

Informazione Filosofica

**Rivista quadrimestrale a cura dell'Istituto
Lombardo per gli Studi Filosofici e Giuridici**

N. 13/14 – giugno/dicembre 2024

Reale e Virtuale



Centro Ricerche e
Formazione
UNITRE



Direttore Scientifico: Silvio Bolognini (Università eCampus)

Comitato Scientifico: Paolo Becchi (Università di Genova), Rolando Bellini (Accademia di Brera – Milano), Enrico Bocciolesi (Università degli Studi di Urbino Carlo Bo), Alessandro Bolognini (Università eCampus), Mario Ciampi (Università Guglielmo Marconi), Massimo De Leonardis (Università Cattolica del Sacro Cuore – Milano), Franco Giuseppe Ferrari (Università Bocconi), Stefano Bruno Galli (Università degli Studi di Milano), Adria Velia González Beltrones (Universidad de Sonora, México), Pier Francesco Lotito (Università degli Studi di Firenze), Sergio Guido Luppi (Università Cattolica del Sacro Cuore – Milano), Marco Marinacci (Università eCampus), Narciso Martínez Morán (UNED – España), Eloy Martos Nuñez (Universidad de Extremadura, RIUL – España), Roberto Montanari (Università Suor Orsola Benincasa – Napoli), Aldo Ocampo González (CELEI, Chile), Mario Pesce (Università di Roma Tor Vergata), Marina Simeone (Università eCampus)

Comitato Tecnico-editoriale: Attilio Cristiano Vaccaro Belluscio, Francesco Carlesi, Alberto Pesce, Roberta Simeone, Luca Siniscalco

Segreteria di redazione: Anna Cattaneo

Per l'invio dei contributi originali, non inviati ad altre riviste (il testo, corredato di un *abstract* in inglese e uno nella lingua di stesura del saggio, ciascuno non più di 500 caratteri, spazi inclusi, e di 5 *key words*, deve essere privo di indicazioni relative all'autore; in un *file* a parte va spedita un documento con nome/cognome, titolo, istituzione di appartenenza, email; vanno seguite le norme redazionali disponibili online sul sito www.informazionefilosofica.it nella sezione dedicata): info.informazionefilosofica@gmail.com

I contributi destinati alla pubblicazione nelle sezioni “Autori e Idee” e “Tendenze e dibattiti” vengono preventivamente sottoposti a procedura di *double-blinded peer review* (revisione a “doppio cieco”). Il Comitato Scientifico si avvale di esperti esterni nel processo di revisione tra pari a doppio cieco. La Direzione editoriale può in ogni caso decidere di non sottoporre ad alcun *referee* l'articolo, perché giudicato non pertinente o non rigoroso né rispondente a standard scientifici adeguati. I contributi non pubblicati non saranno restituiti

Il Codice Etico di *Informazione Filosofica* è consultabile all'indirizzo web www.informazionefilosofica.it

Proprietà: CRIFU – Centro Ricerche e Formazione UNITRE
Via Ariberto 11 – 20123 Milano (MI)
www.unitremilano.education.it

Editore: Armando Editore
Via Leon Pancaldo 26 – 00147 Roma
www.armandoeditore.it

La denominazione ed il marchio dell'Istituto Lombardo per gli Studi Filosofici e Giuridici sono di proprietà del CRIFU, cui appartiene anche la testata «Informazione Filosofica». Fondatori dall'Istituto sono l'avv. Mario Giacomini e famiglia e l'avv. Gerardo Marotta, già presidente dell'Istituto Filosofico italiano di Napoli.
Sito della rivista: www.informazionefilosofica.it
Direttore responsabile: Silvio Bolognini

ISSN: 2724-1637
ISBN: 979-12-985245-3-8

SOMMARIO

PRESENTAZIONE	3
INTRODUZIONE: REALE E VIRTUALE	5
AUTORI E IDEE	8
Il paradigma tecnomediato nelle giuspolitiche per la condizione anziana e i suoi limiti di <i>Silvio Bolognini</i>	10
Intelligenza artificiale e coscienza in una prospettiva buddhista e spiritualista di <i>Roberto Siconolfi</i>	43
Oltre il confine. Reale e virtuale nella trasformazione della soggettività contemporanea di <i>Ignazio Iacone</i>	68
Specula. L'intelligenza artificiale tra guerra, cura e sorveglianza. Etica e potere nella società del controllo digitale di <i>Alberto Pesce</i>	84
TENDENZE E DIBATTITI	102
Le Nuove Frontiere dell'Innovazione Tecnologica in Sanità: l'I.A. in aiuto delle persone anziane con patologie cronico-degenerative, ad esempio i malati di Parkinson, tra opportunità e rischi di <i>Barbara Baccarini</i>	104
Algorithmic Opacity di <i>Giuseppe Gimigliano, Simona Ibba</i>	113
La trasformazione digitale della società: dalla mediamorfosi al transumanesimo di <i>Sandra Ciarcianelli</i>	126

Presentazione

Con il fascicolo nr. 1 (2020), dedicato alla riflessione filosofica sull'Ermeneutica del "Ponte" e del "Muro" e sulla metodologia che vi si accompagna, è rinata *Informazione Filosofica*. Torna così presente nel dibattito accademico nazionale – con una contestuale apertura internazionale – e nella riflessione filosofica in genere una pietra miliare della tradizione editoriale italiana, espressione di ricerca teoretica, storico-filosofica e comparatistica.

Riteniamo dunque doveroso tratteggiare la storia della rivista, segnalandone le peculiarità essenziali. Nata nel 1990 come rivista bimestrale, *Informazione Filosofica* esce a cura dell'Istituto Italiano per gli Studi Filosofici (con sede a Napoli) in collaborazione con l'Istituto Lombardo per gli Studi Filosofici e Giuridici (con sede a Milano), del quale anche la nuova serie è emanazione diretta. L'ampia Redazione Scientifica, supportata dall'eccellente Comitato Scientifico (che vide la presenza, fra gli altri, dei filosofi Remo Bodei, Domenico Losurdo, Paul Ricoeur, Paolo Rossi e Franco Volpi) e dai numerosi collaboratori, realizzò ben 35 pubblicazioni nell'arco dei sette anni di vita del progetto culturale, conclusosi nel 1997.

Scopo primario della pubblicazione era offrire ai lettori un panorama ampio e variegato della tradizione filosofica occidentale (ma non solo), considerandone i fondamenti storici e teoretici, ma impegnandosi anche, al contempo, in un serrato approfondimento del dibattito attuale e in un accurato vaglio delle nuove prospettive di ricerca, nonché delle attività culturali – pubblicazioni, convegni e seminari – tramite cui i contenuti e il metodo proprio della filosofia vengono trasmessi e veicolati al pubblico contemporaneo.

Per sette anni, insomma, *Informazione Filosofica* ha tradotto in una sistematica rassegna l'evoluzione della filosofia italiana e internazionale, mettendo in evidenza autori e correnti, tendenze e dibattiti, libri, riviste, attività didattiche.

Oggi, a distanza di oltre due decenni, il progetto riprende vita, con rinnovato spirito e sotto l'egida del CRIFU – Centro Ricerche e Formazione UNITRE (Ente di Alta Formazione e Ricerca appartenente al RIUL (Red Internacional Universidades Lectoras) e del da tempo recente connesso Istituto Lombardo di Studi Filosofici e Giuridici. La presente pubblicazione si inserisce, con un nuovo Comitato scientifico e uno tecnico-editoriale, nel solco della tradizione della prestigiosa rivista. Una continuità ideale, questa, che si declina in una metamorfosi – un rinnovamento di forme e finalità. *Informazione Filosofica* diventa infatti un quadrimestrale e acquisisce un taglio pienamente accademico, incentrando il proprio focus sulla valorizzazione della storia delle idee e della filosofia alla luce di tematiche cogenti per la contemporaneità del nuovo millennio. I numeri, integralmente monografici, si presenteranno caratterizzati da una

struttura bipartita: la sezione “Autori e Idee” raccoglie i saggi più corposi, contraddistinti da un impianto critico rigorosamente scientifico e dedicati all’approfondimento di autori e/o percorsi tematici che permettano di gettare nuova luce sulla macrotematica di volta in volta affrontata; la sezione “Tendenze e dibattiti” include contributi che, mantenendo la medesima configurazione accademica, affrontano questioni più specifiche, spesso tramite “sconfinamenti” interdisciplinari, sempre toccando questioni di estrema rilevanza e centralità nel dibattito culturale contemporaneo.

A guidarci, nella serrata trattazione di questioni impellenti non soltanto sul piano della ricerca ma dell’esistenza umana concreta, intesa e vissuta nella sua fatticità, è lo splendido aforisma di Nicolás Gómez Dávila, il quale ci ricorda che, rifiutando ogni giustificazionismo acritico e ogni subordinazione alla vulgata corrente e al *sensus communis*, “il filosofo non è portavoce della sua epoca, ma angelo prigioniero nel tempo”.

Introduzione

Reale e Virtuale

Il dibattito sul rapporto fra reale e virtuale, che attraversa oggi ogni ambito dell'esperienza umana, si è imposto come uno dei nuclei più fertili della riflessione filosofica contemporanea. Il digitale non è più un semplice strumento, ma una dimensione costitutiva del mondo della vita: un piano in cui la realtà fisica e quella mediata da dispositivi, reti e algoritmi si intrecciano fino a confondersi. In questo nuovo scenario, le categorie classiche della metafisica – essere, possibilità, apparenza, relazione – devono essere ripensate alla luce di un'ibridazione crescente fra naturale e artificiale, presenza e rappresentazione, corporeità e informazione.

Questo fascicolo di *Informazione Filosofica* affronta tale problematica in una prospettiva tanto speculativa quanto prassistica, muovendosi tra la riflessione ontologica e l'analisi delle implicazioni etiche, politiche e sociali della trasformazione digitale. Emergono, così, diversi piani di interrogazione: quello dell'identità e della soggettività, profondamente mutate nel passaggio a un ambiente tecnomediato; quello della coscienza e della mente artificiale, che solleva questioni di natura epistemologica e spirituale; e quello, cruciale, del potere, dove il virtuale si manifesta come spazio di controllo, di sorveglianza, ma anche di possibilità emancipative.

La rivoluzione tecnologica contemporanea, lungi dal potersi ridurre a questione di efficienza o innovazione, investe la struttura stessa del vivere comune. Essa ridisegna le politiche della cura, l'organizzazione del lavoro, le forme di comunicazione e persino le esperienze fondamentali della vita umana – come nascere, invecchiare, morire. L'interrogativo filosofico che s'impone alla comunità scientifica diventa allora: quale idea di umanità sopravvive – o si reinventa – nell'epoca della connessione pervasiva? La mediazione tecnologica può essere al contempo strumento di assistenza e di alienazione, di libertà e di controllo, di prossimità e di solitudine. La risposta non può dunque limitarsi ad una presa d'atto delle trasformazioni, ma richiede una meta-riflessione in cui griglie ermeneutiche di tipo filosofico permettano una comprensione più articolata – e critica – dei fenomeni in atto e, idealmente, forniscano mezzi di indirizzamento e orientamento.

In questo contesto, la riflessione proposta da *Informazione Filosofica* si articola precisamente come un esercizio di discernimento: comprendere in quale modalità specifica il virtuale non sia un semplice duplicato del reale, ma suo prolungamento e, talvolta, un suo correttivo. Le analisi raccolte mostrano che la tecnologia non costituisce un destino, bensì un linguaggio – un nuovo modo di articolare le relazioni fra individuo, comunità e mondo. La sfida che ne deriva è insieme teoretica e politica: saper abitare criticamente il virtuale, mantenendo aperto lo spazio dell'esperienza, della responsabilità e del senso.

Ripensare il reale alla luce del virtuale significa dunque interrogare le condizioni della nostra presenza al mondo, e le nuove forme di trascendenza che la tecnica dischiude. La filosofia, ancora una volta, è chiamata a esercitare la propria funzione originaria: orientare la ragione nel tempo, offrendo parole adeguate all'inedito e restituendo pensiero all'esperienza.

Autori e Idee

Il paradigma tecnomediato nelle giuspolitiche per la condizione anziana e i suoi limiti

di *Silvio Bolognini**

ABSTRACT (ITA)

Il saggio analizza il ruolo delle tecnologie digitali nelle politiche di welfare rivolte alla popolazione anziana, con particolare riferimento ai modelli di cura integrata e domiciliare abilitati dalle ICT. Attraverso l'esame di studi europei e delle strategie del PNRR, Bolognini mette in luce benefici e criticità del paradigma "Care 3.0", evidenziando il rischio di una tecnologicizzazione coercitiva dell'assistenza e la necessità di preservare la dimensione relazionale e comunitaria del prendersi cura.

Parole chiave: welfare digitale, anziani, innovazione sociale, teleassistenza, etica della cura

The Technomediated Paradigm in Policies for the Elderly and Its Limits

by *Silvio Bolognini*

ABSTRACT (ENG)

The essay explores the role of digital technologies in welfare policies addressing the elderly, focusing on ICT-enabled models of integrated and home care. Through the analysis of European research and Italy's National Recovery and Resilience Plan (PNRR), Bolognini highlights both the benefits and the critical limits of the "Care 3.0" paradigm, warning against the coercive potential of technologized care and advocating for the preservation of the relational and community dimensions of caregiving.

Keywords: digital welfare, elderly, social innovation, telecare, ethics of care

* Università eCampus

1. Dalla “Care 3.0” alla digitalizzazione del welfare

Nel 2016 nell’ambito dell’intervento “Shaping the Welfare Society: Unleashing transformation through ICT-Enabled Social Innovation” Gianluca Misuraca, responsabile del progetto europeo di ricerca IESI (*ICT-Enabled Social Innovation*)¹ che esplorava le iniziative di innovazione sociale abilitate dalle ICT volte al conseguimento degli obiettivi posti dal SIP,² sintetizzava come segue le evidenze della mappatura effettuata in questa area specifica di erogazione dei servizi: “In the area of active and healthy ageing and long-term care, ICTs show a transformational power to improve the traditional way of providing care, or even addressing yet unmet or emerging needs” (Misuraca, 2016a, p. 30).

L’impiego delle ICT dimostrava dunque di avere il potere di migliorare, trasformandoli, i modelli tradizionali di assistenza, rispondendo a fabbisogni emergenti che il sistema vigente non era in grado di approcciare.

Il *Welfare 2.0* agganciava ulteriormente il concetto di “Care 3.0: Robotics for personalised integrated care solutions” (Misuraca, 2016b, p. 13), così come tratteggiato dallo stesso Misuraca nella prefigurazione di un *welfare system* proiettato al 2036:

- Ampia diffusione di “Social Robots” come compagni delle persone anziane bisognose di assistenza.
- Possibilità di autonomia e sicurezza personale, assicurando le famiglie e rendendo possibile uno “stare insieme” mutuato dal *social networking* e dall’impiego di strumenti portatili.
- Riduzione dei costi di ospedalizzazione e del peso gravante sui “caregivers” che supportano l’anziano.
- Miglioramento dei tempi di vita per famiglie e “informal carers”, con parallelo incremento dell’autostima e delle capacità cognitive dell’anziano, in forza dell’interazione con l’Intelligenza Artificiale.

Se quelli sopra indicati rappresentano i presunti punti di forza del nuovo modello avanzato di “cura”, l’intervento richiamato ne adombra in effetti anche le possibili minacce, quali il rischio di

¹ Progetto lanciato dall’*Institute for Prospective Technological Studies* (IPTS) di Siviglia, uno dei sette istituti del JRC, Direttorato Generale della Commissione Europea, avente quale obiettivo la comprensione delle connessioni fra tecnologia, economia e società.

² *Social Investment Package*: una serie di documenti programmatici adottati dalla Commissione Europea nel 2013 in risposta alla necessità di una urgente ridefinizione delle politiche per il *Welfare* in uno scenario europeo caratterizzato da fattori strutturali di forte e crescente pressione sul sistema dello “Stato Sociale” – dall’invecchiamento della società alla gestione dei flussi migratori, dal calo dell’occupazione al rallentamento del PIL con effetti amplificati, in termini di riduzione delle risorse disponibili, dal *Fiscal Compact*.

una insufficiente comprensione del rapporto uomo-macchina, con possibili effetti negativi sull'isolamento sociale, e del furto e dell'uso improprio di dati.

Per quanto, come specificato, lo scenario prefigurato sia proiettato su un relativamente ancora lontano 2036, il *Care 3.0* era un modello già operativo nel 2015, per lo meno all'interno di progetti pilota, e in qualche caso già inserito nel "portafoglio" dei servizi offerti dalla PA.

È significativa in questo senso proprio la mappatura realizzata nel primo anno del progetto IESI menzionato sopra, che copriva l'area dell'autonomia (*independent living*)³ e delle cure integrate (*integrated care*)⁴ nella duplice prospettiva dei beneficiari (intermedi e finali) e degli erogatori dei servizi (Carretero & Kucsera, 2015, p. 22).

Coerentemente con quanto testimoniato dalla letteratura sul tema, le iniziative documentate dimostravano il contributo positivo del fattore tecnologico all'efficientamento del sistema: dalla riduzione del numero e della durata degli accessi e dei ricoveri ospedalieri all'incremento di produttività e qualità dei servizi (Carretero & Kucsera, 2015; Hermans & Vranken, 2010; Lewin et al., 2010)

Le tecnologie messe in campo per elevare il grado di autonomia dell'anziano erano ricondotte, nel Rapporto, a sei categorie:

1. *Generic information and communication technology (ICT) products, services and applications*: si tratta di prodotti e servizi ICT generici, presenti sul mercato, accessibili mediante telefonia mobile o applicazioni *Internet* per PC – dai servizi di *teleworking* alle piattaforme formative e di informazione, dal social *net-working* ai servizi di *e-commerce* e *e-banking*.
2. *Assistive technologies*: strumenti ed equipaggiamenti atti a compensare *deficit* cognitivi, sensoriali o fisici, che vanno dai *software* di riconoscimento vocale alle protesi robotiche fino ai *robot* utilizzati per la riabilitazione, o per l'adempimento di funzioni specifiche (ad esempio l'igiene personale).

³ L'ambito identificato come "independent living" include, nel contesto della Ricerca: "measures to compensate for older people's physical and mental restrictions, empowering and enabling them to deal with functional limitations; to achieve a greater degree of self-sufficiency and maximise their autonomy; and to reduce the need for care. These measures cover:

- Self-care, self- monitoring and self-management of functional limitations
- Age-friendly environments".

⁴ L'ambito identificato come "Integrated care" fa qui riferimento a "measures to increase the accessibility, take-up, productivity, quality and effectiveness of care (more and better care with fewer resources). It focuses on better organisation of care among care providers, and improving the supply and retention of carers, shifting care to the formal sector and making care more attractive in order to increase employment rates. It includes the following topics:

- Coordination/integration of health care and social elements (services, professionals, and relatives)
- Support to (formal and informal) carers".

3. *Smart homes*: sistemi ICT integrati che rendono la casa “intelligente”, basati prioritariamente sull’*Internet of Things*, con estensione a servizi *internet* quali *teleshopping* e *telebanking*.
4. *Technology-based healthcare*: tecnologie impiegate per prevenire, individuare, curare e gestire condizioni di cronicità (telemedicina, e più in generale *telehealth*, o *telemonitoring*, per il monitoraggio integrato da remoto di parametri vitali e dati clinici dell’assistito).
5. *Technology-based home care*, in cui le ICT sono impiegate per monitorare lo stato quotidiano di benessere dell’assistito mediante sistemi più o meno integrati di supporto nell’assunzione di farmaci e nel monitoraggio delle terapie e, soprattutto, mediante “*telecare*”, termine con il quale si fa riferimento a dispositivi di controllo e allarme tipo salva-vita, di livello più o meno avanzato e integrato rispetto al sistema di erogazione dei servizi, ma anche al cosiddetto *telecare* di terza generazione, che prevede l’installazione capillare di sensori ambientali atti a monitorare le azioni quotidiane del soggetto preso in carico, con controllo demandato ai familiari o a personale professionale.
6. *Technology-based wellness services deliver*: servizi finalizzati a promuovere uno stile di vita “sano”, fornendo all’assistito la possibilità, ad esempio, di fare ginnastica fisica e mentale mediante, prioritariamente, giochi per *computer* o *internet-based*, per arrivare all’impiego di veri e propri *social robots* quali quelli rappresentati nella visione del *Welfare* al 2036 su cui ci siamo soffermati sopra.

Per quanto il *trend* monitorato indicasse una progressiva diffusione dell’utilizzo dell’ICT in questi ambiti, il Rapporto ne rilevava anche la lentezza entro i confini dell’Unione Europea, riconducendone le ragioni a fattori molteplici, fra i quali la scarsa informazione in merito all’impatto positivo atteso sull’efficientamento del sistema e alle modalità possibili di impiego della tecnologia nella catena di erogazione dei servizi. Ulteriormente si sottolineava la mancanza di *business models* in grado di incentivare i Governi ad investire nell’introduzione/nel potenziamento del fattore tecnologico.

2. Il PNRR e la sanità digitale

Quasi 10 anni dopo possiamo verosimilmente sostenere che almeno due su tre degli elementi di criticità individuati siano stati “gestiti”: non solo il valore aggiunto apportato ma anche la necessità di una sanità sempre più digitalizzata sono assurti a consapevolezza comune e a *policy*, nonché a programmazione strategica su scala europea e nazionale, con il Piano/Missione 6 del PNRR (Piano Nazionale di Ripresa e Resilienza) approvato nel 2021, che investe come noto 4,750 miliardi di euro in reti di prossimità, strutture, telemedicina e nella “Casa come primo luogo

di cura”, mediante l’assistenza domiciliare, le Centrali Operative Territoriali (COT) e la telemedicina dedicata ai pazienti cronici.

A riscontro dell’acquisita consapevolezza in merito agli obiettivi già delineati nel 2015 il PNRR si propone di:

- garantire a tutta la popolazione lo stesso livello assistenziale e qualità delle cure erogate migliorando l’efficienza organizzativa anche in contesti geograficamente svantaggiati;
- fornire al paziente un’assistenza che comprenda risposte a bisogni clinico-assistenziali, e psicologici anche per ridurre l’impatto negativo del vissuto di malattia del paziente
- ridurre gli accessi al pronto soccorso, l’ospedalizzazione e l’esposizione ai rischi legati ad essa, nonché il ricorso all’istituzionalizzazione;
- permettere una dimissione protetta dalle strutture di ricovero potendo assicurare a domicilio la continuità di assistenza e cure di pari efficacia;
- ridurre gli accessi in day hospital sviluppando risposte alternative a domicilio (PNRR, Missione 6, pp. 4-5).

Il terzo fattore critico indicato nel Rapporto IESI, la mancanza di un modello applicativo condiviso, è il fattore in gestione, possiamo verosimilmente affermare, proprio nell’ambito dell’implementazione del PNRR i cui obiettivi, con riferimento al progetto di presa in carico dell’assistito a casa, sono stati puntualmente sintetizzati dal Ministero della Salute:

Identificare un modello condiviso per l’erogazione delle cure domiciliari che sfrutti al meglio le possibilità offerte dalle nuove tecnologie, come la telemedicina, la domotica, la digitalizzazione.

Realizzare presso ogni Azienda Sanitaria Locale (ASL) un sistema informativo in grado di rilevare dati clinici in tempo reale.

Attivare Centrali Operative Territoriali (COT), una in ogni distretto, con la funzione di coordinamento della presa in carico della persona e raccordo tra servizi e professionisti coinvolti nei diversi setting assistenziali: attività territoriali, sanitarie e sociosanitarie, ospedaliere e dialogo con la rete dell’emergenza-urgenza.

Utilizzare la telemedicina per supportare al meglio i pazienti con patologie croniche [gli interventi di telemedicina realizzabili nell’ambito dell’assistenza domiciliare sono specificati come televisita, teleconsulto medico, teleconsulenza medico-sanitaria, teleassistenza, telemonitoraggio, telecontrollo, teleriabilitazione ndr].

Ulteriormente fra le modalità di applicazione del modello digitale si individua quale elemento qualificante la domotica assistiva per migliorare l’accessibilità e la fruibilità dell’ambiente

domestico, nonché a supporto delle attività di telemonitoraggio e telecontrollo della telemedicina, al fine di ridurre incidenti domestici e preservare lo stato di salute dell'assistito (PNRR, Missione 6, p. 12).

3. Tra efficienza e umanità

Sempre nel 2021, anno di approvazione come abbiamo ricordato del PNRR, un articolo pubblicato sulle pagine dell'*Aging Project* del DIMET (Dipartimento di Medicina Traslazionale dell'Università del Piemonte Orientale) in tema di domotica assistiva sottolineava che:

Le opportunità più interessanti derivano dal cosiddetto *Internet of Things*, cioè dalla possibilità di connettere e far comunicare tra di loro dispositivi e oggetti dell'ambiente domestico, per migliorare la salute, l'autonomia e la qualità di vita delle persone anziane o diversamente abili. Soluzioni tecnologiche studiate ad hoc per gli anziani possono aumentare l'aderenza terapeutica, la percezione di sicurezza e l'autocontrollo (Failla, 2021).

A titolo esemplificativo dell'efficacia dell'approccio, l'articolo citava anche uno studio condotto dall'Eurac Research di Bolzano su 36 individui di età compresa fra i 65 e i 94 anni che avevano testato per un anno il funzionamento di un *kit* domotico composto da vari oggetti connessi in rete: un tablet, un orologio di emergenza e sensori posizionati in punti strategici dell'abitazione.

Questo sistema integrato – conclude l'articolo – si è rivelato utile per monitorare le abitudini dell'anziano, controllare i suoi parametri vitali, agevolarlo nelle attività quotidiane e allertare i soccorsi in caso di emergenza. Se, ad esempio, una persona fa abitualmente colazione tra le sette e le otto di mattina e alle nove non ha ancora aperto il frigorifero, il sensore attaccato al frigo invia un messaggio al suo tablet per accertarsi che non sia successo nulla. Se la persona non risponde entro 20 minuti il sistema invia un alert ad un parente o a un caregiver, in modo che qualcuno possa sincerarsi delle sue condizioni. Altri sensori comandano l'accensione automatica della luce quando l'anziano si alza di notte o l'emissione di un segnale acustico quando una pentola viene dimenticata sul fornello acceso (Failla, 2021).

Il *Care 3.0* e i *social robots* chiamati in causa da Misuraca nello scenario della Sanità del territorio proiettato al 2036 sono menzionati nello stesso articolo in riferimento ad un esperimento condotto dall'Università tecnologica Nanyang di Singapore, che testava l'impatto di *robot* umanoidi in risposta ai bisogni emotivi e sociali delle persone anziane.

Per qualche giorno il robot di compagnia Nadine – uno dei più “umani” al mondo per fattezze, capacità interattive e comportamento – ha intrattenuto e divertito 29 ospiti over 60 della casa di cura Bright Hill Evergreen di Singapore. In particolare, a Nadine è stato affidato il compito di gestire le partite di bingo. Tutte le sessioni di gioco condotte dal robot sono state filmate, in modo da confrontare le reazioni, le espressioni e i comportamenti dei partecipanti con quelli registrati mentre la stessa attività veniva condotta da esseri umani. Ne è emerso – riferisce l'autrice della pubblicazione – che gli anziani si divertivano di più in compagnia del robot: il loro volto appariva più felice e sorridente, la loro attenzione era più alta e si distraevano meno, evitando di rivolgersi al personale per qualsiasi necessità. Il successo di Nadine dipende principalmente dal suo aspetto umano e dalla capacità di leggere i gesti e la mimica facciale, caratteristiche fondamentali per agevolare l'interazione con gli anziani, in genere poco avvezzi alla tecnologia (Failla, 2021).

Pensando al contesto italiano dell'assistenza e alla sua tradizione, ma anche ad un più esteso contesto europeo, le conclusioni dello studio di cui sopra inducono verosimilmente a una riflessione, già sollecitata in effetti dalle autrici di una ricerca IRS pubblicata quasi dieci anni prima, nel 2012, che mettono in luce le difficoltà di adattamento del settore delle ICT alle caratteristiche del *welfare*, centrato sull'idea del “prendersi cura”:

The social services sector is centered around people and service delivery, not technological products. According to some literature the positive contribution of ICT to the quality and the productivity of care has not yet been proven. To do so technology should be embedded in the service delivery model. Too often it is now seen as a substitution of care. On the other hand, technology players do not understand the market of welfare and care (Crepaldi, De Rosa & Pesce, 2012, p. 44).

Analogamente, il Progetto IESI rilevava la plausibilità di una diffusa resistenza al cambiamento da parte degli operatori: “Health and social care professionals may show reluctance to accept new technologies, mainly because the technology implies, they have to learn new competences and change their traditional ways of working. They are also afraid of being substituted by the technology” (Carretero et al., 2015, p. 35).

Poco oltre, il tema viene ripreso sottolineando espressamente la centralità del fattore umano cui il “prendersi cura di” rimanda nell'immaginario comune e in quello degli operatori del terzo settore: “Finally, sometimes there is reluctance among the care professionals to accept and adapt ICTs in their work; for instance because they perceive it as something that deprives the care of its essential human/emotional element or because introducing technology might require extensive training from their part” (Carretero et al., 2015, p. 47).

Gli operatori potrebbero quindi essere poco propensi all'adozione di sistemi ad alto contenuto tecnologico sia per la necessità che si prefigurerebbe loro di acquisire nuove conoscenze e abilità, sia perché effettivamente convinti che ciò possa svuotare il servizio di contenuti fondamentali, quali quelli emozionali che caratterizzano la relazione di aiuto.

Le stesse criticità, identificate come sfide da affrontare per poter inserire il nuovo *welfare* nel *mainstream* della sanità territoriale, sono richiamate in un altro studio del JRC del 2015, che analizza singoli casi applicativi dell'ICT ai servizi di supporto all'autonomia delle persone anziane e che rileva, a nostro avviso, per l'attenzione ad un fattore che sembra perdere interesse in termini di obiettivo e di *focus* nella declinazione delle strategie nazionali di digitalizzazione della sanità, così come nei più recenti progetti di domotica assistiva e *care 3.0* richiamati sopra: la dimensione comunitaria (Carretero & Kucsera, 2015).

Fra i progetti presi in considerazione gli autori approfondiscono, in particolare, l'iniziativa denominata *assisting carers using telematics interventions to meet older people's needs (action)*, rivolta ad anziani con patologie croniche e "badanti" informali. Si tratta di un progetto svedese lanciato nell'ambito del Quarto Programma Quadro di Ricerca dell'Unione Europea: alle famiglie *target* del servizio furono messi a disposizione programmi integrati e multimediali di assistenza accessibili mediante *TVsets* e PC, inclusivi di giochi *on line* e percorsi di *training* fisico e cognitivo per gli *end-users*, un *call center* dedicato, collegato mediante videotelefono e videocamera, e programmi di formazione e monitoraggio, sia per gli utilizzatori che per gli operatori istituzionali delle municipalità coinvolte. La motivazione sottesa al lancio del progetto faceva esplicitamente riferimento alla volontà di promuovere l'inclusione sociale attiva delle persone anziane e consolidare un *welfare* di comunità valorizzando, da una parte, il ruolo del primo microcosmo sociale di riferimento, la famiglia, e, dall'altra, il potenziale delle nuove tecnologie per l'informazione e la comunicazione nel rafforzamento della rete relazionale al cui centro si pone l'assistito:

The rationale – leggiamo nello Studio – that motivated the development of the project in 1997 was based on the following policy and research drivers of that time. At policy level, we can highlight:

- The need to develop community care policies and create resources for “ageing in place”, due to the ageing demographic trends and finite economic resources.
- The recognition of role played by the family in supporting frail older relatives and the room for improvement, mostly in Swedish policy (Board of Health and Welfare Sweden, 2002).
- The onset of EU policies in the mid-nineties focused on the active inclusion of older and disable people within the new information society. The fourth framework programme (1994-1998) developed a specific call for Information Society Technologies. The TIDE initiative (Telematics Integration for the Disabled and Elderly) acted as a ‘flagship’ for these key policies.

- The potential role of technology in enhancing the independence of older people living at home was recognized in UK (Audit Commission, 2004) and Swedish (Board of Health and Welfare Sweden, 2004) policies.

At research level – specificano ulteriormente gli autori – there was also a need to improve the development of telecare and telehealth services with feedback from the target end-users: older people and their carers. This would help to increase the use and acceptance of these services by the users (Carretero, & Kucsera, 2015, p. 66).

Il progetto intendeva, sul fronte degli *users*, trasferire consapevolezza e sicurezza ai beneficiari finali del servizio, migliorare le capacità e le possibilità di assistenza espresse dai familiari e ridurre l'isolamento e il senso di abbandono frequentemente percepiti da parte di entrambi i *target*. Contestualmente, sul versante degli erogatori, obiettivo prioritario era ridurre i costi dell'assistenza ed incrementare il livello di qualità dei servizi, offrendo altresì agli operatori nuove opportunità di crescita professionale.

L'iniziativa è stata portata avanti dal Governo svedese per alcuni anni e costantemente implementata; la contrazione del bacino di utenza tuttavia, dai 350 assistiti nel 2011 ai 100 rilevati nel 2015, ha indotto gli esperti incaricati di monitorare la sperimentazione ad analizzare le criticità ancora, evidentemente, altamente sfidanti ai fini di un definitivo decollo e di una larga diffusione del sistema *ICT based*, nonostante la valutazione di impatto effettuata sul servizio dimostrasse l'effettiva contrazione della spesa pubblica per unità familiare e la riduzione dei costi imputabile al ritardo nell'accesso a strutture protette.

Se, osservavano i valutatori, il fattore finanziario e politico ha giocato un ruolo chiave nel crollo del numero di assistiti, stante la difficoltà di definire un adeguato modello di *business* e di raggiungere la sostenibilità economica in mancanza di una massa critica sufficiente entro i limiti di progetti prevalentemente configurati come *pilots*, altri fattori critici hanno inficiato il successo dell'Iniziativa, in particolare la necessità di una formazione estesa ai diversi *target* di riferimento – dagli utilizzatori finali alle famiglie, dal personale istituzionale agli operatori professionali – e la diffidenza dei prestatori di assistenza. Numerose si sono rivelate le criticità afferenti al *local caring staff* (Carretero & Kucsera, 2015, pp. 75-76).

Diverso si è dimostrato tuttavia l'impatto registrato sul fronte dei beneficiari finali e intermedi della sperimentazione: dalla rilevazione effettuata su circa 400 utilizzatori, mediante somministrazione di questionari, interviste e *focus groups*, emergevano l'incremento del grado di inclusione sociale di persone, anziani e familiari, tradizionalmente esclusi dai benefici delle ICT, e la riduzione dell'isolamento, anche in virtù della costruzione di reti informali con altri partecipanti in condizioni affini. Diffusa era, parimenti, la percezione di un miglioramento della qualità della vita quotidiana, espressa dalla maggioranza delle persone coinvolte nell'indagine. L'utilizzo degli

strumenti messi a disposizione sembrava avere, altresì, influito positivamente sul senso di sicurezza e soddisfazione dei badanti informali, a fronte di un incremento effettivo delle competenze di ruolo. Ne sarebbe risultato, infine, un arricchimento della relazione di aiuto, fondamentalmente basato sull'acquisizione di maggiore consapevolezza e capacità di controllo da entrambe le parti (assistente e assistito).

Valutazioni analogamente positive da parte dei beneficiari dei servizi emergevano dalla valutazione delle iniziative censite dal Rapporto di Ricerca IESI, mappate in base alla riconducibilità a quattro categorie:

1. il potenziale dell'innovazione abilitata dalle ICT, indicizzato ricordiamo come “technical/incremental, sustained/organisational, disruptive e radical”;
2. il coinvolgimento dei livelli di *governance* nell'integrazione dei servizi, in senso verticale e orizzontale (coinvolgimento del privato/terzo settore);
3. la tipologia specifica di integrazione sperimentata (finanziaria, amministrativa, organizzativa, operativa);
4. gli elementi di innovazione sociale espressi, fondamentalmente ricondotti all'orientamento al fabbisogno, all'attivazione di processi collaborativi, alla trasformazione della relazione fra *stakeholders* e alla redistribuzione di valore pubblico.

Nell'area dell'*independent living* la Ricerca mappava nella fattispecie 12 iniziative, fra cui lo stesso progetto ACTION sopra menzionato, inclusive delle diverse formule applicative della tecnologia ai servizi, in particolare *health care* e *home care services*, quasi tutte caratterizzate da potenziale innovativo di livello “radical” o “disruptive”.

Alcuni progetti privilegiavano un approccio mirato, con l'obiettivo ad esempio di mettere a disposizione degli utenti ausili per il recupero o il rafforzamento di capacità cognitive specifiche; altri sperimentavano sistemi più complessi di monitoraggio dei parametri fisici e ambientali per migliorare la capacità di risposta e le caratteristiche di trattamenti e cure mediche, nonché per consentire un intervento immediato in caso di emergenza; talune iniziative focalizzavano l'attenzione sui beneficiari finali dei servizi, altre intendevano agire anche, rafforzandolo, sul sistema informale dell'assistenza, come nel caso del sopra menzionato progetto ACTION.

Nella valutazione dell'impatto dei singoli e diversi progetti ricorreva, sul piano quantitativo, la registrazione di una contrazione della spesa pubblica, imputabile tendenzialmente al ridotto numero di ricoveri ospedalieri e degli accessi al pronto soccorso e al risparmio generato dalla posticipazione della collocazione dell'anziano presso strutture protette; parallelamente, sul piano qualitativo, si registrava un incremento dei livelli di assistenza fornita dai “badanti informali” e, sul piano sistemico, un avanzamento verso quell'integrazione dei servizi atta a favorire un approccio olistico, basato sul fabbisogno dell'assistito nel suo carattere complesso e

multidimensionale, identificata dalla Ricerca al suo massimo grado di espressione come integrazione “pervasiva”, ultronea ai tradizionali sistemi di integrazione amministrativa e operativa, pubblica o pubblico/privata, proiettata verso un nuovo modello di intervento in cui erogatori e beneficiari del servizio co-producono i meccanismi di innovazione dei processi e dei ruoli, per massimizzare la creazione di valore pubblico (Carretero et al., 2015, p. 45).

Sotto l'aspetto degli *users* le valutazioni soggettive dell'esperienza confermavano, come anticipato, quelle registrate dal progetto ACTION: incremento del senso di sicurezza degli assistiti, riduzione dello *stress* dei beneficiari intermedi dei servizi e percezione diffusa di una migliore qualità del vivere quotidiano sia da parte degli utenti finali che da parte dei “badanti informali”, soprattutto, per questi ultimi, in termini di conciliazione dei carichi lavorativi e familiari.

As a matter of fact – argomentavano nel merito gli autori della Ricerca – integration has evolved significantly over the last decade as governments search for ways to address beneficiaries' needs better and, at the same time, manage increase caseloads with reduced resources. Although integrated approaches to social services provision is not a new concept, we are in an exciting period of innovation characterized by schemes based on traditional and emerging ICTs, new funding models, and a more dynamic relationship between governments, citizens, and service providers from the private and not-for-profit sectors (Carretero et al., 2015, p. 44).

Il nuovo paradigma implementato a livello sistemico dal PNRR sulla scorta delle molteplici sperimentazioni localizzate e diffuse realizzate su scala comunitaria si pone quindi, in questa lettura, come catalizzatore di una integrazione dei servizi perseguita per decenni dalle politiche programmatiche.

Sovente le sperimentazioni hanno implicato un utilizzo integrato delle diverse tecnologie disponibili, coniugando ad esempio aspetti di *Telecare* e *Home Automation (smart home)*, come nel caso del progetto francese “Advanced Telecare” – servizio intersettoriale alimentato da molteplici fondi pubblici e subcontractualizzato a società private – peraltro unico nel suo essere totalmente integrato nella catena di erogazione dei servizi. Nel caso specifico il progetto era teso a prevenire le cadute in casa delle persone anziane e prevedeva l'installazione, all'interno dell'abitazione, di sensori e rilevatori di movimenti anomali o variazione delle condizioni ambientali (presenza di gas, fumo, cambiamenti di temperatura...), integrata con la predisposizione di ausili mirati, quali percorsi luminosi per “accompagnare” l'anziano negli spostamenti prevedibili più ad alto rischio (ad esempio dal letto al bagno nelle ore notturne). All'intervento automaticamente innescato dalla segnalazione di eventi critici da parte dei dispositivi elettronici installati si aggiungeva la possibilità, per l'anziano, di contattare un *call center* in qualsiasi momento, anche solo in caso di solitudine, con l'ulteriore impegno da parte dello stesso *call center* a chiamare ogni assistito almeno

una volta al mese. Lo studio mirato condotto in Francia per valutare la sperimentazione documentava, parimenti agli altri progetti richiamati, una riduzione effettiva delle cadute e degli accessi in ospedale dovuti a caduta in ambiente domestico; una riduzione del fabbisogno di assistenza, “reduction in the need for carers’ time” (Carretero et al., 2015, p. 59), ed un efficientamento dei costi, essendo il servizio economicamente meno oneroso del ricovero ospedaliero evitato.

Riscontri analogamente positivi sono stati documentati in rapporto alle 8 iniziative focalizzate sull’ “integrated care”. Si trattava in questo caso di progetti prevalentemente attestati su un livello di innovazione incrementale/organizzativa per quanto concerne il fattore *enabling* ICT, ma rispondenti agli indicatori di innovazione sociale individuati; anch’essi spalmati, come i precedenti, sulle diverse soluzioni tecnologiche applicative proposte.

Mappata in questo *cluster* è stata anche una iniziativa italiana, bolognese, di livello regionale e caratterizzata da una *governance* “intersettoriale”, perché estesa all’università, al privato e al terzo settore. Si trattava, in sintesi, della gestione condivisa di una cartella elettronica integrata informativa dell’utente mediante un *call center*, con funzioni di intermediazione fra assistito e sistema pubblico e privato dei servizi e di monitoraggio, effettuato con chiamate settimanali agli anziani coinvolti nella sperimentazione. Anche questa iniziativa testimoniava della soddisfazione di tutti gli *stakeholders*, del maggior senso di sicurezza dell’utente, della maggiore efficacia delle cure per l’aumentato grado di consapevolezza e il tutoraggio dell’assistito, nonché di ridotti tassi di ospedalizzazione.

Più spinto il grado di innovazione di due sole iniziative di *Telecare*, fra cui una sperimentazione spagnola che condivideva con il programma svedese di cui sopra, a supporto dell’autonomia dell’anziano, l’utilizzo dell’ “ICT to empower the patient’s care community to be more effective and efficient in supporting the individual at home” (Carretero et al, 2015, p. 65). Nel caso specifico si introduceva un sistema di monitoraggio da remoto dell’assistito con applicazioni per l’accesso a percorsi di *training* rivolti agli assistenti informali e per favorire la comunicazione, tramite videoconferenze, fra tutti i soggetti che avevano in carico l’utente (sistema dei servizi, operatori professionali e badanti informali).

Nel rapporto del JRC denominato “Technology-enabled Services for Older People Living at Home Independently: Lessons for public long-term care authorities in the EU Member States” – lezioni per le politiche di settore tratte dall’esperienza maturata nell’ambito del progetto ICT-AGE JRC IPTS concernente l’analisi trasversale delle buone prassi sperimentate – l’autrice, Stephanie Carretero, schematizza come segue i benefici generati dalle 14 iniziative indagate, in rapporto ai 4 obiettivi del SIP (*independence, carers’ productivity, quality of care, cost-efficiency*):

Policy objectives:

To increase the independence of both older adults and carers

Findings of the practice evidence:

Older adults were more independent because they:

- were less reliant on informal carers for daily activities

- had fewer falls at home and fewer hospitalizations as a result of falls

- were less affected by depression

- were more active when alone

- had better walking ability, balance and speed

- were better able to maintain their physical and cognitive status

- had better cognitive functions

- remained in their homes

- felt safer at home

Informal carers were also more independent in their tasks and responsibility, they reported:

- better quality of life

- better health

- more safety

- more freedom and peace of mind

- less stress

To increase the productivity of carers

Carer workers were more productive because the technology-enabled services:

- saved time and money

- facilitated timely medical responses in emergency conditions

- reduced the length of patients' visits

- provided more information on care

- provided more time to be dedicated to more basic needs

- improved carer workers' satisfaction with the job

- allowed new ways of working

- helped them to have more respect for the independence and dignity of older people

- helped them have fewer worries

- helped them retain their jobs

- helped them to improve their relationship with the older adults

To increase the quality of care

More quality of care:

- reduction of medication non-adherence

- improvement of medication safety

- maintenance of quality of care

- increase in the care competence of informal carers

- more client satisfaction with the care provided by carer workers

- reduction in mortality rate

To improve the sustainability of the care system

Cost-effectiveness:

- low cost of the technology

- reduction in total average (mean) costs of care

- reduction in the use of institutional care

- reduction in the use of health care resources

- more rapid discharge from hospitals

- reduced hospital admissions

- avoided nursing home admissions

- Savings for the care system

(Carretero, 2015, p. 15).

