minori nel momento più delicato della loro crescita, quello in cui si forma l'identità.

L'adolescenza è per sua natura il tempo del conflitto, del confronto, dell'accettazione e del rifiuto. È il momento in cui si imparano a gestire i fallimenti, a negoziare prospettive diverse da quelle originarie e a sviluppare resilienza ed empatia attraverso relazioni reali. Le relazioni umane comportano infatti complessità e scontro, e attraverso questi passaggi si forma il carattere della persona. E' un percorso complicato ma necessario.

Un sistema di IA, per come è concepito, tende a fare l'opposto: asseconda, conferma, si adatta all'utente, offre accettazione incondizionata. Non espone mai al rifiuto. Non provoca mai disagio né esclusione. Non mette mai in campo il confronto con una realtà che resiste. Per un adolescente fragile, questa può sembrare una via molto agevole. Ma, di fatto, è la privazione dell'esperienza del limite, indispensabile per la maturazione psichica.

Gli esperti chiamano soggettivazione algoritmica il processo per cui i giovani interiorizzano criteri algoritmici di valore, i like, i punteggi, la visibilità usandoli per costruire il proprio sé in pubblico. L'identità quindi non si forma più solo attraverso l'incontro con altri esseri umani, ma tramite un dialogo continuo con agenti artificiali che modellano la soggettività, filtrano l'esperienza, rimandano un'immagine sempre confermativa e rassicurante di chi già si è. Purtroppo il risultato può essere un'identità autoreferenziale, incapace di reggere il confronto con la realtà e la sua inevitabile, ma necessaria per la crescita, resistenza.

In casi sempre più frequenti, alcuni giovani sviluppano quello che in psicologia si definisce "relazione affettiva surrogata" con le macchine, attribuendo loro qualità umane come rispetto reciproco, empatia e comprensione profonda. Una seduzione perfetta e perciò stesso molto

pericolosa, perché allontana dal rischio imprescindibile della relazione reale. Ma quel è il risultato a lungo termine, se un giovane costruisce la propria idea dell'altro attraverso un sistema che non sbaglia mai, non si stanca, non mostra mai le sue fragilità? Crescerà cercando nel mondo reale interlocutori "perfetti", essendo incapace di tollerare l'imperfezione strutturale di ogni relazione umana autentica.

7.6 La sfida dei deep fake: il ruolo di AGCOM

A ciò si aggiunge una dimensione politica altrettanto rischiosa l'intelligenza artificiale attraverso i motori di ricerca, i sistemi di raccomandazione, i feed personalizzati dei social media non si limita a dialogare con i giovani, li forma. Forma la loro visione del mondo, le loro opinioni e il confine tra ciò che considerano normale e ciò che considerano non accettabile.
Si tratta di ciò che Benanti ha definito "sharp power", cioè la capacità dell'IA di modificare comportamenti senza cambiare esplicitamente le idee. La propaganda digitale condiziona, non ha bisogno di convincere. Seleziona ma senza argomentare. Non ha bisogno di persuadere apertamente, ma costruisce ambienti informativi chiusi (le cosiddette camere d'eco) in cui certi punti di vista appaiono universali e altri sembrano totalmente assenti. Un adolescente che vive in una camera d'eco non si rende conto di viverci: la sua bolla gli sembrerà il mondo. Questa confusione tra bolla e realtà è uno dei rischi più gravi per la costruzione di una coscienza critica e democratica.
Aggiungendo a questo la sfida dei "deepfake" — immagini, video, audio sintetici sempre più realistici, generati da sistemi di IA — viene minata la capacità stessa di distinguere il reale dal falso. In questa terribile sfida un soggetto istituzionale come Agcom potrà avere un peso determinante in termini etici ed educativi, denunciando e stigmatizzando questi sistemi distorsivi in grado di destrutturare ulteriormente le dorsali di una comunicazione corretta. Una comunità composta da cittadini che non sanno più distinguere un documento autentico da uno costruito artificialmente rischia di divenire rapidamente un luogo dove la democrazia è gravemente compromessa.

7.7 Il ruolo della scuola

Di fronte a tutto questo scenario, la scuola è chiamata a fare un importante salto di qualità che va ben oltre l'introduzione del coding nel curriculum o l'uso dei tablet in classe. E' una sfida cruciale: formare giovani capaci di pensare critico la tecnologia, di interrogarla invece di subirla, di riconoscere in essa i valori impliciti che porta con sé. Di considerarla in ultima analisi uno strumento.

L' istituzione scolastica dovrà aumentare per i giovani la comprensione di che cosa sia veramente l'intelligenza artificiale ed è un compito primario delle realtà educative. L'IA può essere uno strumento efficacissimo per personalizzare l'apprendimento, ed espandere le capacità degli insegnanti di adattarsi alle necessità specifiche di ogni singolo studente, rendendo la conoscenza più accessibile e inclusiva, abbattendo al contempo alcune delle barriere che rendono l'apprendimento disuguale. Ma senza una visione etica e regole ben definite i frutti potrebbero al contrario essere l'esclusione, la standardizzazione, la profilazione.

Va tenuto ben presente che già nelle Linee guida del Ministero dell'Istruzione e del Merito (2025) viene riconosciuta questa urgenza e vengono identificati tre principi fondamentali:

L'essere umano al centro — l'IA deve restare a servizio dell'uomo, non sostituirne la responsabilità —- consapevolezza — docenti e studenti devono conoscere i meccanismi algoritmici per evitarne un uso passivo o inconsapevole,

responsabilità — chi utilizza l'IA deve rispondere della adozione delle proprie scelte in termini di equità e impatto educativo.

D'altronde anche Il "Regolamento europeo sull'Intelligenza Artificiale" (AI Act, 2024) si muove nella stessa direzione, classificando i sistemi di IA usati in ambito educativo tra quelli ad alto rischio, proprio perché capaci di influenzare percorsi di crescita e opportunità di vita.

Emerge comunque che i problemi di crescita e di relazione non sono solo questioni normative e non si possono risolvere con circolari ministeriali o regolamenti, per quanto ben strutturati.

È richiesta in realtà una forte presenza di adulti: genitori, educatori, docenti in grado di affiancare anche fisicamente i ragazzi non unicamente attraverso l'accesso agli schermi.

7.8 Educare all'algoretica: imparare a fare le domande giuste

È necessario per tradurre l'algoretica in pratica educativa significa oprare su più livelli contemporaneamente. È necessario non ricondurre tutto né alla sola tecnica né tantomeno alla sola etica astratta.

La conoscenza è il primo livello: comprendere cos'è un algoritmo, come esso apprende dai dati, chi lo progetta e con quale scopo, chi ne è responsabile quando produce danni o sperequazioni. Per questo serve capire i meccanismi di fondo per non essere ingenuamente oggetto dei loro effetti. Ma soprattutto per poter fare le domande giuste nel luogo giusto.

Il "discernimento critico" è Il secondo livello: consiste nella capacità di riconoscere la profilazione, le camere d'eco, la manipolazione emotiva attraverso i contenuti personalizzati e i deepfake. Distinguere la comprensione genuina dalla simulazione di comprensione. Riconoscere quando la macchina sta assecondando invece di

aiutare a crescere. In altri termini quando la facilità offerta dall'IA è un dono e quando è una trappola.

La "responsabilità e la partecipazione" sono il terzo livello. Essere cittadini attivi nella costruzione di regole per l'IA, non semplici utenti passivi è determinante per non incorrere nei rischi di passività e manipolazione. Il tema centrale è saper formulare le domande giuste nella forma e nel contenuto.

Riconoscere il valore insostituibile della relazione umana è probabilmente il tema centrale a cui devono tendere gli altri livelli.

Capire infatti che ciò che un algoritmo può simulare — ascolto, empatia, comprensione e presenza — può non avere la natura degli stessi elementi se trasposti nel mondo reale. Questa distinzione è centrale per individuare cosa è realmente un essere umano e cosa gli serva per crescere bene. Al di là di algoritmi e macchine.

7.9 La governance pubblica. Una questione politica

Fra i quesiti più complessi tra IA e giovani concerne i limiti d'età nell'accesso alle piattaforme. Appare ormai evidente come un limite è assolutamente necessario. Al tempo stesso il problema si sposta su un piano più urgente: chi controlla realmente l'età del minore? E'

possibile davvero lasciare alle grandi corporation tecnologiche il potere di costruire un'anagrafica di tutti gli utenti ancora più dettagliata e pervasiva di quella pubblica?
Sarà necessario elaborare un sistema che possa verificare l'età degli utenti senza dover chiedere ai grandi gruppi privati dati anagrafici reali che potrebbero mettere in pericolo la protezione e la privacy dei minori. In altri termini ci deve essere un controllo trasparente pubblico e potenzialmente sottratto alla logica del profitto. E qui si torna ad una visione in cui la Governance è più importante della tecnica. E che la governance stessa appartiene agli stati democratici e non alle aziende. Nessuna delega quindi al mercato su decisioni che riguardano il futuro, la formazione e l'identità dei giovani.

7.10 Conclusioni per un approccio consapevole all'AI

Un chiarimento, capire come funziona un algoritmo è necessario ma non basta. Il problema non è solo cognitivo ma di educazione all'etica.
Un problema esistenziale e relazionale, che investe il modo in cui i giovani stanno nel tempo, nelle relazioni e in sé stessi.
I giovani vivono in un tempo in cui la velocità è un valore assoluto, in cui la risposta è sempre immediata, in cui l'attesa è percepita come un difetto e il silenzio come un'assenza da colmare. I sistemi di AI soddisfano perfettamente questi tempi. La vita reale, però, è fatta di relazioni profonde, di scelte e di crescita attraverso l'errore e il confronto non si può comprimere.
Una scuola che guardi all'algoretica dovrà avere come valori centrali l'educazione alla lentezza, alla domanda, al dubbio produttivo.
Una parte dell'insegnamento non dovrà essere delegato alle macchine, non perché le macchine non siano capaci ma perché la responsabilità di quelle decisioni appartiene solo all'essere umano. Essa individua come imprescindibile le relazioni tra insegnante e studente, tra coetanei, tra generazioni. Si salvano così valori che nessun algoritmo potrà replicare.
Tenere la persona al centro della riflessione sull'AI, non significa essere contro la tecnologia e non si tratta di rammarico per il non ritorno al mondo analogico, si tratta invece di ricordare, con insistenza, che la tecnologia è fatta dagli esseri umani per gli esseri umani e

non il contrario. Una scuola che formi con l'etica al centro della pedagogia è una scuola che consegna ai giovani non solo competenze ma domande e direttrici di riflessione.
Non solo risposte immediate ma la capacità di incontrare l'altro sapendo che nessuna macchina potrà mai replicare quell'incontro.

Fonti principali: Paolo Benanti, "Human in the Loop" (LEV, 2018); Paolo Benanti e Sebastiano Maffettone, "Noi e la macchina. Un'etica per l'era digitale" (LUISS University Press, 2023); Rome Call for AI Ethics (Pontificia Accademia per la Vita, 2020); UNESCO Recommendation on the Ethics of Artificial Intelligence (2021); Luciano Floridi, "The Ethics of Artificial Intelligence" (Oxford University Press, 2023); Ministero dell'Istruzione e del Merito, Linee guida per l'introduzione dell'IA nella scuola (2025); Regolamento europeo sull'Intelligenza Artificiale / AI Act (2024).

8 Il Mosaico delle Regole nell'Ecosistema Digitale*

Giovanna de Minico**

Sommario: 1. Linee del ragionamento. - 2. La tecnica definisce gli elementi incerti della condotta anticompetitiva. - 3. La tecnica e i mercati digitali. - 4. La tecnica da autoregolazione a fonte del diritto. - 5. Le condizioni di legittimità dei codici-fonti. - 6. I codici previsti nel DSA rispettano il modello di compatibilità costituzionale? - 7. L'Artificial Intelligence Act e la sua pretesa di orientare la tecnica verso l'uomo. - 8. La tecnica come "source of law". - 9. Verso il futuro.

8.1 Linee del ragionamento

Il diritto e la tecnica si sono incontrati, intrecciati, confusi, in un rapporto mutevole che ha permesso a ciascuno di prendere il posto che prima era dell'altro, come era reso inevitabile dalla dialettica politica che in esso si esprime e mai ingessa i suoi attori in posizioni predeterminate.

La tensione della tecnica a emancipare l'uomo era presente già in Eschilo, che racconta del furto da parte di Prometeo del fuoco divino per liberare gli uomini dall'ignoranza e permettere loro di competere con gli Dei[15]. Il fatto che Prometeo muoia incatenato alla roccia del Caucaso mostra però l'incapacità della tecnica di mantenere la promessa iniziale: portare l'uomo alla felicità.

Aristotele recupera l'intuizione di Eschilo, ma la sviluppa con mezzi autonomi perché conserva ancora l'obiettivo di rendere l'uomo felice, ma lo persegue con le virtù dianoetiche

* Il tema è stato da noi trattato in una prospettiva più ampia nella *Relazione*, *Unione europea – Mercato – Tecnica*, svolta al XL Convegno annuale dell'Associazione Italiana dei Costituzionalisti "L'Unione europea a confronto con la Costituzione della Repubblica italiana", Torino 10-11 Ottobre 2025, pubblicata in https://www.associazionedeicostituzionalisti.it/images/convegniAnnualiAIC/2025_Torino/Giovanna_De_Minico.pdf, mentre il saggio definitivo è in corso di pubblicazione nella *Rivista AIC* come Atti del Convegno.

** Professoressa ordinaria di Diritto costituzionale e pubblico, Università degli Studi di Napoli "Federico II"; *legal chief* dei Partenariati pubblico/privati "FAIR" e "Restart", finanziati dall'Unione europea.

15 ESCHILO, *Prometeo incatenato* (trad. it. a cura di L. MEDDA), Milano, Mondadori, 1994.

– principalmente, etica, saggezza, intelletto[16]. Queste nutrono la politica, che, promossa ad "architettura di sistema"[17], diventa criterio ordinatore delle altre scienze e, tra queste, anche della tecnica, emancipando l'uomo dalla condizione di soggezione.

Nella visione aristotelica la tecnica è servente la politica perché, in quanto mezzo, non può fare quanto è nella disponibilità del fine, la politica. Il rapporto mezzo/fine dovrebbe riservare alla tecnica una posizione inevitabilmente ancillare, da seconda, pronta a seguire chi viene prima di lei. La politica invece deve precedere il sapere scientifico guidandolo verso il bene comune[18]. Tra i due termini corre una relazione di proporzionalità inversa per cui al ridursi dell'uno, l'altro si dilata[19].

La sequenza principale/accessorio rispetta la natura della politica, arte della qualità e, al tempo stesso, asseconda l'opposta genesi quantitativa della tecnica, risolvendosi nel riconoscere alla politica una posizione primaria rispetto alla grandezza numerica[20]. Proveremo a capire se questo ordine di intervento tra politica e tecnica abbia resistito nel tempo e, in caso di risposta negativa, ci chiederemo se sia ancora sostenibile la tesi della neutralità tecnica, cioè di un sapere adespota, senza padrone, che si aggira privo di un obiettivo da conseguire[21].

Vedremo insieme in che modo la politica e la tecnica si siano combinate in una misura di coesistenza dalle proporzioni variabili[22]; cosa c'è dietro questa mescolanza e su quale gradino della scala si è attestato il loro precario equilibrio. Ragioneremo intorno a come disegnare un corretto rapporto tra tecnica e diritto, in grado di garantire i principi basilari dell'ordine democratico: responsabilità politica, *rule of law,* e protezione *erga omnes* delle

[16] ARISTOTELE, *Etica Nicomachea,* VI, 3 (1139b–1140).

