

EMILIO TOSI

INTELLIGENZE ARTIFICIALI E DATI PERSONALI TRA AI ACT E GDPR: PRIME RIFLESSIONI SU VULNERABILITÀ DIGITALE E CREPUSCOLO DEL CONSENSO

SOMMARIO: 1. Premessa. — 2. AI Act e GDPR: un dialogo necessario. — 3. La complessità dell'AI Act e della governance dell'AI: il rischio applicativo. — 4. L'eterogeneità delle intelligenze artificiali e il problema definitorio — 5. La classificazione del rischio delle AI: inaccettabile, alto, limitato e minimo. — 6. Dati e governance dei dati delle AI ad alto rischio: l'art. 10 dell'AI Act. — 7. Sorveglianza umana delle AI ad alto rischio: art. 14 AI Act e il principio "human in the loop". — 8. L'AI Act tra iporegolazione e iperregolazione. — 9. L'elasticità applicativa dei principi generali dell'AI Act: una proposta di lettura sistematica e assiologica. — 10. Vulnerabilità digitale della persona nell'epoca delle AI. — 11. Tutela della persona e crepuscolo del consenso nel contesto asimmetrico digitale delle AI: l'emersione progressiva della funzione protettiva della conformazione eteroregolamentare.

1. PREMESSA.

La recente entrata in vigore del Regolamento UE 13 giugno 2024, n. 1689 noto come *Artificial Intelligence Act* — in seguito, per brevità, AI Act — richiede attenta riflessione in ordine alla tutela della persona e dei dati personali: con particolare riferimento alla rilevanza del consenso in tale contesto tecnologico avanzato e all'emersione di nuove vulnerabilità digitali¹.

I dati, in particolare i dati personali, come noto, sono al centro e costituiscono elemento fondante, non da oggi, della *data economy*.

Tale rinnovata centralità si conferma, invero, anche nella pagina tecnologica più recente dell'umanità: quella relativa allo

¹ Si veda sul tema dei complessi rapporti tra AI Act e GDPR la recente raccolta di studi venticinquesimo volume della *Coltana internazionale* di studi giuridici privatistici, *Diritto delle nuove tecnologie*, di-

retta da Vincenzo Franceschelli ed Emilio Tosi: E. Tosi (a cura di), *Persona, dati personali, algoritmi tra GDPR e AI ACT*, Giuffrè Lefebvre, Milano, 2025.

sviluppo e impiego *trasparente, sicuro e responsabile* delle nuove sofisticate e rivoluzionarie entità algoritmiche denominate convenzionalmente *intelligenze artificiali*.

Denominazione convenzionale più descrittiva che rigorosa non essendo univocamente qualificante a fronte, come vedremo, di carenza di uniforme e condivisa definizione.

Lo sviluppo e ottimizzazione delle intelligenze artificiali richiede continuo aggiornamento dei *pattern statistici* e dei parametri su cui esse si fondano: in buona sostanza le AI evolvono, autoapprendono, sviluppando forme più o meno marcate di automazione in taluni casi di autonomia operativa — si pensi alle emergenti AI agentiche — e si addestrano in permanenza fagocitando dati, anche personali, spesso in modo opaco per non dire occulto con buona pace del GDPR.

Autonomia significa capacità di funzionare senza intervento, controllo o supervisione umana che è, come noto, concetto ben diverso dalla semplice automazione che si riferisce ad un processo o a un sistema che solo a determinate condizioni funziona senza intervento umano.

L'antropomorfizzazione delle AI, dettata da preminenti esigenze commerciali oltre che per attrarre ingenti investimenti dei grandi capitali finanziari globali, genera permanente euforia e crea false illusioni: da un lato tende a sovrapporre ingannevolmente autonomia e automazione, concetti ben diversi tra loro; dall'altro tende a rappresentare tali entità come "pensanti" laddove si riscontra solo una simulazione linguistica di pensiero basata su modelli statistici sofisticati quanto inconsapevoli.

In questa fase dirompente dello sviluppo tecnologico emergono nel mercato globale le istanze delle *big tech* di regolazione *soft* basata sull'autodisciplina e l'etica: tale prospettiva pare riduttiva e inaccettabile considerati gli interessi in gioco.

Per la semplice ragione che così come era già successo per la regolamentazione del *world wide web di internet* superata l'euforia iniziale non si può non fare i conti con la realtà normativa cogente a tutela dei diritti fondamentali della persona: il GDPR è pienamente applicabile e richiede puntuale *compliance* anche nel mutato scenario tecnologico digitale agito dalle nuove entità artificiali.

Il tratto comune della politica europea della normazione digitale pare essere ispirato dall'encomiabile tentativo di riequilibrare il problematico rapporto tra persona e tecnologia, libertà e innovazione, Stato e mercato.

Tale principio ispiratore si rinviene non solo nel *General Data Protection Regulation* ma anche nei più recenti interventi normativi della UE in materia di innovazione digitale ormai un vero e proprio *corpus iuris digitalis*: emblematici in tal senso, senza pretese di esaustività, il *Digital Services Act*, *Digital Markets Act*,

Data Governance Act, *Data Act* e da ultimo l'*Artificial Intelligence Act*, la legge europea sull'intelligenza artificiale.

L'AI Act si segnala all'attenzione dei giuristi — e non solo — quale epocale tentativo di disciplinare l'intelligenza artificiale, *recte* le intelligenze artificiali, per renderle strumenti responsabili, affidabili, trasparenti e sicuri nel tentativo di salvaguardare una prospettiva antropocentrica di un nuovo umanesimo digitale fondato sulla tutela della dignità e dei diritti fondamentali della persona ².

Intenti nobilissimi e condivisibili che purtroppo si scontrano con la triste realtà di mercati che abdicano a disciplinare il fenomeno consegnano il genere umano al rischio di sviluppi eterodossi pericolosi per le libertà individuali e democratiche.

Gli Stati sovrani non possono e non devono cedere alla facile scorciatoia di affidarsi esclusivamente all'autodisciplina, consegnando al mercato oligopolistico e al potere autoregulatorio privatistico dell'innovazione concentrato in poche *big tech* e *gatekeeper*, la definizione del perimetro, di rilevanza costituzionale, di diritti e libertà.

Attenzione però ai pericolosi rischi di iper-regolazione e *overlapping regulation* che possono produrre effetti indesiderati in termini di efficacia normativa ed efficienza dei mercati regolati.

L'AI Act dialoga necessariamente con il GDPR: precisamente è sottoposto al GDPR che non intende derogare per nulla e che anzi si prefigge di osservare e applicare pur con tutte le difficoltà del caso.

Il nuovo regolamento europeo non intende, infatti, pregiudicare l'applicazione della disciplina relativa al trattamento dei dati personali posta dal GDPR, inclusi i compiti e i poteri delle Autorità di controllo indipendenti competenti a monitorare la conformità con tali strumenti.

Autorità di controllo che si apprestano a diventare nuovi presidi di legalità digitale — non più solo a tutela dei dati personali ma della persona *tout court* — dalle pratiche vietate dall'AI Act.

Come ci ricorda proprio l'AI Act nei considerando preliminari, la disciplina dell'AI introdotta dal Regolamento in parola “lascia impregiudicati gli obblighi dei fornitori e dei *deployer* dei sistemi di IA nel loro ruolo di titolari del trattamento o responsabili del

² Per una disamina articolata della complessa orditura normativa dell'AI Act si veda A. MANTELERO, G. RESTA, G.M. RICCIO (a cura di), *Intelligenza artificiale commentario. Regolamento (UE) 2024/1689 - AI ACT Linee guida Commissione europea pratiche vietate Disegno di legge sull'intelligenza artificiale*, Torino, 2025. Si vedano inoltre sul tema della regolamentazione giu-

ridica delle AI: G. ALPA (a cura di), *Diritto e intelligenza artificiale: profili generali, soggetti, contratti, responsabilità civile, diritto bancario e finanziario, processo civile*, Pisa, 2020, *passim*; G. FINOCCHIARO, *Diritto dell'intelligenza artificiale*, Bologna, Zanichelli, 2024; A. GENTILI, *Regole per l'intelligenza artificiale*, in *Contratto e Impresa*, 4, 2024, pp. 1043-1069.

trattamento derivanti dal diritto dell'Unione o nazionale in materia di protezione dei dati personali, nella misura in cui la progettazione, lo sviluppo o l'uso di sistemi di IA comportino il trattamento di dati personali”.

E ancora chiarisce che l'interessato dal trattamento dei propri dati continuerà a beneficiare del diritto stabilito dall'art. 22 del GDPR di non essere sottoposto a una decisione basata unicamente sul trattamento automatizzato, compresa la profilazione, che produca effetti giuridici che lo riguardano o che incida in modo analogo significativamente sulla sua persona.

Ma procediamo con ordine.

2. AI ACT E GDPR: UN DIALOGO NECESSARIO.

L'AI Act europeo si propone, in linea di principio, di migliorare il funzionamento del mercato europeo istituendo un quadro giuridico uniforme in particolare per quanto riguarda:

— lo sviluppo, l'immissione sul mercato, la messa in servizio e *l'uso responsabile* di sistemi di intelligenza artificiale, in conformità ai valori dell'Unione Europea;

— la diffusione di un'intelligenza artificiale (IA) *antropocentrica, sicura, trasparente e affidabile* garantendo nel contempo un livello elevato di protezione della salute, della sicurezza e dei diritti fondamentali sanciti dalla Carta dei diritti fondamentali dell'Unione europea;

— prevenendo la disinformazione, tutelando la democrazia, lo Stato di diritto e la protezione dell'ambiente, contro gli effetti nocivi dei sistemi di IA nell'Unione nonché promuovere l'innovazione digitale assicurando un controllo umano di ultima istanza.

Tra le molteplici finalità di questo Regolamento epocale meritano di essere evidenziati i seguenti profili:

— le garanzie per i sistemi di intelligenza artificiale usati per finalità generali e specifiche regole per IA considerate ad alto rischio;

— limiti all'uso dei sistemi di identificazione biometrica da parte delle forze dell'ordine;

— divieto di utilizzo di sistemi di credito sociale o per manipolare e sfruttare le vulnerabilità degli utenti;

— tutela dei consumatori mediante diritto a presentare reclami e ricevere spiegazioni rilevanti

La riforma europea introduce, inoltre, specifici obblighi di trasparenza e spiegazione: per tutelare diritti dei consumatori, dei lavoratori più in generale dei cittadini e in ultima istanza i mercati, quasi un ideale punto di partenza di una nuova governance tecnologica e di un rinnovato contratto sociale.

L'AI Act — in questo richiamando il modello *risk based* delineato dal GDPR — stabilisce obblighi differenziati sempre più

stringenti sulla base dei possibili rischi e del livello d'impatto sui diritti fondamentali della persona.

Emerge il tentativo di contrastare l'elevata opacità, la complessità, la faziosità, un intrinseco grado di strutturale, quanto ineliminabile, imprevedibilità — le *allucinazioni algoritmiche* — e un comportamento tendenzialmente autonomo, anche solo in parte, di taluni avanzati sistemi di IA, onde garantirne la compatibilità con i diritti fondamentali e agevolare l'applicazione delle norme giuridiche europee.

Solo i primi esiti applicativi e i primi pronunciamenti della Corte di Giustizia della UE ci potranno dire qualcosa di più relativamente all'efficacia della nuova disciplina europea.

Come noto la diffusione di tecnologie e sistemi di AI agevola le pratiche manipolative opache rendendo *ipso facto* vulnerabile la persona di fronte all'artificio tecnologico.

A maggior ragione nel nuovo contesto digitale agito da nuove eterogenee e molteplici entità artificiali — generative, predittive, agentiche, sistemi esperti, *stand alone* o contenute in altri prodotti e servizi digitali — in grado di generare contenuti, dialoghi e assumere decisioni nei più svariati settori del vivere umano, con nuove espressioni di “intelligenza stocastica”, “intelligenza algoritmica”, mai del tutto attendibile, che nulla hanno in comune con l'*intelligenza naturale* in quanto non certamente consapevoli, intenzionali né coscienti.

Si comprende, dunque, trattarsi di una mera forzatura commerciale e promozionale finalizzata a solleticare gli umori dei mercati e la curiosità degli utenti finali: l'antropomorfizzazione forzata presente nel sintagma *intelligenza artificiale* è solo un riflesso di *intelligenza apparente* che nulla ha in comune con quella naturale.

Una simulazione linguistica di pensiero — accattivante e compiacente, per tale ragione decettivo — non un pensiero vero e proprio: simulazione basata su modelli statistici complessi, tuttavia, privi di qualsiasi comprensione, consapevolezza e intenzionalità.

Si registra, invero, in tale nuovo contesto digitale futuristico, la frattura definitiva tra *agere* ed *intelligere*, con il rischio che l'umanità metta definitivamente l'*intelligere* al servizio dell'*agere*.

Nei prossimi anni il pensiero rischia di divenire superfluo, un lusso per pochi eletti.

In tale prospettiva distopica la comprensione diviene secondaria rispetto all'azione al risultato ottenuto: non a caso i tanto magnificati *large language model* sono stati definiti causticamente *stochastic parrots*, *pappagalli stocastici*.

Non più, dunque, l'umano *intelligere deinde agere* — come sottolineato dall'attento filosofo italiano Luciano Floridi — ma un

nuovo *pensiero artificiale* che si fonda su un inedito *agere sine intelligere*³.

Il rischio sociale e antropologico è che l'enorme successo di questo *agere artificiale* possa portarci a pensare che l'*intelligere* sia divenuto definitivamente superfluo in un orizzonte di pura efficienza algoritmica dominato da potenti *idiots savant* digitali.

In tale contesto digitale fortemente asimmetrico e disumanizzato (ri)emerge il tema della fragilità umana, della vulnerabilità, dell'adeguata ed efficace tutela della persona e dei diritti fondamentali della personalità: temi certo non nuovi ma che nel contesto artificiale citato assumono contorni giuridici inediti e dirompenti meritevoli di adeguato approfondimento.

Lo scrivente segnala, sommessamente, da tempo — con particolare riferimento al contesto relativo al trattamento dei dati personali sebbene tali riflessioni possano ben valere anche oltre tali iniziali confini speculativi — la crisi dell'effettiva autodeterminazione informativa nel contesto digitale attuale vortiginosamente asimmetrico. Emblematica l'emersione della *la fatica da consenso* che ne indebolisce la funzione⁴, l'inadeguatezza della libera e consapevole autodeterminazione rispetto ai *nudge tech*, *dark pattern* e *AI manipulative*.

Si registra — ormai inesorabilmente — il progressivo *crepuscolo dell'autodeterminazione informativa*, la *crisi del consenso individuale, libero e consapevole*: in buona sostanza del *consenso* espressione del dogma ottocentesco della supremazia della libertà individuale e dell'autonomia contrattuale⁵.

I tempi sono invero maturi per una rilettura, attualizzata al contesto tecnologico digitale più avanzato, della funzione del consenso individuale al trattamento dei dati personali — e più in

³ L. FLORIDI-F. CABITZA, *Intelligenza Artificiale. L'uso delle nuove macchine*, Bompiani, Milano 2021, spec. pp. 166 e 177.

⁴ Persino l'EDPB registra il fenomeno emergente della *consent fatigue* nel documento EDPB, Guidelines 05/2020 on consent under regulation 2016/679, 2020, nota 30, paragrafo 87. Si veda sul punto anche l'interessante iniziativa *Voluntary cookie pledge project* lanciato dalla Commissione UE in data 19 novembre 2023 che ad oggi, tuttavia, non ha incontrato il consenso di *few* degli *stakeholders*.

⁵ I. COFONE, *The Privacy Fallacy: Harm and Power in the Information Economy*, Cambridge University Press, 2023; D. SOLOVE, *The Myth of the Privacy Paradox*, in *George Washington Law Review*, 2021, 1 ss.; D. SOLOVE, *Kafka in the Age of AI and the Futility of Privacy as Control*, in *Boston University Law Review*, 2024, 1021 ss. Sulla *consent fatigue* si veda in partico-

lare D. SOLOVE, *Murkey consent: an approach to the fictions of consent in Privacy Law*, in *Boston University Law Review*, 2024, 594 ss. Spec. 622-623, Secondo cui: "The flip side of not enough choices is too many choices. Too granular an approach is overwhelming. Yet the GDPR requires consent to be quite granular and specific. (...) Privacy choices at many companies have become more granular. Social media settings, for example, used to have a few choices, but now they have more settings than an airplane cockpit. A cost of this trend is that privacy choices are overwhelming and perplexing. (...). Presenting people with so many choices might be appropriate because privacy is complex and does involve many choices. But it can make consent more onerous and lead to mistakes. (...) Obtaining informed consent for every activity involving privacy is impractical. Too much effort must go into educating people, and

generale del consenso contrattuale — sempre più recessivo e inidoneo a tutelare pienamente, di per sé, il soggetto, la persona, in contesto tecnologico digitale fortemente asimmetrico e interferente con la libera manifestazione della volontà, pur senza arrivare alle estreme conseguenze dell'annullabilità per vizi della volontà.

Evenienza che, peraltro, nei casi più gravi, non si può certo escludere, anzi, a tutela della libera e consapevole manifestazione della volontà.

Così come non si può non considerare utilmente *de iure condendo* il ricorso a nullità protettive di nuovo conio, per evitare le incertezze applicative dell'interpretazione protettiva. A tale risultato si può, invero, già oggi approdare tramite lettura assiologica costituzionalmente orientata, finalizzata alla protezione rafforzata della libertà contrattuale delle persone nel caso di contrattazione manipolata dall'AI, manipolata sino a tal punto da potersi considerare non semplice vizio ma mancanza di volontà *tout court*.

Occorre, ad avviso dello scrivente, valorizzare in un crescente contesto disumanizzato e altamente ingannevole in cui finzione e realtà, vero e falso tendono a confondersi — il fenomeno distortivo della realtà dei deepfake ne costituisce significativo esempio — una lettura assiologica costituzionalmente orientata protettiva della persona.

