

ALLA RICERCA DI UNA POSSIBILE EMPATIA ALGORITMICA. MODELLI ETICO-GIURIDICI DELL'IA

Gerardo Marco Bencivenga*

SOMMARIO: 1.- Algoritmo, intelligenza artificiale: tra modello deterministico e modello euristico, tra evoluzione mimetica ed evoluzione protesica; 2.- Due problematiche fondamentali: l'eterodirezione automatica e la delega sistematica; 3.- Empatia algoritmica e rinnovata pedagogia delle intelligenze.

1.- Algoritmo, intelligenza artificiale: tra modello deterministico e modello euristico, tra evoluzione mimetica ed evoluzione protesica.

La tecnologia, oggi, si propone come un “dispositivo complesso” volto a una più o meno completa organizzazione della vita sociale, la quale vede ogni giorno ricadere su di sé la possibilità di utilizzare nuovi dispositivi tecnologici e, di conseguenza, quella di sviluppare delle nuove “costumanze” indotte da essi, con forti implicazioni sia sul piano socio- economico sia su quello più generalmente cognitivo, in un paesaggio sociale nel quale l’artificio tecnologico, inevitabilmente, si proietta sui più delicati ambiti della vita comunitaria, e dunque pure, com’è noto, sul diritto. In realtà, come ben ci spiega Teresa Numerico, tale dispositivo tecnologico introietta, e al contempo proietta modelli di convivenza del tutto progettati e controllati da “forze geo-politicamente e storicamente collocate” di fronte al grande orizzonte dei cosiddetti “Big Data”, gestiti da esse. In simile scenario evolutivo delle tecnologie, lo sviluppo in senso algoritmico di tante funzioni e l’avvento dell’Intelligenza Artificiale pone decisive sfide in gioco: è già davvero molto ampio in questo momento il materiale di riflessioni dedicate all’argomento¹, ma nel dibattito aperto vogliamo provare a offrire un nostro contributo, forse, con qualche motivo di originalità e di interesse al dibattito stesso. Diventa però allora necessario provare a ricostruire, seppur per poche battute, l’essenziale quadro storico all’interno del quale si è tracciata questa evoluzione, poiché, se non è “astratta” la collocazione storica e geopolitica di chi produce/ gestisce l’informatica applicata e la IA, non risulta neppure astratta la maturazione ed evoluzione storica delle tecnologie in quanto tali, le quali si sono in realtà via via arenate, trasformate, abbandonate o riprese nel corso del tempo e a seconda dei casi.

Proprio la citata studiosa ha dedicato un essenziale, quanto ben riassuntivo, resoconto² storico-critico legato al concetto di intelligenza artificiale, dal quale emerge come in effetti tale concetto si sia evoluto attraverso l’utilizzo e sviluppo di diversi suoi contenuti tecnici, e con ciò anche la proiezione di differenti finalità sul piano conoscitivo e sociale. Una storia che comincia alla metà del secolo scorso ma già allora da distinguersi, in effetti, in due diverse linee di studio

* Dottore di ricerca, docente di materie giuridiche ed economiche, avvocato e sindacalista. Docente a contratto di Legislazione Scolastica presso la SSML “Teseo” di Salerno.

¹ Come inquadramento filosofico più ampio, e una fotografia generale del fenomeno disincantata è possibile offrire questo doppio riferimento iniziale: F.J. Varela, E. Thompson, E. Rosch, *The Embodied Mind: Cognitive Science and Human Experience*, Cambridge (Mass.) 1993, e K. Crawford, *Né intelligente né artificiale. Il lato oscuro dell'IA*, Bologna 2021. Come volume pionieristico sul rapporto tra scienza cibernetico-informatica e il diritto, V. FROSINI, *Cibernetica diritto società*, Milano 1968.

² T. Numerico, *Intelligenza artificiale e algoritmi: datificazione, politica, epistemologia*, in *Consecutio Rerum* 3.6 (2019) 241ss.

e applicazione di ricerca: l'una, quella sviluppata da Alan Turing, l'altra quella originata dagli studi della cibernetica, specialmente facente capo a Norbert Wiener. La linea degli studi di Turing si proponeva lo sviluppo di una intelligenza artificiale quale strumento in grado di fingere di saper maneggiare il linguaggio, dunque altamente concentrata sul funzionamento dei sistemi di linguaggio da trasferire, in questo modo, dall'uomo alla macchina: essa viene pensata come un dispositivo capace di “svolgere dei compiti intelligenti”, e il cui battesimo storico nel quale venne proposta per la prima volta l'espressione “intelligenza artificiale” fu la “Summer School” di Dartmouth, nel 1956.

Un percorso complesso, dalla difficile “maneggiabilità”, ma in cui comincia a imporsi l'idea della macchina come “simulazione della mente”, avviato dal famoso Test di Turing con il cosiddetto gioco dell'imitazione³, nel quale veniva proposta la possibilità di un confronto proprio sul piano del “rendimento giuridico”, tra le risposte nascoste date da una macchina o da un vero essere umano. Il fattore etico- filosofico qui significativo da sottolineare, dal dibattito conseguito, è questo fatale “scambio” uomo-macchina posto al funzionamento della macchina stessa, ossia quella “capacità di mentire” così caratteristica dell'uomo e dei suoi linguaggi, nella costitutiva prospettiva d'azione del “fare finta di comportarsi come un essere umano”, in questo suo, tendenziale mostrarsi fondamentale del dispositivo artificiale. Per poter funzionare, cioè, la macchina intelligente per Turing⁴ deve “sembrare umana”, “pretend to be a man”, fino alla possibilità dell'inganno: un circuito nel quale si vede dunque alle origini di tali tecnologie l'intero calco proiettato su di esse dell'intenzione e proiezione squisitamente umane, indovinate non solo nella posizione retrostante, costruttiva alla base delle macchine, ma anche nell'attribuzione successiva di significato nei loro confronti giunta *a posteriori*, nell'uso cioè sociale delle loro risposte, relativo al valore loro attribuito⁵. Comprendiamo già allora, osservando le origini della ricerca in materia, come si annuncii inevitabile un'ottica della relazione uomo-IA da leggersi nel senso peculiare di un'azione di scambio, e non di una relazione di dominio, per esempio della macchina nei confronti dell'uomo. D'altra parte, seguendo per esempio le tesi di Weizenbaum⁶, l'impatto delle tecnologie sembra a sua volta “riprogrammare” la nostra stessa “immagine del mondo”, attraverso un procedimento di “imaginative reconstruction of the world” capace di condizionare il sentimento di ciò che è “accettabilmente umano”, oppure no, nel contesto sociale; un livello ulteriore di declinazione individuabile, dunque questo, nella caratteristica relazione di scambio tra l'uomo e i suoi dispositivi strumentali.

Ma vi è un altro punto essenziale sull'intera questione emergente sin dall'inizio nelle ricerche di Turing, e cioè il compito circoscritto dato alle macchine intelligenti, la cui capacità è pensata in termini soprattutto linguistici e matematici: in tale peculiare funzionamento, le macchine sono costruite “prive di corpo”, o meglio vengono istituite acquisendo tutta una serie di determinate capacità linguistico-razionali degli esseri umani, ma in una fatale assenza al contempo di tutti gli altri attributi dell'uomo, definiti nel complesso di funzioni della sua corporalità vivente, situata in un

³ S. Shieber (cur.), *The Turing Test. Verbal behavior as the hallmark of intelligence*, Cambridge (Mass.) 2004.

