



**SULL'APPROCCIO UMANO-CENTRICO
ALL'INTELLIGENZA ARTIFICIALE.
RIFLESSIONI A MARGINE DEL "PROGETTO EUROPEO DI ORIENTAMENTI
ETICI PER UNA IA AFFIDABILE"**

ARIANNA ALPINI

SOMMARIO: 1. 'Ordine umano-centrico' e tirannia dell'algoritmo. – 2. Rischi tecnologici e paradigma conoscitivo. – 3. Il diritto di non essere sottoposto ad una decisione *totalmente* automatizzata. Limiti e contraddizioni. Il diritto alla spiegazione. – 4. *Segue*. La profilazione. – 5. *Status* giuridico dei *robot*. – 6. Benefici e rischi dell'IA: un bilanciamento da ri-valutare. – 7. L'autonomia dell'agente *software*: necessità di un approccio funzionale compatibile.

1. Il giurista è sempre chiamato a compiere delle scelte: in ciò consiste la sua attività, nel valutare per decidere¹. Tuttavia, vi sono dei momenti storicamente determinati ove queste scelte si rivelano risolutive per la conservazione e l'evoluzione della civiltà. In questi momenti, che segnano indelebilmente la storia dell'uomo, il giurista ha la responsabilità di compiere valutazioni concrete e assumere decisioni ragionevoli, secondo la filosofia che è nel diritto²: non una qualsiasi etica né linee guida, ma valori normativi di civiltà giuridica³.

Il diritto svolge una funzione ordinante e di controllo, finalizzata a preservare e promuovere l'umanità, l'originalità e l'irripetibilità della persona umana, il suo incommensurabile valore racchiuso nella dignità di essere umano. Sì che la scelta del giurista è vincolata concretamente ed effettivamente a tali valori, principi consacrati nell'ordine costituzionale e prima ancora nella storia dell'uomo. Queste considerazioni non sono scontate né possono essere sottovalutate nel momento nel quale ci si affida alla finzione piuttosto che alla realtà. Affidarsi ai dati quantitativi e statistici appare la strada da percorrere per risolvere problemi che però necessitano di una 'intelligenza sensibile'⁴ poiché ogni scelta è destinata a produrre effetti peculiari e spesso imprevedibili sulla persona e sull'ambiente.

¹ L'attività interpretativa è sempre attività di bilanciamento, sul tema per tutti G. PERLINGIERI, *Profili applicativi della ragionevolezza nel diritto civile*, Napoli, 2015, *passim*.

² P. PERLINGIERI, *Il diritto civile nella legalità costituzionale secondo il sistema italo-comunitario delle fonti*, Napoli, 3^a ed., 2006, p. 6 ss., il quale, in particolare, precisa che «sotto il profilo metodologico non è possibile costruire una soggettività, in termini strutturalistici, in forma unitaria, ma è d'obbligo intenderla in forma plurima, distinguendo i problemi della persona umana dai problemi di quei soggetti (sia pure diversificati per scopi e funzioni), che persone umane non sono: le cc.dd. persone giuridiche e ogni altro centro di imputazione soggettiva».

³ P. PERLINGIERI, *Applicazione e controllo nell'interpretazione giuridica*, in *Riv. dir. civ.*, 2010, p. 332 ss.

⁴ Ossia non soltanto intelletto e razionalità ma anche *sentimento*. Vedi le interessanti riflessioni di E. BONCINELLI, *L'anima della tecnica*, Rizzoli, Milano, 2006, *passim*.



Principi etici invocati accanto ai diritti fondamentali propongono un ordine umano-centrico⁵ quale garanzia di una intelligenza artificiale affidabile (*Trustworthy AI*) e, dunque, competitiva; sembra che l'obiettivo sia garantire più che la sovranità della persona la conquista dei mercati. L'uomo, tuttavia, è destinato non a galleggiare al di sopra della storia⁶, ma a prendersi cura del suo inestimabile valore. È ingenuo per il giurista inseguire lo sviluppo tecnologico pensando che è altamente improbabile e, di certo, non imminente la tirannia dell'algoritmo sull'uomo. Se così fosse perché mai la Commissione europea⁷ metterebbe in guardia sulle possibili zone grigie per non dire oscure che si prospettano allorché gli algoritmi assumano una complessità tale da prendere decisioni autonome a prescindere dal controllo umano? Che bisogno ci sarebbe di promuovere un ordine umano-centrico se non quello di scongiurare un nuovo sistema popolato da macchine che potrebbero mettere in discussione la primazia della persona? Di fronte alla nuova rivoluzione industriale si apre la nuova questione sociale⁸.

