

ALESSANDRA ESTHER CASTAGNEDI RAMÍREZ

*Riflessioni a partire da Fernando H. Llano Alonso,
Homo Ex Machina. Ética de la inteligencia artificial y Derecho
digital ante el horizonte de la singularidad tecnológica*

Abstract

This review examines *Homo Ex Machina: Ética de la inteligencia artificial y Derecho digital ante el horizonte de la singularidad tecnológica* by Fernando Llano Alonso, which explores the ethical and legal challenges posed by artificial intelligence (AI) within the advancing horizon of technological singularity. Introducing the concept of “technological humanism”, the book advocates for reconciling scientific advancement with humanistic values while addressing critical issues such as neuro-rights, algorithmic governance, and digital justice.

Structured across six chapters, the book analyzes AI’s impact on human identity, the risks of algorithmic bias, and the integration of AI into legal and social systems. It underscores the need for robust ethical and legal frameworks to preserve human dignity and agency while responsibly incorporating AI into society.

This work stands as a cornerstone in the field of AI ethics and digital law, offering innovative solutions to harmonize human and machine coexistence. By envisioning a future where AI serves humanity without compromising fundamental values, it provides an essential guide for navigating the complexities of the digital age with integrity.

Keywords

- Technological Humanism
- Artificial Intelligence Ethics
- Neuro-rights and Human Identity
- Algorithmic Governance and Bias
- Legal Philosophy in the Digital Era.

Sommario

1. Introduzione
2. Contenuto e struttura del libro
 - 2.1. Etica e neurodiritti
 - 2.2. Diritto e giustizia digitale
 - 2.3. e Governance della IA
3. Un finale ad effetto
4. Conclusioni

1. Introduzione

“Homo Ex Machina” è un’opera scientifica che si inserisce all’interno di una più ampia produzione giuridica, elaborata da Fernando Llano Alonso, contraddistinta da uno studio accurato e approfondito sull’etica dell’Intelligenza Artificiale e sull’influenza che la diffusione delle nuove tecnologie esercita sulla società contemporanea.

Non ci sorprende, dunque, la proposta così rilevante e significativa contenuta nel libro oggetto della presente recensione di attribuire una definizione e di illustrare al lettore le implicazioni che la nuova corrente giuridico-filosofica sta determinando, identificata con il termine “umanesimo tecnologico”. Tale fenomeno prevede la possibilità di conciliare l’avanzamento scientifico-tecnologico con i principi e i valori dell’Umanesimo. Questa iniziativa prende avvio nel 2018 con la pubblicazione del libro “*Homo Excelsior. Los límites éticos-jurídicos del Transhumanismo*”, che propone una profonda riflessione sull’essenza umana, considerata una specie in via di estinzione a causa del rapido progresso tecnologico che non tiene in considerazione i limiti etici che contraddistinguono l’intera esistenza biologica. Al contrario, in “Homo Ex Machina” l’obiettivo è incentrato nell’offrire al lettore una visione integrale sul futuro della “singolarità tecnologica”, ovvero una realtà dominata da esseri umani e cyborg che, indipendentemente dalla loro forma e tangibilità, saranno dotati di un’Intelligenza Artificiale forte e generale. L’importanza di quest’opera, inoltre, non risiede solo nel suo ricco e copioso contenuto, ma anche nella varietà delle aree del diritto che ogni riflessione abbraccia, come il Diritto amministrativo, la Teoria e la Filosofia del Diritto, la Scienza del Diritto, la Sociologia Digitale e l’Informatica Giuridica.

Pertanto, di fronte al nuovo “umanesimo tecnologico”, è necessario stimolare una *consapevolezza tecnologica* che inviti i giuristi a partecipare al dibattito sulle recenti sfide e problematiche che la globalizzazione digitale, attraverso internet, sta suscitando. Prima di addentrarci nella struttura del libro, merita una particolare attenzione il prologo redatto dal noto Professore di Filosofia del Diritto dell’Università degli Studi di Firenze, Stefano Pietropaoli. In questo contesto, l’eventuale fine dell’umanità viene dipinta con le seguenti parole: «*pero hoy nos enfrentamos a un problema totalmente nuevo, o, mejor dicho, a un problema que se ha hecho presente, real, concreto, verdadero, después de haber sido durante*