In particolare nel contesto tecnologico dominato dalle AI tale lettura protettiva può essere efficacemente assicurata tramite conformazione, sempre più stringente, imposta dalle regole digitali europee e più in generale, anche in assenza di un rapporto contrattuale: rafforzamento dei precetti conformativi centrati su responsabilizzazione *ex ante*, il ben noto principio di accountability già statuito dal GDPR, responsabilità, tendenzialmente oggettiva, *ex post* — particolarmente appropriata anche in rela-

there are so many decisions that the effort would put people into a permanent privacy school (with no summer recess). Using the express consent approach is doomed because obtaining such consent does not scale. Thousands of organizations collect, use, and transfer an individual's personal data, so individuals would be deluged with countless requests for consent. In addition, each organization might engage in a wide array of different activities involving personal data at different times, so they would be issuing a stream of consent requests. The result is a tremendous and unwanted burden hoisted upon the individual. (...). Some commentators argue that the fact that many people do not want to think about privacy means

they do not care about privacy. They invoke what is called the "privacy paradox" to contend that people's consent validly reflects their lack of concern about privacy. But as I have argued elsewhere, the privacy paradox is a myth. That people fail to think about privacy does not indicate they do not care". Da ultimo D. SOLOVE-W. HARTZOG, *The Great Scrape: The Clash Between Scraping and Privacy*, California Law, 2025. Si veda in tal senso anche E. CORREN, *The Consent Burden in Consumer and Digital Markets*, 36 Harv. J.L. & Technology, 2023, pp. 561 ss. Sul tema dell'autonomia contrattuale in generale si veda inoltre per tutti: F. Criscuolo, *Autonomia negoziale e autonomia contrattuale*, Napoli, 2008, *passim*.

zione alle responsabilità derivanti dalla creazione e utilizzo di AI in generale e non solo ad alto rischio — e infine sanzioni dissuasive e proporzionate alla gravità dei precetti violati.

Lo sviluppo e l'utilizzo di AI — come si è già rilevato *supra* — comporta il trattamento di ingenti quantità di dati in generale e anche di dati personali sia in fase di addestramento che in fase di utilizzo.

Sono, pertanto, molteplici le intersezioni applicative e normative tra GDPR e AI Act.

Il considerando 10 dell'AI Act ricorda che: *“Il diritto fondamentale alla protezione dei dati personali è garantito in particolare dai regolamenti (UE) 2016/679 e (UE) 2018/1725 del Parlamento europeo e del Consiglio e dalla direttiva (UE) 2016/680 del Parlamento europeo e del Consiglio. La direttiva 2002/58/CE del Parlamento europeo e del Consiglio tutela inoltre la vita privata e la riservatezza delle comunicazioni, in particolare stabilendo le condizioni per l'archiviazione di dati personali e non personali e l'accesso ai dati in apparecchi terminali. Tali atti giuridici dell'Unione costituiscono la base per un trattamento sostenibile e responsabile dei dati, anche nei casi in cui gli insiemi di dati comprendono una combinazione di dati personali e non personali. Il presente regolamento non mira a pregiudicare l'applicazione del vigente diritto dell'Unione che disciplina il trattamento dei dati personali, inclusi i compiti e i poteri delle autorità di controllo indipendenti competenti a monitorare la conformità con tali strumenti”*⁶.

Pure significativo il considerando 68 dell'AI Act secondo cui: *“Il diritto alla vita privata e alla protezione dei dati personali deve essere garantito durante l'intero ciclo di vita del sistema di IA. A tale riguardo, i principi della minimizzazione dei dati e della*

⁶ Considerando 10 AI Act: “Il diritto fondamentale alla protezione dei dati personali è garantito in particolare dai regolamenti (UE) 2016/679 e (UE) 2018/1725 del Parlamento europeo e del Consiglio e dalla direttiva (UE) 2016/680 del Parlamento europeo e del Consiglio. La direttiva 2002/58/CE del Parlamento europeo e del Consiglio tutela, inoltre, la vita privata e la riservatezza delle comunicazioni, in particolare stabilendo le condizioni per l'archiviazione di dati personali e non personali e l'accesso ai dati in apparecchi terminali. Tali atti giuridici dell'Unione costituiscono la base per un trattamento sostenibile e responsabile dei dati, anche nei casi in cui gli insiemi di dati comprendono una combinazione di dati personali e non personali. Il presente regolamento non mira a pregiudicare l'applicazione del vigente diritto dell'Unione che

disciplina il trattamento dei dati personali, inclusi i compiti e i poteri delle autorità di controllo indipendenti competenti a monitorare la conformità con tali strumenti. Inoltre, lascia impregiudicati gli obblighi dei fornitori e dei *deployer* dei sistemi di IA nel loro ruolo di titolari del trattamento o responsabili del trattamento derivanti dal diritto dell'Unione o nazionale in materia di protezione dei dati personali, nella misura in cui la progettazione, lo sviluppo o l'uso di sistemi di IA comportino il trattamento di dati personali. È, inoltre, opportuno chiarire che gli interessati continuano a godere di tutti i diritti e le garanzie loro conferiti da tale diritto dell'Unione, compresi i diritti connessi al processo decisionale esclusivamente automatizzato relativo alle persone fisiche, compresa la profilazione. Norme armonizzate per l'immissione sul mercato, la

protezione dei dati fin dalla progettazione e per impostazione predefinita, sanciti dal diritto dell'Unione in materia di protezione dei dati, sono applicabili nel trattamento dei dati personali. Le misure adottate dai fornitori per garantire il rispetto di tali principi possono includere non solo l'anonimizzazione e la cifratura, ma anche l'uso di tecnologie che consentano di inserire algoritmi nei dati e di addestrare i sistemi di IA senza trasmissione tra le parti o copia degli stessi dati grezzi o strutturati, fatti salvi i requisiti in materia di governance dei dati di cui al presente regolamento". In ogni caso pare fondamentale il richiamo dell'art. 2.7 dell'AI Act che sgombera il campo da ogni dubbio interpretativo in ordine alla gerarchia delle fonti precisando che il GDPR continua ad applicarsi e prevalere sull'AI Act quando le attività delle intelligenze artificiali comportano anche trattamento dei dati personali ⁷:

"Il diritto dell'Unione in materia di protezione dei dati personali, della vita privata e della riservatezza delle comunicazioni si applica ai dati personali trattati in relazione ai diritti e agli obblighi stabiliti dal presente regolamento. Il presente regolamento lascia impregiudicati il regolamento (UE) 2016/679 o (UE) 2018/1725 o la direttiva 2002/58/CE o (UE) 2016/680, fatti salvi l'articolo 10, paragrafo 5, e l'articolo 59 del presente regolamento".

Emblematico, inoltre, il richiamo allo specifico divieto posto dall'Articolo 22 GDPR ai processi decisionali automatizzati relativi alle persone fisiche, compresa la profilazione, che si applicherà anche alle decisioni algoritmiche delle intelligenze artificiali ⁸.

L'interessato ha il diritto di non essere sottoposto a una decisione basata unicamente sul trattamento automatizzato, compresa

messa in servizio e l'uso dei sistemi di IA istituiti a norma del presente regolamento dovrebbero facilitare l'efficace attuazione e consentire l'esercizio dei diritti degli interessati e di altri mezzi di ricorso garantiti dal diritto dell'Unione in materia di protezione dei dati personali nonché degli altri diritti fondamentali".

⁷ L'Art. 10.5 dell'AI Act stabilisce quanto segue: "Nella misura in cui ciò sia strettamente necessario al fine di garantire il rilevamento e la correzione delle distorsioni in relazione ai sistemi di IA ad alto rischio in conformità del paragrafo 2, lettere f) e g), del presente articolo, i fornitori di tali sistemi possono eccezionalmente trattare categorie particolari di dati personali, fatte salve le tutele adeguate per i diritti e le libertà fondamentali delle persone fisiche".

⁸ S. ORLANDO, *Nove certezze su decisioni algoritmiche, diritto alla spiegazione e tutela del consumatore*, in E. Tosi (a cura di), *Persona, dati personali, algoritmi tra GDPR e AI Act. Le nuove frontiere del*

diritto civile nella società delle "intelligenze" artificiali, Milano, 2025, 141 ss. Sul caso specifico del credito scoring come decisione basata su trattamento automatizzato si veda il caso SCHUFA Holding, Corte di Giustizia UE, 7 dicembre 2023 (C-634/21). Tale provvedimento statuisce che un processo di decisioni automatizzate che influisce in modo decisivo sulla stipula, l'esecuzione o la cessazione di un rapporto contrattuale costituisce una "decisione basata unicamente sul trattamento automatizzato" ex art. 22 GDPR. Lo «scoring» è un metodo statistico matematico che consente di determinare una previsione sulla probabilità di un comportamento futuro, come il rimborso di un credito. La Corte dichiara che tale attività deve essere considerata un «processo decisionale automatizzato» in linea di principio vietato dal GDPR, qualora i clienti della SCHUFA, quali le banche, gli attribuiscono un ruolo determinante nell'ambito della concessione di crediti.

la profilazione, che produca effetti giuridici che lo riguardano o che incida in modo analogo significativamente sulla sua persona.

Il divieto non si applica nel caso in cui la decisione:

a) sia necessaria per la conclusione o l'esecuzione di un contratto tra l'interessato e un titolare del trattamento;

b) sia autorizzata dal diritto dell'Unione o dello Stato membro cui è soggetto il titolare del trattamento, che precisa altresì misure adeguate a tutela dei diritti, delle libertà e dei legittimi interessi dell'interessato;

c) si basi sul consenso esplicito dell'interessato.

Si tratta dunque di rafforzare — in prospettiva protettiva rafforzata della persona e dei diritti fondamentali della personalità — anche con riferimento alla disciplina normativa delle AI, i principi di *accountability*, *compliance*, *audit indipendente e responsabilità* — in senso oggettivo — coadiuvati da adeguato *apparato sanzionatorio* pubblico.

A tutela non solo e non tanto dell'interesse della singola persona ma a presidio irrinunciabile e non negoziabile del diritto inviolabile della dignità personale oltre che nella prospettiva metaindividuale di interesse generale dell'ordinamento giuridico alla protezione del genere umano.

3. LA COMPLESSITÀ DELL'AI ACT E DELLA GOVERNANCE DELL'AI: IL RISCHIO APPLICATIVO.

La complessità dell'orditura normativa dell'AI Act fonda le crescenti critiche di *iper-regolazione*: troppi obblighi — in parte sovrapposti a quelli di altri significativi referenti normativi del settore digitale *ex multis Digital Services Act, Data Act e GDPR* solo per richiamare i principali senza pretese di esaustività — rischiano di ostacolare lo sviluppo del settore, ammesso, e non concesso, che ancora vi siano margini per la “vecchia” Europa.

Al tempo stesso la formulazione di alcuni precetti generica, e non sempre univoca, pare risentire di una formulazione più simile a quella di una *cripto* direttiva che di un vero e proprio regolamento.

Forse la fretta di raggiungere l'intesa su un testo normativo condiviso non ha giovato alla chiarezza e sistematicità dell'AI Act e al coordinamento con altri referenti normativi fondamentali del citato *corpus iuris digitalis* europeo.

Essenziale anche in questo caso, come per il GDPR, il ruolo interpretativo e chiarificatore della Corte di Giustizia della UE.

Le nuove norme armonizzate dell'AI Act per l'immissione sul mercato, la messa in servizio e l'uso dei sistemi di IA dovrebbero facilitare l'efficace attuazione e consentire l'esercizio dei diritti degli interessati in materia di protezione dei dati personali nonché degli altri diritti fondamentali.

Già ci sono tutti i segnali globali che non si ripeterà il successo del GDPR: non ci sarà un *Bruxelles effect* analogo a quello straordinario dell'esperienza normativa europea in materia di protezione dei dati personali.

Anzi si rischia semmai il *Washington effect*: incombe il digital omnibus propopsal della Commissione UE, tra semplificazione e deregulation.

L'impianto dell'AI Act è, invero, meno incisivo più disordinato troppo burocratico e a tratti ridondante con una pletora di adempimenti anteriori e successivi all'immissione nel mercato europeo.

A tratti financo poco chiaro in numerosi aspetti della governance dell'AI Act e al riparto di competenze tra Commissione, AI Board e AI Office, da un lato, e Autorità nazionale di notifica e Autorità nazionale di Vigilanza, dall'altro, da raccordare con i poteri delle altre Autorità nazionali di regolazione di specifici settori (concorrenza, mercati finanziari, comunicazioni, consumatori, cybersicurezza oltre a settori vigilati quali bancario, finanziario, assicurativo) solo per menzionare i principali attori: addirittura carente della tassonomia dei poteri della Commissione che impegna l'interprete nella ricostruzione senza una norma *ad hoc* che pare scomparsa nella fretta di approvare il testo per la pubblicazione in GUCE prima della scadenza della Commissione UE in carica *pro tempore*.

Anche la *governance dell'AI* — articolata su tre macrolivelli sistemi di AI, sistemi di AI ad alto rischio, modelli di AI per finalità generali e rischio sistemico — pare inefficiente per chi intenda sviluppare AI: la Commissione UE ha ritenuto preferibile, scelta non condivisibile, di non istituire un'unica *Autorità indipendente per l'AI europea* a favore di molteplici interlocutori fra cui un ufficio parte della burocrazia della Commissione UE oltre ovviamente alla Commissione stessa.

Si aggiunge poi un ulteriore livello di governance per i sistemi di AI ad alto rischio post immissione sul mercato e la conformità dei prodotti mutuata dal Regolamento UE 2019/1020 sulla vigilanza del mercato e la conformità dei prodotti UE.

Le differenze rispetto alla stagione straordinaria della tutela dei dati personali sono d'altra parte evidenti. Con il GDPR si intendeva, esportare a livello globale, la tutela di un diritto fondamentale; con l'AI Act si tenta di esportare un farraginoso modello normativo, dall'orditura assai complessa, e ritagliata su standardizzazione, certificazioni tecniche e controlli di prodotto di stampo squisitamente europeo e non certo di respiro universale.

Non a caso emerge un vero e proprio *paradosso normativo* dell'AI Act sul punto della promozione di valori e diritti fondamentali in contrasto con l'art. 3.5 e art. 21 TUE: troppe eccezioni applicative e limitazioni territoriali *sub* Art. 2 AI Act.

Pare criticabile la scelta dell'*AI Act* — verosimilmente condizionata dalle big tech — per effetto della quale sistemi di intelligenza artificiale sviluppati nell'UE, ma esportati oltre i confini europei, non rilevino ai fini dell'applicazione del regolamento, come se si negasse la potenziale violazione dei diritti fondamentali che certo non viene meno nei mercati extra UE⁹.

Ulteriori considerazioni evidenziano le differenti “ambizioni” del legislatore europeo del GDPR e dell'*AI Act*.

In primo luogo, il Libro bianco della Commissione sull'*AI* del 2020, che annunciava l'intenzione di legiferare sull'*AI*, ha dato il via alla corsa alla disciplina normativa del nuovo fenomeno *disruptive* in molte altre giurisdizioni in tutto il mondo, una vera e propria corsa alla regolamentazione dell'*AI*.

In secondo luogo, l'approccio basato sul rischio dell'*AI Act* e alcuni dei suoi concetti chiave troveranno formulazioni personalizzate in altre leggi di altri Stati.

Questo si vede già chiaramente con il *Colorado AI Act* — CAIA — che è stato adottato il 17 maggio 2024, un paio di giorni prima dell'adozione finale dell'*AI Act* europeo.

L'*AI Act*, è ormai chiaro, non conoscerà la diffusione del modello e ampiezza di altri referenti normativi europei: verosimilmente non diverrà *legal benchmark* globale al pari del GDPR, anzi è sgradito nei rapporti transatlantici con gli Stati Uniti considerato impropriamente alla stregua di una barriera commerciale da rimuovere.

Certo non si può sottacere che l'*AI Act* non è centrato, si ribadisce, sulla tutela di un nuovo diritto fondamentale facilmente esportabile a livello globale come la protezione dei dati personali ma su regole europee per conformare un sistema di *AI* in termini di sicurezza, affidabilità e prodotto industriale.

Il contenuto dell'*AI Act* in termini di sicurezza dei prodotti, compresi gli obblighi di documentazione relativi alle valutazioni di conformità a standard tecnici pensati per il mercato europeo difficilmente si radicheranno altrove come dimostrano le differenti politiche a sostegno dell'*AI* di Stati Uniti e Cina.

L'*AI Act* è basato su un approccio “*risk-based*” che ritroviamo anche in altre normative (prima tra tutte, il GDPR): maggiore è il rischio insito nell'utilizzo di un determinato sistema intelligenza artificiale, maggiori saranno, conseguentemente, le responsabilità di chi sviluppa e usa quel sistema, sino a giungere a un divieto di

⁹ Si veda l'Art. 2.3, ultimo periodo dell'*AI Act*: “Il presente regolamento non si applica ai sistemi di IA che non sono immessi sul mercato o messi in servizio nell'Unione, qualora l'output sia utilizzato nell'Unione esclusivamente per scopi militari,

di difesa o di sicurezza nazionale, indipendentemente dal tipo di entità che svolge tali attività”. Si vedano anche sulla recessività dell'*AI Act* rispetto alla tutela incondizionata dei diritti fondamentali anche il considerando 22 e l'art. 2.4 dell'*AI Act*.

utilizzo delle applicazioni e delle tecnologie il cui rischio è considerato inaccettabile.

Si tratta certamente di un punto di partenza importante per la disciplina del nuovo fenomeno nonostante significative criticità, fra cui la complessità e la carenza di chiarezza di diversi profili definitivi e applicativi che lo rendono più simile a una Direttiva mascherata per la generalità nella formulazione di diversi precetti che richiederanno ampio apparato attuativo per l'effettiva applicazione: ben lungi, pertanto, dall'essere un approdo regolatorio definitivo.

La complessità dell'AI Act risulta ancora più evidente se raffrontata — nella prospettiva comparativa transatlantica — alla sostanziale *deregulation federale* da parte degli Stati Uniti a cui si contrappongono, tuttavia, corre l'obbligo di segnalare, importanti iniziative statali seppur in ordine sparso prive, come si è detto, di un *fil rouge* federale.

Un AI Act più snello e focalizzato sui principi fondamentali e le AI vietate sarebbe stato forse più efficiente dal punto di vista applicativo come punto di partenza iniziale rinviando la normativa di dettaglio a una fase successiva più matura della tecnologia da disciplinare: ma questo è il testo con cui allo stato dobbiamo confrontarci anche se l'auspicio di un intervento di semplificazione e razionalizzazione, a breve termine, è senz'altro condivisibile.