⁵ Nel mese di giugno del 2018 la Commissione Europea ha individuato un gruppo di 52 esperti, provenienti dal mondo accademico, dell'industria e della società civile, con l'obiettivo di supportare l'implementazione della strategia europea sull'intelligenza artificiale. Lo scorso 18 dicembre, attraverso la *European AI Alliance*, è stata pubblicata la bozza conclusiva delle linee guida etiche "Progetto di orientamenti etici per un IA affidabile" (*Draft Ethics Guidelines for Trustworthy AI*) da utilizzare in Europa per la progettazione, lo sviluppo e l'uso di applicazioni di intelligenza artificiale. Tale bozza è stata aperta per un mese a commenti e suggerimenti volti al suo miglioramento, in modo da emettere la versione definitiva dopo la consultazione dei portatori di interessi tramite l'Alleanza europea per l'IA. Le linee guida etiche individuano cinque principi di base che devono sorreggere lo sviluppo e la diffusione dell'intelligenza artificiale. Ai tradizionali quattro principi della bioetica, ovvero il fare del bene (*beneficence*), il non fare del male (*non-maleficence*), l'autonomia (*autonomy*) e la giustizia (*justice*), se ne aggiunge uno specifico, quello della «splicabilità» (*explicability*). Quest'ultimo è stato inserito poiché l'intelligenza artificiale agisce come agente esterno e quindi è necessario che essa operi in modo trasparente, sia nello spiegarne il funzionamento (principio di intellegibilità), sia nel chiarire chi è responsabile del suo funzionamento, il che include anche la comprensibilità del suo risultato. Il documento non si prefigge solo lo scopo di elencare i diritti fondamentali e i corrispondenti principi e valori alla base di un approccio etico, come fa nel primo capitolo, ma quello di fornire linee guida per lo sviluppo e l'operatività delle applicazioni di intelligenza artificiale, tramite la definizione di un *framework* completo. Pertanto, al primo capitolo ne segue un secondo che, dopo aver individuato i dieci requisiti necessari per rendere applicativi i principi etici di base, propone metodi sia tecnici sia non tecnici a supporto dello sviluppo e dell'implementazione di applicazioni di intelligenza artificiale affidabili. I dieci requisiti sono la responsabilità (*accountability*), la *data governance*, la progettazione inclusiva (*design for all*), aperta ad ogni categoria sociale, che include i requisiti di usabilità ed accessibilità, il governo dell'autonomia dell'intelligenza artificiale, la non discriminazione, il rispetto (ed il potenziamento) dell'autonomia umana, il rispetto della *privacy*, la robustezza, la sicurezza e la trasparenza. Il terzo capitolo, infine, propone una lista di controllo, in forma di un *set* di domande o di indicazioni per ciascuno dei dieci requisiti di cui sopra, a supporto della verifica di operatività delle applicazioni di intelligenza artificiale.

⁶ P. GROSSI, *La odierna 'incertezza' del diritto*, in *Giust. civ.*, 2014, 4, p. 921 ss.

⁷ Vedi 5.5 *Potential longer-term concerns*, Progetto di orientamenti etici, cit. Vedi altresì il punto 10 della "Risoluzione del Parlamento europeo del 16 febbraio 2017 recante raccomandazioni alla Commissione concernenti norme di diritto civile sulla robotica", che afferma: «le possibilità di realizzazione personale che derivano dall'uso della robotica sono relativizzate da un insieme di tensioni o rischi e dovrebbero essere valutate in modo serio dal punto di vista della sicurezza delle persone e della loro salute, della libertà, la vita privata, l'integrità, la dignità, dell'autodeterminazione e la non discriminazione nonché della protezione dei dati personali».

⁸ E. CATERINI, *Sostenibilità e ordinamento civile. Per una riproposizione della questione sociale*, Napoli, 2018, *passim*. Cfr. A. SANTOSUOSSO, C. BOSCARATO e F. CAROLEO, *Robot e diritto: una prima ricognizione*, in *Nuova giur. civ. comm.*, 7-8/2012, p. 495.