⁴ A.M. Turing, *Intelligent machinery*, in J. Copeland (cur.), *The essential Turing*, Oxford 2004.

⁵ J.H. Moor (cur.), *The Turing Test. The elusive standard of Artificial Intelligence*, Dordrecht 2006.

⁶ J. Weizenbaum, *Computer power and human reason. From judgement to calculation*, New York 1976. Weizenbaum è stato l'inventore di ELIZA-A, il primo *chatbot*, quale “software” progettato per simulare una conversazione con un essere umano.

conto di stimoli polivalenti. Di fatto, l’idea di Turing sembra essersi fondata su di una vera e propria scissione tra attività del corpo e intelligenza, circoscritta ora esclusivamente alla serie di operazioni del linguaggio. Tale aspetto, l’assenza di corpo del dispositivo artificiale, va considerato anch’esso una chiave fondamentale di lettura quando ci si confronti alla tematica dell’intelligenza artificiale, aspetto sul quale un altro, illuminante contributo storico nel merito si sofferma, quello offerto da Gianfranco Bettetini nel suo *La simulazione visiva*⁷, scritto ancora agli albori della diffusione informatica di massa. Bettetini propone nella sua analisi alcune fondamentali distinzioni, tra cui quella ereditata da Dreyfus⁸ proprio sulla problematica condizione della IA dal momento in cui essa appare priva delle condizioni basilari dell’avere un corpo: ovverosia, la capacità di riconoscere una situazione, di saper scegliere e individuare dei tratti pertinenti agli oggetti e ai fenomeni, e soprattutto di riuscire a concepire una rappresentazione generale del mondo. L’orientamento rigoroso ed esclusivo dell’Intelligenza Artificiale di fatto, nel suo essere priva d’un corpo plurirelazionato al mondo, ne impedisce l’accesso all’universale, a una universalità la quale, a sua volta, sola può permettere la fondamentale attitudine al “buon senso”. A sua volta, la questione della *generalità* connessa alla IA fu affrontata da un altro dei grandi pionieri degli studi informatici, John McCarthy, che cercò per anni un linguaggio in grado di acquisire alle macchine la conoscenza sulla via del senso comune⁹.

Nello studio di Bettetini si distingueva peraltro tra un duplice livello di sviluppo delle intelligenze artificiali, due tipi differenziati di procedimenti trasformativi: una linea di forme deterministiche dell’algoritmo, volte all’esecuzione di un’esclusiva funzione secondo il richiamo di operazioni ripetitive, e una linea di forme euristiche deduttive basate su motori inferenziali e determinantesi attraverso un ragionamento per tentativi. Lo studioso italiano dunque già intravedeva all’orizzonte, attorno al 1992, una duplice intesa e sviluppo riconducibili allo sviluppo delle intelligenze artificiali, da una parte quella rivolta alle “soluzioni” - fatta d’un modello di risposte immediate, subitanee e capaci di organizzare una standardizzazione del proprio rendimento - e dall’altra quella d’un “viaggio cognitivo” nel senso della conoscenza in evoluzione, su quella linea euristica della relazione uomo-macchina meno rigida ma più ampia nel suo procedimento ed orizzonti di funzionamento. Potrebbe allora ipotizzarsi in questo senso anche una differenza qualitativa, emergente tra il modello della risposta e il modello del processo, che la relazione con l’Intelligenza Artificiale, da parte dell’intelligenza umana, sembra poter sviluppare. Fattore decisivo qui da sottolineare è come lo stesso Turing, alle origini di questa avventura di ricerca, avesse ipotizzato di fondare i dispositivi informatici su di una ambivalenza fondamentale del linguaggio, concepiti nel loro poter mimare, imitare l’ambiguità dei linguaggi umani stessi, riproponendo quindi al centro del progetto dell’intelligenza artificiale l’ambivalenza propria della comunicazione interpersonale.

Gli sviluppi storici della ricerca in tale campo paiono in effetti confermare questo campo di oscillazioni, tra determinismo, riduzionismo e apertura di senso comunicativo. In effetti, gli eredi di Turing che giunsero alla fine degli anni ‘60 a sviluppare il programma delle intelligenze artificiali si concentrarono soprattutto su di una matrice di tipo logico-matematico quale base fondamentale di funzionamento delle macchine, in cui “l’idea quindi di addestrare la macchina, come avrebbe

⁷ G. Bettetini, *La simulazione visiva*, Milano 1992, 102ss.

⁸ H. L. Dreyfus, *What Computer can’t do. The Limits of Artificial Intelligence*, New York 1972-1979.

⁹ J. McCarthy, *Generality in artificial intelligence*, in *Communications of the ACM* 30.12 (1987) 1030ss.

desiderato Turing, era stata accantonata”, avviando perciò la ricerca nella peculiare prospettiva delle “macchine capaci di decisioni”, e dunque di un forte orientamento, ora, piuttosto funzionalista di esse. Emerge su di un piano diremmo pedagogico allora in gioco un’ulteriore opposizione, quella tra una concezione comportamentista/costruzionista e una concezione funzionalista, colte a confrontarsi tra loro nel corso della storia scientifica degli studi informatici; un aspetto su cui poco più avanti ritorneremo. Da una parte la relazione con l’intelligenza artificiale concepita come un continuo addestramento costruttivo, diremmo in costante “relazione aperta” con le macchine, dall’altra la costruzione d’un sistema di regole logico-matematiche basilari volte a trasformare le macchine in veri e propri dispositivi capaci di “prendere decisioni”. Quest’ultima concezione di fondo, di fatto meno dominante sul piano scientifico nei decenni più recenti, sembra però permanere e riemergere non raramente negli approcci comuni dell’opinione pubblica e, talvolta, delle stesse istituzioni, e si pone in ogni caso come modello possibile d’utilizzo considerato dell’intelligenza artificiale nel contesto giuridico, dove la fiducia in un dispositivo artificiale capace di prendere decisioni, poiché intimamente in se stesso logico e consequenziale, sembra poter persino prevalere, se non si comprendono adeguatamente una serie di distinzioni concettuali e qualitative come, per come possibile, stiamo tentando di sintetizzare ed illustrare.