L'articolato, suddiviso in tredici capi e corredato da un apparato di tredici allegati, ruota attorno a sette blocchi principali:

- (i) parte generale e definizioni (Capo I)
- (ii) usi vietati dell'AI e gestione del rischio, inclusa quella successiva alla collocazione dei sistemi sul mercato (Capi II, III, VIII, IX);
- (iii) norme specifiche per AI ad alto rischio e per finalità generali, principalmente incentrate sulla trasparenza (Capi IV e V);
- (iv) misure in favore dell'innovazione, ad esempio sandboxes (Capo VI);
- (v) modello di governance, inclusivo delle competenti autorità di supervisione (Capo VII);
- (vi) codici di condotta (Capo X);
- (vii) delegazione legislativa (Capo XI), sanzioni (Capo XII) e previsioni finali (Capo XIII).

Non si può comprendere la rilevanza socioeconomica prima ancora che giuridica dell'AI Act ed il tenore delle previsioni ivi contenute senza tenere a mente il dominio statunitense sui mercati AI, nel rendere disponibili sul mercato business e consumer tecnologie ancora imprecise, o l'uso sistemico da parte di stati totalitari di strumenti di controllo biometrico e sociale.

Per disciplinare il fenomeno non è, tuttavia, sufficiente ricorrere all'elencazione delle categorie di usi vietati dell'AI contenute

nell'AI Act, peraltro caratterizzato da troppe eccezioni e deroghe applicative ad avviso dello scrivente, richiamare le norme sulla trasparenza dei modelli della *General Purpose AI* (GPAI), affrontare il tema della valutazione d'impatto sui diritti fondamentali: occorre contestualizzare la fenomenologia delle intelligenze artificiali sottraendosi alla tentazione, facile quanto inutile, di ragionare solo in termini di astratte categorie e sterili tassonomie peraltro di scarsa utilità in un contesto così in evoluzione come quello di cui trattasi.

La stessa definizione di AI a ben vedere, come si è già rilevato *supra*, non è né semplice né univoca: occorre dunque ancorarsi a principi generali ed evitare la cristallizzazione definitiva angusta che rischia di ingessare l'applicazione dell'AI Act.

Si registra uno iato significativo tra sviluppo delle AI e la disciplina applicabile elaborata dal legislatore europeo.

È preminente, dunque, la necessità di definizioni rigorose e oggettive, precetti normativi chiari e vincolanti corroborati da sanzioni adeguate e dissuasive.

L'AI Act non sempre soddisfa tali aspettative essendo eccessivamente prolifico di definizioni e precetti non sempre di agevole applicazione: si tratta, come si avrà modo di rilevare ulteriormente nel corso del presente approfondimento, di un atto legislativo estremamente complesso.

Il Regolamento in parola pare orientato, si rileva da subito, in discontinuità rispetto al GDPR, principalmente alla sicurezza industriale e alla logica di conformità di prodotto piuttosto che alla tutela dei diritti fondamentali della persona, con una significativa attenzione alla gestione del rischio in termini di valutazione di conformità e con l'attribuzione di un peso significativo agli standard industriali.

Ne emerge, a prima lettura, un testo normativo che pare sfuocato e meno incisivo rispetto al GDPR: l'AI Act cerca una difficile equilibrio tra diritti fondamentali e innovazione tecnologica ma non sempre con la chiarezza d'intenti del GDPR. Anzi pare troppo timido sui principi generali ed eccessivamente burocratico sui singoli precetti differenziati in base al rischio. Inoltre l'intervento regolatorio, come già evidenziato, è caratterizzato da troppe deroghe ed eccezioni: valga per tutti la formulazione dell'ambito applicativo fortemente limitato (art. 2 AI Act) e l'enunciazione delle pratiche vietate (art. 5 AI Act).

È evidente che l'obiettivo del legislatore europeo, nel caso dell'AI Act, non è tanto quello di assicurare — in ogni caso e a tutti i costi — la tutela dei diritti fondamentali, della persona piuttosto quello di bilanciare tutela della persona e del mercato favorendo soprattutto lo sviluppo dell'AI per facilitare un recupero dell'Europa in tale settore specifico, forse più utopico che reale.

Tale tensione intrinseca tra persona e mercato si rivela chiaramente nell'AI Act e spiega la ragione dell'approccio regolatorio focalizzato sulla sicurezza del prodotto e l'equilibrio scelto nella gestione del rischio secondo una prospettiva strategica europea più incline all'accettazione del rischio, considerata la previsione di numerose eccezioni e deroghe applicative, rispetto ad altri interventi legislativi più rigorosi come il GDPR.

Si pensi, a titolo esemplificativo, alle ampie eccezioni dell'art. 111 per le AI già operative prima dell'entrata in vigore dell'AI Act.

Alle ulteriori, eccessivamente ampie, eccezioni relative all'applicazione delle AI ad alto rischio nei casi di *sicurezza nazionale, difesa, ricerca e uso personale*.

4. L'ETEROGENEITÀ DELLE AI E IL PROBLEMA DEFINITORIO.

Un ulteriore profilo da non sottovalutare e che sarà foriero di non pochi problemi interpretativi e applicativi riguarda proprio la definizione univoca di AI.

Le definizioni sono rilevanti dal punto di vista giuridico e soprattutto per essere efficaci richiedono di essere ancorate a parametri condivisi e oggettivi.

Di tali caratteristiche non vi è traccia nell'AI Act.

La definizione di AI risulta problematica in quanto piuttosto generica e descrittiva: si riflette il malessere definitorio della comunità scientifica tecnologica che non ha raggiunto una definizione condivisa e univoca.

Pensiamo alla definizione di AI per nulla semplice e non agganciata ad elementi oggettivi quali parametri di addestramento utilizzati e velocità di calcolo — utilizzati solo per GPAI e rischio sistemico — solo per fare alcuni esempi senza pretesa di esaustività.

Si definisce ai sensi dell'art. 3 dell'AI Act « sistema di IA »: *“un sistema automatizzato progettato per funzionare con livelli di autonomia variabili e che può presentare adattabilità dopo la diffusione e che, per obiettivi espliciti o impliciti, deduce dall'input che riceve come generare output quali previsioni, contenuti, raccomandazioni o decisioni che possono influenzare ambienti fisici o virtuali”*.

E ancora si definisce « modello di IA per finalità generali »: *“un modello di IA, anche laddove tale modello di IA sia addestrato con grandi quantità di dati utilizzando l'autosupervisione su larga scala, che sia caratterizzato da una generalità significativa e sia in grado di svolgere con competenza un'ampia gamma di compiti distinti, indipendentemente dalle modalità con cui il modello è immesso sul mercato, e che può essere integrato in una varietà di sistemi o applicazioni a valle, ad eccezione dei modelli*

di IA utilizzati per attività di ricerca, sviluppo o prototipazione prima di essere immessi sul mercato”.

Persino l'EDPS, a titolo esemplificativo, senza pretese di esaustività, segnala come meritevole di chiarificazione il concetto di livelli di autonomia variabili e precisa che ¹⁰:

“In our view, all IT systems may be considered autonomous, as they do automate calculations and processes. For example: an IT system that includes a machine learning model that shows to a bank clerk a prediction on a loan application is autonomous. But in which sense is that IT system more autonomous than a simple rule based algorithm that will judge some attributes of the application?”.

E ancora sul concetto di adattabilità dopo la diffusione:

“The EDPS considers the notion of adaptiveness should be broadened to include a fundamental element, namely the adaptation and the learning procedure that occur before the deployment, which is a unique element to AI systems. The definition seems to obviate the fact that AI systems learn during their development. AI systems start their development with a set of unadjusted initial parameters. Then, they undergo a learning procedure in order to adjust/tune these parameters to improve their performance in a defined task. AI systems become ‘intelligent’ precisely because of their capability of learning and adapting”.

Sul punto soccorrono, almeno in parte trattandosi di orientamenti non vincolanti che espressamente rinviando alla futura giurisprudenza chiarificatrice della Corte di Giustizia UE, le Linee guida della Commissione UE del 6 febbraio 2025 sulla definizione di un sistema di intelligenza artificiale stabilito dall'AI Act: *Guidelines on the definition of an artificial intelligence system established by Regulation (EU) 2024/1689 (AI Act)*.

Solo determinati sistemi di intelligenza artificiale sono soggetti a obblighi normativi e supervisione ai sensi della legge europea sull'intelligenza artificiale: quali caratteristiche devono presentare per convincere l'interprete all'applicazione dell'AI Act?

L'approccio basato sul rischio dell'AI Act implica che solo i sistemi che danno origine ai rischi più significativi per i diritti e le libertà fondamentali saranno soggetti ai divieti stabiliti nell'articolo 5 dell'AI Act, al regime normativo per i sistemi di intelligenza artificiale ad alto rischio coperti dall'articolo 6 dell'AI Act e ai requisiti di trasparenza per un numero limitato di sistemi di intelligenza artificiale predefiniti stabiliti nell'articolo 50 dell'AI Act.

¹⁰ EDPS comment to the AI Office's consultation on the application of the definition of an AI System and the prohibited

AI practices established in the AI Act - 19 December 2024.

La stragrande maggioranza dei sistemi AI, anche se si qualificano come sistemi di intelligenza artificiale ai sensi dell'articolo 3(1) dell'AI Act, non sarà, invero, soggetta ad alcun obbligo normativo ai sensi della legge europea sull'intelligenza artificiale.

Questo aspetto, come, si è già avuto modo di rilevare, non convince: almeno il precetto generale della sorveglianza umana — seppur differenziata in base al livello di rischio e in attuazione del principio di accountability — dovrebbe essere utilmente applicato in via generale e non limitatamente alle AI ad alto rischio.

D'altra parte tutte le AI sono intrinsecamente rischiose: negare l'evidenza non può che tradursi in deficit protettivo della persona e dei diritti fondamentali della personalità.

L'AI Act si applica, inoltre, anche ai modelli di intelligenza artificiale di uso generale, c.c. *large language model*, che sono regolamentati nel capitolo V della legge sull'intelligenza artificiale.

Centrale per la corretta applicazione della legge sull'intelligenza artificiale è, pertanto, la corretta qualificazione giuridica del sistema di intelligenza artificiale.

Più facile a dirsi che a farsi.

La definizione di un sistema di AI comprende effettivamente un ampio spettro di sistemi non sempre agevolmente qualificabili — in assenza di parametri di legge oggettivi — quale sistema di AI.

La qualificazione di un qualunque software in termini di sistema di AI dovrà basarsi — come suggerito dalla Commissione UE nelle Linee Guida sulla definizione di sistema di AI del 6 febbraio 2025 — sull'architettura e sulla funzionalità specifiche di un dato sistema e dovrà prendere in considerazione i sette elementi principali della definizione stabilita nell'articolo 3(1) dell'AI Act, vale a dire:

- (1) un sistema automatizzato basato su una macchina;
- (2) progettato per funzionare con diversi livelli di autonomia;
- (3) che può mostrare adattabilità dopo l'implementazione;
- (4) e che, per obiettivi espliciti o impliciti;
- (5) deduce, dall'input che riceve, come generare output;
- (6) come previsioni, contenuti, raccomandazioni o decisioni
- (7) che possono influenzare ambienti fisici o virtuali.

Soffermiamoci, a titolo esemplificativo, sul concetto di sistema automatizzato. Tale aspetto qualificatorio si riferisce al fatto che il funzionamento dei sistemi di AI prevede lo sviluppo e l'uso di macchine. Il termine “macchina” può essere inteso in senso ampio come comprendente i componenti sia hardware che software che consentono il funzionamento del sistema di AI.

Tutti i sistemi di AI sono automatizzati, in quanto richiedono macchine per il loro funzionamento, quali l'addestramento dei modelli, l'elaborazione dei dati, la modellizzazione predittiva e il processo decisionale automatizzato su larga scala.

L'elemento di “automatizzazione” nella definizione di sistema di

AI sottolinea il fatto che i sistemi di AI devono essere basati sul calcolo e sul funzionamento delle macchine.

Il termine “automatizzato” comprende, dunque, un’ampia varietà di sistemi computazionali. Ad esempio, oltre ai sistemi informatici *tradizionali* vanno inclusi anche i sistemi di *calcolo quantistico* emergenti, che rappresentano una deviazione significativa dai sistemi informatici tradizionali, così come i sistemi biologici od organici, purché forniscano capacità computazionali.

Non sono, quindi, possibili determinazioni automatiche o elenchi esaustivi di sistemi che rientrano o esulano dalla definizione di un sistema di AI.

Il considerando 97 dell’AI Act non sottovaluta il problema definitorio anzi precisa che: “La nozione di modelli di AI per finalità generali dovrebbe essere chiaramente definita e distinta dalla nozione di sistemi di AI per consentire la certezza del diritto. La definizione dovrebbe basarsi sulle principali caratteristiche funzionali di un modello di AI per finalità generali, in particolare la generalità e la capacità di svolgere con competenza un’ampia gamma di compiti distinti. Questi modelli sono solitamente addestrati su grandi quantità di dati con diversi metodi, come l’apprendimento autosupervisionato, non supervisionato o per rinforzo”.

Secondo il considerando 98 della legge sull’AI, “mentre la generalità di un modello potrebbe, tra l’altro, essere determinata anche da una serie di parametri, si dovrebbe ritenere che i modelli con almeno un miliardo di parametri e formati con una grande quantità di dati utilizzando l’autocontrollo su larga scala mostrino una generalità significativa e svolgano con competenza un’ampia gamma di compiti distintivi”¹¹.

Ancora il considerando 99 della legge sull’AI aggiunge che “i grandi modelli di AI generativa sono un tipico esempio di modello di AI per finalità generali, dato che consentono una generazione flessibile di contenuti, ad esempio sotto forma di testo, audio, immagini o video, che possono facilmente adattarsi a un’ampia gamma di compiti distintivi”.

Invero, solo un poco più agevole, ma solo in apparenza, risulta la definizione di AI per finalità generali.

Soccorre sul punto l’art. 3, paragrafo 63 dell’AI Act definisce un modello di AI per finalità generali “un modello di AI, anche laddove tale modello di AI sia addestrato con grandi quantità di dati utilizzando l’autosupervisione su larga scala, che sia caratterizzato una generalità significativa e sia in grado di svolgere con competenza un’ampia gamma di compiti distinti, indipendentemente dalle modalità con cui il modello è immesso sul mercato, e

¹¹ A titolo esemplificativo si ritiene, dati ufficiali non sono attualmente disponibili, che CHAT GPT 5 ultima release di

agosto 2025 si fondi su un numero impressionante di parametri compresi indicativamente tra 5 e 10 trilioni.

che può essere integrato in una varietà di sistemi o applicazioni a valle, ad eccezione dei modelli di AI utilizzati per attività di ricerca, sviluppo o prototipazione prima di essere immessi sul mercato”.

L’art. 51, paragrafo 2 precisa ulteriormente che “Si presume che un modello di AI per finalità generali abbia capacità di impatto elevato a norma del paragrafo 1, lettera a), quando la quantità cumulativa di calcolo utilizzata per il suo addestramento misurata in operazioni in virgola mobile è superiore a 10^{25} ”.

Inoltre il 10 luglio 2025 è stato pubblicato il *Codice di buone pratiche dell’IA per finalità generali* articolato in tre capitoli: trasparenza, diritto d’autore e sicurezza.

I capi sulla trasparenza e sul diritto d’autore offrono a tutti i fornitori di modelli di IA per finalità generali indicazioni chiare per dimostrare il rispetto dei loro obblighi ai sensi dell’articolo 53 della legge sull’AI.

I capi sulla sicurezza riguardano invece il sottoinsieme ristretto di fornitori dei modelli più avanzati, soggetti agli obblighi della legge sull’AI per i fornitori di modelli di AI per finalità generali con rischio sistemico a norma dell’articolo 55 della legge sull’AI.

In data primo agosto 2025 Commissione UE e Comitato per l’AI hanno rilasciato il prescritto parere favorevole ritenendo trattarsi di strumento volontario adeguato affinché i fornitori di modelli GPAI dimostrino la conformità alla legge sull’AI.

Il Codice di buone pratiche predetto è stato successivamente integrato il 18 luglio 2025, sempre in vista della piena applicazione dal 2 agosto 2025 dei precetti relativi alle AI per finalità generali, dalle *Guidelines on the scope of the obligations for general-purpose AI models established by Regulation (EU) 2024/1689 (AI Act)*.

Nelle citate line Guida della Commissione riemerge il “malessere” definitorio che si tenta di delimitare ricorrendo a parametri computazionali oggettivi: “*Given the wide variety of capabilities and use cases for general-purpose AI models, it is not feasible to provide a precise list of capabilities that a model must display and tasks that it must be able to perform in order to determine whether it is a general-purpose AI model. Instead, the Commission’s approach is to use the amount of computational resources used to train the model (‘training compute’, see paragraph 115 for how this notion should be understood in these guidelines) measured in FLOP as well as the modalities of the model to set an indicative criterion for when it considers a model to be a general-purpose AI model. This approach builds on recital 98 AI Act (omissis). Training compute has the advantage of combining number of parameters and number of training examples into a single number that is reasonably straightforward for providers to estimate. This number is typically proportional to the number obtained by multiplying these two numbers, allowing a single*

threshold to be set rather than separate thresholds for model size and training data size. While training compute is an imperfect proxy for generality and capabilities, the Commission considers setting an indicative criterion which includes a training compute threshold to be the most suitable approach at present. Nevertheless, the Commission's approach may change in the future as technology and the market evolve. Furthermore, the Commission will continue to investigate the availability of other criteria that could be used to assess generality and capabilities with relative ease, especially for smaller actors. Based on the considerations above, an indicative criterion for a model to be considered a general-purpose AI model is that its training compute is greater than 1023 FLOP and it can generate language (whether in the form of text or audio), text-to-image or text-to-video”.

Nonostante tutto la vertigine definitoria persiste e il florilegio di parametri qualitativi e quantitativi non agevola più di tanto non essendo di immediata e univoca lettura nononstate le migliori in intenzioni dichiarate negli atti normativi e attuativi pertinenti.

Ergo il problema definitorio pare permanere e costituirà verosimilmente uno dei principali punti di attrito relativamente all'applicazione — o disapplicazione — del poderoso complesso normativo dell'AI Act.

L'onere della qualificazione dell'AI è certamente una parte fondamentale dell'applicazione dell'AI Act, ma non è affatto semplice per l'interprete l'opera qualificatoria nel complesso contesto predetto e non potrà che essere operata caso per caso in concreto.

Anche sotto tale relevantissimo profilo applicativo sarà, pertanto, decisiva l'attività qualificatoria operata dalla Corte di Giustizia UE.