2. La questione è non soltanto valutare l'adeguatezza delle categorie del diritto civile alle nuove esigenze del mondo virtuale e della intelligenza artificiale⁹, ma anche riscoprire la funzione del diritto e il compito del giurista nella nuova era. L'antica tradizione della cultura giuridica alla quale il giurista contemporaneo ha il dovere di ispirarsi indica nella prudenza la virtù principale, non soltanto come sapienza teoretica ma anche quale elemento fondamentale nella formazione in senso umanistico della persona, scienza che proviene da una attitudine spirituale che trascende i limiti della conoscenza astratta delle norme e che si realizza concretamente nella ricerca del bene e della giustizia¹⁰. Nel momento stesso nel quale i rischi derivanti dalla deriva tecnologica si prospettano in tutta evidenza nei documenti redatti da esperti di diverse discipline, tali rischi devono assumersi come presupposto della riflessione recuperando nel contempo «ciò che nel passato si è conservato come altra possibilità che non ha ancora avuto futuro»¹¹. La filosofia socratica-platonica non rinuncia all'idea che un complesso di regole inattaccabili ha sempre almeno due limiti nascenti dal mondo della vita: l'inventività della vita interiore e la ricerca della verità. Il conflitto generato dalla scelta di separare le regole dalla vita reale, ha finito invece per contrapporre il calcolo alla coscienza, l'arbitrio alla libera ricerca¹². Rinunciando alla discussione sui valori, l'attenzione si è concentrata sui fatti quale unico oggetto possibile di razionale conoscenza. Qui si colloca la genesi di quel processo che ha segnato la storia del pensiero e che si è concretamente realizzato nell'età moderna.

3. Il diritto di non essere sottoposto ad una decisione totalmente automatizzata¹³ non ha di per sé significato concreto tanto che la prospettiva deve essere rovesciata nel momento nel quale per trovarne un qualche senso applicativo si pone attenzione sul corrispondente divieto e sulla sua estensione. Si passa così a discutere sulle parole: decisione totalmente automatizzata nell'ambito della quale non è presente affatto l'intervento umano o decisione nella quale l'intervento umano è minimo? Non sembra che lo spazio dedicato all'intervento umano possa essere misurato né valutato *a priori*, semmai è la *ratio* del singolo intervento a suggerirne l'estensione.

Affinché l'uomo abbia la possibilità di sottrarsi ad una decisione basata unicamente sul trattamento automatizzato, «la decisione deve produrre effetti giuridici che lo riguarda-

⁹ Il problema non può essere affrontato partendo dalle lacune della disciplina della responsabilità civile. Si tratta, invece, di compiere delle scelte di fondo che coinvolgono anche le generazioni future. Vedi E. CATERINI, *op. cit.*

¹⁰ P. PERLINGIERI, *Ricerca, metodologia, valutazione*, in *Rass. dir. civ.*, 2018, p. 1 ss., il quale precisa che: «La ricerca è, quindi, anche la sua storia, dalla quale, in larga misura, dipende il suo futuro». Vedi altresì G. LIMONE, *Ars boni et aequi. Ovvero i paralipòmeni della scienza giuridica*, in G. LIMONE (a cura di), *Ars boni et aequi. Il diritto fra scienza, arte, equità e tecnica*. L'era di Antigone. Quaderni del Dipartimento di Giurisprudenza della Seconda Università degli Studi di Napoli, FrancoAngeli, n. 9, Milano, 2016, p. 10 ss.

¹¹ G. LIMONE, *La macchina delle regole, la verità della vita. Appunti sul fondamentalismo macchinico nell'era contemporanea*, in G. LIMONE (a cura di), *La macchina delle regole, la verità della vita*. L'era di Antigone. Quaderni del Dipartimento di Giurisprudenza della Seconda Università degli Studi di Napoli, n. 8., cit., 2015, pp. 9-61.

¹² G. LIMONE, *La macchina delle regole*, cit.

¹³ Reg. UE n. 679/2016, art. 22. Vedi considerando 71.