Tale assunto d’assimilazione, nel quale un modello di mente logico- matematica s’impone quale sistema ideale di regole e regolazioni passando dagli studi logici umani al funzionamento della macchina - la quale, potenziandosi in questo tipo di prestazione, diventa in prospettiva superiore all’uomo, nel suo rendimento di lettura e calcolo - si accompagna in maniera coerente alla visione e sviluppo storicamente paralleli sviluppati dagli studi sulle intelligenze artificiali, quelli elaborati dalla cosiddetta disciplina cibernetica. Per la cibernetica¹⁰, infatti, l’idea-guida era quella di “guardare alle macchine come a dispositivi indistinguibili da altri agenti organici”¹¹, descendendo come noto all’origine dall’esigenza di far interagire-contrapporre sistema artificiale di macchine contro altro sistema artificiale di macchine, nel contesto bellico della Seconda Guerra Mondiale. In questa prospettiva, dunque, umano e non umano finivano per sovrapporsi e in qualche modo “confondersi” nell’ottica di una possibile, futuribile simbiosi, integrazione anche fisica tra l’uomo e la sua stessa macchina. Qui dobbiamo allora distinguere ancora una volta una duplice sorta di declinazione di tale prospettiva, perché se da una parte l’integrazione fisica uomo-dispositivo artificiale appare oggi sotto potenziali o sperimentali forme - spesso anche inquietanti - di microchip sottocutanei, parti robotizzate del corpo ed ibridazioni transumane, in realtà la linea degli studi legati alla cibernetica ha aperto su di un altro piano una declinazione molto più immediata e sostenibile nell’incontro uomo- macchina, e cioè la trasformazione dei dispositivi informatici in quei computer personal dall’utilizzo massivo e comune, basati sulla formula e il rendimento del cosiddetto user-friendly. Il computer si è trasformato cioè in una zona quotidiana di interfaccia quale oramai ben noto spazio ausiliare delle attività umane¹², fondato su di una programmazione orientata agli oggetti e creando un ambiente di costante interazione col suo utente, nell’ottica progettuale di un altro pioniere dell’informatica come Licklider¹³. Tale peculiare declinazione degli studi cibernetici, come rimarca

¹⁰ N. Wiener, *Cybernetics or Control and Communication in the Animal and the Machine*, II ed., Cambridge (Mass.) 1961.

¹¹ T. Numerico, *Intelligenza artificiale* cit. 254.

¹² L. Manovich, *Il linguaggio dei nuovi media*, Milano 2005.

¹³ J.C.R. Licklider, R.W. Taylor, *The Computer as a Communication Device*, in *Science and Technology* 76 (1968) 21ss.

la Numerico, fece dunque la decisiva scelta di concentrare la propria implementazione e progettazione sulla centralità della comunicazione, oltre o al di fuori di una più rigida concezione funzionalista ed efficientista dell'intelligenza artificiale attribuita alle macchine.

Si tratta perciò di una decisiva svolta storica la quale, in realtà, secondo noi può ben riassumere la linea da seguire nel contesto attuale in cui sistemi di intelligenza artificiale ed esseri umani paiono sempre più chiamati all'interazione quotidiana, poiché già suggerisce nella sua trama storica, sia pratica che teorica, una "praticabilità" interattiva e proficua del mezzo artificiale da parte del suo utilizzatore umano, interazione tendenzialmente aperta quanto evolutiva nel suo costante stimolarsi e qualificarsi. Storicamente, peraltro, è stato proprio questo il processo di ricerca capace di portare alla grande rete interattiva mondiale - lo spazio Internet ormai ben conosciuto, disceso dagli studi dell'IPTO (Information Processing Techniques Office) - così come alle pratiche e modelli di datificazione legati statisticamente a comportamenti e connessioni del pubblico utente di tali reti. Vedremo nei prossimi passaggi meglio più da vicino gli aspetti di interesse, ma anche controversi, legati a tale nuovo ambiente macroinformativo e relazionale, cercando però già di sottolineare come da tale breve ricostruzione evolutiva delle visuali prospettive legate all'intelligenza artificiale debba per noi soprattutto farsi riconoscere come decisiva quella linea a vocazione interattiva, piuttosto relazionale che non predittiva o performativa in senso decisionale e unidirezionale. Si tratta cioè di concepire l'uso delle intelligenze artificiali quale rapporto interattivo - magari, come vedremo, anche a suo modo empatico - e quindi all'interno di una fondamentale prospettiva relazionale e, più ampiamente considerando, pedagogica in tale scambio cognitivo dall'aspetto e modalità di funzionamento inedite. A questo proposito, una felice distinzione posta molto recentemente dal significativo intervento del Garante AGCOM nella sua relazione istituzionale del 2024¹⁴ al Parlamento italiano, firmata da Pasquale Stanzione, può essere qui coerentemente ripresa nel nostro percorso: la distinzione cioè tra prospettiva mimetica e prospettiva protesica dell'evoluzione tecnologica, nella preoccupazione di non vedere la seconda esclusivamente prevalere sulla prima, laddove una concezione e prassi delle tecnologie solo come pure estensioni/protesi delle facoltà umane rischierebbe, davvero, l'esito fatale di una posizione minoritaria *in progress* della persona umana stessa, a ruota delle sue stesse macchine e dei loro responsi - parola d'antica ascendenza giuridica, nel contesto di origine latina - standardizzati e tecnocratici. Al contrario, l'integrazione tra protesi delle intelligenze artificiali e mimesi di scambio da parte di queste con le intelligenze umane appare una prospettiva meno insidiosa e, sotto diversi aspetti, più promettente e performativa.

Su di un piano filosofico-pedagogico, dunque, possiamo meglio procedere dopo questa necessaria ricostruzione storica della ricerca informatica, e alcune distinzioni fondamentali da essa emerse, capaci di instradare il cammino.

2.- Due problematiche fondamentali: l'eterodirezione automatica e la delega sistematica.

L'occupazione sempre più pervasiva della IA nella vita sociale può essere ben testimoniata da dati come quelli presentati dallo stesso documento sopra citato del Garante della Privacy: «Si consideri che circa il 65% dei ragazzi utilizza oggi l'intelligenza artificiale per svolgere i compiti; due studenti

¹⁴ Garante per la protezione dei dati personali, *Regolare il futuro. La protezione dei dati per un'innovazione antropocentrica*, Relazione del Presidente Pasquale Stanzione per l'anno 2023, Roma, 3/7/2024, 4.

su tre avrebbero preparato l'esame di maturità ricorrendo a *Chat Gpt* che peraltro, a quanto pare, non sarebbe riuscita a tradurre correttamente il Minosse, o Della legge, attribuito a Platone. L'intelligenza artificiale è riuscita persino ad arricchire, con effetti visivi e sonori straordinari, la Turandot rappresentata alla Scala. Un' impresa su quattro, nel nostro Paese, ha già integrato l'intelligenza artificiale nei propri processi produttivi ed entro un anno – si stima – il 60% delle aziende la utilizzerà nei procedimenti assunzionali»¹⁵.

Una prospettiva ben pervasiva anche dello stesso diritto, come attestano tanti documenti¹⁶, una vasta serie di interventi e di riflessioni. Sintetizzando questa materia viva di discussione fortemente in trasformazione, ed accompagnandola con alcune osservazioni fondamentali portate da noi sul campo, di tipo diremmo fenomenologico - poiché tale impatto della IA invade realmente il comune territorio quotidiano, risultando perciò osservabile anche più direttamente nelle sue evoluzioni concrete, da diverse posizioni - possiamo quindi riconoscere due problematiche fondamentali incombere quali "relazioni pericolose" tra l'uomo e l'intelligenza artificiale.

