

Markus Krienke*

LA PERSONA AL CENTRO
DELL'INTELLIGENZA ARTIFICIALE.
OLTRE „LO SCACCO DELLE LEGGI DI ASIMOV”**¹

„L'universo fisico classico esiste per la stessa ragione
per cui noi creiamo realtà virtuali con i computer:
per conoscerci meglio”².

„La scoperta scientifica più importante
fu la scoperta dell'ignoranza”³.

Abstract: L'autore in dialogo virtuale e trasversale con diverse posizioni, riflette su come affermare nell'ambiente dell'IA l'unicità e irriducibilità della dignità della persona, nel tentativo di trovare un modo etico-istituzionale per affermare nell'ambito dell'infosfera ciò che “al di fuori di essa” è la realtà della persona come coscienza incarnata. Di fronte allo sviluppo dell'IA abbiamo bisogno „niente di meno che di un'antropologia filosofica in linea coi nostri tempi, ovvero, una comprensione filosofica della natura umana che sia adeguata all'era digitale e alle nostre società dell'informazione”.

Keywords: Intelligenza Artificiale, infosfera, persona, etica.

* Markus Krienke – Facoltà di Teologia di Lugano

** „La Società” 2023 n 1 p. 64-77.

¹ M. Ferraris, *L'esplosione della registrazione*, in: L. Taddio - G. Giacomini (edd.), *Filosofia del digitale*, Mimesis, Milano-Udine 2020, pp. 91-108, qui p. 107.

² F. Faggin, *Irriducibile. La coscienza, la vita, i computer e la nostra natura*, Mondadori, Milano 2022, p. 233.

³ Y.N. Harari, *Homo Deus. Breve storia del futuro*, tr. it. M. Piani, Giunti-Bompiani, Firenze-Milano 2018, p. 263.

1. IL PROBLEMA

Con il dubbio se le famose *leggi di Asimov* hanno portato la filosofia, l'antropologia e l'etica del digitale sulla pista fuorviante di un'implicita antropomorfizzazione delle nuove tecnologie, Ferraris ha senz'altro formulato una sfida non trascurabile per chi se ne occupa in una prospettiva personalistica. A tale confronto si dedica ormai da alcuni anni anche il „gruppo 6” – „Persona e intelligenza artificiale” – dell'associazione *Persona al Centro*⁴. I dispositivi dell'IA non hanno certo intenzionalità e per questo non sono „intelligenti” nel senso umano della parola – e ovviamente il gruppo 6 condivide pienamente questo punto di partenza. Antropomorfizzarla dunque no, ma personalizzarla, e come? La vera differenza specifica tra i due tipi di „intelligenza”, infatti, non sta nell'ambito dell'intelligenza – semplicemente perché nel caso delle macchine non si può parlare di „intelligenza” nella stessa accezione umana del termine – ma nella differenza tra vita digitale e vita biologica⁵. Infatti non a caso la specificità della vita biologica è tornata sistematicamente al centro dei dibattiti all'interno del gruppo di ricerca. E proprio in quanto il digitale si inserisce nell'ambito della vita umana che da sempre si è anche definita come tecnica, e come ogni tecnologia riserba in sé opportunità e rischi per la *giustizia sociale*, gli specifici rischi derivanti da questo tipo di tecnologia costituiscono le preoccupazioni che nutriamo nei suoi confronti.

Oltre questo punto di partenza ontologico, bisogna quindi innanzitutto indagare circa gli effetti della nuova tecnologia sul soggetto in quanto persona cioè nella sua realizzazione sociale. In tale direzione, molti studi puntano sul rischio della libertà e l'ingerenza di potere (economico, politico ecc.), con l'effetto di creare situazioni sociali di ingiustizia: „[i]l soggetto sottomesso nel regime dell'informazione non è docile né ubbidiente. Piuttosto si crede *libero, autentico e creativo: produce e performa se stesso*”⁶. Questa autoproduzione e la messa in scena di sé stessi rischia di portare poi a dipendenza e controllo: „[l]a sorveglianza s'insinua nella quotidianità sotto forma di convenienza”⁷. Il rischio fondamentale per la persona,

