ICONOCRAZIA
RIVISTA DI SCIENZE SOCIALI E FILOSOFIA POLITICA

SAGGIO

Lo svelamento del simulacro: la prospettiva dell'AI Disclosure tra evoluzione normativa e trasparenza tecnologica

ANDREI MIHAI POP

Università degli Studi di Milano

Abstract

Il contributo analizza l'emergere del diritto alla cosiddetta *AI Disclosure*, inteso come l'autonomo diritto soggettivo dell'individuo di essere informato circa la natura artificiale del proprio interlocutore. Partendo dal superamento del *test* di Turing e dai rischi psicologici legati all'*Uncanny Valley*, lo studio evidenzia come l'incapacità di distinguere tra umano e macchina possa ledere la dignità e l'autodeterminazione della persona. L'indagine esamina l'evoluzione normativa europea che trova le sue radici nel *GDPR* e il suo compimento nell'*AI Act*. In particolare, viene approfondito l'articolo 50 dell'*AI Act*, il quale introduce specifici obblighi di trasparenza per *provider* e *deployer*, con particolare riferimento a *chatbot* e *deepfake*. Infine, l'analisi si sposta sul piano nazionale italiano, valutando la legge n. 132/2025 che ha recepito i princìpi di conoscibilità e trasparenza, consolidando il rango quasi-costituzionale di tale tutela informativa.

Parole chiave: *AI Disclosure*, Trasparenza, Autodeterminazione, *AI Act*, *Deepfake*

English version

The paper examines the emergence of the right to AI Disclosure, defined as an individual's right to be informed when interacting with artificial entities. Beginning with the challenges posed by the Turing Test and the psychological implications of the Uncanny Valley, the study highlights how the inability to distinguish between humans and machines may undermine human dignity and self-determination. The research explores the European regulatory framework, tracing its origins from the GDPR to the definitive implementation of the AI Act. Specifically, it analyzes Article 50 of the AI Act, which imposes transparency obligations on providers and deployers regarding chatbots, emotion recognition, and deepfakes. Finally, the paper discusses the Italian legal landscape, focusing on Law No. 132/2025, which enshrines the principles of knowability and transparency, effectively integrating AI Disclosure into the hierarchy of fundamental rights.

Keywords: AI Disclosure, Transparency, Self-determination, AI Act, Deepfake

Introduzione

L'Uomo si è confrontato, fin dalle origini, con esseri animati e non – basilarmente con gli 'oggetti' –, dei quali aveva la percezione e la capacità di distinzione e riconoscimento. Fino a qualche anno fa nessuno avrebbe posto in discussione la capacità di capire se qualcuno si stesse relazionando con un essere umano o con una macchina, ma attualmente tale certezza non è più così solida dinanzi allo sviluppo dell'Intelligenza Artificiale. L'incapacità di discernere tra un proprio simile e una macchina intelligente comporterebbe (in potenza) un netto ribaltamento degli equilibri, costringendo l'Uomo a sentirsi esposto e vulnerabile stante la sua non più indiscussa superiorità intellettuale[1]. Quest'ultimo aspetto è già evidente per determinati settori, sebbene la capacità di distinguere tra reale e artificiale rimanga una prerogativa fondamentale nella nostra psicologia.

Il rapporto di reciproca riconoscibilità tra umano e IA è uno degli argomenti che ha attirato l'attenzione degli esperti fin da Alan Turing. Quest'ultimo, in forza del suo *imitation game*, ha posto una rilevante linea di demarcazione oltre la quale il rapporto Uomo-mondo risulterebbe modificato irrimediabilmente. L'Autore si è fermato al risultato potenziale del suo esperimento, ossia che esso avrebbe potuto essere considerato 'superato' laddove l'IA fosse riuscita a mascherare la sua natura confondendo il proprio interlocutore umano per un periodo sufficientemente lungo. Ciò nonostante, dopo più di 75 anni non vi sono sistemi di IA capaci di celare la propria 'identità' per lunghi periodi di tempo senza fare affidamento su particolari circostanze: distrazione umana; un ambito di applicazione ristretto; particolari condizioni di partenza. Tuttavia, anche senza una esplicita affermazione di Turing, appare evidente che tale risultato appaia comunque foriero di enormi conseguenze: ovverosia, stante l'utilizzo dell'IA (almeno di solito) in contesti differenti rispetto a quello in cui si svolge il noto *test*[2], le persone nella vita quotidiana sono ben più esposte a cadere nel 'tranello'. Infatti, l'attenzione di un utente medio durante una