5. LA CLASSIFICAZIONE DEL RISCHIO DELLE AI: INACCETTABILE, ALTO, LIMITATO E MINIMO.

L'AI Act contiene disposizioni progressivamente più stringenti articolate secondo quattro livelli di *gestione del rischio*:

- inaccettabile
- alto
- limitato
- minimo

Ai sensi dell'AI Act il rischio è la combinazione della probabilità del verificarsi di un danno e la gravità del danno stesso (art. 3, paragrafo 1, n. 2 AI Act).

I sistemi di AI sono sostanzialmente vietati perché considerati troppo pericolosi per i diritti fondamentali della persona e gli altri valori protetti dalla carta dei diritti fondamentali UE.

I sistemi di AI ad alto rischio devono rispettare una serie di stingenti requisiti e obblighi per poter accedere al mercato dell'Unione europea. In particolare, l'AI Act richiede una maggiore

trasparenza per quanto riguarda lo sviluppo e l'utilizzo di tali sistemi.

Invece i sistemi di AI a rischio limitato “*destinati direttamente a interagire con persone fisiche*” sono soggetti a obblighi di trasparenza previsti dall'articolo 50 della legge sull'AI.

Infine, i sistemi di AI a rischio minimo o nullo esulano dall'ambito applicativo dell'AI Act.

Prima di utilizzare un sistema di AI ad alto rischio, i *deployer* che possono enti pubblici che forniscono servizi pubblici, oppure soggetti privati quali banche e compagnie di assicurazione, devono effettuare una valutazione d'impatto sui diritti fondamentali¹²).

Per alcuni altri usi dell'AI — come la manipolazione cognitiva del comportamento, i sistemi di valutazione sociale, il riconoscimento delle emozioni sul posto di lavoro e nelle istituzioni educative e il *social scoring* — i rischi sono invece ritenuti inaccettabili e quindi l'utilizzo di tale tipologia di sistemi è vietato.

L'AI Act vieta espressamente alcune applicazioni di AI che minacciando i diritti fondamentali delle persone comportano un rischio inaccettabile.

L'Art. 5 dell'AI Act individua espressamente i seguenti sistemi di AI vietati:

— i sistemi di categorizzazione biometrica basati su caratteristiche sensibili e l'estrapolazione indiscriminata di immagini facciali da internet o dalle registrazioni dei sistemi di telecamere a circuito chiuso per creare banche dati di riconoscimento facciale;

— i sistemi di riconoscimento delle emozioni sul luogo di lavoro e nelle scuole,

— i sistemi di credito sociale, le pratiche di polizia predittiva basate esclusivamente sulla profilazione o sulla valutazione delle caratteristiche antropologiche di una persona;

— i sistemi che manipolano il comportamento umano o sfruttano le vulnerabilità delle persone.

Si segnalano sul tema relevantissimo delle pratiche di AI vietate le *Guidelines on prohibited artificial intelligence practices established by Regulation (EU) 2024/1689 (AI Act) approvate dalla Commissione UE in data 29 luglio 2025*.

Tuttavia, pur apprezzandosi lo sforzo chiarificatore della *Commissione UE*, 134 pagine esplicative dell'art. 5 dell'AI Act paiono davvero eccessive, evidentemente causate dalla formulazione generica dell'AI Act, già stigmatizzata *supra*, più simile alla direttiva che al regolamento.

¹² Si vedano sul tema G. MALGIERI, C. SANTOS, *Assessing the (Severity of) Impacts on Fundamental Rights*, June 25, 2024), <https://ssrn.com/abstract=4875937>; A. MANTELERO, *The Fundamental Rights Impact*

Assessment (FRIA) in the AI Act: Roots, legal obligations and key elements for a model template, in *Computer Law & Security Review*, 54, 2024, pp. 106 ss.

Il rischio di iperregolazione e di incertezza applicativa si concretizza anche nell'eccesso di Linee Guida che rischiano di disorientare l'operatore prima ancora dell'interprete.

Un secondo livello di rischio contempla le AI ad altro rischio in relazione alle quali sono previsti obblighi stringenti in quanto potrebbero arrecare danni significativi alla salute, alla sicurezza, ai diritti fondamentali, all'ambiente, alla democrazia e allo Stato di diritto.

Rientrano in questa categoria gli usi di AI legati a:

- infrastrutture critiche,
- istruzione e formazione professionale,
- occupazione, servizi pubblici
- servizi privati di base: assistenza sanitaria, bancaria e assicurativa
- alcuni sistemi di contrasto, migrazione e gestione delle frontiere,
- giustizia e processi democratici.

Per questi sistemi vige l'obbligo di valutare e ridurre i rischi, mantenere registri d'uso, essere trasparenti e accurati e garantire sempre la sorveglianza umana: il c.d. *human in the loop*.

I cittadini avranno diritto a presentare reclami sui sistemi di IA e a ricevere spiegazioni sulle decisioni basate su sistemi di IA ad alto rischio che incidono sui loro diritti.

L'AI Act si prefigge lo scopo di migliorare e promuove la protezione dei diritti tutelati dalla Carta: il diritto alla dignità umana (articolo 1), al rispetto della vita privata e alla protezione dei dati di carattere personale (articoli 7 e 8), alla non discriminazione (articolo 21)

Sono introdotte a tal fine alcune necessarie restrizioni alla libertà d'impresa (articolo 16) e alla libertà delle arti e delle scienze (articolo 13) al fine di assicurare il rispetto di motivi imperativi d'interesse pubblico quali la salute, la sicurezza, la tutela dei consumatori e la protezione di altri diritti fondamentali ("innovazione responsabile") nel momento in cui si diffonde e si utilizza una tecnologia di IA.

Tali restrizioni sono proporzionate e limitate al minimo necessario per prevenire e attenuare rischi gravi per la sicurezza e probabili violazioni dei diritti fondamentali.

In relazione alle AI a rischio limitato sono invece previste misure meno stringenti in particolare obblighi di trasparenza che si applicano ai sistemi che: i) interagiscono con gli esseri umani; ii) sono utilizzati per rilevare emozioni o stabilire un'associazione con categorie (sociali) sulla base di dati biometrici; oppure iii) generano o manipolano contenuti ("*deep fake*").

Inoltre, i maggiori obblighi di trasparenza non incideranno in maniera sproporzionata sul diritto alla protezione della proprietà intellettuale (articolo 17, paragrafo 2), dato che saranno limitati soltanto alle informazioni minime necessarie affinché le persone

possano esercitare il loro diritto a un ricorso effettivo e alla necessaria trasparenza presso le autorità di controllo e di contrasto, in linea con i loro mandati.

I set di dati di addestramento, convalida e prova delle AI ad alto rischio devono essere pertinenti, rappresentativi, esenti da errori e completi.

Essi possiedono le proprietà statistiche appropriate, anche, ove applicabile, per quanto riguarda le persone o i gruppi di persone sui quali il sistema di IA ad alto rischio è destinato a essere usato. Queste caratteristiche dei set di dati possono essere soddisfatte a livello di singoli set di dati o di una combinazione degli stessi.

La protezione dei dati personali è protezione della libertà e della dignità della persona, tanto più quando sono coinvolti i minori e i soggetti vulnerabili.

L'applicazione dell'intelligenza artificiale in campo neuroscientifico e, soprattutto i sistemi di *brain reading*, idonei almeno potenzialmente a decodificare il pensiero, così come le controverse pratiche di *neuromarketing* devono, infatti, sempre garantire, come primo dei "neurodiritti", la *privacy mentale*, condizione ineludibile di autodeterminazione, presupposto intangibile di libertà¹³.

Varcata la soglia della lettura del pensiero, la deriva da impedire è rendere la persona un archivio liberamente accessibile dalle AI e altri dispositivi tecnologici, anche se con il consenso dell'interessato, frutto di un'autodeterminazione imperfetta e come tale inidonea a costituire una base giuridica per il trattamento dei propri pensieri, le cui idee siano messe a nudo senza più alcuno spazio per la libertà e l'autonoma volizione.

Siamo davvero disposti ad ammettere, in un prossimo futuro di cui già si discorre, le più evolute tecniche di pubblicità mentale, su impianti cerebrali, solamente affidando la protezione della persona al consenso dell'interessato?

È evidente che la risposta non può che essere negativa essendo il consenso individuale e l'autodeterminazione informativa troppo deboli in determinati contesti tecnologici digitali altamente manipolativi, ora supportati dalle AI, ulteriore innovativo fattore di dirompente asimmetria. Trattasi dell'effetto *nudge tech* che concorre a disegnare con predeterminazione favorevole al contraente forte l'architettura delle scelte del contraente debole più in generale dell'utente digitale spingendo il soggetto interessato a scelte spesso inconsapevoli certamente gradite alla piattaforma digitale-

¹³ Sul tema del neuromarketing si veda per tutti S. Orlando (a cura di), *Profili giuridici del neuromarketing*, Roma University Press, 2025, *passim*.

contraente forte ma non necessariamente protettive della dignità personale ¹⁴.

Le nuove vulnerabilità digitali richiedono, pertanto, barriere precettive conformative inderogabili e rafforzate a prescindere dall'autodeterminazione personale che non può più da sola costituire un solido argine al fenomeno di sfruttamento commerciale dei dati personali da parte delle big tech e più in generale costituire strumento di protezione efficace a presidio dei diritti fondamentali della persona, in ultima istanza della dignità personale ¹⁵.

Infine, criticità non trascurabile, risultano sostanzialmente ignorate dall'AI Act le AI a rischio minimo: in relazione a tale categoria residuale sarebbe stato più opportuno prevedere almeno l'appli-

¹⁴ Sul tema del consenso al trattamento dei dati personali, circolazione dei dati personali e ammissibilità di un consenso contrattuale assorbente dei profili inerenti al trattamento dei dati personali si vedano: V. RICCIUTO, *La patrimonializzazione dei dati personali. Contratto e mercato nella ricostruzione del fenomeno*, in *Dir. Informaz. e informatica*, 2018, p. 689 ss.; P. GALLO, *Il consenso al trattamento dei dati personali come prestazione*, in *Riv. Dir. Civ.*, 2022, pp. 1054 ss.; T.M. UBERTAZZI, *Ripensando alla revoca del consenso nella prospettiva funzionale della privacy*, in *Contratto e impresa*, 2022, pp. 27 ss.; S. ORLANDO, *Per un sindacato di liceità del consenso*, in *Persona e mercato*, 2022, pp. 527 ss.; S. ORLANDO, *Consenso al trattamento e liceità*, in *Persona e mercato*, 2024, pp. 333 ss.; *Libertà e liceità del consenso nel trattamento dei dati personali*, a cura di S. Orlando, *Persona e Mercato* ed., Firenze, 2024; cui adde E. TOSI (A CURA DI), *Privacy Digitale*, Milano, 2019; E. TOSI, *Diritto Privato delle Nuove Tecnologie Digitali. Riservatezza. Contratto, riservatezza, responsabilità tra persona e mercato*, Milano, 2021; E. TOSI, *Dati personali e contratto: un ossimoro apparente*, *European Journal of Privacy Law and Technologies*, 2023, pp. 71 ss. e da ultimo E. TOSI, *Circolazione contrattuale dei dati personali tra contratto e responsabilità*, Milano, 2023, pp. 78 ss. che registra l'emersione della fragilità del consenso individuale nel contesto digitale asimmetrico quale strumento protettivo della persona. *Contra* in ordine alla prevalente natura autorizzatoria, con varie sfumature interpretative, del consenso al trattamento dei dati personali: G. ALPA-G. RESTA, *Le persone fisiche e i diritti della personalità, I*, in *Trattato di Diritto Civile*, diretto da R. SACCO, Torino 2006, *passim*, spec. p. 629 ss.; F. BRAVO, *Lo "scambio di dati personali" nei contratti di fornitura di*

servizi digitali e il consenso dell'interessato tra autorizzazione e contratto, in questa *Rivista*, 2019, p. 34 ss.; Id, *Il "diritto" a trattare dati personali nello svolgimento dell'attività economica*, Padova, 2018, *passim*; P. MANES, *Il consenso al trattamento dei dati personali*, Padova, 2001; e C. IRTI, *Consenso "negoziato" e circolazione dei dati personali*, Torino, 2021, *passim*.

¹⁵ Sulle vulnerabilità digitali si vedano in tal senso: G. MALCIERI, *Vulnerability and Data Protection Law*, Oxford University Press, 2023; C. ROSCA, *Digital arms for digital consumer harms. Mapping Legal and Technical Solutions for Dark Patterns in EU Consumer Law*, Maastricht University Press, 2024; G. VETTORI, *Persona e pluralismo*, in *Persona e mercato*, 2024, 2, p. 331; S. ORLANDO, *Consenso al trattamento e liceità*, in *Persona e mercato*, 2024, 2, p. 347 ss.; S. ORLANDO (a cura di), *Profili giuridici del neuromarketing. Annuario 2023-2024 OGID Osservatorio Giuridico sull'Innovazione Digitale*, Sapienza Università Editrice, 2025; A. GORGONI, *La vulnerabilità nell'ambiente digitale e la protezione della libertà del volere*, in *Profili giuridici del neuromarketing. Annuario 2023-2024 OGID Osservatorio Giuridico sull'Innovazione Digitale*, a cura di S. Orlando, Sapienza Università Editrice, 2025, 142 ss.; N. CRISTIANINI, *La scorciatoia. Come le macchine sono diventate intelligenti senza pensare in modo umano*, Bologna, 2023, p. 99 ss.; F. GALLI, *La pubblicità mirata al tempo dell'intelligenza artificiale: quali regole a tutela dei consumatori*, in *Contr. e impr.*, 2022, 3, p. 919 ss.; D. IMBRUGLIA, *La presunzione delle macchine e il consenso dell'interessato*, in *Riv. trim. dir. proc. civ.*, 2023, p. 921 ss.; A.C. NAZZARO, *Aspetti giuridici del neuromarketing: privacy, profilazione e genuinità del consenso del consumatore*, in M. Caratù e P. Cherubino, *Neuromarketing, società e tutela del consumatore*, Roma, 2023, p. 109; G. SMORTO, *Distribuzione del*

cazione dei principi generali di trasparenza e supervisione umana, seppur attenuati trattandosi di AI a rischio minimo.

6. DATI E GOVERNANCE DEI DATI DELLE AI AD ALTO RISCHIO: L'ART. 10 DELL'AI ACT.

Merita anche un cenno la disciplina riservata dall'AI Act ai sistemi di AI ad alto rischio che utilizzano tecniche che prevedono l'uso di dati per l'addestramento di modelli di AI: devono essere sviluppati sulla base di set di dati di addestramento, convalida e prova che soddisfano i criteri di qualità di cui ai paragrafi da 2 a 5 dell'art. 10 dell'AI Act ogniquale volta siano utilizzati tali set di dati.

rischio e tutela dei diritti nel regolamento europeo sull'intelligenza artificiale. Una riflessione critica, in *Foro it.*, 5, 2024, c. 208, che richiama l'attenzione sul « pericolo che, in molti casi, siano i soggetti più vulnerabili ad essere maggiormente esposti alle conseguenze pregiudizievoli derivanti dall'impiego di strumenti il cui impatto complessivo è considerato positivamente se visto in modo aggregato. Ciò può verificarsi, ad esempio, a causa dei parametri utilizzati per compiere questa valutazione del rischio, i quali prendono a riferimento la figura tipica del c.d. uomo medio, a scapito di una visione più articolata e complessa della società, che tenga conto delle differenze e delle diversità tra individui e tra gruppi di soggetti »; R. CATERINA, *Psicologia della decisione e tutela del consumatore*, in *Analisi Giuridica dell'Economia*, 1, 2012, pp. 67-84; Id., *La libertà di comunicazione: il fenomeno dei social network*, in C. Perlingieri e L. Ruggeri, *Internet e diritto civile*, Napoli, 2015, p. 117 ss. Sottolinea la centralità del ruolo dell'interprete e il momento interpretativo che non può ridursi a mera struttura algoritmica: P. PERLINGIERI, *Struttura algoritmica e interpretazione*, in *Tecnologie e diritto*, 2, 2020, p. 434.

Sulla compatibilità dello sfruttamento commerciale dei dati personali con il GDPR e sulla conseguente complementarietà tra GDPR e Codice del Consumo nella circolazione dei dati personali si vedano da ultimo *Consiglio di Stato* 2 dicembre 2024, n. 9614 e *Consiglio di Stato* 7 gennaio 2025, n. 80. Sul tema della vulnerabilità digitale si veda Cons. Stato, 10 ottobre 2022, n. 8650, in *Foro it.*, V, 2022, c. 542 ss., con nota di C. Bona e N. Bonini, *Tutela del consumatore e nuovi paradigmi scientifici (scienze cognitive e neuroscienze varcano la soglia di Palazzone Spada)*. Con l'ordinanza 10 ottobre 2022 n. 8650, il Consiglio di Stato (Sezione VI), ha rimesso alla Corte di giustizia dell'Unione europea le seguenti questioni

pregiudiziali: “a) Se la nozione di consumatore medio di cui alla Direttiva 2005/29/CE inteso come consumatore normalmente informato e ragionevolmente attento ed avveduto — per la sua elasticità ed indeterminata — non debba essere formulata con riferimento alla miglior scienza ed esperienza e di conseguenza rimandi non solo alla nozione classica dell'homo oeconomicus ma anche alle acquisizioni delle più recenti teorie sulla razionalità limitata che hanno dimostrato come le persone agiscono spesso riducendo le informazioni necessarie con decisioni “irragionevoli” se parametrata a quelle che sarebbero prese da un soggetto ipoteticamente attento ed avveduto acquisizioni che impongono una esigenza protettiva maggiore dei consumatori nel caso — sempre più ricorrente nelle moderne dinamiche di mercato — di pericolo di condizionamenti cognitivi. b) Se possa essere considerata di per sé aggressiva una pratica commerciale nella quale, a causa dell'incorniciamento delle informazioni (framing) una scelta possa apparire come obbligata e senza alternative tenendo conto dell'articolo 6, paragrafo 1, della Direttiva che considera ingannevole una pratica commerciale che in qualsiasi modo inganni o possa ingannare il consumatore medio « anche nella sua presentazione complessiva ». La questione traeva origine da un ricorso di Compass s.p.a. contro una sanzione dell'Agem che è intervenuta per un caso di una pratica commerciale scorretta aggressiva ex art. 24 Codice Consumo, basata sul fatto che Compass s.p.a. offrisse alla propria clientela, oltre ai prodotti finanziari, anche polizze assicurative di diverso genere e tipo. Il Consiglio di Stato in ragione di quanto sopra ha ritenuto di inquadrare la fattispecie nell'ambito delle c.d. vendite abbinate, in cui la scelta di consumo operata dal soggetto appare obbligata e senza alternative, come previsto dall'art. 6, para. 1, dir. 2005/29/UE, a causa dell'incorniciamento (c.d. framing) delle informazioni offerte al consumatore.