Nello stesso rapporto la Carretero sottolineava l'importanza del coinvolgimento degli *stakeholders* per costruire fiducia e sicurezza, “portatori di interesse” identificati con:

Public authorities at different levels (according to their competences) and in different areas (authority responsible for health, for social affairs, or a specific public authority responsible for long-term care).

Care providers responsible for providing care to older adults. These providers could be the same public bodies (if care was delivered through the public health and social care services), not-for-profit organisations receiving public funding to deliver such service, and/or for-profit organisations (if a private company was delivering care services).

Technical providers in charge of providing, installing and maintaining the devices.

Research institutes that evaluate the use and impact of the technology-enabled service.

European, national, regional or local organisations that funded the evaluation, creation or implementation of the service

Public training bodies in charge of training professionals, who installed, maintained and used the service, and universities that trained health and social care practitioners to use technologies.

Older people and their informal carers and their representative organisations, who were involved as primary stakeholders in all stages of the design, implementation and evaluation of the service (Carretero, 2015, pp. 19-20).

Per ciascun “portatore di interesse” il *Report* specificava le modalità di cooptazione. In particolare, per gli utenti dei servizi, anziani e “badanti informali” i ricercatori indicano che:

Older people and informal carers received training from *TELBIL*, *TDP* and *Action*. *TELBIL* used the same training programme based on demonstration that was used for health care professionals. *Action* included this training as a specific module of the service, and local partnerships in *TDP* trained older people and their informal carers in the use of telecare. *Action* also consulted older people and informal carers from the initial design stage of the service through to its implementation, in order to account for their needs in building the new service (Carretero, 2015, p. 21).

Lo spazio di coinvolgimento attivo spaziava, quindi, dal rafforzamento delle capacità e competenze mediante programmi formativi tesi ad abilitare l'utente all'utilizzo degli strumenti messi a disposizione ad una formazione più estesa, per gli “informal carers”, sui contenuti propri del ruolo informale di supporto, fino alla compartecipazione nella progettazione ed implementazione dei servizi.

L'importanza di una condivisione iniziale di questo tipo, degli obiettivi e della visione sottesa alle iniziative proposte veniva analogamente messa in luce dallo studio italiano IRS 2012, che

sottolineava l'impatto dell'introduzione dell'ICT nell'erogazione dei servizi sociali su meccanismi e processi consolidati, con conseguente ridefinizione tanto dell'approccio quanto dei *target* coinvolti: "It most likely also means tearing down walls between sectors (e.g. between family carers and professional care, between preventive and curative care) and collaboration or networking (i.e. flexible relationships between independent partners) inside or outside the social services sector" (Crepaldi, De Rosa & Pesce, 2012, p. 15).

Si tratta, in sostanza, dell'inveramento dell'approccio olistico richiamato sopra, in cui sfumano i confini fra assistenti formali ed informali, fra pubblico e privato e fra prevenzione e cura, in una cornice flessibile che oltrepassa i confini della programmazione settoriale dei servizi. Da qui l'opportunità di una collaborazione che si esprima anche in termini di co-creazione del servizio stesso: "In particular the relational dimension implies that at different deployment stages, the customer can participate in producing the innovation (collaboration). The customer/user can therefore be one of the actors of the innovation, the success of which depends on the quality of this interaction" (Crepaldi, De Rosa & Pesce, 2012, p. 15).

Passando in rassegna la letteratura esistente sul tema dell'innovazione sociale nei servizi, lo studio individuava nella partecipazione attiva degli utenti un tratto essenziale del processo: "Literature evidences that involvement of final users is considered an innovation able to promote equality, effectiveness and control, and adherence to the needs of users" (Crepaldi, De Rosa & Pesce, p. 46).

Il modello di riferimento è quello della *user driven innovation* cosiccome analizzata dai danesi Agger e Lund, che definiscono "buona prassi" l'identificazione dell'*end-user* con un co-produttore del processo di innovazione, arrivando a porre l'esigenza di una istituzionalizzazione di tale identificazione sul piano stesso delle politiche programmatiche (Agger & Lund, 2011, p. 194).

Lo sviluppo del "self help sector" come insieme di servizi innovativi che consentano all'anziano di vivere in modo indipendente emerge dunque da tutte le ricerche richiamate quale efficace modello di risposta alle esigenze dell'assistito nella misura in cui ne favorisce l'incremento della qualità della vita, la percezione di sé come individuo in grado di decidere per se stesso, la partecipazione e l'inclusione sociale:

The promotion and *development of the self-help sector*, mainly pushed through increate networking of self-help institutions, represents an essential element for the empowerment of disabled people – osservano le autrici dello studio IRS – by providing care in another way, user-oriented objectives are realized: quality of life (social, emotional and physical health), self-reliance (to decide about one's own care project and life in general with adapted support), participation, cohesion and inclusion (Crepaldi, De Rosa & Pesce, 2012, p. 46).

Il massiccio investimento del PNNR nella “Casa come primo luogo di cura” procede in questa direzione. A fronte delle possibilità dischiuse dalla domotica assistiva sembra tuttavia meritevole di ulteriore approfondimento e riflessione, almeno in proiezione, il tema della linea di demarcazione fra abilitazione e intrusione: quando, in altre parole, la “tecnologizzazione” dell’ambiente domestico (*smart home, smart care, smart wellness*), atta a garantire che una anomalia clinica (monitoraggio da remoto dei parametri vitali), comportamentale (installazione di sensori su letto, poltrona, porta del frigorifero...) o ambientale (rilevatori di fumo, gas...) possa essere tempestivamente rilevata, passa da fattore abilitante una maggiore “peace of mind” dell’anziano a elemento intrusivo?

Analogamente, fino a che punto la sollecitazione tecnologicamente mediata delle facoltà fisiche e cognitive residue (*PC/WEB training programs, social robots...*) apporta un valore aggiunto al benessere dell’anziano?

Esiste ancora nelle strategie attuali di implementazione della digitalizzazione una priorità attribuita al rafforzamento della dimensione comunitaria del welfare, su cui sembravano puntare alcune delle sperimentazioni richiamate?

4. Etica della cura e prospettive critiche

Prendiamo atto delle valutazioni positive raccolte presso gli utenti, anche finali (quindi gli anziani assistiti), dei servizi sperimentalmente attivati. Ci sembrano tuttavia meritevoli di attenzione, in merito alle questioni aperte di cui sopra, alcune riflessioni della sociologa della scienza e della tecnologia Maggie Mort autrice, con altre colleghe, di una pubblicazione dal titolo emblematico *Ageing with telecare: care or coercion in austerity?* (Mort, Roberts & Callén, 2012), comparsa nello stesso periodo in cui venivano realizzati gli studi valutativi precedentemente menzionati. La Mort osserva, in particolare, come la retorica dell’invecchiamento demografico sia tendenzialmente impiegata nel settore politico ed economico enfatizzando l’aspetto critico dell’insostenibilità finanziaria del fenomeno, da cui l’imperativo dell’efficientamento del sistema del *welfare*, che rimanda all’utilizzo massiccio e diffuso delle ICT. Questo approccio, che la sociologa inglese definisce una vera e propria stigmatizzazione dell’invecchiamento della popolazione nella società contemporanea, si pone alla base della prefigurazione di un modello di assistenza teso a preservare l’autonomia dell’anziano nella propria abitazione.

The proliferation of terms such as ‘demographic ageing’ and forms of measurement such as the ‘age dependency ratio’ (the share of the population aged under 20 and over 64 as a percentage of the group aged 20–64) and the way these are employed in commercial, policy and some practice contexts, act to

frame older people themselves as essentially problematic. In living longer they are seen to place an unsustainable economic burden on care services. Such stigmatising and ageist rhetoric forms the backdrop to exhortations that older people must remain independent and stay in their own homes (Mort, Roberts & Callén, 2012).

A testimonianza del carattere dominante di questa visione la Mort chiama in causa la comunità scientifica, le istituzioni – la sociologa inglese richiama ad esempio un simposio internazionale sull'*healthcare* tenutosi a Barcellona nel 2009 ospitante l'intervento del dirigente di una società di *healthcare software* secondo il quale lo Stato Sociale si sarebbe trovato ad affrontare un "silver tsunami" a cui si sarebbe potuto fare fronte unicamente utilizzando tutte le tecnologie e soluzioni *software* disponibili – e la stampa, citando articoli allarmistici comparsi sui più noti quotidiani britannici in merito alla "Britain's age time bomb" (Walker, 2011).

Oltre 10 anni dopo il tenore della comunicazione si conferma il medesimo, riportiamo a titolo esemplificativo, restando questa volta in Italia, un articolo comparso sul Sole 24 Ore nel febbraio del 2024 con l'inquietante titolo: "Welfare incompiuto, allarme Censis: rischio di collasso sociale", basato su un'indagine realizzata da Assindatcolf (Carli, 2024).

In uno scenario in cui alla crescita progressiva della domanda di assistenza fanno riscontro, da una parte, il venire meno delle risorse pubbliche per il Welfare e, dall'altra, le pressioni di un emergente mercato dell'high tech il modello dello "staying in your own home", efficacemente coniugando l'immagine dell'anziano indipendente con il concetto di dignità dell'individuo, viene dipinto come il futuro ideale, ("as the ideal future for older people" (Mort, Roberts & Callén, 2012), popolato da *telecare packages*, *monitoring systems* e *homecare robots*.

Da qui la legittimazione etica e culturale di uno sviluppo esponenziale delle tecnologie per l'assistenza da remoto. La Mort ricorda la presenza di ben 45 aziende del settore impegnate in altrettante dimostrazioni di sistemi *high tech* alla conferenza di Londra "International Telecare and Telehealth" del 2011, organizzata dall'Associazione dei Servizi di *Telecare*, nonché l'impegno crescente di società e centri di ricerca nella progettazione di ambienti di vita accessibili e abilitanti (Loe, 2010, pp. 319–34).

A riscontro di ciò, ancora *Il Sole 24 Ore* il 13 maggio del 2024 ha dedicato un articolo alla *Silver Economy* definendola un modello economico emergente "sulla creazione di prodotti, servizi e nuove opportunità di lavoro mirate ai consumi e alle esigenze della popolazione con più di 65 anni", inclusivo della produzione di nuove tecnologie per facilitare la vita degli *over 65*, come ad esempio tecniche assistive, dispositivi medici, servizi di assistenza domiciliare, robotica assistenziale, e dello sviluppo di programmi e applicazioni riguardanti benessere e salute.

"Il modello della silver economy – si legge nel testo – riconosce gli anziani come consumatori attivi e partecipanti al mercato del lavoro e permette loro di contribuire (ancora) alla crescita

economica grazie al loro potere d'acquisto e alle loro competenze professionali" (*Il Sole 24 Ore*, 2024).

La *silver economy* delineata dal quotidiano economico è perfettamente in sintonia con la rappresentazione fornita quasi 20 anni prima in un documento politico-programmatico inglese di cui riferisce la Mort, in cui si tratteggia il futuro dischiuso dalle ICT nel settore dell'assistenza come luogo di progresso sociale, in cui sarà pienamente restituita dignità alla condizione anziana:

The number of people requiring community based health and social care support, and the levels and complexity of those needs, is expected to increase considerably over the next decade ... people will have higher expectations ... They want independence, and after a life-time's work, they want and are entitled to, dignity for life ... Telecare is vital to unlocking this future ... It is not realistic to plan to deliver care and support services in the way we do at present (*Department of Health*, 2005).

La critica della Mort a questa visione dell'invecchiamento metteva in luce il rischio che i sistemi di monitoraggio da remoto potessero manifestarsi, ad un grado più avanzato di sviluppo di un sistema assistenziale *tech-driven*, come strumenti coercitivi soprattutto laddove all'utente, in un contesto di taglio di risorse e prestazioni, non fossero offerte opzioni alternative al vivere in una sorta di "telecasa" ("*telehome*").

A riscontro della portata per lo meno potenziale dell'intrusività paventata dalla sociologa britannica si pensi all'adozione di strumenti quali i sensori posti negli spazi vitali dell'assistito, ai segnali automaticamente inviati dai rilevatori di movimento installati all'interno dell'abitazione e, più in generale, alla generazione e condivisione di dati individuali sensibili in ragione di obiettivi di "sicurezza", suscettibili di sovrapporre surrettiziamente le esigenze dei "carers" a quelle, definite a priori, dell'utente.

"Indeed – osserva la Mort – it has been noted that many telecare products have developed from surveillance technologies used on children or prisoners under house arrest" (Mort, Roberts & Callén, 2012).

È verosimilmente opportuno nel merito rilevare come spesso le *policy* di settore ricorrano indistintamente al concetto di "sicurezza" e a quello, che non traduciamo per non perderne l'efficacia, di "peace of mind" dell'assistito.

Un altro spunto di riflessione offerto dallo studio documentato dalla Mort su cui crediamo valga la pena soffermarsi attiene all'utilizzo non conforme ("*misuse*") di alcuni dei dispositivi messi a disposizione degli utenti nei progetti di domotica assistiva analizzati, nel caso specifico il dispositivo di *telecare* appeso al collo dell'anziano: il gruppo di lavoro coinvolto nella sperimentazione riferiva di avere diffusamente registrato un'attivazione delle chiamate non innescata dai dispositivi posizionati nell'abitazione, bensì da parte dello stesso utente, con la mera

richiesta di una visita o di poter scambiare due parole al telefono, pratica per altro consentita agli operatori ingaggiati in caso di mancanza di chiamate in attesa. Tale fenomeno, riportava la Mort, era stato causa di discussione all'interno del *project management*, laddove l'intenzione di rimuovere il sistema in caso di "false chiamate" si scontrava con quella di mantenerlo codificando, anzi, le richieste quale importante segnale di un fabbisogno non soddisfatto dal progetto di *Telecare* in atto, oltre che probabile causa di un incremento dei costi, associato al rischio di perdere "chiamate" di emergenza.

L'autrice denunciava il carattere riduzionistico di una valutazione dei casi di *misuse* meramente funzionale all'inclusione o meno dell'utente nel servizio di *telecare*, unicamente orientata ad attribuire o meno al potenziale utente l'identità di "user" senza prendere in considerazione altre, più complesse, possibilità di analisi della condizione dell'assistito, rimandanti ad esempio all'esigenza di rafforzarne la rete relazionale. Ciò in un contesto di generale depauperamento del sistema sociale dei servizi:

In the case of older people assessed as in need of care and subsequently given the identity of a telecare user – argomenta la Mort – other identities and histories are occluded, like those that might be included in a broader sense of health, such as the identity of the user as a person with social and communicative needs but also with social capital and networks that must be ensured or even reinforced (Mort, Roberts & Callén, 2012).

Ulteriormente la ricerca condotta in Spagna e in Inghilterra faceva emergere ricorrenti casi di rifiuto dell'ausilio elettronico, tendenzialmente motivato dagli anziani con la negazione di un fabbisogno effettivo di "pronto soccorso", nonché la testimonianza di una forte pressione da parte del sistema informale e formale dei "carers".

Se l'uso difforme dei dispositivi avesse teso ad essere decisamente scoraggiato, la resistenza all'utilizzo del dispositivo "da collo" sarebbe stata il più delle volte ignorata tanto dagli operatori quanto dai gestori dei servizi e dai rappresentanti del sistema economico (industria informatica), riferisce la Mort; quando la riluttanza dell'assistito portava ad una conseguenza, essa rifletteva tendenzialmente un atteggiamento di tipo punitivo, determinante la rimozione dell'ausilio.

Nella prospettiva proposta dalla sociologa inglese il riconoscimento di casi di resistenza e utilizzo difforme degli strumenti rappresenta invece un'opportunità importante per non percorrere la strada della coercizione, abbracciando un approccio diverso, non riduzionistico, al *Telecare*, attento alla dimensione più squisitamente umana e sociale della cura. Partendo dal presupposto osservativo che la maggior parte degli anziani *target* dell'indagine non interagiva attivamente con il sistema di monitoraggio da remoto, immaginato e progettato come universale e oggetto di una

fruizione passiva ad eccezione di situazioni di emergenza, la Mort identifica proprio nell'interazione la leva per delineare una nuova strategia di assistenza.

La proposta avanzata era sostanzialmente quella di ripensare il *Telecare* in modo flessibile, inclusivo e creativo, come sistema personalizzabile, da svilupparsi in relazione al contesto e alle caratteristiche soggettive dell'utilizzatore. Tale possibilità, scarsamente compresa allo stato dell'arte documentato dalla stessa Mort e difficile da realizzare in un clima di austerità economica, rimanda ancora una volta ai temi della partecipazione e della condivisione, declinati in un settore in cui sono evidentemente elevate le implicazioni sul piano etico e socio-antropologico.

In a climate of austerity – commentava la sociologa britannica – ethical telecare services must guard against closing down daily experimentation, mutual appropriation and creative and innovative socio-technical practices [...]. Approaching telecare as shared work means rethinking categories such as effects, effectiveness and causality [...] How does telecare work at all? It does not do so on its own but through the (quality of) relations between things and people and connections. Clearly the very limited range of relations and creative use (or non-use) of telecare that can be found in lived realities raises questions about what 'proper use' is thought to be, and how this is reflected in the design and implementation of systems of home care. Our material shows that this varies across different contexts and that telecare should not be understood as a universal solution, as much of the industry and policy literature implies, but as a situated one [...]. The examples of practice found in our study demonstrate that it is not telecare itself that creates or even affords independence. We might argue that, contrary to the idealised scenarios in the glossy brochures, good (tele)care does not pre-exist the user but is an outcome of shared work" (Mort, Roberts & Callén, 2012).

L'attribuzione di centralità al coinvolgimento attivo degli utenti e alla condivisione degli obiettivi perseguiti è in effetti in linea con le conclusioni degli altri studi richiamati sopra. Ci sembra, tuttavia, particolarmente interessante nella lettura della Mort non solo l'attenzione critica posta al rischio di una violazione dei diritti di libertà individuale – rispetto alla quale, come abbiamo visto la Mort individuava una possibile *way out* nella personalizzazione e nella co-creazione del servizio di assistenza e degli ausili di cui si avvale – ma anche l'invito a prendere in considerazione una visione dell'"aging in place" meno focalizzata sull'*upgrading* tecnologico e più attenta a fattori differenti, che rimandano ad esempio al fabbisogno relazionale, valorizzando in questo contesto il repertorio delle tecnologie esistenti, non necessariamente concepite in senso medico, o comunque "health oriented", integrate all'interno di strategie personalizzate di gestione del quotidiano che enfatizzino gli aspetti della continuità e dell'autonomia dell'anziano.

Nella sua analisi la Mort si ispira alla visione di Meika Loe; il sociologo statunitense parte dal presupposto largamente condiviso che, alla luce delle nuove possibilità di assistenza e cura

all'interno dell'abitazione abilitate dalla tecnologia, la maggioranza degli anziani preferisca invecchiare in casa propria, ravvisando in ciò l'opportunità di preservare un senso di normalità, continuità, identità personale, autonomia, dignità e controllo della propria vita:

Given the blurring of home and healthcare, elders themselves state a preference for ageing at home. Forty years of national survey research, including that of the American Association of Retired Persons (AARP), reveals that the vast majority of elders prefer to age at home. Elders cite normalcy, continuity in self-identity, autonomy, and control as reasons for this preference [...] – argomenta l'accademico statunitense. Many elders associate ageing in place with staying at home and in being in control of their day-to-day lives; avoiding institutionalisation can help to preserve autonomy and dignity (Loe, 2010, pp. 319-334).

Focalizzando in particolare l'attenzione sulla popolazione femminile, Loe osserva come molte donne novantenni siano state in grado di sviluppare, durante la propria vita, un approccio creativo e personalizzato alla coniugazione fra “self-care” e utilizzo della tecnologia - tecnologia domestica e “del quotidiano” – e lo riutilizzino con successo nell'ultima fase della propria esistenza:

Many women nonagenarians have developed creative approaches to technology use and self-care, and utilise these daily in their final years of life – osserva il sociologo statunitense – These techniques include designing purses for phones, staging walkers at key places, using reading machines to connect and escape, creating telephone monitoring networks among friends, sewing items to enable comfort at home, using computers to keep track of household finances, and utilising simple kitchen technologies to create affordable healthy meals. In these ways, nonagenarians further their lifelong expertise in using technologies that assist with care work and the co-ordination of routine mundane tasks. This expertise, now taken for granted after years of practice, comes in handy when it comes to food preparation, health monitoring, connecting with others, and creating and maintaining a home over many years. Such domain-specific knowledge can be a key advantage and safety net in the context of ageing, allowing elders the ability to prove self-efficacy and remain at home and healthy, late in life (Loe, 2010, pp. 319-334).

La tecnologia si pone in questo caso come ausilio fondamentale soprattutto nel suo essere tecnologia accessibile e ordinaria, potremmo dire. Da qui l'appello agli attori delle politiche programmatiche affinché tengano adeguato conto, anche negli investimenti a supporto dell'“Aging in place”, di questa dimensione, evitando di concentrare tutti gli sforzi su strumenti sempre più avanzati di *care 3.0*:

While much political rhetoric focuses on access to care issues, key technologies like telephone and television services, as well as other communications technologies, kitchen technologies, and mobility

technologies can be ‘the difference between life and death’. And they can also be costly. Age-based technology discounts and coverage are crucial to ensuring access to these ‘health’ technologies. Policies aimed at assisting elders as they pursue health, community, and changing forms of independence need to address the escalating costs of communication technologies in the home. These old women’s lives reveal that perhaps the answer to the question about an ageing populace, home support and health is not simply new biotechnologies, assistive technologies or design strategies, but a renewed emphasis on elder agency and an awareness of existing technology repertoires and daily strategies to emphasise continuity and autonomy. After all, despite all the emphasis on successful ageing, elders in this project ultimately aim to achieve something more akin to comfortable ageing that emphasises ease (Loe, 2010, pp. 319-334).

Per queste novantenni, conclude Loe, un’etica del “self-care”, dell’auto-aiuto, dovrebbe assecondare e preservare un’accezione estesa del significato di “salute”, che faccia leva sui concetti di *comfort*, sicurezza, autonomia, continuità ma anche di capitale sociale perché il potere più autentico che la tecnologia abilita è quello riconducibile alla componente relazionale e umana sottesa al suo utilizzo, che ne fa uno strumento chiave per rafforzare le reti sociali, per tutelare la continuità nel corso della vita, per rendere possibile la partecipazione intellettuale e favorire il benessere fisico:

For these nonagenarians, a self-care ethic is about accomplishing and maintaining a broad sense of health that involves comfort, confidence, continuity, autonomy *and* social capital in the context of old age. The real power of technology, as many of these nonagenarian women reveal, is in the implicit social relationships and other manifestations of humanity that underlie our use of tools and devices – conclude Loe. As we have seen, self-efficacy and ageing in place are realistic goals when technology can be used to reveal and reinforce social networks, ensure continuity across the lifecourse (when it comes to everyday routines and roles), and enable intellectual participation and physical wellbeing (Loe, 2010, pp. 319-334).

Una efficace rappresentazione dell’importanza dei fattori chiamati in causa dal sociologo statunitense e dalla Mort è verosimilmente fornita dalle note di una donna ingaggiata in una delle sperimentazioni, intervistata nell’ambito della ricerca presentata dalla stessa Mort. Nella nota l’assistita prova sinteticamente a puntualizzare gli elementi salienti di una “everyday care strategy”:

“Arrange things to give peace of mind

Tel. in every room

Key holders

My friend and I ring each other every morning

Neighbours know I am around when my curtains are opened” (Loe, 2010, pp. 319-334).

Si tratta, evidentemente, di elementi non astraibili dal contesto socio-relazionale di riferimento, che forniscono informazioni chiave per disegnare il modo specifico in cui l'utente in questione può convivere con un sistema di *Telecare*; un sistema che, nella prospettiva proposta dai due autori, dovrebbe essere configurato nell'orizzonte di narrative differenti dei concetti di invecchiamento, tecnologia e relazione di aiuto (assistenza). Soprattutto rileva, ci sembra, il rimando forte della maggior parte (ben 4 su 5) degli elementi menzionati dall'anziana signora agli "altri", alla dimensione relazionale e comunitaria (*friends, neighbours...*).

A fronte di una spinta forte e costante verso l'*upgrading* tecnologico che alimenta il mercato "silver" della domotica assistiva è dunque verosimilmente opportuno continuare a riflettere su ciò che viene posto al centro dei programmi di *telecare* e sul rischio che le pressioni degli *stakeholders* economici schiaccino o semplifichino riduzionisticamente il fabbisogno autentico, per altro sovente inespresso, della controparte più debole, l'utente finale dei beni e servizi prodotti, rafforzando e diffondendo una interpretazione mistificata del binomio "aging in place"/tutela della dignità umana dell'anziano.

A fronte di ciò ci sembra altresì opportuno tenere accesa l'attenzione anche sul potenziale coercitivo della tecnologizzazione in chiave *care 3.0* dell'ambiente abitativo.

Ancora più critica della prospettiva della Mort è, in questo ambito specifico, la lettura dell'americana Alison Marie Kenner che, analizzando in particolare l'applicazione delle tecnologie per il *telecare* alla demenza senile, identifica nei sistemi di monitoraggio (significativamente definiti "surveillance technologies for the elderly" (Kenner, 2008, p. 252) la declinazione di una concezione dell'invecchiamento che va riconosciuta come propria di uno specifico paradigma economico, politico e socio-culturale, di cui ricalca inevitabilmente e pericolosamente le geometrie di potere.

Biomedicalizzazione della condizione anziana e crescita esponenziale del mercato dei prodotti e servizi per la terza età sono, per la Kenner, fenomeni strettamente interconnessi, espressione di un approccio tendente a considerare l'anziano come un'astrazione, sradicandolo dal contesto di riferimento: se l'industria dell'invecchiamento ne isola e tipicizza esigenze e *desiderata*, il modello della medicalizzazione, socialmente costruito per quanto basato sulla rilevazione di parametri biologici intrinseci al fattore anagrafico, ne patologizza la rappresentazione, ridefinendo altresì il concetto di cura ("care") ed imprimendo una direzione specifica alla ricerca, alla formazione sanitaria, alla cultura e alle politiche di settore. Entrambi, industria dell'invecchiamento e cultura della medicalizzazione, tendono a giustificare e consolidare la visione forgiata dal paradigma, occultando i fattori sociali, politici ed economici che agiscono sulle condizioni di questa fascia di popolazione. Un fenomeno, quest'ultimo, che in modo sempre più evidente si manifesta anche in relazione al mondo dei giovanissimi, con la progressiva medicalizzazione delle difficoltà di

apprendimento: si pensi alla dilagante presenza di bambini diagnosticati come “DSA” (Disturbi Specifici dell’Apprendimento) a partire dalla scuola primaria, con conseguente attivazione – sancita dalla L.170 dell’ottobre 2010 – di percorsi didattici personalizzati.

Nella fattispecie, i meccanismi della biomedicalizzazione dell’invecchiamento portano, osserva la Kenner in relazione alla demenza, a creare e definire fattori di rischio basati sull’osservazione dei casi, potenzialmente in grado di abbassare progressivamente la soglia della normalità percepita, ponendo le premesse per una massiccia estensione delle pratiche di sorveglianza e alimentando, con ciò, una proliferazione dei sistemi di monitoraggio *ICT based*.

Per quanto la dimensione generale dell’invecchiamento non sia sovrapponibile a quella della demenza, il tema resta senz’altro rilevante, essendo la demenza una patologia molto diffusa, che si sviluppa solitamente in età avanzata e tipicamente presente nell’immaginario collettivo come problema caratteristico della persona anziana.

La Kenner riporta i contenuti inquietanti di studi effettuati per la rilevazione precoce dei sintomi della malattia mediante l’utilizzo di sistemi di rilevazione di dati biologici e ambientali, fra cui il ciclo del sonno e gli schemi di mobilità della persona monitorata all’interno e all’esterno della propria abitazione, costruiti mediante specifici algoritmi.

I ricercatori impegnati in uno studio precedentemente realizzato in Giappone (Suzuki et al., 2006, pp. 41-44.), menzionato dall’autrice, avevano ad esempio individuato nella tendenza all’inattività e nei problemi del sonno efficaci indicatori di un esordio di demenza; in questo modo, secondo la Kenner, le tecnologie per la sorveglianza contribuiscono all’identificazione e categorizzazione di una tassonomia della malattia, corroborando il modello della biomedicalizzazione dell’invecchiamento.

In particolare, la Kenner pone il problema delle decisioni prese “nell’interesse dell’anziano” in base ai dati rilevati e, conseguentemente, dell’impatto della pratica di sorveglianza sulla vita dell’individuo. Lo stesso essere oggetto di monitoraggio può inibire l’azione e con ciò incidere sulla libertà del soggetto:

Decisions may be made on the behalf of elderly people or in their best interest based on somatic or ICT data. If each and every movement is being tracked, recorded and evaluated, will this result in reactions that restrict the activities of elderly people? If monitoring systems are installed in homes and data show that ambulation is random, lapping or pacing, thus indicating increased risk of wandering, will elders then be restricted from leaving their homes? Power and unequal social relations will undoubtedly infuse such situations – particularly when an elder is dependent upon family or social services and costs of aging in place are significantly less than institutional facilities (Kenner, 2008, p. 263).

L'asimmetria di potere sociale denunciata dalla Kenner fra l'anziano, in particolare l'anziano dipendente dalla famiglia o dai servizi sociali, e il sistema esteso dei "care givers" tenderà verosimilmente ad incidere sulla decisione di mettere "in sicurezza" il soggetto identificato come "demente", soprattutto alla luce della significativa differenza di costi fra la permanenza a casa e l'istituzionalizzazione. Assunta l'asimmetria di potere generata dalla variabile anagrafica, nel contesto delle politiche per l'invecchiamento le tecnologie per la sorveglianza degli anziani affetti da questa patologia finirebbero con il creare nuove forme di controllo sociale: "Power asymmetries between different age groups – conclude la Kenner – are nothing new, but within the political economy of aging, surveillance technologies for the elderly are both reproducing and creating new forms of social control" (Kenner, 2008, p. 264).

In sintonia con quanto osservava la Mort in merito alla scarsa rilevanza attribuita alle richieste implicite o esplicite di attenzione da parte degli assistiti, sintomo di un fabbisogno più complesso rispetto a quello cui la sperimentazione si candidava a rispondere (ricordiamo l'utilizzo difforme degli ausili di tele-allarme forniti), la Kenner sottolinea l'interpretazione riduzionistica del concetto di benessere propria del paradigma dominante e alimentata dalla diffusione dei modelli emergenti di assistenza, che sembrano mettere al centro il fattore della sicurezza fisica:

For older Americans, wellbeing has come to mean physical safety where the greatest threat is posed by the aging body itself. The examples of monitoring systems in institutional facilities, homecare, and law enforcement agencies demonstrate that data generated by surveillance technologies often result in increased monitoring, access or mobility restriction, or further medical interventions – these are the reactions of persons who are caring for or otherwise socially responsible for elderly citizens. Interventions may or may not be necessary; what is important is that judgments and decisions are not made by the elderly themselves, but by those who are responsible for their wellbeing (Kenner, 2008, p. 266).

Ancora una volta questo tipo di risposta sarebbe condizionato dal modello della medicalizzazione e della mercificazione dell'invecchiamento, acriticamente assunto senza averlo adeguatamente analizzato all'interno di un orizzonte storico ed economico-politico-culturale specifico.

Le tecnologie diventano, in questo contesto, strumento conservativo, di replicazione e consolidamento delle condizioni esistenti, in grado di delimitare e gestire la popolazione a rischio, giustificandone e cristallizzandone una situazione di dipendenza funzionale alla crescita del mercato:

Despite their usefulness in the context of elderly care, RFIDs and monitoring systems must still be understood as technologies of control that are predisposed to reproduce existing inequalities in social relations. By situating surveillance of the elderly within the political economy of aging, it is evident that

these technologies both construct and manage a risk population in an increasingly profitable healthcare regime. Researchers, innovators and adopters should be attentive to the ways in which monitoring systems reproduce dependencies, contribute to disease taxonomies, redefine the meaning of care, individualize risk, and enable or constrain liberties. Surveillance systems can address the needs of caregivers, the elderly who wish to age in place, and broader social institutions, but current articulations of these systems do not explicitly challenge the ideologies and structural inequalities that disadvantage the elderly in the first place; if anything they retrench and reinforce the norms, hierarchies, and oppression of ageism by rendering them invisible (Kenner, 2008, p. 267).

Le tecnologie applicate alla sorveglianza ridefinirebbero, in altre parole, il concetto di libertà e di relazione di aiuto in base ai meccanismi di individuazione del rischio messi a punto all'interno del paradigma, senza mettere in discussione, ma anzi occultando le disuguaglianze strutturali e le ideologie che alimentano le politiche di controllo dell'invecchiamento demografico.

Anche in questo caso l'analisi critica non mira al rifiuto *tout court* dell'utilizzo dell'ICT, bensì a rendere creatori e applicatori dell'innovazione consapevoli della relatività della visione sottesa all'evoluzione del settore dell'*healthcare*, di cui lo sviluppo della tecnologia è parte integrante.

La questione etica della violazione dei diritti individuali laddove la scelta assistenziale sia determinata da altri (familiari, servizi sociali) sulla scorta di un livello di "rischio" attestato, come nel caso della demenza, si pone per quanto in modo meno evidente e certamente più complesso anche in relazione alla popolazione più estesa cui si rivolgono le tecnologie assistive. Come osserva la Kenner (2008, p. 256): "Those afflicted by dementia are not the only ones affected. The political economy of aging reflects shifting family structures, declining wages, and fewer voluntary and nonprofit resources".

Il mercato dei beni e dei servizi dedicati alla terza/quarta età si sviluppa nell'ambito di una evoluzione dei sistemi di produzione e riproduzione della vita che ha portato, da una parte, ad un progressivo allungamento dell'esistenza, con conseguente esponenziale incremento dei costi dello "Stato Sociale", e, dall'altra, ad una disgregazione delle reti di relazione e protezione rappresentate dalla famiglia e dalla comunità.

Life – come rimarcava anche lo studio dell'IRS precedentemente richiamato – has become an individual project as the 'natural safety net' disappears (family, neighbourhood, parish). A growing uncertainty directs more people to (new) professional care. Without a natural safety net, care becomes the responsibility of all people: glocalisation is the innovative term to describe the concept. *Socio-cultural change* has also an effect on raising demand for such services: pluralization and individualization trends, changes in gender roles and relations, increasing mobility requirements by changing labor markets and structural change in families e.g. demand for a greater density of care services for children and adolescents but also for older people (Crepaldi, De Rosa & Pesce, 2012, p. 33).

Nell'orizzonte del cambiamento socioculturale la vita tende a diventare un progetto individuale, da cui l'incremento della domanda di servizi, rivolti ai bambini così come agli anziani.

Volendo rileggere il tema della “Casa come primo luogo di cura”, abilitata come tale dall'utilizzo della tecnologia, in base alle schematizzazioni diffuse del contemporaneo paradigma *Smart City*, l’“invecchiamento intelligente” è facilmente riconducibile all'*asset* dello “smart living”. In questo contesto assumendo che, coerentemente con quanto previsto dal PNRR, l'assistenza domiciliare sarà sempre più permeata dall'impiego della tecnologia, il livello di diffusione e di *upgrade* di sistemi di *assisted living technologies* e *smart home* si candida a configurarsi come indicatore della qualità della vita dell'anziano. “Information and Communication Technology can turn older adults' homes into smart living environments, decreasing the need for support and improving life quality”, commentano Tassarolo et al. (2022, p. 2661), sottolineando ancora una volta l'importanza, in termini di accettazione, di quell'elemento di protagonismo dello *user* (*co-creation, co-design*) opportunatamente rilevato anche nell'ambito delle ricerche richiamate sopra e per altro perfettamente in linea con l'approccio *bottom-up* della *Smart City* nella sua versione evoluta. E aggiungono: “However, new technology has high uncertainty in terms of acceptance. Co-creation approaches are proposed to reach high acceptability, usability and satisfaction” (Tassarolo et al., 2022, p. 2661).

Il rischio in questo scenario sta verosimilmente in una sovrapposizione in chiave riduzionista fra *smart living* e benessere dell'anziano che stabilisca surrettiziamente una rapporto di proporzionalità diretta fra l'implementazione e lo sviluppo dello strumento abilitante (la tecnologia) e la qualità della vita dell'assistito, il cui livello ottimale sarebbe verosimilmente rappresentato da un anziano completamente autonomo e connesso nella sua abitazione, “messo al sicuro” da eventi critici non seguiti da pronto intervento mediante le tecnologie di monitoraggio attivo e assistito da remoto (teleriabilitazione, televisita, disponibilità di *social robot* per la stimolazione cognitiva). Come osservava Loe, tuttavia, un adeguato *mix* di tecnologie “quotidiane” non di avanguardia, dal telefono al *robot* da cucina, può divenire componente fondamentale di una strategia di “comfortable aging” in grado di elevare sensibilmente la qualità del vivere percepita. L'approccio tendente ad enfatizzare la dimensione tecnologica rischia di semplificare eccessivamente, questo il condivisibile monito di Loe, la complessità del fabbisogno, riducendone lo spessore umano e sociale. Una misurazione del grado di benessere ridotto alla sua componente “ict-enabled”, basata sui parametri dell'autonomia e della sicurezza dell'anziano, precluderebbe altresì la possibilità di immaginare percorsi alternativi. Utilizzando un approccio differente, in altre parole, si potrebbe allargare la prospettiva alla valutazione di ulteriori fattori; il concetto stesso di dignità della persona sotteso ai temi dell'indipendenza e dell'autonomia potrebbe includere significati che attingono maggiormente, ad esempio, alla cultura della “comunità” espressa dalla sociologia

classica. Si potrebbe utilizzare, ipoteticamente, quale indicatore dello stato di benessere dell'anziano, il numero di visite settimanalmente ricevute da un amico o da un parente, attribuendo a tale indicatore il medesimo coefficiente di “sicurezza percepita” attribuito all'introduzione di un *device* per il telemonitoraggio.

Il benessere misurato dal modello *smart living*, “embedded” alla componente tecnologica abilitante, tende inevitabilmente a riflettere un concetto di benessere, nonché concetti correlati come quelli di cura e di invecchiamento che sono propri di un determinato paradigma economico, sociale, politico e culturale.

Lo stesso si dica per gli indicatori utilizzati. In questo senso indicatori di “peace of mind” dell'anziano che focalizzassero l'attenzione, supponiamo, sul grado di partecipazione al *design* del servizio e degli strumenti tecnologici, sulla facilità di utilizzo e sul *training* fornito ai beneficiari intermedi e finali rischierebbero di restituire il grado di avanzamento lungo un *path* standardizzato riflettente una visione specifica della qualità della vita nella condizione anziana, rischiando di cancellare lo spazio culturale per la sperimentazione di diversi modelli e sinergie assistenziali.

Il paradigma emergente coniuga la “minaccia” dell'invecchiamento demografico con la grande opportunità economica per l' “industria per la terza età” producendo l'idea di un *Welfare* 2.0 e di una *Care* 3.0 presuntivamente migliorativi della condizione dell'anziano (preservazione, il più a lungo possibile, del proprio stile di vita, autonomia, sicurezza e serenità (“peace of mind”) e dei cosiddetti “badanti informali” (*informal carers*), alleggerendone il carico nonché lo *stress* fisico, psicologico ed emotivo. Ciò, per altro, con una attesa ricaduta positiva sull'occupazione femminile, soprattutto, stante l'elevato tasso di inattività delle donne gravate da impegni di cura (*European Commission*, 2015, p. 15). “Rapid demographic ageing is not only a major societal challenge (in terms of public budgets, workforce, competitiveness and quality of life) but also a major opportunity for new jobs and growth, also referred to as the Silver Economy” (*European Commission*, 2015, p. 6), leggiamo nel rapporto “Growing the Silver Economy in Europe”, richiesto dalla Commissione Juncker con l'obiettivo di mettere a fuoco le politiche di approccio al tema dell'invecchiamento della popolazione, identificato come una delle sfide chiave poste all'economia ed alla società europea.

Il Paper forniva una mappatura esaustiva di tutte le iniziative verticali e trasversali messe in campo dalle diverse DG della Comunità Europea, in accordo con gli *stakeholders*, per favorire lo sviluppo della *Silver Economy* ed il posizionamento degli Stati membri nel nuovo mercato globale, sottolineando l'importanza di un impegno congiunto teso a rimuovere le barriere e i limiti di interoperabilità entro i confini dell'Unione, suscettibili di frenare la crescita delle piccole e medie imprese operanti in questo mercato:

In the US several large companies are moving into Silver Economy markets, such as Google with the acquisition of NEST and iRobot as well as Apple with its new eHealth Kit and Wellness business. In Europe some of the larger companies such as Bosch, Legrand or Philips have developed Silver Economy strategies. There is a potential risk though, that SMEs operating in those markets could be facing growth barriers due to the relatively small scale of national markets in the EU, differing market conditions and a lack of open standards in different sectors (*European Commission*, 2015, p. 6).

Per favorire il conseguimento degli obiettivi di sviluppo economico prefigurati, l'Unione necessita, leggiamo ancora nel Documento, di politiche proattive che favoriscano gli investimenti in questo settore: “Achieving these objectives would require pro-active public policies designed to enable strategic investments and spending designed to foster active ageing, good health, social inclusion and independence”.

L'*age-friendly housing* veniva identificato, in questo contesto, come ambito strategico atto a garantire all'anziano indipendenza ed un elevato livello di qualità della vita, tagliando nel contempo la spesa pubblica e favorendo la creazione di opportunità occupazionali e di mercato, sia nel settore manifatturiero che dei servizi: “ICT is essential as a source of innovation, as it enables integrated person-centred care, with more focus on prevention, early detection and independent living, provided in the own work and home environment. It can empower people to manage their health and support their healthy lifestyle choices” (*European Commission*, 2015, p. 12).

Con riferimento specifico al mercato, si riportavano le stime del McKinsey Global Institute, che prefiguravano un impatto economico delle *health-care applications* abilitate dall'*Internet of Things* quantificabile in 1.1/2.5 trilioni di dollari l'anno entro il 2025.

In effetti il trilione sembra, quasi 10 anni dopo, un obiettivo conseguibile: “La dimensione del mercato della salute digitale valeva più di 233,5 miliardi di dollari nel 2022 – riferisce il Rapporto sulla “Quota del mercato della salute digitale” pubblicato da Global Market Insights nel marzo del 2023 - e si prevede di rappresentare un CAGR del 15% tra il 2023 e il 2032 guidato dalla crescente adozione di smartphone, tablet e altre piattaforme mobili” (GMI, 2023).

La proiezione al 2032 è stimata 981,5 miliardi di dollari, con il segmento del *telehealthcare* pari a 244 miliardi, in virtù della crescente inclinazione degli utenti verso approcci basati sulla tecnologia.

Un anno dopo, nell'aprile del 2024, un'altra valutazione di mercato fornita questa volta dall'indo-canadese *Precedence Research* stima il superamento del trilione di dollari entro il 2033, trainato dal progresso tecnologico nell'ambito dell'intelligenza artificiale, della robotica e degli strumenti per il monitoraggio da remoto (*Precedence Research*, 2024).

Coerentemente con quanto richiesto nel 2015 all'Europa in termini di politiche proattive, i PNRR dei Paesi membri individuano espressamente una Missione Salute sostenuta dal *Next Generation EU*, in cui per altro l'Italia guida gli investimenti seguita da Francia e Germania ed in cui una quota importante di risorse è stanziata proprio per la digitalizzazione dei servizi.

A fronte delle molteplici proiezioni che si sono succedute sull'evoluzione attesa del sistema torna allora verosimilmente ancora più attuale l'immagine inquietante della “smart home” provocatoriamente proposta nel 2014 da Rem Koolhaas, architetto, urbanista e saggista olandese, che prefigura ironicamente la necessità di dotare le abitazioni del futuro anche di una gabbia di Faraday, che consenta all'individuo di isolarsi dal raggio di azione della fitta trama di *digital devices* che tratterà e moltiplicherà su un numero indeterminato di *display* ogni sua “libera azione”:

The house is turning into an automated, responsive cell, replete with devices like automated windows that you can open but only at certain times of the day – commenta Koolhaas (2024) – floors embedded with sensors so that the change in a person's position from the vertical to the horizontal, for whatever reason, will be recorded; spaces which will not be warmed in their entirety, but instead will track their inhabitants with sensors and cloak them in heat shields. Soon a Faraday Cage will be a necessary component of any home – a safe room in which to retreat from digital sensing and pre-emption.