siglos solo un fantasma agitado por lo que muchos juzgaban mentes perturbadas, afectadas por delirios, fabricantes de pesadillas que parecían destinadas a permanecer evanescentes» (trad. «ma oggi ci troviamo di fronte a un problema del tutto nuovo, o, meglio, a un problema che è diventato presente, reale, concreto, vero, dopo essere stato per secoli solo un fantasma agitato da ciò che molti consideravano menti disturbate, affette da deliri, creatori di incubi destinati a rimanere evanescenti»). L'enfasi che l'autore fiorentino vuole dare alla preoccupazione attuale si manifesta in una descrizione dettagliata e realistica della scena, in cui l'umanità funge sia da spettatore che da vittima. L'unica fonte di salvezza è la definizione di limiti etici che devono essere inseriti nei contesti giuridico-normativo, per evitare di danneggiare parzialmente o integralmente il pensiero umano e conseguentemente eludere la possibilità che il diritto lentamente possa conoscere la sua “preannunciata” morte.

Con estremo rigore, Pietropaoli descrive la visione del Professore di origini cantabrichi, respingendo sia le ideologie dei tecnofili integrati, sia quelle dei tecnofobi apocalittici. La vera questione che si cela dietro a questo scenario non è solo legata all'importanza di mantenere e conservare l'identità umana, valorizzando il concetto di dignità umana, ma anche di assistere eventualmente ad un possibile tramonto della scienza giuridica del diritto intesa come conoscenza sviluppata e alimentata da giuristi umani. Se da un lato, si può affermare con certezza che siamo sommersi sempre più in una realtà digitale, popolata da identità digitali, che allontanano la razza umana dalla propria condizione di *essere* in termini ontologici, evolvendosi in sagome costituite da un insieme di dati interessanti; dall'altro lato, è altrettanto vero che la conseguenza di questo scenario potrebbe essere verosimilmente una fiducia, sempre più in espansione, negli strumenti di Intelligenza Artificiale, lasciando alle spalle ogni tentativo di continuare ad alimentare le riflessioni elaborate dal cervello umano sulla realtà, quale è il diritto per esempio, e lasciando spazio a processi computazionali imparziali, terzi e apparentemente giusti realizzati dagli algoritmi. L'autore sottolinea il valore metafisico di questa scienza giuridica, affermando che «*el derecho no es, si no es humano*» (trad. «il diritto non è, se non è umano»). Il grande dilemma è se si possa considerare il “diritto” alla stregua dell'accezione con cui oggi è identificato, nel caso in cui si decida di ridurlo ad una raccolta degli elaborati imparziali e neutri provenienti delle macchine intelligenti. Fornire una risposta a questa quesito,

al giorno d'oggi, è piuttosto complicato e qualsiasi eventuale soluzione venga adottata, dipenderà dalle decisioni che verranno prese a livello politico, etico e giuridico. Ripercorrendo le parole del professor Pietropaoli, *«el riesgo que corre el derecho de ser invadido por nuevos actores, distintos a los humanos, está en la fuerte tentación de “imputar a las maquinas cualidades que tradicionalmente hemos reservado a los seres humanos, como la inteligencia, autonomía, imparcialidad, conciencia, subjetividad, (...) frente a una erosión progresiva de las mismas cualidades en el ser humano”»* (trad. «il rischio che corre il diritto di essere invaso da nuovi attori, diversi dagli esseri umani, è nella forte tentazione di “attribuire alle macchine qualità che tradizionalmente abbiamo riservato agli esseri umani, come l'intelligenza, l'autonomia, l'imparzialità, la coscienza, la soggettività, (...) di fronte a un'erosione progressiva delle stesse qualità nell'essere umano”»).

2. Contenuto e struttura del libro

Procedendo con un'attenta analisi sulla struttura del libro, osserviamo che è articolato in sei capitoli, nei quali emergono le molteplici dicotomie che devono essere prese in considerazione rispetto allo studio sulla diffusione dell'Intelligenza Artificiale, estrinsecate seguendo tre ordini normativi: l'Etica, il Diritto e la Politica.