Connesso infatti al filone di ricerca logico-matematica applicato alle intelligenze artificiali ne è derivata una corrispondente prospettiva e un approccio spesso privilegiato: quello di attribuire all'interno dei processi computerizzati demandati alle strumentazioni informatiche la possibilità d'una standardizzazione dei rendimenti e una tendenziale, raggiungibile condizione di neutralità delle interpretazioni e dei giudizi, nell'ideale parallela trasposizione in termini di automazione della risposta e, così, d'una miglior "razionalizzazione" delle attività sociali. L'intelligenza artificiale, cioè, concepita quale superiore strumento per così dire produttivo sempre più al centro della società complessa, idealmente aspirante a un'automazione dei propri processi sempre più performativa ed efficiente. Una concezione evidentemente tecnocratica della vita sociale, certo però alquanto connessa con la miriade di attività della società dei consumi e la sua fitta organizzazione dei flussi, e dunque potenzialmente molto influente, in quanto tale, sulle scelte collettive. L'ideale dell'efficienza produttiva si incontra in questa prospettiva con il parallelo ideale d'una pura neutralità delle scelte: una dimensione privilegiata nell'ottica della "miglior performatività", però seguendo la quale il ruolo dell'umano rischia seriamente di diluirsi, da una parte per l'affermazione tendenziale di processi puramente automatizzati - che non hanno più bisogno, evidentemente, di interventi discrezionali e variabili portati dal giudizio umano - del ciclo produttivo come di quello dei consumi, dall'altra per una presunta, miglior affidabilità in termini di "neutralità" delle proprie risposte e delle proprie azioni, performate da parte di una macchina. La prospettiva di un'organizzazione sociale tutta pervasa da un algoritmo sistematico presenta di fatto, accanto ai vantaggi, senz'altro anche il rischio d'una dimidiaione della funzione umana, e la ricerca d'un certo algoritmo definitivo¹⁷, se per diversi aspetti costituisce una stimolante prospettiva di sfida nel miglioramento dell'efficienza organizzativa sociale, per altri potrebbe contribuire a una certa atrofizzazione di tante prerogative e qualità della scelta cognitiva umana. La pretesa della neutralità, a sua volta, sembra più offrirsi come grande equivoco filosofico-morale che non quale credibile soluzione civile: come già mostrato nel 2013 dal famoso caso COMPAS negli Stati Uniti, l'affidamento d'un giudizio a un sistema algoritmico di IA

¹⁵ Id., *Regolare il futuro* cit. 2.

¹⁶ Per esempio, in testi come A. D'Aloia (cur.), *Intelligenza artificiale e diritto. Come regolare un mondo nuovo*, Milano 2020, oppure S. Faro, T.E. Frosini, G. Peruginelli (curr.), *Dati e algoritmi. Diritto e diritti nella società digitale*, Bologna 2020.

¹⁷ P. Domingos, *L'algoritmo definitivo*, Torino 2015.

si è dimostrato non privo¹⁸ di pregiudizi, mostrando al contempo anche un altro, decisivo problema nella segretezza tecnologica dentro cui l'algoritmo stesso rimane avvolto, quale prodotto strategico d'una società privata¹⁹. L'assenza di trasparenza nel funzionamento secretato degli algoritmi costituisce, in realtà, una delle dimostrazioni più lampanti della non neutralità del mezzo, e ancor maggiormente del suo utilizzo strategico nel gioco d'interessi tutt'altro che neutrale di proprietari, governi, potentati in grado di svilupparne e controllarne il processo²⁰.

Nell'ambito giudiziario la questione appare pressoché decisiva, anche perché da diverse parti tale tendenza ad attribuire alle macchine la possibilità diretta di prendere decisioni, in una maniera ritenuta più efficiente e al tempo stesso più "neutrale" sembra spesso riproporsi, come emblematicamente pare profilarsi in alcune proposte influenzate, per esempio, dai resoconti delle neuroscienze. Un orientamento simile tende ad emergere nei contributi del volume *Siamo davvero liberi?*, e in specie nei due saggi *Se non siamo liberi, possiamo essere puniti?*²¹ e *Come quantificare il libero arbitrio*²². In questi due interventi, al centro dei quali campeggia la fondamentale problematica del libero arbitrio, si giunge ad ipotizzare una vera e propria "minorità" destinata in sorte alla libertà umana e al suo stesso operare di giudizio. Ci troviamo nel cuore stesso dell'esercizio del diritto il quale, nel saggio di Sartori e Gnoato, vede concretamente profilarsi la possibilità d'un affidamento alle macchine del delicato tema relativo all'autodeterminazione, nell'accertamento della capacità d'intendere e di volere d'un imputato. Libertà e volizione umane, così come il giudizio dell'esperto psichiatrico-forense, paiono secondo questa prospettiva potersi devolvere a una pura "quantificazione" di stampo tecnico e statistico, realizzata come tale da strumenti standardizzati per una formulazione diagnostica.

Sul piano filosofico, infatti, se tutto vorrebbe automatizzarsi allora inevitabilmente via via ogni libero arbitrio, ogni libertà, e quindi in definitiva ogni responsabilità andrebbe a declinare relativamente alla posizione dell'essere umano. La questione è dibattuta, al momento, sotto diversi fronti, tra cui emerge attualmente ancora una volta con fermezza una certa posizione istituzionale - almeno, nelle forme - riassunta nella citata Relazione del Garante della Privacy per l'anno 2023, ma anche ribadita da un altro intervento²³ dello stesso Stanzione, nel quale il giurista sostiene a chiare lettere l'insostenibilità di qualsiasi rappresentazione "neutralista" dei mezzi cosiddetti artificiali, e la fisiologica necessità, perciò, di legiferare ancora una volta su queste materie da parte del consenso sociale umano, dalla parte di una titolarità e di una responsabilità pienamente riconfermate nel proprio

¹⁸ Si vedano in proposito alcuni degli studi e riflessioni scaturiti dalla discussione su tale caso, per esempio A. Flores, K. Bechtel, C. Lowenkamp, *False Positives, False Negatives, and False Analyses: A Rejoinder to "Machine Bias: There's Software Used Across the Country to Predict Future Criminals. And it's Biased Against Blacks"*, in *Federal Probation* 80.2 (2016) 38ss., nonché F. Donati, *Intelligenza artificiale e giustizia*, in D'Aloia (cur.), *Intelligenza artificiale e diritto* cit. 237ss.

¹⁹ C. Calì, *L'imparzialità del giudicante. Alcune implicazioni etiche dell'utilizzo dell'intelligenza artificiale in giurisprudenza*, in A. Alù, A. Ciccarello (curr.), *La pubblica amministrazione del futuro. Tra sfide e opportunità per l'innovazione del settore pubblico*, Napoli 2021, 121ss.

²⁰ Sull'opacità dell'algoritmo quale fondamentale questione, un interessante contributo proviene da O. Osoba, W. Welser IV, *An Intelligence in Our Image. The Risks of Bias and Errors in Artificial Intelligence*, Santa Monica (Calif.) 2017.

²¹ A. Lavazza, L. Sammicheli, *Se non siamo liberi, possiamo essere puniti?*, in M. De Caro, A. Lavazza, G. Sartori (curr.), *Siamo davvero liberi? Le neuroscienze e il mistero del libero arbitrio*, Torino 2010, 147ss.