⁴ Cfr. il contributo *La persona non è morta. Viva la Persona al centro*, in: *La Società* 30 (2021), n. 4, pp. 176-180. Il presente contributo non è né un riassunto dei dibattiti del „gruppo 6” che non potrebbero essere contenuti in poche pagine, né la formalizzazione di un punto preciso „di arrivo momentaneo” delle riflessioni, in quanto nella molteplicità delle prospettive e anche dei metodi sono difficilmente sintetizzabili. Si tratta piuttosto di una riflessione personale dell'autore in dialogo virtuale e trasversale alle varie posizioni che gli sono sembrate emergere nei tanti mesi di dibattito ed interazione, e quindi a sua volta un contributo per portare avanti tale lavoro in gruppo.

⁵ In questo senso, Floridi individua la differenza specifica tra coscienza umana specificando che „la risposta non è: l'AI non potrà mai diventare cosciente; ma invece: l'AI sarà cosciente come noi se avrà una vita mentale come noi. E questo, per quanto ne sappiamo, è veramente solo fantascienza» (L. Floridi, *In poche battute. Brevi riflessioni su cultura e digitale 2011-2021*, p. 138).

⁶ B.-C. Han, *Infocrazia. Le nostre vite manipolate dalla rete*, tr. it. F. Buongiorno, Einaudi, Torino 2023, p. 4.

⁷ *Ibid.* p. 10.

infatti, consiste proprio nella modificazione della sua autocomprensione e nella modificazione eteronoma della sua autorealizzazione – e sorge l'impressione che molte questioni di *algoritica* o dell'etica del digitale si collocherebbero rispetto a questa domanda fondamentale della persona in secondo rango. Il criterio della *persona al centro*, indica quindi piuttosto un metodo e un ordine di questioni per *come* si affronta la sfida del digitale. Del resto, secondo Ferraris, l'ingente produzione dei dati – si pensi soltanto che secondo studi difficilmente verificabili ma secondo lui ragionevoli negli ultimi due anni si è prodotto il 90% dei dati digitali attualmente disponibili – non rende il mondo più fluido o liquido ma al contrario più solido, proprio perché „salvato” e quindi tracciato. Floridi parla pertanto di una vera e propria ri-ontologicizzazione e ri-epistemologicizzazione del mondo in atto⁸. È questo il modo attraverso cui la tecnologia incide sulla persona: cambiandole completamente l'ambiente, ed essendo la persona relazione, la ripercussione non può che essere sull'autocomprensione della persona stessa. Ciò che deve essere approfondito, quindi, non è tanto il processo tecnologico – anche se senza un sapere materiale dell'oggetto è difficile immaginare una „filosofia del digitale” – quanto l'indagine sul „sostrato umano, il contesto storico, sociale e culturale, ma soprattutto i valori e i modelli ideali con i quali tali cambiamenti dovrebbero confrontarsi”⁹.

Per mettere questo secondo argomento – che si potrebbe chiamare sociale – sul punto si potrebbe quindi dire che il digitale viene ad incidere sull'ambiente che la persona abita e che pertanto la definisce, attraverso le relazioni e interazioni fondamentali della persona stessa¹⁰. In questo modo, la sfida digitale che costringe ad un „allargamento» dell'etica¹¹, come formula Fabris, cioè all'esigenza di non ridurre l'etica sulla questione morale dell'agire. Il potere sull'informazione si traduce in questo modo in un potere sulle forme della realizzazione della persona: come descrive Ferraris, la „svolta copernicana” del Web è il passaggio dall'„eido-centrismo” all'„etocentrismo”: consiste nel „raccogliere, diffondere e capitalizzare” non tanto l'idea quanto „forme di comportamento”¹² che del resto risultano anche

⁸ In termini metafisici, si tratta per lui di un passaggio dalla „*Ur-filosofia* aristotelico-newtoniana” a una „*relazionale*” (L. Floridi, *Il verde e il blu. Idee ingenue per migliorare la politica*, Raffaello Cortina, 2020, pp. 45-46).

⁹ L. Taddio – G. Giacomini, *Introduzione*, in: idd. (edd.), *Filosofia*, cit., pp. 17-23, qui p. 18.