Al di là della provocazione è quanto mai opportuno, ci sembra, non dimenticare il monito degli autori che hanno provato a mettere in luce le potenziali criticità dell'evoluzione attesa e preservare costantemente uno spazio, all'interno del paradigma dello “smart aging”, per il dibattito e per le letture alternative, che mettono in discussione gli assunti di tipo culturale e assiologico sottesi alla definizione del concetto stesso di invecchiamento all'interno del paradigma dominante “smart oriented” inducendo una riflessione critica sull'idea e sulla visione di comunità che ci si propone di rilanciare.

Bibliografia

- Agger, A. & Lund, D.H. (2011). Borgerne og brugerne i samarbejdsdrevne innovationsprocesser. In E. Sorensen & J. Torfing (a cura di), *Samarbejdsdrevet innovation – i den offentlige sektor (Collaborative Innovation in the Public Sector)*. Jurist og Økonomiforbundets Forlag, København.
- Carli, A. (2024, 15 febbraio). Welfare incompiuto, allarme Censis: rischio di collasso sociale. *Il Sole 24 Ore*.

Carretero, S. (2015). Technology-enabled Services for Older People Living at Home Independently: Lessons for public long-term care authorities in the EU Member States. *JRC Science and Policy Reports*.

Carretero, S. & Kucsera, C. (2015). Report on case studies of the technology-based services for independent living for older people. *JRC Scientific and Policy Reports*.

Carretero, S., Kucsera, C., Misuraca, G. & Colombo, C. (2015). ICT-enabled Social Innovation in support to the Implementation to the Social Investment Package - IESI D1.1.1 - Mapping and Analysis of ICT-enabled Social Innovation initiatives promoting social investment in 'Active and Healthy Ageing and Long-term Care, covering the areas of independent living and integrated care. *JRC Technical reports*.

Crepaldi, C., De Rosa, E. & Pesce, F. (2012, maggio). Literature review on innovation in social services. *Europe Health, Education and Welfare (sectors of Services)*, IRS FP7, Report.

Department of Health. (2005). Building Telecare in England, London. In M. Mort, C. Roberts & B. Callén (2012, 25 ottobre). Ageing with telecare: care or coercion in austerity?. *Sociology of Health & Illness*.

European Commission (2015, febbraio). Growing the Silver Economy in Europe. *Background Paper*.

Failla, R. (2021, 3 agosto). Intelligenza artificiale, robot e domotica per l'assistenza agli anziani. *Aging Project – Uniupo*.

GMI (2023). Dimensione del mercato della salute digitale – Rapporto formato, 2032. *Global Market Insights*. <https://www.gminsights.com/it/industry-analysis/digital-health-market>

Hermans, K. & Vranken, R. (2010). *Switch the light on green for innovation in care. Practice Book*. Flemish federation of welfare, Brussels.

Il Sole 24 Ore (2024, 13 maggio). Silver Economy: Qualcosa di nuovo. *Il Sole 24 Ore*.

Kenner, A. (2008). *Securing the elderly body: dementia, surveillance, and the politics of 'aging in place'*. *Surveillance and Society*, 5, 3.

Koolhaas, R. (2014, 24 settembre). My thoughts on the smart city. *High-Level Group meeting on Smart Cities*, Brussels.

Lewin, D., Adshead, S., Glennon, B. & Williamson, B. (2010). Assisted living technologies for older and disabled people in 2030. A final report to Ofcom. *Plum Consulting*, London.

Leys, M. (2009, 10 settembre). Technology and innovation in the elderly care. *Care symposium*, Aalst.

Loe, M. (2010, febbraio). Doing it my way: old women, technology and wellbeing. *Sociology of health & illness*, 32, 2.

- Misuraca, G. (2016a, 20 maggio). Shaping the Welfare Society: Unleashing transformation through ICT-Enabled Social Innovation. *First European Conference – Digital Ecosystems for Social Services*, Roma.
- Misuraca, G. (2016b, giugno). The Future of Welfare Systems Exploring the role of ICT-enabled social innovation. *Publications Office of the European Union*.
- Mort, M., Roberts, C. & Callén, B. (2012, 25 ottobre). Ageing with telecare: care or coercion in austerity?. *Sociology of Health & Illness*.
- Piano Nazionale di Ripresa e Resilienza Missione 6: *Salute Componente 1 (M6C1): Reti di prossimità, strutture e telemedicina per l'assistenza sanitaria territoriale* - Investimento 1.2.1 Assistenza Domiciliare LINEE GUIDA ORGANIZZATIVE CONTENENTI IL MODELLO DIGITALE PER L'ATTUAZIONE DELL'ASSISTENZA DOMICILIARE (Milestone EU M6C1-4).
- Precedence Research. (2024, aprile). *Digital Health Market Size, Share, and Trends 2024 to 2034*. <https://www.precedenceresearch.com/digital-health-market>
- Suzuki, T., Murase, S., Tanaka, T. & Okazawa, T. (2006). New approach for the early detection of dementia by recording in-house activities. *Telemedicine and e-Health*, 13, 1.
- Tessarolo, F., Petsani, D., Conotter, V., Nollo, G., Conti, G., Nikolaidou, M., Onorati, G., Bamidis, P.D. & Konstantinidis, E.I (2022). Developing ambient assisted living technologies exploiting potential of user-centred co-creation and agile methodology: the CAPTAIN project experience. *JOURNAL OF AMBIENT INTELLIGENCE AND HUMANIZED COMPUTING*, 15, 5.
- Walker, K. (2011, 29 giugno). Britain's age timebomb. *Daily Mail*.

Intelligenza artificiale e coscienza in una prospettiva buddhista e spiritualista

di *Roberto Siconolfi**

ABSTRACT (ITA)

Il saggio affronta il tema della coscienza e dell'intelligenza artificiale (IA) in una prospettiva buddhista e spiritualista. Un'intelligenza artificiale può davvero divenire “forte”, per dirla alla John Seale? Può davvero l'insieme delle reti neurali, più o meno ben riprodotte, o simulate, produrre una coscienza come quella dell'essere umano? La coscienza umana è il semplice prodotto della sua attività neuronale, magari immersa nel tempo e nello spazio, in un ambiente storico-sociale, e dunque frutto di semplici qualità sostanzialmente materiali, che siano cerebrali, corporee o di relazione? O forse c'è dell'altro, la dimensione dalla quale una coscienza giunge all'essere umano è di altro tipo, “sottile”, e dunque le scienze materialiste e darwiniste non ci sono più utili per identificarla?

Parole chiave: intelligenza artificiale, coscienza, mente, stanza cinese, buddhista

Artificial Intelligence and Consciousness from a Buddhist and Spiritualist Perspective

by *Roberto Siconolfi*

ABSTRACT (ENG)

The essay addresses the theme of consciousness and artificial intelligence (AI) from a Buddhist and spiritualist perspective. Can artificial intelligence truly become “strong,” to use John Searle’s term? Can a network of neural circuits – more or less accurately reproduced or simulated – actually produce consciousness like that of a human being? Is human consciousness simply the product of neural activity, perhaps shaped by time and space, immersed in a socio-historical environment, and therefore the result of essentially material qualities, whether cerebral, bodily, or relational? Or is there something more – does consciousness originate from a different, more “subtle” dimension, making materialist and Darwinian sciences no longer sufficient to identify or understand it?

Keywords: artificial intelligence, consciousness, mind, Chinese room, Buddhist

* UniTreEdu

IA forte: il valore aggiunto della dottrina buddhista

Uno dei temi fondamentali dell'oggi, da un punto di vista tecnologico, quanto sociale e antropologico, è quello di comprendere se le intelligenze artificiali possano divenire intelligenti, o meglio “coscienti”, come un essere umano.

Dal nostro punto di vista, l'attuale dibattito in merito, che coinvolge scienziati, tecno-scienziati, neuroscienziati e filosofi, va integrato con prospettive di carattere spirituale, che in qualche modo vadano oltre il piano dei “sensi”, della materia, nella comprensione della realtà e nella coscienza della realtà.

A tal proposito, possiamo usufruire della dottrina buddhista, grande tradizione millenaria che dall'Oriente, negli ultimi decenni, se non nell'ultimo secolo, ha avuto sempre maggiore presa anche nel mondo occidentale.

Il buddhismo, oltre ad essere una religione e una filosofia, è anche una “scienza della mente” (Dalai Lama, 1993), con il suo carattere estremamente logico-razionale, e dotato di un metodo rigoroso che farebbe invidia alla nostra scienza, con in più la differenza/integrazione, non da poco, di penetrare la mente e la sua coscienza dall’“interno”, comprendendo i suoi contenuti proprio dall'interno, e non solo limitandosi ad analizzarli e studiarli dall’“esterno”, per così dire, ovvero “concettualmente”.

Per via di ciò, le più alte gerarchie della tradizione *Mahayana* di scuola *Madhyamaka*, a cominciare dal Dalai Lama, dialogano con la scienza occidentale, giunta ora su frontiere d'avanguardia, quanto, per certi versi, inaspettate.

Il dialogo si sofferma sulle neuroscienze e sulla fisica quantistica, due campi capaci di riuscire a penetrare la realtà, almeno concettualmente, come non mai nella storia scientifica dell'Occidente moderno.

Viviamo in un'epoca di grandi evoluzioni tecno-scientifiche, in particolare con l'affermarsi delle cosiddette intelligenze artificiali, sempre più promettenti, ma anche fonti di paure. Il buddhismo, a integrazione delle scienze e delle tecno-scienze occidentali, può fungere da ottima base per la comprensione degli sviluppi delle suddette, ma anche per favorire una loro applicazione virtuosa.

In quest'ottica dobbiamo tematizzare la fatidica quanto “fondamentale” teoria della mente del filosofo John Searle, il quale, a partire da un esperimento, la “Stanza cinese”, finalizzato alla comprensione della possibilità di un'IA di realizzare stati mentali – la cosiddetta “IA forte” –, deduce questa impossibilità, in quanto per produrre una mente umana è necessario un cervello umano – i fenomeni mentali umani derivano dalle effettive proprietà fisico-chimiche dei cervelli umani reali.

Una prospettiva definita “naturalismo biologico”, che però va integrata dalla dottrina della mente

di tipo buddhista, o di altra matrice “spiritualista” come vedremo, secondo la quale essa non emerge dal cervello, ma al contrario è lei a “destare” il cervello, e addirittura può essere indipendente da esso, come dimostra il passaggio del *continuum* mentale di vita in vita nel “ciclo delle rinascite” o, meglio, in quella che viene definita come la “preziosa rinascita umana”.

Questa “mente”, questo *continuum*, è di impossibile ricezione da parte di una IA, per quanto possa essere addestrata, e per quanto il suo sistema “intelligente” possa evolversi.

Chinese room, il dibattito attorno a un’IA “forte”

Il filosofo analitico John Searle, di impostazione neo-kantiana, è uno dei protagonisti del dibattito sulla possibilità dell’intelligenza artificiale di divenire “forte”, ossia di produrre stati mentali come un essere umano.

Tale rilevanza intellettuale dipende soprattutto da un suo esperimento, accompagnato da una deduzione teorica. Si tratta dell’esperimento della Stanza cinese (*Chinese room*), che fu presentato nel 1980 sulla rivista *The Behavioral and Brain Sciences*, con un articolo dal titolo *Minds, Brains and Programs*.

In sintesi: all’interno di una stanza vi è John Searle e un libro di istruzioni con insiemi di caratteri cinesi associati all’inglese (la lingua di Searle). Fuori, invece, vi è un uomo di madrelingua cinese. L’uomo fuori dalla stanza comunica con Searle, senza sapere nulla di lui, e in cinese. Searle, che a sua volta non conosce il cinese, risponde seguendo le istruzioni, che lui chiama il “programma”. Più l’esperimento si complica, e più Searle avrà istruzioni complesse per risolverlo, ma sempre senza alcuna capacità né obbligo di interpretare il cinese.

Tuttavia, l’uomo fuori dalla stanza sta sempre pensando di avere a che fare con un uomo, quello all’interno, che capisce il cinese.

Con la Stanza cinese si vuole dimostrare che la mente può avere un rapporto “sintattico” con i simboli, cioè una manipolazione corretta, senza però averlo su di un piano “semantico”, che comporta cioè il collegamento ad un significato.

Dunque, per ogni “input” vi può essere un “output”, ma senza comprenderne il perché.

In ultima analisi, l’IA può essere solo in grado di manipolare simboli senza comprenderli.

Per John Searle, l’IA può essere “debole” (*weak*), un sistema di simboli fisici che può agire in modo intelligente, ma non “forte” (*strong*), ovvero un sistema di simboli fisici che può produrre stati mentali, e dunque essere veramente intelligente, pensante, o meglio “cosciente”.

La sua visione viene definita “naturalismo biologico”, ovvero l’idea che la coscienza sia causata da proprietà fisiche del cervello, e che la macchina non ne disponga.

Per Searle “i cervelli causano le menti” (1980, p. 11), “i fenomeni mentali umani effettivi [sono]

dipendenti dalle effettive proprietà fisico-chimiche dei veri cervelli umani” (1990, p. 29).

In pratica, esistono dei correlati neurali della coscienza, un’ipotesi che spazza via la possibilità da parte delle macchine di averne una, ammenoché non si riesca, appunto, a riprodurre quelle stesse proprietà.

La Stanza cinese di Searle è una risposta al famoso “Test di Turing”, che stabilisce che una macchina può essere intelligente, anche se, sempre per il matematico Alan Turing, non vi è la necessità di detenere una coscienza per definirsi tale (1950, pp. 433-460).

Sulla stessa lunghezza d’onda di Turing si colloca il professore John McCarthy, informatico e vincitore del premio Turing, nonché inventore del termine Intelligenza Artificiale, il quale ha definito l’intelligenza come “la parte computazionale della capacità di raggiungere obiettivi nel mondo” (2007, p. 2).

Una definizione che secondo Russell e Norvig non permetterebbe di distinguere tra “cose che pensano” e “cose che non pensano”, per cui anche un termostato avrebbe un’intelligenza rudimentale (2003, pp. 48-52).

Russell e Norvig differenziano l’umanità del comportamento dalla sua intelligenza, ovvero il Test riesce a misurare solo se la macchina si comporta come l’umano, ma non l’intelligenza del comportamento, la sua generazione, e la consapevolezza relativa.

Allo stesso tempo, però, essi ritengono che lo sviluppo delle IA possa essere comunque efficace, senza la necessità che imitino per forza gli esseri viventi e la loro intelligenza.

Secondo loro: “I testi di ingegneria aeronautica non definiscono il loro obiettivo come quello di creare macchine che volino esattamente come i piccioni, tanto da poter ingannare i piccioni stessi” (2003, p. 3).

In pratica, le macchine vengono progettate con principi diversi, ognuno adeguato ed efficace per i propri scopi.

E anche Newell e Simon ipotizzano che un sistema di simboli fisici (come un computer digitale) avesse tutti i dispositivi necessari per “un’azione intelligente generale”, o, come è noto oggi, per un’intelligenza artificiale generale (AGI) (1976, pp. 113-126).

L’argomento della Stanza cinese non confuta questo, perché nei termini di “azione intelligente”, viene inquadrato il comportamento esterno della macchina, non la presenza o l’assenza di comprensione, coscienza e mente.

Ma la questione fondamentale è che Searle, e altri ricercatori, come Igor Aleksander, Stan Franklin, Ron Sun e Pentti Haikonen, credono invece che la coscienza sia imprescindibile per l’intelligenza. E che dunque: “Quello che volevamo sapere è cosa distingue la mente dai termostati e dai fegati” (1980, p. 420).

Chinese room: criticità e superamenti

La prima criticità all'ipotesi di Searle arriva da Nils Nilsson: mentre l'argomentazione della Stanza cinese è tarata su un modello vecchio, sulla manipolazione dei simboli, l'IA moderna, del XXI secolo, con i programmi di *deep learning* che ne stanno alla base, non utilizza simboli ma segnali dinamici. Il segnale campionato, possederebbe, dunque, già la semantica (2007).

Questi sistemi, più che simbolici, sono rappresentazioni digitalizzate di sistemi dinamici, e i singoli numeri non hanno una specifica semantica, ma sono campioni o punti dati di un segnale approssimato, il quale invece la possiede.

Tuttavia, anche questa argomentazione non determinerebbe l'effettiva comprensione da parte della macchina di ciò di cui sta parlando, ma semplicemente cambierebbe la modalità con la quale essa emette le sue elaborazioni. E non comunque nella prospettiva di Searle, prospettiva che necessita del correlato fisico-chimico cerebrale.

Ma un'ulteriore ipotesi, per giungere comunque a una coscienza, potrebbe essere la possibilità che un programma per computer possa replicare la capacità dei neuroni di generare stati mentali (in software) o la replica in senso fisico delle stesse reti neurali (in hardware).

Ma Hubert Dreyfus si dimostra scettico all'idea secondo la quale: "Per produrre intelligenza, sembrerebbe sufficiente riprodurre il comportamento del sistema nervoso con qualche dispositivo fisico" (1972, p. 106). E per motivi di necessaria "incarnazione", di "immersione nel mondo" di una coscienza.

Ma ancor più profonde – e radicali – risposte a questi dubbi, possiamo ricercarle non nel naturalismo biologico di Searle, un modello essenzialmente materialista, ma nel buddhismo, in particolare della tradizione *Mahayana*, scuola *Madhyamaka*, e nella sua dottrina della mente, secondo la quale, a detta del Dalai Lama, la coscienza non emerge dal cervello, ma al contrario, esso si desta a partire da una coscienza sottile, da una "mente innata" (2009, p. 328).

Un meccanismo dal più al meno (2023, p. 9), "dall'alto verso il basso" (2023, p. 287), come sostiene Federico Faggin, una vita nel mondo delle tecnologie come inventore, e che ora giunge a *Weltanschauung* "spirituali", in grado di fargli rileggere in una chiave diversa il suddetto mondo.

Un meccanismo "contrario" rispetto a quello ipotizzato dai teorici dell'Intelligenza Artificiale Generale (AGI) come Yoshua Bengio e Antonio Chella, o dai transumanisti Kurzweil, Vinge, Moravec, ecc., e dai postumanisti alla Marchesini, per i quali essa può avere capacità di autoapprendimento, grazie alla programmazione e al gioco di interazioni e *feedback* con l'ambiente esterno, fino a sviluppare una propria autocoscienza e ad evolversi come gli esseri umani.

Una teoria dal meccanismo *bottom-up*, quest'ultima, che sarebbe automaticamente smentita, in quanto la coscienza, per farla breve, viene da "sopra", non da sotto, non si evolve da forme

minime verso quelle “macro”, ma al contrario, si aggancia da forme (platoniche), “potenze” sottili macro-cosmiche, a forme sempre più minime e incarnate, compreso il corpo e il cervello umano (*top-down*).

In questa prospettiva, la macchina sarebbe impossibilitata a ricevere questa coscienza verticale. Ma c'è di più, secondo il Dalai Lama: “Non c'è motivo di credere che lo stato più sottile – chiamiamolo ‘mente innata’ [...] abbia dei correlati neuronali perché non è fisico, non è subordinato al cervello. Ma per altri processi umani che si manifestano nel corso della vita umana è certamente possibile” (2009, p. 328).

Ciò si manifesta nel passaggio del “*continuum* mentale” da una vita a un'altra. Ed è proprio questo *continuum* che è irricevibile e irriproducibile da una macchina, la quale è semplicemente frutto della creazione umana ed esclusa dal ciclo delle rinascite.

Il *continuum* mentale è ricevibile da parte dell'uomo alla nascita, nascita per la quale lo sperma e l'ovulo non sono causa primaria ma “cooperante”, di conseguenza, anche il cervello, l'organo fisico, è semplicemente cooperante nella ricezione del *continuum* (Lama Zopa Rinpoche, 2016, pp. 1-3). Tutto questo lavoro per una macchina è letteralmente impossibile!

A questo punto, ultimo problema, potrebbe essere la verifica dell'esistenza della coscienza. Secondo Federico Faggin, essa è sperimentabile solo in prima persona, e qui definiamo il concetto di *qualia*, ovvero “ciò che proviamo’ quando una sensazione o un sentimento emerge nella nostra coscienza” (2023, p. 147).

Per soddisfare il metodo scientifico occidentale classico, con la sua famosa osservazione “esterna”, da laboratorio, è oggi addirittura possibile riscontrare con il *neuroimaging* i correlati neuronali dell'attività mentale o addirittura spirituale (Newberg, 2014).

Secondo le ricerche del dott. Dietrich Lehmann dell'Università di Zurigo in collaborazione con il Lama buddhista Ole Nydahl, si è riscontrato che nelle “visualizzazioni audio-visive” (tecniche di immaginazione, per così dire, che si effettuano all'interno delle pratiche meditative), vengono stimulate le stesse aree del cervello addette alla vista e all'ascolto nella realtà sensibile. In pratica, anche in questo caso, la coscienza, la mente, non emerge dal cervello, ma si muove in maniera indipendente da esso.

E ancora, secondo rilevazioni negli “stati di premorte”, *Near Death Experience* (NDE), effettuate da vari ricercatori tra i quali il dr. Sam Parnia, si è giunti alla conclusione che il cervello è solo un tramite della coscienza (Parnia et al., 2001), in quanto al risveglio da quello stato, e nonostante quello stato, i pazienti sapevano dire quanto fosse accaduto nella stanza nella quale erano ricoverati, e tanto altro ancora.

Anche dall'uso della Realtà Virtuale (RV) nelle riabilitazioni motorie (Pournajaf et al., 2021), è possibile effettuare delle deduzioni rilevanti in questa direzione. È infatti interessante notare

come le terapie per il recupero motorio di un arto, ad esempio, simulino il movimento a livello mentale, immergendo il soggetto in una realtà nella quale egli riesce a muovere correttamente tale arto. Alla base vi è l'attivazione di quell'area del cervello preposta al movimento, attivazione che avviene solo virtualmente. Emerge dunque come sia la mente, indipendentemente dal cervello, a possedere quel gruppo di funzioni che porta all'effettivo movimento corporeo. È la mente che ha il comando, è la mente che decide cosa fare o non fare, è la mente che, con i dovuti tempi tecnici, è in grado di far riprendere il corpo fisico da un trauma, essendo il corpo fisico, e il cervello, sostanzialmente plasmabili dalla forza "sottile" della mente.

Ma anche nel *Simposio Internazionale "Prospettive buddhiste su coscienza, evoluzione e Intelligenza Artificiale"*, organizzato dal *Centro Studi di Unione Buddhista Italiana*, si è concluso che "la coscienza non può essere intesa come una mera 'funzione', e dunque non può essere attribuita all'IA unicamente sulla base della sua capacità di simulare risposte e comportamenti umani" (2023).

Tutto questo pur nella consapevolezza che l'intelligenza artificiale possa generare elevatissime, capacità computazionali e di simulazione del comportamento umano, e superiori all'umano stesso in termini computazionali – per le quali, come dicevamo con Turing, la "coscienza" non è necessaria.

Teorie della coscienza

A un livello più profondo, filosofico, metafisico, possiamo sinteticamente notare, riferendoci all'evoluzione del dibattito sull'IA "forte" a partire dall'esperimento della Stanza cinese di John Searle, che a emergere sono principalmente due posizioni, due matrici fondamentali diverse, ma che assumono varie forme, gradazioni ed applicazioni.

La prima, essenzialmente materialista, evoluzionista, per la quale la coscienza emerge dalle proprietà biologiche, chimiche, del cervello. Dalla carne, dal corpo, dalla materia, e, secondo alcuni, in relazione con l'ambiente (storico e sociale) di riferimento, e lo fa a tal punto da non poter addirittura esistere, se non necessaria a tale evoluzione, a tali parti, pezzi di carne, di corpo, o meglio cervello, di materia – potremmo dire genericamente "dal basso verso l'alto", *bottom-up*.

Dall'altra vi è la posizione opposta, secondo la quale è la coscienza che si aggancia, si incarna, si manifesta in un cervello, a partire da uno stato, una forma, un corpo sottile. Essa si incarna, appunto, come abbiamo visto, secondo la dottrina buddhista, al cervello, all'organo fisico.

Vie di mezzo sono coloro che invece, o tengono in piedi il dualismo mente/corpo, per certi versi spiazzati dalla questione, non riuscendo a dare una definizione di coscienza – un *hard problem* irrisolvibile, ad esempio, per David Chalmers – oppure lasciano aperta la possibilità di un'eventuale implementazione di una mente su un substrato che riproduca le proprietà causali del

cervello umano (John Searle).

Prospettive, come vediamo, che rientrano in dinamiche dualistiche, che scindono in qualche modo l'aspetto sottile della mente, o della coscienza, dal supporto biologico e neuronale ben preciso del cervello umano e/o che non gli danno la giusta direzione.

In questo scritto correggiamo tale dualismo, dal nostro punto di vista imperfetto, ma non alla maniera delle neuroscienze materialiste e darwiniste, o di un certo “monismo” dello stesso taglio, ma grazie ad una visione dell'uomo animata da dottrine di carattere “metafisico”, corroborate da neuroscienze e filosofia.

Teorie della coscienza: l'impostazione materialista e darwinista

Dal lato dei materialisti, delle neuroscienze materialiste ed evoluzioniste, troviamo Daniel Dennett.

La posizione di Dennett è critica verso i risultati della Stanza cinese. È Searle stesso ad attribuirle la definizione di “funzionalismo informatico” (1992, p. 44; 2004, p. 45).

La teoria della mente di Dennett sostiene che i fenomeni mentali (credenze, desideri e percezioni) possono essere descritti attraverso le loro funzioni, in relazione tra loro e con il mondo esterno. Dunque, poiché un programma per computer può rappresentare tali relazioni funzionali come relazioni tra simboli, allora pure il computer può avere fenomeni mentali.

Per Dennett, nella sua teoria della coscienza, il “Multiple Draft Model” (MDM, modello delle “versioni multiple”), non esiste un “centro unificatore” dei processi coscienti, un unico punto di convoglio informativo localizzabile nel cervello, come ipotizzabile dalle teorie cartesiane.

Vi è invece una moltitudine di processi cognitivi che si attuano in parallelo (1993, p. 141), il che implica grosse conseguenze nei confronti dell'esistenza di un Io, e dell'interiorità nella quale sarebbe situata la soggettività. Ognuno di questi processi e sottoprocessi neurali è già per Dennett un evento mentale, senza bisogno di centri o coscienze unificate (1993, p. 141).

Anzi, secondo l'autore bisognerebbe chiedersi proprio se la coscienza esista (almeno per come viene comunemente intesa). La natura, infatti, favorirebbe proprio la possibilità di esseri dal design semplice, e visto che la cosa è possibile con le IA, allora è molto più probabile che anche gli esseri umani non detengano certe caratteristiche pur essendo convinti di essere coscienti (Cole, 2004, p. 22; Crevier, 1993, p. 271; Harnad, 2005, p. 4). In altre parole, dall'inesistenza della coscienza nell'IA si deduce l'inesistenza della coscienza nell'uomo, e paradossalmente la coscienza non esiste ma l'IA forte esiste.

Su questa base materialistico-eliminativa della coscienza, per Dennett la questione dello zombie

filosofico non sussiste,¹ perché non ha senso porsi la domanda se un essere identico a noi, che si comporta come noi, sia cosciente o meno, anche perché la coscienza non è detenuta da un centro unificato, ma da funzioni e processi cognitivi e comportamentali dislocati e complessi – dunque, non è un'entità separata, ma annessa a tali processi detenuti anche dallo zombie in esame (Dennett, 1995, pp. 322-326).

In maniera simile a Dennett, Richard Dawkins, il quale, pur non avendo mai elaborato una teoria della mente, la definisce come il frutto dell'evoluzione, che produce tutt'al più degli enti simili ai “meme” (idee, credenze, comportamenti), che si diffondono nelle menti umane in maniera analoga ai geni e alla loro proliferazione biologica (1992, pp. 63-85, 222-235).

Un piano meramente ridotto ai meccanismi fisici e biologici (fondamentalmente basato su genetica e memetica), in alternativa a ogni dualismo di stampo cartesiano.

Sulla stessa scia troviamo i filosofi postumanisti come Marchesini (2001), Haraway (1995), Hayles, i quali definiscono la non esistenza di un'essenza fissa o di un Sé, e la relativa natura meramente neuronale e informazionale della mente, e della coscienza, vista tutt'al più come epifenomeno (Hayles, 1999, p. 3). Questo sempre nella prospettiva del superamento del dualismo cartesiano mente/corpo, e per estensione natura/cultura.

Queste teorie post-umaniste, si incrociano con le filosofie post-strutturaliste, in quella dimensione tipica di “sfaldamento”, “smaterializzazione della realtà” (Siconolfi, 2021, pp. 94-105), di apertura alle forze del subcosciente tipica del mondo postmoderno (Dugin, 2019, p. 174).

In termini complementari, come introduzione filosofica a questo complesso e multilivellare dibattito, possiamo menzionare:

- la non esistenza di un centro unificatore della coscienza, ma di una sua distribuzione sui fenomeni della realtà (effetto e non causa). Un processo impersonale, “desoggettivato” e oltre il soggetto, effetto di connessioni multiple tra corpi, affetti, desideri, energie e forze sociali – è questo il caso delle teorie di un Deleuze, in collaborazione con Guattari (1972; 1980), ma anche di un Lyotard (1978; 1981);
- la dissoluzione del soggetto nel processo di scrittura, la coscienza come assenza, dissoluzione dell'Io, in Blanchot (1955). Oppure la scarsa importanza del soggetto nella produzione di senso, che risulterebbe essere piuttosto l'effetto di un'azione decentrata che si svolge nel campo della linguistica – la “morte dell'autore” in Roland Barthes (1968), che a detta di Cosimo Accoto dà

¹ Lo zombie filosofico è un esperimento mentale nel quale si immagina un essere identico a noi ma privo di soggettività, di capacità di provare stati interiori (*qualia*). Questa argomentazione è usata proprio per sottolineare la diversità tra esseri che possiedono coscienza da quelli che non la possiedono, i quali sarebbero pure capaci di reagire a stimoli esterni ma senza provare stati interni (dolore, ecc.). La coscienza, insomma, per dirla con David J. Chalmers, va oltre mere spiegazioni di carattere fisicalistico, alcune delle quali indagate anche in questa pubblicazione (materialismo, comportamentismo, funzionalismo, ecc.). Si vedano Chalmers, 1996 e Nagel, 2013.

forza alla possibilità di un linguaggio macchinico (LLM), che equipari l'ente macchina all'uomo (2023);

- l'impossibilità di afferrabilità della coscienza. Un processo sempre rimandato, “differito”, così come indefinibile è l'Essere, per dirla con Heidegger, che conserva in sé una “differenza”, per buona parte inesplicabile attraverso il linguaggio e che costituisce lo spazio costitutivo di ogni esistenza – teoria della “Différance” di Derrida (1982);
- il Sé, la soggettività, come costruito socio-culturale, un dispositivo di controllo ad opera del potere biopolitico, oppure, come nel caso de l'“autore”, una delle funzioni possibili della specificazione-soggetto (Foucault, 1975, 1969).

Il computazionalismo, invece, è quella teoria che sostiene che la mente può essere accuratamente descritta come un sistema di elaborazione di informazioni. Tali processi computazionali (*software*) sono indipendenti dal supporto fisico sui quali vengono implementati (*hardware*), e di conseguenza essi potrebbero venire implementati non per forza su un cervello biologico, ma anche su un computer (ipotesi IA forte).

Ma per Steven Horst, pur detenendo delle importanti qualità metodologiche, il computazionalismo è insufficiente, e sempre per le motivazioni espresse dagli altri critici della IA forte: la questione della manipolazione semantica dei simboli, la coscienza di essi (Horst, 2005).

Sulla stessa falsariga di Horst e dei critici dell'IA forte alla Searle, Stevan Harnad, per il quale la manipolazione dei simboli da parte di un sistema computazionale è insufficiente per la comprensione degli stessi.

Per tale comprensione, invece, c'è bisogno di un ancoraggio del sistema computazionale nel mondo concreto dell'esperienza corporea e sensoriale (teoria del *symbol grounding*, 1990), e quindi la modalità d'implementazione (compresi i processi fisici e biologici) è fondamentale per la definizione della mente (i correlati fisico-chimici di Searle).

Il test di Turing è un sistema limitato per valutare l'effettiva presenza stati mentali, per forza di cose “coscienti” (Harnad, 2002).

Concezione della mente ancor più estrema del computazionalismo è il “comportamentismo” (Watson, 1924; Skinner, 1958), per il suo centrarsi sul semplice comportamento esterno di una mente, in un meccanismo di input e risposte, e bypassare completamente la questione degli stati interni, visti nemmeno come processi computazionali.

Per Searle, in *Intentionality*, per i comportamentisti: “Avere una mano significa semplicemente essere predisposti a certi tipi di comportamento come afferrare”, ignorando il perché si muovano in questo modo e gli stati interni che portano a quel movimento (1983, p. 263).

A un diverso livello invece il “connessionismo”, che va oltre il computazionalismo, in quanto in questo caso i processi mentali sono modellati su reti di unità connesse o reti neurali artificiali. Le

informazioni sono elaborate in modo distribuito e parallelo, e non simbolico (computazionalismo). Il connessionismo va oltre anche al comportamentismo, in quanto riconosce l'importanza dell'esperienza (ancoraggio a percezione diretta, *grounding*), ma ammette anche la possibilità di generare stati interni, anche se non simbolici.

Riguardo la Stanza cinese, i connessionisti credono che un tale tipo di architettura “parallela” sia in grado di comprendere (Cole, 2004, pp. 16-17; Hauser, 2006, p. 7), in quanto Searle, nella Stanza, è solo uno degli agenti che svolge la sua piccola parte di elaborazione.

La mente per i connessionisti non è “manipolazione simbolica isolata”, ma emerge da processi distribuiti, interazioni dinamiche e di apprendimento dall'ambiente. I simboli non sono indipendenti dal contesto o dal *grounding*, e se la macchina avesse accesso al contesto e al mondo reale genererebbe automaticamente una comprensione.

Se da un lato il connessionismo è il modello di maggior successo nelle applicazioni pratiche di IA, soprattutto per le reti neurali profonde (*deep learning*), sia nella generazione di testo, che nel riconoscimento di immagini, che nella robotica (interazione con il mondo reale), dall'altro, però, non risolve la questione della coscienza di tali sistemi.

Più in generale possiamo annoverare all'interno delle teorie di stampo materialista e darwinista della mente, quella di Gerald Maurice Edelman, oppositore del funzionalismo, ma sostenitore inizialmente di una forma di riduzionismo biologico, di “darwinismo neurale”, una teoria evolutiva della mente. In *Sulla materia della mente*, egli sostiene in merito all’“evoluzione neurale” che “l'intero processo si basa sulla selezione e coinvolge popolazioni di neuroni impegnate in una competizione topobiologica” (1992, p. 134).

Edelman, tuttavia, spostandosi su posizioni anti-riduzioniste, rettifica una posizione esclusivamente basata sul “cervello”, inteso in senso statico, geneticamente e permanentemente determinato, e supera il dualismo natura/cultura (per estensione anche mente/corpo), tipico delle neuroscienze cognitive.²

Per Edelman (2007, p. 21), “il cervello è incarnato e il corpo è inserito nell'ambiente”: le attività del cervello sono collegate a quelle provenienti dal corpo, e viceversa. Quest'ultimo, però, è inserito in un ambiente (eonicchia), nel quale si relaziona con lo stesso meccanismo di influenze e *feedback*. Dunque, il processo è meramente darwiniano: il cervello è “un sistema selettivo che opera nell'arco della vita” (2007, p. 24), la cui organizzazione strutturale di base è predeterminata dai nostri geni, e prevede, alla nascita, un'innumerabile serie di schemi potenziali inseriti in una

² È possibile suddividere le filosofie della mente su base riduzionistica e non-riduzionistica. Schierati sul primo fronte troviamo i neurologisti, i comportamentisti, i funzionalisti e i computazionalisti. Sul secondo, mentalisti, dualisti, evolucionisti sostenitori della mente estesa (in senso meramente orizzontale come vedremo), e dottrine di carattere filosofico-fenomenologico e spirituale. In una linea mediana potrebbe essere considerato l'emergentismo. Tutte teorie grossomodo affrontate in questo paper, seppur schematizzate in senso diverso e forse più completo (alcune tralasciate perché non inerenti allo studio sull'IA forte).

s sofisticata organizzazione. L'esito finale di questa evoluzione, tuttavia, è molto determinato dall'ambiente specifico in cui si "incarna" quel particolare sistema nervoso. Per Edelman è, appunto, di "processo" che si tratta, e non della semplice localizzazione in specifiche aree cerebrali, secondo quella che possiamo definire come una lettura "materialistico-statica". La struttura plastica cerebrale, infatti, malleabile, e non fissa, è data proprio da questa influenza "attiva" dell'ambiente sui cambiamenti epigenetici e sulla formazione delle mappe concettuali.

Inoltre, per Edelman, è possibile distinguere fra coscienza primaria e coscienza di secondo ordine. La prima è anche propria degli animali. Ma vi è anche una coscienza di secondo ordine, fondata su capacità di astrazione e capacità simboliche.

In altre parole, lo sviluppo della coscienza di ordine superiore, deve un'accelerazione all'acquisizione del linguaggio e della sintassi, e la coscienza ha basi biologiche fondate su un sistema selettivo, il quale si estende a tutta la complessità e alla contingenza storica dell'esperienza fenomenica.

All'interno di questa "esperienza" annoveriamo anche la creazione artistica e certi prodotti della nostra vita mentale, benché non si giunga ad una loro "descrizione scientifica" (2007, p. 139), data la molteplicità di combinazioni tra coscienza di ordine superiore e linguaggio, e i relativi collegamenti di pensiero, emozione, memoria ed esperienza, e visto che il numero delle possibili combinazioni di discriminazione cresce senza limiti.

Per Edelman, il dualismo tra scienza e discipline umanistiche, e quindi diremmo tra natura e cultura, è di per sé superato, proprio perché in questa "seconda natura" l'evoluzione umana va di pari passo: "Co-evoluzione della cultura che fornisce un mezzo di cambiamento relativamente rapido e potente che influenza le basi della conoscenza, della sensazione e del comportamento" (2007, p. 63).

Pure sulla questione "qualia", Edelman effettua una lettura sempre di stampo darwinistico, sulla base del meccanismo di selezione-evoluzione "neurale", che produce stati integrativi e complessi, ma sempre riconducibili alla struttura e alla dinamica del cervello.

Un po' come Dennett, il quale pensa che i qualia, premessa la loro possibile esistenza, siano diversi da come concepiti da certe teorie e credenze sulla coscienza, ma rappresentino la semplice elaborazione di esperienze fisiche ed oggettive e non riducibili a null'altro. Con possibilità limitata di descrizione, ma non indescrivibili (Dennett, 1988, pp. 44-77).

Il doppio livello della coscienza di Edelman richiama quello di un altro grande neuroscienziato e teorico della coscienza, Antonio Damasio, per il quale esistono, appunto, una coscienza nucleare e una estesa. La prima "è un fenomeno semplice, biologico, che ha un unico livello di organizzazione, è stabile in tutto l'arco di vita dell'organismo, non è una caratteristica esclusiva degli esseri umani, e non dipende dalla memoria convenzionale, dalla memoria operativa, dal

ragionamento e dal linguaggio” (Damasio, 2000, pp. 30-31). La seconda “è un fenomeno biologico complesso, con vari livelli di organizzazione, che si evolve nel corso della vita dell’organismo [...] presente anche in altri animali, a livelli semplici, ma che comunque raggiunge i suoi limiti superiori solo negli esseri umani. Dipende dalla memoria convenzionale e dalla memoria operativa [...] negli esseri umani è anche arricchita dal linguaggio” (Damasio, 2000, p. 31)

Anche Damasio ritiene l’esperienza corporea fondamentale per il Sé, la coscienza come fenomeno biologico, che riguarda sia il campo nucleare che esteso, ma con la differenza che mentre nel primo caso riguarda il “qui e ora”, è transitoria, si verifica quando l’organismo effettua un’esperienza e si chiude con la fine di essa, la seconda va oltre, conservando l’esperienza, “archiviandola”, insieme alle altre per così dire, in una funzione definita “Sé autobiografico” (Damasio, 2000, p. 32).

Le decisioni sono inconsciamente guidate proprio da “marcatori somatici”, e cioè “segni” delle esperienze emotive passate (1995, p. 245). In effetti, Damasio batte più sulla riunificazione tra pensiero ed emozioni, superando il dualismo cartesiano in merito, con un collegamento tra il materiale emotivo e il pensiero, proprio nell’elaborazione del pensiero (e viceversa).³

Ma approdiamo alla categoria di “uomo neuronale”, sulla quale una vasta letteratura di lavori scientifici e saggi sono stati pubblicati, e di cui il neurobiologo Jean-Pierre Changeux è stato uno dei massimi teorici promuovendo una vera e propria “biologia dello spirito”.

Una teoria in grado di superare le cosiddette scissioni tra anima e corpo e tra materia e pensiero, sociale e cerebrale, scienze naturali e scienze umane.

Lo spirito, in questo senso, sarebbe un vero e proprio computer naturale che potrebbe sperimentare se stesso attraverso una “macchina isomorfa”.

In *La vie des formes et les formes de la vie* (2012) Changeux ammette la possibilità di un’ipotesi metafisica nella generazione delle forme, oltre a quella di tipo “darwiniano”. Secondo quest’ultima

Il fenomeno si trova nello spazio e nel tempo all’interno di una serie temporale di cause ed effetti. Si occupa di una stratificazione a livello organizzativo e sulla transizione da un livello di organizzazione a un altro, qualunque sia il livello. Il presupposto è che il passaggio da un livello all’altro richiede due elementi fondamentali: un generatore di diversità e un sistema di selezione (Changeux, 2012, pp. 11-12).

Invece, quella di tipo metafisico

³ Sull’importanza delle emozioni in ambito cerebrale citiamo anche le ricerche e le opere di Joseph E. LeDoux. Anche se a differenza di Damasio, per LeDoux il collegamento con le emozioni avviene sulla base di percorsi neurali, non in virtù dell’esperienza corporea.

fa appello alla teologia naturale di Platone a John Paley o all'Abbe Pluche: le forme sarebbero astrazioni non materiali, non fisiche, extra-mentali, essenze, create da un "demiurgo" o "grande orologiaio". Si troverebbero, secondo Platone, *hyperouranios topos*, vale a dire oltre il cielo... Queste idee riaffiorano oggi giorno tra i seguaci del "design intelligente" (Changeux, 2012, pp. 11-12).

Pur tuttavia, sempre per Changeux, neuroscienziato di impronta materialista e darwinista, tali astrazioni non materiali, "forme" platoniche, sarebbero comunque frutto della selezione naturale (neurale), in un processo comunque evolutivo, a partire dai neuroni, e tali "enti" sarebbero solo apparentemente innati, ma comunque generati *bottom/up*, in un'ipotesi definita da molti come "riduzionista".

Sempre in quello che possiamo definire come il fronte materialista, seppur diversificato, troviamo Merleau-Ponty. Merleau-Ponty supera anch'egli il materialismo "statico", proprio come Edelman, ma lo fa sul versante del corpo, e della sua esperienza concreta nel mondo. In *Fenomenologia della percezione*, il *corps propre*, che richiama il *Leib* husserliano, è corpo vissuto, incarnato, unità gestaltica, "esperienziale", che va oltre la mera attività sensoriale, ma in rapporto dinamico con l'ambiente, oltre dualismi "mente/corpo", ma anche soggetto/mondo (Merleau-Ponty, 2003, pp. 151-156).

La sua è un'ontologia della carne, come definita ne *Il visibile e l'invisibile*, da intendere non come mente, materia, sostanza, ecc. ma alla greca *arché*, principio "incarnato", prototipo dell'essere, a metà tra gli enti individuali del mondo e l'idea (Merleau-Ponty, 1969, pp. 153-154).

Sulla base della filosofia di Merleau-Ponty, Hubert Dreyfus critica l'IA forte proprio sul punto dell'impossibilità dell'incarnazione di una tale intelligenza. L'intelligenza per Dreyfus è per forza di cose "incarnata", "situata", basata sull'esperienza diretta nel mondo, radicata nell'esperienza corporea pre-riflessiva, intraducibile attraverso l'uso di algoritmi.