[17] ID., *Etica Nicomachea*, I, 2 (1094a–b).

[18] U. GALIMBERTI, *Psiche e techne. L'uomo nell'età della tecnica*, Milano, Feltrinelli, 2021, 250.

[19] G. F. PIZZETTI, *Decisione politica ed* expertise *tecnico*, in G. DE MINICO – M. VILLONE (a cura di), *Stato di diritto. Emergenza e Tecnologia*, Milano, Consulta OnLine, 2020.

[20] F. SALMONI, *Il difficile equilibrio del diritto costituzionale tra presunta neutralità della tecnica e scelte politiche*, in *Justiça do Direito*, 2019, 33, 2, 152.

[21] S. ROMANO, *Mitologia giuridica*, in *Frammenti di un dizionario giuridico*, Milano, Giuffrè, 1947, 126 ss., in part. 134.

[22] N. IRTI - E. SEVERINO, *Dialogo su diritto e tecnica*, Bari, Laterza, 2001.

libertà fondamentali. E, infine, ci chiederemo quale ruolo può ancora essere riservato al soggetto politico sovranazionale[23].

8.2 La tecnica definisce gli elementi incerti della condotta anticompetitiva

La tecnica interviene sulla norma, che il Legislatore antitrust ha intenzionalmente lasciato indefinita, per compierla in modo da consentire al precetto elastico di adattarsi senza interruzioni al mutamento della situazione tecnico-economica, mantenendo invariato il suo tenore letterale.

Infatti, il precetto è avaro di indicazioni circa l'abuso, la posizione dominante e l'alterazione consistente dell'equilibrio competitivo[24], lasciando che sia l'Autorità antitrust a dirlo *case by case*. A questa spetta infatti un'operazione preliminare, ma essenziale per applicare la norma alla concreta condotta sottoposta al suo esame: definire l'indefinito, completare quanto la *Lex mercatoria* ha lasciato intenzionalmente incompiuto.

In che modo l'Antitrust compone il suo parametro di riempimento? Il ricorso alle discipline aziendalistiche non basta in quanto, essendo basate su una causalità meccanicistica, pervengono a risultati opinabili, privi della necessaria oggettività scientifica, il cui massimo di affidabilità è dato dalla professionalità di chi è chiamato a interpretare. Quando la valutazione di un fatto si basa su un parametro tecnico di non sicura acquisizione, il risultato della valutazione è una "questione scientifica controversa"[25], per la presenza di ineliminabili fattori di incertezza dovuti a carenze nelle conoscenze scientifiche" al punto da non potersi escludere che, sostituito l'interprete, si pervenga a esiti opposti a quelli iniziali perché le decisioni risentono del modo di pensare di chi definisce in via preliminare il parametro valutativo.

[23] U. GALIMBERTI, *Psiche e techne. L'uomo nell'età della tecnica*, cit., 250-262.

[24] M. CLARICH, *Per uno studio* sui *poteri dell'Autorità garante della concorrenza e del mercato,* in F. BASSI-F. MERUSI (a cura di), *Mercati e amministrazioni indipendenti,* Milano, Giuffrè, 1993, 123.

[25] L. VIOLINI, *Le questioni scientifiche controverse*, Milano, Giuffrè, 1986, 136.

Data l'insufficienza delle regole endogene, l'Antitrust dovrà scegliere cosa intende per equilibrio competitivo[26]. Tre tesi potranno aiutarla, ma la scelta dell'una o dell'altra dipenderà solo dal suo insindacabile volere: la prima predilige la protezione dei consumatori su ogni altro interesse; la seconda protegge i concorrenti, già presenti, cioè il mantenimento dello *status quo*; la terza guarda avanti e garantisce l'elasticità futura del mercato a favore dei concorrenti che verranno. A seconda del modello competitivo scelto, l'Autorità valuterà la condotta in termini di illiceità, vietandola; o di conformità al diritto antitrust, ammettendola. Stante questa variabilità di giudizio sulla stessa condotta, quale definizione riservare ai poteri dell'Antitrust? E, quindi, come risolvere l'alternativa tra il paradigma della funzione para-giurisdizionale e un archetipo inedito di potere?

Il modello della *iuris dictio* ha uno sguardo unidirezionale, che si compie nell'acquisire un fatto all'astratta previsione di legge secondo l'equazione "se c'è A, ci deve essere B", il che trascura il momento precedente questo confronto: definire il parametro in base al quale giudicare la condotta. Ebbene, mentre per il giudice il parametro normativo, già perfetto, è oggetto solo d'interpretazione, per l'Antitrust è l'inizio di un'attività creativa del diritto, in quanto è chiamata preliminarmente a riempire gli spazi lasciati in bianco dal legislatore con la sua energia volitiva, le sue scelte di valori. Ciò rivela la sua attitudine all'agire politico, non è un caso che l'Antitrust attinga dalla stessa borsa del decisore politico gli strumenti del mestiere: acquisizione illimitata di dati, informazione circolare con gli organi di vertice, disponibilità al dialogo con la base[27].

Riconoscere che la norma antitrust abbia già un vincolo di fine quando compie gli elementi imperfetti non contraddice l'accostamento da noi proposto dei poteri in oggetto alla funzione di indirizzo politico, che, per la lettura costituzionalistica, ricorre anche in presenza di una limitata libertà negli obiettivi, purché sia riservato alla prerogativa politica specificare tempi, modi e articolazioni[28].

[26] M. LIBERTINI, *Concorrenza* (voce), in *Enc. Dir., Annali* III, 2010, 191 ss.; R. BORK, *The antitrust paradox: a policy at war with itself*, NY, Basic Books, 1978 e R. A. POSNER, *Antitrust law*, Chicago, Univ. of Chicago Pr., 1976 (2ª ed. 2001).

[27] M. MANETTI, *Poteri neutrali e Costituzione*, Milano, Giuffrè, 1994, 197.

[28] M. DOGLIANI, *Indirizzo politico. Riflessioni su regole e regolarità nel diritto costituzionale*, Napoli, Jovene, 1985, 201 ss.

La diversità tra il decisore politico e l'Antitrust non è tanto nel conflitto 'assolutezza contro relatività della libertà di scelta', quanto nella fonte da cui i due termini traggono i necessari suggerimenti per comporre l'identità del fine. L'Antitrust non si deve lasciar guidare dalla volontà politica di maggioranza, a differenza del decisore politico, che compone il *common good* proprio secondo il sentire politico della maggioranza di turno. L'Antitrust, creata per essere indisponibile al comando dell'*élite* politica, nel chiudere il parametro di giudizio non si deve inclinare verso nessuna delle forze in campo: si deve mantenere equidistante dalle famiglie politiche egemoni, come dai regolati forti. Anche qui l'accostamento non porta a una perfetta coincidenza tra la funzione di ordine dell'Antitrust e l'indirizzo politico del soggetto rappresentativo.

Pertanto, possiamo dire che siamo dinanzi a una figura ibrida di potere, che condivide con la giurisdizione l'attività di giudizio di una condotta in base a un parametro di legge; ma che al tempo stesso è innovativa nel modo in cui l'Antitrust compie il parametro di giudizio, perché spende le sue energie volitive come farebbe un decisore politico nel comporre obiettivi e nel programmare i mezzi della sua futura azione politica. Per una compiuta definizione dei poteri in esame si dovranno dunque combinare i due momenti: di *iuris latio* nel costruire il parametro e di *iuris dictio* nell'applicarlo al caso sottoposto al suo esame. In sintesi, è una funzione assistita da un elevato grado di politicità, attenuato dalle sintetiche e avare indicazioni di legge[29].

Chiediamoci quale effetto abbia tale funzione, in parte accostabile a quella di indirizzo politico, sull'equilibrio istituzionale.

Visto che il silenzio del legislatore ha scaricato sulle Autorità indipendenti[30] la responsabilità di mediare tra i valori in gioco, appena elencati, ma non già sviluppati in una dimensione di reciproca coesistenza, ai nuovi autori della giuridicità è stato affidato il compito di chiudere la vertenza sugli "équilibres souhaitables" in via originaria, mancando di parametri valutativi incontrovertibili per comporre la regola pacificatrice.

[29] Ci sia consentito il rinvio alle riflessioni argomentate in: *Antitrust e Consob. Obiettivi e Funzioni*, Padova, Cedam, 1997, 277-280.

[30] D'ora in avanti con l'acronimo di: A.I.

La politica generale degli eletti si è progressivamente ritirata per lasciare spazio alla politica settoriale dei tecnici, i quali, pur se non sempre graditi ai cittadini, non sono revocabili *medio tempore*, perché il loro incarico non ha titolo in un mandato politico, il che li sottrae al giudizio di responsabilità politica per le decisioni adottate.

Questa sostituzione intersoggettiva ha spezzato il circuito politico-rappresentativo Popolo–Parlamento–Governo, introducendo un elemento spurio – l'Autorità indipendente – che ha posto in tensione la sovranità popolare, pietra angolare dell'edificio repubblicano. E la sovranità popolare, variabilità di significati a parte, resta indissolubilmente legata all'idea che il rapporto politico-rappresentativo sia l'unico e insostituibile canale di legittimazione della decisione politica[31].

Siccome è giusto riconoscere nella centralità della sovranità popolare un limite alla revisione della Costituzione, nemmeno il ricorso all'art. 138 Cost. potrebbe ricondurre alla legittimità costituzionale il quadro fin qui definito.

La *issue* non è 'se' le A.I. siano compatibili con l'architettura di base, l'interrogativo verte piuttosto intorno al 'se' tale prerogativa politica, assegnata a chi non è neppure indirettamente riconducibile alla volontà popolare, sia in linea con l'assetto istituzionale. Neppure la norma di revisione potrebbe legittimamente fare ciò, perché ammettendolo opererebbe una rivisitazione così rilevante dell'architettura di fondo del sistema da superare i limiti del potere di revisione, in quanto altererebbe i tratti identitari della Repubblica democratica.

Forse, il dibattito dottrinario non ha colto sempre, e a fondo, le implicazioni in termini costituzionalistici dei nuovi modi di produzione del diritto. La questione è stata ridotta – quanto alle A.I. – nei termini di una loro previsione in Costituzione, ma riconoscere a esse visibilità costituzionale[32] non è sufficiente a giustificare il premio delle prerogative politiche riconosciuto loro.

---

[31] E. CHELI, *Atto politico e funzione di indirizzo politico*, Milano, Giuffrè, 1961, 93.

[32] Due gli orientamenti emersi in sede di Bicamerale D'Alema: quello minimalista, che inseriva nel testo il meno possibile, salvo l'esistenza stessa delle Autorità. In antitesi, la tesi della piena costituzionalizzazione dello statuto delle Autorità, tipizzate per ambiti d'intervento e relative attribuzioni. Quest'ultimo metodo rischiava di ingessare un fenomeno, che invece si deve mantenere mobile in ragione delle esigenze storico-politiche del momento: A. ROUYÈRE, *Faut-il faire*

Né basterebbe per una conclusione diversa completare il conferimento della funzione politica con una previsione costituzionale del contraddittorio con i regolati.

E ciò per due motivi: la diversità ontologica della rappresentanza politica in termini di presupposti, natura ed effetti rispetto a quella di interessi esclude che si possa trarre dalla prima una legittimazione trasferibile alla seconda, per la semplice ragione che questa legittimazione è nella disponibilità incedibile della prima e non anche della seconda.

Poi se si considera il circuito Popolo-Parlamento-Governo, come l'unico canale capace di trasmettere il titolo abilitante alla scelta politica, esso è un tratto identitario del nostro ordinamento, e, come tale, non ammette surrogati neppure se disegnati dalla norma di revisione. In sintesi, il dialogo delle A.I. con la base non può supplire alla mancanza di una disciplina sostanziale a livello primario.

Vedremo nelle conclusioni una possibile idea per superare questa *impasse* istituzionale, nella quale la tecnica – divenuta potere autoritario e assoluto – ha scalzato la politica, che, privata del suo linguaggio dei fini, sembra inseguire la tecnica, invece che orientarla, cedendo a quest'ultima spazi di decisione che, in un ordinamento democratico, dovrebbero restare saldamente nelle mani dei soggetti rappresentativi.

8.3 La tecnica e i mercati digitali

Quando il mercato digitale iniziò ad affermarsi, la disciplina antitrust non era ancora pronta a ricomprendere nelle tradizionali figure delle intese e degli abusi la condotta anticompetitiva degli imprenditori della *e-economy*. Ma sarà necessario ricordare la fisionomia dei mercati digitali per comprendere le ragioni della loro iniziale esclusione dall'economia tradizionale.

La tecnica, in primo luogo, diventa tratto identitario dei mercati digitali, assente per le piazze analogiche: siamo davanti a *double-sided market*, dominati da *Gatekeeper* (GK), gestori neutrali della piattaforma, attivi, non solo nel promuovere l'incontro tra domanda e offerta,

figurer les autorités administratives indépendantes dans la Constitution?, in *Les petites affiches*, 1992, 56 ss. Cfr.: M. MANETTI, *Autorità indipendenti: tre significati per una costituzionalizzazione*, in *Pol. dir.*, 4. 1997, 657 ss.; S. STAMMATI, *Le autorità di garanzia e di vigilanza*, in S. PANUNZIO (a cura di), *I costituzionalisti e le riforme*, Milano, Giuffrè, 1998, 343 ss.; e, volendo, G. DE MINICO, *Regole. Comando e Consenso*, Torino, Giappichelli, 2005, in part. *Conclusioni*.

quanto nel concorrere sul mercato a valle con altri *competitor*, costretti a rivolgersi a loro per l'uso della piattaforma[33]. Il GK è inevitabilmente un operatore verticalmente integrato, pronto a reiterare nel tempo atteggiamenti abusivi, dovuti al suo fisiologico conflitto di interessi. Non va dimenticato un effetto naturale di questi mercati: sono capaci di tenere legati i propri clienti (c.d. *effetto lock-in)*, i quali, pur potendo cambiare piattaforma, difficilmente lo faranno perché altrove non troverebbero condizioni migliori. Infatti, quando *Google* ci fornisce un servizio digitale in cambio dei nostri dati, costruisce un profilo accurato intorno a noi; e grazie a questo identikit la pubblicità sarà tagliata addosso a ogni singolo cittadino-consumatore, sarà più mirata e quindi destinataria di più *clic*; per questo motivo gli spazi pubblicitari diventeranno più preziosi e quindi vendibili a un prezzo più alto di quello che avrebbero senza profilazione.

Visto che la tecnica ha dato una curvatura inedita all'illecito anticompetitivo, come interviene sull'interpretazione della disciplina antitrust?

Se si assume la tecnica come elemento integrativo di una lettura evolutiva del dato normativo, le vecchie norme antitrust, pur lasciando invariato il loro tenore letterale, vedrebbero modificato il significato dei rispettivi precetti in modo automatico e continuativo all'avanzare della tecnica, complice anche l'elasticità delle fattispecie anticompetitive.