²² G. Sartori, F. Gnoato, *Come quantificare il libero arbitrio*, in De Caro, Lavazza, Sartori (curr.), *Siamo davvero liberi?* cit. 167ss.

²³ P. Stanzione, *Intelligenza artificiale e decisioni politiche*, in A. Patroni Griffi (cur.), *Bioetica, diritti e intelligenza artificiale*, Milano-Udine, 2023, 215ss.

fondamentale, ineludibile ruolo. Un intero volume di interventi²⁴, d'altronde, è stato recentemente dedicato a tale, urgente necessità declinata secondo un'ottica intimamente bioetica: di fronte alla sfida dell'automatizzazione/deresponsabilizzazione umane, una rinnovata riflessione in termini bioetici diventa infatti un bisogno morale e sociale, come affrontato da diversi contributi a cominciare dall'Introduzione di Patroni Griffi, il citato contributo di Stanzione o quello di Anna Papa, così come il pregnante intervento di Alessandra Modugno, in cui i concetti di sostanza, relazione e libertà costituiscono i cardini di una “padronanza delle proprie azioni” riconducibile alla possibile esercitazione del libero arbitrio da parte umana, nella sua ampia capacità di relazione personale vissuta in termini ontologici: «il respiro della libertà, secondo eccellenza e iniziativa di bene e per il bene, si coglie anche nella capacità immaginativa e progettuale, come nella percezione dei bisogni ed enucleazione dei desideri cui ogni vita si connota, elementi tutti da considerare nel valutare le scelte morali da compiersi o compiute»²⁵.

Una visione votata all'idea d'una pluridimensionalità della persona umana, sulla quale ritorneremo anche più avanti, ma che risuona altresì nelle ferme indicazioni, concepite sotto il forte segno filosofico del personalismo, provenienti da un altro, importante osservatore come Patroni Griffi quando ribadisce la centralità della “dignità umana” nel suo poter libero scegliere, quale metavalore inaggirabile, incardinato nella stessa Costituzione²⁶. Si apre in ogni caso una indispensabile stagione di algoretica all'interno della stessa bioetica, a scongiurare una potenziale, parallela stagione di “tirannia dell'algoritmo”.

Nello stesso volume *Siamo davvero liberi?* sopra citato non mancano peraltro contributi fortemente critici verso tale tendenziale prospettiva tecnocratica, come emerge nelle riflessioni di Roberta De Monticelli, in cui l'intera prospettiva, sotto diversi aspetti, viene come ribaltata, poiché per la filosofa italiana lunghi dal doversi considerare come un percorso unilaterale quello dalla IA all'uomo si tratta invece di fare emergere in questo passaggio storico la forza d'una nuova prova e del linguaggio e delle relazioni costruttive dell'uomo nel mondo²⁷. L'incontro con “un'altra intelligenza” da parte umana chiama l'uomo stesso a un'opera di riconfigurazione personale, in una nuova impresa sulla via come dice la studiosa “del diventare persona”. Tale spunto, evidentemente, può ben dare sostegno alla prospettiva che stiamo provando a delineare, ossia quello d'un privilegio - sia teorico, sia pragmatico - da concedere alla relazione costruttiva tra uomo in mezzo, nel terreno comune d'una certa, possibile apertura nel senso dell'empatia algoritmica.

Certamente, lo specchio dell'alterità tecnologica, della macchina in cui a diverso titolo rispecchiarsi ed in qualche modo scoprirsì e riscoprirsì da parte dell'uomo, potrebbe costituire la via reale d'uno scoprimento di sé, d'un costruirsi della persona, in cui la “prova” d'una intelligenza parallela e concorrenziale a quella umana costituirebbe in questo senso l'occasione d'un percorso di arricchimento da parte di essa. La stessa chiamata alla responsabilità, sotto questo aspetto, non elimina ma, semmai, ne complica, arricchisce e riconfigura il percorso d'azione e decisione. Simile lettura richiama un altro importante spunto, stavolta venuto dal pensiero di Luciano Floridi, il

²⁴ Patroni Griffi (cur.), *Bioetica, diritti e intelligenza artificiale* cit.

²⁵ A. Modugno, *Intelligenza della realtà e azione responsabile: il “fattore umano” come meta- criterio*, in Patroni Griffi (cur.), *Bioetica, diritti e intelligenza artificiale* cit. 123ss.

²⁶ F. Patroni Griffi, *Le regole della bioetica tra legislatore e giudici*, Napoli 2016, 15s.

²⁷ R. De Monticelli, *Che cos'è una scelta? Fenomenologia e neurobiologia*, in De Caro, Lavazza, Sartori (curr.), *Siamo davvero liberi?* cit. 109ss.

cui studio su tale materia²⁸ - tra i più approfonditi - interpreta il profilo fondamentale della intelligenza artificiale quale dispositivo capace di un “*agere sine intelligere*” cui sembra tendenzialmente, se non esclusivamente indirizzato: l’essenza di tale operare artificiale infatti è da considerarsi estraneo alla complessità decisionale umana, poiché se quest’ultima è capace, con l’azione a posteriori della propria presa di coscienza, di spiegare il perché d’una decisione, la macchina non è in grado di fare altrettanto – essa, di fatto, non sa ricostruire i propri processi decisionali. Ecco allora una nuova prospettiva d’interpretazione dello scenario: il rapporto tra l’uomo e la sua macchina intelligente “capace di decidere” dovrebbe farsi intendere nei termini di un suggerimento di decisione possibile, lasciando all’essere umano una certa, definitiva capacità di “decidere di decidere ancora”. Come aggiunge a sua volta Calì, più che di IA si potrebbe a questo punto parlare piuttosto di AA, ovverossia di “*artificial agency*”, poiché è sul piano dell’*agere*, e non dell’*intelligere*, che tali dispositivi sembrano essenzialmente offrirsi e declinarsi²⁹.