¹⁰ Ha pertanto ragione Harari a dire che „[l]e idee cambiano il mondo soltanto quando cambiano il nostro comportamento” (*Homo Deus*, cit., p. 476).

¹¹ A. Fabris, *L'orizzonte umano nell'epoca della rivoluzione digitale*, in: *Annali del Centro Studi Filosofici di Gallarate* 1 (2021), n. 1-2, pp. 53-62, qui p. 54. „Si tratta [...] di un processo d'interazione, complesso e articolato. Su questo terreno avviene l'allargamento del campo dell'etica – vera e propria novità relativa a tale disciplina – al quale stiamo assistendo” (Ibid. p. 60).

¹² M. Ferraris, *L'esplosione*, cit., p. 106; cfr. Id., *Documanità. Filosofia del mondo nuovo Laterza*, Roma-Bari 2021. La stessa cosa afferma Benanti quando sottolinea che „gli artefatti che l'uomo

sempre più de-relazionalizzate o dis-intermediate¹³. In questo modo, l'IA „non mira alla comprensione e a ciò che ne deriva, finalità e volontà di potenza comprese, bensì alla automazione di processi”¹⁴, e proprio per questo un approccio etico che è determinato dall'idea di risolvere la sfida del digitale con le „leggi di Asimov” è strutturalmente limitato.

Esattamente la stessa prospettiva si trova nella proposta floridiana circa un'*infraetica* („infrastruttura etica”)¹⁵: è solo tramite un'etica dell'infrastruttura – etica sociale – che si riesce a tematizzare la persona non solo a livello teoretico ma soprattutto etico-pratico nei confronti delle nuove sfide tecnologiche. Lo specifico dell'etica nell'infosfera consisterà, secondo la caratteristica della tecnologia di cui si tratta, nel fatto che il potere non si determina attraverso il possesso dei mezzi di produzione, ma il possesso e il controllo delle informazioni. E qui siamo in una terza dimensione, di carattere etico, della sfida in atto.

Al fine di poter esplicitare meglio la prospettiva persona nell'ambito dell'intelligenza artificiale come emerge dalla triplice dimensione ontologica, sociale ed etica (riassumibile in uno sguardo antropologico generale), nell'ambito digitale, bisogna certamente orientarsi ad un modello che – come già accennato – non è tecnologico (sebbene la conoscenza tecnologica è indispensabile) ma in grado di inquadrare il rapporto tra tecnologia e persona. Il gruppo 6 di *Persona al Centro* si è concentrato – in modo senz'altro critico – pertanto su quello proposto di Federico Faggin dal nome *Irriducibile*, ipotizzando un confronto di tale prospettiva „idealistica” con un contro modello „realista” per il quale qui si propone – per motivi di molti temi in comune – *Homo Deus* di Harari.

2. IL CONFRONTO CON FAGGIN

Per inquadrare la prospettiva di Faggin, bisogna partire dal giudizio complessivo di Pianca secondo cui la scienza dell'IA „si è nutrita per decenni” da „materialismo monista, fisicalismo e meccanicismo”¹⁶. L'esperienza interiore e la coscienza sono fenomeni *non algoritmici*, e Faggin intende dimostrarlo facendo comprendere

produce generano e mediano una cultura” (P. Benanti, *Digital age. Teoria del cambio d'epoca. Persona, famiglia e società*, San Paolo, Cinisello Balsamo (MI) 2020, p. 51).

¹³ „L'impoverimento delle relazioni sociali [...] ha come contropartita per l'individuo [...] una perdita di sostanza” perché „l'umano richied[e] l'apertura agli altri per sviluppare la propria soggettività” (J.-M. Besnier, *L'uomo semplificato*, tr. it. D. Frontini, Vita e Pensiero, Milano 2013, p. 53).

¹⁴ Ferraris, *L'esplosione*, cit., p. 108.

¹⁵ L. Floridi, *Ipersotira, l'emergenza dei sistemi multiagente e il design dell'infraetica*, in: L. Taddeo - G. Giacomini (edd.), *Filosofia*, cit., pp. 141-169, qui p. 163.