[1] La sfida è storica, poiché mai l'umanità ha dovuto prendere coscienza dell'esistenza di entità (artificiali in questo caso) che possano pareggiare, se non addirittura superare, le proprie capacità. Purtuttavia, ciò non deve far insorgere un atteggiamento antitecnologico, essendo gli scenari catastrofici (in cui le macchine si ribellano all'Uomo – ad esempio come in pellicole quali *Terminator* o *2001: Odissea nello Spazio*) più fantasia che realtà.

[2] Si intende un contesto protetto in cui tutti i partecipanti sono consapevoli di 'giocare'.

attività quotidiana – in cui non si può pretendere di *default* una diligenza particolarmente elevata – lo espone a rischi rilevanti.

Su tale questione, uno degli ambiti più analizzati è quello delle comunicazioni automatizzate di breve durata; a titolo esemplificativo, si pensi agli assistenti vocali utilizzati in alcune – ad oggi – attività quotidiane. Inoltre, appare importante rilevare che il loro vertiginoso 'miglioramento' non potrà che incrementare il pericolo di utenti ignari di interagire e interloquire con 'non umani'. Si pensi alla diffusione di *bot* nel contesto dei *social media*, all'interno dei quali un numero considerevole di persone trascorre una rilevante percentuale della propria esistenza. Se tale fonte di informazione risultasse contaminata da *bot* strutturalmente[3] incubatori di notizie inesatte/false (c.d. *bias effect*), vi potrebbero essere conseguenze rilevanti. Peraltro, si consideri l'ipotesi di un assistente virtuale (dotato di IA) che consigli un certo investimento rispetto a un altro; oppure la tecnica del c.d. *deepfake*, ossia l'ipotesi di immagini, video e audio manipolati dall'IA tali per cui, in assenza di una effettiva *disclosure*, l'occhio umano molto probabilmente non distinguerebbe il vero/reale dal falso/artificiale (rischio che non potrà che aumentare con il progressivo sviluppo tecnologico)[4].

Recentemente, quest'ultimo fenomeno ha attirato l'attenzione anche del Legislatore italiano; invero, è stato introdotto l'articolo 612-*quater* c.p. che andrebbe a punire proprio quelle alterazioni della realtà che abbiano cagionato un nocumento (ingiusto) alla vittima. Questi scenari, pur solo sinteticamente tratteggiati, comporterebbero incredibili vantaggi per chi ne traesse beneficio in

---

[3] Va evidenziato che, tralasciando ipotesi di immissioni dolose, vi potrebbe essere una ragione strutturale per la circolazione di dati inesatti/falsi. Dopotutto, il carburante dell'IA sono i dati che si sono stratificati nel corso di decenni di attività da parte degli esseri umani. Questi ultimi hanno, per esempio, trasposto alcuni propri *bias* all'interno di dati solo apparentemente oggettivi. Dunque, in tale ultima ipotesi, l'inserimento di simili informazioni in un sistema di IA non può essere ascritto al comportamento (doloso o colposo) di una persona stante l'impossibilità di pretendere un controllo assoluto su ogni singolo dato prodotto dall'umanità. Ciononostante, l'impossibilità di pretendere un comportamento alternativo non esime, secondo una certa tesi, dalla responsabilità laddove l'IA abbia causato dei danni – c.d. "responsabilità oggettiva".