Se un’insidia fondamentale viene dalla “serietà” logico-procedurale attribuita all’IA nel senso d’una automazione efficientista “competitiva” e neutralizzata in vista d’un rendimento standardizzato e ripetibile - nei termini quasi d’una ideale rappresentazione produttiva - vi è anche una minaccia “ludica” indovinabile ad insinuarsi nel più comune contesto quotidiano. Nell’interfaccia performativa uomo-computer, infatti, si sviluppa un caratteristico intreccio di funzioni, in cui il fattore utilitario, l’utility ricercata dal percorso “input-output” di domanda e risposta, sembra coniugarsi con un emergente fattore ludico, derivante da tale peculiare formula di comunicazione e “relazione” tra soggetto umano e strumentazione hardware-software informatica. La dinamica di azione- reazione connessa ai dispositivi elettronici-informatici capace di tramutarsi in funzione ludica è stata ben dimostrata in questi decenni dal grande orizzonte del mondo “gaming”, coinvolgente in profonda partecipazione le generazioni giovanili e divenuto voce significativa nel mercato dell’economia digitale, ma anche dal contesto social in genere dal consumo on-line, nel quale milioni di persone tendono a spendere quantità non indifferenti, ormai da anni, del proprio tempo personale. Ora, si vede come tale lato ludico della relazione digitale cominci a declinarsi anche nel senso della vera e propria “intelligenza artificiale”, ossia tramite l’interazione coi famosi, emergenti programmi a ciò dedicati quali ChatGpt, Gemini e simili: va rilevata perciò la crescente diffusione d’una “conversazione digitale” con l’intelligenza artificiale, da interpretarsi in quel caratteristico, originale sviluppo in cui istanza utilitaria e istanza ludica si compenetrano, nel gioco più o meno utile di domande e risposte attivato da simile conversazione. Se su di un piano psicologico andranno osservate, nel tempo, attitudini e tendenze comportamentali che tale devoluzione potrebbe indurre, prima di tutto monitorandone il non improbabile rischio di dipendenza o condizionamento - come avviene sovente con l’eccessivo legame al gioco - andrà però altresì rilevato il vero, più nascosto pericolo insinuato in simile modello di conversazione e sollecitazione info-algoritmica continua: poiché in tale sollecitazione-richiesta continua alla macchina, in questo domandare all’alterità artificiale operazioni e soluzioni nel gustoso gioco delle risposte immediate e subito orientate, si cela il potenziale rischio

²⁸ L. Floridi, *Etica dell’intelligenza artificiale. Sviluppi, opportunità, sfide*, Milano 2022. Cfr. anche Id., F. Cabitza, *Intelligenza artificiale. L’uso delle nuove macchine*, Milano 2021.

²⁹ C. Calì, *Algoritmi e processi decisionali. Alle origini della riflessione etico-pratica per le IA*, in *Scienza&Filosofia* 27 (2022) 69ss.

di un certo, progressivo habitus delegante, come in realtà già avviene sin troppo spesso con tante operazioni e microscelte quotidiane, già delegate in gran parte alle macchine.

Attraverso quindi le maglie per diversi aspetti ingannevoli della piacevolezza ludica connessa all’interpellazione “dialogante” dell’uomo coi dispositivi digitali rischia di imporsi una via parallela di rinuncia alla responsabilità, una devoluzione in termini di sintesi decisionale affidata ai dispositivi stessi, e ai loro peculiari algoritmi. Intere sequenze dei procedimenti operativi capaci di portare dalla conoscenza e selezione dei singoli dati alla loro sintesi, più analitica o spesso intuitiva, negli esseri umani vanno a devolversi e ad affidarsi così a una selezione e sintesi algoritmiche, in un processo non banale di rinuncia alla scelta, derivante da una precedente mancata selezione e mancata, successiva sintesi, in milioni di individui intenti di continuo a sollecitare i propri “device” alla ricerca di risposte o soluzioni. Incontriamo così i ben noti fenomeni³⁰ del “recommender systems”, del “microtargeting” e della “scelta per te” preordinata e programmata dall’algoritmo nel vasto mondo dei consumi digitalizzati, ossia mediati dal digitale, nel grande gioco di lettura e combinazione di “big data” capace di tradursi in guida direzionale dei consumi stessi e di tanti orientamenti della vita quotidiana, ma anche perciò a superare la frontiera dell’abitudine alla scelta e ancora una volta, così, anche quella della deresponsabilizzazione delle proprie selezioni ed orientamenti socio- culturali.

Ma proprio ragionando attorno al concetto di “selezione” potrebbe aprirsi un’ulteriore, significativa prospettiva convocabile in gioco, laddove si riconosca all’interno di esso un coacervo d’implicazioni nella chiamata in causa dell’intelligenza umana e del suo, necessario lavoro di elaborazione, visto dalla parte stavolta della pedagogia applicata e, più ampiamente, della sociologia dell’apprendimento. Osservando il fenomeno connesso alla relazione uomo-intelligenza artificiale attraverso questa ulteriore prospettiva psicopedagogica potremo così forse ancor meglio accompagnare la nostra idea di una possibile relazione empatica dell’uomo, coi suoi algoritmi, come ora vedremo nel successivo, conclusivo paragrafo di questo saggio.

3.- Empatia algoritmica e rinnovata pedagogia delle intelligenze.

Sembra sottovalutarsi, in effetti, una dimensione in realtà fondamentale che sempre entra in azione in tutti i processi di apprendimento e scambio con l’ambiente da parte degli esseri umani, e cioè quella prospettiva più eminentemente pedagogica rivolta alla coltivazione delle facoltà cognitive e alla capacità di sviluppo della conoscenza. Infatti dietro al rischio di sistematica delega - colta, come visto, nel caratteristico incrociarsi d’utilità e divertimento - nell’attuale configurarsi della relazione quasi-ludica tra uomo ed intelligenze artificiali sembra profilarsi non solo una fatale imposizione del progresso tecnologico ma anche, se non soprattutto, una più semplice crisi della capacità di selezione, analisi e sintesi negli esseri umani, da leggersi quale “crisi pedagogica” delle propensioni intellettive e cognitive diffuse di fronte al “dovere delle scelte” d’ogni giorno. Ogni scelta umana evidentemente chiama in causa un doveroso percorso di conoscenze a proprio fondamento, conoscenze studiate, comparate ed infine selezionate, ogni volta, per arrivare all’esito di una preferenza, e dunque d’una scelta; un percorso di conoscenze geografiche o culinarie, per esempio, viene in qualche modo compromesso o direttamente eluso nel sistematico attuale e diffuso affidarsi ai navigatori stradali o a certe app di indirizzo enogastronomico. La datificazione, col suo sistema di “data mining” e

³⁰ Si vedano le preoccupate considerazioni di Stanzione, *Intelligenza artificiale* cit. 216ss., nonché quelle di Numerico, *Intelligenza artificiale* cit. 260ss.

“recommender systems” impone la sua logica sfruttando - ma anche in ciò determinando - tale crisi della capacità di selezione, connessa come detto con una difficoltà del ciclo analisi-sintesi nel gioco delle cognizioni intellettive da attivare.