I set di dati di addestramento, convalida e prova sono soggetti a pratiche di governance e gestione dei dati *adeguate alla finalità prevista del sistema di AI ad alto rischio*. Tali pratiche riguardano in particolare:

- a) le scelte progettuali pertinenti;
- b) *i processi di raccolta dei dati e l'origine dei dati, nonché la finalità originaria della raccolta nel caso di dati personali;*
- c) le operazioni di trattamento pertinenti ai fini della preparazione dei dati, quali annotazione, etichettatura, pulizia, *aggiornamento*, arricchimento e aggregazione;
- d) la formulazione di ipotesi, in particolare per quanto riguarda le informazioni che si presume che i dati misurino e rappresentino;
- e) una valutazione della disponibilità, della quantità e dell'adeguatezza dei set di dati necessari;
- f) *un esame atto a valutare le possibili distorsioni suscettibili di incidere sulla salute e sulla sicurezza delle persone, di avere un impatto negativo sui diritti fondamentali o di comportare discriminazioni vietate dal diritto dell'Unione, specie laddove gli output di dati influenzano gli input per operazioni future;*
- g) *le misure adeguate per individuare, prevenire e attenuare le possibili distorsioni individuate conformemente alla lettera f);*
- h) l'individuazione di lacune o carenze pertinenti nei dati tali da pregiudicare il rispetto del presente regolamento e il modo in cui tali lacune e carenze possono essere colmate.

I set di dati di addestramento, convalida e prova sono pertinenti, *sufficientemente rappresentativi e, nella misura del possibile*, esenti da errori e completi *nell'ottica della finalità prevista*.

Essi possiedono le proprietà statistiche appropriate anche, ove applicabile, per quanto riguarda le persone o i gruppi di persone *relativamente ai quali il sistema di AI ad alto rischio è destinato a essere usato*. Queste caratteristiche dei set di dati possono essere soddisfatte a livello di singoli set di dati o a livello di una combinazione degli stessi.

I set di dati tengono conto, nella misura necessaria per la finalità prevista, delle caratteristiche o degli elementi particolari dello specifico ambito geografico, *contestuale*, comportamentale o funzionale all'interno del quale il sistema di AI ad alto rischio è destinato a essere usato.

Nella misura in cui ciò sia strettamente necessario al fine di garantire il rilevamento e la correzione delle distorsioni in relazione ai sistemi di AI ad alto rischio *in conformità del paragrafo 2, lettere f) e g), dell'art. 10 dell'AI Act*, i fornitori di tali sistemi possono *eccezionalmente* trattare categorie particolari di dati personali, fatte salve le tutele adeguate per i diritti e le libertà fondamentali delle persone fisiche.

Oltre alle disposizioni di cui ai regolamenti (UE) 2016/679 e (UE)

2018/1725 e alla direttiva (UE) 2016/680 devono essere soddisfatte, affinché tale trattamento avvenga, tutte le condizioni seguenti:

a) *il rilevamento e la correzione delle distorsioni non possono essere realizzati efficacemente mediante il trattamento di altri dati, compresi i dati sintetici o anonimizzati;*

b) *le categorie particolari di dati personali sono soggette a limitazioni tecniche relative al riutilizzo dei dati personali, nonché a misure più avanzate di sicurezza e di tutela della vita privata, compresa la pseudonimizzazione;*

c) *le categorie particolari di dati personali sono soggette a misure tese a garantire che i dati personali trattati siano resi sicuri e protetti nonché soggetti a garanzie adeguate, ivi compresi controlli e documentazione rigorosi dell'accesso, al fine di evitare abusi e garantire che solo le persone autorizzate e sottostanti a opportuni obblighi di riservatezza abbiano accesso a tali dati personali;*

d) *le categorie particolari di dati personali non devono essere trasmesse, trasferite o altrimenti consultate da terzi;*

e) *le categorie particolari di dati personali vengono cancellate dopo che la distorsione è stata corretta oppure i dati personali hanno raggiunto la fine del loro periodo di conservazione, a seconda di quale delle due condizioni si verifica per prima;*

f) *i registri delle attività di trattamento a norma dei regolamenti (UE) 2016/679 e (UE) 2018/1725 e della direttiva (UE) 2016/680 comprendono i motivi per cui il trattamento delle categorie particolari di dati personali era strettamente necessario per rilevare e correggere distorsioni e i motivi per cui tale obiettivo non poteva essere raggiunto mediante il trattamento di altri dati.*

Per lo sviluppo di sistemi di AI ad alto rischio che non utilizzano tecniche che prevedono l'addestramento di modelli di AI, i paragrafi da 2 a 5 si applicano solo ai set di dati di prova.

Anche in questo caso, avendo sempre a mente la tutela dei diritti fondamentali della persona, pare recessivo applicare i precetti qualitativi sui dati limitatamente alle AI ad alto rischio: non essendo meno importante la qualità del dato anche con riferimento alle altre AI non ad alto rischio ma soprattutto essendo i precetti generali sulla correttezza e qualità dei dati personali non derogati né derogabili dall'AI Act e pertanto applicabili a tutte le AI in forza dell'art. 5 del GDPR.

7. SORVEGLIANZA UMANA DELLE AI AD ALTO RISCHIO: ART. 14 AI ACT E IL PRINCIPIO "HUMAN IN THE LOOP".

Particolarmente significativo in prospettiva antropocentrica protettiva dei diritti fondamentali della persona l'innovativo *principio di sorveglianza umana*.

L'AI Act statuisce invero all'art. 14 che i sistemi di AI ad alto rischio sono progettati e sviluppati, anche con strumenti di interfaccia uomo-macchina adeguati, in modo tale da poter essere

efficacemente supervisionati da persone fisiche durante il periodo in cui sono in uso.

La sorveglianza umana mira a prevenire o ridurre al minimo i rischi per la salute, la sicurezza o i diritti fondamentali che possono emergere quando un sistema di AI ad alto rischio è utilizzato conformemente alla sua finalità prevista o in condizioni di uso improprio ragionevolmente prevedibile, in particolare qualora tali rischi persistano nonostante l'applicazione di altri requisiti di cui alla presente sezione.

Le misure di sorveglianza sono commisurate ai rischi, al livello di autonomia e al contesto di utilizzo del sistema di AI ad alto rischio e sono garantite mediante almeno uno dei tipi di misure seguenti:

a) misure individuate e integrate nel sistema di AI ad alto rischio dal fornitore prima della sua immissione sul mercato o messa in servizio, ove tecnicamente possibile;

b) misure individuate dal fornitore prima dell'immissione sul mercato o della messa in servizio del sistema di AI ad alto rischio, adatte ad essere attuate dal *deployer*.

Ai fini dell'attuazione dei paragrafi 1, 2 e 3, il sistema di AI ad alto rischio è fornito al *deployer* in modo tale che le persone fisiche alle quali è affidata la sorveglianza umana abbiano la possibilità, ove opportuno e proporzionato, di:

a) comprendere *correttamente* le capacità e i limiti *pertinenti* del sistema di AI ad alto rischio ed essere in grado di monitorarne debitamente il funzionamento, anche *al fine di individuare e affrontare* anomalie, disfunzioni e prestazioni inattese;

b) restare consapevole della possibile tendenza a fare automaticamente affidamento o a fare eccessivo affidamento sull'output prodotto da un sistema di AI ad alto rischio ("distorsione dell'automazione"), in particolare in relazione ai sistemi di AI ad alto rischio utilizzati per fornire informazioni o raccomandazioni per le decisioni che devono essere prese da persone fisiche;

c) interpretare correttamente l'output del sistema di AI ad alto rischio, tenendo conto ad esempio degli strumenti e dei metodi di interpretazione disponibili;

d) decidere, in qualsiasi situazione particolare, di non usare il sistema di AI ad alto rischio o altrimenti di ignorare, annullare o ribaltare l'output del sistema di AI ad alto rischio;

e) intervenire sul funzionamento del sistema di AI ad alto rischio o interrompere il sistema mediante un pulsante di "*arresto*" o una procedura analoga *che consenta al sistema di arrestarsi in condizioni di sicurezza*.

Tasto digitale? pulsante analogico? Disconnessione dalla rete di alimentazione?

Il precetto del *pulsante d'arresto* risulta particolarmente suggestivo ma non sappiamo ancora quanto efficace in concreto visto che parrebbero essere già stati riscontrati dai ricercatori primi

inquietanti segnali di autoprotezione di taluni più avanzati sistemi di ai di fronte ai tentativi di disattivazione da parte umana.

Tale previsione normativa — ai limiti del *naïf* — ricorda i tentativi, infruttuosi, di bypassare Hal9000 del protagonista di Odissea nello Spazio e di tanta filmografia fantascientifica americana: anche se il pulsante di arresto si rivela effimera tutela in quanto viene quasi sempre neutralizzato dall'AI e gli esiti non sono mai prevedibili per l'uomo...

8. L'AI ACT TRA IPOREGOLAZIONE E IPERREGOLAZIONE.

Un primo aspetto positivo dell'AI Act è che disponiamo di un quadro di regole, marcatamente antropocentrico, cui fare riferimento in un momento tecnologico di grande fermento.

Profilo che, tuttavia, non è esente da critiche. Invero, troppe regole in un contesto non ancora consolidato e in continua forte evoluzione, come si è già accennato, rischiano di disorientare non solo l'intrapresa ma anche l'interprete. Meglio sarebbe stato elaborare pochi principi alti da osservare subito a protezione della dignità della persona rinviando la normativa di dettaglio a una fase successiva e tecnologicamente più matura.

Apprezzabile è il superamento del mero piano etico, certamente importante ma non cogente, con l'introduzione di precetti normativi certi e vincolanti: integrati, ma non sostituiti, da codici di condotta coregolamentati e non meramente autodisciplinati da vaghi codici autoreferenziali non effettivamente vincolanti ma lasciati all'autonomia discrezionale del soggetto che li ha adottati.

Ulteriore aspetto positivo che emerge dall'analisi comparata dei due testi normativi — GDPR e AI ACT — è che la protezione dei dati personali si conferma centrale e non recede rispetto all'AI: prevale dunque sempre e comunque il GDPR.

Semmai si pone un problema di effettiva applicazione e compliance del GDPR da parte degli sviluppatori e *deployer* di AI.

Un terzo aspetto, almeno parzialmente, positivo riguarda certamente la statuizione del principio *human in the loop* per le AI *ad alto rischio* (art. 14 AI Act).

Parzialmente perché manca un principio generale di sorveglianza umana — che nella formulazione attuale è per legge limitato alle sole AI *ad alto rischio* — a tutte le attività decisorie algoritmiche delegate ad intelligenze artificiali.

Non sfuggono tuttavia diverse criticità, anche a prima lettura.

Innanzitutto, una critica all'arco temporale di entrata in vigore eccessivamente differenziato e dilatato nel tempo.

La maggior parte delle nuove regole andrà a regime dopo 24 e addirittura 36 mesi per le AI *ad alto rischio*: un periodo davvero troppo lungo, inaccettabile, che mette a rischio di obsolescenza una normativa così dipendente dall'evoluzione tecnologica.

Almeno i divieti relativi a pratiche vietate sono entrati in vigore solo 6 mesi dopo l'entrata in vigore.

I codici di buone pratiche dopo 9 mesi.

Le norme sui sistemi di AI per finalità generali (*General Purpose AI Systems*), incluse quelle relative alla governance, saranno applicabili dopo 12 mesi dall'entrata in vigore del 1° agosto 2024.

Una seconda criticità — riguarda ad avviso dello scrivente — l'ambito di applicazione che con un errore di prospettiva di non poco momento esclude due settori strategici piuttosto pericolosi:

— sistemi utilizzati esclusivamente per scopi militari, di difesa o di sicurezza nazionale;

— sistemi di intelligenza artificiale utilizzati solo a scopo di ricerca e innovazione.

Qualche dubbio anche per l'esclusione dell'utilizzo dell'AI per scopi non professionali: forse sarebbe stato il caso di precisare espressamente, sebbene possa apparire ovvio, ma in questo contesto estremamente nuovo e per certi aspetti imprevedibile non è opportuno dare nulla per scontato, purché trasparente e lecito considerate le molteplici finalità illecite che potrebbero non sempre rilevare sotto il profilo penalistico.

Proprio l'utilizzo di AI per scopi militari è stato giustamente stigmatizzato anche dal Santo Padre al G7 ospitato dall'Italia con un fermo monito al divieto di armi automatizzate per uccidere senza sorveglianza umana.

Così pure non convince del tutto l'assenza di principi generali applicabili minimi anche alle AI a rischio minimo — come le AI ludiche — che paiono disinteressare, a torto visto i potenziali danni per i minori, il legislatore europeo, forse con un difetto di regolazione.

Anche l'approccio del legislatore ai profili della tutela dei diritti fondamentali pare ben più timido rispetto al GDPR.

Inoltre le esigenze e i rischi dell'uso dell'intelligenza artificiale variano in modo significativo a seconda che ciò avvenga in un ambiente privato o in un contesto accessibile al pubblico.

Le eccezioni di cui all'articolo 5, paragrafo 1, lettera d), per l'uso di applicazioni biometriche di identificazione remota in tempo reale in spazi accessibili al pubblico a fini di contrasto sono di portata eccessiva e non corrispondono un'ingerenza proporzionata nei diritti fondamentali dei cittadini.

Anche l'uso di applicazioni biometriche di identificazione remota a fini di contrasto costituisce un'ingerenza intensiva nei diritti fondamentali dei cittadini e avrebbe pertanto dovuto essere incluso nell'elenco delle pratiche vietate di cui all'articolo 5.

La classificazione come applicazione di AI ad alto rischio non pare sufficiente al potenziale di rischio associato all'uso di tali applicazioni.

L'articolo 54, paragrafo 1, prevede un'autorizzazione generale,

indifferenziata e orizzontale al trattamento dei dati personali nell'ambito di test reali.

Dal punto di vista della protezione dei dati, questa disposizione è troppo vaga e non può costituire una base giuridica per il trattamento dei dati compatibile con il GDPR.

Il riutilizzo dei dati personali raccolti per uno scopo specifico, per finalità che non sono in alcun modo correlate allo scopo della raccolta, non è in alcun modo prevedibile per l'interessato.

Nella misura in cui la disposizione è intesa come una forma di « riutilizzo compatibile » ai sensi dell'articolo 6, paragrafo 4, del RGPD, l'articolo 54, paragrafo 1, non è una misura necessaria e proporzionata in una società democratica per proteggere gli obiettivi di cui all'articolo 23, paragrafo 1, conformemente all'articolo 6, paragrafo 4, del RGPD.

Si rileva ancora che la norma dell'AI Act citata non distingue tra categorie particolari di dati personali ai sensi dell'articolo 9, paragrafo 1, del GDPR e altri dati personali.

L'articolo 54, paragrafo 1, pare difficilmente compatibile con il principio minimizzazione dei dati personali ai sensi dell'articolo 5, paragrafo 1, lettera c), del GDPR, in quanto né l'ambito né le categorie di dati personali potenzialmente trattati in laboratori reali sono limitati in alcun modo.

E ancora ai sensi dell'articolo 47 della legge europea sull'AI, i fornitori di sistemi di AI ad alto rischio sono tenuti a redigere vari documenti denominati congiuntamente “dichiarazione di conformità UE”. Questi documenti saranno tenuti a disposizione delle autorità nazionali competenti 10 anni¹⁶.

L'aspetto interessante è che l'Allegato IV spiega cosa deve includere questa “dichiarazione di conformità UE”. Il paragrafo 5 prevede quanto segue:

“Se un sistema di AI comporta il trattamento di dati personali, una dichiarazione che attesti che tale sistema di AI è conforme ai regolamenti (UE) 679/2016 e (UE) 1725/2018 e alla direttiva (UE) 680/2016”.

È chiaro che questa norma è stata redatta con una logica di conformità alla sicurezza dei prodotti e senza valutare una piena compatibilità con la protezione dei dati personali.

Non può essere il “sistema” — che sia AI o meno — ad essere oggetto di valutazione di compatibilità preventiva con il GDPR, se non in linea del tutto teorica, quindi, priva di rilevanza, ma

¹⁶ Art. 47 AI Act: “Il fornitore compila una dichiarazione scritta di conformità UE leggibile meccanicamente, firmata a mano o elettronicamente, per ciascun sistema di IA ad alto rischio e la tiene a disposizione delle autorità nazionali competenti per dieci anni dalla data in cui il sistema di IA ad alto

rischio è stato immesso sul mercato o messo in servizio. La dichiarazione di conformità UE identifica il sistema di IA ad alto rischio per il quale è stata redatta. Su richiesta, una copia della dichiarazione di conformità UE è presentata alle pertinenti autorità nazionali competenti”.

semmai solo i trattamenti dei dati personali effettivamente operati, in concreto, con quel sistema.

La conformità al GDPR è sempre riferibile ai trattamenti di dati personali effettuati. Risulta evidente che non si può garantire con una dichiarazione generale *ex ante* che un sistema di AI sarà sempre conforme al GDPR senza prevedere le tipologie di trattamenti che potranno essere effettuati e le categorie di dati personali utilizzate (comuni oppure anche particolari).

Sarebbe stato forse più opportuno formulare la norma, fermo restando che la conformità operativa dell'AI al GDPR non può che verificarsi in relazione ai trattamenti di dati personali effettivamente operati, richiamando espressamente l'osservanza già in fase di sviluppo ai precetti conformativi generali del *principio di minimizzazione dei dati personali* (art. 5 lett. c) GDPR) e della *privacy by design e by default* (art. 25 GDPR) nello sviluppo e utilizzo delle AI.

Per analizzare se le attività di trattamento dei dati personali sono conformi, le diverse fasi di vita di un sistema di AI devono essere suddivise (ad esempio, sviluppo iniziale, formazione, implementazione, utilizzo dopo l'implementazione) e analizzate dettagliatamente.

L'AI Act pare meno incisivo rispetto al GDPR nella protezione dei diritti fondamentali e più incline ad assecondare le logiche di mercato.