[4] Si pensi alla potenzialità diffamatoria/ricattatoria della diffusione di video o immagini manipolate (ad esempio il *deepfake porn*). Su tale tematica offre precisazioni e delucidazioni anche il 'considerando' numero 134 dell'*AI Act*, il quale impone che si apponga un marchio distintivo – c.d. *watermark* – che renda chiara la natura *fake*-artificiale del contenuto generato. Si evidenzia che tale normativa entrerà effettivamente in vigore solamente nel 2026, ma la Commissione Europea ha già pubblicato il 17/12/2025 una prima bozza di un Codice di Condotta relativo all'etichettatura dei prodotti-servizi generati attraverso l'IA. Si veda il suddetto documento al *link*: https://digital-strategy.ec.europa.eu/en/library/first-draft-code-practice-transparency-ai-generated-content.

caso di loro commercializzazione (ipotesi che già è realtà). Detto altrimenti, risulta evidente il rapporto di proporzionalità, ossia: più cresce la verosomiglianza di un sistema di IA rispetto a un Uomo e maggiore sarà la fiducia verso quest'ultimo. Dunque, il vantaggio di utilizzare l'IA rispetto all'Intelligenza Umana (c.d. IU) apparirebbe ulteriormente accentuato: meno costi di gestione del personale, una pressoché infallibilità e una capacità operativa superiore (basilarmente senza necessità di riposo). Purtuttavia, è altresì noto che tale fiducia e gradimento rispetto al proprio interlocutore calano bruscamente laddove si riveli la sua natura artificiale. Questo rapporto è stato teorizzato, e poi osservato, già nel 1970 con il c.d. *Uncanny Valley.* Quindi, già da queste brevi riflessioni, si può comprendere come l'interesse per una *AI disclosure* – in sintesi, l'essere consapevoli di interagire con una entità artificiale – non sia generalizzato, bensì discusso e mutevole in base al centro di interesse di osservazione. Pertanto, i diritti delle persone al cospetto delle forze del mercato richiedono una tutela che solamente lo *ius* potrebbe garantire, ovviamente a condizione che tale strumento venga adoperato nei modi e nei termini corretti. In altre parole, risulta essenziale – a prescindere dal contesto e dallo scopo di utilizzo – che venga enucleato e riconosciuto il diritto a essere informati circa la natura artificiale del proprio interlocutore. Infatti, laddove tale requisito non venisse soddisfatto, vi sarebbe il concreto e attuale pericolo di una lesione della capacità di autodeterminazione del singolo individuo e ciò lederebbe altresì la sua dignità umana[5]. Per tale ragione, l'*AI Disclosure* può essere considerata una versione innovativa del c.d. 'consenso informato'.

Il nuovo diritto all'*AI Disclosure*

L'UE[6] ha deciso di farsi carico di tale esigenza esplicata nel primo paragrafo potendosi, ad oggi, considerare il diritto all'*AI Disclosure* come un diritto

[5] Sul punto, si afferma nel *Report of COMEST on robotics ethics* che: «Dignity is inherent to human beings, not to machines or robots. Therefore, robots and humans are not to be confused even if an android robot has the seductive appearance of a human, or if a powerful cognitive robot has learning capacity that exceeds individual human cognition». Si consenta di rimandare al lavoro della World Commission on the Ethics of Scientific Knowledge and Technology, *Report of COMEST on robotics ethics*, 2017, p. 50, reperibile al *link*: https://unesdoc.unesco.org/ark:/48223/pf0000253952.

[6] Purtuttavia, non è solo l'UE che si è interessata a questa tematica. Invero, anche negli Stati Uniti d'America è in discussione l'introduzione di un diritto a un'*AI Disclosure*. Infatti, due membri della camera dei rappresentanti, Ritchie Torres e Daniel S. Goldman, hanno proposto (e il tutto risulta