2.1. Etica e neurodiritti

La tematica legata alle riflessioni sull'Etica dell'Intelligenza Artificiale è stata sviluppata nel primo e nel terzo capitolo del libro, denominati rispettivamente *“Singularidad Tecnológica, Metaverso e Identidad personal: del Homo Faber al Novo Homo Ludens”* y *“De Máquinas y Hombres. Tres cuestiones ético-jurídicas sobre la Inteligencia Artificial”* (trad. *“Singularità Tecnologica, Metaverso e Identità personale: dal Homo Faber al Novo Homo Ludens”* e *“Sulle Macchine e Uomini. Tre questioni etico-giuridiche sull'Intelligenza Artificiale”*).

Le questioni presentate nel primo capitolo sono intimamente legate al riconoscimento del diritto all'identità umana e ai limiti che devono essere stabiliti in relazione alle terapie di miglioramento. Questo dibattito si estende al ruolo centrale che giocano le neurotecnologie e all'urgente necessità di formulare una nuova generazione di diritti umani, meglio conosciuti come “neurodiritti”, in grado di rispondere accuratamente ai progressi tecnologici attuali e parallelamente permettano una reale integrazione tra

biologia e tecnologia, come si può notare negli ultimi sviluppi annunciati da Elon Musk. L'evoluzione dell'umanità prende avvio a partire dall' *homo faber*, un'entità che nella sua accezione originale si definiva come l'uomo "misura di tutte le cose". Nel corso del tempo, si è assistito ad una evoluzione del genere umano, identificando quest'ultimo con il termine latino *homo ludens*: un essere che trascende i limiti del suo ambiente naturale e si addentra in una realtà aumentata dalla tecnica, conosciuta come "realtà sovranaturale tecnica". L'ultima fase dell'evoluzione della specie umana è associata al termine *phone sapiens*, caratterizzata da una profonda interdipendenza con la tecnologia, che manifesta i suoi effetti in una duplice direzione: da un lato, si sviluppa un legame quasi inestricabile con gli artefatti tecnologici; dall'altro, si percepisce una crescente sfumatura dell'identità e un graduale dissolvimento dell'essenza umana. Potrebbe essere definita come una vera e propria dissociazione tra corpo e spirito, provocata dalla proiezione della vita personale di ogni individuo in una realtà digitale parallela, dominata da una felicità apparente. In questo spazio, l'essere umano affronta una crisi di identità acuta, alimentata dal crescente dominio degli algoritmi sulla propria vita. Questa influenza crescente compromette la sua autonomia ed erode significativamente la libertà di prendere decisioni e agire liberamente. La stessa crisi si manifesta anche nel campo delle neurotecnologie, in particolare nelle *Brain-Computer Interface* (BCI), capaci di decifrare le attività neuronali delle persone. I brillanti risultati di questo fittizio progresso, tuttavia, nascondono la faccia oscura della medaglia: le molteplici e ingiustificate violazioni dei diritti umani, quali il diritto all'identità personale, il libero arbitrio, la privacy mentale e la pari opportunità di accesso a qualsiasi modifica neurologica, che vengono generate quotidianamente e spesso sotto l'occhio ignaro della società. Di conseguenza, la proposta presentata dallo spagnolo Rafael Yuste e dalla *Neurorights Foundation* si concentra sulla formulazione di un insieme di principi noti come "neurodiritti", che hanno avuto particolare eco in Cile. In questo contesto, prosegue la discussione tra coloro che considerano prioritario riconoscere una nuova generazione di diritti e coloro che ritengono opportuno interpretare i diritti umani già esistenti alla stregua dei valori etico-morali che dominano la società moderna. Con l'introduzione dall'emergente metaverso digitale è essenziale salvaguardare l'identità, l'autonomia e l'autenticità umana. Alcuni esperti interpretano l'immersione nel mondo digitale come un trasferimento

delle esperienze umane dal piano fisico a molteplici virtuali. In questo processo si osserva una tendenza generale a trascurare la coscienza umana, causato dallo stato in cui le persone versano in quel momento, paragonabile ad un momento di *trance*.