¹⁶ F. Pianca, *La narrativa dell'intelligenza artificiale: soggettività etica e riduzionismo antropologico*, in: L. Sandonà - F. Zaccaron (edd.), *Noi techno-umani. Pandemia, tecnologie, comunità (antropologica 2020)*, Meudon, Trieste 2021, pp. 83-95, qui p. 87.

come esse possono essere descritte soltanto tramite la fisica quantistica e non quella classica che invece si applica a stati causati nel mondo fisico esteriore. Ciò non significa, però, che l'esperienza della coscienza stia nel modello quantistico il quale serve semplicemente a „spiegarla” (non „comprenderla”): oltre la quantistica come modello base di descrizione, c'è quindi un'eccezione a sua volta non rappresentabile in un modello razionale.

Faggin avanza la prospettiva – del resto non sorprendente a chi ha studiato il neoplatonismo, l'idealismo tedesco o altre espressioni di monismo spiritualistico – che la coscienza è precedente ai sistemi complessi fisico-naturali: lungi dall'essere una proprietà emergente di un sistema complesso, è al contrario ciò da cui emerge un sistema complesso. Per lui, è precisamente „una proprietà fondamentale, al pari dell'elettricità che non può sorgere da particelle elementari”, per cui „tutto nell'universo deve allora essere cosciente”¹⁷. Questo „*panpsychismo*” – presentato in riferimento a D'Ariano – vede il mondo biologico non come epifenomeno, ma come avente un lato interiore che è l'esperienza della vita che si esprime, appunto, nella stessa realtà esteriore. Nel suo modello QIP (*Quantum Information-based Panpsychism*) la realtà viene sperimentata, nella coscienza che consiste di base nella „*seity*”¹⁸ ossia coscienza di sé, secondo i suoi *qualia* che sono gli specifici contenuti dell'esperienza coscienziale. Mentre la natura fisica viene descritta dalla teoria classica¹⁹, quella quantistica fa capire la dimensione della „*seity*” ma solo in chiave rappresentativa e descrittiva, non „sostitutiva”, come Faggin si esprime, e forse nemmeno esplicativa e comprensiva. Ciò lo porta all'affermazione finalistica tipica del *panpsychismo* immanentista che – se la coscienza sta all'inizio e non come un fenomeno emergente alla fine – „ci dev'essere un Principio creativo che dia scopo, significato e direzione all'Universo. Penso che tale Principio possa essere questo: *Uno vuole conoscere se stesso per autorealizzarsi*, e quindi per gioire della propria esistenza e amarla. Pertanto la coscienza, il libero arbitrio e la vita rappresentano i mezzi necessari, ovvero le proprietà che Uno deve possedere, per conoscere sé stesso”²⁰. Così le *seity* – al plurale in quanto esse producono le esperienze

¹⁷ Faggin, *Irriducibile*, cit., p. 132.

¹⁸ Essa „implica che la nostra esistenza non dipende dal corpo. Il corpo permette alla *seity* di percepire e di operare in un mondo fisico che è solo una piccola porzione della realtà più vasta in cui esiste” (Ibid. p. 159).

¹⁹ „un sistema classico è deterministico e riduzionistico, e quindi il suo comportamento è la somma dei comportamenti delle sue parti, e non può esistere un tutto che sia indipendente dalla somma delle parti. È questa la ragione per cui un sistema classico potrebbe essere cosciente soltanto se i suoi atomi fossero coscienti” (Ibid. pp. 197-198).

²⁰ Ibid. p. 169. Non a caso, il riferimento è a Schelling, e sfocia nell'affermazione che „riproduzione crea generazioni di *seity* che, ampliando la loro conoscenza, espandono la conoscenza di Uno, dato che Uno non è separato dalle sue emanazioni. *Uno è l'interiorità di tutto ciò che esiste; è ciò che connette «da dentro» tutte le sue creazioni*” (Ibid. p. 171).

individuali – comunicano e producono *sintatticamente* le leggi fisiche, per cui il libero arbitrio risulta difeso²¹.