no o deve incidere in modo analogo significativamente sulla sua persona» (*similarly significant effects*). Verrebbe da chiedersi quale potrebbe essere un effetto derivante dal trattamento dei dati il quale, incidendo sulla persona, non sia giuridicamente rilevante. Il divieto del legislatore europeo non soltanto non è chiaro ma prevede delle eccezioni: le decisioni basate unicamente su trattamenti automatizzati sono consentite «se la decisione è necessaria per l'esecuzione o la stipula di un contratto tra l'interessato e il titolare del trattamento; se la decisione si basa sul consenso esplicito dell'interessato; se la decisione è autorizzata dal diritto dell'Unione o dello Stato membro al quale è soggetto il titolare del trattamento»¹⁴. Nei primi due casi «il titolare del trattamento attua misure appropriate per tutelare i diritti, le libertà e i legittimi interessi dell'interessato, almeno il diritto di ottenere l'intervento umano da parte del titolare del trattamento, di esprimere la propria opinione e di contestare la decisione»¹⁵. Nell'ultimo caso «misure adeguate a tutela dei diritti, delle libertà e dei legittimi interessi dell'interessato saranno stabiliti dal legislatore». L'interessato ha diritto di ricevere informazioni significative sulla logica utilizzata, nonché sull'importanza e le conseguenze previste da tale trattamento per l'interessato.

Secondo parte della dottrina tali misure implicherebbero in capo all'interessato un 'diritto alla spiegazione' autonomo e ulteriore rispetto ai diritti di informazione e accesso¹⁶. Come dire che l'informazione di per sé non garantisce alcunché. Tuttavia, le Linee guida del Comitato europeo¹⁷, pur sottolineando la necessità di spiegare e motivare la decisione automatizzata, non forniscono una risposta sulle modalità per mezzo delle quali si dovrebbe realizzare tale diritto, atteso che si tratterebbe di spiegare il funzionamento di un algoritmo. Il nodo della questione risiederebbe nella cd. opacità degli algoritmi¹⁸: opacità con riferimento ora al segreto industriale e al vantaggio concorrenziale, ora all'alta specializzazione tecnica necessaria a decifrare l'algoritmo, ora al processo di autoapprendimento e implementazione dell'algoritmo stesso. In realtà ispezionare il codice sorgente, ammesso che se ne abbiano le competenze, non renderebbe comunque trasparente la motivazione della decisione, poiché sarebbe in grado soltanto di esporre il metodo utilizzato e non anche le ragioni. Per spiegare la *ratio* di una decisione vi deve essere necessariamente un ordito motivazionale, il bilanciamento fra interessi, l'argomentazione, la sintesi, tutti elementi estranei al metodo algoritmico.

4. Il diritto dell'interessato di non essere sottoposto ad una decisione basata unicamente sul trattamento automatizzato comprende la profilazione ossia «qualsiasi forma di trattamento automatizzato di dati personali consistente nell'utilizzo di tali dati per valutare determinati aspetti personali relativi ad una persona fisica, in particolare per analizzare o prevedere aspetti riguardanti il rendimento professionale, la situazione economica, la salute, le preferenze personali, gli interessi, l'affidabilità, il comportamento, l'ubicazione o gli spo-

¹⁴ Art. 22, par. 2, Reg. cit.

¹⁵ Art. 22, par. 3 Reg. cit.

¹⁶ Per una analisi critica dell'attuale dibattito, E. PELLECCIA, *Profilazione e decisioni automatizzate al tempo della black box society: qualità dei dati e leggibilità dell'algoritmo nella cornice della responsible research and innovation*, in *Nuove leggi civ. comm.*, 5/2018, p. 1210 ss.

¹⁷ *Guidelines on Automated individual decision-making and Profiling for purposes of Regulation 2016/679*.

¹⁸ Per una indagine antropologica sul tema vedi M. MAZZOTTI, *Per una sociologia degli algoritmi*, in *Rass. it. sociol.*, 3-4, 2015, p. 465 e ss.



stamenti di detta persona fisica»¹⁹. Valutazioni queste già utilizzate nei sistemi di selezione del personale e nell'erogazione di prestiti da parte di istituti di credito²⁰. Tuttavia tali sistemi si basano sull'ipotesi, per nulla fondata, che i dati raccolti siano neutri e perciò affidabili e che, pure là dove non lo fossero completamente la soluzione sarebbe incrementare il numero dei dati raccolti. Sennonché i dati raccolti non soltanto riflettono gli stessi pregiudizi e le stesse discriminazioni presenti nella società, ma possono anche crearne di nuovi aumentando situazioni di diseguaglianza e discriminazione²¹. La soluzione indicata dal Regolamento europeo sarebbe il diritto alla spiegazione, che ripropone i problemi già indicati come, ad esempio, la circostanza che per fornire una qualche spiegazione sarebbe necessario, almeno, diffondere i dati utilizzati per creare il modello, dati i quali, tuttavia, contengono informazioni sensibili.