Il terzo capitolo si addentra in questo fenomeno, intrecciando la realtà tangibile con racconti di fantascienza, motivando il lettore a mantenere uno sguardo critico e discernente. L'autore invita il suo pubblico ad esplorare consapevolmente le implicazioni della "singolarità tecnologica", ponendo una particolare enfasi sulla dualità dei suoi effetti, nonché i rischi e i benefici connessi all'eventuale possibilità che le macchine, in un futuro prossimo, siano capaci di replicare o superare la coscienza umana. Questa analisi ci conduce verso un labirinto di questioni e riflessioni, stimolando una profonda deliberazione sul futuro che, tramite la nostra presa di posizione, stiamo costruendo. Dalle quattro leggi di Isaac Asimov, tra cui è inclusa anche la Legge zero, come riportato nella sua opera più esemplare *Robots and Empire* del 1985, si giunge a una profonda analisi della distinzione tra coscienza funzionale e fenomenica di Margaret A. Boden. Questa è il *passapartout* che apre tutte le porte al temuto contrasto tra gli apocalittici, contrari all'avanzamento della scienza e della tecnologia, e gli integrati, che ripongono tutta la loro fiducia nello scientismo. Al centro si colloca l'opinione dell'autore di questa magnifica opera, sostenitore dell'"umanesimo tecnologico", il quale delinea con estrema precisione le direttive educativo-culturali ed etico-giuridiche che lo contraddistinguono e che devono essere seguite a mo' di manuale di istruzioni. Pertanto, poter accettare che in un futuro prossimo l'essere umano possa trovarsi circondato da robot super intelligenti, solleva ulteriori dubbi quali il possibile riconoscimento giuridico di agenti morali artificiali nel quadro normativo di ciascun paese. Questo dilemma non sembra così scapestrato se si esaminano le tre importanti teorie professate da Peter Singer, John Rawls e dagli scienziati italiani Stefano Nolfi e Davide Floreano. Il riconoscimento di una personalità giuridica e di un diritto specifico per i robot non avrebbe come fine principale regolare il trattamento dei robot come entità autoconsapevoli, titolari di diritti e doveri, bensì di costruire un nuovo istituto giuridico che abbia finalità risarcitorie, in linea con i principi e i valori dei diritti umani riconosciuti sia nella Carta dei Diritti Fondamentali dell'Unione Europea, sia nella Convenzione Europea per la Protezione dei Diritti Umani e delle Libertà Fondamentali.

2.2. Diritto e giustizia digitale

Il secondo blocco tematico, trattato nei capitoli due e quattro, intitolati rispettivamente “*Justicia digital, Algoritmos y Derecho: De la predictibilidad del Big Data al mito del Juez Robot*” y “*Homo Ex Machina: el Ser Humano ante la Inteligencia Artificial y el Gobierno de los Algoritmos*” (trad. “*Giustizia digitale, Algoritmi e Diritto: Dalla prevedibilità del Big Data al mito del Giudice Robot*” e “*Homo Ex Machina: l’Essere Umano di fronte all’Intelligenza Artificiale e al Governo degli Algoritmi*”), contiene un’analisi giuridico-filosofica sulle nuove sfide che il diritto sta affrontando, incluso l’importanza di costruire una buona *governance* degli algoritmi.

Il secondo capitolo si concentra esclusivamente sulla progressiva evoluzione della giustizia digitale nell’ambito legale-giudiziario e sulle tecniche che dovrebbero essere adottate per evitare che la trasformazione della giustizia rappresentativa e umana comporti una denaturalizzazione del diritto. Il punto di partenza della riflessione è la creazione di sistemi di *machine learning* applicati alla pratica giudiziaria, che potrebbero essere meglio descritti come programmi informatici che funzionano tramite l’applicazione di formule algoritmiche, diffondendo così un modello di *smart justice* che non tiene in considerazione l’attività tipica di un giudice-filosofo, capace di risolvere specifici problemi legali combinando i principi e gli *standard* giuridici con i fattori emotivi ed empatici. È evidente, come pone in luce l’autore, che di fronte a questa presa di coscienza, una delle questioni con cui i giuristi devono confrontarsi è se le macchine saranno effettivamente in grado di sostituire il giudice umano. Nel testo viene presentato un eventuale progresso, partendo da una prima fase dominata dall’automazione, passando poi per una seconda caratterizzata dall’innovazione tecnologica, con la possibilità di sfociare in una terza, determinata dalla presenza del giudice-computer, che agirebbe in sostituzione dei giudici-umani. L’autore adotta una posizione neutrale sull’argomento, riconoscendo da un lato l’incapacità delle macchine di replicare la complessità del cervello umano e, dall’altro, che i programmi informatici attuali possano già replicare alcuni aspetti del processo giudiziario. Sottolinea, tuttavia, che ciò che è realmente rilevante non è il raggiungimento della perfezione nell’elaborazione giurisprudenziale contenuta negli atti giudiziari, bensì l’abilità nell’utilizzo della retorica in termini di abilità persuasive. Per quanto riguarda la giustizia predittiva, il secondo aspetto da considerare è la (non) prevedibilità degli algoritmi. Diverse