Se dunque si volesse sviluppare un modello di relazione tra intelligenza umana e intelligenza artificiale improntato a una certa “empatia produttiva” uno dei punti cruciali su cui agire sembra essere proprio quello d’una rinnovata pedagogia mirata alla coltivazione cognitiva, e concepita in un preciso senso costruttivistico, se non precisamente costruzionista. La dinamica di scambio tra uomo e intelligenze artificiali nel gioco tra “deep researching” e “deep thinking” oggi attivata con i software di IA può davvero mantenersi virtuosa solo se riuscirà a ritrovarsi la via d’una simile pedagogia: l’uomo “costruttore di strutture di conoscenze” del costruttivismo si coniuga con la capacità, perseguita dal costruzionismo, di farlo tramite un “feedback” con oggetti o situazioni pragmaticamente concrete, che ora sono le esperienze di risposta dinamica dell’uomo di fronte al “soggetto informatico”. L’ispirazione pedagogica qui richiamata, risalente fino a Piaget, vede ogni apprendimento come un processo inteso non secondo una banale accumulazione di dati, o lineare trasmissione di conoscenze, ma come una costante ricostruzione dei significati. Sotto diversi aspetti, il confronto avvenuto nel dibattito pedagogico in decenni lontani tra istruzionismo e costruttivismo pare oggi riproporsi al cospetto dei nuovi mezzi informatici, in un gioco delle parti dove però, adesso, il ruolo del “bambino” chiamato all’apprendimento è ricoperto dall’umanità stessa quando si colloca di fronte al mezzo da lei stessa creato capace di rispondergli. Se più generalmente una nuova pedagogia di tipo digitale-costruttivistico appare uno scenario da recuperare, poiché improntata come tale al duplice, virtuoso realizzarsi d’uno studio costruttivo-ricostruttivo quanto basilarmemente “empatico”, il concetto degli “artefatti cognitivi” proposto a sua volta dal costruzionismo sembra alquanto calzare con l’ambiente strumentale dell’interfaccia digitale, il quale si offre realmente come un “set di costruzione” in cui l’apprendimento ogni volta si concretizza, a suo modo “scambiandosi” con la realtà, generando un esito in tale dinamica ulteriormente empatico tra l’uomo, il suo strumento e le risposte giunte come “beni cognitivi” da simile interazione. L’avventura di scoperta pedagogica vissuta da uno degli allievi di Piaget, Seymour Papert³¹, può risultare molto significativa a questo proposito: l’apprendimento diventa in questa prospettiva sintonico, basandosi cioè su tre principi fondamentali da integrare, quello della continuità con le conoscenze pregresse, quello della potenza in relazione ai progetti caricati di significato, e quello della risonanza culturale pertinente alle conoscenze stesse. Attraverso questa pedagogia, non solo dunque le intelligenze artificiali sono chiamate ad apprendere, ma anche quelle umane, le quali seguendo questi principi costruzionisti nella propria relazione con i mezzi digitali sono chiamati a svilupparsi sul piano cognitivo secondo continuità, potenza, risonanza. Un percorso virtuoso che il costruzionismo e la sperimentazione di Papert hanno testato, in anni recenti, direttamente confrontandosi con il contesto informatico e la relazione con il computer, nel dichiarato obiettivo pedagogico di vedere “il bambino che programma il computer” e non “il computer a programmare il bambino”. In quest’ottica si capovolge la lineare didattica dell’imparare per usare, in un più empatico, ma anche produttivo, usare per imparare. Il “passaggio empatico” nella relazione uomo-intelligenza artificiale passa, a nostro avviso, da questo preciso crinale di rinnovata visuale pedagogica, concepita in senso dinamico e costruzionista, nella

³¹ S. Papert, *Mindstorms. Children, Computers, and Powerful Ideas*, New York 1980, tr. it. *Mindstorms: bambini computers e creatività*, Milano 1984.

quale far entrare a sua volta la pragmatica consapevolezza delle cosiddette intelligenze multiple, ossia modelli differenziati d'intelligenza applicata individuati dalla psicologia³², i quali aprono perciò a paralleli modelli differenziati nell'approccio relazionale con la stessa intelligenza artificiale. Soprattutto i “modelli di mente” sintetica e creativa possono probabilmente applicarsi a tale scenario relazionale, nel quale è auspicabile l'avvento più generalmente inteso d'una visuale multiprospettica attiva sul piano psicologico, ed in cui proprio la mente sintetica può ergersi, infine, a reale bussola capace di guidare la peculiare relazione ludica, empatica e conoscitiva dell'uomo di fronte alle sue intelligenze artificiali. La visuale del “farsi persona” nell'originale rapporto coi mezzi informatici considerata più sopra assume in tale paradigma pedagogico-costruzionista ed empatico un significato concreto, nel bambino-uomo capace di crescere assieme ai suoi mezzi, e la stessa idea del meta-criterio del fattore umano quello tecnologico trova una sua adeguata traduzione e misura di confronto. Traducendo tutto questo percorso nell'ottica d'una visuale giuridica, si vede come esso possa ben coniugarsi con i seguenti, significativi spunti concentrati attorno all'idea dello “*ius informatico*”. Proprio la forza della mente sintetica, della sua capacità di giudizio, delle peculiari qualità della flessibilità e della adattabilità psicologiche possono caratterizzare l'inimitabile capacità dell'intelligenza umana applicata al diritto. A questo proposito, offre un significativo contributo anche un recente intervento³³ di Danilo Ceccarelli Morolli, impegnato a ragionare proprio attorno alla relazione IA - intelligenza umana. Lo studioso italiano risale alla celebre affermazione di Celso, passata alla storia quale sintesi d'essenza del compito giuridico, appunto delineato dall'antico giurista quale *ars boni et aequi*, un'espressione la cui pregnanza semantica - come la definisce Ceccarelli Morolli – richiama direttamente alla “essenza ontologica” del diritto. Ma se il diritto è un'arte del buono e dell'equo, come si relaziona, in quanto tale, con l'avvento delle nuove tecnologie? È in questione, per lo studioso, la centralità stessa del giudizio umano, quel che noi abbiamo riassunto nella dialettica autonomia-eteronomia dell'interpretazione giurisprudenziale, oggi appunto posta in questione: secondo Ceccarelli Morolli, si tratta di comprendere più a fondo il concetto stesso di “equità” risalendo perciò alla definizione che ne dava Aristotele, il quale sosteneva che la chiamata in causa dell'equità fosse da relazionarsi in quanto tale con la giustizia e nella peculiare presa in considerazione del caso concreto, poiché necessaria nell'azione applicativa proprio per mitigare la rigidità in astratto delle norme e delle leggi. L'intervento umano, dunque, risulta dunque cruciale nell'esercizio del primo principio dell'arte giuridica, l'equità così inestricabilmente collegata alla concretezza dei casi e alla contestualità, anche la più complessa, delle singole situazioni. Per Ceccarelli Morolli l'altro principio-base del diritto, il *bonum*, costituisce a sua volta il fine stesso al quale dovrebbe tendere l'attività giuridica come tale, “stella polare che orienta l'interprete e l'operatore del diritto” nel loro agire e così “rendere giustizia”.

Di fatto si registra oggi un fenomeno inedito anche sul piano della elaborazione giuridica, il quale profila all'orizzonte una sorta di campo nuovo disciplinare in cui dovrebbero incontrarsi le logiche della potenza informatica con le forme del diritto, e definito dal neologismo “guscibernetica”. Ma rimane evidente, seguendo ancora Ceccarelli Morolli, come improprio una volta di più attribuire ad

³² Per esempio, nelle sintesi proposte dai lavori di H. Gardner, *Le intelligenze multiple*, Milano 2009, e W. McKenzie, *Intelligenze multiple e tecnologie per la didattica*, Trento 2005.