Ciò porta alla conseguenza che „la coscienza non può magicamente emergere dagli algoritmi, ma è già presente nei campi delle particelle elementari di cui tutto è fatto”²². La coscienza (umana) è poi in grado di trasformare l’informazione in conoscenza semantica: ciò che succede a livello semantico, infatti, è quindi sempre già *dentro* la coscienza, e solo la dimensione „prevedibile” o „automatizzabile” può essere demandato a una macchina mentre „la vera intelligenza è la capacità di giudicare correttamente una situazione e di trovare una soluzione creativa qualora la situazione la richieda. La vera intelligenza richiede *comprensione*”²³.

Accanto alla coscienza, anche la vita non si spiega se non attraverso questo lato interiore di esperienza che la coscienza ha aperto: in questa prospettiva, essa è „la strategia fondamentale con cui Uno può conoscere se stesso”²⁴. In tale contesto, la probabilità acquisisce un significato fondamentale per la teoria di Faggin: a differenza di un monismo deterministico, il modello quantistico inserisce la probabilità – una categoria della coscienza e non della fisica – in modo costitutiva nella realtà, anziché ridurla, come succede nel determinismo, alla semplice mancanza di sufficienti informazioni²⁵. Così diventa nuovamente chiaro come „la materia è l’aspetto simbolico di enti coscienti che comunicano, ossia trasmettono significato tra di loro”²⁶. Della macchina, infatti, possiamo „comprendere” solo il comportamento, non possiamo „vedere” intelligenza, per cui è un determinato comportarsi o un determinato funzionamento a cui attribuiamo il nostro termine „intelligenza”.

3. IL CONFRONTO CON HARARI

Secondo la visione tecno-umanista di Harari, il potenziamento tecnologico dell’uomo porta il capitalismo alla sua piena realizzazione. L’„umanesimo evolutivista” condurrà alla creazione di „narrazioni più potenti e religioni [forse meglio: teologie politiche] più totalitarie che in qualsiasi epoca precedente”²⁷, tra cui per Harari fungono anche capitalismo e comunismo, entrambi basati sulla crescita

²¹ Cfr. Ibid. pp. 163-164 e 174.

²² Id., *Silicio. Dall’invenzione del microprocessore alla nuova scienza della consapevolezza*, Mondadori, Milano 2019, p. 319.

²³ Ibid. p. 321.

²⁴ Id., *Irriducibile*, cit., p. 211.

²⁵ Cfr. Ibid. pp. 216-217.

²⁶ Ibid. p. 231; nell’orig. in corsivo.

²⁷ Y. N. Harari, *Homo Deus*, cit., p. 219.

economica²⁸: si tratta semplicemente di sistemi diversi di elaborazione dei dati²⁹. Pertanto, la scienza potrebbe – aggirando l’immaginario fantascientifico dei robot umanoidi – condurre alla superintelligenza bypassando la coscienza: „[p]er milioni di anni l’evoluzione organica ha navigato lentamente lungo la rotta della consapevolezza. L’evoluzione dei computer inorganici potrebbe aggirare completamente questi stretti vincoli”³⁰. Il tecno-umanesimo diventa datismo dal momento in cui nell’evoluzione tecnologica non rimane più nulla dell’esperienza umana, riconnessa a volontà e libertà, e si riconosce che „l’universo consiste di flussi di dati e che il valore di ciascun fenomeno o entità è determinato dal suo contributo all’elaborazione dei dati”³¹. La posizione radicalmente opposta a quella di Faggin emerge proprio qui, quando Harari spiega il significato del datismo ossia che „la Quinta sinfonia di Beethoven, una bolla finanziaria e il virus dell’influenza sono soltanto tre pattern di un flusso di dati che può essere analizzato usando gli stessi concetti di base e gli stessi strumenti”³².

Con questa teoria Harari elabora in realtà anche il nesso tra informazioni e capitalismo, ricorrendo alla teoria di Hayek, secondo il quale il capitalismo è un sistema decentralizzato di governance delle informazioni, e per questo è più efficace e di successo del comunismo. Per lo stesso motivo non solo le democrazie superano secondo Harari i sistemi autocratici, ma sono addirittura destinate a scomparire per inefficienza (essendo legate a processi partitici e parlamentari lenti)³³: infatti – come conferma anche Floridi – la dinamica dell’internet supera qualsiasi confine politico e giuridico, caratterizzato dal „sistema di Westfalia”³⁴. Sistemi di controllo e anche di discriminazione più efficaci dei totalitarismi politici diventano così possibili³⁵, e questa minaccia porta per la democrazia l’urgenza di „reinventarsi in una forma radicalmente nuova” per evitare che „gli esseri umani finiranno per vivere in una «dittatura digitale»”³⁶. L’evoluzione tecnologica, infatti, porta il datismo – la gestione dei dati con metodo capitalistico – al rango di una „nuova religione” di cui il „valore supremo” è il „flusso di informazioni” che però non produce nessun „*Homo Deus*” né conferisce alcuna dignità speciale all’umano bensì vede gli umani come „meri strumenti per creare «Internet-di-Tuttele-Cose»” per realizzare un „sistema cosmico di elaborazione dati” il quale sarebbe „come