sono le posizioni dottrinali che si sono manifestate in tal senso, suddivise tra gli arguti sostenitori dell'uso del computer-legale, in grado di risolvere pratiche giudiziarie grazie alla sua capacità rigidamente deduttiva e il gruppo di soggetti che identificano nel diritto un sistema logico non formale, caratterizzato da giudizi di comprensione, verità, buon senso e arte. Se fossimo del tutto favorevoli al *tecno-paradigma*, dovremmo accettare che il potere discrezionale dei giudici che esercitano nelle loro attività di interpretazione e applicazione del diritto, verrebbe sostituito dalle micro-direttive aritmetiche algoritmiche. Le obiezioni relative a questa ultima idea si concretizzano con l'interrogativo: Chi controlla l'assenza di pregiudizi e chi vigila sull'imparzialità del creatore algoritmico? Chi garantisce il principio di trasparenza e legalità nello sviluppo del risultato algoritmico? In verità, come suggerisce il Professore cantabrico, l'ideale sarebbe combinare la scienza giuridica con il futuro tecnologico, evitando una trasfigurazione del diritto in una dimensione digitale e immediata. Di fronte alla smaterializzazione e alla disintegrazione del mondo, è necessario pertanto proteggere l'essenza umana, punto di partenza di qualsiasi ragionamento finora esposto.

Che tipologia di relazione dovrebbe, dunque, essere scelta per governare l'interazione tra macchine e uomini? Questo è il quesito con cui si apre il quarto capitolo. I principali modelli sviluppati oggi, in base al grado di coinvolgimento e alla natura dell'intervento umano, sono quattro, rispettivamente *Human-in-the-Loop*, *Human-in-the-Loop for Exceptions*, *Human-on-the-Loop* e *Human-out-the-Loop*. Il primo approccio conferisce all'utente il controllo sull'attività realizzata da un sistema di *machine learning*, consentendo il trasferimento delle abilità umane al sistema di Intelligenza Artificiale, attraverso diverse tecniche. Un esempio è l'uso delle BCI che consentono ad un esperto di guidare ed insegnare alla macchina, come accade quando si utilizza un telecomando. Grazie agli attuali progressi nel settore della robotica, questo metodo di insegnamento bi-direzionale non solo conferisce alle macchine conoscenza umana, ma consente anche agli umani di ricevere e assimilare abilità dalle macchine stesse attraverso dei *feedback* multimodali: a tal proposito, conviene menzionare i *wearables*, dispositivi indossabili che comunicano dati preventivamente o immediatamente all'utente. Il secondo approccio prevede l'estensione dell'automatizzazione a tutto il processo decisionale, riservando agli esseri umani la possibilità di revisione delle eccezioni. Nella configurazione *Human-on-the-Loop*,