5. Se lo sviluppo e l'utilizzo dell'IA devono essere considerati non quale obiettivo in sé, ma quale mezzo per aumentare il benessere umano, allora il principio guida non può che essere la primazia della persona umana²². Di conseguenza lo sviluppo e l'utilizzo dell'intelligenza artificiale dovrebbero svolgersi secondo i canoni di proporzionalità e ragionevolezza, non lasciandosi ispirare da una finalità etica quale mero riflesso dei diritti fondamentali, ma applicando i principi supremi dell'ordine costituzionale.

La persona umana non può essere ridotta alla prospettiva antropocentrica dei dati. Che la filosofia dell'ordinamento vigente basata sul personalismo riconosca alla persona la posizione più alta nella gerarchia dei valori è cosa ben diversa dal collocarla al centro di un sistema di dati e di processi. Se il progettista deve garantire che un robot funzioni in modo conforme ai principi etici e giuridici locali, nazionali e internazionali allora perché, nonostante si riconosca che il *robot* è un prodotto e dunque una cosa, si prospetta l'istituzione di uno *status* giuridico per i *robot*²³ affinché siano considerati come persone elettroniche con il

¹⁹ Artt. 4 e 22 Reg. UE 679/2016. Vedi Considerando 71 del Reg. cit.

²⁰ Profilare la persona è un obbligo secondo le direttive europee Mifid a tutela dell'investitore per valutare l'adeguatezza delle operazioni che gli sono preposte. La profilatura dell'identità digitale della persona, invece, nasconde pericoli più gravi ed insidiosi di acquisti di prodotti finanziari inadeguati rispetto alla sua esperienza finanziaria. Sulla identità digitale e la tutela della persona vedi G. ALPA, *L'identità digitale e la tutela della persona. Spunti di riflessione*, in *Contratto e Impresa*, 2017, 3, p. 723 ss., il quale afferma che non è accettabile e si dovrebbe dunque proibire la cessione volontaria di dati personali particolarmente sensibili.

²¹ E. PELLECCIA, *Profilazione e decisioni automatizzate al tempo della black box society*, cit., p. 1210 ss.

²² Vedi il punto 13 della "Risoluzione del Parlamento europeo del 16 febbraio 2017 recante raccomandazioni alla Commissione concernenti norme di diritto civile sulla robotica": «il quadro etico di orientamento dovrebbe essere basato sui principi di beneficenza, non maleficenza, autonomia e giustizia, nonché sui principi sanciti all'articolo 2 del Trattato sull'Unione europea e nella Carta dei diritti fondamentali dell'Unione europea – quali la dignità umana, l'uguaglianza, la giustizia e l'equità, la non discriminazione, il consenso informato, la vita privata e familiare e la protezione dei dati, così come sugli altri principi e valori alla base del diritto dell'Unione come la non stigmatizzazione, la trasparenza, l'autonomia, la responsabilità individuale e sociale – e sulle pratiche e i codici etici esistenti».

²³ Vedi lett. F) punto 59 della Risoluzione del Parlamento europeo del 16 febbraio 2017, cit. Spunti di riflessione in P.M. ASARO, *Robots and responsibility from a legal perspective*. Proceedings of the IEEE, 2007, in www.peterasaro.org/writing/asaro%20legal%20perspective.pdf, p. 2.; M. LAUKYTE, *Artificial agents among us: should we recognize them as agents proper?*, in *Ethics and Information Technology*, 2017; L. B. SOLUM, *Legal personhood for artificial intelligences*, in *North Carolina Law Review*, 1992. Al riguardo anche G. SARTOR, *Gli agenti software: nuovi sogget-*



conseguente riconoscimento della personalità giuridica? Anche la società giapponese per l'intelligenza artificiale ha formalizzato le linee guida etiche che devono essere applicate ai suoi membri²⁴. L'art. 9 riconosce una soggettività alle creazioni di IA, una autonomia così elevata rispetto ai suoi programmatori tale da rendere necessario che le stesse applicazioni di IA debbano a loro volta rispettare le linee guida, sí che il prodotto della ricerca viene riconosciuto come autonomo membro della società. A tale riconoscimento consegue altresì il riconoscimento in capo al robot di diritti e di doveri non soltanto con riferimento alla legge sul *copyright* ma anche in ambiti particolarmente delicati quali il diritto di famiglia.