Alcuni spunti di riflessione in tal senso:

— divieti limitati in relazione, per esempio, alla sorveglianza biometrica nelle aree accessibili al pubblico ma ignorati nelle aree private;

— la maggior parte dei requisiti si applicano fondamentalmente solo all'AI ad alto rischio, ad eccezione dell'articolo 6(3) delle deroghe e alcune AI già immesse sul mercato;

— tra l'AI ad alto rischio, abbiamo qualche AI controversa, come i rilevatori di menzogne; noleggio-tech con riconoscimento delle emozioni; polizia predittiva basata sui big data profiling.

— la marcatura CE non significa di per sé 'AI Legale', dal momento che il *deployer* deve ancora garantire il rispetto delle leggi applicabili, compresa, *ex multis*, la protezione dei dati personali ai sensi del GDPR.

Infine, l'AI Act prevede, all'articolo 111, dei termini entro i quali i sistemi di AI già immessi sul mercato o messi in servizio devono conformarsi ai requisiti e agli obblighi ivi previsti.

In particolare, fatta salva l'applicazione delle disposizioni di cui all'articolo 5 (*Pratiche di IA vietate*), a partire da 6 mesi dalla data di entrata in vigore dell'AI Act ossia a decorrere dal 1° agosto 2024:

— i sistemi di AI che costituiscono componenti di sistemi IT su larga scala istituiti dagli atti giuridici elencati nell'Allegato X e che sono stati immessi sul mercato o messi in servizio prima di 36 mesi

dalla data di entrata in vigore dell'AI Act, devono essere resi conformi all'AI Act entro il 31 dicembre 2030;

— l'AI Act si applica agli operatori di sistemi di AI ad alto rischio, diversi dai sistemi descritti nel precedente punto, che sono stati immessi sul mercato o messi in servizio prima di 24 mesi dalla data di entrata in vigore della legge stessa, solo se, a decorrere da tale data, tali sistemi sono soggetti a modifiche significative della loro progettazione.

- Nel caso di sistemi di AI ad alto rischio destinati ad essere utilizzati da parte delle autorità pubbliche, i fornitori e i *deployer* di tali sistemi dovranno adottare le misure necessarie per conformarsi ai requisiti e agli obblighi previsti dall'AI Act entro 6 anni dalla sua entrata in vigore.

Infine, l'articolo 111 prevede altresì che i fornitori di modelli GPAI immessi sul mercato prima di 12 mesi dalla data di entrata in vigore dell'AI Act adottino le misure necessarie per conformarsi agli obblighi della legge stessa entro 36 mesi dalla sua entrata in vigore.

9. L'ELASTICITÀ APPLICATIVA DIFFERERENZIATA ULTRATESTUALE DEI PRINCIPI GENERALI DELL'AI ACT: UNA PROPOSTA DI LETTURA SISTEMATICA E ASSIOLOGICA.

Si registra, infine, nell'AI Act la deprecabile assenza di una norma riepilogativa dei *principi generali* applicabili alle AI sulla falsariga dell'art. 5 del GDPR.

I principi generali per le AI non sono formulati dalla legge europea sull'intelligenza artificiale in modo esplicito e organico come nel GDPR ma vanno estrapolati dalla disamina delle singole norme e da una lettura sistematica del testo complessivo anche alla luce di quanto statuito da altri referenti normativi dell'ormai affollato *corpus iuris digitalis* europeo: si richiamano a tal fine, limitandoci ai principali, gli emblematici modelli delineati dal *Digital Services Act*, *Digital Markets Act*, *Data Governance Act* e *Data Act*.

Si menzionano qui di seguito quelli che, ad avviso dello scrivente, si ritengono più significativi.

Innanzitutto, considerando le correlazioni essenziali tra principi ordinatori su cui si fonda tutto l'impianto normativo dell'AI Act ovvero:

— *principio di accountability*, fornitori e *deployer* sono tenuti ad autovalutare il rischio correlato all'AI sviluppata e immessa sul mercato dal *fornitore*, documentando il processo valutativo aziendale, o impiegata sotto la propria autorità (*deployer*) e il corollario

— *principio di gestione del rischio*, che richiede di applicare i precetti differenziati previsti dall'AI Act in misura minimale per le AI a rischio limitato e massima per le AI ad alto rischio.

Il principio di accountability e gestione del rischio non sono certo, come noto, una novità: essendo, invero, già stati codificati, seppure con riferimento a differente, ma strettamente correlato, settore applicativo della tutela e circolazione dei dati personali, nei ben noti artt. 5, 2, art. 24 e 32 del GDPR.

Corre l'obbligo di rilevare che l'*accountability* dell'AI Act si presenta temperata rispetto a quella del GDPR: nel caso che qui ci occupa è il legislatore dell'AI Act a predeterminare in via generale ed astratta le rischiosità differenziate nell'ambito delle quali il produttore e il *deployer* è chiamato a conformarsi secondo logiche di *compliance differenziate*. Nel GDPR, invece, l'analisi del rischio, non predeterminata dalla legge secondo livelli differenziati d'intensità, è lasciata integralmente all'autovalutazione, autonoma, del titolare del trattamento.

In secondo luogo, occorre evidenziare — pur in assenza di espresso richiamo in tal senso ma in ossequio a lettura sistematica e assiologica costituzionalmente orientata — l'applicabilità ultratestuale dei principi generali, mutuati da altri modelli normativi del contesto digitale o di nuovo conio, applicabili *elasticamente* a tutte le AI — con il *caveat* di applicazione minimale per le AI a rischio minimo e rafforzata per le AI ad alto rischio — a partire dai seguenti:

— *principio di formazione*, c.d. *AI Literacy*: novità introdotta dall'art. 4, AI Act;

— *principio di sicurezza intrinseca ed estrinseca*: tanto più elevato è il rischio tanto più elevate sono le misure di sicurezza e gli obblighi specifici da implementare: ma qualsiasi AI in ossequio a lettura assiologica protettiva, costituzionalmente orientata, della persona deve, in ogni caso, essere sicura in se e per sé in quanto resistente agli attacchi e non pericolosa per le persone;

— *principio di trasparenza*: art. 50, AI Act;

— *principio di sostenibilità ambientale*: novità introdotta dall'art. 1, AI Act.

Iniziamo dall'art. 4 dell'AI Act *Il principio di AI literacy: alfabetizzazione in materia di IA* che statuisce un ampio e rilevantissimo obbligo formativo utile anche per misurare la responsabilizzazione e la responsabilità di fornitori e deployer: “*I fornitori e i deployer dei sistemi di IA adottano misure per garantire nella misura del possibile un livello sufficiente di alfabetizzazione in materia di IA del loro personale nonché di qualsiasi altra persona che si occupa del funzionamento e dell'utilizzo dei sistemi di IA per loro conto, prendendo in considerazione le loro conoscenze tecniche, la loro esperienza, istruzione e formazione, nonché il contesto in cui i sistemi di IA devono essere utilizzati, e tenendo conto delle persone o dei gruppi di persone su cui i sistemi di IA devono essere utilizzati*”.

Segue, non certo per importanza, il *principio di trasparenza* relativo ai sistemi di IA destinati a interagire direttamente con le

persone fisiche, a prescindere dal rischio, integrato dal corollario del *principio di equità e non discriminazione*: art. 13 *Trasparenza e fornitura di informazioni ai deployer* e 50 *Obblighi di trasparenza per i fornitori e i deployers di determinati sistemi di IA dell’Ai Act*. Il considerando 27 precisa che: “Con «trasparenza» si intende che i sistemi di IA sono sviluppati e utilizzati in modo da consentire un’adeguata tracciabilità e spiegabilità, rendendo gli esseri umani consapevoli del fatto di comunicare o interagire con un sistema di IA e informando debitamente i deployer delle capacità e dei limiti di tale sistema di IA e le persone interessate dei loro diritti. Con «diversità, non discriminazione ed equità» si intende che i sistemi di IA sono sviluppati e utilizzati in modo da includere soggetti diversi e promuovere la parità di accesso, l’uguaglianza di genere e la diversità culturale, evitando nel contempo effetti discriminatori e pregiudizi ingiusti vietati dal diritto dell’Unione o nazionale”.

Il considerando 72 aggiunge che: “Per rispondere alle preoccupazioni relative all’opacità e alla complessità di determinati sistemi di IA e aiutare i deployer ad adempiere ai loro obblighi a norma del presente regolamento, è opportuno imporre la trasparenza per i sistemi di IA ad alto rischio prima che siano immessi sul mercato o messi in servizio”.

Il principio di sostenibilità ambientale: specialmente art. 1 *Oggetto dell’Ai Act* che statuisce quanto segue: “Lo scopo del presente regolamento è migliorare il funzionamento del mercato interno e promuovere la diffusione di un’intelligenza artificiale (IA) antropocentrica e affidabile, garantendo nel contempo un livello elevato di protezione della salute, della sicurezza e dei diritti fondamentali sanciti dalla Carta dei diritti fondamentali dell’Unione europea, compresi la democrazia, lo Stato di diritto e la protezione dell’ambiente, contro gli effetti nocivi dei sistemi di IA nell’Unione, e promuovendo l’innovazione”. La sostenibilità ambientale europea — almeno nelle statuizioni di principio — diviene, dunque, parametro rilevante per disincentivare lo sviluppo di AI energivore e inquinante. Invero con scarsi risultati pratici a livello globale essendo — almeno allo stato — le istanze ambientali del tutto secondarie rispetto alla frenesia da sviluppo tecnologico e supremazia nel settore dell’AI.

Seguono poi principi applicabili solamente alle AI *ad alto rischio* cui pare prevalentemente orientato in larga parte lo sforzo regolatorio delle AI.

Primus inter pares non si può non partire dal fondamentale che, a parere di chi scrive, avrebbe potuto più coraggiosamente essere espressamente esteso dal legislatore nella sua portata generale oltre i limiti ristretti delle AI *ad alto rischio*, come fatto per l’art.4 in materia di *AI literacy*:

— *principio del rispetto dell’autonomia decisionale umana*: già perverso statuito nell’art. 22 del GDPR e nell’art.25 del *Digital*

Service Act emerge dagli artt. 13 *Trasparenza e fornitura di informazioni ai deployer* e 86 *Diritto alla spiegazione dei singoli processi decisionali dell'AI Act*; ulteriormente rafforzato dal complementare e innovativo

— *principio di sorveglianza umana*: art. 14 AI Act di cui si è già detto *amplius supra*;

— *principio di conformità*: art. 47 AI Act per effetto del quale il fornitore di AI ad alto rischio è obbligato sotto la propria responsabilità a rilasciare: (i) un'attestazione secondo cui il sistema di AI è conforme ai requisiti di cui alla sezione 2 e riporta le informazioni di cui all'allegato V dell'AI Act e, ove applicabile, a qualsiasi altra disposizione pertinente del diritto dell'Unione che preveda il rilascio di una dichiarazione di conformità UE; (ii) se un sistema di AI comporta il trattamento di dati personali, una dichiarazione attestante che tale sistema di AI è conforme ai regolamenti (UE) 2016/679 e (UE) 2018/1725 e alla direttiva (UE) 2016/680.

Emerge, inoltre, il *principio di sicurezza rafforzato per le persone e di sicurezza intrinseca delle AI ad altro rischio (safety and security)*: il *principio di accountability* si traduce nei sistemi ad alto rischio nell'adozione di misure protettive, di sicurezza, di conformità, trasparenza e tracciabilità rafforzata: si vedano in particolare, senza pretese di esaustività, gli artt. 9 *Sistema di gestione dei rischi in continuo*, 11 *Documentazione tecnica*, 12 *Conservazione dei log degli eventi*, 15 *Accuratezza, robustezza e cbersicurezza*, 18 *Conservazione dei documenti* e 25 *Responsabilità lungo la catena del valore dell'AI*. Analoghe considerazioni valgono per le misure protettive, di sicurezza, di conformità, trasparenza e tracciabilità rafforzata richieste per i modelli di AI per finalità generali con rischio sistemico (art. 55 AI Act).

Così pure prima di utilizzare sotto la propria autorità un AI ad alto rischio in attuazione dei principi ordinatori di accountability e gestione differenziata del rischio occorre predisporre un'accurata preliminare *Valutazione d'impatto dei diritti fondamentali per i sistemi ad altro rischio* ai sensi dell'art. 27 dell'AI Act che statuisce quanto segue: “prima di utilizzare un sistema di IA ad alto rischio di cui all'articolo 6, paragrafo 2, ad eccezione dei sistemi di IA ad alto rischio destinati a essere usati nel settore elencati nell'allegato III, punto 2, i deployer che sono organismi di diritto pubblico o sono enti privati che forniscono servizi pubblici e i deployer di sistemi di IA ad alto rischio di cui all'allegato III, punto 5, lettere b) e c), effettuano una valutazione dell'impatto sui diritti fondamentali che l'uso di tale sistema può produrre”.

Last but not least, il *principio di tutela della riservatezza, qualità dei dati personali e data governance* relativo alle AI ad alto rischio: art. 10 AI Act. Si veda in particolare il considerando 67 che precisa quanto segue: “Il requisito secondo cui i set di dati dovrebbero essere, per quanto possibile, completi ed esenti da

errori non dovrebbe incidere sull'uso di tecniche di tutela della vita privata nel contesto dello sviluppo e della prova dei sistemi di IA. In particolare i set di dati dovrebbero tenere conto, nella misura necessaria per la finalità prevista, delle caratteristiche o degli elementi particolari dello specifico contesto geografico, contestuale, comportamentale o funzionale nel quale il sistema di IA ad alto rischio è destinato a essere usato. I requisiti relativi alla governance dei dati possono essere soddisfatti ricorrendo a terzi che offrono servizi di conformità certificati, compresa la verifica della governance dei dati, dell'integrità dei set di dati e delle pratiche di addestramento, convalida e prova dei dati, purché sia garantita la conformità ai requisiti in materia di dati di cui al presente regolamento"; con il corollario dell'innovativo principio di standardizzazione dei dati di addestramento.

Emerge, infine, il *principio di standardizzazione* dei dati di addestramento che si applica alla conformità normativa nell'ambito dell'AI Act e riguarda la trasformazione dei dati in un formato coerente e uniforme per garantire una buona qualità e affidabilità nel processo di training dei modelli di intelligenza artificiale.

In particolare l'innovativo *principio di standardizzazione dei dati di addestramento* nell'AI Act si evince dall'art. 10 dell'AI Act che richiede per le AI ad alto rischio che i set di dati per addestramento, convalida e test siano soggetti a pratiche di governance e gestione che assicurino la qualità, l'origine, la pertinenza e la rappresentatività dei dati in modo appropriato agli scopi del sistema e dall'art. 53, comma 1, lett. d) che richiede ai fornitori delle AI per finalità generali di fornire una sintesi sufficientemente dettagliata dei dati utilizzati per l'addestramento dei modelli di intelligenza artificiale ¹⁷.

I set di dati di addestramento devono rispettare requisiti rigorosi di qualità, trasparenza e gestione durante tutte le fasi del ciclo di vita del dato, dal pre-training al fine-tuning del modello. La standardizzazione è parte integrante di questo processo, in quanto permette di eliminare *bias*, fornire dati coerenti e affidabili, fondamentali per l'efficacia e la correttezza dei modelli di AI.

¹⁷ L'art. 40 dell'AI Act statuisce che: "I sistemi di IA ad alto rischio o i modelli di IA per finalità generali che sono conformi alle norme armonizzate o a parti di esse i cui riferimenti sono stati pubblicati nella *Gazzetta ufficiale dell'Unione europea* conformemente al regolamento (UE) n. 1025/2012 si presumono conformi ai requisiti di cui alla sezione 2 del presente capo o, se del caso, agli obblighi di cui al capo V, sezioni 2 e 3, del presente regolamento, nella misura in cui tali requisiti o obblighi sono contemplati da tali norme". Si sottovaluta spesso

che l'AI Act si inserisce nella cornice del New Legislative Framework (NLF) UE (Reg. UE 765/2008). Questo quadro normativo è stato pensato per armonizzare le normative tecniche, semplificare il controllo di conformità dei prodotti all'interno del Mercato Unico Europeo e introdurre una politica comune sulla sorveglianza del mercato, focalizzando l'attenzione sull'applicazione delle normative durante l'intero ciclo di vita dei prodotti. Allo stato tali norme armonizzate non sono ancora state pubblicate creando problemi applicativi di non poco momento.

Il principio di standardizzazione è essenziale per assicurare che i processi di addestramento AI siano conformi all'AI Act e al GDPR, trasparenti, e affidabili, e che rispettino gli obblighi di protezione dei dati e di equità previsti dalla normativa europea.

La standardizzazione dei dati è inoltre funzionale ad agevolare l'oggettivazione delle metriche degli audit interni e da parte di revisori terzi indipendenti.

Tuttavia, come si è rilevato, gli apparenti limiti normativi letterali relativi al rispetto dei *principi dell'autonomia decisionale umana, della trasparenza, della riservatezza, dei dati personali e della sicurezza*, paiono ancora una volta, in attuazione dei generali principi di ragionevolezza e proporzionalità, superabili in via interpretativa in ossequio a lettura sistematica assiologica italo-europea, costituzionalmente orientata, a tutela dei diritti fondamentali della persona.

Il rispetto dell'autonomia decisionale umana, il principio di trasparenza, la tutela della riservatezza, la tutela dei dati personali devono essere giustamente osservati, sempre e comunque, in tutti i contesti di sviluppo, immissione sul mercato e impiego sotto la propria autorità di sistemi di AI, parendo irragionevole il contrario, a prescindere dalla classificazione del rischio dell'AI come alto oppure limitato o minimo e a prescindere dall'*ombrello* protettivo in ogni caso offerto dalla sovrapposizione prevalente dei precetti inderogabili del GDPR.

Emerge, corre l'obbligo di sottolineare, la peculiare caratteristica di *elasticità applicativa* propria di tali *principi generali* che richiedono di essere applicati con *intensità differenziata* in ragione del diverso livello di rischio dell'AI in conformità ai summenzionati *principi ordinatori* dell'AI Act di *accountability* e *gestione del rischio*.

10. VULNERABILITÀ DIGITALE DELLA PERSONA NELL'EPOCA DELLE AI.

L'AI Act offre il destro per approfondire il tema correlato della vulnerabilità digitale.: condizione umana che nel contesto digitale presenta profili meritevoli di approfondimento nella prospettiva, ma non solo, del consumatore ¹⁸.