gli esseri umani assistono e convalidano le conclusioni raggiunte dalle macchine. Viene utilizzata un'architettura algoritmica avanzata che consente ai sistemi di Intelligenza Artificiale di suggerire modifiche o aggiustamenti normativi, successivamente valutati e approvati da un operatore umano. Infine, il quarto e ultimo modello si contraddistingue per assegnare alla macchina un ruolo decisivo, occupandosi di tutte le micro-decisioni quotidiane. In questo scenario, l'intervento umano si limita solo a definire i nuovi limiti strategici e gli obiettivi da raggiungere, lasciando le restanti attività quotidiane nelle mani dell'Intelligenza Artificiale. I quattro modelli si caratterizzano per la presenza di una “riserva di umanità”, elaborata per la prima volta da Ernesto Garzón Valdés. Creatore della filosofia del *coto vedado* (trad. area riservata), l'autore argentino ritiene che ci siano beni primari, considerabili come pietre miliari nelle costituzioni democratico-liberali che favoriscono lo stato sociale, che sono innegoziabili e sono assolutamente irrinunciabili per il rispetto della dignità e la coscienza umana di fronte all'irruzione delle nuove tecnologie. La soluzione, come suggerisce Luciano Floridi, consiste nel diffondere una “cultura algoritmica” più responsabile, il cui obiettivo, riprendendo le parole del filosofo italiano, è quello di diffondere una «*responsabilidad moral distributiva entre todos los agentes morales*» (trad. «responsabilità morale distributiva tra tutti gli agenti morali») che coincidono con le persone fisiche o giuridiche.

2.3. e Governance della IA

In conclusione, la terza e ultima sezione tematica è illustrata ampiamente nei capitoli cinque e sei, rispettivamente denominati “*Globalización y Revolución tecnológica ante la crisis del Paradigma Humanista*” y “*Inteligencia Artificial, Discriminación y Sesgos Algorítmicos*” (trad. “*Globalizzazione e Rivoluzione tecnologica di fronte alla crisi del Paradigma Umanista*” e “*Intelligenza Artificiale, Discriminazione e Bias Algoritmici*”).

Il blocco filosofico-politico si ramifica in due parti: una prima sezione focalizzata sulle riflessioni relative a quale sistema politico adottare con l'avvento di un eventuale modello tecnocratico digitale e una seconda che sorprende il lettore con una profonda riflessione sulle fondamenta che dovrebbero essere elaborate per il sostegno dei diritti umani, con il fine di evitare qualsiasi fenomeno di discriminazione e qualsiasi cambio di paradigma determinato dalle nuove tecnologie prive di valori etici.

L'*incipit* del quinto capitolo si distingue per lo studio dettagliato degli effetti determinati dalla quarta rivoluzione industriale tecnologica. L'unione tra globalizzazione e fenomeno del transumanesimo ha creato un nuovo paradigma tecno-scientifico che ripone piena fiducia nella ricerca, idealizzandola e idolatrandola al punto tale da battezzarla come unico strumento di risoluzione dei potenziali rischi globali, celando tuttavia alla società la vera necessità di affrontare responsabilmente le attuali sfide su scala mondiale. L'effetto collaterale nefasto è un generale disorientamento della popolazione, causato dal cambiamento dei valori tipici del paradigma umanista, quali lo stato di diritto, i diritti fondamentali e la dignità umana. Klaus Schwab annuncia la teoria del *Big Reset*, come potenziale fenomeno in grado di definire, tra le plurime conseguenze, un ipotetico cambiamento strutturale nello sviluppo della nostra personalità. Lo scopo di questa dottrina si allontana dall'idea kantiana del progresso, definito con le seguenti e testuali parole da parte del suo ideatore «*para catalizar un nuevo renacimiento cultural que nos permitirá sentirnos parte de algo mucho más grande que nosotros mismos*» (trad. «per catalizzare un nuovo rinascimento culturale che ci permetterà di sentirci parte di qualcosa di molto più grande di noi stessi»). Come soluzione all'espansivo e crescente globalizzazione, nel 2021 Schwab ha proposto un "capitalismo degli *stakeholders*", in cui tutti i soggetti che entrano in gioco (i governi, la società civile, le imprese e la comunità internazionale) sono interconnessi e condividono come obiettivi primari la felicità, il benessere di tutte le persone e l'armonia planetaria. Questa soluzione rappresenta una via di fuga rispetto al regresso della globalizzazione, a cui abbiamo assistito durante la pandemia COVID-19, il cui scopo non si potrà mai celebrare se si valuta il collegamento esistente tra quest'ultima e l'economia mondiale. D'altro canto, questo ritorno all'idea nazionalista dello Stato-Nazione si mostra contrario all'esistenza dei mercati globali. Questo dilemma potrebbe essere risolto con la soluzione proposta da Rodrik, grande sostenitore dell'idea di sacrificio della globalizzazione in favore dell'instaurazione di un certo grado di sovranità nazionale e di democrazia. Il rifiuto manifestato nei confronti di una *governance* globale democratica nasconde un potenziale rischio di sostituzione dell'attuale forma di governo, in accordo con le direttive del Diritto Internazionale, in un'oligarchia che concentrerebbe tutta la ricchezza e il potere tecnologico nelle mani di pochi, comportando una possibile perdita dei diritti naturali così come professati da John Locke; si parla, in questo ambito, della