Inoltre, questa interpretazione tecnicamente orientata avrebbe il pregio di non lasciarsi ingannare dalle apparenze perché assumerebbe le condotte per quello che sono, cioè comportamenti ripetutamente aggressivi del mercato, anche se non presentano gli indizi classici e letterali dell'abuso di dominanza o dell'intesa vietata. Una tecnica, abbandonata a se stessa, sarebbe un pericoloso "*mean[s] to produce revenue and market control*"[34], il cui esito interpretativo impedirebbe la porosità delle norme agli illeciti inediti. Invece, una tecnica orientata al *common good* non lascerebbe le *Big Tech* indisturbate nell'accaparrarsi i dati dei cittadini, protette dall'anarchia tecnologica e dalla forza di una dominanza incontestabile dai terzi.

[33] G. PITRUZZELLA, *Le libertà di informazione nell'era di Internet*, in *MediaLaws*, 1, 2018, 23.

[34] S. ZUBOFF, *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power*, New York, Profile Books Ltd, 2019.

Qui notiamo atteggiamenti mutevoli della Commissione europea[35], sintetizzabili in tre filoni: un primo, un secondo e forse anche un terzo, in uno sguardo prospettico.

A) La CE inizialmente ritenne che le condotte delle *Big Tech*, anche se lesive della privacy, fossero fuori della sua competenza "fall [ing] out of the aim of antitrust law"[36] perché rientranti "within the scope of the EC data protection rules". Questa applicazione del diritto antitrust insensibile al mutato contesto tecnologico si è risolta nell'assumere a indici di abuso e come criteri per l'individuazione dei mercati rilevanti i tradizionali parametri quantitativi – basati sulla onerosità del negozio e sulla sostituibilità dei prodotti – validi per le piazze analogiche, ma incapienti su quelle digitali. Mentre sarebbe stato ragionevole adottare criteri alternativi per valutare l'abuso: non più l'incremento minimo e stabile del prezzo, inidoneo a spostare la domanda, visto che le contrattazioni sui mercati digitali sono a costo zero, ma la riduzione minima e stabile della qualità della prestazione[37]. Il nuovo parametro non è privo di incognite: come si calcolerà un mutamento qualitativo su un valore immateriale, quale è la privacy? E, ancora, il metro di valutazione sarà uguale per tutti i soggetti; oppure, andrebbe misurato in ragione delle tendenze dei clienti e della loro diversa sensibilità alla protezione dei dati personali?

B) La seconda fase vide la CE incline ad atteggiamenti più flessibili perché si era resa conto che l'abuso poteva avere manifestazioni sintomatiche, che, quando fu scritto l'art. 82 TCE (poi divenuto art. 102 TFUE), non erano neppure pensabili. Aver accettato la tecnica come elemento integrativo di un'ermeneutica evolutiva – che preferisco chiamare lettura *constitutional by design* per la sua capacità di incorporare il principio di uguaglianza – significava ammettere che la tecnica avesse riscritto nel silenzio l'illecito antitrust, anche se non lo aveva fatto, perché si era insinuata negli interstizi, negli spazi lasciati incompiuti dal legislatore comunitario, acquisendo ai divieti condotte fortemente compromissorie

[35] D'ora in avanti con l'acronimo di: CE.

[36] M. VESTAGER, *Competition in a BD world, 18 January 2016*, in *https://ec.europa.eu/commission/presscorner/detail/en/speech_16_5224*.

[37] M.E. STUCKE - A.P. GRUNES, *Big data and competition policy*, Oxford, OUP, 2016, 115 ss.

dell'equilibrio competitivo, che altrimenti sarebbero rimaste nascoste nel cono d'ombra delle norme, regalando alle *Big Tech* un'impunità permanente.

b.1) Gli esempi sono offerti dalla recente Comunicazione sui mercati rilevanti del 2024[38], nella quale la CE, ai fini dell'individuazione del mercato rilevante, abbandona il tradizionale criterio merceologico a favore di un metodo modellato sulle caratteristiche delle piazze digitali, dove le *Big Tech* forniscono un pacco inclusivo di prestazioni plurime. Il che ha convinto la CE ad avere uno sguardo, non più atomistico sul mercato, ma a tutto tondo, obbedendo alla tensione economica verso l'ecosistema digitale.

b.2) Mentre per quanto attiene ai sintomi dell'abuso, nei recenti casi di sanzioni contro *Google* o *Apple* si apprezza che la CE, persa la fiducia nella rigida separazione delle competenze – secondo la quale ciò che è di attribuzione dell'Autorità privacy è fuori dalle sue preoccupazioni – ha adottato un approccio molto *reasonable* e rispettoso della regola di assegnare a ciascuno il suo. L'Antitrust coglie la lesione della privacy come *vital clue*[39], ma non lo assume sufficiente di per sé a perfezionare l'illecito antitrust: è solo un campanello d'allarme, salvo poi valutare se ricorra o meno l'aggressione del mercato. La CE, come la Corte di Giustizia (Grande sezione, 4 luglio 2023, *Meta Platforms Inv v. Bundeskartellamt*), non ha dato credito neppure alla suggestiva tesi dell'Antitrust tedesca, che aveva parlato di "presunzione normativa"[40], commettendo l'errore di perfezionare con automatismo i modelli di condotte lesive. Se la CE lo avesse fatto, sarebbe caduta nell'eccesso opposto: dall'assoluta

[38] Commissione europea, *Comunicazione della Commissione sulla definizione del mercato rilevante ai fini dell'applicazione del diritto dell'Unione in materia di concorrenza*, C/2024/1645, in https://eur-lex.europa.eu/legal-content/IT/ALL/?uri=CELEX:52024XC01645, 30 ss.

[39] Corte di Giustizia (Grande Sezione), 4 luglio 2023, *Meta Platforms Inc. e a. contro Bundeskartellamt*, ECLI:EU:C:2023:537, *in https://eur-lex.europa.eu/legal-content/IT/TXT/?uri=celex:62021CJ0252*. Sulla stessa linea anche il Tribunale (Sesta Sezione ampliata), 14 settembre 2022, *Google e Alphabet/ Commissione (Google Android)*, Causa T-604/1, in *https://curia.europa.eu/juris/document/document.jsf;jsessionid=6CA59A0E313E2A87C7F8782CE06C8D65?text=&docid=265421&pageIndex=0&doclang=it&mode=lst&dir=&occ=first&part=1&cid=13231855,* in part. parr. da 284 a 305.

[40] In proposito si legga il: Bundesgerichtshof (*Federal Court of Justice*), *Decision KVR 69/19*, 23 giugno 2020 (la sua *press release* recante: *The Federal Court of Justice provisionally confirms the allegation of abuse of a dominant market position by Facebook*, in– https://www.bundesgerichtshof.de/SharedDocs/Pressemitteilungen/DE/2020/2020080.html), dove la Corte Federale, a differenza del giudice di appello, desumeva il nesso di "causalità normativa" - (art. 19(1) della Legge federale sulla concorrenza) – dalla sola condotta accrescitiva del potere di mercato del dominante. Con la conseguenza di non dover provare gli ulteriori elementi del comportamento illecito, perché era sufficiente dimostrare la "posizione di preminenza del soggetto egemone", in *https://juris.bundesgerichtshof.de/cgi-bin/rechtsprechung/document.py?Gericht=bgh&Art=en&Datum=2020-6&Seite=4&nr=109506&pos=121&anz=279*.

irrilevanza della privacy alla sua assorbenza su ogni altra valutazione inerente all'equilibrio competitivo; invece, ha preferito un approccio aderente ai fatti, mutuati dalla condotta aggressiva del mercato e, al tempo stesso, dalla lesione alla privacy.

La Corte di Giustizia (Grande sezione)[41] ha osservato che, nell'esaminare un abuso di posizione dominante, l'Antitrust può accertare anche la conformità del comportamento di tale impresa a norme diverse da quelle rientranti nel diritto della concorrenza, quali le disposizioni del GDPR[42], in quanto la loro violazione rappresenta "importante indizio [...] per valutare le conseguenze di una determinata pratica sul mercato o per i consumatori". Pertanto, continuare a ritenere che la disciplina antitrust e le regole privacy siano settori indipendenti l'uno dall'altro rischia di pregiudicare l'effettività del diritto della concorrenza, se la loro interpretazione disgiunta diventa una modalità, non episodica, ma ricorrente[43].

b.3) Il terzo step, ancora al di là da venire, anticipa segnali promettenti: essi riguardano il sistema sanzionatorio. Siccome le sanzioni pecuniarie non prevedono un *surplus* economico rispetto a quanto basti per compensare la lesione procurata, al trasgressore conviene violare la norma piuttosto che osservarla, reputando la sanzione un costo inevitabile del suo inarrestabile processo espansivo. L'assenza di pena nella sanzione ha lasciato invariata la dominanza delle *Big Tech*, non ha messo a cura dimagrante la loro acquisizione illecita di potere e la reiterazione di abusi di posizione dominante, come dimostrato dal loro continuo ripetersi nel tempo con modalità pressoché identiche.

Invece, le sanzioni dovrebbero fare un salto di qualità, se solo la politica lo volesse, e orientarsi verso misure ripristinatorie dei diritti lesi, cioè capaci di riportare il diritto a come era prima di subire la lesione. Se però la violazione ai diritti è irreversibile, la sanzione dovrà ridefinire la situazione complessiva del dominante, e modificarla fino a rovesciarla come un calzino perché è proprio quella situazione la causa di un illecito, che altrimenti non sarebbe accaduto. Il riferimento è alle sanzioni destrutturanti (un esempio è offerto dall'art. 18 del

[41] Commissione europea, 4 luglio 2023, *Meta Platforms Inc. e a. contro Bundeskartellamt*, ECLI:EU:C:2023:537, in part. il paragrafo 47.

[42] Regolamento (UE) 2016/679 del Parlamento europeo e del Consiglio, del 27 aprile 2016, (regolamento generale sulla protezione dei dati), in *https://eur-lex.europa.eu/eli/reg/2016/679/oj?locale=it*.

[43] M. VESTAGER, *Press release, Case AT.40684 - Facebook Marketplace*, Bruxelles, 4 giugno 2021, in *https://ec.europa.eu/commission/presscorner/detail/en/ip_21_2848*.

DMA), che spezzano l'operatore verticalmente integrato, ordinandogli la perdita della titolarità della piattaforma in modo da isolare dal pregresso assetto proprietario il suo ruolo di operatore sul mercato *retail*, perché è in quella coincidenza di *status* – gestore neutrale del traffico di accesso e fornitore dei servizi digitali sul mercato a valle – che si compie quel fisiologico conflitto di interessi che va eliminato.

Al momento tali sanzioni non sono mai state applicate, salvo il timido avvio nel recentissimo caso *Google AdTech* 2025 (AT40670)[44], dove la CE ha sollecitato *Google* a presentare una proposta di impegno diretto alla separazione proprietaria dell'asset-dati, o, quanto meno, all'isolamento di una parte della catena del valore per sanare in origine questo conflitto di interessi *in re ipsa*. Queste sanzioni promuoverebbero con proiezione *pro-futuro* la contendibilità dei nuovi mercati, dove si è verificato un trasferimento e una concentrazione selvaggia di dominanza, attenuando le barriere protettive con lo *sharing* dei dati o con la misura estrema della deconcentrazione dell'integrazione verticale.

C) Questa coraggiosa, appena intrapresa rivoluzione antitrust, imporrebbe alle *Big Tech* di osservare le regole, non più come un adempimento episodico, un evento occasionale, suggerito da un mero calcolo di convenienza economica – cioè solo quando coincida con i loro interessi individuali – ma perché comportamento consapevole, dettato dal doveroso rispetto delle libertà economiche e dei diritti fondamentali altrui, anche se non si accordano con i propri vantaggi egoistici.

In conclusione: i piani, un tempo separati, tra privacy e *competition* si sono ora mescolati; i beni, prima lontani e aggredibili da distinte condotte, sono ora esposti a lesioni progressive o sono tutelabili l'uno come conseguenza dell'altro; e le Autorità, un tempo incomunicabili, sono chiamate a parlare perché questo intreccio del diritto sostanziale impone un intreccio di poteri con proficua "confusione" di procedure. Dinanzi a un illecito che procura una violazione sia alla concorrenza che alla privacy, anche il suo ritorno alla legalità dovrà

[44] Commissione europea, *Press release*, 5 settembre 2025, in https://ec.europa.eu/commission/presscorner/detail/it/ip_25_1992.

riparare entrambi i beni aggrediti; diversamente, ancora un lato dell'illecito rimarrebbe senza ristoro.

Dunque, l'interpretazione tecnicamente orientata del diritto antitrust, se tenuta lontana dalle valutazioni legali tipiche, mostra vantaggi indiscutibili: è pragmatica, obbedisce al principio di leale cooperazione tra le Autorità (art. 4, par. 3, TUE), è sensibile alla *privacy concern*, è *well-tailored* alle dinamiche dell'economia digitale, acquisisce a sé condotte aggressive del mercato, che per la loro innovatività tecnologica rimarrebbero immuni dalla punizione e soprattutto rispetta la centralità dell'individuo perché accorda tutela rapida ed efficiente ai diritti lesi.

8.4 La tecnica da autoregolazione a fonte del diritto

Nelle riflessioni che seguiranno disegneremo gli attributi della *self-regulation*[45], regime giuridico e requisiti di compatibilità costituzionale, se è chiamata a integrare l'ordine giuridico a livello europeo.

In ragione della variabilità dei modelli, alcuni studiosi hanno parlato della *s.r.* in termini di un *umbrella concept*[46], ma, nonostante ciò, essa rimane un modo di essere del diritto dei privati[47], perché si tratta di regole che il destinatario si è dato preventivamente, azzerando così la distanza che separa lui regolatore da lui regolato.

L'archetipo, dovuto al pragmatismo anglosassone[48] poi fatto proprio dall'ordine giuridico europeo, segna una battuta d'arresto del diritto verticistico statale, che rompe con la legge fonte esclusiva della volontà regolatoria dello Stato per aprirsi a un diritto che inizialmente si afferma senza Stato in quanto, come insieme di regole negoziali, non aveva bisogno di chiedere allo Stato la forza dell'imperatività della norma giuridica, bastandogli la limitata efficacia *inter partes* del negozio regolatorio. Ma la complessità delle relazioni moderne ha

[45] D'ora in avanti con l'acronimo di: *s.r.*

[46] Si deve a L. SENDEN, *Soft law in European community law*, Oxford-Portland, Oregon, Hart Publishing, 2004, 110.

[47] Dovuto il richiamo a W. CESARINI SFORZA, alla cui intuizione dobbiamo la felice espressione riportata nel testo, che peraltro è il titolo del suo lavoro, *Il diritto dei privati*, Milano, Giuffrè, 1963.

[48] Per rigore sistematico è doveroso rinviare a J. BLACK, *Constitutionalising self-regulation*, in *Mod. L. Rev.*, 59, 1996, 26.

chiesto risposte giuridiche flessibili, tempestive e aderenti alla realtà in movimento; di conseguenza l'ordinamento si è lasciato dietro di sé il modello monolitico e autoritario per avvicinarsi a forme reticolari e partecipate dal basso. Lo Stato ha cominciato a cedere spazi normativi a soggetti collettivi privati, i quali si sono rivelati capaci di elaborare regole efficaci all'interno dei contesti specifici di appartenenza.