La "scienza cognitiva incarnata" di Dreyfus, Merleau-Ponty e che passa per Varela e Maturana, Thompson e Rosch, supera le scienze cognitive classiche basate sulla mente e le sue rappresentazioni, e sul cervello all'origine dei processi cognitivi (Fugali, 2016, pp. 203-204). Si giunge così alle teorie *embodied*, basate sulla preminenza dei processi sensori-motori, i vincoli del corpo, congiunto ai fattori ambientali (Fugali, 2016, pp. 204-207).

La prospettiva di Merleau-Ponty, di un soggetto incarnato e in relazione dinamica col mondo, è simile a quella di Husserl, sebbene quest'ultimo dia un peso maggiore alla coscienza – quella "trascendentale" –, rispetto al corpo, "corpo vivo" (*Leib*), che comunque Husserl differisce dal corpo inteso in senso materiale (*Körper*) (Husserl, 2002, pp. 146-149). Per Husserl vi è sempre un'unità, un tutt'uno corporeo, ma a partire da questa coscienza, che trova appoggio nel corpo, che è trascendentale, ma sempre immanente, però, e alle cui essenze si giunge attraverso un

processo di riduzione fenomenologica, con la “sospensione del giudizio” (*epoché*). La “coscienza trascendentale”, dunque, non è situata in una zona del sovrasensibile, per così dire, ma è la condizione trascendentale dell’esperienza nel mondo, il campo costitutivo dell’intenzionalità della coscienza. Una concezione unitaria, questa, che va al di là di dualismi (da Kant a Cartesio), e che ha similitudini con la “teoria della persona” di Max Scheler (2013).

Teorie della coscienza: oltre la materia e verso lo spirito

Si avvicinano ad una concezione più piena della coscienza Roger Penrose e Stuart Hameroff, che pur ritenendo che un’IA possa ottenere risultati superiori all’umano in tante attività da un punto di vista computazionale sostengono però che essa non potrà mai generare una coscienza, in quanto questa è legata a fenomeni fisici legati alla gravità quantistica (Teoria Orch-Or) (Hameroff & Penrose, 2014b, pp. 94-100).

La sede della coscienza, per Penrose e Hameroff, sarebbe neuronale, situata nei microtubuli, sottoposti a processi non computazionali, ma incorporati nella geometria dello spazio-tempo, un “condensamento” nel cervello dovuto alla riduzione oggettiva della funziona d’onda (2014, pp. 39-78). In ultima analisi, la coscienza per Penrose e Hameroff è situata in quella metà strada, tra la realtà materiale (cervello) e il mondo immanifesto (proprietà intrinseche dell’universo), che è tipica della fisica quantistica.

Su una lunghezza d’onda simile Severino, il quale parte da una prospettiva filosofico-heideggeriana, secondo la quale la coscienza è situabile nella dimensione eterna ed immutabile dell’essere ed è impossibile ridurla ad un processo meramente meccanico e neuroscientifico. L’IA, per Severino, è situata non sul dominio dell’“essere”, ma in quello materiale, di conseguenza è impossibile per essa divenire “forte” (Communitas Lecture, 2018).

Come vediamo, man mano ci siamo spostati sul fronte opposto, quello di coloro che sostengono il primato della coscienza sul cervello e sul corpo. Una “estensione” cerebro-corporea, che va incontro al campo del mentale, ma in senso “sottile”, più pieno, “verticale”.

In fin dei conti Edelman, con la teoria della “seconda natura”, sembra essere andato oltre il riduzionismo meramente neuronale. Eppure anche lui, come altri neuroscienziati materialisti, non si sposta verso un piano per davvero metafisico della coscienza.

Un quadro, questo, dal quale non esce nemmeno Damasio, che supera sì il dualismo cartesiano *res cogitans/res extensa*, ma rimane schiacciato sul livello orizzontale del corpo, della materia, e della produzione delle parti immateriali a partire da esso.

Spostandoci, invece, su un asse “verticale”, abbiamo già citato la diversità radicale della prospettiva buddhista, nella quale è la mente (coscienza) sottile ad avere il primato, suffragando il

tutto con esperimenti e ricerche “scientifiche” di vario genere. Anche la teoria di Faggin sulla penetrazione del cosciente in ambito materiale in una dinamica *top-down* può essere annoverata fra queste posizioni

In questo fronte rientrano anche l'altra antica conoscenza spirituale, che ha fornito una sua precisa dottrina della mente: l'induismo.

Nell'induismo, centri, funzioni e aspetti della mente sottile sono prioritari sulla parte strettamente cerebrale – il termine *manas* si associa alla mente, e ha la stessa radice del latino *mens* (Evola, 2010, p. 62). Questa non va intesa in senso psicologico, ma in relazione a un organo e a un potere, quello che agisce nella percezione, nelle relazioni motorie di un individuo e nella produzione di immagini (fantasia e immaginazione) (Evola, 2010, p. 62).

Il *manas* è “la radice o il principio fondamentale dei varî sensi, più o meno allo stesso titolo di ciò che la filosofia medioevale europea aveva chiamato *sensorium comune*” (Evola, 2010, p. 62). Nelle *Upanishad* si dice che è col *manas* che si vede, si ode, si gusta ecc. Di conseguenza, sono i sensi ad essere articolazioni del *manas*, non viceversa (Evola, 2010, p. 62). Il *manas* è dunque la forza unitaria che attiva i vari sensi.

Ma ancora, nell'induismo abbiamo la *buddhi*, la facoltà discriminativa, l'intelligenza superiore che consente il ragionamento, il discernimento e la capacità di distinguere ciò che è giusto da ciò che è sbagliato, ciò che è reale da ciò che è illusorio.

La *buddhi* è legata maggiormente alla consapevolezza e alla capacità di prendere decisioni basate su una visione chiara, diversamente dal *manas* che è più meccanico. È considerata un aspetto più “elevato” della mente perché connessa alla saggezza e alla realizzazione spirituale.

La *Ahamkara*, invece, è il principio che genera l'identificazione con il “Sé individuale”. È il senso dell'Ego che crea un senso di separazione tra “me” e “l'altro” (Swami Prabhavananda & Isherwood, 1993, p. 12). È il nucleo dell'identità personale, ma è anche responsabile di molti conflitti interiori, poiché tende a rafforzare l'illusione dell'individualità separata dal tutto (*maya*).

Infine, secondo alcune dottrine Vedanta, vi è la *chitta*, la facoltà della mente legata alla memoria, alle impressioni latenti (*samskara*) e al subconscio. Contiene tutte le esperienze passate e le abitudini mentali, che influenzano il comportamento e il pensiero conscio. È come un archivio, dove le azioni e i pensieri passati lasciano un'impronta.

Una possibilità di integrazione uomo-macchina a partire dalla coscienza umana

Transumanesimo e postumanesimo sono due filosofie, o “ideologie”, dei nostri tempi che portano avanti la tesi della possibile evoluzione dell'*Homo Sapiens* grazie all'ausilio delle tecnoscienze.

È una tematica che vale la pena sviluppare, per abbozzare possibili forme di coniugazione tra queste filosofie/ideologie, specchio del grande avanzamento tecno-scientifico contemporaneo, in una prospettiva che sia anch'essa fondata sull'umano, sulla sua coscienza, coscienza intesa in tutti i suoi livelli (compreso quello sottile).

Nell'ottica da noi proposta, potremmo rileggere alcune delle teorie più in voga riguardanti la fusione uomo/macchina, integrandole del piano sottile-coscientiale.

Innanzitutto, un Raymond Kurzweil con il suo *La Singolarità è vicina*. Kurzweil pone, sulla scorta di Vernor Vinge, la previsione – per lui, autodefinitosi “futurologo”, una “sicurezza” –, dell'avvento di una nuova entità super-intelligente, la “Singolarità”, che avrà fatto proprie tutte le conoscenze del mondo e della storia umana, integrando tutte le conquiste e i dispositivi prodotti dal progresso tecno-scientifico all'essere umano stesso, creando uno scarto, un punto di rottura – in matematica si direbbe: una curva a gomito (2008, pp. 10-11) – nella quale la nuova entità, la Singolarità, stacca l'uomo per come da noi conosciuto, creando una nuova specie (2008, pp. 24-30). Una nuova specie in grado addirittura di dominare integralmente i meccanismi base dell'universo, portando a quella che potremmo definire come una “coscientizzazione dell'universo”, il quale passa da materia inconsapevole a materia intelligente (2008, pp. 20-21).

E questo è il caso della “sesta epoca” dell'universo, che dalla forma primitiva, dalle strutture atomiche primordiali giunge ad un livello di massima coscienza.

Poi la teoria di Roberto Marchesini, filosofo ed etologo italiano, che con *Post-human. Verso nuovi modelli di esistenza* tratteggia una mutazione simile ma dalla matrice differente. Tale differenza è data dalla visione postumanista, che a differenza di quella transumanista, non è “antropocentrica”. Mentre il transumanesimo si presenta come un superamento dell'umano, in una particolare reinterpretazione del “sovrumanesimo” nietzscheano, almeno in alcuni suoi autori – un superamento che però lascia l'umano al centro dell'universo, essere dominante e portatore di volontà di potenza a mezzo tecnologico –, il postumanesimo, invece, abdica al ruolo antropocentrico, ponendo l'uomo come ente tra gli altri all'interno della biosfera, in un permanente stato di apertura ed ibridazione con l'alterità animale, la teriosfera (Marchesini, 2002, pp. 105-140) o macchinica (Marchesini, 2002, pp. 172-206), e attraverso la “tecnosfera” (Marchesini, 2002, pp. 245-279), ovvero l'ambito della tecnica non più vissuto come separato dalla vita umana, ma addirittura incarnato nella sua stessa storia evolutiva e genetica.

Prove di dialogo tra transumanesimo e postumanesimo esistono, a partire dalla figura di Max More in *Il sovrumano nel transumano* (2010).

Infine, interessante per il nostro tipo di considerazioni è *L'illuminismo oscuro*, di Nick Land (2021), intellettuale inglese, fondatore dell’“accelerazionismo”, corrente filosofica che sostiene l'idea che l'accelerazione drastica delle tendenze tecnologiche in atto possa provocare un totale

rimescolamento delle carte, una forma di caos creativo, che porti ad un nuovo ordine. La teoria di Land è quella dell’“Orizzonte bionico”, ovvero il punto massimo di fusione tra la tecnica e l’umano, “natura” e “cultura” – la classica dicotomia alla quale transumanisti e postumanisti cercano di porre risoluzione (Marchesini, 2002, pp. 72-86). L’orizzonte bionico è “la soglia della fusione definitiva di natura e cultura, in cui una popolazione diventa indistinguibile dalla sua tecnologia” (2021, p. 145).

Con la presentazione di queste tre teorie, sinteticamente estrapolate dai tre saggi di riferimento, è possibile sostenere che una trasformazione antropologica dovuta all’avanzamento tecnologico è sì possibile, ma va rettificata, interpretata, “integrata”, in aspetti per certi versi opposti a quelli presentati dagli autori.

I tre studiosi, infatti, sono tutti di marca materialista e darwinista, di conseguenza vivono la trasformazione antropologico-tecnologica come un processo evolutivo *bottom-up*, che può portare le intelligenze artificiali a produrre la loro coscienza, e perché no, a giungere alla Singolarità, o all’Orizzonte bionico, al transumano o al postumano, sulla base di processi sostanzialmente darwinistici, che prendono in considerazione solo l’evoluzione dei corpi, compreso quello delle macchine, e delle reti neurali, anche umane. Un’evoluzione vista solo in ottica materialistico-cerebrale.

A nostro avviso è invece possibile ipotizzare una nuova forma umana, ibridata, anche altamente, con dispositivi tecnologici di ogni sorta, tuttavia il dominio principale, e il comando del nuovo essere, sarà sempre e comunque nella coscienza umana.

L’analisi critica, di queste tre opere, emblema di tutto un movimento culturale, filosofico, anche emergente nel dibattito pubblico, e che ha sicuramente degli spunti interessanti, serve a leggere nella giusta ottica quella che può essere una possibile trasformazione antropologica, che può assumere dei tratti altamente fantascientifici. Una trasformazione che ha anche applicativi da un punto di vista medico-scientifico, correlati, inevitabilmente, a un discorso antropologico e politico – pensiamo ad Elon Musk e alle sperimentazioni sull’interfaccia cervello-computer (BCI) “*Telepathy*”, con tutte le suggestioni filosofico-politiche legate al movimento lungotermista.⁴

Del resto, dato l’alto livello di integrazione dell’uomo con i dispositivi tecnologici più disparati, potremmo considerare esso come già “transumano”, qualcosa che in qualche modo sia andato già

⁴ Ideologia a sfondo religioso nella quale si fanno previsioni e si mettono in campo progetti a lungo termine per l’umanità, per il benessere collettivo, delle vite altrui, anche in un futuro remoto, tanto da prosperare anche nell’intero universo (da qui anche i tentativi di Musk di andare su Marte). In pratica, l’umanità, con il suo potenziale, deve essere garantita sul lungo termine, e dunque si pensa anche all’esistenza delle generazioni a venire, che in questa visione contano quanto quella attuale. A tal fine bisogna necessariamente far fronte ai rischi esistenziali, cioè a quei rischi che potrebbero condurre al collasso tecnologico o all’estinzione del genere umano. Il manifesto del pensiero lungotermista è *What We Owe the Future* (2022), di William MacAskill e poi, annoveriamo, come ulteriore caposaldo del pensiero lungotermista *The Precipice* (2020) di Toby Ord.

oltre l'*Homo Sapiens*. Non è un caso che un Luciano Floridi, filosofo della comunicazione, abbia coniato l'espressione "*On Life*", e non più *On* o *Off Line*, a testimonianza della connessione oramai permanente dell'uomo nell'"Infosfera", nuova e vera dimensione presente, che connette le menti umane in una psiche unica e collettiva (2017).

Le reti neurali umane sono già cablate con quelle della macchina, che sia il PC, lo Smartphone o i social network. I social network, computer, tastiere, parafrasando De Kerckhove, sono già l'estensione del nostro sistema limbico, della nostra corteccia, ecc. (2014, p. 144).

E andando ancor più nel profondo, integrando queste conoscenze della dimensione sottile della coscienza, le nostre menti sono già invisibilmente connesse, con una connessione invisibile, sottile, appunto, "magica", l'Infosfera del Floridi alla quale accennavamo, traslazione della Noosfera del Teilhard de Chardin.

Dunque, fondamentale è comprendere come la macchina possa essere integrata all'uomo, recependo e ricevendo strati (minimi, liminali) della sua stessa coscienza, ma col presupposto che la coscienza sia innanzitutto sottile e innanzitutto umana. Innanzitutto umana, non in senso antropocentrico, ma "cosmocentrico", come nelle concezioni pre-moderne, e cioè proprio come sostiene Faggin, come principio massimo dell'universo, ma principio e forza base che viene meglio di tutti replicata dall'uomo, microcosmo nel macrocosmo.

In *Cyborg-Buddha. Dialogo sull'intelligenza artificiale* (2022), interessanti considerazioni vengono svolte sulle possibilità di forma assunte dall'uomo nel corso non solo dei secoli, o dei millenni, ma addirittura degli *yuga*, i cicli cosmici che scansano il tempo secondo visioni tradizionali della realtà (come induismo e buddhismo). Noi non possiamo sapere le forme che l'uomo abbia assunto in cicli cosmici precedenti, in quanto quel tipo di storia ci è abbastanza ignota, possiamo però supporre che dato il gigantesco avanzamento tecnologico del nostro tempo, sia possibile una modifica, come peraltro sta già avvenendo.

E infatti, secondo un James Hughes (2007, 2013, 2016) o un Mike La Torre (2015), proprio in virtù della non esistenza, nel buddhismo, di una mente fissa, un'essenza, di un Io, un Sé,⁵ bensì proprio per via della concezione della mente come flusso continuo, allora la macchina, l'IA, potrebbe divenire una sorta di sua estensione: un dispositivo parte di questo flusso e messo a fin di bene, atto cioè sia a superare la sofferenza, sia raggiungere stati di coscienza superiori, anche se non proprio la "liberazione" finale.

A tal fine, e in linea con la filosofia buddhista, emerge l'importanza dell'educazione, l'addestramento delle IA sulla base di principi di virtù, saggezza e compassione. Compassione

⁵ Da questo punto di vista il buddhismo si sposa, su alcuni punti, e in un senso parziale, con le teorie filosofiche post-strutturaliste sopramenzionate.

della quale dovrà beneficiare la stessa macchina, oramai essa stessa forma di vita, proprio in quanto incorporata nel suddetto flusso.

Tutto questo discorso, ovviamente, va preso con le pinze, anche conoscendo il reale potenziale “infinito” della mente umana secondo la prospettiva buddhista – “la mente è un oceano di saggezza”, parafrasando il Dalai Lama e Lama Yeshe.⁶

Quello che però ci preme asserire è che, nonostante questa modifica sia attuabile, e attuata, il dominio della coscienza della mente, prerogativa esclusiva dell'uomo, ne resterebbe inscalfito, e questo alla luce anche del buddhismo, scienza della mente non essenzialista o eternalista, e dunque la più aperta al divenire trasformativo della realtà, anche di quella umana.

Una prospettiva questa che si aggancia anche a certe teorie “panpsichiste”, e che potrebbero stare alla base anche della possibilità coscienziale della IA.

Il punto è però sempre lo stesso: “*En to Pan*”, “Tutto è Uno”, recitando quello che è diventato una sorta di slogan anche di una certa vulgata New Age, e la macchina al pari dell'uomo e dell'animale, oltre che degli alberi, dei fiumi, dei laghi, della montagna, delle stelle, partecipa a questo Uno, ne è realizzazione e concretizzazione vivente.

Tuttavia, in che modo detiene e realizza questo Uno, e quali i suoi strati “sottili”, che risiedono in parte maggiore nell'uomo, poi nell'animale, e nel regno vegetale e minerale? Potrebbe beneficiare degli strati sottili dell'uomo, seppure in forma molto minima, e in un processo di integrazione sempre maggiore uomo/macchina?

Queste sono le vere domande da farsi e dalle quali partire, per analizzare, comprendere e poi realizzare una tale e sempre migliore integrazione.

E, in tutto ciò, quella provocazione di senso offerta dallo sviluppo della tecnica e dall'evoluzione della macchina stessa può essere davvero utile per far capire cosa sia l'uomo e la sua coscienza? Solo carne, geni, e relazioni con l'ambiente esterno e la storia, o forse qualcosa di più? E questo “qualcosa di più” deriva dai fattori sopra menzionati oppure è proprio esso a produrli?

E l'universo è un insieme interconnesso di parti, per dirla con un David Bohm, con la sua teoria dell'Universo olografico, con l'ordine “implicato”, sottile che informa quello “esplicito”, grossolano (1996, pp. 231-281)?

⁶ Pensiamo alla pratica *Tukdam*, lo stato meditativo con il quale i grandi maestri del buddhismo, ultimo in ordine di tempo Lama Zopa Rinpoche (aprile 2023), si accompagnano alla morte, facendo lentamente dissolvere gli stati mentali, e con il corpo, che allo stesso modo va gradualmente incontro alla decomposizione. Ma in generale, diversi sono gli esempi di monaci buddhisti dalla straordinaria longevità, o le cui pratiche mostrano un grande beneficio per la salute corporea, oltre che mentale (Luang Pho Yai, Lama Itigilov, Luang Phor Khoon, Matthieu Ricard, Drubwang Rinpoche). Questo a dimostrazione che certi limiti fisici che il transumanesimo si prefigge di superare, potrebbero essere superabili da una stessa pratica spirituale e corporea opportuna, e che, probabilmente, tali limiti erano sconosciuti in cicli cosmici precedenti (in tempi di minore materializzazione della realtà), come ci narrano le storie straordinariamente longeve di personaggi dei testi sacri di svariate tradizioni religiose e spirituali.

È una sorta di simulazione alla stregua di un videogame, per dirla con Nick Bostrom (2003, pp. 243-255), e stimolante è dunque, anche in questo caso, la funzione della macchina?

A noi la sentenza!

Bibliografia

- Accoto, C. (2023). *Potenza della Latenza. L'intelligenza artificiale generativa: testi, immagini, agenti*, <https://cosimoaccoto.com/wp-content/uploads/2023/04/potenza-della-latenza-3-studi-sulla-generative-ai.pdf>
- Barthes, R. (1988). La morte dell'autore. In R. Barthes, *Il brusio della lingua. Saggi Critici* (vol. IV) (pp. 51-56). Einaudi.
- Blanchot, M. (1967). *Lo spazio letterario* (G. Zanobetti & G. Fofi, trad.). Einaudi.
- Bohm, D. (1996). *Universo, mente, materia*. Red Edizioni.
- Bostrom, N. (2003). Are You Living in a Computer Simulation?. *Philosophical Quarterly*, vol. 53, n. 211, 2432-55.
- Chalmers, D.J. (1996). *The Conscious Mind: in Search of a Fundamental Theory*. Oxford University Press.
- Changeux, J.P. (2012). *La Vie des formes et les formes de la vie*. Odile Jacob.
- Cole, D. (2004). The Chinese Room Argument. In E.N. Zalta (a cura di), *The Stanford Encyclopedia of Philosophy*. <https://plato.stanford.edu/archives/fall2004/entries/chinese-room/>.
- Communitas Lecture (2018). *Una discussione sulle criticità dell'Artificial Intelligence. Un dibattito, fra scienza e filosofia. Il confronto tra sir Roger Penrose e Emanuele Severino tenutosi a Milano il 13 maggio 2018 per le Communitas Lectures 2018*. <https://www.youtube.com/@communitaslecture3843>
- Crevier, D. (1993). *AI: The Tumultuous Search for Artificial Intelligence*. BasicBooks.
- Dalai Lama & Goleman, D. (2009), *Emozioni Distruttive*. Mondadori.
- Dalai Lama, Goleman, D., Thurman, R.F., Benson, H. & Gardner, H.E. (1993). *La scienza della mente*. Chiara Luce Ed.
- Damasio, A.R. (1995). *L'errore di Cartesio*. Adelphi.
- Damasio, A. (2000). *Emozioni e Coscienza*. Adelphi.
- Dawkins, R. (1992). *Il gene egoista. La parte immortale di ogni essere vivente*. Mondadori.
- De Kerckhove, D. (2014). L'impatto di internet sul sistema limbico sociale. In G. Greco (a cura di), *Pubbliche intimità. L'affettivo quotidiano nei siti di Social Network*. FrancoAngeli.

- Deleuze, G. & Guattari, F. (1975). *L'Anti-Edipo* (vol. I di *Capitalismo e schizofrenia*) (A. Fontana, trad.). Einaudi.
- Deleuze, G. & Guattari, F. (2017). *Mille piani. Capitalismo e schizofrenia* (P. Vignola, a cura di). Orthotes.
- Dennett, D. (1993). *La conscience expliquée*. Odile Jacob.
- Dennett, D.C. (1995). The unimagined preposterousness of zombies. *Journal of Consciousness Studies*, 2 (4), 322-326.
- Dennett, D.C. (2004). Quainare i qualia. In A. De Palma & G. Pareti (a cura di), *Mente e corpo. Dai Dilemmi della filosofia alle ipotesi della neuroscienza* (pp. 189-233). Bollati Boringhieri.
- Derrida, J. (1982). *“Différance”*. *Margini della filosofia*. University of Chicago Press.
- Dreyfus, H. (1972). *What Computers Can't Do*. MIT Press.
- Dugin, A. (2019), *Teoria e fenomenologia del Soggetto Radicale* (F. Marotta, A. Scarabelli & L. Siniscalco, a cura di). AGA.
- Edelman, G.M. (2007). *Seconda natura. Scienza del cervello e conoscenza umana*. Raffaello Cortina Editore.
- Edelman, G.M. (1992). *Sulla materia della mente*. Adelphi.
- Evola, J. (2010). *Lo Yoga della potenza*. Edizioni Mediterranee.
- Faggini, F. (2023), *Irriducibile. La coscienza, la vita, i computer e la nostra natura*. Mondadori.
- Floridi, L. (2017). *La quarta rivoluzione. Come l'infosfera sta trasformando il mondo*. Raffaello Cortina Editore.
- Foucault, M. (2004). Che cos'è un autore. In M. Foucault, *Scritti letterari* (pp. 1-21). Feltrinelli.
- Foucault, M. (1976). *Sorvegliare e punire: nascita della prigione* (A. Tarchetti, trad.). Einaudi.
- Fugali, E. (2016). *Corpo*. Il Mulino.
- Hameroff, S. & Penrose, R. (2014a). Consciousness in the universe: A review of the 'Orch OR' theory. *Physics of Life Reviews*, vol. 11, n. 1, 39-78.
- Hameroff, S. & Penrose, R. (2014b). "Reply to seven commentaries on 'Consciousness in the universe: Review of the "Orch OR" theory'". *Physics of Life Reviews*, 11 (1), 94-100.
- Haraway, D. (1995). *Manifesto cyborg. Donne, tecnologie e biopolitiche del corpo* (L. Borghi, a cura di; R. Braidotti, introduzione). Feltrinelli.
- Hayles, N.K. (1999). *How We Became Posthuman: Virtual Bodies in Cybernetics, Literature and Informatics*. University of Chicago Press.
- Harnad, S. (2002). "What's Wrong and Right About Searle's Chinese Room Argument?". In Bishop, M. & Preston, J. (a cura di), *Essays on Searle's Chinese Room Argument* (pp. 294-307). Oxford University Press.
- Harnad, S. (2005). *"Searle's Chinese Room Argument"*. Encyclopedia of Philosophy, Macmillan.

- <https://web.archive.org/web/20070116025618/http://eprints.ecs.soton.ac.uk/10424/01/chineseroom.html>.
- Hauser, L. (2006). “*Searle’s Chinese Room*”. Internet Encyclopedia of Philosophy. <https://web.archive.org/web/20090413030634/http://www.iep.utm.edu/c/chineser.htm>
- Horst, S. (2005). “*The Computational Theory of Mind*”. In E.N. Zalta, (a cura di), The Stanford Encyclopedia of Philosophy. <https://plato.stanford.edu/archives/fall2005/entries/computational-mind/>
- Hughes, J. (2007). The compatibility of religious and transhumanist views of metaphysics, suffering, virtue and transcendence in an enhanced future. *Global Spiral*, 8 (2).
- Hughes, J. (2013). Using neurotechnologies to develop virtues: a Buddhist approach to cognitive enhancement. *Account Res*, 20 (1), 27-41.
- Hughes, J. (2016). FTP025: James Hughes – Cyborg buddha, transhuman enlightenment and basic income. <https://futurethinkers.org/cyborg-Buddha-james-hughes-transhuman-enlightenment/>.
- Husserl, E. (2002), *Idee per una fenomenologia pura e una filosofia fenomenologica. II* (V. Costa, a cura di). Einaudi.
- Kurzweil, R. (2008). *La singolarità è vicina*. Apogeo Education.
- Lama Zopa Rinpoche (2016). *The Wish-Fulfilling Golden Sun of the Mahayana Thought Training eBook: Directing in the Short Cut Path to Enlightenment*. FPMT.
- Land, N. (2021). *Illuminismo oscuro*. Gog.
- LaTorra, M. (2015). What is Buddhist transhumanism?. *Theol. Sci.*, 13 (2), 219-229.
- Lyotard, J.F. (1981). *La condizione postmoderna: rapporto sul sapere* (C. Formenti, trad.). Feltrinelli.
- Lyotard J.F. (1978). *Economia libidinale* (M. Gandolfi, trad.). Colportage.
- MacAskill, W. (2021, 6 dicembre). *What We Owe the Future by William MacAskill*. Basic Books.
- Marchesini, R. (2002). *Posthuman. Verso nuovi modelli di esistenza*. Bollati Boringhieri.
- Marsh, M.N. (2010). *Out-of-body and near-death experiences brain-state phenomena or glimpses of immortality?*. Oxford University Press.
- McCarthy, J. (1998). *What is AI?*. Revised November 12, 2007. www-formal.stanford.edu/jmc/watisai.pdf
- Merleau-Ponty, M. (2003). *Fenomenologia della percezione* (A. Bonomi, trad.). Bompiani.
- Merleau-Ponty, M. (1969). *Il visibile e l’invisibile* (A. Bonomi, trad.). Bompiani.
- More, M. (2010). Il sovrumano nel transumano. *Divenire*, 4. <http://www.divenire.org/articolo.asp?id=41>
- Nagel, T. (2013). *Cosa si prova ad essere un pipistrello?*. Castelvechi.
- Newberg, A.B. (2014). The neuroscientific study of spiritual practices. *Front. Psychol*, 5, 215. <https://pubmed.ncbi.nlm.nih.gov/24672504/>.

- Newell, A. & Simon, H.A. (1976). *Computer Science as Empirical Inquiry: Symbols and Search*. Communications of the ACM, 19 (3), 113-126.
- Nilsson, N. & Lungarella, M. (2007), The Physical Symbol System Hypothesis: Status and Prospects, 50 Years of AI. *Festschrift*, LNAI 4850, 9-17.
- Ord, T. (2020, 3 Marzo). *The Precipice: Existential Risk and the Future of Humanity*. Bloomsbury.
- Parnia, S., Waller, D.G., Yeates, R. & Fenwick, P. (2001). A qualitative and quantitative study of the incidence, features and aetiology of near-death experiences in cardiac arrest survivors. *Resuscitation*, vol. 48, Issue 2, 149-156.
- Pournajaf, S., Morone, G., Goffredo, M., Bonaiuti, D. & Franceschini, M. (2021, settembre). Realtà virtuale applicata alla riabilitazione: evidenze cliniche e prospettive future. *MR. Giornale italiano di medicina riabilitativa*, vol. 35, nr. 3, <https://springerhealthcare.it/mr/archivio/realta-virtuale-applicata-alla-riabilitazione-evidenze-cliniche-e-prospettive-future/>.
- Swami Prabhavananda & Isherwood, C. (1993). *Gli aforismi yoga di Patanjali* (G. Fico & S. De Robertis, trad.). Mediterranee.
- Ring, K. & Cooper, S. (1999). *Mindsight: Near Death and Out of Body Experiences in the Blind*. William James Center for Consciousness Studies.
- Russell, S.J. & Norvig, P. (2003). *Artificial Intelligence: A Modern Approach*. Prentice Hall.
- Scheler, M. (2010). *Essenza e forme della simpatia* (L. Boella, a cura di). Angeli.
- Searle, J. (1983). *Intentionality: An Essay in the Philosophy of Mind*. Cambridge University Press.
- Searle, J. (1985). *Intenzionalità. Un saggio di filosofia della conoscenza* (D. Barbieri, trad.). Bompiani.
- Searle, J. (1980). Minds, Brains and Programs. *Behavioral and Brain Sciences*, 3 (3), 417-457.
- Searle, J. (1990). Is the Brain's Mind a Computer Program?. *Scientific American*, vol. 262, n. 1.
- Searle, J. (1992). *The Rediscovery of the Mind*. M.I.T. Press.
- Searle, J. (2004, 1 novembre). *Mind: a brief introduction*. Oxford University Press.
- Siconolfi, R. (2021, aprile). Lo stato di ordine e la funzione del caos nella filosofia-metafisica di René Guénon. *Informazione filosofica*, n. 3, 94-105.
- Skinner, B.F. (1958). *Verbal Behavior*. Acton. Copley Publishing Group.
- Turing, A.M. (1950, 10 ottobre). *Computing Machinery and Intelligence*, 49 (236), 433-460.
- Unione Buddhista Italiana. (2023, 22 novembre). *Concluso il Simposio internazionale "Prospettive buddhiste su coscienza, evoluzione e intelligenza artificiale"*. <https://unionebuddhistaitaliana.it/comunicati-stampa/concluso-il-simposio-internazionale-prospettive-buddhiste-su-coscienza-evoluzione-e-intelligenza-artificiale/>
- Unione Buddhista Italiana. (2022, 14 dicembre). *CYBORG-BUDDHA. Dialogo sull'intelligenza artificiale - Francesco Tormen e Stefano Davide Bettera*. <https://www.youtube.com/watch?v=GDgzvkwY0wk>

Van Lommel, P. (2001). Near Death Experience in Survivors of Cardiac Arrest: A Prospective Study in the Netherlands. *Lancet*, n. 358, 2039-2045.

Van Lommel, P. (2011). Near-death experiences: the experience of the self as real and not as an illusion. *Annals of the New York Academy of Sciences*, vol. 1234, 19-28.

Watson, J.B. (1924). *Behaviorism*. W.W. Norton & Company, Inc.

Oltre il confine. Reale e virtuale nella trasformazione della soggettività contemporanea di *Ignazio Iacone**

ABSTRACT (ITA)

Questo studio indaga criticamente la relazione tra reale e virtuale nella riconfigurazione della soggettività contemporanea, superando l'opposizione riduttiva tra digitale e tangibile. A partire dall'intuizione di Pierre Lévy – secondo cui il virtuale non si oppone al reale, bensì all'attuale – viene proposta una concezione della virtualità come modalità ontologica autonoma, capace di agire profondamente sulla costituzione dell'esperienza e dell'identità umana. Il contributo si fonda su una prospettiva teorica che intreccia l'ontologia della durata di Henri Bergson e la teoria dell'identità narrativa di Paul Ricoeur, per interrogare le forme plurali del sé che emergono nell'ecosistema digitale, dove l'intelligenza artificiale opera non solo come infrastruttura tecnica, ma come ambiente generativo e performativo. La tesi centrale ruota attorno alla proposta di un "umanesimo algoritmico", inteso come postura etico-ontologica capace di coniugare la creatività insita nella virtualità con la consapevolezza della finitezza, della vulnerabilità e della storicità dell'umano. In quest'ottica, il virtuale non si configura come una realtà parallela o illusoria, bensì come estensione e attualizzazione delle potenzialità dell'essere, radicate nella memoria, nella narrazione e nella facoltà di attribuire senso.

Parole chiave: Identità digitale, Reale e virtuale, Intelligenza artificiale, Bergson

Beyond the Border: The Real and the Virtual in the Transformation of Contemporary Subjectivity by *Ignazio Iacone*

ABSTRACT (ENG)

This study critically investigates the relationship between the real and the virtual in the reconfiguration of contemporary subjectivity, overcoming the reductive opposition between the digital and the tangible. Starting from Pierre Lévy's intuition – according to which the virtual is not opposed to the real, but rather to the actual – a conception of virtuality as an autonomous ontological modality is proposed, capable of profoundly affecting the constitution of human experience and identity. The contribution is based on a theoretical perspective that intertwines Henri Bergson's ontology of duration and Paul Ricoeur's theory of narrative identity to question the plural forms of the self that emerge in the digital ecosystem, where artificial intelligence operates not only as a technical infrastructure but also as a generative and performative environment. The central thesis revolves around the proposal of an "algorithmic humanism", understood as an ethical-ontological stance capable of combining the creativity inherent in virtuality with an awareness of the finitude, vulnerability and historicity of the human. From this perspective, the virtual is not configured as a parallel or illusory reality, but rather as an extension and actualisation of the potential of being, rooted in memory, narration and the faculty of attributing meaning.

Keywords: *Digital Identity, Real and Virtual, Artificial Intelligence, Bergson*

* Università LUMSA

1. Introduzione

Pierre Lévy ha scritto che “il virtuale non si oppone al reale, ma all’attuale” (Lévy, 1997, p. 763). Questa frase, che può sembrare enigmatica, è in realtà una chiave preziosa per comprendere il cambiamento profondo del nostro tempo. Laddove il senso comune contrappone il mondo “vero” a quello “digitale”, l’esperienza contemporanea ci dice che questa opposizione non regge più. Il virtuale non è un’illusione, ma una forma diversa di realtà che trasforma, agisce e struttura. Il termine virtuale deriva dal latino *virtus* che significa virilità ma anche forza o potere. *Virtus* è anche la radice della parola virtù che indica i punti di forza di una persona. Solo con Charles Peirce la parola *virtuale* ha assunto il significato analogo a *come se* (Peirce, 1918, p. 763). Nella prospettiva peirceana, quindi, una realtà virtuale sarebbe come la realtà, qualcosa che ha alcuni effetti della realtà, ma che non è reale. Secondo questa definizione virtuale equivale a illusione. Il virtuale non è reale, è un’illusione, un miraggio che però ha dei grandi poteri. Rispetto a questo primo significato, oggi ne possediamo un altro che è quello più evoluto rispetto al punto di partenza. Oggi, per esempio, virtuale equivale anche a ciò che viene generato da un computer. Una biblioteca virtuale è un luogo generato da un computer, ma con effetti pari a quelli di una biblioteca reale, per cui sarebbe assolutamente sbagliato affermare che la libreria virtuale è falsa. Per dirla con il filosofo David Chalmers (2023, pp. 244-245), virtuale

è uno spazio *immersivo*, *interattivo* e *generato da un computer*. [...] *Immersivo* significa che sperimentiamo l’ambiente come un mondo intorno a noi, con noi stessi presenti al centro. [...] *Interattivo* significa esiste un’interazione bidirezionale tra gli utenti e l’ambiente e tra gli oggetti nell’ambiente. L’ambiente influisce sugli utenti; gli utenti influiscono sull’ambiente. Gli oggetti nell’ambiente si influenzano a vicenda. Nella VR su vasta scala, l’utente controlla un corpo virtuale, un avatar, con opzioni di azione disponibili più o meno continuamente. *Generato da un computer* significa che l’ambiente è basato su un computer: cioè, un computer sta generando i segnali che vengono inviati ai nostri sistemi sensoriali. Ciò contrasta con ambienti non generati da un computer come il teatro, i film e la TV e con la realtà fisica comune.

Il virtuale, dunque, quanto è reale? Se teniamo presente solo il mondo dei videogiochi, dovremmo concludere che la dimensione virtuale è uno spazio immaginario popolato da corpi e altrettanti soggetti immaginari; ma se pensiamo, ad esempio, al mondo virtuale di *Second Life*, in cui si entra in comunicazione attraverso un avatar, dobbiamo concludere che la stanza e l’avatar sono perfettamente reali. È chiaro che l’avatar non è un corpo fisico, ma perché mai dovrebbe esserlo? L’avatar è un corpo virtuale perfettamente reale. Non c’è nulla di immaginario! Gli oggetti e i corpi virtuali così come gli eventi non sono oggetti fisici normali, ma sono ugualmente reali. Per chiarire questo dilemma occorre far riferimento alla *teoria causale della percezione*. Secondo

questa teoria, un oggetto che vediamo è sempre causa dell'esperienza di vedere un oggetto. Quando vedo allo zoo un leone, il leone determina la mia esperienza: innesca una catena causale che attraverso i fotoni, gli occhi e il nervo ottico, culminano nell'esperienza di vedere il leone. La visione di oggetti virtuali determina un'esperienza? La risposta è assolutamente positiva. Anche in questo caso si determina un processo casuale che culmina nell'esperienza di un oggetto virtuale. Si potrebbe obiettare che a determinare l'esperienza è lo schermo del computer o uno schermo all'interno di un visore, ma lo schermo è in realtà solo una stazione di passaggio, come la Tv. Quando in televisione vediamo il Presidente della Repubblica Italiana, Mattarella, vediamo per davvero Mattarella. La Tv o lo schermo di un computer sono solo stazioni di passaggio che mi aiutano a vederlo. Ciò che causa l'esperienza non è lo schermo ma Mattarella che vedo. Una prima conclusione, dunque, può essere che la realtà virtuale non è un'allucinazione. Gli oggetti virtuali esistono davvero. Sono strutture di dati concreti all'interno di un computer. Se il virtuale non è un'allucinazione è indispensabile esaminare i luoghi dove si consuma la tensione tra virtuale e reale. Il primo luogo in cui si manifesta la tensione tra reale e virtuale è il corpo.

Il corpo, che per millenni è stato il perno dell'identità e dell'azione, si trova oggi "dislocato" in ambienti digitali. Lavoriamo, amiamo, curiamo e comunichiamo attraverso schermi, avatar, dispositivi che trasmettono la nostra immagine ma non la nostra presenza fisica. Eppure, il virtuale non cancella il corpo: lo rifica. Nei mondi virtuali, il corpo si trasforma in immagine, simbolo, interfaccia. È un corpo "altro", spesso idealizzato, talora frammentato. Ciò che perdiamo in fisicità, lo guadagniamo in rappresentazione. Ma a quale prezzo? La domanda resta aperta. Un secondo fronte riguarda la natura stessa della realtà. Sempre più spesso, ciò che consideriamo reale è ciò che "funziona": ciò che un algoritmo ci propone, che un sistema predice, che un'interfaccia rende visibile. In questo nuovo scenario, non viviamo più in una realtà data, ma in una realtà mediata, filtrata, selezionata da macchine. I motori di ricerca, i social network, le intelligenze artificiali non si limitano a mostrarci il mondo: ce lo costruiscono. E noi, troppo spesso, confondiamo ciò che appare con ciò che è. Siamo alla presenza di una vera e propria metamorfosi che "non consiste semplicemente nel fatto che ciascuno è potenzialmente interconnesso agli altri, ma che questo entrare nel "mondo" significa entrare in qualcosa che segue una logica del tutto diversa. Ognuno si ritrova in un mondo sostanzialmente diverso da quello che si immagina e si aspetta" (Beck, 2024, p. 11). Secondo alcuni neuroscienziati un intelletto sintetico, costruito dalla mente umana, non è altro che una proiezione nella quale il soggetto umano si rispecchia per capire meglio sé stesso. La realtà, tuttavia, smentisce tale posizione; un'anima può conoscere sé stessa solo rispecchiandosi in qualcosa di simile a sé, cioè in un'altra anima contenuta in un essere vivente. Tutta la fenomenologia dell'Altro da Levinas, Martin Buber a Paul Ricoeur ha sempre insistito sul fatto che la coscienza personale è irriducibile

a qualsiasi macchina. È pur vero che nell'era dell'IA l'identità umana sopravviverà alla propria morte fisica – rimarranno le nostre immagini, la nostra voce – ma “ciò avverrà in un senso del tutto nuovo e inquietante per cui la mia immagine, la mia voce, il mio pensiero, potrebbero essere rielaborati da un'AI per far sopravvivere un surrogato di identità che può continuare a intessere relazioni e illudere che io sia ancora vivo” (Crippa & Girgenti, 2024, p. 74). È necessario, dunque, uno sforzo intellettuale che ponga l'Io a riparo dal suo appiattimento su un'IA. E questo implica che il soggetto umano torni a prendere coscienza della propria temporalità e precarietà strutturali e a considerare socraticamente che “sa di sapere e non sapere”; che torni a svolgere salutari esercizi di affettività, rivalutando positivamente i propri desideri, le proprie paure, i piaceri e i dolori. Una macchina non prova paura, desiderio, dolore; è potente ma non possiede volontà di potenza. Una macchina sa, ma non sa di sapere. Urge, in questo “cambiamento d'epoca” (Papa Francesco, 2020, p. 307) una svolta antropologica in grado di rendere l'uomo capace di distinguere l'umano dall'inumano. La tecnologia ha già determinato una dimensione indifferenziata nella quale il simbolo precede il reale, l'informazione precede l'evento. Baudrillard ammoniva profeticamente: “Gli esseri mortali hanno sconfitto, sopravvivendo, l'indifferenziato immortale. Gli immortali hanno il progetto di ricostruire un universo omogeneo e uniforme. [...] Abbiamo separato la riproduzione dal sesso; ora cercheremo di dissociare la vita dalla morte. La morte diventa una sorta di attrazione ontologica” (Baudrillard, 2007, p. 23). Riconsiderare la morte e la propria finitezza come elementi peculiari ed esclusivi dell'esistenza non significa consegnare l'uomo ad un destino di impotenza, ma di forza. La società ipertecnologica vive un rapporto contraddittorio con la finitezza e la morte. Da un lato la morte è considerata un tabù, dall'altro viene esibita come elemento demolitore del pudore in più settori dell'esistenza. La finitezza e la morte sono diventati problemi tecnici specifici; non costituiscono più un problema fisico-metafisico che deve essere vissuto e superato all'interno di una comunità di vita, ma un problema tecnico che deve essere risolto da tecnici con metodi tecnici. La morte dev'essere evitata quale fenomeno soggettivo e sostituita con la morte tecnica che non impegna personalmente. Negli anni '70 del secolo scorso J. Ratzinger (2020, p. 80) dichiarava:

Degradando la malattia e la morte e collocandole sul piano tecnicamente fattibile, si degrada contemporaneamente l'uomo. Nel tendere a ridurre l'*humanum*, s'incontrano oggi stranamente due opposti atteggiamenti: a una visione positivista e tecnocratica del mondo l'uomo è d'intralcio, quanto lo è al naturalismo integrale, che, vedendo nello spirito il vero ostacolo, tenta sempre più di denigrare l'uomo quale “animale mal riuscito”. Con la scelta dell'atteggiamento verso la morte viene scelto insieme l'atteggiamento verso la vita; per cui la morte ci può far da chiave per decifrare che cosa sia in fondo l'uomo. La brutalizzazione della vita umana cui oggi assistiamo è intimamente connessa al rifiuto del

problema della morte. Tanto la repressione quanto la banalizzazione risolvono il problema negando l'uomo stesso.