proprietà privata, della libertà e della felicità reale. Ciò mette in discussione la capacità delle attuali strutture di cooperazione internazionale e suggerisce che il modello del “capitalismo degli *stakeholder*” potrebbe offrire una soluzione alternativa, anche se la sua implementazione pratica potrebbe essere piena di sfide e critiche, specialmente per quanto riguarda la sua rappresentatività democratica e la sua legittimità. Il grande dibattito si concentra su come una possibile inversione di rotta verso il capitalismo degli *stakeholder* possa impattare sulla democrazia, sulla sovranità nazionale e sulla *governance* globale, ed, inoltre, se tale modello possa veramente rappresentare gli interessi di tutta l’umanità di fronte alle sfide e alle crisi globali odierne. La discussione attinge anche all’influenza ed al controllo delle grandi società digitali sulla nostra realtà virtuale, capace di cancellare l’identità individuale delle persone. Si menziona, inoltre, l’argomentazione di Martha Nussbaum su come il *marketing* globale e il transumanesimo stiano esercitando una pressione non indifferente sulla società al punto tale da estraniarci dalla nostra stessa essenza umana e da provocare una erosione della cultura democratica. Il testo prosegue esaminando l’idea di Shoshana Zuboff sui pericoli del capitalismo di sorveglianza, che congiuntamente alle pratiche transumaniste, origina un nuovo modello capitalista che mette a rischio la democrazia e la libertà individuale. La proposta mira a rafforzare la governance globale democratica e a creare un quadro normativo per le nuove tecnologie, mettendo al centro la democrazia e il rispetto dei diritti umani.

Nell’ultimo capitolo si affronta la questione della discriminazione nei confronti di categorie che si distinguono nella società per motivi di sesso, razza, età e disabilità. Il mito dell’infallibilità dell’Intelligenza Artificiale è un’utopia se si considera il *report* dell’Agenzia dei Diritti Fondamentali dell’Unione Europea, denunciante l’onnipresenza di pregiudizi e discriminazioni nella società globale, radicati in dinamiche psicologiche, sociali e culturali. Da qui l’immediata e celere comprensione delle ragioni per cui la predizione dell’Intelligenza Artificiale è suscettibile di essere erranea. Sono stati pubblicati vari studi scientifici che si oppongono all’uso di strumenti di valutazione del rischio basati su algoritmi nel contesto giudiziario o poliziesco. Gli schemi algoritmici con cui lavora l’Intelligenza Artificiale sono noti come *black box*, la cui mancanza di trasparenza rende impossibile esercitare un controllo di equità e di trasparenza, oltre che di rappresentatività dei dati di addestramento degli stessi software da parte del consumatore. La risposta a

questo problema deve basarsi sul rispetto dei diritti fondamentali. Da questa base si origina e si estrinseca la lotta dell'Unione Europea per la protezione dei valori della privacy, della non discriminazione, della sicurezza e dell'applicazione della Legge. Particolarmente rilevante è l'Emendamento 36 della Legge sull'Intelligenza Artificiale del 2023 (*IA Act*), che ha introdotto la necessità di alfabetizzazione nell'ambito dell'Intelligenza Artificiale per evitare la discriminazione in termini di sesso ed età, incitando così i fornitori a promuovere strumenti che garantiscano un sufficiente livello di quest'ultima. Una maniera interessante per contrastare il crescente dominio degli algoritmi all'interno dello Stato di Diritto è la soluzione ideata dagli italiani Andrea Simoncini ed Erik Longo, individuata con il termine "costituzionalismo ibrido", il cui obiettivo è garantire che la dignità e la libertà delle persone diventino parte integrante del programma di formazione degli esperti di Intelligenza Artificiale. In particolare, si tratta di incorporare i valori, i principi e i diritti costituzionali nella fase di progettazione del linguaggio delle macchine.