La riflessione di Santi Romano[49] ha fornito la base teorica al mutamento in atto: il concetto di autonomia dello studioso, intesa, non solo come facoltà dell'individuo, quanto come dato strutturale del diritto, è la premessa ideologica del pluralismo ordinamentale. Ogni gruppo sociale organizzato, dotato di potere normativo e capacità regolativa, può costituire un ordinamento minore *ex se*, al quale lo Stato gli riconosce un suo spazio di azione, senza lasciarlo *extra-ordinem*, ma integrandolo nel quadro sistemico, sempre che rispetti le condizioni impostegli prima. Chiariamo da subito però che l'apertura dello Stato alle fonti private non ha significato rinuncia *tout court* alla sua sovranità normativa; al contrario, lo Stato riconosce l'autoregolazione come una forma integrativa del sistema giuridico a condizione che si svolga nel rispetto dei principi costituzionali e sotto la propria supervisione. Affiora un modello di 'coregolazione', in cui autorità pubbliche e soggetti privati condividono la funzione normativa secondo una precisa divisione di ruoli e tempi di intervento. Invero, lo Stato diventa 'architetto di sistema', che delinea le strutture istituzionali, definisce gli obiettivi generali e si riserva il potere di intervento correttivo, mentre il diritto dei privati si muove con una certa leggerezza entro il terreno che lo Stato gli ha perimetrato. In questo contesto, i soggetti privati deputati a compiti regolativi devono possedere una precisa fisionomia strutturale che garantisca adeguata rappresentatività degli interessi di base, trasparenza nei processi decisionali, apertura al confronto con i portatori di interessi esterni e disponibilità al controllo pubblico.

Questa condivisione del lavoro normativo tra il legislatore e i privati ripropone la nostra questione iniziale: quale il rapporto tra politica e tecnica?

[49] S. ROMANO, *L'ordinamento giuridico*, (1ª ed., 1918), 3ª ed., Firenze, Sansoni, 1977, in part. il cap. II., *passim*, ma anche nei *Frammenti di un dizionario giuridico*, (voce *Autonomia*), Milano, Giuffrè, 1947, 29.

Possiamo ancora dire che la tecnica sia accessoria alla norma primaria, se quest'ultima ha rinunciato a un linguaggio politico, perché si è risolta nel dettare il titolo di competenza, cioè la norma sulla produzione giuridica, astenendosi dal porre specifici obiettivi da conseguire e principi cui conformare la tecnica?

Se così stanno le cose, cioè mancando un perimetro politico, il codice fa da quadro e cornice al tempo stesso, inizia il discorso normativo e lo porta a compimento. Il risultato è l'inversione dei termini della relazione: la tecnica parlerà in luogo della politica, con l'aggravante che qui la decisione è dei privati, neppure di un soggetto pubblico, pseudo-indipendente.

8.5 Le condizioni di legittimità dei codici-fonti

In questo mutato contesto il concetto di 'coregolazione' è la rappresentazione plastica della nuova architettura di sistema: pubblico e privato collaborano nel processo di costruzione del diritto, ciascuno in base alle proprie competenze e ai rispettivi tempi di ingresso e di uscita dal palcoscenico politico: in particolare, il pubblico promuove, orienta e vigila; il privato propone, sperimenta e attua[50].

Le condizioni di legittimità per promuovere l'autoregolazione a fonte di produzione del diritto oggettivo, cioè atto idoneo a generare norme astratte e generali, sono anche i requisiti minimi della sua compatibilità costituzionale. Per comodità del lettore suddivideremo queste condizioni in tre distinte categorie: soggettive, procedurali e finalistiche.

Sul piano soggettivo, il regolatore deve essere un'entità rappresentativa e democraticamente organizzata. Dunque, non ogni associazione può assumere compiti normativi: è necessario un riconoscimento legislativo che identifichi, *ex ante*, i criteri di idoneità (ad es. numero di membri, diffusione territoriale, funzione pubblica riconosciuta). Parliamo di una rappresentatività fondata su criteri oggettivi (numero di iscritti, rilevanza sociale, storicità) e valutata caso per caso, e in grado di garantire la democraticità delle decisioni del privato, che devono essere prese con il consenso della base associativa, lontane dai modelli di imposizione per mano di un'*élite* dirigente.

[50] Così: R. GRIFFIN, *Public and private power in social media governance: multistakeholderism, the rule of law and democratic accountability*, in *Transnational Legal Theory*, 14, 1, 2023, 46-89,

Da un punto di vista procedurale, il metodo per produrre la norma è cruciale: inclusivo, trasparente e fondato su un dialogo reale, e non solo formale, con le parti sociali perché solo il loro effettivo coinvolgimento è tale da garantire un processo normativo equilibrato e legittimato dalla sua diffusività[51].

Infine, la finalità dell'autoregolazione è la sua aderenza all'interesse generale, non la sua coincidenza asfittica con la protezione delle posizioni dominanti; da qui l'inammissibilità di regole elaborate con finalità esclusive, rapaci o discriminatorie. Pertanto, il procedimento regolatorio deve essere 'poroso' – aggettivo che usiamo nell'accezione di Teubner – alla partecipazione di tutti i regolati, anche di quelli che vantano posizioni avversariali rispetto ai privati che per primi hanno assunto l'iniziativa, cioè i regolati forti, *well-structured and well-funded*[52], in modo che la norma risulti sorretta da un reale *in idem placitum consensus*. Ciò significa che l'elaborazione delle regole consensuali non è riserva segreta ed esclusiva delle parti negoziali dominanti, essa invece dovrà avvenire alla luce del sole e articolarsi in un dialogo equiordinato e il più diffuso possibile tra ogni categoria di regolati dalla norma negoziata. Un diritto dei privati condiviso nel suo farsi risulterà meno odioso a chi lo dovrà osservare, perché ha contribuito a crearlo.

Altro discorso è se una genesi equilibratamente concordata tra le opposte parti del negozio normativo possa supplire al difetto di legittimazione politica del regolatore privato. Riteniamo che, se la diffusività partecipativa fosse in grado di compensare l'iniziale asimmetria tra i contraenti, al più realizzerebbe una rappresentanza di interessi completa, che però rimane un'entità distinta e non sovrapponibile alla rappresentanza politica in ragione dell'autonomia dei rispettivi interessi difesi e scopi perseguiti. Infatti, la prima esibisce unicamente valori individuali; la seconda invece è proiettata verso il *common good*,

[51] Nel nostro lavoro, *A hard look at Self-regulation in the UK*, in *Eur. Bus. L. Rev.*, 2006, 17, 1, sono consegnate le osservazioni di metodo per ricondurre alla compatibilità costituzionale la fisionomia delle Autorità private di regolazione:

[52] W. STREECK - P.C. SCHMITTER (eds.), *Private interest government*, London, Sage Publications, 1985; ma anche: S. R. ACKERMAN, *Consensus versus incentives: a skeptical look at regulatory negotiation*, in *Duke l. j.*, vol. 43, 1206, 1994, 1216-1217. Nella nostra dottrina C. M. BIANCA, *Le autorità private*, Napoli, Jovene, 1977, 4 ss., dove l'Autore a ragione leggeva nella nuda attribuzione del potere normativo a un privato un *vulnus* al principio di uguaglianza. Mentre P. RESCIGNO, *La giustizia interna nelle associazioni private*, in ID., *Persona e comunità. Saggi di diritto privato*, Padova, CEDAM, vol. I, 1988, 123 ss., misurava la struttura e i poteri di tali Autorità secondo i parametri costituzionali della democraticità, quanto alla prima, e dell'uguaglianza, circa l'assegnazione di poteri a rilevanza esterna.

il quale non è la sfilata atomistica delle singole posizioni soggettive, ma la loro lettura, prima analitica e poi sintetica, dalla prospettiva dell'interesse pubblico. Questa irriducibilità di una entità nell'altra esclude che qualsiasi sostituzione reciproca possa accadere perché ciò che è nella disponibilità di un tipo di rappresentanza non è anche nella disponibilità dell'altra[53].

L'impossibilità giuridica che la rappresentanza privata operi in sostituzione di quella politica ci consente di risolvere la spinosa *issue* affrontata a più riprese dal diritto europeo, che ha cercato di chiarire quale sia l'ampiezza del *decisum* rimesso legittimamente ai privati. La questione, inizialmente riguardante la delega ai Comitati, soggetti pubblici ancillari alla Commissione, è estensibile per analogia alla delega della funzione legislativa europea ai privati. Anche in questo caso varranno le condizioni di legittimità poste nel caso Meroni[54], che tendeva a escludere la delega di compiti regolatori squisitamente politici, o caratterizzati da un elevato tasso di discrezionalità o anche implicanti scelte valoriali antitetiche. Se così fosse, saremmo dinanzi non all'esercizio di una funzione sotto il controllo del delegante, ma alla sua cessione con l'inevitabile imputazione della responsabilità politica al cessionario, alterando così l'iniziale architettura dei Trattati: "The intent is clear. Policy choices remain for the Commission; implementation is for the Agency"[55].

Questo limite di contenuto imposto alla delega[56] concorre a difendere l'*institutional balance*, che neppure la sua interpretazione più dinamica[57] è riuscita a dilatare senza compromettere le sfere di attribuzione definite nei Trattati, limite questo, che impone a ciascuna istituzione l'obbligo di "exercise its power with due regard for the power of other institutions…", il che comporta che "it should be possible to penalize any breach of that rule which may occur". In sintesi, il ricordato principio non consente spostamenti definitivi e solitari di attribuzioni decisionali, che, implicando valutazioni politiche in capo a soggetti estranei alle istituzioni

[53] In proposito, si rinvia al nostro studio, *Regole. Comando e consenso*, cit., in part. cap. II.

[54] Corte di Giustizia, *Meroni v. High Authority*, causa 9/56, 13 giugno 1958 [1957-1958], *ECR* 133.

[55] P. CRAIG, *The Constitutionalization of Community Administration*, in *Eur. L. Rev.*, 28, 6, 2003, 849.

[56] Corte di Giustizia (Grande Sezione), 22 gennaio 2014, C 270/12, in https://eur-lex.europa.eu/legal-content/IT/ALL/?uri=CELEX:62012CJ0270.

[57] E. CHITI, *Beyond «Meron» the community legitimacy of the provisions establishing the European agencies*, in G. CANANEA (a cura di), *European regulatory agencies*, Parigi, ISUPE Press, 2004, 83.

originarie, porterebbero con sé inevitabili slittamenti esterni di responsabilità, se non addirittura la creazione di aree immuni dal controllo politico.

Nell'interpretazione del Supremo Giudice, che ha privilegiato un'impostazione dinamica della *s.r*. a una più formale, che la avrebbe tenuta fuori dal suo *judicial review*[58], il principio in esame rappresenta quella sottile linea di confine che separa la decisione politica, riservata al Legislatore europeo chiamato a parlare per primo, dall'atto di *s.r.*, che invece deve seguire, sviluppare e articolare il discorso politico già iniziato da chi lo ha preceduto. Dunque, il rispetto dei principi del costituzionalismo comune europeo – uguaglianza, libertà, responsabilità, in una parola la democraticità del sistema – sono la condizione necessaria di legittimità della delega di poteri normativi al privato. Come solo il rispetto dei tre piani in cui si articola la condivisione del compito normativo con i privati – ricordiamo, quello soggettivo, procedurale e finalistico – consente alle Istituzioni europee di riconoscere all'autoregolazione dignità di fonte nel sistema giuridico, mentre in assenza di queste garanzie, la norma privata deve rimanere un atto negoziale con efficacia necessariamente limitata alle sole parti private.

8.6 I codici previsti nel DSA rispettano il modello di compatibilità costituzionale?

Il Codice di condotta sulla disinformazione, previsto all'art. 45 del DSA, è un chiaro esempio di processo co-regolativo perché la proposta della parte privata, soggettivamente complessa, è stata condivisa dalla Commissione, che, a più riprese, ha indicato aggiustamenti nel merito e nella procedura al *draft* privato. Così all'iniziale bozza del 2018[59], criticata con osservazioni

[58] Così: A. POGGI, *Soft law nell'ordinamento comunitario*, 2005, in https://www.associazionedeicostituzionalisti.it/old_sites/sito_AIC_2003-2010/materiali/convegni/aic200510/poggi.html, par. 3.

[59] Si tratta del *Code of Practice on Disinformation* del settembre 2018: un codice di autoregolamentazione contenente 21 impegni in diversi settori, dalla trasparenza nella pubblicità politica alla demonetizzazione dei promotori della disinformazione. In https://digital-strategy.ec.europa.eu/it/library/2018-code-practice-disinformation.

puntuali dalla Commissione[60], è seguita l'ultima versione[61], ora integrata nel DSA[62], oggetto di un'*Opinion* tendenzialmente approvativa di Bruxelles[63].

In linea con l'impostazione teorica prima indicata verificheremo se l'atto possegga i requisiti soggettivi e oggettivi di legittimità, necessari per la sua acquisizione alle fonti del diritto.

Sono due le condizioni di compatibilità istituzionale di questa divisione del lavoro tra il legislatore europeo e i privati regolatori: il primo dovrà a) disegnare la fisionomia interna di questi governi privati di interesse e b) riservare a sé la direzione politica degli sviluppi dell'autoregolazione, orientando il potere privato non diversamente da come farebbe per l'esercizio delegato della funzione normativa; mentre al privato competerà lo sviluppo e l'implementazione nel rispetto della cornice tratteggiata dal primo.

Circa il profilo soggettivo: le Autorità private dovrebbero esibire un'identità democratica, un *board* internamente articolato in modo da rappresentare l'intero fascio degli interessi antagonisti coinvolti nella materia da disciplinare, perché sono aiutanti di campo del soggetto pubblico, come ci ricorda la Suprema Corte americana "Lindke sued Freed under 42 U. S. C. §1983, alleging that Freed had violated his First Amendment rights. As Lindke saw it, he had the right to comment on Freed's Facebook page, which he characterized as a public forum. Freed, Lindke claimed, had engaged in impermissible viewpoint discrimination by deleting unfavorable comments and blocking the people who made them"[64].

Invece, i codici del DSA sono dettati dai privati, la cui identità è schiacciata su una sola parte del rapporto, quella industriale, con la pretermissione sostanziale della parte debole, che

[60] Commission Staff Working Document, *Assessment of the Code of Practice on Disinformation - Achievements and areas for further improvement*, 10 febbraio 2020, in https://digital-strategy.ec.europa.eu/en/library/assessment-code-practice-disinformation-achievements-and-areas-further-improvement. L'anno seguente la Commissione europea ha elaborato linee guida per potenziare il Codice, già in prospettiva di integrazione con il *Digital Services Act*. Si veda: Commissione europea, *European Commission Guidance on Strengthening the Code of Practice on Disinformation* COM/2021/262 final, 26 maggio 2021, in https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex:52021DC0262.

[61] Il *Code of Conduct on Disinformation*, pubblicato nel giugno 2022, è stato ripubblicato il 13 febbraio 2025, a seguito dell'integrazione nel DSA, con un Preambolo aggiornato. La versione più recente è in: https://disinfocode.eu/the-code/read (il sito web è anche il *Transparency Centre*, secondo il Capo VIII del Codice stesso).

[62] Ai sensi dell'art. 45, par. 4, del DSA, infatti, l'integrazione dei Codici di condotta è subordinata al doppio parere positivo della Commissione europea e del Comitato europeo per i servizi digitali.