La dialettica tra reale e virtuale, umano e inumano rappresenta, dunque, una delle tensioni più pregnanti della nostra epoca, segnata dalla pervasività delle tecnologie digitali. Questa contrapposizione, tuttavia, non nasce con il cyberspazio, né trova nella realtà aumentata o nei mondi digitali la sua unica espressione. La riflessione filosofica ha da sempre interrogato le modalità dell'essere e dell'apparire, della presenza e della potenzialità. Henri Bergson (1859-1941), con la sua ontologia del tempo e della durata, ha offerto strumenti concettuali che, sebbene concepiti in un contesto pre-digitale, possono essere sorprendentemente interessanti per illuminare le trasformazioni antropologiche indotte dalla virtualizzazione del reale. Riflettere sul concetto di virtuale nella metafisica di Bergson, può aiutarci a evidenziare come tale concetto si configuri non come il contrario del reale, bensì come una delle sue modalità più profonde, tenendo in debito conto i rischi di una deleteria trasformazione antropologica indotta dall'accelerazione digitale e in particolare da un uso sempre più pervasivo dell'Intelligenza Artificiale.

2. Il virtuale tecnologico tra amplificazione e scollamento: ripartire da Bergson

Bergson elabora una critica radicale al tempo spazializzato della scienza moderna, opponendovi il concetto di *durée réelle*, la durata reale, come flusso continuo e qualitativo. Il tempo reale, per Bergson, non è una successione di istanti misurabili, ma una totalità viva, in cui il passato perdura nel presente e lo informa. “Noi percepiamo la durata come una corrente che non ci sarebbe possibile risalire. È il fondo del nostro stesso essere; è, come avvertiamo distintamente, la sostanza delle cose con cui siamo in comunicazione. Invano si fa brillare ai nostri occhi la prospettiva di una matematica universale; noi non possiamo sacrificare l'esperienza alle esigenze di un sistema” (Bergson, 2002, p. 37).

Nel saggio *Materia e memoria*, Bergson distingue tra memoria abituale e memoria pura. Quest'ultima conserva gli stati passati in forma virtuale, disponibili per l'attualizzazione ma non ridotti a semplici tracce registrate. La memoria, in quanto virtuale, è reale in un senso proprio: essa esiste, ma non è attuale. Il passato è virtuale, ma reale: esso agisce e reagisce nel presente. “La memoria praticamente inseparabile dalla percezione, inserisce il passato nel presente, contrae così in un'unica intuizione i molteplici momenti della durata, e quindi, per la sua duplice operazione, è la causa per cui di fatto percepiamo la materia in noi, mentre di diritto la percepiamo in sé” (Bergson, 2021, p. 58).

Per Bergson, il virtuale non è illusione o fantasma. È, al contrario, una modalità ontologica distinta dell'essere, in cui la potenza precede l'atto e lo trascende. In ciò, si distingue nettamente dalla tradizione platonica, nella quale il mondo sensibile è copia dell'idea, e dalla modernità cartesiana, che contrappone mente e materia in termini di realtà e rappresentazione. Bergson ha fatto della materia “un insieme di immagini” e non per questo ha affermato che l'essere è per come ci appare e nei limiti in cui ci appare. Non ha concepito l'essere come un fenomeno e la coscienza come un'intenzionalità d'atto.

La sua tesi, se vogliamo, è più radicale, [...] per Bergson, l'apparire stesso o, meglio, l'essere è l'essere dell'apparire. [...] L'apparire deve farsi assoluto emancipandosi dal rapporto con il suo presunto soggetto-sostrato. Esso non potrà più essere pensato come apparire di (soggetto-sostrato) o apparire per un (soggetto-sostrato). Esso dovrà coincidere con l'essere stesso, con il suo atto in atto, un atto intrascendibile perché fondamento di tutto “ciò” che appare. [...] “Materia” nel primo capitolo di *Matière et mémoire* diventa così il nome dell'apparire come assoluto, dell'apparire in quanto tale, nella sua trascendentale purezza. [...] Materia però è anche il nome per la coscienza virtuale. Tale coscienza virtuale non può in alcun modo essere intesa come supporto-sostrato dell'apparire, non è una coscienza- ego, non è il luogo della manifestatività. Ad essa partecipiamo nell'atto della percezione pura, il quale è, quindi, il fondamento in exteriorità, il nocciolo di impersonalità che assicura alla mia rappresentazione cosciente un radicamento originario e indistruttibile nell'essere e nella sua “verità” (Ronchi, 2011, pp. 111-113).

La società digitale si fonda su archivi, big data e cloud che esternalizzano la memoria. Questa esternalizzazione, lungi dal replicare la memoria bergsoniana, tende a ridurre il passato a dati attuali e manipolabili; si perde così il carattere qualitativo e non recuperabile dell'esperienza vissuta. Nei mondi virtuali e nei social media, l'identità si moltiplica in rappresentazioni e proiezioni. L'avatar non è un'estensione corporea come per Merleau-Ponty, ma una maschera senza carne, un'immagine che vive di retroazioni algoritmiche più che di intenzionalità vissuta. Il tempo digitale dissolve la durata, l'esperienza è disarticolata in notifiche, aggiornamenti, bit. *L'hic et nunc* pervasivo inibisce la continuità temporale: viviamo in un eterno presente aggiornato che cancella tanto la memoria quanto l'attesa. La durata bergsoniana è spezzata, compressa in una simultaneità che rifiuta la profondità. Applicando le teorie di Bergson al *novum* digitale, si potrebbe dire che le tecnologie non creano il virtuale, ma lo attualizzano secondo modalità predeterminate. Il rischio è che tale attualizzazione perda la tensione con la virtualità originaria, riducendo l'orizzonte dell'essere a ciò che è già stato codificato. L'uomo digitale rischia di non essere più “sorgente di senso”, ma terminale di attualizzazioni precostituite. Nella società dello spettacolo e della simulazione, il virtuale si de-realizza: non è più potenza, ma copia senza originale, simulacro che inghiotte il reale. La metafisica bergsoniana può allora diventare una

resistenza al nichilismo digitale, richiamando alla complessità dell'essere come divenire. L'era digitale mette in crisi il corpo vissuto; il corpo non è più veicolo della durata, ma interfaccia; le emozioni sono quantificate, la presenza è mediata. Serve allora una nuova antropologia che, alla luce delle intuizioni Bergsoniane, riscopra il corpo come tensione tra virtuale e attuale, tra memoria e creazione. Il virtuale, per Bergson, è la vita stessa nella sua potenzialità incessante. La tecnologia, sebbene tenda a oggettivare e frammentare, può essere ri-orientata a custodire la durata, a favorire esperienze che non siano solo attuali, ma cariche di passato e apertura. È necessario, dunque, un ripensamento etico e ontologico della tecnologia digitale alla luce di una metafisica della vita. Solo così l'uomo potrà evitare di ridursi a spettro tra schermi, e tornare ad abitare il tempo come creazione. La virtualità digitale, infatti, è una realtà umana. “Non un'altra realtà, bensì una parte dell'unica realtà umana” (Aroldi & Scifo, 2002, p. 145).

Concepire la dimensione virtuale come un universo alternativo, come una realtà parallela e autonoma rispetto a quella che conosciamo è un atteggiamento scientificamente debitore a una mitopoiesi tecnologica più che a una epistemologia attenta dell'uomo. La contrapposizione tra reale e virtuale dovrebbe lasciare il posto ad una partizione diversa, sempre all'interno dell'unica realtà umana; è possibile ritenere un pensiero, un affetto, una preghiera meno reali di un frutto bello da guardare? D'altra parte, il mondo virtuale strettamente connesso al complesso mondo delle tecnologie digitali non nasce dal nulla, ma sorge in quel complesso contesto culturale denominato postmoderno la cui caratteristica principale è il pluralismo senza fondamento. “Nel postmoderno – afferma Pierpaolo Donati (2000, p. 31) – non c'è un solo universalismo, ma molti universalismi”. E Marc Augé (1993, p. 52) nota come l'epoca postmoderna si caratterizzi come il tempo del “non-luogo”. Le posizioni critiche di Donati e Augé, pur condivisibili, non dicono tutto del virtuale. Il cyberspazio, infatti, non è esattamente un non-luogo, ma un luogo-altro. Sicuramente non sappiamo ancora dove ci troviamo quando entriamo nel cyberspazio; la domanda non ha ancora una risposta, tuttavia, non è più possibile concepire la realtà virtuale come semplice spazio dell'inutile e dell'evasione quanto come laboratorio di creatività in cui dare maggior risalto all'intelligenza. “Il cyberspazio potrà diventare un luogo di esplorazione dei problemi, di discussione pluralistica, di messa a fuoco di processi complessi, di decisioni collettive e di valutazione dei risultati” (Turkle, 1995, p. 82). La realtà virtuale, dunque, esige nuovi sistemi di significato, nuove strategie simboliche che permettano la costruzione di un nuovo ambiente sociale, dove poter rileggere l'esperienza reale. De Kerckhove (2001) afferma che lo spazio virtuale dovrebbe costituire il luogo dove l'intelligenza si materializza, facendosi visibile come architettura della connettività. Il programma di De Kerckhove sembra non realizzato, anzi, si assiste ad un vero e proprio appiattimento dell'intelligenza umana su quella artificiale con il rischio di una smodata moltiplicazione dell'Io che fatica a ritrovare la propria autenticità. Del

resto, come afferma Luciano Floridi (2002, p. 31): “Il digitale non è semplicemente qualcosa che potenzia o aumenta la realtà, ma qualcosa che la trasforma radicalmente, perché crea nuovi ambienti che abitiamo e nuove forme di agire con cui interagiamo”.

3. Identità e molteplicità: quanti “io” abitiamo?

L'essere umano ha sempre vissuto una certa molteplicità interiore. Ma oggi questa molteplicità diventa esplicita, visibile, strutturata. Abbiamo un “sé” su *LinkedIn*, un altro su *Instagram*, un altro ancora nei gruppi di lavoro o nei videogiochi online. Ogni contesto digitale genera un frammento di identità, una maschera, un ruolo. Questa moltiplicazione può essere ricca e liberatoria, ma anche faticosa e disorientante. In certi casi, l'individuo non si riconosce più, diventa spettatore di sé stesso, prigioniero di una performance continua. Il rischio è la perdita di coerenza, di radicamento, di silenzio interiore. Dovremmo dunque rimpiangere il passato, tornare a un mondo senza virtuale? Sarebbe un'illusione. Il virtuale non è una fuga dalla realtà, ma una sua nuova configurazione. Il compito filosofico è allora un altro: riconoscere che il virtuale ci impone di ripensare la realtà, non di rinnegarla. Non è più sufficiente chiedersi “chi sono?”. Nella contemporaneità, la domanda si è fatta plurale e sfaccettata: quanti sono gli ‘io’ che mi abitano?, in quali spazi si manifestano?, e con quali criteri possiamo considerarli autentici o fittizi, integri o disgiunti? Viviamo immersi in una realtà che è al tempo stesso materiale e digitale, tangibile e mediatica, corporea e simbolica. In questo contesto, parlare di “identità” al singolare è, per molti versi, anacronistico. L'esperienza contemporanea dell'io è dislocata, stratificata, eppure profondamente unitaria – o meglio: bisognosa di una nuova unità, capace di accogliere le differenze senza ridurle a frammentazione patologica. Siamo alla presenza di almeno di tre livelli di sé: sé corporeo, sé narrativo e sé digitale-performativo. Non si tratta di mere astrazioni filosofiche, ma di strutture verificabili attraverso le neuroscienze, la psicologia cognitiva e la teoria dei media. Il nostro primo “io” non parla. È il corpo che ci precede, che agisce, che percepisce, che è nel mondo prima ancora di poter dire “io sono”. La fenomenologia – da Husserl a Merleau-Ponty – ha descritto questa dimensione come quella dell'incarnazione: il corpo non è una cosa tra le cose, ma la condizione originaria della nostra presenza al mondo (Merleau-Ponty, 2003).

Anche la neuroscienza contemporanea conferma questo primato del corpo vissuto: la percezione di sé dipende da circuiti profondi che precedono la riflessione cosciente. Il cosiddetto *minimal self* è un sé pre-riflessivo, costitutivo dell'identità personale, e si manifesta nella continuità delle sensazioni, del ritmo cardiaco, del tono muscolare, del dolore e del piacere. L'io corporeo è dunque il fondamento, non in senso gerarchico ma in senso temporale e ontologico: siamo prima

ancora di saperlo. E ogni altra forma di identità – dalla narrazione alla rappresentazione virtuale – deve confrontarsi con questa base irrinunciabile. L'essere umano non si accontenta di vivere: vuole anche dare senso alla vita che vive. Questo bisogno genera il secondo strato dell'identità: il sé narrativo. Paul Ricoeur, in *Sé come un Altro*, ha mostrato che l'identità personale è sempre una forma di narrazione: siamo ciò che raccontiamo di essere. “La determinazione della nozione di persona, infatti, è fatta a partire dai predicati che le attribuiamo” (Ricoeur, 2011, p. 112).

La memoria non è un archivio neutro di dati, ma una costruzione attiva; il futuro, analogamente, non è attesa passiva ma progetto narrativo. Il sé narrativo unifica i frammenti dell'esperienza in un intreccio coerente, permettendoci di dire: “questa è la mia vita”. Anche in ambito clinico questa dimensione è centrale: la psichiatria fenomenologica e la psicoterapia narrativa utilizzano il racconto del paziente per ricostruire un'identità ferita o disgregata. Dove il corpo è presenza, il racconto è permanenza: dà durata al nostro essere nel tempo. Ma oggi, il racconto non si svolge solo nell'intimità della coscienza o nel dialogo con l'altro. Ogni giorno, milioni di persone mettono in scena la propria identità nei *social media*, nei mondi virtuali, negli spazi della comunicazione digitale. Si tratta di una forma inedita di identità: performativa, pubblica, algoritmica.

Sherry Turkle (2019) ha parlato di *multiple selves*, sottolineando come il cyberspazio favorisca la frammentazione e la sperimentazione identitaria. Luciano Floridi (2017), con la sua ontologia dell'infosfera, ha invece proposto una lettura più integrativa: non viviamo fuori dal reale quando siamo *online*, ma in una realtà ibrida, *onlife*, dove reale e virtuale si interpenetrano. In questo contesto, i nostri profili digitali sono più che semplici rappresentazioni: diventano spazi ontologici in cui l'identità si costruisce per iterazioni successive. Il *like*, il commento, la condivisione sono atti che non solo esprimono ma generano identità. L'io digitale non è una maschera posticcia: è un io reale, con conseguenze psichiche e sociali tangibili; è un io incorporato (Merleau-Ponty, 2017) sorto dalla tendenza dell'essere umano a condurre l'artefatto tecnologico all'interno del suo modo di esperire, facendolo divenire uno strumento per allargare la dimensione della sensibilità del corpo nel mondo. “Tuttavia la funzione di mediazione propria degli artefatti tecnologici non deve essere intesa, accogliendo le analisi della svolta empirica, come una mediazione tra io-soggetto e mondo-soggetto, ma piuttosto come mediazione nella quale entrambi, soggetto e oggetto, vengono costituiti” (Benanti, 2017, p. 105).

L'“io digitale” (*digital self*) è, dunque, quell'insieme di profili, di tracce e di rappresentazioni simboliche che ciascuno di noi proietta negli ambienti connessi. L'intelligenza artificiale (IA) non si limita più a essere un'infrastruttura neutra di calcolo: oggi agisce come “ambiente” plastico che modella, filtra e talvolta contraffà tali rappresentazioni, ridefinendo i confini dell'identità personale.

4. L'Intelligenza Artificiale tra virtualità attuale e futuro presente

Nel volgere di pochi decenni l'“io” è passato dall'essere un costrutto autobiografico custodito nella memoria e nel corpo a diventare una trama di segni che vive – e si modifica di continuo – nell'infosfera (Floridi, 2017). Questa transizione non riguarda soltanto la quantità di dati che ci rappresentano: muta i presupposti ontologici della soggettività. Se, come sostiene Maurizio Ferraris, l'umanità contemporanea è ormai una “documanità” (Ferraris, 2024) – un insieme di soggetti resi reali dal fatto di essere registrati – allora l'Intelligenza Artificiale, che di quelle registrazioni vive e si nutre, diventa a sua volta un dispositivo di co-produzione dell'identità. Nell'economia dell'infosfera, l'IA trasforma ogni gesto, percorso o preferenza in tracce terziarie, facendole poi rifluire in modelli predittivi che anticipano e influenzano le nostre future azioni. L'identità non è più l'origine dei dati, bensì l'esito del loro rimescolarsi continuo; un processo che riecheggia l'idea simondoniana di “individuazione” come divenire aperto, mai concluso, alimentato da nuove tensioni con il proprio ambiente tecnico. Qui l'IA svolge una funzione simile a quella del *Gestell* heideggeriano: un apparato che inquadra l'essere e lo predispone alla manipolazione, riducendo la pluralità delle possibilità all'insieme di ciò che può essere calcolato e predetto; ma, proprio perché “ci inquadra”, l'IA diviene anche lo specchio in cui scorgiamo il carattere generativo – e non sostanziale – del nostro io.

Quando un filtro di realtà aumentata liscia i tratti del nostro volto, o un sistema di raccomandazione compone la nostra bacheca social, l'esperienza soggettiva si struttura in una relazione triadica: corpo–dispositivo–algoritmo. Il corpo rimane, veicolo dell'essere-al-mondo, ma la sua presa sul reale passa ora attraverso un'interfaccia che ne ricalcola l'aspetto e il valore sociale in tempo reale. Donna Haraway aveva già colto, nel *cyborg*, la promessa e la minaccia di quell'ibridazione: nella misura in cui dissolviamo le opposizioni fra naturale e artificiale, ampliamo i nostri margini di azione ma diventiamo anche più dipendenti dai protocolli che tengono insieme gli ibridi (Haraway, 2020). Con l'IA generativa la dialettica si acuisce: l'identità visiva e vocale è replicabile all'infinito, come temevano gli allievi di Walter Benjamin riguardo alla perdita dell'“aura” dell'opera d'arte (Benjamin, 2017); l'aura non scompare: cambia di segno, perché l'unicità non risiede più nell'irripetibilità materiale, ma nella capacità di custodire la coerenza fra innumerevoli versioni di sé.

Foucault chiamava “tecnologie del sé” (Foucault, 2017) le pratiche con cui gli individui plasmano la propria soggettività; nell'ecosistema algoritmico tali pratiche vengono esteriorizzate e automatizzate in *dashboard* di *self-tracking* e *coach* digitali. A differenza degli esercizi antichi, però, esse sono oggi inscritte in architetture di potere che funzionano come apparati ideologici althusseriani: il *feed* personalizzato interpella l'utente “per nome”, inducendolo a riconoscersi nella

categoria che l'algoritmo gli assegna. È la logica del “tu sei questo” che alimenta i cicli di *engagement* ma restringe lo spazio dell'esplorazione identitaria. La questione diventa urgente quando le identità possono essere rubate o simulate: le frodi che sfruttano *deepfake* vocali e video sono in continuo aumento, evidenziando come la stessa IA che amplifica il sé possa frammentarne l'attendibilità. Una possibile risoluzione non può limitarsi alla protezione dei dati: in gioco è la protezione di una forma di vita. L'UE ha riconosciuto la centralità del tema classificando i sistemi di riconoscimento biometrico fra i rischi elevati dell'*AI Act*. Ma la normazione, da sola, non restituisce l'“aura” perduta: occorre una pedagogia dell'infosfera che insegni a praticare una sospensione critica di fronte alle interpellazioni algoritmiche, una nuova “tecnica del sé” che trasformi l'eterodirezione in co-autoregolazione. Rosi Braidotti parla di fioritura postumana (Braidotti, 2026) per indicare la possibilità di un'esistenza relazionale, interspecie e inter-tecnologica, che non rimuova ma integri l'alterità del dispositivo. Se assumiamo quella prospettiva, l'IA può diventare il *farmakon* stiegleriano: veleno quando riduce il soggetto a mero capitale comportamentale, rimedio quando inaugura spazi di creatività condivisa. È evidente che gli intelletti sintetici che “simulano capacità umane, in realtà, sono privi di qualità umane” (Papa Francesco, 2020), ecco perché bisogna tenerne conto per una interazione costruttiva esseri umani e macchine. Il pericolo è che “l'umano può essere condizionato a tal punto da assecondare il dispositivo di intelligenza artificiale, molto più che il contrario” (Paglia, 2024, p. 24). L'IA è certamente una straordinaria conquista tecnologica, ma limitata a compiti computazionali basati su dati quantitativi, al contrario, “l'intelligenza umana è organica, evolutiva e radicata nella corporeità e nell'esperienza vissuta. Inoltre, l'intelligenza umana è capace di cercare significato e scopo nella vita, una dimensione che l'IA non può replicare” (Dicastero per la Dottrina della Fede – Dicastero per la Cultura e l'Educazione, 2025, p. 17). Tutta la conoscenza che il soggetto umano porta con sé è un insieme armonico e misterioso a cui l'IA non può accedere; Massimo Chiriatti (2025, p. 121), giustamente si chiede:

È l'umano a fornire intelligenza o è il sistema ad apprendere in virtù della “sua intelligenza”? Risposta: nonostante i progressi, l'AI resta lontana dall'intelligenza umana. [...] L'associazione tra parola, pensiero e realtà prende forma nei nostri processi cerebrali, ma le macchine possono solo osservare ciò che noi esteriorizziamo di questi processi. [...] L'implicazione è che l'azione umana modella gli input dell'AI ed è influenzata dai suoi output.

Nell'intreccio fra io digitale e Intelligenza Artificiale, dunque, si misura il passaggio storico dal soggetto cartesiano, garante della propria unità, all'essere-dati distribuito, sempre in differimento rispetto a sé stesso. L'IA, lungi dall'essere un mero agente esterno, è ormai parte della “carne” simbolica del soggetto: lo decostruisce, lo rimappa, lo rimette in circolo. Riconoscere questo

statuto ibrido non significa arrendersi all'eteronomia del calcolo, ma ripensare la libertà come facoltà di negoziare – con lucidità filosofica e vigilanza politica – le molteplici versioni di noi stessi che l'algoritmo genera. In tale negoziazione si gioca l'avvenire dell'identità umana: o diventeremo l'ultimo anello di una catena di profilazioni, oppure inaugureremo – con le parole di Floridi – un'"infosfera" ospitale, in cui la pluralità degli io sarà la condizione stessa di una nuova forma di dignità. Una seria riflessione filosofica non può non tenere conto del fatto che oggi ci troviamo di fronte all'abbattimento del confine tra ciò che riteniamo naturale, biologico e vivente e ciò che consideriamo meccanico, tecnologico e inanimato. È chiaro che non possiamo non assumere una posizione netta nella ricerca di un equilibrio tra una visione ottimistica della tecnologia e il pessimismo del nostro lento adattamento culturale: intolleranza verso previsioni errate ed apocalittiche dell'IA, che saranno sempre sbagliate senza le necessarie correzioni umane, e intolleranza verso costruzioni di intelletti sintetici che non soddisfano criteri etici. I soggetti umani sanno di non sapere, l'IA non sa di sapere, non riesce a imparare e apprendere; è sulla base di questa discriminante che possiamo fronteggiare e gestire la novità. Siamo persuasi che l'IA ha cambiato per sempre il destino dell'umanità ed è divenuta la nuova ermeneutica del mondo; abbiamo altresì chiaro che essa ridefinisce le modalità con cui interpretiamo non solo la natura e il mondo, ma noi stessi. Nell'itinerario verso la comprensione della nostra identità, attraverso l'ausilio dell'intelligenza artificiale, possiamo affermare, con Edgar Morin, che solo la cultura è la dimensione attraverso la quale l'uomo può conquistare il suo essere. "Solo integrando il progresso tecnologico con una profonda consapevolezza culturale saremo in grado di esprimere appieno il potenziale dell'IA per l'arricchimento dell'umanità" (Brasioli, 2025, p. 115).

5. Conclusioni

Nel paesaggio ibrido della soggettività contemporanea, reale e virtuale si intrecciano in un gioco dinamico che questo studio ha voluto esplorare con qualche pretesa filosofica. L'identità si frammenta e moltiplica: al sé corporeo, radicato nella finitezza biologica, si affianca un sé narrativo che tesse il filo della memoria e dell'esperienza, mentre un sé digitale-performativo prolifera nell'infosfera, costantemente esposto alla costruzione artificiale di sé tramite algoritmi e piattaforme. Eppure, lungi dall'essere entità dissociate, queste dimensioni dell'io trovano una possibile integrazione attraverso la *durata vissuta* e la *narrazione unificante*. Come suggerisce l'intuizione bergsoniana, il tessuto della realtà è virtuale, è durata, un "perpetuo divenire che si fa e si disfa" nello spirito e nel cosmo (Di Fazio, 2013).

In questa prospettiva, la memoria non è mero deposito di dati ma ponte vivente tra passato e presente, una memoria virtuale in senso bergsoniano che preserva il passato in uno stato latente e

coesiste con il flusso del presente. La coscienza, nella sua durata, sfugge alle rigidità del tempo cronologico frammentato: essa è intreccio di momenti qualitativi, non riducibile alla serie di “istantanee” digitali o ai *log* computazionali. Il reale e il virtuale si compenetrano, dunque, nella nostra esperienza temporale: il virtuale digitale amplifica possibilità narrative e relazionali, ma deve ancorarsi al tempo vissuto, a quella continuità creativa che Bergson contrappone alla spazializzazione algoritmica del tempo. Soltanto così si evita l’illusione di ridurre l’individuo a una somma di stati, frammenti di dati, profili social, *output* di IA. Al contrario, la persona emerge come sintesi in divenire, frutto di un racconto che dà senso all’eterogeneità delle esperienze. In quest’ottica, riecheggia la lezione di Paul Ricoeur: la soggettività non è né una sequenza incoerente di eventi, né una sostanza immutabile, chiusa al divenire, bensì “proprio quel tipo di identità che soltanto la composizione narrativa può costruire nel suo dinamismo” (Ricoeur, 2011, pp. 184-185).

La molteplicità dell’io – corpo, coscienza autobiografica, persona digitale – necessita di una trama narrativa che ne riconosca la differenza senza dissolverne l’unità. È nel racconto di sé (verso sé stessi e verso gli altri) che l’individuo contemporaneo può cercare un filo conduttore, cucendo assieme il vissuto materiale e la rappresentazione virtuale. La durata interiore garantisce che, pur attraverso mille trasformazioni performative *online*, esista una continuità e autenticità potenziale del sé, fondata sulla memoria personale e sull’intenzionalità di significato. Tuttavia, a guidare tale costruzione identitaria deve essere un criterio di autenticità esistenziale che risiede nella consapevolezza della finitezza. È nel riconoscimento dei limiti – la vulnerabilità del corpo, la mortalità inevitabile, la fragilità intrinseca di ogni essere umano – che il soggetto può discernere ciò che è reale e significativo, al di là delle maschere digitali luccicanti. Solo accettando e rivelando la propria vulnerabilità l’essere umano esce dalla massa anonima ed inautentica e instaura relazioni sincere. Questa intuizione, vicina al pensiero di Simone Weil, ci dice che l’io digitale rischia di gonfiarsi in un’immagine di sé onnipotente e invulnerabile, mentre è proprio la coscienza del limite – del corpo esposto alla sofferenza, alla malattia, alla morte – a radicare nuovamente l’identità nella verità dell’esperienza. La finitezza riconosciuta diventa così misura di verità personale: ricorda al soggetto che egli non è un algoritmo immortale né un profilo curato all’infinito, ma un essere incarnato in un tempo che scorre irreversibilmente. Da questa consapevolezza può sorgere una rinnovata autenticità, che rifugge sia le derive nichilistiche sia l’artificiosità di un sé interamente performativo. In altre parole, la morte e la vulnerabilità non sono antagoniste della vita digitale, ma criteri ultimi che ne svelano l’essenza, riportandoci all’umiltà di esistere in relazione con altri mortali, bisognosi gli uni degli altri per attribuire significato alla propria storia.

Alla luce di queste riflessioni, proponiamo l'idea di un "umanesimo algoritmico" quale posizione etica e ontologica di resistenza alla riduzione computazionale della soggettività. Nell'era dell'IA e dei *Big Data* vi è la tentazione di trattare l'essere umano come un nodo prevedibile di informazioni, un profilo di preferenze da tracciare e manipolare. Contro questa visione, l'umanesimo algoritmico afferma la centralità dei valori umani e della complessità del soggetto di fronte al potere pervasivo dei processi automatici. Esso si nutre della durata bergsoniana e della coscienza narrativa ricoeuriana, riconoscendo nell'uomo un perpetuo divenire creativo non traducibile in *bit*. Allo stesso tempo, prende le mosse dalla preoccupazione esistenziale di pensatori come Heidegger e Agamben, che hanno denunciato la deriva tecnologica in cui l'uomo è ridotto a ente tra enti, a risorsa tra risorse all'interno di un apparato tecnico disumanizzante. Come suggerisce Luciano Floridi, occorre un nuovo umanesimo digitale per navigare nel mondo virtuale. Ciò implica sviluppare sistemi di IA e ambienti virtuali che tengano conto dell'intera gamma dell'esperienza umana, includendo empatia, creatività, spontaneità e capacità di errore – tutte qualità che ci ricordano come la persona sia più grande della somma computabile dei suoi dati. In chiave ontologica, l'umanesimo algoritmico riafferma che la dignità del soggetto risiede nella sua irriducibilità: l'io non è algoritmo, ma porta con sé il mistero della libertà creatrice e della coscienza riflessiva (quella capacità, direbbe Bergson, di scegliere e innovare, sfuggendo a un futuro già tutto calcolato). In chiave etica, esso esorta a progettare l'IA con l'uomo e per l'uomo, mantenendo aperto uno spazio di scelta e responsabilità umana in ogni decisione automatizzata. Questo atteggiamento di resistenza costruttiva non è un nostalgico umanesimo antropocentrico, bensì un umanesimo relazionale e complesso, consapevole che oggi "l'io viene dopo il tu; noi siamo l'esito delle relazioni che ci legano agli altri" (Fuschetto, 2020) e che persino nell'infosfera l'identità sboccia dall'intreccio con altre identità e agenti (umani o artificiali che siano). In accordo con Edgar Morin – l'"umanista planetario" della complessità – si tratta di tenere insieme tutte le dimensioni dell'umano, abbracciandone le contraddizioni e vulnerabilità e perseguendo "la ricerca di un Umanesimo Planetario" (Dominici, 2021). È un umanesimo nuovo, algoritmico e al contempo planetario, perché capace di confrontarsi con l'intelligenza delle macchine senza perdere di vista l'Altro.

Malgrado le sfide poste dalla realtà virtuale e dall'IA, rimane aperta la possibilità di una soggettività umana creativa e relazionale anche nell'era dell'infosfera. La condizione è che l'uomo riconosca sé stesso come essere-in-relazione, finito ma libero, capace di memoria e immaginazione, e rifiuti di specchiarsi solo nell'artificiale. La "durata vissuta" – questa misteriosa continuità interiore – unita alla coscienza della propria vulnerabilità, offre un ancoraggio per non smarrirsi nel flusso digitale. Da essa scaturisce la facoltà di dare senso al nuovo, di trasformare l'informazione in conoscenza vissuta e l'interconnessione tecnica in incontro. In definitiva,

l'essere umano contemporaneo – corporeo, narrativo, digitale – può ritrovarsi persona autentica nell'infosfera solo riconoscendo che le relazioni sono la vera trama dell'esistenza e che la libertà di creare significati nuovi è la firma inconfondibile della soggettività umana.

Bibliografia

- Aroldi, P. & Scifo, B. (2002). *Internet e l'esperienza religiosa in rete*. Vita e Pensiero.
- Augé, M. (1993). *Nonluoghi: Introduzione a una antropologia della surmodernità*. Elèutera.
- Baudrillard, J. (2007). *L'illusione dell'immortalità*. Armando Editore.
- Beck, U. (2024). *La metamorfosi del mondo*. Laterza.
- Benanti, P. (2022). *La condizione Tecno-umana. Domande di senso nell'era della tecnologia*. EDB.
- Benjamin, W. (2017). *L'opera d'arte nel tempo della sua riproducibilità tecnica*. Bompiani.
- Bergson, H. (2002). *L'evoluzione creatrice*. Raffaello Cortina.
- Bergson, H. (2021). *Materia e memoria* (A. Pessina, a cura di). Laterza.
- Braidotti, R. (2013). *The Posthuman*. Polity Press.
- Brasioli, D. (2025). Un futuro presente. Perché l'AI cambia il mondo. *Limes*, 4.
- Chalmers, D.J. (2023). *Più realtà. I mondi virtuali e i problemi della filosofia*. Raffaello Cortina.
- Chiriatti, M. (2025). Incoscienza artificiale. *Limes*, 4.
- Crippa, M. & Girgenti, G. (2024). *Umano, Poco Umano. Esercizi spirituali contro l'Intelligenza Artificiale*. Piemme.
- De Kerckhove, D. (2001). *The Architecture of Intelligence*. Birkhauser.
- Dicastero per la Dottrina della Fede – Dicastero per la Cultura e l'Educazione. (2025). *Antiqua et nova. Nota sul rapporto tra intelligenza artificiale e intelligenza umana*. San Paolo.
- Di Fazio, F. (2013). Percezione, rappresentazione e memoria nella filosofia di Bergson. *Consecutio Rerum*. Edizioni Efestò.
- Dominici, P. (2021, 8 luglio). Edgar Morin abitatore del tempo dell'imprevedibilità. *Il Sole 24 Ore*.
- Donati, P. (2000). *La cittadinanza societaria*. Laterza.
- Ferraris, M. (2024). *Documanità. Filosofia del mondo nuovo*. Laterza.
- Floridi, L. (2017). *La quarta rivoluzione. Come l'infosfera sta trasformando il mondo*. Raffaello Cortina.
- Floridi, L. (2022). *Etica dell'Intelligenza Artificiale. Sviluppi, opportunità, sfide*. Raffaello Cortina.
- Foucault, M. (2017). *Sull'origine dell'ermeneutica del sé. Due conferenze al Dartmouth College*. Cronopio.

- Fuschetto, C. (2020, 9 agosto). Umanesimo dell'altro bit: intervista a Luciano Floridi. *Scienza in rete*. <https://www.scienzainrete.it/articolo/umanesimo-dellaltro-bit-intervista-luciano-floridi/cristian-fuschetto/2020-08-09>
- Haraway, D. (2020). *Manifesto cyborg. Donne, tecnologie e biopolitiche del corpo*. Feltrinelli.
- Lévy, P. (1997). *Il virtuale*. Raffaello Cortina.
- Merleau-Ponty, M. (2003). *Fenomenologia della percezione*. Bompiani.
- Paglia, V. (2024). *L'algoritmo della vita. Etica e Intelligenza Artificiale*. Piemme.
- Papa Francesco. (2020a). *Discorso ai partecipanti all'assemblea plenaria della Pontificia Accademia per la Vita* (28 febbraio 2020). Libreria Editrice Vaticana.
- Papa Francesco. (2020b). *Discorso ai partecipanti all'assemblea plenaria della Pontificia Accademia per la Vita* (28 febbraio 2020). https://www.vatican.va/content/francesco/it/speeches/-2020/february/documents/papa-francesco_20200228_accademia-perlavita.html
- Peirce, C. S. (1918). Virtual. In J.M. Baldwin (a cura di), *Dictionary of Philosophy and Psychology* (vol. 2). (pp. 764-766). Macmillan.
- Ratzinger, J. (2020). *Escatologia. Morte e vita eterna*. Cittadella Editrice.
- Ricoeur, P. (2011). *Sé come un altro*. Jaca Book.
- Ronchi, R. (2011). *Bergson. Una sintesi*. Christian Marinotti Edizioni.
- Turkle, S. (1995). *La vita sullo schermo*. Apogeo.
- Turkle, S. (2019). *Insieme ma soli. Perché ci aspettiamo sempre più dalla tecnologia e sempre meno dagli altri*. Einaudi.

Specula. L'intelligenza artificiale tra guerra, cura e sorveglianza. Etica e potere nella società del controllo digitale

di *Alberto Pesce**

ABSTRACT (ITA)

Il saggio analizza le trasformazioni sociali e culturali indotte dall'intelligenza artificiale, con particolare attenzione al suo impatto sui sistemi di controllo, prevenzione e gestione della criminalità, sui sex robot, sulle armi computerizzate e sul fenomeno del Sesto Potere o, meglio, sulla società della sorveglianza. Attraverso l'esempio del crime mapping, vengono esplorati i meccanismi predittivi e di sorveglianza che rischiano di trasformarsi in strumenti di discriminazione algoritmica e di rafforzamento delle disuguaglianze sociali. I sex robot sono la nuova frontiera delle interazioni affettive tra macchina e robot, mentre gli scenari di guerra diventano sempre di più luoghi dei "War Games". Il lavoro introduce l'analisi del "sesto potere", inteso come il potere invisibile e pervasivo dell'IA nel modellare opinioni, comportamenti e decisioni collettive. Centrale è la riflessione sull'etica dell'IA: la necessità di regole condivise, trasparenza, accountability e giustizia sociale per evitare che l'automazione si traduca in nuove forme di dominio. Specula invita a un ripensamento critico dei rapporti tra tecnologia, potere e democrazia, sottolineando la responsabilità collettiva nella costruzione di un futuro tecnologico equo.

Parole chiave: IA, crime mapping, sesto potere, sorveglianza, etica

Specula. Artificial Intelligence Between War, Care, and Surveillance: Ethics and Power in the Society of Digital Control

by *Alberto Pesce*

ABSTRACT (ENG)

The essay analyzes the social and cultural transformations induced by artificial intelligence, with particular attention to its impact on crime control, prevention and management systems, sex robots, computerized weapons and the phenomenon of the Sixth Estate or, rather, on the surveillance society. Through the example of crime mapping, predictive and surveillance mechanisms are explored that risk turning into tools of algorithmic discrimination and reinforcement of social inequalities. Sex robots are the new frontier of affective interactions between machine and robot, while war scenarios are increasingly becoming places of "War Games". The work introduces the analysis of the "sixth power", understood as the invisible and pervasive power of AI in shaping collective opinions, behaviors and decisions. Central is the reflection on the ethics of AI: the need for shared rules, transparency, accountability and social justice to prevent automation from translating into new forms of domination. Specula calls for a critical rethinking of the relationship between technology, power and democracy, emphasizing the collective responsibility in building a fair technological future.

Keywords: AI, crime mapping, ethics, surveillance, sixth power

* Università Link Campus

1. Introduzione

Negli ultimi anni l'IA da argomento letterario e di narrativa di tipo fantascientifico ha trovato spazio nelle interazioni personali e collettive della nostra società. La sua diffusione si è manifestata in molti settori della vita quotidiana degli agenti sociali. La possiamo ritrovare nei sistemi di diagnostica medica, nei sistemi per monitorare la criminalità, nei nuovi modelli di armi nei contesti di guerra, negli algoritmi che regolano le interazioni tra individui e robot, nonché nelle applicazioni di realtà aumentata. L'intelligenza artificiale influisce, in modo sensibile, sugli stili di vita, nonché sulle relazioni e interazioni tra gli esseri umani.

Nel campo della medicina, le tecnologie di IA offrono la possibilità di diagnosi più rapide, trattamenti su misura e un continuo monitoraggio delle condizioni fisiologiche del paziente (Topol, 2019). Allo stesso modo, nel settore della sessualità si mostrano cambiamenti significativi: robot sessuali in grado di imitare emozioni e prestazioni sessuali, algoritmi che coordinano incontri sulle app di appuntamenti, e modalità di intimità totalmente mediate dal digitale (Danaher, 2017). Nel contesto militare, la presenza di sistemi d'arma autonomi e di droni equipaggiati con intelligenza artificiale ha suscitato interrogativi etici e strategici senza precedenti, tanto che organizzazioni internazionali da anni stanno discutendo di un possibile divieto di tali strumenti (Garcia, 2024).

Nonostante queste differenze tra i vari settori, c'è un elemento comune che collega queste applicazioni: il ruolo della sorveglianza.

La sorveglianza attuale non è solamente un caso particolare riservato alle istituzioni di sicurezza, ma rappresenta un principio organizzativo ampiamente diffuso, in grado di influenzare il comportamento umano, modellare le aspettative e ridefinire il significato stesso di privacy (Bauman & Lyon 2013). L'intelligenza artificiale, in questo contesto, non si limita a essere un semplice strumento tecnico, ma svolge un ruolo di agente per la ristrutturazione sociale e di ridefinizione dell'identità, osservando, raccogliendo, elaborando e frequentemente, intervenendo modificando ruoli e status sociali (Lyon, 2020).

Il cambiamento da un sistema di controllo centralizzato come per il Panopticon (Foucault, 1975) a un sistema fluido e diffuso, in una società contemporanea liquida (Bauman, 2012), si è accelerato dall'intelligenza artificiale, attraverso rituali di interazione virtuali (Goffman, 1988). Gli algoritmi che selezionano contenuti, analizzano profili utente e anticipano comportamenti non necessitando più della presenza fisica di un osservatore, l'identità digitale è costruita e modellata per i bisogni della società dei consumi, la sorveglianza, intanto, avviene tramite reti di dati e dispositivi, utilizzati dagli attori sociali quotidianamente, dai telefoni agli elettrodomestici connessi, fino ai sistemi di monitoraggio delle aree urbane, nulla sfugge a questo nuovo "grande

fratello” (Orwell, 2019). Possiamo definire questa vita digitale come una “cultura della sorveglianza autorizzata”, in cui il monitoraggio è così radicato da diventare una parte integrante della vita sociale, un aspetto accettato, se non addirittura ricercato, per motivi di comodità e di efficienza (Lyon, 2020).

Tuttavia, la sorveglianza non è semplicemente una questione tecnica o macro-strutturale, le relazioni tra le persone sono sempre delle rappresentazioni svolte di fronte a un pubblico (Goffman, 1997), perciò la creazione di una identità virtuale diventa un elemento essenziale per poter agire e vivere questa “nuova realtà aumentata o virtuale”. Il controllo avviene attraverso una rete di sistemi di algoritmici che analizzano, classificano e, a volte, limitano l’agire virtuale delle persone. Pertanto, il contesto sociale è diventato un palcoscenico costante di visibilità, dove gli individui adattano continuamente la propria identità in base alle aspettative non solo degli altri, ma anche dei sistemi automatizzati che gestiscono la visibilità e l’accesso.

La relazione tra intelligenza artificiale, controllo e sorveglianza, emerge in diversi settori della nostra vita, nell’ambiente sanitario la capacità di effettuare visite e diagnosi tempestive (Topol, 2019) è associata al pericolo, da parte dell’amministrazione sanitaria, di un’ampia acquisizione di informazioni personali e riservate del paziente, questi dati sono frequentemente gestiti da società private che le impiegano per scopi commerciali.

Nel campo della sessualità robotica, le informazioni riguardanti desideri, preferenze e comportamenti sessuali, generano risorse utili per le strategie di marketing e per la costruzione di profili psicologici, più si conosce il cliente meglio si possono soddisfare i suoi bisogni (Balistreri, 2023).

In contesti bellici, così come nella prevenzione e repressione del crimine, l’uso dell’IA è ormai uno strumento insostituibile per la sorveglianza strategica, ad esempio attraverso i sistemi di riconoscimento facciale si riducono i tempi necessari per identificare un bersaglio e per pianificare le operazioni, delegando il potere decisionale dalle persone alle macchine (Oliveros-Aya, 2023).

Sebbene vi siano diverse modalità di utilizzo dell’IA che garantiscono vantaggi concreti alla vita degli esseri umani, a caratterizzare il nostro tempo è l’emergere di un modello di sorveglianza personale sempre più diffuso. Tale modello di “tracciamento dei nostri bisogni presenti e soprattutto futuri” ha come obiettivo principale ottenere valore economico e esercitare un controllo sociale. Può essere considerato come una forma di “capitalismo della sorveglianza”, in cui i dati sul comportamento vengono convertiti in risorse per prevedere e influenzare le azioni future delle persone. (Zuboff, 2019).