definitivamente appartenente all'ordinamento giuridico nazionale e sovranazionale. Tale evoluzione europea ha avuto il suo principio nel *GDPR* e il suo punto di arrivo nell'*AI Act*[7]. Il *General Data Protection Regulation*, dal combinato disposto degli articoli 13 comma 2 lettera f), 14 comma 2 lettera g), 15 comma 1 lettera h) e dal 'considerando' numero 71, prevede importanti obblighi informativi, potendosi concludere che già nel 2016 si potevano intravedere le prime tracce del diritto all'*AI Disclosure*. Purtuttavia, il *GDPR* assumeva una prospettiva limitata ai soli processi decisionali automatizzati che attengano a dati personali[8]. Di conseguenza, come si può agevolmente intuire, rimanevano scoperti ambiti in cui, pur non essendoci trattamenti automatizzati, ovvero laddove gli stessi non riguardassero i dati di interesse per il *GDPR*, l'IA poteva arrecare nocumento. A tale *vulnus* ha posto rimedio l'*AI Act* recentemente entrato in vigore. In particolare, il nuovo regolamento europeo, al titolo IV, articolo 50, è intervenuto innovando il panorama giuridico con una normativa di dettaglio (ben sei articolati commi) che si rivolge sia al *provider* sia al *deployer*[9]. Questo approccio risulta conforme alla tecnica

---

ancora in discussione) l'introduzione della seguente norma: *H.R.3831 – AI Disclosure Act of 2023*. L'elemento più interessante è il "come" si dovrebbe adempiere a tale nuovo obbligo, ossia tramite un avviso dal seguente contenuto: «This output has been generated by artificial intelligence». Per la proposta, si veda: https://www.congress.gov/bill/118th-congress/house-bill/3831/text?s=2&r=153.

[7] Doveroso segnalare anche il contributo dell'*European Group on Ethics in Science and New Technologies* (*EGE*), il quale pone proprio l'accento su tale questione: «[…] we may ask whether people have a right to know whether they are dealing with a human being or with an AI artefact». Così l'European Group on Ethics in Science and New Technologies, *Statement on Artificial Intelligence, Robotics and 'Autonomous' Systems*, Bruxelles, 9/3/2018, p. 11, reperibile *online* al seguente indirizzo: https://www.unapcict.org/sites/default/files/2019-01/EC_AI-%20Robotics-%20and%20Autonomous%20Systems.pdf.

[8] Peraltro, vi sono alcune opinioni che negherebbero la compatibilità tra l'IA e il *GDPR* laddove il sistema di IA continuasse a rimanere opaco, discriminatorio e soggetto a *bias*. Tali caratteristiche non sarebbero conciliabili con i princìpi introdotti da tale regolamento europeo in tema di protezione dei dati personali. Condivide i limiti con il *GDPR*, in tema di *AI Disclosure*, anche la *Directive on Automated Decision-Making* canadese. In particolare, il comma 6.2.1 dispone: «Providing notice through all service delivery channels in use that the decision rendered will be undertaken in whole or in part by an Automated Decision System […]». Il testo di tale disposizione è consultabile al *link*: https://www.tbs-sct.canada.ca/pol/doc-eng.aspx?id=32592&section=html. Purtuttavia, si segnala la proposta di riforma in Canada – *Artificial Intelligence and Data Act (AIDA)* – che poteva innovare questo aspetto, così come ha fatto l'*AI Act* in Europa. Purtuttavia, tale progetto si è arrestato anche alla luce delle ultime elezioni canadesi e delle resistenze di vari partiti.

[9] Per una definizione ai fini del presente regolamento di '*deployer*', si veda l'articolo 3, n. 4 dell'*AI Act*, il quale prevede: «'Deployer' means a natural or legal person, public authority, agency or other body using an AI system under its authority except where the AI system is used in the course of a personal non-professional activity». Inoltre, si rileva anche il 'considerando' numero 13 dell'*AI Act*, in cui si afferma – circa il *deployer* – che: «[…] any natural or legal person, including a public authority, agency or other body, using an AI system under its authority, except where the AI system is used in the course of a personal non-professional activity».

dell'*AI Act* che punta a una continua suddivisione di responsabilità (c.d. *distributed responsibility*) tra tali due categorie. Infatti, l'articolo 50[10] dell'*AI Act* – rubricato 'Obblighi di trasparenza per i fornitori e i *deployer* di determinati sistemi di IA' – risulta essere paradigmatico, necessitando di una analisi approfondita. Ovverosia, tale disposizione prevede che:

> «1. I fornitori garantiscono che i sistemi di IA destinati a interagire direttamente con le persone fisiche sono progettati e sviluppati in modo tale che le persone fisiche interessate siano informate del fatto di stare interagendo con un sistema di IA, a meno che ciò non risulti evidente dal punto di vista di una persona fisica ragionevolmente informata, attenta e avveduta, tenendo conto delle circostanze e del contesto di utilizzo. Tale obbligo non si applica ai sistemi di IA autorizzati dalla legge per accertare, prevenire, indagare o perseguire reati, fatte salve le tutele adeguate per i diritti e le libertà dei terzi, a meno che tali sistemi non siano a disposizione del pubblico per segnalare un reato.
> 2. I fornitori di sistemi di IA, compresi i sistemi di IA per finalità generali, che generano contenuti audio, immagini, video o testuali sintetici, garantiscono che gli *output* del sistema di IA siano marcati in un formato leggibile meccanicamente e rilevabili come generati o manipolati artificialmente. I fornitori garantiscono che le loro soluzioni tecniche siano efficaci, interoperabili, solide e affidabili nella misura in cui ciò sia tecnicamente possibile, tenendo conto delle specificità e dei limiti dei vari tipi di contenuti, dei costi di attuazione e dello stato dell'arte generalmente riconosciuto, come eventualmente indicato nelle pertinenti norme tecniche. Tale obbligo non si applica se i sistemi di IA svolgono una funzione di assistenza per l'*editing standard* o non modificano in modo sostanziale i dati di *input* forniti dal *deployer* o la rispettiva semantica, o se autorizzati dalla legge ad accertare, prevenire, indagare o perseguire reati.
> 3. I *deployer* di un sistema di riconoscimento delle emozioni o di un sistema di categorizzazione biometrica informano le persone fisiche che si sono esposte in merito al funzionamento del sistema e trattano i dati personali in conformità dei regolamenti (UE) 2016/679 e (UE) 2018/1725 e della direttiva (UE) 2016/680, a seconda dei casi. Tale obbligo non si applica ai sistemi di IA utilizzati per la categorizzazione biometrica e il riconoscimento delle emozioni autorizzati dalla legge per accertare, prevenire o indagare reati, fatte salve le tutele adeguate per i diritti e le libertà dei terzi e conformemente al diritto dell'Unione.
> 4. I *deployer* di un sistema di IA che genera o manipola immagini o contenuti audio o video che costituiscono un «*deep fake*» rendono noto che il contenuto è stato generato o manipolato artificialmente. Tale obbligo non si applica se l'uso è autorizzato dalla legge per accertare, prevenire, indagare o perseguire reati. Qualora il contenuto faccia parte di un'analoga opera o di un programma manifestamente artistico, creativo, satirico o fittizio, gli obblighi di trasparenza di cui al presente paragrafo si limitano all'obbligo di rivelare l'esistenza di tali contenuti generati o manipolati in modo adeguato, senza ostacolare l'esposizione o il godimento dell'opera. I *deployer* di un sistema di IA che genera o manipola testo pubblicato allo scopo di informare il pubblico su questioni di interesse pubblico rendono noto che il testo è stato generato o

[10] Tale norma, benché probabilmente la più significativa, non è l'unica a richiamare tale principio. Invero, ad esempio, i 'considerando' numero 27 (in cui si sostiene, in un inciso, la necessità di: «[…] making humans aware that they communicate or interact with an AI System […]»), 120, 132, 134 e 136 dell'*AI Act* citano la c.d. *AI Disclosure* e questi richiami evidenziano l'importanza di tale diritto.