²⁸ Per Harari religione è „una qualsiasi narrazione globale che conferisce legittimità oltreumana a leggi, norme e valori umani. Essa legittima le strutture sociali esistenti con l’argomento che esse riflettono leggi che trascendono gli uomini storicamente determinati” (Ibid. p. 224).

²⁹ Cfr. Ibid., p. 451.

³⁰ Ibid., p. 379.

³¹ Ibid., p. 449.

³² Ibid., p. 450.

³³ Cfr. Ibid., pp. 452 e 456-457.

³⁴ Cfr. L. Floridi, *Il verde*, cit., p. 27.

³⁵ Cfr. Y.N. Harari, *21 lezioni per il XXI secolo*, tr. it. M. Piani, Ariccia (Roma) 2020, p. 104.

³⁶ Ibid., p. 104.

Dio". L'uomo si distingue dagli animali per il fatto che in lui „l'informazione fluisce secondo schemi molto più complessi” rispetto agli animali³⁷. Da questa prospettiva risulta l'imperativo che „[n]on dobbiamo lasciare alcuna parte dell'universo disconnessa dalla grande rete della vita”, e il „più grande peccato” sarebbe appunto l'impedimento del flusso dei dati³⁸.

Nella misura in cui lasciamo alle spalle un tradizionale concetto di proprietà privata, perché capiamo che al posto dell'averci interessa partecipare a dei servizi (ad esempio il mondo dello sharing), avviene la realizzazione della trasformazione della realtà in dati, perché le auto condivise (per fare un esempio) devono trovarsi a un determinato momento in un certo luogo, a secondo delle nostre intenzioni di utilizzo. Nello stesso momento, gli individui diventano parte „di qualcosa di molto più grande” di loro stessi, cioè del flusso di dati, e si realizza la dimensione religiosa del datismo. „La religione dei dati ora dice che ogni parola e ogni azione è parte del grandioso flusso dei dati, che gli algoritmi vi stanno guardando costantemente e che essi si preoccupano di qualsiasi cosa facciate e di qualsiasi sentimento proviate”³⁹. A questo punto qualsiasi progetto umanistico, afferma Harari, sarà irrilevante e ogni antropocentrismo sarà superato.

La consapevolezza a cui ci porta il progresso tecnologico e i benefici che da esso emergono, si esprime anche dal fatto che ad esempio in epoche precedenti, anche pre-tecnologiche, l'umanità estinse tantissime specie animali, senza accorgersene⁴⁰. Mentre quindi tale estinzione non è da attribuire originariamente al progresso tecnologico, è anzi esso stesso che promette di poter risolvere questa e altri problemi umanitari, da „la penuria di cibo, la diffusione delle malattie e i conflitti bellici” fino alla morte stessa⁴¹. E se è proprio vero che la „vera nemesis della moderna economia è il collasso ecologico”⁴², è altrettanto chiaro che per Harari a questo punto si evince l'effettiva mancanza di alternative al finalismo datistico dell'evoluzione dell'umanità⁴³.

4. SEITY O DANTE? PERSONA E DIGNITÀ UMANA NELL'INFOSFERA

Entrambi i nostri autori sopravvalutano la tecnica o meglio una spiegazione tecnoscientifica della realtà – l'altra faccia del loro riduzionismo tecnologico che è stato criticato da molte parti e anche dal gruppo 6. Anche una spiegazione quantistica

³⁷ Id., *Homo Deus*, cit., p. 466.

³⁸ Ibid., p. 467.