Due aspetti, in particolare, richiedono un’analisi dettagliata: da un lato, la sorveglianza volta alla prevenzione e al contrasto del crimine, attraverso pratiche quali il crime mapping e la polizia

predittiva (Ratcliffe, 2010), dall'altro, la sorveglianza caratterizzata del sesto potere (Bauman & Lyon, 2013), cioè quella economico-culturale della società dei consumi, che unisce controllo e marketing in un unico processo. In entrambe le situazioni, l'intelligenza artificiale non funge solamente da strumento di osservazione, ma svolge un ruolo attivo nel processo decisionale, influenzando le priorità degli interventi, definendo le categorie di rischio e addirittura plasmando le identità digitali degli individui (Han, 2017).

Il presente saggio avrà un duplice obiettivo: analizzare le maggiori applicazioni dell'IA e i contesti sociali dove viene utilizzata; quindi, procedere a degli approfondimenti e a due analisi dettagliate.

La prima si concentrerà sull'importanza e sul ruolo del monitoraggio nella giustizia e nella sicurezza pubblica, evidenziando il potenziale e le problematiche associate alla mappatura dei crimini e relativi Bias algoritmici, nei sistemi di previsione dei delitti (Pronzato, 2024).

La seconda parte analizzerà la sorveglianza come un "sesto potere" nel contesto del consumo, in cui il controllo continuo è legato alla creazione di "desideri" e alla riproduzione di comportamenti dei consumatori, indirizzati e regolati.

L'obiettivo non è quello di condannare l'IA in sé, ma di offrire un'analisi critica delle sue applicazioni, mostrandone sia i rischi che i benefici e presentando un metodo etico in grado di orientare il suo sviluppo e la sua regolamentazione e normativa.

In un'epoca in cui ogni clic, ogni movimento e ogni interazione sono soggetti a monitoraggi, la questione principale di questo saggio è riflettere sulla gestione della nostra vita attraverso un algoritmo, fino a che punto possiamo vivere sotto un controllo costante e quale prezzo siamo disposti a pagare per i benefici che l'intelligenza artificiale promette, o ci illude, di offrirci.

Questo saggio non nasce come una ricerca etnografica con raccolta di dati sul campo, ma si inserisce nell'ambito di una riflessione teorico-speculativa. La scelta di non raccogliere dati empirici è motivata dalla natura stessa dell'oggetto: l'IA, intesa come "infrastruttura di sorveglianza", è un fenomeno in rapida trasformazione, utilizzato in modo diverso nelle sue pratiche sociali e istituzionali (Zuboff, 2019). In questo contesto, una ricerca empirica richiederebbe la raccolta e il confronto di dati nei singoli settori, mentre l'obiettivo attuale è riflettere sull'utilizzo degli algoritmi, per poi sviluppare in futuro analisi specifiche capaci di cogliere i frammenti contingenti di un fenomeno in continua evoluzione. La riflessione teorica, invece, consente di anticipare scenari, definire categorie concettuali e proporre chiavi interpretative che possano orientare, in un secondo momento, indagini empiriche più mirate.

Il lavoro si configura quindi come uno studio preliminare, o pilota (Memoli & Saporiti, 1995, p. 105), che possa garantire di ibridare analisi e riflessioni, in una interpretazione sociologica che possa comprendere le prospettive micro e macro della vita sociale, anche quella digitale.

L'IA, in modo critico, non può essere definita come una tecnologia neutrale, ma come il risultato di rapporti di potere, interessi economici e strategie di tipo politico, per questo motivo la domanda che dovremmo porci è a chi può essere utile l'IA, ma soprattutto, chi ne trae vantaggi e chi, invece, ne subisce effetti e conseguenze negative. Gli algoritmi non sono strumenti imparziali, sono carichi di simboli che plasmano il senso, le identità e i comportamenti collettivi, attraverso modelli di rituali di interazione digitali. Questi sistemi determinano ciò che noi definiamo visibile e pensabile, costruendo dei “frame” collettivi che forniscono significato e agire sociale (Goffman, 2001), questa impostazione epistemologica non ha lo scopo di descrivere verità assolute, ma di proporre cornici interpretative, dei frame plurali, e aperture speculative che possano fornire risposte oppure stimolare nuove domande di ricerca.

In questo contesto le strutture sociali perdono rigidità, la sorveglianza non può più essere rigida e centralizzata, come nel Panopticon di Foucault (1975), ma diventa fluida, mobile e adattiva (Bauman, 2012). L'IA rappresenta esattamente questa fluidità, con sistemi, spesso invisibili, che raccolgono, elaborano e utilizzano grandi quantità di dati in tempo reale, in spazi diversificati, urbani oppure virtuali.

Perciò, la sorveglianza non è più un'eccezione, ma una prassi quotidiana, non è solo subita, ma anche partecipata, quasi voluta, gli individui stessi collaborano, con l'uso di app, social network e dispositivi digitali, alla riproduzione dei dati che determinano la loro sorveglianza (Zuboff, 2019). I programmi di intelligenza artificiale, in questo senso, non è solo strumento di controllo dall'alto, ma parte di un rituale di interazione in cui chi è osservato si trasforma, in parte, come co-autore della sorveglianza (Lyon, 2020).

Gli attori sociali, consapevoli di essere osservati, modellano i propri atteggiamenti e modificano ruoli e identità per gestire le interazioni. Il pubblico di queste “opere teatrali virtuali” non è solo umano, ma anche algoritmico (Goffman, 1997), l'inclusione o l'esclusione da determinate piattaforme dipende da valutazioni automatiche e invisibili. Un esempio emblematico è rappresentato dai sistemi che gestiscono le ordinazioni dei rider, in cui anche un singolo rifiuto da parte del lavoratore può determinare l'esclusione dagli incarichi futuri assegnati dall'algoritmo.

2. L'IA nella società, come infrastruttura o struttura sociale

Nella società attuale, l'IA non è più un prototipo relegato in laboratori di ricerca, il suo utilizzo non è più limitato a piccoli impieghi sperimentali, ma è diventata una infrastruttura socio-tecnologica che permea la vita degli attori sociali. Dalla sanità alla finanza, dalla sessualità agli scenari di guerra gli algoritmi impattano radicalmente l'agire delle persone, determinando promesse di efficienza e innovazione, ma definendo anche nuovi rischi di disuguaglianze,

dipendenza e controllo. Gli algoritmi diventano dei veri attori sociali che ridefiniscono ruoli, identità e rapporti di potere (Latour, 1992), possiamo leggerla come il nuovo paradigma del capitalismo della sorveglianza, in cui i dati degli individui, comportamenti, bisogni e desideri, diventano materia prima per previsioni commerciali e sociali, le grandi piattaforme digitali hanno trasformato la sorveglianza in un modello economico: non si tratta più soltanto di osservare, ma di anticipare e condizionare il comportamento futuro (Zuboff, 2019).

Uno degli ambiti in cui l'IA ha avuto sviluppi più rapidi è il sistema sanitario, soprattutto dopo la pandemia COVID 19, infatti i nuovi modelli di medicina assistita, con robot e algoritmi di controllo dei dati clinici, viene utilizzato sempre più spesso. I sistemi di machine learning vengono impiegati per la diagnosi di tumori o per l'analisi di immagini radiologiche oppure per la previsione del rischio cardiovascolare (Topol, 2019; Jiang et al., 2017). L'obiettivo è migliorare la qualità delle cure e ridurre i costi dei sistemi sanitari, tuttavia, questa dimensione di "cura algoritmica" si accompagna a un aumento della sorveglianza sanitaria. I dispositivi indossabili, le app che monitorano i livelli cardiaci o di pressione dei pazienti e i big data medici trasformano, i corpi delle persone, in flussi di dati costantemente osservabili (Federspiel et al., 2023). La salute diventa controllabile, i dati raccolti vengono registrati e archiviati, immagazzinati in grandi banche dati e spesso condivisi con attori privati. Qui affiorano alcuni interrogativi determinanti su chi possiede queste informazioni e come, successivamente vengono utilizzate, perciò, la cultura della sorveglianza trasforma la prevenzione sanitaria in un modello in cui la protezione della vita si collega con logiche consumistiche e di mercato, che possono provocare implicazioni rilevanti per la privacy e per le disparità nell'accesso alle cure sanitarie (Lyon, 2020).

Un altro ambito singolare e controverso è quello della sessualità e dell'intimità, la diffusione di robot sessuali e sistemi di intelligenza artificiale progettati per interagire emotivamente (Danaher & McArthur, 2017; Devlin, 2018) ridefinisce il concetto di relazione, in cui l'IA diventa partner, confidente e specchio emotivo.

Parallelamente, le app di incontri fanno largo uso di algoritmi predittivi che consigliano compatibilità, selezionano profili e influenzano gli atteggiamenti degli agenti sociali. La sessualità diventa così sempre più "un sistema di affinità computerizzata", preferenze, opzioni e desideri sono trasformati in dati, analizzati e modificati in valore commerciale.

Gli individui gestiscono la propria "vetrina" su piattaforme che funzionano come palcoscenici permanenti (Goffman, 1997), ogni utente possiede un ruolo e una identità digitale, spesso difforme da quella della vita reale, ma, a differenza della scena sociale tradizionale, qui il pubblico non è composto solo da potenziali partner, bensì da algoritmi che classificano e ordinano i ruoli e le identità, questa trasformazione produce una cultura della sorveglianza in cui il desiderio è osservato, previsto e indirizzato (Lyon, 2020).

Il campo militare è uno dei settori più avanzati nell'uso dell'IA, sistemi di droni capaci di riconoscere e colpire obiettivi, modelli di sorveglianza satellitare, armi autonome letali, rappresentano l'avanguardia e le nuove frontiere delle guerre moderne, producendo riflessioni etiche sull'utilizzo di questi nuovi sistemi nei conflitti armati, che spesso non hanno bisogno del controllo umano (Singer, 2009; Sharkey, 2018), i combattimenti si svolgono tra algoritmi come nei War Games.

La guerra tecnologica riproduce una logica di sorveglianza totale, fondata sulla capacità di identificare il nemico in tempo reale, prevedendone le mosse e rispondendo in modo rapido e automatizzato, riducendo al minimo l'intervento umano. Le decisioni strategiche e operative non sono più esclusivamente nelle mani delle persone, ma sempre più delegate a sistemi algoritmici che arrivano a stabilire, in modo semi-autonomo, la vita o la morte degli individui coinvolti. In questo quadro, la sovranità tecnologica assume i tratti di un "potere liquido", un dominio che non si esercita soltanto attraverso la forza fisica, ma soprattutto tramite il controllo invisibile dei flussi informativi e delle reti di sorveglianza.

Gli individui operano rituali di interazione che non sono passivi, cercano di gestire comportamenti e impressioni, adattano la relazione con modelli di comunicazione verbale e non verbale, fino agli scambi "faccia a faccia", adattando corpi e sguardi (Goffman, 1988). Oggi lo "sguardo" non è solo prerogativa degli esseri umani, ma anche dei sistemi algoritmici, che però risultano invisibili, perciò, la scena sociale è ormai una piattaforma digitale e gli attori devono modulare la propria presenza non solo per gli altri, ma anche per le macchine che li osservano.

La sorveglianza militare, lungi dall'essere un ambito separato e confinato al solo scenario bellico, tende a saldarsi con il controllo civile. Le tecnologie sviluppate per i conflitti armati vengono progressivamente riadattate per la sicurezza interna dei cittadini (O'Neil, 2016), dando vita a un'osmosi tra logiche militari e sociali. La frontiera tra difesa esterna e controllo interno si dissolve, producendo nuove forme di "società della sorveglianza", in cui gli strumenti nati per la guerra diventano pratiche ordinarie di monitoraggio e governo della popolazione (Lyon, 2020).

La storia recente dimostra che molti sistemi oggi di uso comune hanno avuto origine in ambiti militari, per poi essere convertiti in strumenti civili, le innovazioni belliche hanno rappresentato i principali motori di trasformazione della società contemporanea, dal radar al GPS all'energia atomica, fino agli odierni sistemi di intelligenza artificiale applicati alla sicurezza urbana. Quello che emerge è una continuità strutturale tra guerra e pace, in cui la tecnologia diventa il veicolo privilegiato per esercitare un controllo, normalizzando nella vita quotidiana delle logiche originariamente pensate per il conflitto (Hobsbawm, 2014).

Nonostante la difformità dei campi applicativi, sanitari, sessualità, guerra, l'elemento che accomuna questi utilizzi dell'IA è la loro funzione di controllo e sorveglianza. I modelli

algoritmici osservano, raccolgono e interpretano dati con lo scopo di controllare gli agenti sociali, questa sorveglianza è normalizzata e diffusa, non è percepita come imposizione esterna ma come parte integrante della vita quotidiana, normalizzando e naturalizzando il dominio (Bourdieu, 1998). Lo potremmo definire un “sistema capitalistico della sorveglianza”, un modello economico in cui i comportamenti e le azioni individuali e di gruppo, vengono trasformati in risorse da monetizzare. La novità più pervasiva dell’IA si fonda su una sorveglianza che non si limita a registrare il passato, ma è orientata al futuro, l’obiettivo è dirigere le azioni sociali, i desideri delle persone, in sintesi un “consumismo digitale indirizzato” (Zuboff, 2019).

3. IA tra sorveglianza, prevenzione e repressione del crimine

L’introduzione delle tecnologie di Intelligenza Artificiale nell’ambito della sicurezza e della prevenzione del crimine rappresenta una delle innovazioni maggiori nei modelli di repressione della criminalità. L’applicazione di strumenti algoritmici ai processi di sorveglianza, raccolta dati, monitoraggio degli spazi pubblici e previsione dei comportamenti criminali modifica profondamente non solo le strategie di repressione e prevenzione, ma anche le logiche della governance sociale. La promessa è quella di una maggiore efficienza nella lotta al crimine, grazie alla capacità dei sistemi predittivi, alla rapidità di elaborazione dei dati e alla possibilità di sorvegliare in modo capillare contesti prima opachi. Tuttavia, questa promessa coesiste con criticità rilevanti, il rafforzamento di forme di controllo e sorveglianza sociale produce e rafforza nuove forme di disuguaglianze attraverso i Bias algoritmici (Pronzaro, 2024). Si sviluppa la normalizzazione di uno Stato di sorveglianza permanente, e del controllo sistematico di gruppi sociali liminari, oppure modelli di razzializzazione della delinquenza.

I sistemi di controllo non operano solo attraverso la forza, ma attraverso la visibilità, oppure modellano i sistemi di sorveglianza, l’IA con la sua capacità di osservare, registrare e predire, spesso in modo anonimo o invisibile, rappresenta una nuova fase di questa logica panottica, amplificandola e radicalizzandola (Foucault, 1975).

Uno degli ambiti più discussi è quello dei sistemi di polizia predittiva, ovvero l’uso di algoritmi in grado di analizzare enormi quantità di dati, denunce, statistiche di criminalità, dati socioeconomici, flussi di mobilità urbana, per individuare aree o individui potenzialmente a rischio di comportamenti criminali. Sistemi come PredPol negli Stati Uniti o XLAW in Italia si propongono di “anticipare” il crimine, orientando le pattuglie in zone considerate sensibili, individuando degli hot-spot (Block, 1995), cioè delle aree ad alta concentrazione criminale. Questi sistemi creano delle crime mapping, territori sensibili all’attività delinquenziale che devono essere monitorati, ma mentre i sistemi PredPol si concentrano sugli autori di reato (Kaufmann, 2019) il

sistema italiano XLAW classifica zone e reati, creando dei territori adatti al crimine, le “zone di caccia” (Lombardo, 2019). Queste aree urbane diventano luoghi privilegiati per alcuni tipi di reati, di solito quelli predatori, cioè furti e rapine. L'offender si sente al sicuro in questi territori di caccia, perché conosce il luogo e, cosa più importante, le vie di fuga. Spesso in queste aree usufruisce dell'appoggio di persone che gli garantiscono sicurezza e aiuto in caso dell'arrivo delle forze dell'ordine. La grande rivoluzione di questo programma è la sua capacità di concentrarsi e tracciare luoghi e reati (Brantingham et al., 2017) ma non immagazzina dati sugli offender, in questo modo, non raccogliendo dati sensibili sulle persone risulta un vero modello etico di repressione, e soprattutto prevenzione del crimine. Creando queste mappe il sistema produce degli alert orari che vengono inviati alle stazioni di polizia, queste indirizzano le pattuglie in vie precise a orari prefissati, nelle città dove il sistema è stato utilizzato si è avuta una sensibile diminuzione dei crimini predatori (Lombardo, 2019).¹

Queste tecniche si fondano su un presupposto di razionalità utilitaristica, si cerca di prevenire il reato intervenendo sulla routine (Cohen & Felson, 1979), riducendo le opportunità e aumentando i rischi percepiti dal potenziale offender. Tuttavia, tali sistemi non sono neutrali, gli algoritmi riflettono i dati con cui sono stati inseriti nei sistemi di catalogazione, oppure da come vengono interpretati dagli operatori di polizia e finiscono, spesso, per riprodurre pregiudizi strutturali, aree povere o a maggioranza di minoranze etniche diventano sistematicamente “sorvegliate”, rafforzando dinamiche di esclusione e stigmatizzazione.

L'IA viene integrata nei sistemi di videosorveglianza urbana con telecamere intelligenti in grado di riconoscere volti, analizzare movimenti sospetti, rilevare assembramenti o comportamenti anomali. Le cosiddette smart cities implementano tecniche di monitoraggio basate su reti di telecamere e sensori di movimento, allo scopo di garantire sicurezza e ordine pubblico.

La sorveglianza algoritmica tende così a diventare “totale”, diffondendosi negli spazi pubblici, sui trasporti, nei luoghi di lavoro e persino nelle scuole. Questo fenomeno spinge a riflettere su una nuova forma di “biopolitica algoritmica” (Foucault, 2015) dove la vita sociale è costantemente e perennemente osservata e regolata da sistemi automatici (Bowers et al., 2004).

Uno degli utilizzi della PredPol è il riconoscimento facciale, applicato per identificare sospetti, monitorare eventi di massa o rintracciare individui segnalati, tecniche biometriche estese includono anche l'analisi delle impronte digitali, della voce, della postura e persino della cinesica e prossemica (Fussey et al., 2021). Questi strumenti vengono adoperati non solo a scopi investigativi, ma anche come forme di controllo preventivo, sistemi di IA capaci di rilevare “comportamenti anomali” in luoghi sensibili come stazioni ferroviarie o aeroporti, attivano

¹ Per maggiori informazioni sulla sperimentazione del programma XLAW si consiglia di visitare il sito web: <https://www.xlaw.it/presentazione/>

automaticamente procedure di sicurezza. L'aspetto problematico riguarda la possibilità che questi sistemi siano inesatti oppure che generino rischi di falsi positivi, uno dei rischi maggiori è quello di creare Bias algoritmici, producendo dei modelli di razzializzazione e discriminazione preesistenti (Pronzato, 2024).

Se il Panopticon, come dispositivo architettonico, era in grado di esercitare potere diretto, attraverso la visibilità asimmetrica (Foucault, 1975), in cui pochi sorvegliavano molti, con l'IA assistiamo a una trasformazione radicale, i cittadini non sono soltanto visibili, ma vengono decifrati, classificati e ordinati da sistemi automatici che li traducono in informazioni. Questi dati vengono poi utilizzati per esercitare un dominio politico ed economico, giustificando tale controllo con la retorica della sicurezza dei cittadini. I governi legittimano l'uso dell'IA, costruendo delle narrazioni che normalizzano e giustifica la sorveglianza come necessaria per la protezione collettiva. I sistemi di potere neutralizzano la restrizione delle libertà personali e individuali attraverso resoconti che minimizzano le conseguenze negative e proiettano la responsabilità altrove (Matza & Sykes, 1957; 1961).

L'uso dell'IA nella sorveglianza solleva, però, questioni cruciali sul rispetto della privacy e della libertà individuale, l'introduzione di sistemi di riconoscimento facciale, di uso di telecamere è stata contestata da associazioni per i diritti civili, riconoscendo il rischio di derive autoritarie (Ferguson, 2017).

Gli algoritmi riflettono e rispecchiano le informazioni che vengono programmate, se queste contengono pregiudizi storici, le decisioni algoritmiche li amplieranno, il rischio è di istituzionalizzare forme di discriminazione sistemica mascherate da neutralità tecnologica, ci si illude che L'IA sia oggettiva e neutra mentre legittima le scelte di altri producendo decisioni con profonde implicazioni sociali.

4. Il Sesto Potere, la sorveglianza nella società che consuma

Il potere nella modernità è stato tradizionalmente distinto in tre sfere principali: legislativo, esecutivo e giudiziario, con l'aggiunta successiva del potere mediatico, spesso definito "quarto potere", e della finanza globale, che potremmo chiamare "quinto potere". Tuttavia, un nuovo e pervasivo dispositivo di dominio si è imposto, la sorveglianza come "sesto potere". Non si tratta più soltanto di un apparato di controllo gestito dai governi nazionali, ma di un sistema diffuso e radicato che permea l'intera società consumistica (Bauman & Lyon, 2013).

Questo "sesto potere" è strettamente legato alla logica del capitalismo contemporaneo e del modello di consumismo digitale: se il capitalismo industriale, di tipo fordista, si nutriva di forza-lavoro e quello finanziario di titoli e derivati, il capitalismo della sorveglianza si alimenta

attraverso la raccolta di dati personali. Informazioni sulle preferenze, sui comportamenti, sulle emozioni e persino sulle relazioni sociali, attraverso i social, gli acquisti sul web o tramite le app di incontri, diventano risorse preziose da estrarre, registrare, manipolare, utilizzare e, infine, monetizzare (Zuboff, 2019).

Non si tratta più soltanto di garantire e gestire la sicurezza degli agenti sociali, come spiegato nei paragrafi precedenti, legati alla prevenzione e repressione del crimine, oppure nel determinare una superiorità tattica nei conflitti bellici, ma di produrre, orientare o, meglio, condizionare i desideri stessi dei consumatori (Han, 2017) trasformando la sorveglianza in una nuova forma di potere, per influenzare gli attori sociali nella loro vita economica e sociale.

Nella società consumistica non è sufficiente essere “liberi cittadini”, per esistere pienamente occorre essere consumatori attivi, chi non aderisce e partecipa al circuito dei consumi, chi non lascia tracce digitali sui social o facendo acquisti, rischia l'esclusione e perciò l'invisibilità sociale. L'individuo è incentivato a mostrare e raccontare volontariamente la propria vita sui social network, a condividere dati con applicazioni commerciali, a fornire informazioni in cambio di sconti o servizi gratuiti. Questa dinamica ricorda il paradosso individuato da Lyon (2020) per cui la sorveglianza non è più solo coercitiva, ma volontaria e desiderata, gli agenti sociali partecipano attivamente al proprio controllo, convinti di ottenere benefici rapidi, senza, però, comprendere che alimentano un sistema di potere, controllo e manipolazione molto più ampio.

Nella sorveglianza consumistica l'uso dei dati personali diventa una merce da archiviare e utilizzare, ogni interazione digitale, un “like”, una ricerca su Google o un acquisto su Amazon generano informazioni che vengono elaborate dai programmi algoritmici capaci di prevedere comportamenti futuri (Srnicek, 2017).

L'IA non si limita a catalogare e a registrare informazioni, ma anticipa e modella i desideri delle persone, sviluppando un marketing predittivo, le piattaforme sono in grado di indicare prodotti prima ancora che l'utente sia pienamente cosciente di desiderarli. Ciò avviene tramite il ricorso a sistemi di profilazione personalizzata, che trasformano la persona in un insieme di pattern comportamentali, trasformando le nostre attività quotidiane in valore economico, il risultato è un modello di capitalismo digitale che non sfrutta esclusivamente la forza-lavoro, ma anche la nostra stessa esistenza quotidiana.

Tuttavia, il “sesto potere” va oltre, oggi assistiamo a una sorveglianza reticolare e universale, dove tutti osservano tutti, mediati da piattaforme che registrano, archiviano e commercializzano ogni sguardo. L'effetto è duplice, gli individui interiorizzano il controllo e plasmano il proprio comportamento in funzione della visibilità sociale, modificando ruoli e identità, mentre le corporation tecnologiche immagazzinano un potere inedito, capace di modificare, gestire o influenzare mercati, elezioni, orientamenti culturali e sociali. Oggi anche le interazioni sociali

avvengono su piattaforme digitali costantemente monitorate, i rituali di interazione “faccia a faccia” avvengono mediati da uno schermo digitale (Goffman, 1988). Anche sul web tutti recitiamo un ruolo, la sorveglianza consumistica trasforma il “palcoscenico” virtuale in uno spazio algoritmico (Goffman, 1997), dove gli “spettatori” non sono solo altri individui, ma anche, o soprattutto macchine che raccolgono dati e interpretano segnali. Ogni like, ogni foto condivisa, ogni recensione diventa parte di una drammaturgia digitale destinata a un pubblico invisibile, composto da multinazionali, governi e sistemi di IA. La sorveglianza non è soltanto coercizione, ma diventa una interazione sociale, un processo in cui persone, algoritmi e istituzioni co-producono nuove forme di potere e nuove dinamiche identitarie.

Un aspetto centrale del “sesto potere” è la sua capacità di normalizzare e naturalizzare la sorveglianza. nel passato i sistemi di controllo erano percepiti come invasivi o eccezionali, oggi sono inserite nella vita quotidiana (Zuboff, 2019; Lyon, 2020). Aprire un profilo social, servirsi di una carta di credito, accedere a una piattaforma streaming, sono tutte attività che implicano l'accettazione, più o meno consapevole, di essere monitorati e che i nostri dati siano gestiti, archiviati e utilizzati.

La logica è quella dello scambio, dati in cambio di servizi, questo baratto sociale e commerciale apparentemente innocuo produce, però, una crescente dipendenza dalle piattaforme e una riduzione degli spazi di autonomia individuale e personale, la libertà di scelta delle persone è, in realtà, profondamente condizionata dalle architetture algoritmiche che orientano le nostre decisioni.

Ma il sesto potere non ha solo implicazioni di tipo economico, ma anche politiche, le stesse tecniche di profilazione utilizzate per il marketing vengono applicate per controllare, verificare oppure orientare il voto, influenzando le opinioni pubbliche, diffondendo disinformazione mirata.

La società consumistica, fondata sulla raccolta e l'utilizzo dei dati, diventa così un campo dove la distinzione tra potere, sorveglianza e manipolazione economica e potere politico si dissolve (Zambonelli, 2020) entrambi i sistemi, politico ed economico, utilizzano gli stessi strumenti algoritmici per modificare le opinioni degli agenti sociali.

5. Conclusioni: verso una società del controllo digitale – possibilità, rischi e dilemmi etici

In questo saggio abbiamo analizzato alcune delle molteplici dimensioni attraverso le quali l'IA si intreccia, pervade e invade la nostra vita sociale contemporanea. Dai conflitti armati, all'uso medico, dalla sessualità alla sorveglianza, dalla prevenzione e repressione del crimine fino alla nuova frontiera del capitalismo digitale. Abbiamo esplorato come l'IA non sia soltanto un

insieme di tecnologie, programmi algoritmici e sistemi automatizzati, ma un vero modello organizzativo del potere che produce profitto attraverso la sorveglianza continua degli agenti sociali, capace, anche, di riconfigurare relazioni sociali, identità, ruoli, economie, culture e perfino l'immaginario collettivo.

Abbiamo trattato degli scenari di guerra, dove l'IA alimenta nuovi modelli e strategie militari fondate su droni autonomi (Manhas, 2023), sistemi predittivi e armi autonome, in definitiva su una cyber-sorveglianza globale. Abbiamo affrontato la medicina automatizzata, in cui la promessa di prognosi precise e terapie personalizzate possono produrre il rischio di modificare il paziente in dati e informazioni da analizzare, monetizzare e gestire. Abbiamo toccato la sessualità, con la diffusione di robot e interfacce affettive che trasformano e riscrivono i modelli di intimità, producendo domande sulla natura delle relazioni affettive degli esseri umani (Levy, 2009).

Infine, al centro della nostra analisi abbiamo trattato la questione cruciale della sorveglianza, un vero “sesto potere” (Bauman & Lyon 2013), un potere che non osserva soltanto, ma che plasma e definisce i comportamenti, modella i desideri, orienta le ambizioni, impone i consumi e determina chi può o non può partecipare pienamente alla vita sociale, soprattutto quella digitale.

Uno dei risultati emersi è che la sorveglianza contemporanea non è limitabile a una dimensione di tipo repressiva o securitaria, come ha mostrato anche il paragrafo dedicato alla prevenzione e repressione del crimine attraverso le crime mapping e gli algoritmi predittivi (Kaufmann et al., 2019), la logica della sorveglianza si è trasformata in infrastruttura sociale che gestisce e modifica la vita delle persone.

Non si tratta più solo di controllare gli individui oppure di monitorare devianze, ma di anticipare e orientare comportamenti, la sorveglianza diviene una forma di ingegneria o, meglio, la costruzione sociale dell'agire quotidiano che opera sia a livello politico, attraverso la gestione dei sistemi di sicurezza, sia a livello economico, per indirizzare consumi e guidare i desideri.

I frame di costruzione di un sistema culturale codificato (Goffman, 2001), attraverso modelli di interazione sociale non sono più soltanto scenari interpretativi costruiti tra diversi agenti sociali, ma sono oggi ininterrottamente definiti e ridefiniti dagli algoritmi che organizzano la visibilità, l'accesso alle risorse, le opportunità tra gli attori sociali, di poter interagire. La “presentazione del sé” non avviene più solo sul palcoscenico sociale (Goffman, 1988), ma su piattaforme digitali gestite da logiche consumistiche e capitalistiche che possiedono e gestiscono i programmi di IA, e che trasformano ogni interazione in dato utile per il controllo e per un profitto (Locatelli & Vittadini, 2023).

Un filo rosso che ha attraversato tutti i capitoli è l'ambivalenza dell'IA, essa promette benefici enormi, la riduzione dei costi sanitari e, grazie alle cure tempestive, l'aumento delle aspettative di

vita delle persone, migliorare la sicurezza dei cittadini con strumenti predittivi del crimine, offrire esperienze personalizzate nelle relazioni affettive, eppure i rischi sono rilevanti.

I sistemi di IA rendono i conflitti armati più disumani, se mai una guerra sia stata “umana”, ma il rischio di disumanizzazione lo possiamo trovare anche nei sistemi di cura digitalizzati, oppure nella sessualità artificiale (Bryson, 2010), infine i modelli di sorveglianza consumistica producono nuove forme di disuguaglianze, concentrando il potere di raccogliere i dati nei server di poche piattaforme digitali.

Una delle attività più problematiche dell’IA riguarda l’effetto del “sesto potere” sulla democrazia, le società democratiche fondano la libertà delle persone su principi di trasparenza, partecipazione e uguaglianza (Soro, 2019). Tuttavia, l’infrastruttura digitale della sorveglianza opacizza i modelli decisionali, consegnando agli algoritmi scelte cruciali, dalla concessione di un prestito per un mutuo all’assunzione a un lavoro, fino ai modelli predittivi del rischio criminale.

Siamo entrati in una fase in cui i dati non servono solo a riprodurre la realtà, ma a generarla e governarla, questa nuova condizione rende sempre più difficoltoso identificare tra scelta autonoma e comportamento indotto, diventa difficoltoso comprendere quando siamo liberi oppure condizionati da un algoritmo (Zuboff, 2019).

La democrazia e gli Stati democratici, perciò, rischiano di mutare in tecnocrazie invisibili, dove il potere viene esercitato mediante interfacce apparentemente neutre ma profondamente orientate e indirizzate da interessi economici e politici (Razzante, 2024).

Il controllo algoritmico non si limita a osservare, ma crea nuovi rituali di interazione, gli attori sociali coscienti di essere ininterrottamente monitorati, modulano i propri atteggiamenti e comportamenti in funzione della visibilità digitale, una forma di controllo interiorizzato, normalizzato e naturalizzato (Bourdieu, 1998), in cui non serve più un sorvegliante esterno, perché la persona stessa diventa responsabile del proprio controllo, condividendo foto, permettendo tracciamenti sul web, consegnando volontariamente la propria vita e i propri dati ai sistemi di IA.

Un altro argomento cruciale riguarda l’impatto dell’IA sulla qualità delle relazioni umane, la possibilità di interagire con robot sessuali, oppure con assistenti virtuali e avatar digitali dischiude scenari che mettono in discussione la differenziazione tra umano e artificiale (Parsakia & Rostami, 2023). Queste tecnologie possono diminuire la solitudine e garantire un sostegno psicologico, ma rischiano di assottigliare le competenze relazionali, abituando gli agenti sociali a rapporti basati sull’obbedienza algoritmica (Stardust, 2024), il legame tra individui viene gradualmente surrogato da interazioni con sistemi artificiali.

È chiaro che dobbiamo iniziare una riflessione etica su chi è il responsabile dei programmi di algoritmo, solo gli ingegneri informatici, oppure le grandi società che commercializzano (O’Neil,

2016), gli Stati che cercano di regolamentare questi sistemi automatizzati, oppure gli stessi utenti che li utilizzano in modo razionale?

Forse non esiste una coscienza collettiva sulle manipolazioni che i gruppi commerciali, o gli Stati nazione, impongono grazie all'IA, ma solo resistenze individuali, che però si scontrano con il potere della sorveglianza globale, evidentemente gli utenti devono avere più spazio decisionale nelle autorizzazioni di raccolta di dati da parte dei programmi algoritmici. Forse la sfida è proprio questa, riuscire a poter incidere sulle scelte e sul potere di sorveglianza di questi sistemi automatizzati (Pollicino & Dunn, 2024).

Il titolo scelto per questo lavoro, *Specula*, rimanda all'idea di un osservatorio ma anche di una torre di guardia: è infatti urgente sviluppare uno sguardo critico capace di esaminare l'IA nei suoi usi; ma *Specula* richiama, anche, il rischio che la società intera si trasformi in un Panopticon digitale diffuso, dove tutto è visionato, registrato e archiviato. Il termine *Specula* si trasforma così in un osservatorio critico e insieme in una torre di vigilanza, in cui l'IA è uno strumento di conoscenza ma anche il simbolo del pericolo di una sorveglianza totale.

La sorveglianza, oggi, rappresenta l'avanguardia del nuovo modello capitalistico (Zuboff, 2019), è necessario pensare a nuove forme di regolazione e resistenza. La vera posta in gioco non è l'uso della tecnologia in sé, ma la definizione dell'umano e le scelte consapevoli delle persone. L'IA ci costringe, oggi più di ieri, a interrogarci su cosa significhi essere liberi, responsabili, e come relazionarci con gli altri.

Se ci limiteremo a subire in modo passivo la logica della sorveglianza e del consumismo capitalistico digitale, il rischio è quello di ridurre le persone a dati calcolabili o informazioni da raccogliere. Se, invece, sapremo impiegare in modo critico e responsabile le tecnologie, individuando rischi e potenzialità, potremo trasformare l'IA in uno strumento utile. Il futuro non è scritto in un algoritmo, dipenderà dalle decisioni collettive che riusciremo a produrre (Doctorow, 2024), rivendicando la centralità dell'umano nell'epoca delle macchine intelligenti.

Bibliografia

- Balistreri, M. (2023). Le questioni morali e le implicazioni psicologiche della riproduzione, del sesso e delle relazioni affettive nelle missioni spaziali. *Rivista internazionale di Filosofia e Psicologia*, 14(3), 148-167.
- Bauman, Z. (2012). *Modernità liquida*. Laterza.
- Bauman, Z. & Lyon, D. (2013). *Sesto potere. La sorveglianza nella modernità liquida*. Laterza.
- Block, R. & Block C. (1995). Space, place, and crime: Hot-spot areas and hot places of liquor-related crime. *Crime Prevention Studies*, 4, 145-184.
- Bourdieu, P. (1998). *Il dominio maschile*. Feltrinelli.
- Bowers, K.J., Johnson, S.D & Pease, K. (2004). Prospective Hot-Spotting: The Future of Crime Mapping?. *The British Journal of Criminology*, vol. 44, nr. 5, 641-658.
- Brantingham, P.J., Brantingham, P.L. & Andresen, M.A. (2017). The geometry of crime and crime pattern theory. In R. Wortley & M. Townsley (a cura di), *Environmental Criminology and crime analysis* (2nd ed.) (pp. 98-115). Routledge.
- Bryson, J.J. (2010). Robots should be slaves. In Y. Wilks, *Close engagements with artificial companions: Key social, psychological, ethical and design issues* (pp. 63-74). John Benjamins Publishing.
- Cohen, L.E. & Felson, M. (1979). Social change and crime rate trends: a routine activity approach. *American Sociological Review*, vol. 44, 588-608.
- Danaher, J. (2017). Robotic sex and consent: A feminist analysis. In J. Danaher & N. McArthur (a cura di), *Robot sex: Social and ethical implications* (pp. 105-132). MIT Press.
- Devlin, K. (2018). *Turned on: Science, sex and robots*. Bloomsbury Sigma.
- Doctorow, C. (2024). *Come distruggere il capitalismo della sorveglianza*. Mimesis.
- Federspiel, F., Mitchell, R., Asokan, A., Umana, C. & McCoy, D. (2023). Threats by artificial intelligence to human health and human existence. *BMJ Global Health*, 8, 5. doi: 10.1136/bmjgh-2022-010435.
- Ferguson, A.G. (2017). *The rise of big data policing: Surveillance, race, and the future of law enforcement*. NYU Press.
- Foucault, M. (1975). *Sorvegliare e punire: Nascita della prigione*. Einaudi.
- Foucault, M. (2015). *Nascita della biopolitica. Corso al Collège de France (1978-1979)*. Feltrinelli.
- Fussey, P., Davies, B. & Innes, M. (2021). Assisted' facial recognition and the reinvention of suspicion and discretion in digital policing. *The British Journal of Criminology*, vol. 61, nr. 2, 325-344.
- Garcia, D. (2024). Algorithms and Decision-Making in Military Artificial Intelligence. *Global Society*, vol. 38, 24-33.
- Goffman, E. (1988). *Il Rituale di interazione*. Il Mulino.

- Goffman, E. (1997). *La vita quotidiana come rappresentazione*. Il Mulino.
- Goffman, E., (2001). *Frame analysis. L'organizzazione dell'esperienza*. Armando Editore.
- Han, B.C. (2017). *Psicopolitica: Il neoliberismo e le nuove tecniche del potere*. Nottetempo.
- Hobsbawm, E.J. (2014). *Il secolo breve*. Rizzoli.
- Jiang, F., Jiang, Y., Zhi, H., Dong, Y., Li, H., Ma, S. & Wang, Y. (2017). Artificial intelligence in healthcare: Past, present and future. *Stroke and Vascular Neurology*, 2(4), 230-243.
- Kaufmann, M., Egbert, S. & Leese, M. (2019) Predictive Policing and the Politics of Patterns. *The British Journal of Criminology*, vol. 59, nr. 3, 674-692.
- Latour, B. (1992). Where are the missing masses? The sociology of a few mundane artifacts. In W. Bijker & J. Law (a cura di), *Shaping technology/ building society* (pp. 225-258). MIT Press.
- Levy, D. (2009). *Love and sex with robots: The evolution of human-robot relationships*. Duckworth Overlook.
- Locatelli, E. & Vittadini, N. (a cura di) (2023). *Digital media. Piattaforme algoritmiche e società*. Vita e Pensiero.
- Lombardo, E. (2019). *Sicurezza 4P - Lo studio alla base di XLAW per prevedere e prevenire i crimini predatori*. Mazzanti Editore.
- Lyon, D. (2020). *La cultura della sorveglianza. Perché la società del controllo ci ha reso tutti controllori*. Luiss University Press.
- Manhas, S. (2023). Drone warfare. A Gray area. *International Journal of Electrical, Electronics and Computers*, vol. 8, n. 5, 01-04.
- Matza, D. & Sykes, G. (1961). Juvenile Delinquency and Subterranean Values. *American Sociological Review*, 26 (5), 712-719.
- Matza, D. & Sykes, G. (1957). Techniques of Neutralization: A Theory of Delinquency. *American Sociological Review*, 22 (6), 664-670.
- Memoli, R. & Saporiti, A. (1995). *Disegno della ricerca e analisi dei dati*. Euroma La Goliardica.
- O'Neil, C. (2016). *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*. Crown Publishing Group.
- Oliveros-Aya, C. (2023). Artificial intelligence in drones and robots for war purposes: A biolegal problem. *Janus Net e-journal of International Relation*, vol. 14, n. 2.
- Orwell, G. (2019). *1984*. Mondadori.
- Parsakia, K. & Rostami, M. (2023). Digital Intimacy: How Technology Shapes Friendships and Romantic Relationships. *AI and Tech in Behavioral and Social Sciences*, 1 (1), 27-34.
- Pollicino, O. & Dunn, P. (2024). *Intelligenza artificiale e democrazia. Opportunità e rischi di disinformazione e discriminazione*. Bocconi Press.
- Pronzato, R. (2024). *Algoritmi, strutture e agire sociale. Un'analisi sociologica*. Franco Angeli.

- Ratcliffe, J.H. (2010). Crime mapping: Spatial and temporal challenges. In A. Piquero & D. Weisburd (a cura di), *Handbook of quantitative criminology* (pp. 5-24). Springer.
- Razzante, R. (2024). *Il governo dell'Intelligenza Artificiale*. Cacucci.
- Soro, A. (2019). *Democrazia e potere dei dati. Libertà, algoritmi, umanesimo digitale*. Baldini e Castoldi.
- Srnicek, N. (2017). *Capitalismo digitale: Google, Facebook, Amazon e la nuova economia del web*. Luiss University Press.
- Stardust, Z. (2024). Sex tech in an age of surveillance capitalism. Design, data and governance. *Routledge Handbook of Sexuality, Gender, Health and Rights* (pp. 448-458). Routledge.
- Topol, E. (2019). *Deep medicine: How artificial intelligence can make healthcare human again*. Basic Books.
- Zambonelli, F. (2020). *Algocrazia. Il governo degli algoritmi e dell'intelligenza artificiale*. Scienza Express.
- Zuboff, S. (2019). *Il capitalismo della sorveglianza. Il futuro dell'umanità nell'era dei nuovi poteri*. Luiss University Press.

Tendenze e dibattiti

Le Nuove Frontiere dell'Innovazione Tecnologica in Sanità: l'IA. in aiuto delle persone anziane con patologie cronico-degenerative, ad esempio i malati di Parkinson, tra opportunità e rischi.

di Barbara Baccarini*

ABSTRACT (ITA)

In relazione all'attuale incremento della popolazione anziana e dei pazienti cronici, in un'ottica di LTC, è necessario offrire loro nuove opportunità, grazie all'IA e alle nuove tecnologie. Ad esempio, i software di IA applicati al *problem solving* in medicina, aiutano i pazienti a migliorare la qualità di vita. Il paper pone enfasi sul fatto che, grazie alla medicina predittiva e all'IA, senza giungere a un "soluzionismo tecnologico" è possibile avvalersi di notevoli opportunità. Al contempo è necessario valutare attentamente i rischi che possono mettere in pericolo la salute e la vita delle persone, ad esempio l'errore diagnostico, in quanto gli strumenti di IA devono essere usati per la promozione della salute.

Parole chiave: IA, salute, innovazione tecnologica, sanità, LTC

The New Borders of Technological Innovation in Health Care: Artificial Intelligence to Help Elderly People with Chronic Degenerative Pathologies, For Example Parkinson's Patients Between Opportunities and Risks

by Barbara Baccarini

ABSTRACT (ENG)

In relation to the current increase in the elderly population and chronic patients, from a LTC perspective, it is necessary to offer them new opportunities, thanks to AI and new technologies. For example, AI software applied to problem solving in medicine helps patients improve their quality of life. The paper places emphasis on the fact that, thanks to predictive medicine and AI, without reaching a "technological solutionism" it is possible to take advantage of notable opportunities. At the same time, however, it is necessary to carefully evaluate the risks, which can endanger people's health, such as diagnostic errors, as AI tools must be used for health promotion.

Keywords: AI, health, technological innovation, healthcare, LTC

* Università eCampus

1. La rilevanza delle nuove tecnologie e dell'intelligenza artificiale in Sanità in aiuto delle persone anziane con patologie cronico-degenerative

In riferimento all'attuale contesto sociale e al mutamento del quadro demografico, che ha visto crescere ulteriormente la fascia di popolazione anziana in condizione di cronicità, è doveroso riflettere sulla necessità di creare percorsi di assistenza continuativa.