> manipolato artificialmente. Tale obbligo non si applica se l'uso è autorizzato dalla legge per accertare, prevenire, indagare o perseguire reati o se il contenuto generato dall'IA è stato sottoposto a un processo di revisione umana o di controllo editoriale e una persona fisica o giuridica detiene la responsabilità editoriale della pubblicazione del contenuto.
> 5. Le informazioni di cui ai paragrafi da 1 a 4 sono fornite alle persone fisiche interessate in maniera chiara e distinguibile al più tardi al momento della prima interazione o esposizione. Le informazioni devono essere conformi ai requisiti di accessibilità applicabili.
> 6. I paragrafi da 1 a 4 lasciano impregiudicati i requisiti e gli obblighi di cui al capo III, così come gli altri obblighi di trasparenza stabiliti dal diritto dell'Unione o nazionale per i *deployer* dei sistemi di IA.
> 7. L'ufficio per l'IA incoraggia e agevola l'elaborazione di codici di buone pratiche a livello dell'Unione per facilitare l'efficace attuazione degli obblighi relativi alla rilevazione e all'etichettatura dei contenuti generati o manipolati artificialmente. La Commissione può adottare atti di esecuzione per approvare tali codici di buone pratiche secondo la procedura di cui all'articolo 56, paragrafo 6. Se ritiene che il codice non sia adeguato, la Commissione può adottare un atto di esecuzione che specifichi norme comuni per l'attuazione di tali obblighi secondo la procedura d'esame di cui all'articolo 98, paragrafo 2».

I primi commi si rivolgono al *provider* (comma 1 e 2) e al *deployer* (comma 3 e 4) imponendo, in sintesi, un obbligo informativo nei confronti dell'utente finale nel momento in cui quest'ultimo dovesse relazionarsi con un sistema di IA[11]. Tuttavia, lo stesso articolo fa salve quelle ipotesi in cui sia già lapalissiano, secondo le capacità di una persona media[12], l'utilizzo dell'Intelligenza Artificiale[13]. Tale dovere informativo è disciplinato – da un punto di vista procedurale – anche al comma 5, in cui si afferma la necessaria chiarezza di tale comunicazione almeno nel momento di primo contatto con l'IA[14]. Infine, il comma 7 dell'articolo 50

[11] Per quanto riguarda il *deployer* (e più ampiamente l'*AI Disclosure*), si noti altresì il 'considerando' numero 93, il quale – riferendosi ai sistemi di IA ad alto rischio – dispone che: «Deployers of high-risk AI systems listed in an annex to this Regulation also play a critical role in informing natural persons and should, when they make decisions or assist in making decisions related to natural persons, where applicable, inform the natural persons that they are subject to the use of the high-risk AI system. This information should include the intended purpose and the type of decisions it makes».

[12] La *natural person* citata dal regolamento appartiene a quelle categorie che la migliore dottrina ha definitivo come "concetti elastici" necessitanti di una continua interpretazione. Sulle clausole generali (o sui concetti elastici) vi è una sterminata manualistica. Per esempio, Alpa G. (2001), *Istituzioni di diritto privato. Nozioni*, Torino: Utet, p. 61.

[13] Inoltre, si escludono dalla *disclosure* quei contatti con i sistemi di IA che siano previsti per legge (ed estensivamente sottoposti concretamente al controllo di una autorità giudiziaria) al fine di prevenire, indagare e accertare reati.

[14] La possibilità di evitare un continuo *reminder* circa la natura artificiale del sistema nel momento in cui l'utente lo utilizzi per un lungo periodo di tempo, senza interruzioni, contempera la necessaria tutela dei diritti dei singoli con la celerità e il buon funzionamento del sistema. Concretamente, tale dovere potrebbe essere adempiuto mediante una semplice comunicazione che compaia sullo schermo dell'utente nel momento in cui si accede, per esempio, a una applicazione che utilizzi l'IA. L'utente, per poter proseguire, dovrebbe solamente "dar per letta" tale informazione senza la formale necessità di una 'accettazione'.