¹⁸ Si vedano le utili riflessioni in materia di *consumatore frastagliato* in base al livello di vulnerabilità della persona fisica, consistente nel *gap* conoscitivo che lo rende maggiormente esposto al rischio di sopraffazione da parte del professionista: R. SENICACIA, *Consumatore, vulnerabilità e principio di trasparenza tra razionalità e bias*

cognitivi, in *La Nuova Giurisprudenza Civile Commentata*, 2024, 1522 ss.; M. RABITTI, *Il consumatore vulnerabile e la fragilità del diritto. Brevi considerazioni*, in *Dialoghi dir. ec.*, 2023, 351; S. PAGLIANTINI, *Il consumatore frastagliato. (Istantanee sull'asimmetria contrattuale tra vicende circulatorie e garanzie)*, Pacini, 2021, *passim*.

La vulnerabilità è una caratteristica intrinseca degli esseri umani che è destinata ad amplificarsi con la diffusione progressiva delle AI in ecosistema digitale già fortemente asimmetrico.

Come noto in determinati contesti, la nostra dipendenza dal contesto familiare, economico e sociale ci espone a squilibri di potere enfatizzando la nostra vulnerabilità.

Ora anche la rivoluzione digitale concorre a determinare tale complesso fenomeno, esponendo le nostre vite ad architetture tecnologiche che creano dipendenza, a pratiche commerciali mentalmente manipolative, alla raccolta abusiva di dati comportamentali, a forme occulte di discriminazione e molto altro ancora.

Il GDPR è la prima “legge digitale” che riconosce la maggiore vulnerabilità della persona nei contesti asimmetrici digitali e in rapporto con il potere tecnologico dominante dell’algoritmo.

La nozione di “consumatori vulnerabili” — introdotta nella legislazione dell’UE nel 2005 con la direttiva sulle pratiche commerciali sleali — presenta alcune analogie con il concetto di “soggetti vulnerabili”¹⁹.

Nel testo finale dell’AI Act ci sono ben 16 riferimenti al concetto

Sul concetto di vulnerabilità in dottrina si vedano P. E. CORRIAS - E. PIRAS (a cura di), *I soggetti vulnerabili nell'economia, nel diritto e nelle istituzioni*, Napoli, 2021 *passim*; P. E. CORRIAS, (a cura di), *I soggetti vulnerabili nella disciplina comune e nei mercati regolamentati*, Napoli, 2022. In particolare: E. Battelli, *Dal concetto di debolezza alla nozione di vulnerabilità*, *op. cit.*, pp. 35-58; A. FUSARO, *L'atto patrimoniale della persona vulnerabile*, *I* 2019, 3 ss.; A. GENTILI, *La vulnerabilità sociale. Un modello teorico per il trattamento legale*, in *Riv. crit. dir. priv.*, 2019, 46 ss. Anche la Corte di Giustizia UE segnala il progressivo abbandono della del modello astratto di consumatore medio e del giudizio di scorrettezza della pratica commerciale in ragione delle scelte svantaggiose (o meno vantaggiose) poste in essere dallo stesso, a favore della progressiva rilevanza ai bias cognitivi che incidono sulla libertà di decisione. In giurisprudenza si veda *ex multis* Corte giust. UE, 14.11.2024, C-646/22, Compass Banca Spa c. Autorità Garante della Concorrenza e del Mercato. La Corte evidenzia che, secondo il considerando 18 della direttiva 2005/29, « il consumatore medio è una persona, da un lato, normalmente informata e, dall'altro, ragionevolmente attenta ed avveduta. Tuttavia, dato che, conformemente all'art. 7 della direttiva 2005/29, spetta al professionista fornire ai consumatori le informazioni rilevanti di cui questi ultimi hanno bisogno, nella fattispecie concreta,

per prendere la loro decisione, tale caratteristica del consumatore medio, di essere « normalmente informato », deve essere intesa come riferita alle informazioni che si possono ragionevolmente presumere note ad ogni consumatore, tenendo conto dei pertinenti fattori sociali, culturali e linguistici, e non alle informazioni proprie dell'operazione di cui trattasi. Di conseguenza, detta caratteristica non esclude che una pratica commerciale possa falsare in misura rilevante il comportamento economico di tale consumatore virtuale a causa di una carenza informativa di quest'ultimo ».

¹⁹ Sul consumatore vulnerabile in generale si veda la Comunicazione del 2016 *Understanding Consumer Vulnerability in the EU's key markets* della Commissione Europea, la quale ha definito il consumatore vulnerabile come un consumatore che per caratteristiche (a) socio-demografiche, (b) comportamentali, (c) per una sua situazione personale o (d) un contesto di mercato: (i) è maggiormente soggetto al rischio di subire esiti negativi sul mercato; (ii) ha una limitata capacità di massimizzare il proprio benessere; (iii) ha difficoltà ad ottenere o ad assimilare l'informazione; (iv) è meno in grado di acquisire scegliere o accedere a prodotti adeguati o (v) è più suscettibile a determinate pratiche di marketing. In proposito, v. anche la relazione annuale sull'attività svolta dell'Autorità Garante della Concorrenza e del Mercato del 31 marzo 2023, in cui, oltre a sottolineare la particolare attenzione che occorre avere nei confronti

di vulnerabilità umana, sebbene il Regolamento rinunci a dare una definizione normativa espressa di vulnerabilità limitandosi a presupporla.

Può essere utile richiamare qualche norma dell'AI Act per meglio esemplificare.

I sistemi di AI che sfruttano determinate vulnerabilità umane sono vietati [articolo 5, paragrafo 1, lettera b)].

Inoltre, le vulnerabilità umane sono un parametro per aggiornare l'elenco dei "sistemi di IA ad alto rischio" in futuro (articolo 7, lettera h)).

Tali vulnerabilità devono essere analizzate e attenuate nella nuova *valutazione d'impatto sui diritti fondamentali* da parte di coloro che utilizzano l'AI ad alto rischio (art. 27 AI Act) e considerate con "particolare attenzione" dalle autorità di vigilanza del mercato quando si tratta di sistemi di IA che presentano rischi (arti. 79, par. 2).

L'art. 27 dell'AI Act stabilisce che prima di utilizzare un sistema di AI ad *alto rischio*, ad eccezione dei sistemi di IA ad alto rischio destinati a essere usati nel settore elencati nell'allegato III, punto 2, i *deployer* che sono organismi di diritto pubblico o sono enti privati che forniscono servizi pubblici e i *deployer* di sistemi di AI ad alto rischio di cui all'allegato III, punto 5, lettere b) e c), effettuano una valutazione dell'impatto sui diritti fondamentali che l'uso di tale sistema può produrre²⁰.

A tal fine, i *deployer* effettuano una valutazione che comprende gli elementi seguenti:

a) una descrizione dei processi del *deployer* in cui il sistema

dei consumatori vulnerabili, si evidenzia l'emersione di nuove forme di vulnerabilità: « In un brevissimo arco temporale la vertiginosa accelerazione dell'evoluzione tecnologica e digitale da un lato, e la crisi socio-economica innescata dalla pandemia e acuita dal conflitto russo-ucraino con le conseguenti dinamiche inflattive dall'altro, hanno generato nuove forme di vulnerabilità legate a situazioni endogene e di contesto, che hanno posto in termini nuovi il problema della tutela dei consumatori, spingendo legislatori e istituzioni pubbliche a farsi carico dei nuovi bisogni di protezione. Il riferimento è in primis alla vulnerabilità digitale e alla vulnerabilità economica-finanziaria. (...) La direttiva Omnibus (2019/2161/UE), entrata in vigore ad ottobre 2022, nasce nell'ambito del c.d. New Deal for Consumer con l'intento di rimediare alle lacune e alle incongruenze dell'acquis in materia di tutela del consumatore, assicurando che le regole vigenti rispondano in maniera efficace alle sfide dell'economia digitale ».

²⁰ L'art. 27 dell'AI Act stabilisce inoltre che: "2. L'obbligo di cui al paragrafo 1 si applica al primo uso del sistema di IA ad alto rischio. Il *deployer* può, in casi analoghi, basarsi su valutazioni d'impatto sui diritti fondamentali effettuate in precedenza o su valutazioni d'impatto esistenti effettuate da un fornitore. Se, durante l'uso del sistema di IA ad alto rischio, ritiene che uno qualsiasi degli elementi elencati al paragrafo 1 sia cambiato o non sia più aggiornato, il *deployer* adotta le misure necessarie per aggiornare le informazioni. 3. Una volta effettuata la valutazione di cui al paragrafo 1 del presente articolo, il *deployer* notifica all'autorità di vigilanza del mercato i suoi risultati, presentando il modello compilato di cui al paragrafo 5 del presente articolo nell'ambito della notifica. Nel caso di cui all'articolo 46, paragrafo 1, i *deployer* possono essere esentati da tale obbligo di notifica. 4. Se uno qualsiasi degli obblighi di cui al presente articolo è già rispettato mediante la valutazione d'impatto sulla protezione dei dati effettuata a norma dell'articolo 35

di IA ad alto rischio sarà utilizzato in linea con la sua finalità prevista;

b) una descrizione del periodo di tempo entro il quale ciascun sistema di IA ad alto rischio è destinato a essere utilizzato e con che frequenza;

c) le categorie di persone fisiche e gruppi verosimilmente interessati dal suo uso nel contesto specifico;

d) i rischi specifici di danno che possono incidere sulle categorie di persone fisiche o sui gruppi di persone individuati a norma della lettera c), del presente paragrafo tenendo conto delle informazioni trasmesse dal fornitore a norma dell'articolo 13;

e) una descrizione dell'attuazione delle misure di sorveglianza umana, secondo le istruzioni per l'uso;

f) le misure da adottare qualora tali rischi si concretizzino, comprese le disposizioni relative alla governance interna e ai meccanismi di reclamo.

L'AI Act chiarisce che la valutazione d'impatto dei diritti fondamentali — meglio nota come FRIA —, acronimo di *fundamental rights impact assesment* — *non sarà applicabile indistintamente a tutti gli utilizzi di AI ma solamente nei seguenti casi*:

a) solo per aspetti non coperti da altri obblighi, come la valutazione d'impatto sulla protezione dei dati personali meglio nota come DPIA acronimo di data *protection impact assessment* ai sensi dell'art. 35 del GDPR;

b) se uno qualsiasi degli obblighi relativi alla FRIA è già soddisfatto attraverso la DPIA condotta ai sensi dell'art. 35 del GDPR, la FRIA deve integrare la DPIA (art. 27, par. 4, AI Act).

La FRIA è un adempimento ulteriore e distinto dalla Valutazione di conformità di cui all'art. 47 AI Act che si vedrà *infra*. Essa non si limita alla verifica documentata dei requisiti previsti per i sistemi di AI ad alto rischio, ma comporta un'analisi approfondita dell'impatto sui diritti fondamentali derivante dall'utilizzo di AI ad alto rischio.

L'obbligo di svolgere la FRIA diventerà esecutivo 24 mesi dopo la pubblicazione dell'AI Act nella Gazzetta ufficiale dell'UE: purtroppo, dunque, non prima della primavera del 2026.

I “codici di condotta” dell'AI Act dovranno valutare e prevenire l'impatto negativo dei sistemi di IA sulle persone vulnerabili (art. 90).

Nell'ambito degli spazi di sperimentazione normativa nell'AIA,

del regolamento (UE) 2016/679 o dell'articolo 27 della direttiva (UE) 2016/680, la valutazione d'impatto sui diritti fondamentali di cui al paragrafo 1 del presente articolo integra tale valutazione d'impatto sulla protezione dei dati. 5. L'ufficio per l'IA

elabora un modello di questionario, anche attraverso uno strumento automatizzato, per agevolare i *deployer* nell'adempimento dei loro obblighi a norma del presente articolo in modo semplificato”.

gli interessati in condizione di vulnerabilità a causa dell'età o della disabilità devono essere "adeguatamente tutelati" (art. 60).

Nonostante tutti questi riferimenti alla vulnerabilità umana, c'è ancora una notevole incertezza sulla definizione del concetto di vulnerabilità.

Infatti, l'articolo 3 dell'AI Act contiene 68 definizioni di concetti e termini, nessuna delle quali riguarda la vulnerabilità umana: la scelta non è causale ma discende dalla difficoltà di formulare una definizione univoca generale ed astratta essendo, come noto, le situazioni di vulnerabilità scrutinabili caso per caso in concreto in ragione di molteplici ed eterogenei fattori ambientali.

Anche per tale ragione l'utilizzo del concetto di vulnerabilità risente di tale vaghezza anche nello stesso AI Act in cui l'utilizzo non risulta rigoroso e univoco.

Il riferimento più interessante e vicino a una definizione di vulnerabilità può essere rinvenuto nell'articolo 7, lettera h):

"la misura in cui esiste uno squilibrio di potere o le persone che potrebbero subire il danno o l'impatto negativo si trovano in una posizione vulnerabile rispetto al *deployer* di un sistema di IA, in particolare a causa della condizione, dell'autorità, della conoscenza, della situazione economica o sociale o dell'età".

Se la Commissione europea vuole aggiornare l'elenco dei sistemi di IA ad alto rischio in futuro, dovrà considerare, tra gli altri parametri, proprio anche il parametro della vulnerabilità nei termini suindicati.

L'AI Act utilizza in questo caso un concetto di vulnerabilità come elemento graduale e contestuale. Il termine "persone... in una posizione vulnerabile" trasmette l'idea, in linea con quanto *supra* rilevato, che la vulnerabilità sia una *condizione relazionale* che non può essere predefinita e cristallizzata per legge.

Si può anche osservare che la vulnerabilità è condizione relazionale ("posizione vulnerabile in relazione a") basata su uno "squilibrio di potere", che può essere generato da caratteristiche personali della persona debole ("conoscenza, età, status"), della parte forte ("autorità"), o da fattori sociali ("circostanze economiche o sociali").

L'indicazione non è tassativa: invero queste sono le principali ma non sono le uniche possibili fonti di vulnerabilità ammesse dall'AI Act, poiché l'articolo 7, lettera h), dice "in particolare", ammettendo conseguentemente altre condizioni strutturali, esterne o interne, che potrebbero generare vulnerabilità da valutare caso per caso, in concreto.

Rientrano, a parere di chi scrive, tra le situazioni di vulnerabilità anche il digital divide e l'analfabetizzazione digitale.

Il preambolo dell'AI Act contiene altri tre casi specifici di vulnerabilità umana enunciati espressamente dal legislatore europeo.

In particolare, essa comprende: le vulnerabilità online dei minori (considerando 48); le persone che richiedono o ricevono prestazioni o servizi pubblici essenziali a causa della loro tipica “dipendenza” da tali prestazioni (considerando 58); e le persone che sono soggette a sistemi di AI nella gestione della migrazione, dell’asilo e del controllo delle frontiere (considerando 60), in quanto “dipendono dall’esito delle azioni delle autorità pubbliche competenti”.

In tale prospettiva antropocentrica e protettiva delle situazioni di vulnerabilità della persona pare emblematico il divieto generalizzato di sfruttamento delle situazioni di vulnerabilità della persona o di un gruppo specifico di persone.

L’art. 5, par. 1, lett. b), dell’AI Act che vieta la commercializzazione o l’uso di un “sistema di IA che sfrutta una qualsiasi delle vulnerabilità di una persona o di un gruppo specifico di persone a causa della loro età, disabilità o una specifica situazione sociale o economica, con l’obiettivo, o l’effetto, di falsare materialmente il comportamento di tale persona o di una persona appartenente a tale gruppo in un modo che causi o sia ragionevolmente probabile che causi tale persona o un danno significativo a un’altra persona”.

Sebbene sembri riferirsi a “una qualsiasi delle vulnerabilità di una persona o di un gruppo”, chiarisce le fonti specifiche di vulnerabilità (“la loro età, disabilità o una specifica situazione sociale o economica”).

Mentre potremmo facilmente interpretare i casi di vulnerabilità in base a “età e disabilità”, la nozione di “specifico situazione sociale o economica” sembra più vaga. Il considerando 29 cita alcuni esempi non esaustivi, vale a dire “persone che vivono in condizioni di estrema povertà, minoranze etniche o religiose”.

La “situazione sociale o economica” di cui all’articolo 5, paragrafo 1, lettera b), è tutt’altro che chiara.

Tuttavia, l’analisi testuale dell’AI Act potrebbe offrire un elenco più completo di casi, comprese le persone con redditi più bassi o che appartengono a specifici gruppi emarginati, ad esempio minoranze politiche, linguistiche o razziali, migranti.

Alcuni esempi eclatanti di situazioni di vulnerabilità escluse dalla formulazione testuale dell’articolo 5 dell’AI Act paiono poter essere individuate nelle persone vittime di violenza di genere, i consumatori, i dipendenti, i soggetti senza disabilità ma portatori di gravi malattie, i giocatori d’azzardo e le persone dipendenti dai social media o con altre dipendenze specifiche.

Il richiamo ai fattori di vulnerabilità nell’AI Act sono molto più ampi nell’articolo 7, paragrafo 2, lettera f), che nell’articolo 5, paragrafo 1, lettera b), dove non vi è alcun riferimento specifico allo squilibrio di potere, all’autorità o all’asimmetria della conoscenza.

La ragione di questa discrepanza è probabilmente che l'articolo 5 individua alcuni sistemi di AI vietati, mentre l'articolo 7 fornisce invece istruzioni all'autorità di regolamentazione.

Particolarmente utili anche nel contesto specifico delle AI possono considerarsi le riflessioni consolidate in relazione alla nozione di vulnerabilità interpersonale, relazionale, contestuale e sbilanciata dal punto di vista del potere elaborata dalla dottrina in materia di protezione dei dati personali ²¹.

Sarà, in ogni caso, fondamentale, anche in questo caso, l'elaborazione giurisprudenziale della Corte di Giustizia UE sul punto specifico della tutela dei soggetti vulnerabili e in particolare del divieto di approfittamento di talune situazioni di vulnerabilità — che si auspica la più ampia possibile in conformità a lettura assiologica costituzionalmente orientata nel sistema italo-europeo delle fonti — per cogliere l'effettiva ampiezza protettiva di tale fondamentale previsione posta a presidio dei diritti fondamentali della persona nel contesto strutturalmente decettivo delle intelligenze artificiali.

In conclusione, sebbene l'AI Act si possa certamente considerare uno degli atti legislativi più avanzati al mondo per quanto riguarda il riconoscimento e la protezione delle vulnerabilità umane, l'applicazione protettiva concreta di tale concetto elastico è, allo stato, ancora complesso e problematico non privo di dubbi interpretativi che solo il tempo e l'elaborazione dottrinale e giurisprudenziale potranno chiarire.