³³ D. Ceccarelli Morolli, *Il concetto di 'limite' nel diritto (romano). Brevi riflessioni*, in *Roma e America online* 1 (2023) 81s.

apparecchiature informatiche il peso di quella che va riconosciuta come una vera e propria *ars*, cioè lo stesso diritto: quella specifica capacità di interpretazione caratteristica dell'identità giurisprudenziale nella sua stessa essenza, e volta a comprendere le norme tramite un'applicazione flessibile e contestualizzata di esse. Come dice lo studioso, il diritto è un'arte di "sopprimere i principi etici e di bilanciare gli interessi in gioco", e i suoi rappresentanti devono perciò "sentirlo" intuitivamente, razionalmente ed emotivamente quando sono chiamati ad applicarlo, citando in ciò il contributo storico di François Gény³⁴. L'ausilio delle "intelligenze" algoritmiche andrà perciò collocato nell'ordine di quello che esso realmente costituisce, ovverosia uno strumento, un instrumentum semplicemente a disposizione della decisione umana, come già avviene con tanti altri strumenti, "protesi" che la storia dell'umanità si è messa a propria utilità - una dimensione protesica, in accordo con una parallela dimensione mimetica nel rapporto con le tecnologie, come detto poco più sopra. Cosicché il dilemma andrà superato per così dire ritornando alla base stessa delle due concezioni, quella strumentale e quella critico-interpretativa, e di conseguenza rilevando l'ineluttabilità della seconda nella gestione sostanziale della prima. La visione sulla IA di Ceccarelli Morolli viene così a coincidere con quella sopra considerata di Floridi, vista quale *auxilium* d'azione non al di sopra, ma piuttosto a sostegno dell'umano procedere.

Ricomponendo arricchendolo questo quadro interpretativo l'originale lettura proposta da Bernard Stiegler, nel suo dialogo con Antoinette Rouvroy³⁵: per il filosofo francese, il regime della "governamentalità algoritmica" si basa sul rischio d'un equivoco, quello d'una eliminazione della *ragione* a favore di un esclusivo dominio dell'intelletto, per dirla in termini kantiani, intelletto chiamato a diventare una "macchina di calcolo" nella visione leibniziana. In realtà, se l'intelletto opera da macchina computazionale in termini analitici, ciò che infine raccoglie i dati al fine di discernere è invece la ragione, ossia la capacità di prendere delle decisioni, di fare critica, di lavorare su delle distinzioni - e, in coerenza con quello sostenuto più sopra a proposito della mente sintetica, di ricomporre dei processi sintetici, divenuti razionali perché, finalmente, non solo analitici. Su tale distinzione fondamentale tra intelletto e ragione si costruisce l'altrettanto fondamentale differenza tra fatto e diritto, laddove l'esercizio della giurisprudenza e quindi del diritto non sarà mai riducibile alla pura rappresentazione e semplice collocazione del fatto. La distinzione tra fatto e diritto richiama altresì quella tra essere e dover essere, come ricordato e sottolineato da un intervento di Tommaso Edoardo Frosini³⁶: il rischio di confonderle, invece che di tenere differenziate queste posizioni, va una volta di più denunciato, specie di fronte alle pretese d'un nuovo, presunto statuto del "giurista tecnologico".

Infine, può chiudere questo percorso un ulteriore, illuminante spunto ereditato da un altro importante pensatore contemporaneo, Hans Jonas, dal suo volume sull'orlo dell'abisso. Per Jonas, impegnato da sempre a pensare un accordo tra ragione, libertà e responsabilità, la moderna tecnologia informatica costituisce "un'attività meccanica eseguita prima da un essere vivente e poi da una macchina", laddove la possibilità di attribuire gli qualcosa di simile alla consapevolezza critica umana

³⁴ F. Gény, *Metodo di interpretazione e fonti del diritto privato positivo*, a cura di G. Tarello, Torino 1955.

³⁵ A. Rouvroy, B. Stiegler, *Il regime di verità digitale. Dalla governamentalità algoritmica a un nuovo Stato di diritto*, in *La Deleuziana* 3 (2016) 6ss.

³⁶ T.E. Frosini, *L'orizzonte giuridico dell'intelligenza artificiale*, in Patroni Griffi (cur.), *Bioetica, diritti e intelligenza artificiale* cit., 201ss. Importanti pure le riflessioni portate da A. Simoncini, *L'algoritmo incostituzionale: intelligenza artificiale e il futuro delle libertà*, in D'Aloia (cur.), *Intelligenza artificiale e diritto* cit. 167ss.

e del tutto fuorviante, perché «chi considera seriamente questa ipotesi paragona anzitutto il calcolo al pensiero e poi il pensiero alla coscienza». Infatti, bisogna comprendere bene questo aspetto: «che l'automa possa divenire attraverso l'esecuzione dei compiti esso stesso vivente, acquisire una psiche con una propria volontà non è altro che una sfrenata speculazione» (...) Come non si può ascrivere sensibilità ad un termostato perché è sensibile agli stimoli esterni, tantomeno abbiamo motivo di ascrivere il pensiero ad un processo meccanico perché, predisposto da noi a farlo, opera in modo corrispondente al pensiero»³⁷.

Jonas, dunque, distingue il calcolo dal pensiero, così come il pensiero stesso dalla coscienza: l'opera del computer così si rivela per quello che è, ossia un'operazione di calcolo capace di realizzare con grande solerzia i passi meccanici necessari d'un processo - ma, in realtà, destinata tale stessa operazione a essere sempre poi assunta da un atto di coscienza, da una coscienza che, sola, ne dà il significato.

Abstract.- Il presente contributo prova a ricomporre lo scenario delle finalità talvolta differenti, se non conflittuali, all'origine delle tecnologie di intelligenza artificiale. Ri emerge dal confronto originario tra prospettiva linguistica e visuale cibernetica un duplice, tendenziale esito storico, delle intelligenze artificiali quali spazi d'affidamento tecnocratico preteso, capace di superiori automatismi e di neutralità delle decisioni, oppure quali territori d'uno scambio virtuoso, concepito in quanto *auxilium*: nella visuale giuridica, perciò, *ars boni et aequi*. Seguendo allora questa seconda prospettiva, il contributo prova a elaborare ed offrire una categoria originale di relazione, quella di empatia algoritmica, pensata come “spazio pedagogico” ludico-educativo di scambio continuo tra le macchine e l'umanità che le utilizza, dove infine è la seconda a dover sempre controllare la prima - a maggior ragione, nell'azione giurisprudenziale e del diritto, dove ancora, oggi come sempre, dovrà ribadirsi una distanza tra fatto e diritto, tra l'essere e il dover essere nell'esercizio quotidiano giuridico.

This paper attempts to recompose the scenery of the sometimes different, if not conflicting, purposes at the origin of artificial intelligence technologies. From the original comparison between linguistic perspective and cybernetic vision, a dual, tendential historical outcome emerges, of artificial intelligences as spaces of claimed technocratic trust, capable of superior automatisms and neutrality of decisions, or as territories of a virtuous exchange, conceived as *auxilium*: in the legal vision, therefore, *ars boni et aequi*. Following this second perspective, the contribution tries to elaborate and offer an original category of relationship, that of algorithmic empathy, conceived as a playful-educational “pedagogical space” of continuous exchange between machines and the humanity that uses them, where ultimately it is the latter that must always control the former - even more so, in the action of jurisprudence and law, where still, today as always, a distance between fact and law, between being and having to be in the daily exercise of law, must be reaffirmed.

³⁷ H. Jonas, *Sull'orlo dell'abisso. Conversazioni sul rapporto tra uomo e natura*, Torino 2000, 45ss.