richiama l'*AI Office*, il quale dovrà rivestire un ruolo propulsivo incentivando l'implementazione di codici di condotta che agevolino l'individuazione di contenuti generati con l'IA. Questa disposizione, laddove osservata nella sua complessiva struttura, rappresenta chiaramente l'enucleazione del diritto alla c.d. *AI Disclosure* rientrante, quindi, tra il catalogo dei diritti fondamentali su di un piano europeo e, conseguentemente, anche nel piano sub-costituzionale all'interno dei singoli Stati membri (compresa l'Italia). Tale valore giuridico si deve riconoscere in forza dell'articolo 11 della Costituzione, poiché le fonti dell'UE, come noto, assumono un valore pari alla Carta fondamentale, salvo il rispetto di certi limiti fondamentali e inviolabili nazionali (c.d. 'teoria dei contro-limiti affermata'). Quest'ultima precisazione risulta necessaria, poiché affermando che le fonti unionali hanno quasi lo stesso rango della Costituzione, risulta chiaro che queste prevarranno anche sulle leggi ordinarie dello Stato. Infatti, in caso di contrasto normativo dovrà prevalere il diritto unionale (così già dalla sentenza *Simmenthal* della CGUE). Dunque, il legislatore italiano non potrà emanare una legge che contrasti con il diritto a una *AI Disclosure* sancito all'articolo 50 dell'*AI Act*, ovverosia, e in caso contrario, i giudici potranno disapplicare una simile norma al fine di garantire quella "superiorità" del diritto dell'UE negli ambiti di sua competenza.

L'Italia, allo stato attuale, parrebbe aver adottato una prospettiva rispettosa dei princìpi e delle norme unionali[15]. Invero, il disegno di legge del Governo n. 1146/2024 sull'IA, approvato definitivamente il 17/9/2025 e pubblicato il 25/9/2025 con la legge n. 132/2025, ha introdotto il principio di conoscibilità dell'Intelligenza Artificiale nel momento in cui si entri in contatto con la stessa. Questo impone, sia al settore pubblico sia a quello privato, la necessità di rispettare determinati princìpi giuridici nell'ideazione, sviluppo, sperimentazione e utilizzo di sistemi dotati dell'Intelligenza Artificiale. In primo luogo, l'articolo 3 al comma 3 (rubricato 'Princìpi generali') di tale recente legge dispone che:

> «I sistemi e i modelli d'Intelligenza Artificiale per finalità generali devono essere sviluppati e applicati nel rispetto dell'autonomia e del potere decisionale dell'Uomo, della prevenzione del danno, della conoscibilità, della trasparenza,

[15] A prescindere dalla già assodata applicabilità 'automatica' di un regolamento dell'UE all'interno degli Stati membri, la legge del 2025, all'articolo 1 comma 2, afferma esplicitamente che qualsiasi disposizione contenuta nel testo dovrà essere interpretata conformemente al diritto unionale (una c.d. "interpretazione unionalmente orientata").

> della spiegabilità e dei princìpi di cui al comma 1, assicurando la sorveglianza e l'intervento umano».

In secondo luogo, si ribadisce tale principio all'articolo 4 comma 3 – dedicato ai 'Princìpi in materia di informazione e di riservatezza dei dati personali' – in cui si prevede che:

> «Le informazioni e le comunicazioni relative al trattamento dei dati connesse all'utilizzo di sistemi d'Intelligenza Artificiale sono rese con linguaggio chiaro e semplice, in modo da garantire all'utente la conoscibilità dei relativi rischi e il diritto di opporsi ai trattamenti autorizzati dei propri dati personali».

In definitiva, alla luce dell'*AI Act* e dell'intervento nazionale italiano, parrebbe fondato ritenere l'esistenza pacifica all'interno del tessuto giuridico del diritto all'*AI Disclosure*. Per tali ragioni, il suo rispetto sarà imposto sia ai privati sia alle autorità pubbliche nel momento in cui utilizzano sistemi di IA.

Conclusioni

L'indagine condotta permette di affermare che l'emergere del diritto alla cosiddetta *AI Disclosure* non rappresenti meramente un nuovo tassello nel mosaico degli obblighi informativi, bensì costituisca il presidio fondamentale per la sopravvivenza del patto di fiducia tra consociati e istituzioni nell'era della simulazione totale. Come si è avuto modo di analizzare, la capacità della macchina di abitare la c.d. *Uncanny Valley* con una verosimiglianza sempre più raffinata non è un fenomeno neutro: essa incide direttamente sulla dignità della persona e sulla sua capacità di autodeterminazione.