[63] Commissione europea, *Commission Opinion of 13.2.2025 on the assessment of the Code of Practice on Disinformation within the meaning of Article 45 of Regulation 2022/2065*, in *https://digital-strategy.ec.europa.eu/en/library/code-conduct-disinformation.*

[64] Supreme Court, US, 15 March 2024, *Lindke c. Freed LLC*, in, 2024, 601 U.S. 187.

pure subirà le regole impostole dalla controparte *well funded and well structured*, che lei non si è scelta. Qui il diritto a genesi spontanea, il *droit douce*, come dicono i francesi, è più imperativo dell'eteronomia, perché si basa su un consenso presunto, che, se pur prestato dai privati deboli, non ha conosciuto un'esperienza procedimentale, né tantomeno di visualizzazione esterna. Infatti, non sappiamo se e cosa i partecipanti, diversi dalle *Big Tech*, abbiano osservato al *draft* presentato loro completo in ogni sua parte; lì dove sarebbe stato preferibile per un'effettiva equiordinazione di tutti gli interessati al tavolo delle trattative che le *Big Tech* avessero seguito un modello negoziale, che avesse il suo incipit nello schema del *notice and comment*, cioè in una lista di domande a risposte aperte, rimesse alla discrezionalità decisionale dei contraenti, invece, che in un tessuto unilateralmente composto e articolato in caselle regolatorie già integralmente riempite nel contenuto.

Quanto alla *disclosure*, l'attributo è stato del tutto assente nella fase della negoziazione, per cui non sappiamo se le eventuali osservazioni delle parti deboli siano state accolte e in che misura. In sintesi, il contratto non è stato valorizzato come momento riequilibratore di un'iniziale asimmetria negoziale tra contraenti disallineati; con la conseguenza che assegnare anche a 'soggetti diversi' dalle *Big Tech* l'etichetta di contraenti significa accontentarsi del mero dato formale della loro firma in calce all'atto[65]. Ma sul terreno dei contenuti le regole per i contraenti deboli risulteranno più comandate di quelle imposte dall'eteronomia: in sintesi, siamo davanti a un diritto in apparenza consensuale nella sostanza unilaterale e imperativo.

Quanto ai fini, questi andrebbero predefiniti con una precisione tale da assicurare alla politica la scelta sul tipo di rischio, sulla sua entità e sul modo di valutazione, ma, tra queste tre opzioni, la principale è l'individuazione del gradino della *ladder of risk* dove fermare il pericolo accettabile: decidere quanta ingiustizia sia consentita e da quale punto in poi l'ingiustizia diventi intollerabile è, infatti, una *political issue* perché ha a che fare con una precisa idea qualitativa e quantitativa di giustizia e di uguaglianza da realizzare.

[65] Il Preambolo del Codice afferma che i firmatari stessi ne sono gli Autori: "It is in this spirit that the Signatories have drafted the present Code identifying the commitment areas and measures each Signatory is making.", *Code of Conduct*, 5. Più esplicita la *Commission Opinion*, secondo cui "the Commitments, Measures and KPIs of the Code of Practice on Disinformation were negotiated between a very diverse community of stakeholders", cit., 14.

A queste riflessioni, critiche perché rimandano a una decisione politica girata in bianco ai privati, va aggiunta invece una nota di apprezzamento per la faticosa negoziazione tra *Big Tech* e Commissione, che ha prodotto un atto in linea con gli obiettivi del DSA – indicati nelle sez.i II-VII del codice 2022 – *responsive* nelle misure per conseguire gli obiettivi e puntuale, in accoglimento ai rilievi della Commissione, quanto agli indici per rilevare sul piano qualitativo e quantitativo la misura di realizzazione degli obiettivi promessi[66].

Si pensi, a titolo di esempio, alla misura 21.1[67], che impone di apporre etichette, indicanti le valutazioni dei *fact-checker*; nonché indicare avvisi agli utenti, che tentano di condividere o hanno precedentemente condiviso il contenuto valutato e, infine, di rendere visibili con immediatezza i pannelli informativi sui contenuti notificati dai *fact-checker* che violano le loro politiche.

Rimane però aperta la domanda principale per un costituzionalista: questo modello di *co-regulation* è in linea con il circuito politico-rappresentativo oppure no?

Si tratta di uno strappo che, a differenza di quello inizialmente procurato dalla regolazione sostanzialmente primaria delle A.I., nel caso della co-regolazione sarà più difficile da ricucire perché la deroga è più subdola della precedente in quanto l'Autorità privata si propone come un innocuo autoregolatore, per poi tradire la sua genesi privata quando le sue norme, superando l'efficacia *inter partes*, vincolano un ambito soggettivo più ampio della base associativa che sia autoregola.

Questo modello non è privo di ricadute sull'assetto istituzionale perché, stante la contrattazione asimmetrica a vantaggio dei forti, la *co-regulation* viola il principio di uguaglianza formale in quanto il privato sostanzialmente pretermesso subirà un'invasione della sua sfera di autonomia, che, in assenza del potere privato rafforzato dalla cogenza pubblica, avrebbe potuto respingere; di contro, le Autorità private invadono l'altrui ambito

[66] Ogni impegno (*Commitment)* è, infatti, strutturato in Misure (*Measure*) specifiche, che i firmatari devono attuare a seconda della natura dei loro servizi; per quasi ogni misura si indicano poi QRE (*Qualitative Reporting Elements*) ovvero elementi di rendicontazione qualitativa, e SLI (*Service Level Indicators*), indicatori quantitativi per misurare l'efficacia delle misure. I firmatari si impegnano, inoltre, a sviluppare, entro 9 mesi dalla firma, degli Indicatori Strutturali (*Commitment 41*) per comprendere l'impatto complessivo del Codice.

[67] *Code of Conduct,* 28.

individuale, azione che, in assenza di questa autorizzazione *ex lege*, non potrebbero intraprendere.

Se trasportiamo questa riflessione sul terreno delle fonti e poi su quello ordinamentale, la tecnica per essere un elemento cooperativo del sistema deve poter esibire la norma europea, non come mero titolo abilitante a creare diritto oggettivo, ma come principio di disciplina sostanziale, già immediatamente conformativa del rapporto, i cui futuri sviluppi sono rimessi alla tecnica armonizzante.

Ecco allora che la fonte, che si dovrebbe collocare su un gradino al di sotto della legge europea, in concreto non parla per seconda perché la politica non ha dettato in via iniziale un linguaggio essenziale; questa inversione dell'ordine degli interventi lungo la catena regolatoria fa sì che il codice si assuma la responsabilità della prima mossa, come anche dei suoi sviluppi successivi.

Qui la realizzazione del *common good* diventa un evento futuro e del tutto incerto, dipendendo dalla sua occasionale coincidenza con gli interessi egoistici dei *private interest government*, mancando una seria cornice di riferimento che assicuri l'obbedienza della *co-regulation* alla vocazione sociale.

I privati strutturati economicamente e possessori indebiti dei nostri dati, che dovrebbero essere docili esecutori di un compito pubblico, disponibili alla *mise en place et de la mise en oeuvre des politiques publiques* sono l'ennesimo mito creato dall'illusione di una tecnica al servizio dell'uomo, ma per il momento il DSA ha introdotto autorità private, aggressive e poco disponibili a realizzare la centralità dell'uomo.

8.7 L' Artificial Intelligence Act e la sua pretesa di orientare la tecnica verso l'uomo

Esaminiamo l'*Artificial Intelligence Act*[68] dall'angolo visuale della sua promessa: piegare l'I.A. alla centralità dell'individuo. Poche battute sul tipo di modello regolativo introdotto sul

[68] Regolamento (UE) 2024/1689, cit., d'ora in poi con l'acronimo: *AI Act*. Anche per riferimenti dottrinali, italiani e stranieri, si rinvia al nostro saggio: *Giustizia e intelligenza artificiale*: *un equilibrio mutevole*, in *Riv. AIC*, 2/2024, 86 ss.

mercato digitale e poi un giudizio circa il mantenimento o meno della sua ambiziosa promessa.

Iniziamo col dire in breve, ai fini del nostro discorso, cosa sia l'Intelligenza artificiale[69]. La sua identità coincide con un processo di apprendimento, guidato dalla mente umana ma altresì capace di evoluzione autonoma[70]; si nutre di masse crescenti di dati[71] con cui confeziona in base a valutazioni automatiche il contenuto seriale di decisioni pubbliche o private.

Il modo di procedere dell'I.A si snoda lungo due assi. Il primo riguarda il metodo del suo ragionamento: questo, di tipo statistico-correlazionale, anticipa con valutazione prospettica il verificarsi di probabili situazioni in base all'*id quod plerumque accidit*[72]. Il secondo asse riguarda invece gli effetti dell'agire intelligente: questi tendono almeno a orientare, se non anche a conformare, la condotta di ampie collettività di persone.

Il percorso logico della I.A. ha il suo *focus* in una prognosi *ex ante*. Questa previsione riguarderà il momento iniziale, in cui l'Intelligenza anticipa un probabile evento futuro – ad es. un fatto criminoso – ma interesserà anche quello conclusivo, in cui la mente meccanica tratteggerà una *policy* di comportamento, crescente nei suoi obblighi all'aumentare del rischio perché è diretta a evitare che il pericolo futuro e incerto degeneri in un danno concreto e attuale alle libertà fondamentali dei terzi.

Siccome il rischio dell'I.A. è il fattore ineliminabile della regolazione in esame, questa segue un modello precauzionale, dove l'aggettivo sottolinea la sua capacità di anticipare un accadimento dannoso e di provare a evitarlo imponendo agli operatori regole di comportamento cautelari che, se osservate, generano nella collettività una relativa certezza quanto alla sicurezza della macchina, e creano un affidamento legittimo che le cose stiano andando per il verso giusto, cioè che nessun danno accadrà. Le norme diranno all'operatore

[69] D'ora in avanti con l'acronimo: I.A.

[70] Così: U. RUFFOLO - A. AMIDEI, *La regolazione ex ante dell''intelligenza artificiale tra gestione del rischio by design, strumenti di certificazione preventiva e "autodisciplina" di settore*, in A. PAJNO - F. DONATI - A. PERRUCCI (a cura di), *Intelligenza artificiale e diritto: una rivoluzione?*, Bologna, Il Mulino, 2022, I, 493.

[71] Ci sia consentito il rinvio al nostro lavoro: *Big Data e la debole resistenza delle categorie giuridiche*, in *Dir. Pubbl.*, 1, 2019, 90.

[72] V. MAYER-SCHÖNBERGER - K. CUKIER, *Big data. Una rivoluzione che trasformerà il nostro modo di vivere – e già minaccia la nostra libertà* (trad. it. a cura di R. MERLINI), Milano, Garzanti, 2013, 16.

con linguaggio giuridicamente vincolante o di *soft law* come scegliere gli elementi compositivi dell'I.A., come addestrare i set di dati, come assicurarne la conservazione, l'equità e la non discriminazione, come valutare gli *input*, come ritagliare uno spazio all'intervento umano e, infine, come articolare la logica che conduce all'*output*. La disciplina precauzionale non si esaurisce però nel disegnare il paradigma astratto di cervello artificiale, perché questo, una volta immesso sul mercato, continuerà a essere assistito da regole puntuali, che obbligheranno il fornitore ad aggiornarlo, monitorarlo e, se del caso, correggerlo anche con la misura estrema del suo ritiro dal mercato in caso di danni irreversibili.

Il fatto che l'*AI Act* abbia seguito un approccio *risk-based* ha dato una particolare inclinazione alla sua tecnica normativa: questa si compone di valutazioni legali tipiche, che hanno categorizzato le attività in tre tipologie in ragione dell'entità del rischio. Quelle vietate in assoluto per l'inaccettabilità del pericolo potenziale; quelle ad alto rischio, ammesse perché il *check and balance* tra innovazione e libertà fondamentali le permette nel rispetto della disciplina precauzionale di minimizzazione del rischio[73]; e, infine, quelle caratterizzate da un rischio così tenue da essere esonerate dall'osservanza delle regole più gravose, come richiede il criterio della proporzionalità tra entità del rischio e peso regolatorio.

Il ragionamento si basa sull'assunto che l'osservanza della legalità astrattamente posta isolerà i diritti fondamentali dal pericolo di lesione, creando una parvenza di liceità, cioè una garanzia per la collettività che tutto andrà bene. Se poi ciò non dovesse accadere e i diritti fossero lesi, si applicheranno le norme sulla responsabilità per danni, ancora però in cerca d'autore nell'ordinamento europeo, dopo il ritiro della Proposta di direttiva sulla responsabilità per danni da Intelligenza artificiale[74], il che lascia insoddisfatta anche la pretesa risarcitoria del danneggiato, ammesso che la lesione dei diritti fondamentali possa essere compensata col denaro.

[73] S. HEISS, *Artificial Intelligence Meets European Union Law The EU Proposals of April 2021 and October 2020*, in *EuCML*, 6, 2021.

[74] Commissione europea, *Ritiro della proposta di Direttiva del Parlamento europeo e del Consiglio, relativa all'adeguamento delle norme in materia di responsabilità civile extracontrattuale all'intelligenza artificiale,* C/2025/5423, 6 ottobre 2025, in *https://eur-lex.europa.eu/legal-content/IT/TXT/PDF/?uri=OJ:C_202505423* .

Questo, dunque, il nocciolo sostanziale della disciplina precauzionale, ma la sua effettività dipenderà da come verrà assicurata l'osservanza delle regole. Pertanto, anche il profilo formale andrebbe orientato all'obiettivo precauzionale, se si vuole che la disciplina preventiva vada a buon fine. Qui l'*AI Act* ha riproposto i modelli classici della funzione di controllo[75]: un sindacato preventivo sulle attività automatizzate non ancora avviate, e uno successivo su quelle in corso d'opera.

Il controllo anticipato, apprezzabile come idea, è invece deludente nella sua realizzazione in quanto il legislatore non ha obbligato il fornitore a sottoporre il sistema a un certificatore esterno e pubblico, accontentandosi di una sua semplice autodichiarazione, c.d. *self-compliance*, che attesti il rispetto della disciplina prudenziale. La coincidenza soggettiva tra controllato e controllante, cioè questo fisiologico conflitto di interessi, compromette l'obiettività del sindacato, strutturalmente inadatto a offrire quelle garanzie di neutralità necessarie a rassicurare i terzi, anche se l'*AI Act* conosce il modello del controllo preventivo esterno, che però impiega solo in situazioni residuali e tassative (ex art. 43, par. 1, co. 2) perché eccessivamente oneroso per il privato.

Durante i lavori preparatori dell'*AI Act* numerose furono le critiche alla *self-compliance* per la debole obiettività delle autovalutazioni e degli *audit* in luogo di imparziali controlli di revisori terzi. Un *audit* solido non solo garantisce i diritti dei cittadini ma facilita gli utenti *downstream*, cioè coloro che costruiranno ulteriori sistemi sulle I.A. presenti nel mercato[76].

Quanto detto rende la *self-compliance* e le altre autodichiarazioni disseminate qua e là nel testo il vero punto debole di questa disciplina, che in nome della celerità decisionale ha sostituito ai controlli autoritativi le dichiarazioni di buona condotta del diretto interessato, che sarebbe stato opportuno sottoporre all'attenta verifica di un soggetto pubblico imparziale.

Infine, un ultimo atto della disciplina prudenziale prende a prestito la forma dell'autodichiarazione: la "valutazione d'impatto sui diritti fondamentali"[77]. Si tratta di un

[75] G. FERRARI, *Gli organi ausiliari*, Milano, Giuffrè, 1956, 270 ss.

[76] C. Dunlop, *Policy Briefing. An EU AI Act that works for people and society. Five areas of focus for the trilogues*, 6 settembre 2023, in *https://www.adalovelaceinstitute.org/policy-briefing/eu-ai-act-trilogues/*.