6. L'approccio etico dell'Unione europea alla intelligenza artificiale è volto a sollecitare una riflessione etico-umanistica sul progresso tecnologico mondiale²⁵. Tuttavia si tratta di un documento non vincolante e contraddittorio nella sua premessa. La bozza del codice etico parte da un assunto tutt'altro che pacifico, ossia che i benefici dell'IA superano i suoi rischi. Tale affermazione è smentita dallo stesso documento là dove si sofferma sui *software* che consentono il riconoscimento facciale, su quelli che fingono con successo di essere umani, come ad esempio gli assistenti vocali, sugli algoritmi di profilazione con finalità di *scoring*, sui sistemi *lams* ovvero i c.d. *robot killer*. Inserendo tra i requisiti per l'affidabilità dell'IA il rispetto (e miglioramento) dell'autonomia umana, con l'affermazione che i sistemi

ti del ciberdiritto?, in *Contr. impr.*, 2002, vol. II, p. 465 ss.; G. TADDEI ELMI, *I diritti dell'intelligenza artificiale tra soggettività e valore: fantadiritto o jus condendum*, in L. LOMBARDI VALLAURI (a cura di), *Il meritevole di tutela*, Milano, 1990, pp. 685-711; G. TADDEI ELMI, *Logos e intelligenza artificiale*, in L. LOMBARDI VALLAURI (a cura di), *Logos dell'essere. Logos della norma*, Bari, 1999, pp. 603-652; L. LOMBARDI VALLAURI, *Neuroni, mente, anima, algoritmo: quattro ontologie*, in L. LOMBARDI VALLAURI (a cura di), *Logos dell'essere. Logos della norma*, cit., pp. 571-601. Sulla possibilità di creare un nuovo soggetto di diritto, titolare di situazioni giuridiche soggettive attive e passive, potrebbero richiamarsi le argomentazioni relative all'attribuzione della soggettività giuridica e della personalità giuridica agli enti collettivi. Vedi P. PERLINGIERI, *La personalità umana nell'ordinamento giuridico*, Napoli, 1972, *passim*; F. GALGANO, *Le associazioni, le fondazioni, i comitati*, Padova, 1987, *passim*; P. RESCIGNO, *Persone e gruppi sociali*, Napoli, 2006, *passim*. V. SCALISI, *Categorie e istituti del diritto civile nella transizione al postmoderno*, Milano, 2005, *passim*. Occorre tenere presente che la personalità giuridica non richiede necessariamente il riconoscimento di uno *status* giuridico dei *robot*. Forse l'approccio più adeguato è quello che guarda non tanto ai diritti e ai doveri dei robot quanto all'attività prodotta dall'uomo e dagli algoritmi che rappresenta un fatto del tutto nuovo e peculiare.

²⁴ Il Comitato etico della Società Giapponese per l'Intelligenza Artificiale (*Japanese Society for Artificial Intelligence-JSAI*), nel mese di settembre 2017 ha approvato le linee guida sull'intelligenza artificiale che devono essere osservate dai membri di JSAI.

²⁵ Sul binomio uomo-robot vedi: I.R. NOURBAKHSH, *Robot fra noi. Le creature intelligenti che stiamo per costruire*, Torino, 2014, *passim*; R. CINGOLANI e G. METTA, *Umani e umanoidi. Vivere con i robot*, Bologna, 2015, *passim*; L. DE BIASE, *Homo pluralis. Esseri umani nell'era tecnologica*, Torino, 2015, *passim*; B. HENRY, *Dal golem ai cyborgs. Trasmigrazioni nell'immaginario*, Livorno, 2013, *passim*. Più di recente, L. PALAZZANI, *Il potenziamento umano. Tecnoscienza, etica e diritto*, Torino, 2015, *passim*. L'impiego di agenti robotici, proprio in rapporto al contesto prettamente umano, suscita preoccupazioni di matrice etica; la Roboetica, infatti, altro non sarebbe che l'etica applicata alla robotica. Interessanti riflessioni in tal senso, *ex multis*, da parte di P. LIN, K. ABNEY e G.A. BEKEY, *Robot Ethics. The ethical and social implications of robotics*, Cambridge, 2012; R. BOGUE, *Robot ethics and law*, in *Industrial Robot, An International Journal*, 2014; A. SHARKEY, *Robots and human dignity: a consideration of the effects of robot care on the dignity of older people*, in *Ethics and Information Technology*, 2014; G. TADDEI ELMI e F. ROMANO, *Robotica: tra etica e diritto*. Un seminario promosso dal Dipartimento Identità Culturale e dall'ITTIG del CNR, in *Inform. e dir.*, 1-2/2010, pp. 145-152.