11. TUTELA DELLA PERSONA E CREPUSCOLO DEL CONSENSO NEL CONTESTO ASIMMETRICO DIGITALE DELLE AI: L'EMERSIONE PROGRESSIVA DELLA FUNZIONE PROTETTIVA DELLA CONFORMAZIONE ETEROREGOLAMENTARE.

Come già rilevato l'AI Act si segnala — pur nella sua complessità per certi aspetti ridondante e preludio di un rischio di *iper-regolazione* in fase attuativa — per la marcata prospettiva antropocentrica pur senza rinunciare ad assicurare lo sviluppo e la diffusione delle AI.

Si tratta, invero, di uno degli atti legislativi più avanzati al mondo per quanto riguarda il riconoscimento e la protezione delle vulnerabilità umane nel contesto relazionale tecnologico digitale.

L'inarrestabile evoluzione tecnologica digitale fortemente asimmetrica segna un cambio di paradigma della modalità di controllo: non più centralizzato — evocato dal modello distopico orwelliano di 1984 — ma sempre più ubiquitario e immersivo di tipo kafkiano sul modello distopico de *Il processo* per intenderci.

²¹ SI VEDA SUL TEMA G. MALGIERI, *Vulnerability and Data Protection Law*, cit., *passim*.

Non un solo centro, che crea oppressione attraverso la sorveglianza stile Grande Fratello, ma un controllo digitale diffuso — dirompente e (pre)potente — che mira a disorientare la persona creando un senso di impotenza e incertezza strutturali ²².

Il primo modello, anche se invasivo, può essere contrastato, perché più facilmente individuabile; ma il secondo, risulta insuperabile dal singolo poiché non si ha più un soggetto centrale individuato e da combattere ma un controllo diffuso, opaco, tra più soggetti — privati — straordinariamente potenti che esercitano un controllo labirintico più difficile da arginare efficacemente.

L'autodeterminazione della singola persona non è più efficace, da sola, in quanto soggiace ai ben noti meccanismi manipolativi del *nudge tech* e *dark pattern* che agevolano le scelte gradite al contraente forte (un solo bottone da premere) e rendono disagiati quelle meno gradite (molteplici finestre e scelte defatiganti da operare).

È, dunque, certamente più facile per le persone trovarsi soggiate dall'assetto tecnologico digitale immersivo dominante con inesistente, o limitato, effettivo esercizio dei propri diritti, dell'autodeterminazione informativa, della tutela dei propri dati e in ultima istanza della dignità della persona.

Per tale ragione il tema della tutela della persona e dei diritti fondamentali della personalità alla luce delle nuove vulnerabilità risulta particolarmente fecondo nel contesto digitale pervaso dalle AI.

In conclusione, i presidi a tutela della dignità della persona oltre che dei suoi dati personali e degli altri diritti fondamentali, nel contesto tecnologico digitale, presente e ancor più in futuro, devono essere sempre meno affidati esclusivamente alle scelte di autodeterminazione individuale e sempre più a una dimensione pubblica di eteroconformazione dettata dai precetti di legge inderogabili e di tutela collettiva di categorie di soggetti vulnerabili oltre che dalle competenti autorità di regolazione.

Per tal via giungendo persino ad ipotizzare una vulnerabilità digitale strutturale nei contesti digitali agiti dalle intelligenze artificiali, potenzialmente manipolative della libertà e consapevolezza del consenso individuale.

Tali scelte, anche di un soggetto capace di agire, saranno solo apparentemente libere ma invero fortemente condizionate, talvolta anche inconsciamente, dal substrato digitale fortemente asimmetrico anche a causa, si ribadisce, della dirompente diffusione delle AI.

La vulnerabilità dei soggetti — meritevoli di protezione raffor-

²² Si veda per tutti P. Stanzone, I "poteri privati" delle piattaforme e le nuove frontiere della privacy, Torino, 2022, *passim*.

zata da parte del soggetto forte — pare una condizione peculiare del contesto digitale e ancora di più nella prospettiva ormai attuale della diffusione generalizzata delle intelligenze artificiali.

In tale peculiare contesto asimmetrico digitale emerge inesorabilmente la fragilità del consenso individuale quale esclusivo e prevalente strumento di tutela della persona.

●Occorre, dunque, a prudente avviso di chi scrive, si ribadisce ancora una volta, in tale peculiare contesto tecnologico digitale vortiginosamente asimmetrico, demitizzare il ruolo giuridico delle scelte individuali dell'autodeterminazione informativa e registrare l'inesorabile e progressivo *crepuscolo del consenso* individuale e dell'autonomia contrattuale.

Pare, invero, necessario in tale prospettiva ermeneutica spostare l'attenzione sempre più verso tutele e rimedi metaindividuali, secondo una visione ordinamentale integrata privatistica-pubblicistica, mutuata dall'esperienza del diritto privato regolatorio, prevedendo sanzioni amministrative e rafforzando obblighi generalizzati di *compliance*, *audit indipendenti* e *accountability* dei *big tech* a prescindere dalle scelte individuali del singolo soggetto ²³.

Il contesto digitale asimmetrico offre una straordinaria occasione teorica per ridisegnare a livello europeo i limiti all'autonomia privata alla luce della prevalente necessità di eteroconformazione protettiva inderogabile a tutela della persona a prescindere dalla supposta libertà del volere: perverso tutta da dimostrare in contesti asimmetrici e opachi del digitale manipolativo — o semplicemente occulto — e delle intelligenze artificiali agentiche.

Il rischio — già segnalato da Kafka nei suoi romanzi seppur riferiti a contesto analogico diverso ma comparabile, per certi aspetti, sotto il profilo della soverchiante e diffusa asimmetria di potere — è sempre quello della sconfitta della libertà individuale che non risulta più in grado, si ribadisce, di operare scelte autonome, consapevoli e affidabili in un contesto artificiale caratterizzato dal dominio tecnologico digitale fortemente asimmetrico e concentrato nel potere decisionale di poche grandi multinazionali: le "big tech" della *silicon valley*.

La persona — in condizioni di forti asimmetrie di potere pervasive — tende, infatti, ad arrendersi alle scelte più confortevoli e semplici, ad uniformarsi alle regole del sistema rinunciando ad esercitare scelte libere e consapevoli difforni dal *mainstream* imposto dal substrato tecnologico dominante.

Per tale ragione, corre l'obbligo di sottolineare ancora una volta, occorre spostare il baricentro protettivo della persona dalle

²³ Si vedano *ex multis* R. NATOLI, *Il Diritto privato regolatorio*, in *Rivista della regolazione dei mercati*, 2020, 134 ss.; M.

ANCELONE - M. ZARRO (a cura di), *Diritto civile della regolazione*, Napoli, 2022, *passim*.

scelte individuali della medesima, rivelatesi assai fragili nel contesto asimmetrico digitale, alle protezioni ordinamentali inderogabili e conformative che assicurino automaticamente livelli protettivi essenziali per la collettività, per gruppi di persone e non solo per l'individuo ²⁴.

Si registra, come già nel contesto del GDPR, anche e soprattutto per la regolazione dell'AI a protezione della persona e dei diritti fondamentali, secondo una lettura assiologica italo-europea costituzionalmente orientata, la necessità di rafforzare le regole di *accountability, compliance, audit indipendenti, responsabilità e sanzionatorie*, anche applicato al *rapporto privatistico* come già utilmente operato nell'esperienza consumeristica delle pratiche commerciali scorrette ²⁵.

Il genere umano deve sottrarsi all'effimera seduzione commerciale, non richiesta, della *privatizzazione della conoscenza* da parte dei nuovi *oracoli digitali*: AI magnificate per creare rapidamente opere d'arte di ogni genere — musicale, grafico-pittorico e letterario — costringendoci a continuare a svolgere attività manuali ripetitive e noiose, in ultima istanza per “pensare” al nostro posto, minacciando il libero arbitrio.

Una rivoluzione tecnologica antropocentrica e sostenibile deve semmai proteggere e valorizzare la dignità umana e tutelare l'ambiente in cui l'uomo vive: consentire di liberare risorse speculative a favore della persona, aumentare il tempo libero dedicato ad attività intellettuali creative in quanto le AI si sostituiranno all'uomo per ottimizzare l'esecuzione di attività manuali, seriali e banali e non consumare inutilmente risorse ambientali preziose per inutili quesiti teorici gestibili con l'intelligenza naturale, senza rinunciare a pensare.

Non il contrario: la creatività, la conoscenza e il sapere devono rimanere patrimonio comune dell'umanità.

²⁴ Si vedano *ex multis* sui contratti regolamentati e conformati: G. BERTI DE MARINIS, *Contratti dei mercati regolamentati: norme imperative e conformazione*, Napoli, 2019; C. SOLINAS, *Il contratto amministrato. La conformazione dell'operazione economica agli interessi generali*, Napoli, 2018; e da ultimo V. RICCIUTO, *L'equivo della privacy*, Napoli, 2022, *passim*.

²⁵ All'AI Action Summit di Parigi sono state evidenziate le seguenti priorpubblicità nelle dichiarazioni conclusive — riferite esclusivamente alle applicazioni per uso civile — sottoscritte in data 11 febbraio 2025 da 61 paesi, Cina compresa, con la rilevante eccezione di Stati Uniti e Regno Unito, e precisamente:

— *Promoting AI accessibility to reduce digital divides;*

— *Ensuring AI is open, inclusive, transparent, ethical, safe, secure and trustworthy, taking into account international frameworks for all;*

— *Making innovation in AI thrive by enabling conditions for its development and avoiding market concentration driving industrial recovery and development;*

— *Encouraging AI deployment that positively shapes the future of work and labour markets and delivers opportunity for sustainable growth;*

— *Making AI sustainable for people and the planet;*

— *Reinforcing international cooperation to promote coordination in international governance.*

Non si tratta di demonizzare questi innovativi e potenti strumenti di calcolo statistico, certamente utili in determinati contesti produttivi e non necessariamente in tutti i settori del vivere civile: si tratta tuttavia di comprendere il significato dell'utilizzo di tali strumenti e soprattutto ponderare i risultati.

Trattandosi di risultati stocastici basati su pattern complessi e sempre più elevati parametri di calcolo sono destinati a generare risultati che sono solo tendenzialmente corretti mai certamente veri che richiedono sempre prudente controllo umano.

Questo tema riguarda proprio la diffusione della cultura dell'intelligenza artificiale che allo stato pare ancora povera al di fuori del novero ristretto degli addetti ai lavori: l'*AI literacy* non è solo un tema di *compliance* ai precetti normativi europei ma è la vera sfida culturale posta dall'art. 4 dell'*AI Act*.

Dal 2 febbraio 2025 trovano applicazione fornitori e *deployer* di sistemi di AI sono tenuti ad attuare fattivamente un obbligo fondamentale: garantire che tutte le persone coinvolte nel funzionamento e nell'uso dei loro sistemi di intelligenza artificiale (IA) abbiano competenze, conoscenze e consapevolezza adeguate attraverso specifici piani e azioni di alfabetizzazione sulle intelligenze artificiali, appunto *AI Literacy*.

Soprattutto nei giovani, ma non solo, vi è l'errata percezione che le intelligenze artificiali — in particolare i principali *large language model*, LLM, diffusi presso il grande pubblico, sviluppati principalmente dalle *big tech* della *silicon valley* Chat GPT di Open AI, Claude di Anthropic, Gemini di Google Deepmind, Grok di X, Llama di Meta, Deepseek dell'omonima tech company cinese — siano davvero “pensanti”: nuovi *oracoli digitali* onniscienti e infallibili a cui è lecito chiedere tutto e attendersi risposte sempre affidabili. Non è così.

Corre l'obbligo di ribadire che le AI attualmente disponibili sul mercato si basano su sofisticati modelli statistici probabilistici con miliardi di pattern: ma non pensano nulla, invero, simulano — con modalità linguistiche accattivanti — il modo di pensare umano.

Le AI che oggi imperversano sul mercato non sono dunque né consapevoli, né intenzionali né tantomeno coscienti: tuttavia — pur senza comprendere agiscono con grandi rischi operativi e responsabilità per produttori e utilizzatori.

Anche per superare tali *false illusioni commerciali* ed erronei convincimenti occorre non solo formare adeguatamente le persone all'utilizzo corretto e consapevole ma soprattutto coltivare, è il caso di dire con sapienza, un nuovo *umanesimo digitale* a servizio e promozione della libertà, solidarietà e dignità personale come ci insegna la nostra Costituzione e la Carta dei diritti fondamentali dell'Unione Europea: ricordiamocelo sempre.

Occorre tutelare, dunque, in ultima istanza, la *dignità inviolabile della persona* nel contesto del mercato asimmetrico digitale

delle intelligenze artificiali aderendo pienamente a una lettura sistematica funzionale protettiva rafforzata, in conformità al metodo assiologico costituzionalmente orientato nel sistema italo-europeo delle fonti particolarmente fecondo nello specifico contesto di cui trattasi.

Rifuggendo dalla *deregulation* federale statunitense ma senza incorrere nei rischi opposti della soffocante e controproducente iper-regolazione europea ridondante e priva di coordinamento tra molteplici referenti normativi.

Non sfugge allora che la formazione programmatica obbligatoria pur essenziale non sia sufficiente a colmare il divario di alfabetizzazione digitale, *bias* cognitivi, vulnerabilità, non solo a causa della diffusione delle intelligenze artificiali.

Debolezze e vulnerabilità della persona sono strutturalmente ineliminabili nei mercati digitali *a fortiori* se progressivamente “agiti” dalle pervasive — non sempre percepibili — intelligenze artificiali, dominati da *algoritmi* opachi e manipolativi: vuote *imago sine re* del pensiero umano.

Occorre controbilanciare la diffusione delle AI quali inconsapevoli, autonomi e imprevedibili agenti: illusori e ingannevoli simulacri di pensiero, vere e proprie *imago sine re* dell'intelligenza naturale.

Per tali ragioni protettive non si può, invero, non constatare che le vicende contrattuali del contesto asimmetrico digitale e delle AI sono sempre meno affare esclusivo dei privati e sempre più rilevanti dal punto di vista pubblico e dell'interesse generale alla correttezza e trasparenza dei mercati.

Il giusto equilibrio, la giusta regola, il giusto rimedio, devono indirizzare la lettura sistematica e assiologica costituzionalmente orientata dell'interprete in ossequio ai generali principi di ragionevolezza, trasparenza, proporzionalità e solidarietà²⁶.

Si registra, pertanto, in tale prospettiva funzionale protettiva, la progressiva emersione della *conformazione contrattuale ex lege* dei contratti digitali e più in generale *la conformazione eterore-*

²⁶ L'espressione *giusto rimedio*, frutto dell'interpretazione sistematica e funzionale, in applicazione del metodo assiologico costituzionalmente orientato, che segna il superamento dell'interpretazione esegetica rigidamente formale e letterale, si deve a P. PERLINCHIERI, *Il « giusto rimedio » nel diritto civile*, in *Il giusto proc. civ.*, 2011, p. 1 ss. Recentemente si vedano S. POLIDORI, *Ragionevolezza, proporzionalità e «giusto rimedio»: le tendenze evolutive e un'occasione mancata (dalla Cassazione)*, in G. Perlinchieri e A. Fachechi (a cura di), *Ragionevolezza e proporzionalità nel diritto contemporaneo* (due tomi), Napoli, 2017; F. DELLA

ROCCA, *Il contratto tra tecnologia e diritti fondamentali*, in *Tecn. dir.*, 2022, p. 272 ss., il quale rileva, da un lato, che « un controllo postumo di ragionevolezza giudiziale degli effetti algoritmici rispetto al caso concreto dovrebbe essere sempre possibile, a garanzia della giustizia del contratto e dei valori costituzionali presenti », dall'altro lato, auspica « il superamento del dato lessicale [già] nei classici contratti per garantire una ragionevole applicazione della disciplina e del giusto rimedio »; e da ultimo F. LONGOBUCCO, *Smart contract e 'giusto rimedio civile' del re-coding tra Rule by Design e Rule of Law*, in *Giornate della*

golamentare dei rapporti tecnologici privatistici: via necessaria ed elettiva per perseguire l'obiettivo meritevole della protezione rafforzata della persona, non solo dei dati personali, e in ultima istanza della dignità inviolabile della persona stessa.

Precetti normativi conformativi inderogabili quali rimedi normativi elettivi per fronteggiare adeguatamente — a fronte del crepuscolo del consenso individuale — la vertigine asimmetrica dell'artificio digitale: funzionali a proteggere la persona, a colmare non solo le *debolezze contrattuali* ma anche e soprattutto le nuove insidiose e sfuggenti *vulnerabilità digitali*.

Abstract

L'AI Act si propone di migliorare il funzionamento del mercato istituendo un quadro giuridico uniforme finalizzato a disciplinare un *uso responsabile* di sistemi di intelligenza artificiale, in conformità ai valori dell'Unione Europea e a promuovere la diffusione di un'intelligenza artificiale *antropocentrica, sicura, trasparente e affidabile*. Il presente saggio intende formulare prime riflessioni civilistiche relativamente al tale recente epocale intervento normativo europeo in prospettiva protettiva dei diritti fondamentali della persona. In tale contesto asimmetrico digitale in continua evoluzione e segnato da nuove vulnerabilità digitali e dal crepuscolo del consenso, essenziale risulta lumeggiare il necessario dialogo normativo con il GDPR al fine di assicurare un'effettiva protezione della persona e dei dati personali dalla bulimia di dati necessaria per addestrare e ottimizzare le AI. In tale prospettiva ricostruttiva si evidenziano anche i principi ordinatori dell'AI Act. Lettura dottrinale sistematica operata — in prospettiva assiologica costituzionalmente orientata — avendo sempre come stella polare la tutela inviolabile della dignità della persona e dei diritti fondamentali.

The AI Act aims to improve the functioning of the market by establishing a uniform legal framework to regulate the responsible use of artificial intelligence systems, in accordance with the values of the European Union, and to promote the spread of human-centric, secure, transparent, and reliable artificial intelligence. This essay aims to offer initial civil law considerations regarding this recent landmark European regulatory intervention from the perspective of protecting fundamental human rights. In this ever-evolving asymmetric digital context, marked by new digital vulnerabilities and by the twilight of consent, it is essential

to highlight the necessary regulatory dialogue with the GDPR to ensure effective protection of individuals and their personal data from the data overload required to train and optimize AI. From this reconstructive perspective, the key principles of the AI Act are also highlighted. This systematic doctrinal analysis — through a constitutionally oriented axiological perspective — is conducted with the inviolable protection of human dignity and fundamental rights as its guiding principle.