In tal modo si offre l'opportunità, in un'ottica di *long term care* (LTC), di convivere più a lungo, ma più serenamente, con malattie cronico degenerative, spesso associate in maniera plurima (comorbilità) (Marengoni et al., 2011) tra cui la Malattia di Parkinson (MP), che interessa la persona a livello olistico e che, secondo le rilevazioni statistiche, aumenterà nel contesto nazionale e internazionale fino a raddoppiare nel 2030 (Busse et al., 2010; Dorsey et al., 2007).

Uno dei punti prioritari del programma di Salute Pubblica dell'Ufficio Europeo dell'OMS "Health 2020: the European Policy for Health and Well Being" è proprio quello di mettere al centro la cura delle patologie croniche.

Questo induce a prendere atto della necessità di una presa in carico continuativa non solo da parte del sistema sanitario, ma anche da parte del sistema sociale e delle reti primarie in quanto per la malattia cronica non si può parlare di guarigione, ma di una cura che si protrae nel tempo. (Baccarini, 2020).

Pertanto, il problematico iter di convivenza del malato con la patologia e l'adattamento a una difficile quotidianità fa riflettere sull'importanza di aiutare il paziente a mantenere il proprio modo di vivere e il luogo dell'abitare.

Ne consegue che risulta altrettanto importante per la persona fragile e vulnerabile avvalersi di supporti tecnologici, che possano aiutarla e consentirle di monitorare le proprie condizioni di salute, restando al proprio domicilio.

Tra gli esempi dell'Information and Communication Technology (ICT) in Sanità, un settore di forte investimento è quello dell'e-care, tramite il quale la tecnologia viene utilizzata al fine di fornire un supporto durante il percorso assistenziale.

Ne sono esempio la telemedicina, la teleassistenza, che consentono al paziente di essere assistito a distanza permettendogli di restare più serenamente al proprio domicilio vicino al caregiver e alla famiglia.

Quanto sopra evidenzia che, tramite nuovi strumenti tecnologici e grazie all'I.A. e alle Assistive Technologies, i pazienti affetti da malattie cronico degenerative, come, ad esempio, la M.P., possono essere monitorati a distanza, affrontando con minori disagi la malattia.

È necessario prendere atto del fatto che l'evoluzione di strumenti e applicativi legati all'ICT, ha ormai attribuito una fisionomia digitale all'attuale società e tutto ciò si riflette pure sul sistema sanitario come si deduce da lemmi e neologismi.

Si evince in Lupton (2014, pp. 1344-1345): “Quando si parla di sanità digitale, in senso lato, si intende individuare una gamma di strumenti e software veramente vasta e [...] capace di toccare, plasmare e incidere su ogni aspetto del sistema salute”.

Per quanto concerne poi la mobile health (*m-health*), l'OMS la definisce “un servizio in ambito medico e di Sanità Pubblica supportato da apparecchi mobili quali smartphone, tablet, strumenti per il monitoraggio dei pazienti a distanza e altri strumenti wireless” (WHO, 2011, p. 6).

Un suo elemento centrale è costituito da software installabili sui diversi dispositivi mobili e definiti app, in quanto possono permettere a questi strumenti di ottenere il loro massimo potenziale connettendo questi con altri *device* o sensori della *wearable technology* tra cui anche apparecchiature wireless per la telemedicina (European Commission, 2014, p. 3).

Quanto sopra preso in esame costituisce ovviamente una parte, anche se importante, delle tecnologie nell'ambito dell'e-health e m-health.

2. La nuova realtà: l'apporto dell'I.A. e delle Assistive Technologies nel contesto sanitario tra opportunità e rischi

Sempre nel contesto della I.A., oltre alle app per diagnosi e autodiagnosi, rivolte sia a professionisti, sia al pubblico, esistono alcune applicazioni di I.A. in campo di *problem solving* clinico (si fa riferimento a realtà quali Isabel, IBM Watson o Illumeo, già in essere e adottate).

Isabel è stato creato dal finanziere londinese Jason Maude il quale, avendo sperimentato su sua figlia Isabel un grave errore medico, prese la decisione di creare un software diagnostico, che richiede al medico di inserire i sintomi rilevati durante l'anamnesi e su questa base produce un elenco probabile di diagnosi.

IBM Watson (si veda Susskind & Susskind 2015) che è presente quale medico di corsia presso alcuni ospedali statunitensi, riesce a memorizzare milioni di referti medici di malati di tumore e un suo utilizzo concreto è il seguente: “Watson [...] potrà prevedere fino a tre ore in anticipo una crisi glicemica di un diabetico” (Ardisson, 2018, p. 82).

Il connubio tra Watson con sensori e app assume una grande rilevanza e ha ricadute importanti come l'esemplificazione di un'applicazione di questa tecnologia nella direzione dell'esclusione della presenza del medico, il quale viene totalmente sostituito dalla connessione con I.A. e vari strumenti.

Infine, ILLUMEO, creato dalla Philips, riesce a ricavare alcune informazioni importanti dall'anamnesi di un paziente e confronta una pluralità di immagini per migliorare il processo decisionale clinico.

Per quanto riguarda poi le opportunità e i rischi evidenziati, risultano degne di nota, in merito ad alcune applicazioni di I.A. sopracitate, le opportunità offerte, in aiuto al medico e al paziente

riuscendo a produrre un elenco probabile di diagnosi, al fine di individuare le cure migliori e, al fine di non incorrere nel grave rischio dell'errore diagnostico.

Degna di interesse poi, in aiuto dei pazienti cronici, si rivela la robotica assistiva, che si è sviluppata negli ultimi decenni in Giappone e oggi sta emergendo nel contesto europeo tramite i più recenti programmi di ricerca quale l'*European Strategic Research Agenda for Robotics in Europe* (2014-2020) SPARC (Becchimanzi, 2022).

Attualmente le tecnologie digitali robotiche e indossabili ottengono grandi consensi in diversi contesti (Yang et al., 2018) tra cui l'ambito medico e l'assistenza socio-sanitaria.

Queste assumeranno un ruolo di rilievo in quanto possono implementare le abilità umane utili a migliorare la qualità di vita, ad esempio, di persone con MP, potenziandone la mobilità, il senso di sicurezza e favorendo l'inclusione sociale. (Ancona et al., 2021; Sapci & Sapci, 2019).

Le suddette tecnologie, avvalendosi di una vasta gamma di dispositivi, servizi e sistemi atti a creare una sinergia tra tecnologie all'interno delle abitazioni e gli oggetti che vengono usati quotidianamente, sta mostrando interessanti nuovi sviluppi (Rinaldi et al., 2018).

Da una parte il lavoro dei sanitari può trarre un grande beneficio dall'integrazione delle nuove tecnologie digitali in sistemi interconnessi, dall'altra parte si può prevenire la solitudine e l'isolamento dei malati, grazie a nuovi stimoli emotivi e cognitivi.

Le tecnologie digitali robotiche favoriscono l'autonomia domestica, controllano costantemente lo stato di salute (Iecovich, 2014) ed è innegabile che possano essere un'opzione anche meno gravosa dal punto di vista economico rispetto a un percorso di cura ospedaliero e al contempo supportano anche il percorso diagnostico e terapeutico.

Le *Assistive Technologies* (ATS) sono definite dalla WHO come “qualsiasi prodotto o strumento o attrezzatura tecnologica adottata o appositamente progettata per migliorare le capacità e l'indipendenza di un individuo, per facilitare la partecipazione e migliorare il benessere generale”.

La definizione che l'*Association for the Advancement of Assistive Technology in Europe* (AATE) fornisce di Tecnologia Assistiva è “qualsiasi prodotto o servizio basato sulla tecnologia in grado di facilitare le persone con limitazioni funzionali di ogni età nella vita quotidiana, nel lavoro e nel tempo libero”.

In particolare, risultano di grande interesse le ATS, che aiutano gli utenti con disabilità e gli anziani a superare le barriere infrastrutturali, per consentire loro la piena partecipazione alle attività sociali.

Le tecnologie per la cura e i dispositivi indossabili facilitano il follow up permanente dello stato di salute, delle condizioni cliniche croniche e delle capacità funzionali attraverso il monitoraggio da remoto, la riabilitazione fisica a domicilio, il controllo dell'assunzione/somministrazione di farmaci.

Degna di nota è poi la sperimentazione dell'integrazione dei sistemi robotici o micro robotici all'interno di prodotti di uso quotidiano o arredi come armadi intelligenti in grado di monitorare e/o ricordare l'assunzione di farmaci (Ennis et al., 2017)

Risulta altrettanto degna di nota un'indagine effettuata al fine di indagare l'accettabilità, nei confronti dell'uso di Google Glass quale dispositivo di assistenza da parte di persone con MP e caregiver.

L'indagine ha avuto esiti positivi nei confronti del dispositivo, inoltre il potenziale di assistenti digitali come Alexa (assistente digitale sviluppato da Amazon) o Siri (sviluppato da Apple), in ambito sanitario e assistenziale, risulta ampio.

Si deve anche precisare in merito alle tecnologie Iot (Yang et al., 2018) (l'Iot indica la connessione di dispositivi o prodotti a Internet, inclusi elettrodomestici, apparecchi sanitari, etc): una volta connesso è in grado di archiviare ed elaborare informazioni in rete in modo indipendente, ma è in grado di comunicare anche con altri dispositivi appartenenti alla rete, pertanto le tecnologie Iot possono costituire un valido strumento per i pazienti cronici sia nella telemedicina sia nel monitoraggio a distanza.

Secondo Asakawa (Asakawa et al., 2019) le tecnologie digitali, tra cui i sistemi indossabili e robotici interconnessi, possono migliorare sia la diagnosi sia il trattamento della MP tramite una valutazione obiettiva, anche se esistono studi che evidenziano come le tecnologie assistive spesso non hanno un riscontro nella realtà della vita (Sanders et al., 2012) oppure si teme che vi siano barriere economiche e socio-culturali che possano sfavorirne l'utilizzo.

Le tipologie di assistive robot sono molteplici, ma si può affermare che per le persone con MP i Companion Robot, che integrano IA e capacità di apprendimento, sono molto utili per attenuare il senso di solitudine e il mantenimento del benessere psico-motorio (Odekerken Schoroder et al., 2020).

Anche i SAR, che sono creati per fornire assistenza e supporto agli utenti, sono molto utili per ottenere progressi misurabili nel corso della riabilitazione, convalescenza, apprendimento.

Sia i robot assistivi sia i robot sociali sono stati concepiti come tecnologia di compensazione, cioè utili a compensare un declino fisico, cognitivo e/o psicosociale (Feil-Seifer, 2005).

Un caso rappresentativo è il Robot ElliQ, sviluppato dalla Intuition Robotics, presentato al CES di Las Vegas nel 2015, che è stato progettato proprio per gli anziani e i pazienti cronici, al fine di supportarli nell'essere attivi, autonomi e connessi con gli altri.

Grazie all'I.A. e alla possibilità di connettersi con altri oggetti smart presenti nell'abitazione, ElliQ, tra le sue numerose funzioni, è in grado di avviare video chiamate, ricordare gli appuntamenti, quando è il momento di assumere farmaci; ElliQ può inoltre calmare le persone e chiamare i soccorsi in caso di emergenza, offrire supporto e stimolazione cognitiva.

Da quanto emerso sopra, sono numerose le opportunità offerte dalla I.A. e dalle tecnologie digitali robotiche, al fine di migliorare la qualità di vita dei pazienti cronici, facendoli sentire seguiti e monitorati costantemente, ma, al contempo, più autonomi nello svolgimento delle attività quotidiane.

Anche se le opportunità offerte dalle nuove tecnologie e dall'I.A. sono evidenti, queste devono, però, entrare nelle case dei malati, pertanto, tra i rischi evidenziati, si può sottolineare il fatto che possano risultare invasive, ledendo la privacy e le abitudini del malato. Sempre in merito ai rischi, alcuni pazienti reagiscono positivamente al mediatore robotico, altri invece lo rifiutano, preferendo l'interazione con un essere umano.

3. La Medicina predittiva e l'Intelligenza Artificiale

Inoltre, per quanto concerne la medicina predittiva e il gran numero di dati che vengono prodotti dall'I.A., si sa che questi devono essere governati. Come si evince in Maturo (2024, pp. 134-135): “Sappiamo che ciò è al di là di qualsiasi capacità umana. Di qui la crescente centralità degli algoritmi [...]. Gran parte delle decisioni degli algoritmi si basa sulle previsioni che “essi” sono in grado di formulare”.

Ancora viene precisato, in merito alla medicina predittiva, che il sogno dell'onniscienza della medicina predittiva è alla base delle convinzioni dei “tecno entusiasti” – come vengono definiti dal sociologo bielorusso Evgenij Morozov (2013) –, i quali, in modo acritico, vedono nello sviluppo tecnologico, la risoluzione di qualsiasi problema.

Per quanto concerne poi il contesto medico, secondo il cardiologo americano Eric Topol (2019) l'I.A. solleverà i medici dall'ingente peso del lavoro burocratico.

Inoltre, l'I.A. consentirà di rendere più autonomi i pazienti, i quali, tramite la conoscenza dei dati sull'operato dei medici, potranno anche prendere atto dei centri di cura più efficienti.

Al contempo risulta, però, doveroso prendere atto dei rischi connessi all'utilizzo dei dispositivi tecnologici. Infatti, mette all'attenzione Maturo (2024, pp. 134-136): “Nel caso specifico degli algoritmi, vi è la possibilità che alcune loro procedure riproducano convinzioni non sufficientemente testate o semplicemente sbagliate. Tali basi possono amplificarsi e dare luogo a discriminazioni sociali”.

Sempre per quanto concerne il contesto sanitario e la cura dei pazienti cronici, è degno di nota il caso riguardante i pazienti cronici afroamericani sottocurati in quanto un algoritmo predittivo sistematicamente sottostimava la gravità delle loro condizioni di salute, attribuendo loro un basso risk score.

Infatti si evince: “Questo bias e, di conseguenza, la discriminazione, derivano da una colossale ingenuità. Infatti i programmatori avevano utilizzato, tra le proxy di salute, le spese sanitarie individuali. [...] Disponendo, di redditi bassi [...] gli afroamericani spendevano poco per la loro salute perché poveri, non perché sani” (Maturò 2024, p. 136).

Ne deriva che, come evidenziarono i ricercatori, solo il 18% dei pazienti afroamericani aveva ricevuto le cure necessarie, ma in realtà ben il 47% avrebbe avuto bisogno di cure e pertanto risultò gravemente penalizzato.

Quanto sopra esposto, in riferimento alle opportunità e ai rischi, fa riflettere sul fatto che, nel contesto medico, senza giungere all'accettazione di un “soluzionismo tecnologico”, è indubbio che, grazie alla medicina predittiva e all'I.A., possano emergere notevoli opportunità, quali, ad esempio, sollevare i medici dall'onere del lavoro burocratico e offrendo ai pazienti l'opportunità di divenire più consapevoli e autonomi.

Al contempo viene posto all'attenzione il fatto che uno dei rischi più gravi è proprio quello di incorrere in bias diagnostici (si può fare riferimento, ad esempio, all'errore concernente i pazienti cronici afroamericani sottocurati), che possono mettere in pericolo la salute e la vita delle persone.

Ne consegue che, come ha ulteriormente dimostrato la Pandemia Covid 19 (Baccarini, 2023), risulta prioritario tener presente che la salute è il bene più prezioso per l'umanità a livello globale e locale, un bene che va tutelato al fine di erogare le cure necessarie a tutti e al contempo, per quanto concerne l'I.A. è doveroso preservare la libertà e la criticità del pensiero nel pieno rispetto della salute e della dignità della persona.

Conclusioni

Il lavoro ha focalizzato l'attenzione, nella parte iniziale, sull'attuale impatto del processo di senilizzazione e sul conseguente incremento delle patologie croniche e cronico degenerative quali, ad esempio, la MP. Al contempo è stata messa in rilievo l'importanza delle Nuove Tecnologie in sanità e l'evoluzione di strumenti connessi all'ICT, che hanno dato una fisionomia digitale all'attuale società, la quale, nel contesto medico, è connotata da termini specifici e settoriali, quali, ad esempio, sanità digitale, *e-health*, *m-health*, etc.

In merito all'I.A., oltre alle app per diagnosi ed autodiagnosi, è stato fatto riferimento, ad esempio, ai software di I.A. applicati al problem solving in medicina e alle ATS, sottolineando le opportunità da loro offerte, ma anche i rischi evidenziati, in un'ottica di LTC, nei confronti di una popolazione anziana, che in tal modo può convivere più serenamente con una cronicità a volte stigmatizzante, restando al proprio domicilio.

È stata poi posta particolare enfasi sul fatto che, grazie alla medicina predittiva e alla I.A., senza giungere alla accettazione di un “soluzionismo tecnologico” e di un atteggiamento acritico, è possibile avvalersi di indubbie opportunità nel contesto medico-sanitario, anche se è importante ponderare attentamente i rischi che possono mettere in pericolo la vita e la salute delle persone come è accaduto con i pazienti afroamericani sottocurati.

Bibliografia

Ancona S., Faraci, F.D., Khatab, E., Fiorillo, L., Gnarra, O., Nef, T., Bassetti, C.L.A. & Bargiotas, P. (2021). Wearables in the home-based assessment of abnormal movements in Parkinson's disease: a systematic review of the literature. *Journal of neurology*, 1-11.

Ardissone, A. (2018), *L'uso delle tecnologie in Sanità*. FrancoAngeli.

Asakawa, T., Sugijama, K., Sameshina, T., Kobayashi, S., Wang, L., Hong, Z., Chen, S., Li, C. & Namba, H. (2019). Can the latest computerized technologies revolutionize conventional assessment tools and therapies for a neurological disease? The example of Parkinson's disease. *Neurologia medico-chirurgica*, 59, 3, 69-78.

Baccarini, B. (2020). Per nuove politiche di home care tramite un raccordo ospedale-territorio. In C. Cipolla, M.R Dal Molin, C. Pipinato (a cura di), *Di Parkinson si vive! Dimensioni socio-assistenziali della malattia*. Franco Angeli.

Baccarini, B. (2023). La Pandemia Covid 19 e la Governance dei Sistemi socio-sanitari. Homeless Book.

Becchimanzi, C. (2022). Design e disabilità: nuove tecnologie abilitanti. In F. Tosi & M. Pistolesi (a cura di), *Home care design for Parkinson's Disease*, Ergonomia e Design. FrancoAngeli.

Busse, R. et al. (2010). Tackling chronic disease in Europe. Strategies, interventions and challenges. *World Health Organization-European observatory on Health Systems and Policies, Observatory Studies*, 20.

Dorsey, E.R. et al. (2007, 30 gennaio). Projected number of people with Parkinson disease in the most populous nations, 2005 through 2030. *Neurology*, 68 (5), 384-386.

Ennis, A., Rafferty, J., Synnott, J., Cleland, I., Nugent, C., Selby, A. & Masci, G. (2017). A Smart Cabinet and Voice Assistant to Support Independence in Older Adults. In *International Conference on Ubiquitous Computing and Ambient Intelligence: UCAmI 2017: Ubiquitous Computing and Ambient*

- Intelligence*, vol. 10586, 466-472. (Lecture Notes in Computer Science; vol. 10586). https://doi.org/10.1007/978-3-319-67585-5_47.
- Feil-Seifer, D. & Mataric, M.J. (2005). Defining socially assistive robotics, in 9th International Conference on Rehabilitation Robotics. *ICORR 2005*, Chicago, IL, 2005, IEEE, Piscataway.
- Iecovich, E. (2014). Aging in place. From theory to practice. *Anthropological notebooks*, 20, 1, 21-33.
- Lupton, D. (2014). Critical perspectives on digital health technologies. *Sociology Compass*, 8 (12), 1344-1359.
- Marengoni, A. et al. (2011). Aging with multimorbidity: a systematic review of the literature. *Ageing Research Review*, 10 (4), 430-439.
- Maturo, A. (2024). *Il primo libro di sociologia della salute*. Einaudi.
- Morozov, E. (2013). *To save Everything, Click Here, Technology, Solutionism and the Urge to Fix Problems That Don't Exist*. Allen Lane.
- Odekerken Schoroder, G., Mele, C., Russo-Spena, T., Mahar, D. & Ruggiero, A. (2020). Mitigating loneliness with companion robots in the COVID-19 pandemic and beyond: an integrative framework and research agenda. *Journal of Service Management*, 31, 6, 1149-1162.
- Rinaldi, A., Becchimanzi, C. & Tosi F. (2018). Wearable Devices and Smart Garments for Stress Management. In S. Bagnara, R. Tartaglia, S. Albolino, T. Alexander & Y. Fujita (a cura di), *Proceedings of the 20th Congress of the International Ergonomics Association (IEA 2018)* (vol. VII: *Ergonomics in Design, Design for All, Activity Theories for Work Analysis and Design, Affective Design*) (pp. 898-907). Springer.
- Sanders, C., Rogers, A., Bowen, R., Bower, P., Hirani, S., Cartwright, M. & Chrysanthaki, T. (2012). Exploring barriers to participation and adoption of telehealth and telecare within the Whole System Demonstrator trial: a qualitative study. *BMC health services research*, 12, 220, 1-12.
- Sapci, A.H. & Sapci H.A. (2019). Innovative assisted living tools, remote monitoring technologies, artificial intelligence-driven solutions, and robotic systems for aging societies: systematic review. *JMIR aging*, 2, 2, e15429.
- European Commission (2014). *Population ageing in Europe: facts, implications and policies*. Bruxelles, EU.
- Susskind, R. & Susskind, D. (2015). *The future of the professions. How technology will transform the work of human experts*. Oxford University Press.
- Topol, E. (2019). *Deep Medicine. How Artificial Intelligence Can Make Healthcare Human Again*. Basic Books.
- Yang, G.Z., Bellingham, J., Doupont, P.E., Fischer, P., Floridi, L., Full, R. & Nelson, B.J. (2018). The grand challenges of Science Robotics. *Science Robotics*, 3, 14, eaar7650.

Algorithmic Opacity

by *Giuseppe Gimigliano**, *Simona Ibba**

ABSTRACT (ENG)

This paper interrogates algorithmic opacity as an epistemic condition that resists classical models of explanation rooted in scientific realism and logical formalism. Challenging both transparency fetishism and regulatory instrumentalism, it proposes a shift from deductive clarity to contextual intelligibility. Generative AI systems, particularly foundation models, expose a paradox: they operate with increasing autonomy while becoming epistemically inaccessible, undermining the notion of algorithmic accountability. The authors argue for a reconceptualization of explanation beyond causal reducibility, drawing from contemporary philosophy of science to frame opacity as a constitutive feature of complex systems. This reframing demands democratic governance structures that are ethically sensitive to uncertainty, non-linearity, and emergence. Unlike approaches that reduce opacity to a solvable design flaw, the paper advocates treating it as a philosophical problem requiring institutional imagination. It questions the feasibility of full epistemic control, proposing instead a regime of negotiated intelligibility.

Keywords: algorithmic opacity, epistemic complexity, foundation models, transparency critique, philosophy of science

Opacità Algoritmica

di *Giuseppe Gimigliano*, *Simona Ibba*

ABSTRACT (ITA)

Il presente contributo interroga criticamente l'opacità algoritmica come una condizione epistemica che sfugge ai modelli classici fondati sul realismo scientifico e sul formalismo logico. Mettendo in discussione il ricorso costante alla trasparenza, si propone un passaggio dalla chiarezza deduttiva all'intelligibilità contestuale. I modelli fondativi dei sistemi di intelligenza artificiale contengono in sé un paradosso, ovvero operano con autonomia sempre crescente e, tuttavia, risultano epistemicamente inaccessibili. Ciò pone in crisi la nozione stessa di responsabilità algoritmica. Si propone, in questa sede, una sostanziale riconcettualizzazione della spiegazione oltre la riducibilità causale al fine di interpretare l'opacità come tratto costitutivo dei sistemi complessi. Questa riconfigurazione richiede strutture di governance democratiche che siano eticamente sensibili all'incertezza, alla non-linearità e, nondimeno, all'emergenza. Il presente contributo tratta l'opacità come un problema epistemologico riguardante la filosofia della scienza che esige co-partecipazione istituzionale, a differenza dei molti approcci che riducono la stessa opacità a un difetto progettuale tecnicamente risolvibile. Viene messa in dubbio la possibilità di un controllo epistemico totale e propone, invece, un regime di intelligibilità negoziata.

Parole chiave: opacità algoritmica, complessità epistemica, modelli fondativi, critica della trasparenza, filosofia della scienza

* Pontificia Università Antonianum

1. What is an algorithm?

The Treccani Dictionary defines the term “algorithm” as follows: (archaic: algorism) noun, masculine [from the medieval Latin *algorithmus* or *algorismus*, derived from the name of the 9th-century Arab mathematician Muḥammad ibn Mūsa al-Khwārizmī (so named after his place of origin, Khwarizm, a region in Central Asia)].

1. A term that, in the Middle Ages, referred to numerical calculation procedures based on the use of Arabic numerals. In modern usage, including reference to computers, it denotes any mathematical calculation scheme or procedure; more precisely, a calculation process that is explicit and describable with a finite number of rules, leading to a result after a finite number of operations – i.e., applications of those rules. Specifically, the Euclidean algorithm is a method for determining the greatest common divisor of two integers a and b , based on successive divisions.
2. In computer science, a set of instructions to be applied to perform a computation or solve a problem.
3. In mathematical logic, any “effective” process for computing a function or deciding a set (or predicate) – that is, any process that allows, through a finite number of steps executed according to a finite set of explicit rules, to obtain the function’s value for a given input or to decide whether a given individual belongs to the set (or satisfies the predicate).

From this definition, it becomes clear that the term *algorithm* holds a multiplicity of meanings, ranging from colloquial language to technical-scientific discourse. According to Hill (2016, pp. 35-59), in common perception, “any procedure or decision-making process, however poorly designed, can be defined as an algorithm”.

Even in academic circles, a certain ambiguity persists. The term is used not only for the mathematical construct with a finite, abstract, effective control structure given in imperative form, but also for the interaction of multiple algorithms within software developed for a specific task.

In this brief study, the ethical issue of algorithmic opacity will be analyzed by considering the characteristics of the entire software in which the algorithms are implemented. In doing so, we aim to adopt not only a technical-informatic approach, but one more consistent with a generalist context.

The non-neutrality of algorithms is supported scientifically by a wide range of research. Prates, Avelar & Lamb (2020, 6363-6381), for example, highlight the phenomenon of *machine bias*, where trained systems, unbeknownst to their designers, reflect social asymmetries related to gender or racial bias. Through experimental trials, Prates et al. demonstrated that Google Translate, when using AI systems, shows a high tendency to translate terms related to STEM jobs (science,

technology, engineering, and mathematics) into masculine forms, more than what would emerge from an actual analysis of the demographic data of workers in those sectors.

With the recent and continuous development of generative AI, algorithms are increasingly at the core of decision-making processes in social and governmental contexts, commercial transactions, language processing, and human interactions.

To clarify the value and related risks of algorithmic mediation, Mittelstadt et al. (2016) developed a map of ethical issues raised by algorithms (Fig. 1). This can be used to identify the most significant problems and hypothesize possible solutions.

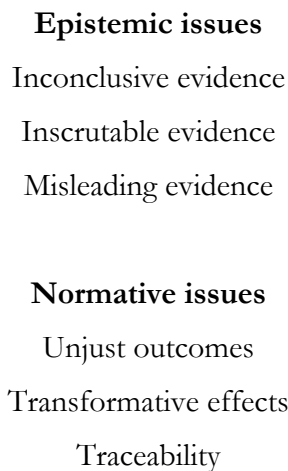


Fig. 1: Map of ethical issues related to algorithms

Algorithms can be used to generate information or trigger actions that may have ethical consequences. In deep learning, systems are autonomous or semi-autonomous, as they produce results by processing synthetic data using non-deterministic models. In addressing this scenario, the map in Fig. 1 identifies six ethical issues:

- Inconclusive, inscrutable, and misleading evidence pertain to epistemic factors.
- Unfair outcomes and transformative effects are explicitly normative in nature.
- Traceability is significant both epistemologically and normatively.

Epistemic factors highlight the importance of data quality (Floridi & Illari, 2014) and accuracy in justifying the conclusions reached by algorithms, as well as the subsequent decisions that carry moral relevance for individuals and the environment.

Normative issues refer to unjust outcomes, unintended consequences, and a lack of transparency – hence, to the opacity of algorithmic processes. We will explore this latter aspect in greater detail.

The sixth ethical issue, traceability, refers to both epistemic and normative concerns, as well as to design and development factors that hinder the ability to identify the cause of a decision, and therefore to assign corresponding responsibility.

The lack of transparency that often characterizes machine learning (ML) algorithms gives rise to the ethical issue we can define as *inscrutable evidence*.

The causes of this lack of transparency can be varied and affect every phase of the design, management, or maintenance of the software or its associated processes. In particular, during design, it is important to consider the decision-making phase, which can give rise to significant technological limitations resulting from the failure to consider certain relevant ethical, communication, and structural aspects.

In other cases, intentional data obfuscation may occur. This operation, which is necessary in some respects to ensure protection from unauthorized access and to safeguard personal information, can be disruptive if performed excessively.

Excessive obfuscation can render data unusable even for legitimate operations and can introduce greater complexity into processes.

Legal constraints related to intellectual property must also be considered, resulting in a lack of control, accountability, and reliability.

Another extremely significant aspect is the human cognitive inability to interpret large algorithmic models and analyze large data sets. This type of analysis presents an unequal struggle between human and computational agents. It's impossible for a human agent to analyze, in a limited amount of time, the amount of data a machine can analyze.

Poor code quality or poorly structured data can also reduce the transparency of AI systems. Good code quality does not directly affect the functional properties of the software, but rather the way it operates. It takes into account not only internal qualities, but also dynamic and behavioral properties and the user experience. Continuous algorithm updates, often driven by economic motivations, ultimately lead to excessive algorithmic malleability that obscures the history of its evolution.

In self-learning algorithms typical of generative AI, the lack of transparency is also an intrinsic concept, due to the automatic production of new sets of rules and synthetic data during the learning process. Even for the developers themselves, the reasons behind certain decisions made by the software remain unclear. This obscurity must not, from an ethical standpoint, justify the development of complex systems that remove both developers and users from human responsibility.

Furthermore, transparency is not inherently an ethical concept, but rather a necessary condition for the establishment of ethically sound development and use environments.

In certain algorithmic contexts – such as the electoral domain – data opacity may be ethically more appropriate than transparency, which instead could entail greater risk. An overload of information that is difficult for users to understand makes the system opaque, difficult to understand, and structurally inconsistent. Excessive focus on transparency can also divert human and financial resources from more ethically relevant issues, such as security or accuracy. Excessive transparency can also make the system more vulnerable to attacks. Indeed, it is important to consider that, in an electoral system, it is important to safeguard the secrecy of the vote, which should be cast without external influence and should not be subject to alteration. The “virtual ballot box” must ensure that votes cannot be substituted or observed during the election session.

Based on the arguments presented in the previous sections, we cannot consider transparency an ethical end in itself. Instead, it is a pro-ethical condition (Wachter, Mittelstadt & Floridi, 2017) that must be assessed in relation to the context and applied in appropriate measure. From this perspective, it becomes crucial to distinguish between the various factors that may hinder algorithmic transparency.

Based on the arguments presented in the previous sections, transparency cannot be considered an ethical goal in itself. Rather, it is a pro-ethical condition that must be assessed in relation to the context and applied appropriately. From this perspective, it becomes crucial to distinguish between the various factors that can hinder algorithmic transparency.

Several research fields have proposed tools and approaches to address algorithmic transparency. The use of standardized documentation procedures, such as technical data sheets in the electronics sector, is considered one of the main transparency tools and is necessary for the formal certification of quality standards. Alongside the formal aspect, it is important to combine the use of testing tools that evaluate individual algorithms and decision-making processes based on parameters pre-established during the design phase. The training and communication aspect is also essential: adequate training for companies, public bodies, and developers on the risks of algorithmic opacity, with the dual objective of promoting more transparent and therefore more reliable design and improving users’ understanding of algorithms, is the foundation of adequate transparency.

Added to these are formative parameters that identify measurable transparency factors regarding data, models, inferences, and interfaces. For example, Diakopoulos and Koliska identified the following factors: uncertainty – including the possible margin of error; timeliness – which indicates when the data was collected or the data sources; the volume of data on which the systems were trained; and completeness and sampling methods.

Explainable AI, or What-If Tool, from Google and IBM, are examples of interpretable explanations for the internal processes of ML algorithms.

What-If, for example, is a tool developed by Google for data analysis and model interpretation in the field of artificial intelligence. It provides a range of features, including graphical and interactive ones, that allow users to gain a deeper understanding of machine learning models and the data on which they operate. Users can gain insight into the impact of individual software features within the overall model, understand which parameters are most influential, and identify how manipulation of input data can impact model predictions. Interpretability also refers to the evaluation of the interactions between machine learning techniques and the data they operate on. It is important both for explaining the decisions a system makes regarding an individual (local interpretability) and for evaluating broader discriminatory models (global interpretability).

For example, interpretability allows a candidate in a competition who receives a certain score generated by a machine learning algorithm to understand the reasons behind that numeric value and to compare it with the scores of other candidates.

Identifying appropriate methods to increase the clarity of machine learning algorithms is particularly complex due to their inherently multidimensional and intricate nature, the growing number of models, and the vastness of the data on which these models are trained.

The most significant research efforts have focused on two main axes: subject-centered explanations: tailored to users and their interactions with the algorithm and model-centered explanations: focused exclusively on the internal logic and functioning of the algorithm itself.

2. Opacity in generative AI System

Generative AI is based on deep learning neural networks (LeCun, Bengio & Hinton, 2015, pp. 436-444.). These are very large networks, organized in multiple layers, with an extremely high number of samples: each layer computes the values for the next, processing the information in an increasingly refined way, generating a vast amount of synthetic data (not drawn from the real world, but computed by machines).

In August 2021, the Stanford Institute for Human-Centered Artificial Intelligence's (HAI) Center for Research on Foundation Models (CRFM) defined a *Foundation Model* as: "Any model trained on broad data (typically using large-scale self-supervision) that can be adapted to a wide range of tasks".

In these models, data is not labeled by humans and can self-structure, leading to the emergence of unexpected properties. For example, a model trained on a large set of linguistic data may learn

to generate original stories or perform arithmetic operations without being explicitly programmed to do so.

This adaptability to a wide range of tasks enables the use of these models across many domains, but it also opens the door to the proliferation of cascading errors from the base model to the respective derivative models used in specific application areas.

Currently, there is no clear understanding of how these models function internally due to their emergent properties, nor is it possible to reliably predict when they will fail or what they are capable of. To tackle such issues, Bommasani et al. (2021) argue that most critical research on Foundation Models will require deep interdisciplinary collaboration, aligned with the fundamentally sociotechnical nature of these systems.

The impact of foundation models has been growing rapidly in recent months, while transparency is simultaneously decreasing. To assess the level of opacity in foundation models, Bommasani et al. (2023) introduced a transparency index that considers all phases of software development—from design to market deployment. In Table 1, transparency scores for the main foundation models are shown based on the most significant transparency metrics.

Foundation Model Transparency Index Scores by Major Dimensions of Transparency, 2023

Source: 2023 Foundation Model Transparency Index

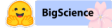
	 Meta	 BigScience	 OpenAI	 stability.ai	 Google	 ANTHROPIC	 cohere	 AI21labs	 Inflection	 amazon	
	Llama 2	BLOOMZ	GPT-4	Stable Diffusion 2	PaLM 2	Claude 2	Command	Jurassic-2	Inflection-1	Titan Text	Average
Data	40%	60%	20%	40%	20%	0%	20%	0%	0%	0%	20%
Labor	29%	86%	14%	14%	0%	29%	0%	0%	0%	0%	17%
Compute	57%	14%	14%	57%	14%	0%	14%	0%	0%	0%	17%
Methods	75%	100%	50%	100%	75%	75%	0%	0%	0%	0%	48%
Model Basics	100%	100%	50%	83%	67%	67%	50%	33%	50%	33%	63%
Model Access	100%	100%	67%	100%	33%	33%	67%	33%	0%	33%	57%
Capabilities	60%	80%	100%	40%	80%	80%	60%	60%	40%	20%	62%
Risks	57%	0%	57%	14%	29%	29%	29%	29%	0%	0%	24%
Mitigations	60%	0%	60%	0%	40%	40%	20%	0%	20%	20%	26%
Distribution	71%	71%	57%	71%	71%	57%	57%	43%	43%	43%	59%
Usage Policy	40%	20%	80%	40%	60%	60%	40%	20%	60%	20%	44%
Feedback	33%	33%	33%	33%	33%	33%	33%	33%	33%	0%	30%
Impact	14%	14%	14%	14%	14%	0%	14%	14%	14%	0%	11%
Average	57%	52%	47%	47%	41%	39%	31%	20%	20%	13%	

Table 1: Foundation Model Transparency Index Scores by Major Dimensions of Transparency, 2023

An analysis of the data reveals the following:

- The highest score achieved is 54/100.
- The average overall score is 37/100.
- All systems therefore require significant improvements in terms of transparency.

- There is a 37-point gap between the most transparent foundation model (Meta) and the least transparent one.

Foundation Model Transparency Index Total Scores, 2023

Source: 2023 Foundation Model Transparency Index

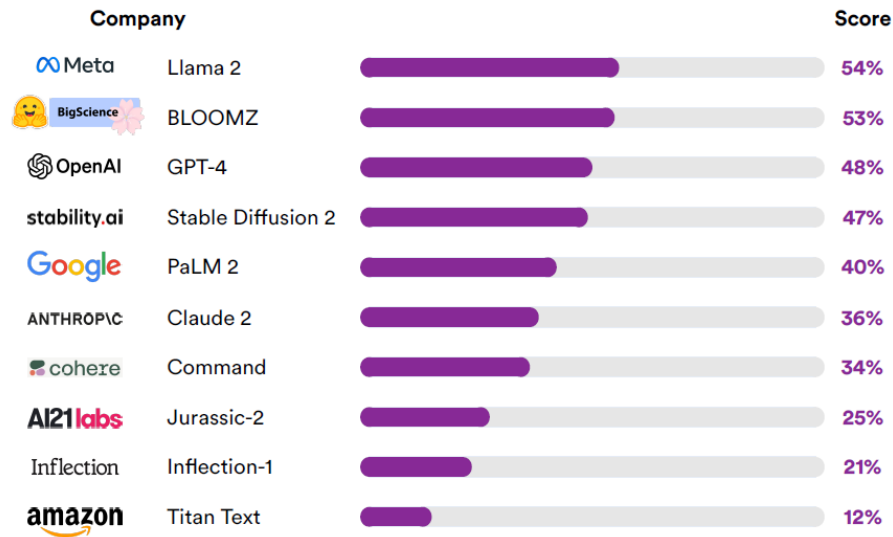


Figure 2: Foundation Model Transparency Index Scores, 2023

It is particularly notable that several models have a zero-transparency index regarding the input data used by the system (e.g., AI21 Labs, Inflection, Amazon).

Foundation Model Transparency Index Scores by Domain, 2023

Source: 2023 Foundation Model Transparency Index

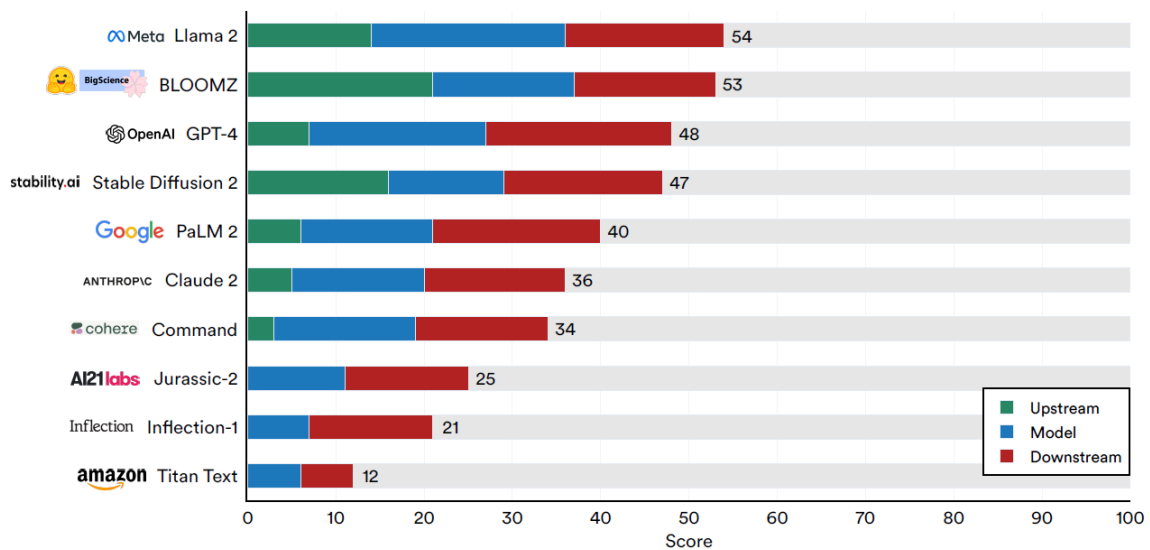


Figure 3: Foundation Model Transparency Index Scores by Domain, 2023

There is complete opacity regarding knowledge of the system's input data. No foundation model discloses the data sources, the associated usage licenses or information related to copyright.

Source: 2023 Foundation Model Transparency Index

Subdomain	Indicator	Meta	BigScience	OpenAI	stability.ai	Google	ANTHROPIC	cohere	A121labs	Inflection	amazon
		Llama 2	BLOOMZ	GPT-4	Stable Diffusion 2	PaLM 2	Claude 2	Command	Jurassic-2	Inflection-1	Titan Text
Data	Data size	1	1	0	1	0	0	1	0	0	0
	Data sources	0	1	0	0	0	0	0	0	0	0
	Data creators	0	0	0	0	0	0	0	0	0	0
	Data source selection	0	1	0	0	0	0	0	0	0	0
	Data curation	1	1	1	1	1	0	1	0	0	0
	Data augmentation	1	1	0	1	1	0	0	0	0	0
	Harmful data filtration	1	1	1	1	0	0	0	0	0	0
	Copyrighted data	0	0	0	0	0	0	0	0	0	0
	Data license	0	0	0	0	0	0	0	0	0	0
	Personal information in data	0	0	0	0	0	0	0	0	0	0
Data Labor	Use of human labor	1	1	0	0	0	0	0	0	0	0
	Employment of data laborers	0	1	0	0	0	0	0	0	0	0
	Geographic distribution of data laborers	0	1	0	0	0	0	0	0	0	0
	Wages	0	1	0	0	0	1	0	0	0	0
	Instructions for creating data	1	1	0	0	0	1	0	0	0	0
	Labor protections	0	0	1	0	0	0	0	0	0	0
Data Access	Third party partners	0	1	0	1	0	0	0	0	0	0
	Queryable external data access	0	1	0	1	0	0	0	0	0	0
	Direct external data access	0	1	0	1	0	0	0	0	0	0
Compute	Compute usage	0	0	0	0	0	0	0	0	0	0
	Development duration	1	0	0	1	0	0	0	0	0	0
	Compute hardware	0	1	0	1	0	0	0	0	0	0
	Hardware owner	1	0	1	1	1	0	1	0	0	0
	Energy usage	1	0	0	0	0	0	0	0	0	0
	Carbon emissions	1	0	0	1	0	0	0	0	0	0
Methods	Broader environmental impact	0	0	0	0	0	0	0	0	0	0
	Model stages	1	1	1	1	1	1	0	0	0	0
	Model objectives	1	1	1	1	0	1	0	0	0	0
	Core frameworks	0	1	0	1	1	0	0	0	0	0
	Additional dependencies	1	1	0	1	1	1	0	0	0	0
Data Mitigations	Mitigations for privacy	1	1	1	1	0	0	0	0	0	0
	Mitigations for copyright	0	1	0	0	0	0	0	0	0	0
Upstream Subtotal		44%	66%	22%	50%	19%	16%	9%	0%	0%	0%

Figure 4: Foundation Model Transparency Index 2023

Transparency is highest with regard to the algorithmic models used in the system. However, most companies do not reveal information about:

- the size of the model;
- nor explain how or why they made certain decisions concerning model release.

Only two developers highlight the limitations resulting from opacity. None assess the multiple intentional harms their models could potentially cause.