La conclusione di questo capitolo è rappresentata da una presa di coscienza da parte dell'Unione Europea dei rischi che l'umanità sta affrontando, lasciando ai cittadini la scelta di poter gestire i propri dati. Il legislatore europeo ha quindi sviluppato un quadro normativo che comprende una triplice dimensione: etica, giuridica e politica. Il minimo comun denominatore di queste tre dimensioni è un approccio antropocentrico e antropogenico nell'applicazione dell'Intelligenza Artificiale. Ciò giustifica la necessità di sottoporla ad un giudizio e ad una valutazione etica, che sia posteriormente accompagnato da uno sviluppo etico-giuridico dell'Intelligenza Artificiale all'interno del quadro normativo dell'Unione Europea. Corrisponde alle attività dell'organizzazione internazionale *sui generis* la responsabilità di qualsiasi Stato di diritto di vigilare sul rispetto degli obiettivi, dei valori, dei principi umanistici e umanitari e, più in generale, di valutare un buon uso dei sistemi di Intelligenza Artificiale.

3. Un finale ad effetto

L'autore sorprende il suo pubblico con l'epilogo di questo capolavoro. Il Professor Fernando Llano Alonso esplora la tensione tra il desiderio umano di conoscenza e la prudenza necessaria per navigare tra i pericoli che hanno dominato lo sfondo dell'incontro tra Ulisse e le Sirene in un celebre canto dell'Odissea di Omero. La menzione a questa scena si rea-

lizza per la reale somiglianza con la realtà a cui assistiamo ogni giorno, in cui l'uomo contemporaneo si trova di fronte alle sfide dell'era tecnologica. Questa discussione si estende ad un'analisi critica su come la tecnologia, e in particolare l'Intelligenza Artificiale, debba essere incorniciata all'interno di un contesto umanista. Si difende senza alcun dubbio l'avvento del "tecno-umanesimo", un approccio che si colloca in equilibrio tra l'accettazione dei progressi tecnologici e la preservazione dei valori umanistici. La tecnologia deve servire l'umanità, rispettando principi etici e morali, allontanando qualsiasi idea di dominio o sostituzione delle capacità umane. Nella conclusione, l'autore sviluppa definitivamente il suo punto di vista, argomentando che l'umanesimo tecnologico, lungi dall'essere obsoleto, è essenziale per integrare in modo efficace l'innovazione tecnologica nella società, garantendo che tutti ne possano beneficiare e non solo alcuni. Il testo invita a riflettere su come l'umanità possa affrontare le sfide dell'era digitale, mantenendo come impegno saldo il rispetto dei principi umanistici che da sempre hanno guidato il suo progresso nel corso degli anni.

4. Conclusioni

“Homo Ex Machina. Etica dell'Intelligenza Artificiale e Diritto digitale di fronte all'orizzonte della singolarità tecnologica” rappresenta un vero e proprio punto di svolta nell'evoluzione del pensiero umanistico di fronte all'avanzata tecnologica virtuosa e vertiginosa alla quale stiamo assistendo. Potrebbe essere considerato un pilastro nel campo della ricerca, avente ad oggetto l'etica dell'Intelligenza Artificiale, poiché l'opera approfondisce ampiamente la coesistenza tra umani e macchine, proponendo soluzioni etico-filosofiche che cercano di armonizzare questo binomio. Dal mio punto di vista, questo lavoro supera il suo predecessore, *“Homo Excelsior. Los límites ético-jurídicos del transhumanismo”*, non solo per l'ambizione che dimostra, ma anche per la capacità di proiettare un futuro in cui le nuove tecnologie, nel pieno rispetto delle tre famose leggi di Isaac Asimov, servano allo sviluppo umano senza violare i nostri valori fondamentali. Di fronte alle considerazioni e le soluzioni che l'autore ci suggerisce, soprattutto per quanto riguarda il ruolo che il cittadino dovrebbe avere rispetto alle nuove sfide tecnologiche da affrontare, questa opera potrebbe essere considerata una bussola morale per navigare nell'attualità, dominata dalla così tanto temuta singolarità tecnologica.