di intelligenza artificiale dovrebbero contenere in se stessi le contromisure per evitare possibili abusi molto gravi da essi derivanti; illustrando le criticità e i rischi che generano preoccupazioni nell'utilizzo dei sistemi di intelligenza artificiale, quali le armi letali autonome, i sistemi di punteggio di massa e i sistemi di intelligenza artificiale nascosti, il documento svela le sue contraddizioni. È evidente che l'assunto secondo il quale «i benefici dell'IA superano i suoi rischi», va sottoposto a revisione critica. Appare necessaria una 'impostazione predefinita di salvaguardia dei diritti della persona umana' che di certo trova fondamento non in una Carta *soft* e di autoregolamentazione ma nelle stesse Costituzioni e nei Trattati internazionali.

7. Uno dei principi etici (*Ethical Principles in the Context of AI and Correlating Values*) individuati dalla Commissione europea è quello di autonomia²⁶ inteso come non obbligatorietà dell'uso di sistemi di intelligenza artificiale, che si traduce nel divieto di rendere obbligatorio il ricorso a tali sistemi per lo svolgimento di determinate azioni, compiti o mansioni, al fine di bilanciare l'imposizione dei mercati di *standard* di prestazioni sempre più elevati. A questo si collega il principio di trasparenza²⁷ che, esigendo il pieno controllo e la comprensibilità umana dei processi di IA, sembra di fatto escludere la configurabilità della c.d. intelligenza artificiale forte (*strong AI*). Sì che le *machine learning* che evolvono apprendendo dalla propria esperienza e in modo non sempre prevedibile, o i più complessi sistemi adattivi in grado di modificare il proprio codice autonomamente per adattarlo alle circostanze esterne appaiono in contrasto con tali principi.

Da questa prospettiva la proposta di riconoscere ai *robot* uno *status* giuridico di persone elettroniche dal quale deriverebbe una piena personalità giuridica e dunque persino l'esercizio di diritti fondamentali appare una scelta non compatibile con i valori costituzionalmente garantiti. Se infatti la macchina è servente, ausiliaria alle attività umane, non si può sostenere che l'autonomia del *software* giunga al punto di escludere il controllo umano e di agire nel proprio interesse. Ciò significherebbe accettare tutti i rischi connessi alla imprevedibilità dei comportamenti degli algoritmi poiché si tratterebbe non di errori di calcolo del programmatore ma di decisioni imprevedibili di algoritmi.

Tuttavia non si può non riconoscere all'algoritmo una qualche capacità²⁸. Gli agenti *software*²⁹ altrimenti definiti come attanti³⁰, potrebbero violare obblighi contrattuali previsti

²⁶ *The Principle of Autonomy: "Preserve Human Agency"*, par. 4, p. 9 del Progetto di orientamenti etici, cit.

²⁷ *The Principle of Explicability: "Operate transparently"*, par. 4, p. 10 del Progetto di orientamenti etici, cit. Si veda altresì la nota corrispondente che precisa: «The literature normally speaks of "explainability". The concept of "explicability" to refer both to "intelligibility" and to "explainability" and hence capture the need for transparency and for accountability is introduced in L. Floridi, "AI4People – An Ethical Framework for a Good AI Society: Opportunities, Risks, Principles, and Recommendations", *Minds and Machines*».

²⁸ C. PERLINGIERI, *L'incidenza dell'utilizzazione della tecnologia robotica nei rapporti civilistici*, in *Rass. dir. civ.*, 4/2015, p. 1235 ss.

²⁹ G. SARTOR, *Gli agenti software: nuovi soggetti del ciberdiritto?*, cit., p. 465. Tali agenti agiscono come fossero dei rappresentanti dell'utente, determinandosi con un certo grado di autonomia in ordine alla possibilità di stipula di un contratto con relativi contenuti e condizioni: F. BRAVO, *Contratto cibernetico*, in *Dir. e inform.*, 2/2011, p. 169 ss.