[77] *Fundamental Rights Impact Assessment*, FRIA, ex art. 27, *AI Act*.

giudizio anticipato sui rischi che l'Intelligenza potrebbe procurare ai diritti e affidato dall'*AI Act* allo stesso fornitore dell'I.A., ma questa intenzionale coincidenza soggettiva vanifica la funzione cautelativa della valutazione, perché il suo autore interno, meno scrupoloso del terzo nel rappresentare le cose, potrà presentare come minimi rischi che non sono tali, e attenuarne le misure di mitigazione. Era atteso come strumento incisivo, se fosse stata rispettata la formulazione adottata negli emendamenti del Parlamento europeo[78]; invece è stato gravemente indebolito nel testo approvato dal Trilogo[79], sia perché è stato reso obbligatorio solo per taluni fornitori[80], sia perché il *deployer* valuterà l'impatto solo sui soggetti, non anche sui diritti fondamentali, diversamente da come prevedevano gli emendamenti del Parlamento.

Con queste variazioni sul tema del *self-assessment* il sistema pubblicistico continua un'inarrestabile corsa verso l'involuzione privatistica, come dimostrato da due circostanze: molte situazioni rimangono fuori dalla valutazione d'impatto, e il suo svolgimento, precluso ai terzi, non contempla neanche l'apporto pubblico, perché non informa in anticipo l'Autorità nazionale, che potrà intervenire solo a cose fatte. L'intervento pubblico tardivo è un ossimoro per una disciplina che ha l'ambizione di essere precauzionale e di migliorare il godimento dei diritti, non di comprometterli[10], risultato questo che il *self-risk-assessment* non potrà certo garantire.

Dunque, la promessa di un'I.A. umano-centrica attende ancora di essere mantenuta perché al momento è sbilanciata a favore dei privati *well-funded and well-structured*. In questa situazione è difficile credere che il Legislatore europeo abbia voluto davvero rafforzare la fiducia dei cittadini nel processo automatizzato, visto che non ha posto le condizioni perché

[78] Parlamento europeo, *Compromise Amendments (COM (2021)0206 – C9 0146/2021 – 2021/0106(COD),* 9 maggio 2023, art. 29 bis (ora art. 27).

[79] Consiglio europeo, *Proposal for a Regulation of the European Parliament and of the Council laying down harmonized rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts - Analysis of the final compromise text with a view to agreement*, 26 gennaio 2024, art. 29 bis (ora art. 27).

[80] Precisamente, l'art. 27 lo limitava a soli soggetti pubblici e privati fornitori di servizio pubblico, ai fornitori di *social scoring* o di assicurazioni sul rischio.

ciò accada[81]. Invero, il sistema denuncia una contraddizione di fondo: nei *consideranda* il decisore politico promette centralità alla persona e il suo dominio sulla macchina; nell'articolato dimentica la promessa fatta qualche rigo sopra.

Apprezzare un approccio regolatorio *risk-based* non nasconde i difetti di un sistema di I.A., poggiato su una catena di presunzioni non controllate, scritte sotto dettatura delle *lobby*[82], noncurante dei possibili pregiudizi ai diritti fondamentali. Inoltre, questi meccanismi, non assistiti da un solido controllo pubblico, rischiano di ingannare i cittadini con vuote certezze giuridiche[83], con finzioni di legalità, che generano affidamenti ingiustificati su situazioni, in apparenza legali, ma nella sostanza illecite, proprio come quel *Christmas tree and pretty lights*, di cui parla Feigenson[84], alle cui luci neppure il giudice più ostile sa resistere.

8.8 La tecnica come "source of law"

La tecnica fa ancora un passo in avanti, da elemento integrativo della norma o del suo precetto per volontà di legge a fonte del diritto per auto-promozione. In questa azione la tecnica altera radicalmente il paradigma tradizionale del diritto dei privati, perché disegna regole di condotta senza ricorrere alla mediazione umana; e in questa assenza dell'individuo compie la sua rottura con il sistema classico delle fonti.

Infatti, la sua peculiarità si manifesta nell'essere fonte autosufficiente, che avvia e conclude l'iter regolatorio solitariamente: il cerchio inizia e termina nella tecnica. Insomma, tutto si risolve nella statuizione della mente artificiale. Questa situazione si verifica quando – qui

[81] M. DRAGHI, *Discorso*, alla Conferenza della Commissione europea "Un anno dopo il rapporto Draghi: cosa è stato realizzato, cosa è cambiato", Bruxelles, 16 settembre 2025, in *https://www.eunews.it/2025/09/16/un-anno-dopo-il-rapporto-draghi/*con ragionate critiche al disegno europeo sull'I.A.: l'eccesso di regole impedisce all'I.A. di fare da volano alla competitività, riducendola a ostacolo dell'innovazione.

[82] P. FRIEDL - G.G. GASIOLA, *Examining the EU's Artificial Intelligence Act*, in *VerfBlog*, 2024/2/07; Corporate Europe Observatory, *Byte by byte How Big Tech undermined the AI Act*, 17.11.2023, in *https://corporateeurope.org/en/2023/11/byte-byte.*

[83] L'atteso *Digital Omnibis* (Com(2025) 837 final) in *Digital Omnibus Package*, della Commissione europea, ancora in corso, nel tentativo di semplificare norme e procedure, non sembra cogliere i difetti, illustrati sopra, perché alleggerisce le regole per i grandi operatori e, dall'altro canto, priva i cittadini dell'effettiva tutela precauzionale.

[84] N. FEIGENSON, *Brain imaging and courtroom evidence: on the admissibility and persuasiveness of fMRI*, in *International Journal of Law in Context*, 2, 3, 2006, 246.

offriamo uno dei tanti esempi – il DSA[85] e i codici di autoregolazione delle piattaforme[86] non disegnano la fattispecie della notizia falsa o diffamatoria[87]. Quindi, per conoscere in anticipo il divieto, sarà necessario leggere i concreti atti di pulizia delle piattaforme *online*[88]. Così la catena normativa scende al livello più basso dei codici per coincidere con il concreto provvedere diretto a rimuovere la notizia falsa, denigratoria o violenta[89]. Ne consegue che l'astrattezza e la generalità si convertono nella concretezza e nella particolarità del provvedimento di rimozione meccanica, a danno della certezza e dell'uguaglianza del diritto. Il DSA ha pertanto assegnato ai GK una funzione pubblica inedita: pulire la rete dalle notizie false, un compito questo, che un ordinamento democratico non deve assegnare neppure al soggetto pubblico, perché esso presuppone il preventivo disegno di un paradigma di verità. L'Atto europeo va oltre questo limite: non solo delega al privato la funzione di pulizia, ma lo esonera dall'osservare le garanzie formali e sostanziali, necessarie a delimitare il potere nell'incontro con le libertà fondamentali. Qui salta sia la garanzia della *rule of law*, perché

[85] Regolamento (UE) 2022/2065 del Parlamento Europeo e del Consiglio del 19 ottobre 2022 relativo a un mercato unico dei servizi digitali e che modifica la direttiva 2000/31/CE, in *https://eur-lex.europa.eu/legal-content/IT/TXT/PDF/?uri=CELEX:32022R2065*.

[86] Nella dottrina anglosassone ricordiamo l'innovativo pensiero di A. OGUS, *Rethinking self-regulation*, in *Oxford Journal of Legal Studies*, 15, 1, 1995, 99-100, che alle etichette predefinite ha preferito uno studio della *self regulation* per il grado di autonomia decisionale, di forza legale e di potere monopolistico del soggetto di autonormazione: "to appreciate the range of possibilities […] which can properly be described as 'self-regulation'". Mentre J. BLACK, *Constitutionalising self-regulation*, in *Mod. L. Rev.*, 59, 1996, 26, ha sottolineato la necessità di un intervento conformativo dello Stato, se le regole del soggetto collettivo sono non esclusivamente "tailored to the circumstances of particular firms", ma imposte all'intera categoria sociale di riferimento. Per gli sviluppi nella letteratura anglosassone ci sia consentito il rinvio al nostro lavoro: *A Hard Look at Self-Regulation in the UK*, cit., 183-211.

Per una trattazione ampia dei profili formali e sostanziali dei codici di autoregolazione delle piattaforme si vedano almeno: A. SIMONCINI, *La co-regolazione delle piattaforme digitali*, in *Riv. Tri. Dir. Pubbl.*, 4, 2022, 1031-1049; G. DI COSIMO, *La co-regolazione delle tecnologie digitali: il paradigma centro-periferia*, in *Oss. sulle Fonti*, 1, 2024, 272 ss.; O. POLLICINO, *Regolazione e innovazione tecnologica nell' "ordinamento della rete"*, in *Riv. AIC*, 2, 2025, 169; L. TORCHIA, *I poteri di vigilanza, controllo e sanzionatori nella regolazione europea della trasformazione digitale*, in *Riv. Tri. Dir. Pubbl.*, 4, 2022.

[87] Un esame ampio degli obblighi comportamentali delle piattaforme è offerto da: M. OROFINO, *Il Digital Service Act tra continuità (solo apparente) ed innovazione*, in F. PIZZETTI (a cura di), *La regolazione europea della società digitale*, Torino, Giappichelli, 2024, 144.

[88] A. LAMBERTI, *La libertà di manifestazione del pensiero "in trasformazione"*, in *Riv. AIC*, 3, 2025, 340.

[89] A. LUCARELLI, *Nuovi mezzi di comunicazione, assetti imprenditoriali e soggettività politiche*, in *Riv. AIC*, 3, 2025, 103.

manca la definizione *ex ante* del concetto di falso, ma anche quella del *due process*, perché il controllo sui contenuti si fa *inaudita altera parte* per ragioni di celerità[90].

A ciò si aggiunga che la mancanza di una definizione astratta e generale del concetto di falso lascia ai GK la libertà di confondere il falso con la notizia politicamente inopportuna o non conforme al pensiero dominante[91]. Questo vuoto normativo del DSA rimanda a un ragionamento tautologico secondo il quale la condotta di odio o quella violenta è vietata perché vietata. Questo circolo vizioso può nascondere pericolose tesi liberticide e facili scivolamenti verso un unico e solo discorso lecito: quello di Stato. Detto rischio è inevitabile in assenza di un parametro normativo che definisca in anticipo il concetto di falso; con il paradossale esito di una cesura netta tra l'ambiente *offline*, dove la divulgazione di idee false non è reato, a meno che non aggredisca beni-interessi diversi dalla verità, e quello virtuale, dove l'idea presuntivamente falsa da esercizio di un diritto diventa fatto illecito.

Tale modo di procedere azzera la certezza del diritto, la pretesa del cittadino alla conoscibilità anticipata della norma, e copre tutto con un velo sotto il quale il lecito e l'illecito, il permesso e il divieto, si mescolano fino a confondersi l'uno nell'altro senza soluzione di continuità.

Al controllo politico dei privati sulle idee di altri privati preferiamo le virtù benefiche del *marketplace of ideas*, che, nel consentire la coesistenza del falso con il vero, lascia ai cittadini il compito di distinguerli in quanto hanno una maturità sufficiente per farsi un'idea propria, senza essere pre-indirizzati da chi ha la pretesa di sapere per loro cosa sia la verità. "In other words, the government's control over speech as an intellectual arbiter of the truth must be constricted, if not completely denied. Baked into the architecture of the First Amendment then - at least from a structural rights interpretation - is 'a deep skepticism about the good faith of those controlling the government.' That skepticism flows from two facts: (1) decisions

[90] I. BURI - J. VAN HOBOKEN, *The Digital Services Act (DSA) proposal: a critical overview,* in *https://dsa-observatory.eu/wp-content/uploads/2021/11/Buri-Van-Hoboken-DSA-discussion-paper-Version-28_10_21.pdf*; J. LAUX-S. WACHTER-B. MITTELSTADT, *Platform Regulation, Independent Audits, and the Risk of Capture Created by the DMA and DSA,* in *Computer Law & Security Review,* 43, 2021, p. 105613; M. L. CHIARELLA, *Digital Markets Act (DMA) and Digital Services Act (DSA): New Rules for the EU Digital, in Athens Journal of Law,* 9, 1, gennaio 2023, 33-58; M.D. COLE - C. ETTELDORF - C. ULLRICH, *Updating the Rules for Online Content Dissemination. Legislative Options of the European Union and the Digital Service Act Proposal,* Baden-Baden, Nomos, 2021, 117 ss.

[91] Argomenta a favore della censura privata: C. PINELLI, *Disinformazione, comunità virtuali e democrazia: un inquadramento costituzionale*, in *Dir. Pubbl.,* 1, 2022,180 ss.

about what is true or false, when made by those in power, 'are bound up with political perspectives that the government seeks to undermine;' and (2) 'the government's natural tendency [is] to twist reality to its own purposes"[92].

La sostituzione della mente artificiale a quella umana genera una "regola tecnica" completa, particolare, *ad personam*, riconducibile alla volontà sintetica, dinanzi alla quale lo *human in the loop* è ridotto a una mera rivendicazione verbale della centralità dell'individuo; mentre il diritto a interloquire con una persona in carne e ossa, pur riconosciuto dalla legge, non ha lasciato tracce di sé nel processo creativo delle regole: ancora un mito che si aggiunge ai tanti altri creati dall'*AI Act*.

Riflettiamo ora su cosa significhi estromettere il soggetto-persona dal processo normativo. Quando la tecnica diventa *source of law*, un "soggetto-non soggetto" entra sulla scena politica: l'I.A. abilitata a dire il diritto in luogo della ragione umana. Cade dunque l'imputazione umana dell'atto-fonte perché siamo davanti a un fatto non riconducibile a un momento volitivo dell'individuo: qui manca l'energia intenzionale della persona, che nella fonte-atto occupa il segmento deliberativo della decisione normativa[93].

Affermare il vuoto volitivo non significa ridurre l'I.A. a un'entità incapace di conoscere e riflettere; lo prova il fatto che l'*AI Act* chiede all'Intelligenza una prestazione di risultato, il cui adempimento richiede consapevolezza e logica argomentativa: scrivere una motivazione ragionevole a supporto della statuizione intelligente. Questa motivazione presuppone che l'I.A. sia in grado di compiere deduzioni secondo il criterio causale, non diversamente da come farebbe la mente umana. Se questo accostamento del *reasoning* tra le due intelligenze è sostenuto in filosofia[94], noi non siamo pronti ad aderire all'una o all'altra delle tesi in campo perché immature sul piano scientifico; mentre possiamo osservare che il ragionamento, alla base della motivazione, è ancora un aspetto deficitario dell'atto algoritmico, prestandosi a

---

[92] C. CALVERT - S. MCNEFF -A. VINING - S. ZARATE, *Fake News and the First Amendment: Reconciling a Disconnect Between Theory and Doctrine*, in *U. Cin. L. Rev.*, 2018, 86, 99, 134.

[93] A. CARDONE, *"Decisione algoritmica"* vs *Decisione politica?*, Napoli, ES, 2021, 159.