³⁹ Ibid., pp. 471-472. Cfr. Benanti, *Digital age*, cit., pp. 90-93.

⁴⁰ Cfr. Y.N. Harari, *Homo Deus*, cit., p. 97.

⁴¹ Ibid., p. 7.

⁴² Ibid., p. 263.

⁴³ Mentre Harari usa la parola „datismo”, è Benanti a parlare di „dataismo”, in riferimento allo stesso Harari, ma „giocando sull'assonanza con *dadaismo*” (Benanti, *Digital age*, cit., p. 90).

della realtà, resta tale, sebbene sappia aprire decisamente più possibilità di dialogo con le discipline umanistiche rispetto ad altri presupposti di carattere della teoria della fisica. Tuttavia, vale anche per Faggin che „[o]gni concezione della natura umana che pretenda di descriverla in termini positivi (e quindi, implicitamente, di normarla) porta a risultati paradossali e indifendibili sul piano dell’esperienza”⁴⁴. Bisogna quindi come secondo risultato di questo confronto, inquadrare bene la sfida: non con le visioni apocalittiche o fantascientifiche di future superintelligenze, ma circa le ripercussioni che le tecnologie digitali effettuano già oggi sull’agire della persona e sulla sua autocomprensione⁴⁵. E su questo livello, tale tecnologia va affrontato nel modo in cui si sono gestite anche le trasformazioni sociali del passato, cioè tramite regolamenti giuridici, che certamente si basano a loro volta su una visione antropologica ed etica.

Ciò su cui deve quindi vertere il confronto personalistico con questi modelli etici dell’IA, tra Faggin e Harari – pur nella grande differenza dei loro approcci – è come affermare nell’ambiente dell’IA quello che pur tutti e tre cercarono di esplicitare, ossia l’unicità e irriducibilità della dignità della persona in un medio per definizione antropo-decentrico⁴⁶. In altre parole, bisogna trovare un modo etico-istituzionale per affermare nell’ambito dell’infosfera ciò che „al di fuori di essa” è la realtà della persona come coscienza incarnata. Dunque, quello che emergerà sarà probabilmente una nuova articolazione dei diritti, della politica, della giustizia sociale, dell’ordine internazionale, che ha la potenzialità di essere più funzionale rispetto alle sfide del futuro che ci attendono. Forse non sarà nemmeno così diversa da quanto ci immaginiamo e da come è l’assetto istituzionale attuale nel „mondo analogo”. Altri strumenti saranno nuovi cogliendo la specificità della sfida, a partire dalla garanzia e da un’etica della privacy: proprio dalla questione del diritto incentrato sulla persona inizia ogni etica sociale istituzionale. In tale modo, gli strumenti e le istituzioni nuove contribuiscono senz’altro anche a formare una nuova *consapevolezza* umanitaria e creano un livello di *problem solving* universale come nessuna politica e tecnologia finora è riuscita. Tali regole, incentrate sulla persona, devono inoltre essere di base *liberali* e giustificate *democraticamente*, per non portare a nuove forme di statalismo, di cui anche una società dell’infosfera non è per sé sicura⁴⁷. Una prospettiva personalistica si traduce, quindi, meno in „leggi etiche per la macchina” come per Asimov o in un’etica traducibile in algoritmi:

⁴⁴ A. Caronia, *Dal cyborg al postumano. Biopolitica del corpo artificiale*, a cura di L. Borrelli e F. Malagnini, Meltemi, Milano 2020, p. 234.

⁴⁵ Oddenino descrive come „gli algoritmi sono divenuti una mediazione continua e sostanzialmente necessaria di ogni attività” (A. Oddenino, *Intelligenza artificiale, potere computazionale, tutela dell’umano*, in: *Crederci Oggi* 42 [2022], n. 252, pp. 53-66, qui p. 55).

⁴⁶ Cfr. G. Razzano, *Il primato dell’essere umano nell’era dell’intelligenza artificiale*, in: A. Patroni Griffi (ed.), *Bioetica, diritti e intelligenza artificiale*, Mimesis, Milano-Udine 2023, pp. 103-121.