³⁰ G. TEUBNER, *Ibridi ed attanti. Attori collettivi ed enti non umani nella società e nel diritto*, trad. di L. Zampino, Mimesis, Milano-Udine, 2015, p. 21 ss.



dalle parti. In questo contesto ciò che assume maggiore rilievo è l'associazione uomo-macchina ossia la relazione, le connessioni che corrono tra l'agente *software* e il contraente che lo delega all'esecuzione del contratto: idea che prende origine dalla teoria degli ibridi³¹. Gli orientamenti etici proposti dal recente documento europeo vanno inseriti in questa riflessione. Come si dovrebbe conseguire l'affidabilità dell'intelligenza artificiale? La risposta europea è costituire un rapporto di fiducia tra l'uomo e l'agente *software*. Dunque il presupposto è che quest'ultimo abbia una propria autonomia. Tuttavia questa va calibrata e plasmata a seconda della funzione che legittimamente può essere riconosciuta a tali agenti nel contesto dell'ordinamento vigente³². Tale operazione non può essere lasciata ai codici di autoregolamentazione, né a principi etici, ma va riconosciuta alla scienza giuridica e sottoposta al controllo di legalità. Questa non viene meno a causa della mancanza dell'elemento territoriale inteso come sedimentazione spaziale delle garanzie, ma semmai nel contesto dematerializzato e virtuale essa può trovare nuove forme espressive. Del resto, proprio la caratterizzazione del post-moderno come di un *locus* senza centro avvalorava la necessità di rinviare all'ordine costituzionale³³. Il rischio da tenere in conto è che dietro il rassicurante approccio etico e, dunque, attraverso la c.d. etica applicata, proliferi un processo di ibridazione tra norme tecniche e norme giuridiche, una sorta di diritto eticizzato che assume quale paradigma conoscitivo il modello macchinico anziché il valore della persona. Tale supremo valore rischierebbe di essere compromesso dalla incontrollabile contaminazione artificiale, rischio al quale non si può mostrare indifferenza salvo voler rimanere ciecamente ancorati al mondo della finzione.

³¹ Vedi ancora G. TEUBNER, *o.l.c.*

³² In risposta ad un *responsability gap*, infatti, la creazione di una nuova personalità giuridica per i robot, al fine di poterli rendere (quali centri di imputazione di atti ed effetti giuridici) direttamente responsabili nei riguardi dei terzi per i danni cagionati, rappresenterebbe la soluzione più coraggiosa. Percepibile risulta il rischio di equivoci nell'attribuire soggettività ad un agente non umano ma meccanico. Tuttavia, la possibilità di riconoscere una soggettività dei *robot* – seppur nella sua inconfutabile particolarità – sarebbe una risposta adeguata ad un embrionale ma crescente *standard* di autonomia. La configurazione di un tale e nuovo soggetto di diritto sarebbe tendente e strumentale a favorire il meccanismo di imputazione di effetti giuridici in capo al *robot*, permettendo così una ripartizione degli oneri economici (derivanti dal risarcimento da corrispondere) tra i soggetti coinvolti nel finanziamento di una massa patrimoniale associata alla macchina. Una responsabilità così imputata al *robot*, soggetto di diritto, ma di fatto ripartita tra i soggetti che operano nel funzionamento economico del *robot* medesimo, non costituisce un deterrente al progresso – come viceversa ne deriverebbe se dovesse essere il solo produttore, ad esempio, a sopportare i costi di un risarcimento per danni causati – e, al tempo stesso, risulta adeguatamente soddisfattiva dell'interesse risarcitorio del danneggiato. Il riconoscimento di soggettività giuridica, pertanto, può giustificarsi in modo strumentale: una finzione giuridica per meglio organizzare la gestione dei rischi e l'allocazione della responsabilità. Cfr. sul punto E. PALMERINI, *Robotica e diritto: suggestioni, intersezioni, sviluppi a margine di una ricerca europea*, in *Resp. civ. e prev.*, 6/2016, pp. 1816-1850 cit.; D.J. CALVERLEY, *Imagining a non-biological machine as a legal person*, in *Artificial Intelligence & Society*, 2008, p. 22; G. TEUBNER, *Rights of non-humans? Electronic agents and animals as new actors in politics and law*, in *Journal of Law & Society*, 2006, pp. 497-521; A. BENSOUSSAN, *Plaidoyer pour un droit des robots: de la "personne morale" à la "personne robot"*, in *La lettre des juristes d'affaires*, n. 1134, 28 ottobre 2013.

³³ A. VENANZONI, *Intersezioni costituzionali. Internet e Intelligenze Artificiali tra ordine spontaneo, natura delle cose digitale e garanzia dei diritti fondamentali*, in www.forumcostituzionale.it/wordpress/?p=10709, p. 1 ss.