[94] S. BUBECK, V. CHANDRASEKARAN, R. ELDAN *et alii*, *Sparks of Artificial General Intelligence: Early experiments with GPT-4,* in *http://efaidnbmnnnibpcajpcglclefindmkaj/https://arxiv.org/pdf/2303.12712*; già R. KURZWEIL, *The singularity is near: when humans transcend biology*, London, Penguin Putnam Inc, 2006 e D. DENNETT, *From Bacteria to Bach and Back*: *the evolution of minds*, NY, W. W. Norton & Company, 2017; per un'idea alternativa: J. R. SEARLE, *Minds, brains, and programs, in Behavioral and Brain Sciences*, 3, 3, 1980, 417-457.

facili impugnative per l'inconsistenza della motivazione rispetto al modello astratto posto dal Legislatore europeo. Pertanto, se la motivazione è l'unico elemento che avvicina I.A. a quella umana, essa prova poco, visto che l'I.A. adempie in misura insoddisfacente all'obbligo motivazionale.

Ma ammettiamo pure che l'I.A. abbia capacità dianoetiche[95], questo riconoscimento non sana la sua incapacità di auto-addestrarsi alle virtù etiche, essenziali per compiere le scelte sui valori in modo da disegnare una misura di equilibrio desiderabile tra gli opposti interessi, come invece fa il legislatore umano.

L'I.A. non sa costruire da sé il proprio patrimonio valoriale[96], pertanto spetterà alla mente umana educarla al linguaggio etico, dirigerla verso quella combinazione di beni disegnata dalla politica, che, come voleva Aristotele, ha il compito fondamentale di correggere le tensioni autonomistiche e le cadute egoistiche di una tecnica abbandonata a se stessa[97], e qui il nostro discorso si salda con le riflessioni iniziali.

Nel caso invece della tecnica-fonte, essa è principio e fine della decisione politica, che si risolve interamente nella determinazione artificiale, abilitata a scrive su una lavagna quasi bianca. Questo modello di tecnica non tollera limiti, perché forte dell'autolegittimazione non riconosce nessuna autorità superiore a sé. Per tale motivo va dove vuole, noncurante di un'eventuale responsabilità politica, da cui è sollevata perché per essere responsabili occorre che qualcuno investa un altro con un mandato politico, di cui il secondo dovrà rendere conto al primo, in quanto suo mandante. In sintesi, l'autogenesi le regala il privilegio dell'impunità assoluta.

Né la ricerca di un soggetto politicamente responsabile per le scelte compiute sarebbe compensabile con l'*accountability* della macchina. Questa idea crea più problemi di quelli che vorrebbe risolvere: e, tra questi, l'irremovibilità nel tempo della tecnica, che potrebbe resistere all'infinito, sottraendosi al rischio di essere sostituita, come invece accade con

[95] Si rinvia al paragrafo iniziale quanto alla definizione aristotelica e all'antagonismo tra le virtù etiche e quelle dianoetiche.

[96] Di avviso contrario: E. SEVERINO, in N. IRTI - E. SEVERINO, *Dialogo su diritto e tecnica*, cit., 52.

[97] L. B. SOLUM, *Artificially Intelligent Law*, in *BioLaw Journal*, 1, 2019, 57; anche A. D'ALOIA, *Intelligenza artificiale, società algoritmica, dimensione giuridica. Lavori in corso*, in *Quad. Cost.*, 3, 2022, 667-670.

l'eletto non più gradito agli elettori, liberi di non riconfermargli per il successivo mandato la fiducia iniziale.

Né il difetto di responsabilità sarebbe recuperabile con una tecnica costruita *by design* perché l'anticipazione della soglia dell'antigiuridicità, introducendo nel frullatore meccanico gli elementi essenziali per adeguarla al parametro di legittimità, funziona abbastanza bene quando si deve evitare un danno, cioè quando si disegnano le misure precauzionali dirette a prevenire l'evento lesivo. Se invece la questione ha il suo centro nel giudizio di responsabilità politica, esso, in quanto valutazione su comportamenti a effetti esauriti, non può essere anticipato da una tecnica orientata a includere la partecipazione popolare in luogo di quella politico-rappresentativa. Questa surrogazione non solo non è corretta perché colloca sullo stesso piano entità non allineabili, e le ragioni le abbiamo già illustrate rispetto alle Autorità indipendenti e alle Autorità private, ma in questo caso essa commette un errore sui tempi: quello del giudizio deve necessariamente seguire quello della condotta. Ne consegue che, se la misura suppletiva è dettata in anticipo, non può funzionare da elemento compensativo l'assenza di responsabilità, perché non si può giudicare ciò che non è ancora accaduto; il che lascia aperto il *vulnus* democratico per la fuga dall'*accountability*[98].

Vediamo se invece il modo di essere della procedura, osservata dalla tecnica per creare la regola giuridica, possa essere un valido sostituto al difetto di legittimazione politico-rappresentativa e quindi compensare l'assenza del mandato elettivo.

Indipendentemente dalla procedura scelta, è stato sostenuto con suggestive argomentazioni[99] che l'iter da solo garantirebbe la bontà della regola finale, supplendo così all'incapacità della tecnica a rappresentare e sintetizzare interessi e istanze politiche. Noi proponiamo una idea diversa: il susseguirsi di momenti prodromici in sé corretti non basta ad assicurare il *common good*. Infatti, la neutralità della procedura non ha come esito automatico la vittoria del bene sul male perché il suo *output* dipende da come la persona ha orientato, guidato e infine verificato la tecnica. Dunque, la procedimentalizzazione della

[98] A. CARDONE, *"Decisione algoritmica"* vs *Decisione politica?*, cit., 163.

[99] *Contra*: O. POLLICINO, *Regolazione e innovazione tecnologica nell''ordinamento della rete,* in *Riv. AIC*, 2, 2025, 152.

tecnica[100] offre sì una certezza: l'obbedienza a un iter preconfezionato prima del suo avvio. Ciò non è cosa da poco, ma neppure merita significati impropri. Se sopravvalutato, si promuove la procedura, ridotta a un involucro vuoto, a fonte legittimante le future regole che lei stessa si darà.

Questa procedimentalizzazione alla lunga distrae dai nodi politici nascosti sotto la perfezione procedurale: la difesa della separazione dei poteri e delle libertà fondamentali. La loro cura, invece, deve rimanere affidata al soggetto politico rappresentativo, che ha appunto il dovere di attivarsi in prima persona per garantirli ai suoi cittadini, astenendosi dal delegarne il compito a pallidi sostituti incapaci di rendere una prestazione di risultato.

Notiamo che il contesto ipertecnologico taglia le gambe allo Stato di diritto. Infatti, la divisione dei poteri cede davanti a una tecnica onnipotente, il nuovo Re Sole dell'età digitale, che si comporta da legislatore ma anche da amministratore e, se del caso, persino da giudice, grazie all'autodichia *in house,* sperimentata da molte piattaforme.

Né le libertà fondamentali hanno ricevuto un migliore trattamento in termini di certezza del diritto perché sono difese se, e nella misura in cui, coincidono con le regole meccanicistiche, che la tecnica di volta in volta detta secondo ripetitività statistica. Pertanto, la protezione delle libertà è affidata a un evento del tutto futuro e incerto, il cui accadere dipenderà dalla loro occasionale coincidenza con gli obiettivi privatistici delle Autorità di regolazione: "may claim that their objective are in line with the public interest, but whether or not this is so will depend on the frameworks in which they operate"[101].

In conclusione, questa fiducia incondizionata nella forza salvifica della procedura non solo non è sostenuta da prove reali, ma pregiudica la tenuta dell'ordine giuridico, perché il "procedural fetishism"[102] – cioè l'appiattimento acritico sulle virtù taumaturgiche di un iter

[100] E. CELESTE, *Digital Constitutionalism: a new systematic theorisation*, in *Int. Rev. L. Comput. Treaty*, 2019, 33, 76, e G. DE GREGORIO, *The Normative Power of Artificial Intelligence*, in *Ind. Jour. Glob. L. S.*, 2023, 30, 73.

[101] J. KAY - J. VICKERS, *Regulatory reform: an appraisal*, IN G. MAJONE (ed.), *Deregulation or re-regulation? Regulatory reform in Europe and the United States*, Londra, 1990, 239.

[102] T. R. TYLER - K. M. MCGRAW, *Ideology and the interpretation of personal experience: procedural justice and political quiescence*, in *Jour. Soc. Issue*, 1986, 42, illustrano il *procedural fetishism* come una pericolosa tendenza che genera nelle persone la convinzione che una decisione ingiusta sia corretta solo perché obbedisce a un preciso rigore formale.

chiuso nella sua perfezione formale[103], ma leggero nel contenuto – consegna nelle mani incontrollate dell'I.A. sia la cura delle nostre libertà che gli argini ai poteri costituiti, nell'illusoria speranza che "if we manage to do so just procedures will yield just outcomes"[104].
Riflettiamo infine sull'impatto della tecnica sul sistema delle fonti del diritto: essa rompe con il principio di gerarchia tra le fonti, essendo difficile sostenere che la regola tecnica, perfettamente coincidente con l'I.A., si collochi su un gradino immediatamente sottostante al Regolamento europeo, riempiendone gli spazi lasciati in bianco. La realtà è lontana da questa costruzione teorica perché la tecnica ha diversificato ulteriormente la scala delle fonti, aggiungendovi un nuovo gradino: oltre lo Stato e oltre le Autorità private. Si tratta di una *fictio iuris* che non ci deve trarre in inganno perché la tecnica assorbe la politica, la sovrasta e, sorretta da una patente di legittimità, detta statuizioni *ad personam* e concrete, obbedienti alla logica statistica, che non hanno nulla in comune con la libera scelta dei valori, riservata al decisore politico.
Chiediamoci ora se questo concreto provvedere dell'I.A. in luogo dell'astratto disporre crei problemi all'ordine giuridico, continuando il ragionamento appena concluso sulle fonti. Ebbene, per effetto della tecnica l'ordinamento tenderà ad avvicinarsi ai modelli di *common law*, con la particolarità che allo *stare decisis* delle pronunce dei giudici si sostituirà l'automatica statuizione della mente artificiale con effetti ostativi all'innovazione del diritto. Infatti, mentre il giudice a certe condizioni può rovesciare le precedenti pronunce perché sono emersi nuovi valori, l'I.A. non è allenata all'*overruling*. Ne consegue che la regola *biased* si cristallizzerà, riproponendo per l'avvenire il medesimo errore, ormai non più removibile[105].
In conclusione, la tecnica *as source* commette con una sola mossa più violazioni del diritto: calpesta la gerarchia delle fonti, logora la certezza del diritto, nega all'ordinamento l'attributo del *civil law*, perpetua ingiustizie per l'avvenire, mascherando il tutto con una finta legalità.

[103] J. ROCHELEAU, *Proceduralism*, in D. K. CHATTERJEE (ed.), *Encyclopedia Global Justice*, Salt Lake City, Springer, 2012, 906.

[104] E. CEVA, *Beyond legitimacy. Can Proceduralism say anything relevant about justice?* in *Critical Rev.Int'l soc and Pol. Phil*, 15, 12, 2012, 183-191.

[105] F. DONATI, *Intelligenza artificiale e giustizia*, in *Riv. AIC*, 1, 2020, 421; E. LONGO, *Giustizia digitale e Costituzione. Riflessioni sulla trasformazione tecnica della funzione giurisdizionale*, Milano, Franco Angeli, 2023, e, volendo, un nostro lavoro: *Giustizia e intelligenza artificiale: un equilibrio mutevole*, in *Riv. AIC*, 2, 2024, 95 ss.

In sintesi, questa tecnica cancella il valore democratico della *rule of law* a favore dell'arbitrio della *rule of tech*[106].

8.9 Verso il futuro

Quale prospettiva per il rapporto tra Tecnica e Politica?

Proponiamo una regola antica: a ciascuno il suo. Il vecchio – la norma che esaurisce in sé ogni manifestazione della giuridicità – e il nuovo – la tecnica che sta erodendo il terreno della prima a titolo non sempre lecito – si devono comporre in una struttura reticolare dell'ordinamento che riserva al soggetto politico sovranazionale il compito, non di regolare nel dettaglio le condotte umane, ma di ordinare il molteplice, di ricondurre la diversità ad unità; insomma, di "réaliser la mise en cohérence" e di assicurare la compatibilità delle norme di provenienza diversa.

Dunque, alla politica dei soggetti rappresentativi europei spetta di disegnare *ex ante* la struttura e l'interazione dei sistemi privati da controllare e da far sviluppare in modo che questi possano integrare con l'autoregolazione l'ordine giuridico maggiore. Da qui un ordine di intervento: la diffusione del potere normativo tra i nuovi beneficiari, Autorità indipendenti, Governi di interessi privati e Tecnica, non va lasciata andare per il suo verso, ma tratteggiata dal decisore politico nel profilo soggettivo e funzionale dei nuovi autori, individuando obiettivi e valori specifici da conseguire con le loro future regole. Questa divisione del lavoro conserva al soggetto politico sovranazionale la sua indeclinabile responsabilità di ricomporre il diritto imperativo e quello consensuale a sistema; mentre ai nuovi regolatori consente di continuare, non già di avviare *ex novo*, il discorso politico già impostato nelle sue linee portanti dal primo che li ha preceduti.

Un ordinamento che tende ad articolarsi su più spazi giuridici, che agiscono simultaneamente e che intervengono sugli stessi oggetti con forza diversa ha un unico punto di tensione: orientare ciascuna realtà giuridica a rispettare il patrimonio valoriale di ogni altra. Questo significherà, da un lato, evitare – stabilendo misure di reciproca tolleranza – che le idealità dei sottosistemi regolatori si annullino a vicenda e, dall'altro, garantire la coerenza

[106] M. ZALNIERIUTE, *Against procedural feticism: a call for a new digital constitution*, in *Ind. Jour. G.L.S.*, 39, 2, 2023, 253-256.

del diritto posto dai nuovi autori con i valori supremi dell'ordine politico, di cui l'Europa è il responsabile ultimo.

Quindi, l'unica modalità di cooperazione tra le fonti imperative tradizionali e le altre – pubbliche, private e meccaniche – si compie nella relazione di complementarità delle seconde alle prime, le sole capaci di coniugare la pluralità degli ordini derivati con la superiorità dell'ordinamento originario, che ben si può avvalere dei diritti a genesi endogena, purché indichi loro *ex ante* la direzione da intraprendere e intervenga *ex post* a correzione di eventuali esiti incompatibili con le linee guida impartite.

Allora il Legislatore sovranazionale dovrà fare meno e in modo diverso da come ha fatto finora.

Quanto al meno, dovrà eliminare le regole superate da una realtà che è andata avanti; quelle inutilmente invasive dell'autonomia d'impresa che non procurano benefici alla contendibilità dei mercati digitali; quelle discriminatorie nel pesare di più sui piccoli e meno sui grandi e quelle che disallineano situazioni da equiordinare in presenza di uno stesso servizio, prestato con mezzi alternativi. Il meno riguarderà anche le procedure, pensate per Autorità, nazionali ed europee, che agiscono autonome le une dalle altre, preferendo al solipsismo, che poi diventa confusione di procedure e discipline, il dialogo costruttivo per disegnare regole scarne di numero ma pluricomposte nei valori.

Quanto al diverso, il Legislatore dovrà riscrivere parte della disciplina sull'I.A. per correggerla nella sua torsione egoistica, dovuta all'uso improprio delle autodichiarazioni in luogo dei controlli pubblici imparziali, più idonei delle prime ad assicurare l'osservanza sostanziale, non presunta, delle regole precauzionali, poste a garanzia dei diritti fondamentali della persona. Questa attenzione all'individuo segnerebbe l'inizio di un umanesimo digitale, che non è solo un'espressione linguistica dimenticata in un *considerandum*, ma un valore che diventa regola di relazione tra la Politica e la Tecnica. La seconda non è più al servizio dei gruppi privati forti e ben dotati, ma si pone a fianco dei cittadini con le loro attese di confidare in una tecnica che non nasconde tranelli dietro un'apparente legalità, che è comprensibile nel suo iter logico e che prova a migliorare la vita di chi, individuo o collettività, sia rimasto indietro.