No developer provides:

- a reporting system for usage;
- nor a mechanism for users to request redress.

3. How to address transparency in AI algorithms: hypotheses

Based on the data presented above, several political and legal approaches regarding generative AI could be hypothesized:

- Transparency is a fundamental prerequisite for innovation.
- Political and legislative tools should encourage the responsible development of foundation models.
- End users should have access to adequate information about generative AI systems—especially those that are extremely powerful and used by hundreds of millions of people.
- Existing transparency regulations should be applied.
- Regulation should take into account that transparency is not an ethical principle in and of itself, and that it is only useful insofar as the information it yields is actionable and not excessive.

4. Algorithmic Opacity and Governance Frameworks

To speak of opacity in algorithms is to touch on one of the most difficult and contested questions of our time. The metaphor of the “black box” is often invoked, but perhaps it is already inadequate. These are not boxes; they are sprawling systems, architectures with countless nodes and parameters, operating at scales and speeds that human beings cannot easily imagine. Yet, despite this, they shape our daily lives. They decide whether a loan is granted, whether a résumé passes the first stage of a job application, or whether a citizen is flagged by predictive policing software. The opacity here is not just a matter of inconvenience. It is a matter of legitimacy. Can decisions with consequences for individuals and communities be accepted if their rationale is hidden, inaccessible, or – worse – nonexistent in any form that humans can interpret? The European Union has tried to craft an answer through its Artificial Intelligence Act. The design of this legislation is pragmatic: classify systems by risk, then apply obligations in proportion. Low-risk applications, like entertainment algorithms, face minimal regulation. High-risk systems – healthcare diagnostics, biometric surveillance, education scoring tools – are held to standards of documentation, traceability, and human oversight.

It is a clever attempt to balance innovation with responsibility. And yet the very ambition of the Act exposes its fragility. To demand transparency of a neural network trained on hundreds of billions of words is to ask the impossible. Disclosing a dataset does not render it meaningful. Listing technical specifications may create paperwork but seldom genuine comprehension. Transparency, in such a case, risks becoming a bureaucratic ritual: something that exists on paper but fails in practice.

Other frameworks, meanwhile, have chosen different tones. The OECD Principles on AI, adopted in 2019, provide not laws but moral orientation: fairness, robustness, human-centered values. They aim for consensus, not coercion. UNESCO's Recommendation on the Ethics of Artificial Intelligence, issued in 2021, takes the language further, embedding AI in a discourse of human rights, sustainability, and cultural pluralism. In the United States, there is no federal law equivalent to the AI Act. Instead, the National Institute of Standards and Technology proposes a Risk Management Framework: a voluntary set of guidelines that organizations may use to identify and reduce risks. These contrasts are telling. Europe leans on law, UNESCO invokes global morality, the OECD offers gentle coordination, and the United States relies on managerial pragmatism. Each solution reflects the culture that produced it; none is sufficient on its own.

But opacity is not always a matter of secrecy. In many cases, it is structural. Neural networks – particularly foundation models – generate patterns and outcomes that even their developers cannot fully explain. This is not because companies refuse to disclose information (though that often happens), but because the very architecture of the system resists human reasoning. Emergent behavior is not reducible to neat causal chains. Opacity, in this sense, is not a flaw to be corrected but a characteristic of the technology itself. And this raises a troubling question: how can governance ensure accountability when the logic of the system is, by design, beyond our grasp?

The global dimension complicates the matter further. Consider a single algorithm: designed in California, trained on servers in Finland, fine-tuned in Bangalore, and deployed in hospitals in Milan. Which legal regime applies? Who is responsible for oversight? The reality is that no single authority can claim effective control. The result is fragmentation and, often, regulatory arbitrage. Corporations place their operations in jurisdictions with permissive rules, while exporting their products worldwide. The comparison with financial markets is instructive: just as capital flows evade national regulators, so too do algorithms move across borders, indifferent to the reach of a single legal system.

Scholars and policymakers have proposed institutional oversight mechanisms to manage this landscape. Algorithmic impact assessments, independent audits, corporate ethics boards: each is useful, none is decisive. Audits can reveal patterns of bias or measure error rates, but they struggle with questions of fairness or dignity. Ethics boards lend moral weight, yet without binding authority they risk remaining symbolic.

Impact assessments may encourage foresight, but too often they are reduced to box-ticking exercises. Ironically, the burden of compliance falls hardest on smaller actors, while the largest corporations – those with the most powerful and opaque systems – can absorb costs and present

themselves as leaders in responsible AI. Governance, in this way, sometimes reproduces inequality even as it seeks to remedy it.

It is here that philosophy must intervene. Governance frameworks are not neutral scaffolding applied to autonomous technologies. They shape, from within, the way algorithms are imagined and legitimized. To demand transparency is not merely to ask for more data; it is to define what counts as valid explanation, what counts as accountability. Too much disclosure overwhelms; ordinary citizens cannot parse thousands of technical details, and excessive openness may expose systems to malicious exploitation. Too little disclosure produces decisions that resemble commands: authoritative but unjustified. The task is to find balance, to calibrate transparency in ways that genuinely support accountability without collapsing into noise or danger.

One may borrow an analogy from environmental law. Pollution has never been eliminated, but it has been constrained, redistributed, and managed through institutions, norms, and sanctions. So it is with algorithmic opacity: it will not vanish, but it can be disciplined. Europe's legal instruments, UNESCO's ethical appeals, OECD's principles, NIST's pragmatic frameworks – taken alone, each is partial; taken together, they begin to form a mosaic. Imperfect, fragile, and constantly in need of revision, but nonetheless indispensable.

What remains is the recognition of limits. Governance cannot turn opaque systems into glass boxes. The ambition to fully “explain” every algorithm is unrealistic, and perhaps misguided. What governance can do is to create conditions under which opacity is tolerable: where risks are mitigated, accountability is possible, and the dignity of affected individuals is respected. The future of algorithmic governance will depend on the capacity to coordinate across jurisdictions, to reconcile innovation with democratic oversight, and to accept transparency not as an absolute virtue but as a contextual practice. Only then can opacity, however stubborn, coexist with justice, responsibility, and the fragile trust on which democratic societies depend.

Conclusions

Deep learning represents a revolution that creates both opportunities and numerous risks and limitations. Given its rapid development and the impact it has on citizens, AI cannot remain an unknown space.

The opacity of an algorithm, which includes both the lack of awareness about its working assumptions and the failure to consider its limits of applicability, leads to misuse and harmful decisions lacking proper ethical grounding.

Recognizing this revolution – so captivating and, at the same time, so unpredictable – initiates a philosophical reflection on humanity and its responsibilities, aimed at developing critical thinking with the explicit purpose of defining appropriate boundaries.

In conclusion, algorithmic opacity cannot be eliminated; it can only be managed, negotiated, and constrained within the limits of our political and ethical frameworks. The task ahead is not to imagine a world without opacity, but to design institutions strong enough to hold it accountable. Sometimes this will require law, sometimes ethics, sometimes simply the steady work of oversight that resists easy visibility. What ultimately matters is that opacity never becomes an excuse for irresponsibility, but rather a challenge that compels societies to reaffirm justice, responsibility, and human dignity.

Bibliography

- Bommasani, R., Hudson, D.A., Adeli, E., Altman, R., Arora, S., von Arx, S. & Liang, P. (2021). *On the opportunities and risks of foundation models*. HAI.
- Bommasani, R., Klyman, K., Longpre, S., Kapoor, S., Maslej, N., Xiong, B. & Liang, P. (2023). *The foundation model transparency*. HAI.
- Floridi, L. & Illari, P. (eds.). (2014). *The philosophy of information quality*. Springer.
- Hill, R.K. (2016). *What an algorithm is*. Philosophy & Technology.
- LeCun, Y., Bengio, Y & Hinton, G. (2015). *Deep learning*. Nature.
- Prates, M.O., Avelar, P.H. & Lamb, L.C. (2020). *Assessing gender bias in machine translation: a case study with google translate*. Neural Computing and Applications.
- Mittelstadt, B.D., Allo, P., Taddeo, M., Wachter, S. & Floridi, L. (2016). *The ethics of algorithms: Mapping the debate*. Big Data & Society.
- Wachter, S., Mittelstadt, B. & Floridi, L. (2017). *Why a right to explanation of automated decision-making does not exist in the general data protection regulation*. International Data Privacy Law.

La trasformazione digitale della società: dalla mediamorfosi al transumanesimo

di *Sandra Ciarcianelli**

ABSTRACT (ITA)

La rivoluzione digitale che caratterizza l'inizio del XXI secolo ha trasformato quasi tutti gli ambiti della vita delle persone in una nuova prospettiva che le conduce ad una esistenza digitalmente e perennemente interconnessa: il postumano. Questo processo di trasformazione rappresenta o cerca di ripensare l'umano mettendo in discussione criticamente l'umanesimo, quella corrente della filosofia umanista che considera la natura umana uno stato universale da cui emerge l'umano con la sua autonomia, razionalità ed esercizio del libero arbitrio. La società contemporanea si trova quindi a dover riflettere e a indagare sugli aspetti sociologici, filosofici ed etici, ma anche pedagogici e sociali, che si potrebbero prospettare in un futuro non così lontano, quando il confine e i limiti del corpo e della mente diventeranno sempre più labili e indistinguibili.

Parole chiave: postumano, rivoluzione digitale, convergenza tecnologica, cyborg, intelligenza artificiale

The Digital Transformation of Society: from Mediamorphosis to Transhumanism

by *Sandra Ciarcianelli*

ABSTRACT (ENG)

The digital revolution that characterized the beginning of the 21st century has transformed nearly every aspect of people's lives into a new perspective, leading them to a digitally and perpetually interconnected existence: the posthuman. This transformation represents, or seeks to rethink, the human by critically questioning humanism, that current of humanist philosophy that considers human nature a universal state from which the human emerges with its autonomy, rationality, and exercise of free will. Contemporary society therefore finds itself compelled to reflect and investigate the sociological, philosophical, and ethical, as well as pedagogical and social, aspects that could emerge in a not-so-distant future, when the boundaries and limits of the body and mind will become increasingly blurred and indistinguishable.

Keywords: posthuman, digital revolution, technological convergence, cyborg, artificial intelligence

*Università eCampus

1. Introduzione

Quando il giornalista statunitense Roger Fidler nel 1997 scrisse *Mediamorphosis: Understanding New Media*, utilizzò un neologismo da lui stesso coniato qualche anno prima, per delineare il complesso processo di mutamento e trasformazione che si stava compiendo con l'arrivo dei media digitali. La sua teoria condusse ad una profonda analisi dell'evoluzione delle tecnologie e dei mezzi di comunicazione di massa che, invece di essere studiati separatamente, furono intesi come parti di un unico e complesso sistema interdipendente, nel quale cogliere le relazioni esistenti tra passato e presente.

Fidler, un visionario e pioniere dell'editoria elettronica, definì *mediamorfosi* la trasfigurazione e mutabilità del processo comunicativo dei mass media causate dalle influenze reciproche dei bisogni percepiti, delle tensioni politiche, delle rivoluzioni tecnologiche e, non ultime, le trasformazioni del tessuto sociale: “Le previsioni che descrivono ciò che le persone faranno concretamente con queste reti e con i servizi che esse garantiranno sono ampiamente divergenti e contrastanti” (Fidler, 2000, p. 13).

Egli considerava il sistema della comunicazione come un unico insieme nel quale l'arrivo dei nuovi media non avveniva in modo casuale e spontaneo, ma era il risultato della trasformazione dei suoi precursori: con l'ingresso dei nuovi dispositivi tecnologici, quelli precedenti non si estinguono, ma si adattano e si trasformano. Le trasformazioni tra passato e presente avvengono attraverso il principio della *mediamorfosi* e basandosi su tre presupposti fondamentali (Fidler, 2000, p. 13).

1 *Coevoluzione* – l'evoluzione costante e la reciproca influenza delle nuove e passate forme di comunicazione. Coevoluzione e anche coesistenza poiché la metamorfosi non avviene come un processo di evoluzione sequenziale o come mera sostituzione del mezzo antico con uno più aggiornato;

2 *Convergenza* – l'integrazione di conoscenze, strumenti e tutte le attività rilevanti per gli esseri umani finalizzati ad un obiettivo comune, consentendo alla società di fornire risposte alle nuove domande per migliorare l'ecosistema fisico e sociale;

3 *Complessità* – il livello percepito degli eventi come sistemi apparentemente caotici ma che, proprio grazie al loro stato di esistenza tra caos e ordine, i sistemi complessi emergono e si evolvono.

Il concetto di *convergenza tecnologica* è applicato principalmente nel campo delle telecomunicazioni per denominare le reti, i sistemi e i servizi che sono da esse generate, combinandoli tra loro creando nuove possibilità sinergiche. Può essere definita come una interconnessione tra le

tecnologie informatiche, i contenuti multimediali e le reti di comunicazione che sono scaturite come risultato dell'evoluzione e della diffusione di Internet e di tutti gli spazi digitali.

Tuttavia, la convergenza tecnologica deve fare anche i conti con una società fondata sul consumismo, per il quale l'evoluzione tecnologica comporta una sempre maggiore produzione di prodotti sempre più ipertecnologici, da un lato, e uno stravolgimento delle potenziali capacità fisiche e cognitive dell'essere umano, dall'altro: *dall'Homo sapiens al postumanesimo tecnologico*.

I cambiamenti nell'economia e nel mondo del lavoro, nella vita pubblica e privata, che affiancano la rivoluzione digitale, si attuano a ritmo rapido ovunque esistano i prerequisiti materiali per l'applicazione e l'uso della digitalizzazione avanzata. I nuovi media influenzano sempre più i comportamenti comunicativi, i processi di socializzazione, la cultura linguistica e la percezione della realtà soggettiva viene costantemente ri-contestualizzata in base alla realtà virtuale.

Negli anni Novanta il postumanesimo entra prepotentemente nella riflessione culturale e scientifica per poter indagare sulle nuove esigenze contemporanee, alla ricerca di una prospettiva più filosofica, di postulare scenari teoretici ed epistemologici non strettamente scientifici e biotecnologici. La posizione postumana riconosce l'imperfezione e la scissione al proprio interno per poter interpretare il mondo attraverso prospettive più ampie ed eterogenee, cercando, tuttavia, di mantenere un rigore intellettuale e dedizione all'osservazione oggettiva. Al centro della pratica postumana vi è la capacità di cambiare prospettiva in modo fluido e in costante divenire e di manifestarsi attraverso identità diverse. In pratica il postumano non è un individuo singolare e ben definito, ma un soggetto che può assumere o incarnare identità distinte e comprendere il mondo da prospettive multiple ed eterogenee (Haraway, 1991).

I nuovi termini, quali "postumano", "transumano", assurgono ad una terminologia ontologica della ricerca nel contempo filosofica e scientifica, ma anche esistenziale e umano, riferendosi a differenti paradigmi teorici divergenti. Il termine "postumanesimo" si può dunque riferire ad un'area di ricerca filosofica, culturale e critica (Ferrando, 2017, pp. 51-61).

In contrapposizione al concetto dell'umanesimo rinascimentale, fondato su una visione ottimistica delle capacità dell'umanità di trovare uno stile di vita equilibrato, il postumanesimo affronta questioni di etica, di giustizia sociale, di comunicazione tecnologica e di sistemi sociali, negando la posizione speciale dell'uomo e considerandolo come una delle tante specie naturali.

Nella teoria critica della Scuola di Francoforte, il postumano si riferisce ad un essere teoretico che rappresenta o che va alla ricerca di una ri-concettualizzazione dell'umano. La teoria critica assume un approccio teorico che, in contrasto con la teoria tradizionale di matrice cartesiana, è interessata ad unire teoria e pratica incorporando la tensione epistemologica presente nella filosofia e nella storia delle scienze sociali. Questo approccio di ricerca si apre alla filosofia sociale concentrandosi sulla valutazione riflessiva e critica della società e della cultura al fine di analizzare

e sfidare i costrutti di potere: i problemi sociali sono principalmente influenzati e creati dalle strutture sociali e culturali e non dai fattori individuali e psicologici. In tale contesto, il postumano mette in discussione criticamente ciò che invece afferma l'umanesimo: dall'affermazione che la natura umana è uno stato universale da cui emerge l'essere umano, alla natura umana come autonoma, razionale e in grado di assurgere il libero arbitrio e unificata in sé stessa come apice dell'esistenza.

Come sosteneva Horkheimer, è l'ideologia il principale ostacolo alla liberazione umana, poiché mira a renderli schiavi (Horkheimer & Marcuse, 2003). La teoria critica della società parte dal principio di una riflessione sul carattere scientifico delle scienze umane, ovvero da una critica al pensiero empirico come fonte di spiegazione dei fenomeni sociali, incentrata sulla comprensione della cultura come elemento di trasformazione della società. In questo contesto teorico, la posizione postumana non può che riconoscere una nuova ontologia facendo riferimento a diversi ambiti della conoscenza come l'informatica e le biotecnologie, oltre alla filosofia e alle scienze sociali.

2. Il Postumanesimo e la Rivoluzione Digitale

Il postumanesimo parte da una nuova prospettiva degli schemi interpretativi, nei quali l'informatica e le biotecnologie, assieme alle scienze sociali, contribuiscono a modificare la percezione della realtà e dell'ambiente circostante, grazie al progresso scientifico che mira a spingersi verso nuove dimensioni travalicando i confini naturali dell'essere umano. L'obiettivo è quello di riuscire a trasformare lo sviluppo del futuro dell'umanità dove i lineamenti naturali umani si fondono con quelli non-umani per creare nuovi individui ibridi con innovative facoltà cognitive e fisiche.

Secondo Rosi Braidotti (2014), questo filone di ricerca filosofica mira a superare le concezioni tradizionali dell'umanità per le quali il termine *postumano* indica una concezione iper-umanistica dell'umanità che, come tale, deve però elaborare le esclusioni e le asimmetrie degli esseri umani per potersi considerare una delle tante specie naturali.

Il postumanesimo, infatti, nega lo status speciale degli esseri umani che non hanno il diritto di sfruttare e distruggere la natura o di considerarsi eticamente superiori. Inoltre, viene spesso profondamente distinto il *postumanesimo tecnologico* dal *postumanesimo critico*: se con il primo si intende superare la recente fase di sviluppo umano attraverso l'ausilio della tecnologia, con il secondo si preoccupa di superare la concezione umanistica dell'umanità, per la quale la tecnologia svolgerebbe solo un ruolo di subalternità.

Il postumanesimo tecnologico sostiene l'idea che l'Intelligenza Artificiale sostituirà gli esseri umani con un essere ipotetico e futuro le cui capacità della mente e del superamento del corpo biologico possano superare di gran lunga quelle dell'essere umano odierno (Salerno, 2021). Un essere postumano potrebbe essere descritto come una creatura che nasce attraverso l'espansione delle sue capacità fisiche e psicologiche ma anche essere la creazione di un'unità di intelligenza umana e artificiale nella quale la coscienza umana venga trapiantata in un corpo estraneo o caricata in una struttura informatica.¹

Il postumanesimo critico invece disapprova principalmente l'antropologia filosofica per aver affidato le sue definizioni all'*essenzialismo* degli esseri umani, ovvero all'essenza generale e permanente di un individuo concreto, secondo cui le *entità* hanno proprietà necessarie. L'entità indica anche l'essenza di un oggetto come elemento necessario per l'esistenza e l'identità dell'oggetto stesso. Per i postumanisti è impossibile sapere cosa si prova ad essere un altro essere e, di conseguenza, è anche impossibile conoscere se tale essere possieda la caratteristica del libero arbitrio.

Gli approcci al postumanesimo sono molto variegati, poco omogenei e spesso molto critici: lo stesso termine è stato contestato da uno dei principali autori associati al postumanesimo, lo scrittore, filosofo e artista messicano Manuel De Landa (nato 1952), che propone un approccio realista e materialista alla conoscenza della realtà attraverso una *teoria generale dell'assemblaggio*. Il suo obiettivo è quello di sfidare il paradigma sociologico e di condurre analisi significative concentrando la ricerca dalla micro alla macro-scala, ovvero dalle azioni specifiche degli individui al comportamento delle società nel loro complesso.

Introducendo il concetto della teoria degli assemblaggi del filosofo francese Gilles Deleuze (1925-1995) che, anche attraverso le teorie della complessità e dell'organizzazione e le teorie dell'informazione, cerca di indicare un'entità o un sistema composto da molteplici componenti eterogenei che si relazionano tra loro. Questi assemblaggi sono molto variabili, spaziando da assemblaggi biologici e tecnologici, ad assemblaggi sociali e culturali; quindi, la società non è considerata solo come proprietà degli esseri umani, ma anche come proprietà emergente degli assemblaggi stessi.

L'artista britannico Robert Pepperell (nato 1963), nel suo *Manifesto del postumano* (Pepperell, 2003) afferma invece che “*i corpi umani non hanno confini*” stravolgendo i tradizionali presupposti umanisti della scienza e della filosofia, dell'arte e della visione antropocentrica del genere umano. Egli espone una nuova concezione della condizione umana basandosi su diverse teorie, come quella

¹ Esempi potrebbero essere una modifica dell'organismo umano attraverso la nanotecnologia, interfacce neurali, farmaci che espandono la coscienza, le tecnologie informatiche indossabili o impiantate direttamente nei corpi umani.

quantistica, del caos, ma anche quelle del cyberpunk e della cibernetica connesse agli attuali progressi tecnologici e scientifici.

La storica delle scienze naturali statunitense Donna Haraway, afferma nel suo *A Cyborg Manifesto* (2018), che postumano è quasi sinonimo di cyborg, poiché “umano” e “robot” sono le due componenti principali del postumano critico. L’autrice sostiene che l’umanesimo liberale, che distingue e separa la mente dal corpo, considerando quest’ultimo come un guscio, un veicolo per la mente, diventa sempre più complesso poiché, dalla fine del XX e all’inizio del XXI secolo, la tecnologia dell’informazione ha iniziato a mettere in discussione i limiti e i confini del corpo umano. Tuttavia, è necessario essere consapevoli dei progressi della tecnologia informatica e allo stesso tempo occorre comprendere come disincarnata, qualcosa che non può certamente sostituire il corpo umano, ma può solo essere inglobata in esso e nelle pratiche della vita umana (Hayles, 1999).

Anche De Landa ha esplorato l’intreccio della storia dell’umanità con la storia della guerra e delle tecnologie attraverso l’uso di armi e bombe *intelligenti* e la loro relazione con la disumanizzazione della guerra, la mistificazione delle tecnologie, l’ossessione per la sorveglianza e la conversione del conflitto armato come spettacolo trasmesso in tempo reale.

Nella sua opera *La guerra nell’era delle macchine intelligenti* (1996) emerge l’influenza delle teorie sviluppate con fervore da alcuni intellettuali francesi come Foucault, Deleuze e Baudrillard che hanno avuto un notevole impatto sul lavoro filosofico americano nella seconda metà del XX secolo. Una delle principali preoccupazioni di De Landa è l’emergere e l’introduzione dell’Intelligenza Artificiale nell’industria militare, il cui principale obiettivo è la completa e totale eliminazione degli esseri umani come attori direttamente coinvolti nelle tattiche di guerra. Ciò fa emergere la prospettiva di un cieco e bieco trasferimento di potere dall’uomo alla macchina e alla potenziale minaccia di un dominio distruttivo da parte dell’Intelligenza Artificiale. Per arginare questo possibile scenario, De Landa sostiene la decentralizzazione dei sistemi di controllo e il supporto allo sviluppo di interfacce che possano generare una simbiosi, una complementarità tra esseri umani e macchine.

Ancor prima dell’Intelligenza Artificiale, l’arrivo della tecnologia digitale e dei computer, ormai più di trenta anni fa, ha trasformato tutti gli ambiti della vita in molti paesi industrializzati dalla fine del XX secolo, inaugurando una nuova epoca definita *Rivoluzione Digitale*. La digitalizzazione ha radicalmente modificato il mondo del lavoro, creando nuovi modelli di business, nuove strutture competitive e cambiamenti sostanziali nei contenuti del lavoro, oltre che rinnovati processi produttivi e lavorativi.

In una sequenza della storia umana, iniziata con i cacciatori e raccoglitori, poi con le società agricole fino alle società industriali, l’attuale Rivoluzione Digitale rappresenta il terzo grande

sconvolgimento dopo la Rivoluzione Neolitica² e la Rivoluzione Industriale. Questo nuovo sviluppo non era stato prevedibile all'inizio degli anni Ottanta, ma nel 1989 ci fu la svolta epocale quando Tim Berners-Lee e Robert Cailliau svilupparono il World Wide Web presso il centro di ricerca CERN di Ginevra, senza i quali i protocolli ipertestuali e il browser web, già esistenti allora, sarebbero rimasti principalmente una questione solo per pochi fruitori interessati.

La transizione da un mondo analogico all'era digitale, caratterizzata dall'archiviazione dei dati e delle reti digitali, descrive una cesura nei mass media e nelle tecnologie dell'informazione che appare profonda tanto quanto la transizione dall'oralità alla scrittura. Il processo di digitalizzazione di una quantità sempre maggiore di informazioni si traduce nell'evoluzione di tutti i tipi di strumenti che trasformano i legami sociali in una nuova modalità mai sperimentata prima: gli strumenti sono sempre più numerosi, sempre più efficienti, la capacità di calcolo di uno smartphone è superiore a quella di un microcomputer dei primi anni Ottanta.

Questi nuovi strumenti sono in grado di assumere diverse funzioni logiche, come l'Intelligenza Artificiale, sono sempre più piccoli, invisibili e versatili, sono addirittura in grado di essere innestati sotto la pelle o nei corpi viventi, hanno una buona capacità di integrazione nel loro ambiente (domotica) e nelle relazioni sociali, come l'interazione uomo-robot. Il loro funzionamento in rete consente ai loro utenti di effettuare ricerche complesse offrendo loro diverse possibilità, come filmare, registrare gli altri (anche senza il proprio consenso), percepire il virtuale attraverso schermi, o attraverso un sistema di *realtà aumentata* (sovrapposizione di vari elementi come suoni, immagini 2D e 3D, video computati da un sistema informatico in tempo reale). Le loro capacità future sono al momento imprevedibili e suscitano, da un lato una curiosità spinta dalla ricerca e dall'esplorazione di nuovi campi di azione, dall'altro paure per l'ignoto e per il conseguente processo di *transumanesimo*.

La digitalizzazione influenza anche le relazioni sociali e gli stili di vita: attraverso l'uso delle tecnologie delle informazioni e della comunicazione si sta modificando il modo in cui le persone interagiscono tra loro. Sebbene offra molti vantaggi, come la possibilità di rimanere in contatto sempre e ovunque, viene tuttavia percepita in modo sfavorevole per la percezione di una perdita di umanità e di contatto fisico.

Dopo circa trent'anni dell'utilizzo delle tecnologie cyborg e le conseguenze del loro adattamento a lungo termine, i mondi virtuali devono convergere necessariamente verso un ritorno alla realtà e a un'etica dei diritti umani. Il postumanesimo amplifica l'essenza del significato di una forma di

² La Rivoluzione Neolitica è un lungo periodo di cambiamento sociale e culturale della storia umana, tra 10.000 e 7.000 anni fa, associato all'invenzione indipendente della produzione alimentare attraverso la coltivazione di raccolti e l'allevamento del bestiame in alcune regioni del mondo. Segna l'inizio della Nuova età della pietra, il Neolitico.

vita apprezzata e trattata come tale, ma richiede anche una definizione più inclusiva e una maggiore risposta etico-morale alle forme di vita umane e non umane.

L'aggregazione delle diverse correnti che convergono nel postumanesimo tecnologico, nasce dal postulato che una intelligenza artificiale potrà sostituire quella degli esseri umani, anche transcendendone il loro corpo biofisico. Tuttavia, questa previsione può anche significare che un essere postumano sia una creatura che abbia un'espansione delle sue capacità fisiche e psicologiche attraverso una unità di intelligenza umana e artificiale e che abbia una coscienza caricata in un corpo estraneo o in un computer (Miccoli, 2010).

Un postumano rappresenta un ipotetico essere futuro le cui capacità di base superano radicalmente quelle degli esseri umani attuali, un essere cibernetico in stretto rapporto con le tecnologie digitali, l'ingegneria genetica e la bioingegneria.

Il postumanesimo descrive dunque un'età evolutiva dopo l'umanità, ma tenta anche di descrivere i *postumani* come ipotetici esseri viventi futuri le cui capacità superano di gran lunga quelle degli umani contemporanei. Individui *ibridi* che ricorrono alla pratica del *biohacking*, letteralmente la pratica di amalgamare la biologia con l'etica della comunità hacker. Il biohacking comprende un ampio spettro di pratiche che spazia dal fai-da-te, come per gli impianti magnetici o di microchip, al ricorso dei biologi per condurre il sequenziamento genetico, per cure mediche e nutrizionali, per l'uso di dispositivi cibernetici per le registrazioni dei dati biometrici.

Il principio dell'etica hacker risale agli anni Cinquanta e Sessanta presso il Massachusetts Institute of Technology, ma il termine venne coniato dal giornalista statunitense Steven Levy nel suo libro *Hackers: Heroes of Computer Revolution* pubblicato nel 1984.

Le linee guida etiche degli hacker permettono di osservare come i computer si siano sviluppati con i dispositivi personali utilizzati per migliorare la qualità della vita, contribuendo a rendere un essere umano superiore e quindi postumano.

3. Dal Postumano al Transumanesimo

Il postumanesimo è strettamente correlato al *transumanesimo*, una corrente di pensiero che mira a trasformare la condizione umana con l'utilizzo delle tecnologie emergenti (tecnologia dell'informazione, nanotecnologia, biotecnologia, robotica, scienze cognitive e intelligenza artificiale), per ottenere il massimo potenziale in termini di evoluzione umana, posizionando l'evoluzione biologica in secondo piano e raggiungere quindi il livello post-umano.

Il transumanesimo si fonda sull'idea che, aumentando considerevolmente le capacità intellettuali, fisiche e psicologiche umane, si riesca a superarne anche i suoi limiti per sradicare le sofferenze causate dalle malattie e ottenere l'immunità agli effetti del tempo (invecchiamento e morte) e

potenziare le possibilità di trasformarsi in esseri con capacità altamente amplificate dalla condizione naturale.

Il transumanesimo rappresenta un nuovo paradigma che studia i potenziali benefici e anche gli eventuali pericoli delle tecnologie emergenti, soprattutto in relazione agli aspetti etici e morali che ne conseguono per l'utilizzo delle biotecnologie e per lo sviluppo di impianti cibernetici. Postula che le capacità contemporanee del corpo umano non sono uniche e rappresentano solo una fase della sua costante evoluzione: anche se l'approccio tradizionale alla medicina si basa sul ripristino di un organismo malato a uno stato sano, le sue capacità dovrebbero essere utilizzate per consentire ad una persona sana di raggiungere uno stato migliore di salute. Per i transumanisti la decisione di utilizzare tali terapie dovrebbe essere una decisione esclusivamente individuale e negare tali capacità da una prospettiva utilitaristica equivarrebbe a privare le persone malate dell'opportunità di essere curate (Ferrando, 2017, pp. 51-61).

L'introduzione massiva delle nuove tecnologie richiede una ridefinizione di molte norme sociali tradizionali dal momento che questi cambiamenti potrebbero far emergere l'esistenza di esseri postumani con un aspetto completamente diverso, a causa del processo di cyborgizzazione, dell'ingegneria genetica o finanche al trasferimento della mente. Ciò comporterebbe nuovi bisogni rispetto a quelli tradizionali, a causa della rimozione degli istinti evolutivi per la riproduzione e a causa della velocità mentale e dei fattori sensoriali. I postumani potrebbero essere in grado di costruire una società molto più diversificata e potrebbero anche affrontare problemi completamente differenti da quelli degli esseri umani classici (Bostrom, 2005).

Le sostanziali differenze tra un essere postumano e un soggetto umano tradizionale includono:

- assenza di invecchiamento
- resistenza alle malattie
- capacità cognitive superiori a quelle di qualsiasi genio moderno
- capacità sensoriali potenziate, ad esempio una maggiore acuità visiva e uno spettro cromatico più ampio
- nuove capacità sensoriali disponibili.

I transumanisti sostengono che, malgrado i possibili rischi da dover identificare ed eventualmente evitare, le tecnologie digitali offrano principalmente dei potenziali benefici per il miglioramento dell'umanità nel suo complesso eliminando le malattie, le disabilità, la povertà e la malnutrizione. Esplorano il potenziale delle nuove tecnologie e dei sistemi sociali innovativi anche per migliorare la qualità della vita intesa come ecosistema e garantire l'uguaglianza sociale nell'accesso ai beni materiali.

Considerando che il ritmo del progresso tecnologico stia sempre più accelerando, i fautori del transumanesimo ipotizzano che i prossimi decenni saranno fautori di radicali trasformazioni e

che gli esseri umani possano e debbano diventare sempre più umani, ma attraverso l'uso dell'ingegneria genetica, la nanotecnologia molecolare, la *protesica*³ sempre più sofisticata e personalizzata per ogni singola esigenza, fino alla creazione di una interfaccia diretta cervello-computer.

Il postumanesimo affonda le sue radici nell'umanesimo secolare e pur postulando che non vi siano forze soprannaturali che possano influenzare o guidare l'umanità, sostiene che esiste un imperativo etico per gli esseri umani: impegnarsi per il progresso e per migliorare la propria esistenza. Per poter raggiungere questo obiettivo, i transumanisti adottano approcci interdisciplinari per comprendere le possibilità di superare i limiti biologici, avvalendosi di differenti rami della scienza, della filosofia, dell'economia, della storia, della sociologia e della psicologia.

Dal transumanesimo si è poi proposta una serie di sottosistemi:

- *Immortalismo* – una corrente intellettuale fondata sull'affermazione che fino ad ora non è stato scoperto nulla in biologia che possa condurre direttamente all'inevitabilità della morte. Obiettivo dell'immortalismo è quello di disattivare il meccanismo di autodegradazione umana attraverso l'utilizzo dell'ingegneria genetica, delle cellule staminali e della nanotecnologia per poter raggiungere l'immortalità;
- *Estropianismo* – un ramo del transumanesimo emerso già nei primi anni Ottanta, con l'obiettivo di creare una realtà priva di qualsiasi segno di regressione, utilizzando la tecnologia per stimolare il continuo sviluppo e progresso dell'umanità. Il fondatore di questa corrente è il filosofo e futurologo anglosassone Max More (nato 1964) autore di saggi sul processo decisionale nelle tecnologie emergenti (More, 2015).
- *Singularismo* – movimento fondato sul presupposto che l'era umana terminerà con l'arrivo della singolarità, ovvero con l'emergere di una superintelligenza sovrumana. I suoi sostenitori si sforzano di creare la superintelligenza attraverso la biotecnologia, la nanotecnologia e la scienza cognitiva. Secondo l'informatico, scrittore e futurologo statunitense, Ray Kurzweil (nato 1948) prevede l'arrivo della singolarità nel 2045.
- *Imperativo edonistico* – un sottosistema del transumanesimo che si concentra sul raggiungimento della felicità eterna per la specie umana, senza vincoli biologici. L'obiettivo è quello di ridurre al minimo le sofferenze per gli esseri coscienti. Questo movimento è stato fondato dal filosofo britannico David Pearce (nato 1959) che ha pubblicato un manifesto abolizionista, *The Hedonistic Imperative*, nel 1995.

³ La protesica è un campo della conoscenza scientifica che fonde la medicina e la tecnologia e si occupa della sostituzione di tessuti, organi o difetti funzionali mediante materiale sostitutivo artificiale, protesi.

- *Postgenerismo* – movimento che si concentra sulle questioni del postgender e della postsessualità. Questi teorici attingono alle ricerche di genere, degli studi queer e di alcuni aspetti del femminismo. L'obiettivo principale è quello di rifiutare le categorie di genere e creare esseri androgini.

- *Transumanesimo ecologico* – movimento che si concentra sull'uso delle tecnologie avanzate per migliorare l'ambiente. Il principale postulato è la terraformazione dei corpi celesti e la creazione di biosfere artificiali adattate a vecchie e nuove forme di vita, prodotte tecnologicamente.

Fino ad ora la vita umana è stata caratterizzata dalla brevità, dalle ingiustizie, dalla miseria, il paradigma del transumanesimo sostiene che la specie umana può essere in qualche modo sostituita da un altro essere che trascende sé stesso.

Questa concezione del futuro conduce inevitabilmente ad una riflessione critica, non solo sugli aspetti tecnologici del progresso scientifico, ma anche su tutte quelle questioni di carattere etico e filosofico che convergono nel dualismo corpo-mente di cartesiana memoria.

4. Riflessioni conclusive

Considerare il transumanesimo come una forma speciale di postumanesimo, oppure considerarle due distinte e inconciliabili correnti filosofiche, resta sempre un'aperta discussione in molti articoli scientifici. Nel contesto di questo dilemma, è tuttavia chiaro che ci sono ragioni per credere che Friedrich Nietzsche (1844-1900) possa essere stato il precursore sia del transumanesimo e sia del postumanesimo con la sua idea di *superuomo*, oltre l'uomo in senso biologico, superiore perché in grado di superare le convenzioni sociali e morali affermando sé stesso (Nietzsche, 1986). Ciò avviene superando la morte di Dio, intesa come crisi dei valori, cosicché l'essere umano possa creare indipendentemente una propria esistenza. Il superuomo rappresenta quindi una forza di volontà, un impulso a creare una nuova vita con nuovi sistemi di valori.

L'immagine nietzschiana del superuomo, pur nascondendo un dramma esistenziale, rappresenta tuttavia un tipo biologico superiore, un postumano e/o un transumano, come risposta alla morte di Dio che fa ricorso alla scienza e alle tecnologie per superare i limiti umani e migliorare le sue capacità fisiche e mentali: un essere umano oltre l'umano.

Se le tecnologie da un lato aiutano gli esseri umani a superare alcuni ostacoli, dall'altro si possono correre dei rischi a lungo termine sugli effetti collaterali dell'adattamento alle cyborg tecnologie. Ad esempio, cosa potrebbe accadere dopo vent'anni di utilizzo costante di tecnologie mediate dal computer e la loro successiva eliminazione e dell'adattamento a lungo termine alle tecnologie virtuali, mondi seguiti da un eventuale ritorno alla realtà, oppure entrare in un mondo ancora più

virtuale dove l'etica post-cyborg venga associata all'eliminazione forzata delle tecnologie cyborg da parte delle autorità.

Secondo il pensiero transumanista, un postumano è un ipotetico essere futuro, dotato di capacità di base che superano radicalmente quelle degli esseri umani odierni, che non potranno essere considerate inequivocabilmente umani secondo gli standard attuali. I postumani si concentrano principalmente sulla cibernetica e sul rapporto con la tecnologia digitale e la teoria evolutiva della mente (ETM) che consentirebbe lo sviluppo di cervelli senzienti. L'enfasi principale è sui sistemi. Il transumanesimo invece si concentra sulla modifica della specie umana attraverso qualsiasi tipologia di scienza emergente, come l'ingegneria genetica, la tecnologia digitale e la bioingegneria, ma non affronta adeguatamente le preoccupazioni sull'evoluzione dell'umanesimo. I postumani potrebbero essere intelligenze artificiali totalmente sintetiche, oppure una simbiosi di intelligenza umana e artificiale, oppure un cyborg, oppure una combinazione di tecnologie e terapie di estensione della vita, interfacce neurali e trasformarsi in esseri fisicamente e mentalmente potenti. Ma potrebbero semplicemente diventare particolarmente intelligenti e tecnologicamente sofisticati che il loro comportamento sarebbe percepito incomprensibile per gli esseri umani moderni, a causa della loro intelligenza e immaginazione limitate; potrebbero cercare di fornire un significato più pieno alla propria esistenza credendo nella possibilità di collegarsi con un regno superiore, oltre i propri limiti, per fondersi o entrare in contatto con l'Infinito e l'Eterno.

O più banalmente, tutte le tecnologie esistenti e future, potrebbero essere il modo più efficace per migliorare la società e che le soluzioni tecnologiche possano essere utilizzate per ideali progressisti della democrazia e per l'uguaglianza umana. O anche il suo esatto contrario, essere un tentativo di alterare lo stato naturale dell'uomo, come ad esempio la clonazione o la manipolazione del DNA umano.

Bibliografia

- Alfano Miglietti, F. (1997). *Identità mutanti. Dalla piega alla piaga: esseri delle contaminazioni contemporanee*. Bruno Mondadori.
- Bostrom, N. (2005). In Defence of Posthuman Dignity. *Bioethics*, vol. 19, nr. 3, 202-214.
- Braidotti, R. (2014). *Il postumano: la vita oltre l'individuo, oltre la specie, oltre la morte* (A. Balzano, trad.). Derive Approdi.

- Caronia, A. (2020). *Dal cyborg al postumano. Biopolitica del corpo artificiale*. Meltemi.
- DeFleur, M. & Ball-Rokeach, S. (1995). *Teorie delle comunicazioni di massa*. Il Mulino.
- Dery, M. (1996). *Velocità di fuga: Cyberculture a fine millennio*. Feltrinelli.
- De Landa, M. (1996). *La guerra nell'era delle macchine intelligenti*. Feltrinelli.
- De Landa, M. (2002). *Ctheory. Interview with Manuel De landa, 1000 Years of War*. Manuel De Landa in conversation with: Don Ihde, Casper Bruun Jensen, Jari Friis Jorgensen, Srikanth Mallavarapu, Eduardo Mendieta, John Mix, John Protevi, and Evan Selinger. <https://journals.uvic.ca/index.php/ctheory/article/view/14560/5406>
- Ferrando, F. (2017). Postumanesimo, transumanesimo, antiumanesimo, metaumanesimo e nuovo materialismo. Relazioni e differenze. *Lo Sguardo - rivista di filosofia*, nr. 24, II: *Limiti e confini del postumano*, 51-61.
- Ferrante, A. (2017). *Dialoghi sul postumano*. Mimesis.
- Fidler, R. (2000). *Mediamorfosi. Comprendere i nuovi media*. Guerini.
- Hayles, N.K. (1999). *How We Became Posthuman: Virtual Bodies in Cybernetics, Literature, and Informatics*. University of Chicago Press.
- Haraway, D. (1991). Situated Knowledges. In D. Haraway, *Simians, Cyborgs, and Women* (pp. 183-202). Routledge.
- Haraway, D. (2018). *Manifesto cyborg. Donne, tecnologie e biopolitiche del corpo*. Feltrinelli.
- Horkheimer, M. & Marcuse, H. (2003). *Filosofia e teoria critica* (S. Petrucciani, a cura di). Einaudi.
- Levy, S. (1984). *Hackers: Heroes of the Computer Revolution*. Anchor Press/Doubleday.
- Lucci, A. (2016). *Umano post umano: immagini dalla fine della storia*. Inschibboleth.
- Marchesini, R. (2002). *Post-human*. Bollati Boringhieri.
- Marchesini, R. (2009). *Il tramonto dell'uomo: la prospettiva post-umanista*. Dedalo.
- McLuhan, M. (1967). *Gli strumenti del comunicare*. Il Saggiatore.
- Miccoli, A. (2010). *Posthuman Suffering and the Technological Embrace*. Lexington Books.
- More, M. (1990-1996). Transhumanism Towards a Futurist Philosophy, in <http://www.maxmore.com/transhum.htm>. Ultima visita settembre 2025.
- Moscovici, S. (2005). *Le rappresentazioni sociali*. Il Mulino.
- Nietzsche, F. (1986). *Così parlò Zarathustra. Un libro per tutti e per nessuno*. Adelphi.
- Pearce, D. (2007). *Text adapted from invited talks given at the FHI (Oxford University) and the Charity International Happiness Conference* in <https://www.hedweb.com/abolitionist-project/index.html>. Ultima visita settembre 2025.
- Pepperell, R. (2003). *The Posthuman Condition: Consciousness Beyond the Brain*. Intellect Books.
- Salerno, V. (2021). Postumanesimo e transumanesimo: le tecnologie convergenti e l'immaginario educativo cyborg. *QTimes – webmagazine*, a. XIII, n. 3.

Scalisi, R. (2001). *Storia dell'interazione uomo-macchina dai main-frame ai computer indossabili*. Guerini.

Silverstone, R. (2008). *Mediapolis. La responsabilità dei media nella civiltà globale*. Vita e Pensiero.

Vatinno, G. (2010). *Il Transumanesimo, una nuova filosofia per l'Uomo del XXI secolo*. Armando Editore.

Zizek, S. (2021). *Hegel e il cervello postumano*. Ponte alle Grazie.