⁴⁷ Cfr. V. Frosini, *L’orizzonte giuridico dell’intelligenza artificiale*, p. 208.

una macchina non sarà mai un soggetto morale e capace di fungere come soggetto di dignità, diritti o doveri. Ciò che siamo chiamati a fare, è definire le norme (giuridiche) sociali per la gestione dell'intelligenza artificiale e per decidere quale spazio vogliamo dare a tale innovazione tecnologica nella nostra società, compatibilmente con la nostra idea di persona: infatti, come sottolinea anche Harari, „gli umani [...] impongono vincoli severi al loro utilizzo [delle nuove tecnologie]”, come nell'esempio del rifiuto del movimento eugenetico⁴⁸.

Si acquisisce quindi la consapevolezza che l'affermazione della dignità a livello politico-sociale ossia istituzionale, non può fare a meno delle regole giuridiche, e che significativamente „la minore presenza di regole a protezione dei dati e della privacy si registra in Paesi con regimi che si possono definire meno democratici”⁴⁹. In un'analisi trasversale Floridi approfondisce i sei documenti più importanti attualmente esistenti che hanno proposto una lista di principi etici nell'affrontare l'evoluzione dell'IA e il suo impatto sulla società, rilevando che nei 47 principi complessivi dimostrano una significativa convergenza su quattro principi base: beneficenza, non-maleficenza, autonomia e giustizia, ai quali Floridi ritiene indispensabile aggiungere quello della „esplicabilità”⁵⁰. Inoltre, egli specifica che ciascuna iniziativa è espressione di „global scope” oppure sono stati formulati all'interno di „democrazie occidentali”⁵¹. Ovviamente sono senz'altro nella linea dei principi etici per un'„IA affidabile” proposti dall'Unione Europea che a sua volta ha influito sui principi dell'OECD e i 42 Paesi qui connessi.

EPILOGO

Nonostante tutte le „svolte” o “rivoluzioni” che sono state associate con l'avvento dell'IA, non bisogna dimenticare che di base l'IA „deve essere trattata come una tecnologia normale”⁵². E, forse, il lato veramente pericoloso delle nuove tecnologie, che allo stesso momento ci dimostra che non possiamo non affrontarle che in modo „antropocentrico” o – se si vuole – „personalistico”, è che „faremmo meglio ad aver paura di noi stessi”⁵³. Se ci vuole pertanto l'esercizio della „responsabilità per il potere”, e quello si esercita tramite la „memoria” come capacità di „accogliere il dato della nostra vita per proiettarlo in avanti attraverso la cura e dunque

⁴⁸ Y.N. Harari, *Homo Deus*, cit., p. 73.

⁴⁹ M. Bellini – M. T. Della Mura, *Umano digitale. Verso un'etica dell'innovazione, in dialogo*, Milano 2023, p. 113.

⁵⁰ Cfr. L. Floridi – J. Cowls, *A Unified Framework of Five Principles for AI in Society*, in: L. Floridi (ed.), *Ethics, Governance, and Policies in Artificial Intelligence*, Springer 2021, pp. 5-17, qui, p. 9.

⁵¹ *Ibid.* p. 13.

⁵² L. Floridi, *Introduction – The Importance of an Ethic-First Approach to the Development of AI*, in: Id. (ed.), *Ethics*, cit., pp. 1-4, qui p. 3.

⁵³ Y.N. Harari, *Homo Deus*, cit., p. 95.

attraverso la relazione⁵⁴, ciò avrà quindi delle conseguenze per la legislazione e la governance delle nuove tecnologie. Ciò che abbiamo bisogno, infatti, è „niente di meno che un’antropologia filosofica in linea coi nostri tempi, ovvero, una comprensione filosofica della natura umana che sia adeguata all’era digitale e alle nostre società dell’informazione⁵⁵. Sarà il compito del gruppo 6 di Persona al Centro delinearne i principi e le linee di sviluppo.

⁵⁴ F. Miano, *Sfide e possibilità della condizione tecno-umana*, in: C. Caltagirone – L. Cucurachi (edd.), *La condizione tecno-umana tra eccesso ed eccedenza*, Morcelliana, Brescia 2021, pp. 35-48, qui p. 48.

⁵⁵ L. Floridi, *In poche battute*, cit., p. 80.