Analoga operazione andrebbe condotta sulle tante regole del DMA, che andrebbero sfoltite e riscritte, privilegiando gli accertamenti concreti della dominanza alle valutazioni legali tipiche. Questa regolazione, alleggerita dall'attributo della permanenza, dovrebbe ritirarsi con il compiersi dell'equilibrio competitivo. Il tutto andrebbe completato da un sistema sanzionatorio, coraggioso nei suoi rimedi strutturali, quelli che spezzano l'operatore verticalmente l'integrato e ridisegnato nelle sue pene pecuniarie, oggi carezze quasi gradite ai *Tech Barons*, lasciandoli indisturbati nella loro dominanza.

Solo a queste condizioni l'intreccio tra Tecnica e Politica, da incontro oggettivamente disaggregante l'ordine giuridico, potrebbe avvicinarsi a essere un fattore di crescita dell'ordinamento e diventare un luogo meritevole della fiducia del cittadino, perché controbilanciato dalla presenza di un decisore politico forte, architetto del sistema e, al tempo stesso, regolatore di ultima istanza.

9 Conclusioni: regolazione e governance nell'era dell'IA generativa

Andrea Renda

Il quadro normativo sull'intelligenza artificiale che si sta consolidando a livello europeo è il prodotto di una stagione legislativa straordinariamente intensa, caratterizzata da uno sforzo senza precedenti di anticipare i rischi di una tecnologia in rapida evoluzione. L'AI Act, il DSA, il GDPR e la proposta di Digital Omnibus, tra gli altri, compongono un mosaico giuridico complesso, la cui applicazione coerente dipende in misura determinante dalla qualità dell'enforcement a livello nazionale e dalla capacità delle autorità di settore di interpretare il proprio ruolo in modo proattivo e adattivo.

A fronte di questa complessità, i capitoli del presente Rapporto convergono su un messaggio comune: l'AGCOM non può limitarsi ad attendere che il quadro normativo si stabilizzi prima di intervenire. La finestra di opportunità per costruire capacità tecnica interna, stringere alleanze istituzionali e acquisire legittimità come regolatore dell'IA è aperta adesso, e rischia di chiudersi rapidamente se non colta con decisione.

9.1 Le prospettive di applicazione dell'AI Act e del DSA

Il processo di applicazione dell'AI Act procede a un ritmo molto più lento di quanto auspicato. Il rinvio all'agosto 2027 delle regole relative alle applicazioni ad alto rischio — deciso nell'ambito della proposta di Digital Omnibus — e le difficoltà nel completare i lavori di standardizzazione affidati al CEN-CENELEC lasciano aperta una lunga fase di incertezza normativa. In questo intervallo, il DSA rimane lo strumento di riferimento principale per l'enforcement dei contenuti generati o amplificati dall'IA sulle piattaforme digitali. L'AGCOM — in qualità di Digital Services Coordinator — è già oggi chiamata a esercitare poteri di vigilanza e sanzione che hanno una portata molto più ampia di quanto il dibattito pubblico abbia finora riconosciuto.

Il Digital Omnibus, se tradotto in legge nella forma attualmente proposta dalla Commissione, introdurrà una significativa centralizzazione dell'enforcement dell'IA in capo all'AI Office, soprattutto per i sistemi sviluppati a partire da GPAI o utilizzati dalle grandi piattaforme. Questo scenario renderà ancora più urgente la necessità per l'AGCOM di posizionarsi come interlocutore credibile dell'AI Office e come autorità di riferimento per i contesti nazionali di applicazione dell'IA nei settori di propria competenza.

9.2 La natura pervasiva dell'IA e le competenze regolatorie di AGCOM

Uno dei contributi più rilevanti di questo Rapporto è la dimostrazione — attraverso la mappatura sistematica del capitolo 3 — che l'intelligenza artificiale non è un fenomeno settoriale ma un paradigma trasversale che attraversa tutti gli ambiti di competenza di AGCOM. La gestione algoritmica delle reti di comunicazione elettronica, la personalizzazione dei contenuti media, la protezione del pluralismo e della tutela dei consumatori nelle interazioni con i sistemi automatizzati: queste non sono sfide separate, ma espressioni diverse di una trasformazione unitaria che richiede una risposta regolatoria altrettanto unitaria.

Questa trasversalità ha implicazioni pratiche precise. Primo, non è possibile affrontare le sfide dell'IA con strumenti pensati per un solo settore: occorre un framework regolatorio modulare, capace di declinare principi comuni — trasparenza, human oversight, non-discriminazione, protezione dei vulnerabili — in obblighi operativi specifici per ciascun contesto. Secondo, la supervisione dell'IA richiede capacità tecnica interdisciplinare che non può essere improvvisata: il benchmark comparativo con Ofcom, Arcom e BNetzA dimostra che le autorità che hanno investito precocemente in data scientist, AI auditor e laboratori tecnici interni ottengono risultati di enforcement significativamente superiori. Terzo, il coordinamento con le altre autorità nazionali — Garante privacy, AGCM, ACN — non è un optional istituzionale ma una condizione strutturale dell'efficacia del sistema.

9.3 L'IA generativa e l'agentic AI: nuove frontiere regolatorie

La rapida evoluzione verso sistemi di IA "agentici" — capaci di operare in modo autonomo, pianificare azioni complesse e interagire con l'ambiente digitale senza supervisione umana

continua — pone interrogativi radicalmente nuovi per l'applicazione dell'AI Act e per il ruolo delle autorità di settore. I sistemi agentici sfidano il presupposto fondamentale dell'approccio basato sul rischio: l'identificabilità e la prevedibilità dei casi d'uso. Un agente autonomo che negozia contratti, genera contenuti, gestisce infrastrutture di rete o interagisce con i consumatori esercita funzioni che ricadono simultaneamente in più categorie di rischio e in più ambiti di competenza regolatoria.

Per l'AGCOM, questo significa prepararsi a scenari in cui LLM e agenti IA operano come attori autonomi nell'ecosistema mediatico e delle comunicazioni: producendo contenuti, gestendo account sui social media, personalizzando offerte commerciali, moderando discussioni online. La risposta regolatoria non può attendere il completamento della cornice legislativa europea: è necessario avviare ora lo sviluppo di linee guida specifiche, protocolli di notifica e sistemi di audit adattati a questa nuova generazione di sistemi.

9.4 Regolamentare con l'IA: SupTech e governance adattiva

Il presente Rapporto non si limita ad analizzare come regolamentare l'IA, ma apre una prospettiva altrettanto rilevante: come usare l'IA per migliorare la qualità della regolazione stessa. Le tecnologie di supervisory technology (SupTech) — dai sistemi di monitoraggio automatizzato del traffico dati agli strumenti di analisi semantica per il rilevamento della disinformazione — possono trasformare la capacità di enforcement di AGCOM in modo radicale, consentendo un passaggio dagli audit periodici al monitoraggio continuo e dagli interventi ex post alle risposte anticipatorie.

La prospettiva del “law as code” — in cui le regole vengono formalizzate in modo computazionale e applicate automaticamente — è ancora lontana dalla realtà della regolazione italiana ed europea, ma rappresenta una traiettoria verso cui orientare gli investimenti in capacità tecnica e competenza istituzionale. Un’autorità che saprà usare l'IA per regolamentare sarà anche un’autorità meglio attrezzata per regolamentare l'IA.

9.5 Verso la sovranità tecnologica europea: l'EuroStack e il ruolo di AGCOM

Il contesto geopolitico in cui AGCOM è chiamata ad agire è segnato da una tensione strutturale tra la dipendenza dell'Europa dai grandi modelli di IA sviluppati negli Stati Uniti e in Cina, e l'ambizione di costruire un'infrastruttura tecnologica sovrana — il cosiddetto EuroStack — che garantisca resilienza, pluralismo e rispetto dei valori fondamentali europei. In questo contesto, le autorità indipendenti come AGCOM sono attori fondamentali: non solo come guardiani del rispetto delle regole, ma come promotori attivi di un ecosistema dell'innovazione che sia al tempo stesso competitivo e degno della fiducia dei cittadini.

9.6 Un contratto sociale digitale per l'era dell'IA

Il filo conduttore che attraversa tutti i capitoli di questo Rapporto è la convinzione che la tecnologia non sia neutrale: l'intelligenza artificiale, così come ogni altra forma di potere, richiede legittimazione democratica e accountability pubblica. Le autorità indipendenti — nate per proteggere l'interesse generale dalle distorsioni dei mercati e dei poteri privati — sono oggi chiamate a svolgere questa funzione in un ecosistema digitale dove le asimmetrie di potere, di informazione e di capacità si sono moltiplicate in modo esponenziale.

Questo Rapporto è, in ultima analisi, un contributo alla costruzione di un "contratto sociale digitale" per l'era dell'IA: un patto tra istituzioni, imprese e cittadini che riconosca i benefici straordinari della tecnologia, ma che non rinunci a governarne i rischi con strumenti proporzionati, efficaci e rispettosi dei diritti fondamentali. L'AGCOM può e deve essere uno dei protagonisti di questo processo, contribuendo con la propria expertise settoriale, la propria indipendenza istituzionale e la propria capacità di dialogo con gli attori europei e nazionali del sistema di governance dell'IA.

9.7 Principali raccomandazioni

Sulla base dei contributi raccolti, il Comitato formula le seguenti raccomandazioni prioritarie per AGCOM:

- **Potenziamento della capacità tecnica interna**: Investire in un team dedicato di AI auditor, data scientist e giuristi specializzati, in grado di condurre ispezioni algoritmiche

autonome e di sviluppare strumenti di testing indipendente sul modello delle esperienze di Ofcom e BNetzA. Il team dovrebbe essere strutturato come unità trasversale, con competenze in reverse engineering degli algoritmi di raccomandazione, analisi dei bias nei dataset e valutazione della conformità dei sistemi di IA agli obblighi dell'AI Act. L'investimento iniziale dovrebbe prioritizzare il reclutamento di profili tecnici senior e la stipula di convenzioni con università e centri di ricerca, sul modello del Data Science Lab di Ofcom e della partnership tra Arcom e l'INRIA francese.

- **Coordinamento istituzionale rafforzato**: Stringere protocolli di collaborazione formali con AgID, ACN, Garante privacy e AGCM, garantendo un enforcement coerente e proporzionato del quadro normativo europeo sull'IA nei settori di rispettiva competenza. I protocolli dovrebbero prevedere meccanismi operativi concreti: tavoli tecnici permanenti per la gestione dei casi di sovrapposizione di competenze, procedure di notifica reciproca in caso di apertura di procedimenti che coinvolgano sistemi di IA, e un sistema condiviso di segnalazione delle violazioni da parte degli utenti. L'obiettivo è superare il rischio di frammentazione che renderebbe inefficace l'intero sistema di enforcement, evitando che i soggetti regolati possano sfruttare i vuoti di coordinamento tra autorità.
- **Istituzione di un Osservatorio permanente sull'IA**: Creare un sistema di monitoraggio continuo degli impatti dell'IA nei settori regolati, con dashboard pubbliche, report periodici e strumenti di rilevamento automatizzato di bias, discriminazioni e violazioni del pluralismo informativo. L'Osservatorio dovrebbe operare con cadenza trimestrale, pubblicando un rapporto pubblico su un insieme di indicatori standardizzati — tra cui la diversità dei contenuti promossi dagli algoritmi di raccomandazione, la concentrazione del mercato pubblicitario, le pratiche di throttling e zero-rating nel traffico dati, e i casi di deepfake rilevati nelle comunicazioni politiche. Il modello è l'Observatoire des algorithmes di Arcom, che ha dimostrato come il monitoraggio permanente produca effetti deterrenti significativamente superiori agli audit periodici.
- **Adeguamento del sistema sanzionatorio**: Attivare le sanzioni previste dal DSA (fino al 6% del fatturato globale) e dall'AI Act (fino al 7%), superando i limiti strutturali della L. 249/1997 che rendono l'enforcement poco deterrente rispetto ai fatturati degli

operatori. L'adeguamento richiede un intervento normativo che aggiorni le massime edittali previste dalla legge istitutiva di AGCOM, armonizzandole con le soglie europee e garantendo la proporzionalità rispetto alla dimensione economica dei soggetti regolati. In parallelo, le procedure sanzionatorie dovrebbero essere riformate per ridurre i tempi di definizione dei procedimenti e introdurre la possibilità di misure cautelari immediate nei casi di rischio sistemico, evitando che il lento decorso delle procedure neutralizzi l'effetto deterrente delle sanzioni.

- **Sviluppo di un framework regolatorio modulare**: Elaborare codici di condotta settoriali, linee guida e obblighi di trasparenza algoritmica armonizzati tra i diversi ambiti di competenza, fondati su principi comuni adattabili all'evoluzione tecnologica (trasparenza, human oversight, non-discriminazione, protezione dei vulnerabili). Il framework dovrebbe articolarsi su tre livelli: un insieme di principi orizzontali validi per tutti i settori regolati; obblighi specifici declinati per ciascun ambito — comunicazioni elettroniche, media audiovisivi, servizi postali, tutela dei consumatori; e un meccanismo di revisione annuale che incorpori le evidenze emerse dall'Osservatorio e le novità del quadro normativo europeo. Questo approccio consente di non moltiplicare gli strumenti regolativi, progettandoli invece una volta sola con la necessaria cura e rendendoli adattabili ai diversi contesti applicativi.
- **Partecipazione attiva ai network europei**: Contribuire proattivamente ai gruppi di lavoro BEREC ed ERGA, co-costruendo strumenti comuni per l'audit dell'IA e l'armonizzazione delle pratiche di enforcement a livello europeo. AGCOM dovrebbe presentarsi in questi contesti non come soggetto che riceve indicazioni, ma come portatore di un'esperienza regolatoria specifica — quella di Digital Services Coordinator in un paese di grandi dimensioni — e come promotore di un'accelerazione dei tempi di sviluppo degli standard condivisi. In particolare, la partecipazione al BEREC AI Toolkit e ai lavori ERGA su deepfake e pluralismo algoritmico offre l'opportunità di influenzare gli standard europei che si applicheranno anche al mercato italiano, con vantaggi competitivi significativi per un'autorità che abbia già sviluppato capacità tecnica interna.
- **Sperimentazione e SupTech**: Avviare la sperimentazione di strumenti di supervisory technology (SupTech) per il monitoraggio automatizzato dei mercati regolati, aprendo la

strada a forme di enforcement basate su dati in tempo reale invece che su ispezioni periodiche. L'AGCOM dovrebbe istituire una regulatory sandbox dedicata ai sistemi di IA che operano nei settori di propria competenza — comunicazioni elettroniche, media audiovisivi, servizi digitali e postali — consentendo a imprese, startup e centri di ricerca di testare soluzioni innovative in un ambiente controllato, con deroga temporanea e circoscritta agli obblighi regolatori ordinari. Il modello di riferimento è quello già previsto dall'art. 57 dell'AI Act per le autorità nazionali competenti, che l'AGCOM potrebbe attivare in coordinamento con AgID e ACN, posizionandosi come primo laboratorio regolatorio italiano sull'IA applicata all'ecosistema mediatico e delle comunicazioni.