Il percorso normativo europeo, culminato nell'*AI Act*, e il recepimento nazionale italiano tramite la legge n. 132/2025 segnano una linea di demarcazione netta. L'articolo 50 dell'*AI Act*, imponendo obblighi di trasparenza per la generazione di contenuti sintetici e l'interazione con sistemi di IA, riconosce implicitamente che la 'verità' della fonte è un presupposto essenziale per l'esercizio delle libertà democratiche. In un contesto comunicativo dominato da immagini e discorsi, la mancata segnalazione dell'origine artificiale di un contenuto non è solo un'omissione tecnica, ma una forma di inquinamento della sfera pubblica che altera il processo di formazione del consenso.

L'introduzione del principio di conoscibilità nell'ordinamento italiano (ex articolo 3, legge n. 132/2025) rafforza questa visione, elevando la trasparenza algoritmica a parametro di legittimità per lo sviluppo tecnologico, sia nel settore pubblico che in quello privato. Tale scelta legislativa sottrae la tutela dell'individuo alle sole logiche del mercato, ancorandola ai princìpi costituzionali di cui all'articolo 11, attraverso il dialogo costante con le fonti unionali.

Tuttavia, la sfida che si profila all'orizzonte non è esclusivamente normativa, ma culturale e semiotica. Se l'iconocrazia contemporanea si fonda sul potere dell'immagine di farsi realtà, l'*AI Disclosure* funge da dispositivo di 'svelamento'. Il diritto di (ri)conoscere l'artificiale diventa, dunque, il diritto a non essere ingannati da un simulacro che, pur superando il *test* di Turing nella quotidianità delle relazioni digitali, rimane ontologicamente privo di responsabilità e coscienza.

In conclusione, il diritto all'*AI Disclosure* si configura come una nuova declinazione del consenso informato. Esso non serve a limitare il progresso tecnologico, ma a garantire che l'essere umano rimanga il centro di gravità dei processi decisionali e relazionali. Solo attraverso una rigorosa applicazione degli obblighi di trasparenza sarà possibile evitare che la 'valle perturbante' si trasformi in un abisso in cui l'individuo perde la bussola della propria realtà, vedendo compromessa la propria autonomia intellettuale davanti all'IA. La sfida per i giuristi e per i teorici della società sarà ora quella di vigilare affinché queste norme non restino mere enunciazioni di principio, ma diventino strumenti effettivi di una nuova ecologia della verità digitale.

Bibliografia

Alpa G. (2001). *Istituzioni di diritto privato. Nozioni*, Torino: Utet.

Bianca M. C., Patti S. & Patti G. (2001). *Lessico di Diritto Civile*, Milano: Giuffré.

Dathathri S. *et al*. (2024). Scalable watermarking for identifying large language model outputs, *Nature*, 634, pp. 818-823.

Donati F. (2020). Intelligenza Artificiale e giustizia, *AIC*, 1, pp. 424-425.

Geller T. (2008). Overcoming the Uncanny Valley, *IEEE Computer Graphics and Applications*, XXVIII (4), pp. 11-17.

Łabuz M. (2024). Deep fakes and the Artificial Intelligence Act – An important signal or a missed opportunity?, *Policy & Internet*, pp. 1-18.

Marasà F. (2023). Intelligenza Artificiale e tutela dei dati personali. Quali riflessi sulla giustizia predittiva?, *Osservatorio del diritto civile e commerciale*, 1, pp. 100-109.

McIntire J. P., McIntire L. K. & Havig P. R. (2010). Methods for chatbot detection in distributed text-based communications, *International Symposium on Collaborative Technologies and Systems*, pp. 463-472.

Mori M. (1970). Bukimi No Tani, *Energy*, 4.

Pollicino O. & Amitrano D. (2023). AI e responsabilità civile, ecco le regole per non avere guai, in *Agenda Digitale*, 12/12/2023, consultato il 30/12/2025 (https://www.agendadigitale.eu/cultura-digitale/ai-e-responsabilita-civile-ecco-le-regole-per-non-avere-guai/).

Simoncini A. (2019). L'algoritmo incostituzionale: l'Intelligenza Artificiale e il futuro delle libertà, *Rivista di Biodiritto*, pp. 77-79.