

ARCHIVIO TEOLOGICO TORINESE



2024/1

gennaio-giugno 2024 • Anno XXX • Numero 1

Rivista della FACOLTÀ TEOLOGICA DELL'ITALIA SETTENTRIONALE
SEZIONE DI TORINO

INTELLIGENZA ARTIFICIALE (E DINTORNI)
ALLA PROVA DI FILOSOFIA E TEOLOGIA

Nerbini

ARCHIVIO TEOLOGICO TORINESE

A cura della Facoltà Teologica dell'Italia Settentrionale – Sezione di Torino
Anno XXX – 2024, n. 1

Proprietà:

Fondazione Polo Teologico Torinese

Facoltà Teologica dell'Italia Settentrionale – Sezione di Torino

Via XX Settembre, 83 – 10122 Torino

tel. 011 4360249 – fax 011 4319338

istituzionale@teologitorino.it

e-mail Segreteria: donandrea.pacini@gmail.com

Registrazione n. 1 presso il Tribunale di Torino del 27 gennaio 2015

Direttore responsabile: Mauro Grosso

Redazione: Andrea Pacini (direttore), Gian Luca Carrega e Antonio Sacco (segretari), Oreste Aime, Dino Barberis, Roberto Carelli, Ferruccio Ceragioli, Carla Corbella, Mauro Grosso, Pier Davide Guenzi, Luca Margaria, Paolo Mirabella, Alberto Nigra, Alberto Piola

Editore:

Edizioni Nerbini - Prohemio Editoriale srl

via G.B. Vico 11 - 50136 Firenze - ROC n. 34429 (10.6.2020)

e-mail: edizioni@nerbini.it

www.nerbini.it

Realizzazione editoriale e stampa: Prohemio Editoriale srl - via G.B. Vico 11 - 50136 Firenze

Amministrazione e ufficio abbonamenti:

abbonamenti@nerbini.it

ABBONAMENTO 2024

Italia € 44,50 – Europa € 64,50 – Resto del mondo € 74,50

Una copia: € 27,00

Per gli abbonamenti e l'acquisto di singoli fascicoli dal 2022 in poi:

Versamento sul c.c.p. 1015092776

intestato a Prohemio Editoriale srl, Firenze

Sommario

Intelligenza artificiale (e dintorni) alla prova di filosofia e teologia

Introduzione

Mauro Grosso – Luca Peyron » 7

Uomo e tecnica.

Spunti per una riflessione nel pensiero medievale
Amos Corbini » 13

Dal mondo al dato, dal dato al codice.

*Sulla necessità di una teoria della conoscenza
e del linguaggio nel rapporto con il mondo*
Luca Margaria » 35

Tra umano e digitale: un contributo dalla metafisica

Mauro Grosso » 55

Senza entrare in competizione:

intelligenza umana e intelligenza artificiale
Alberto Piola » 73

La teologia morale alla prova del mondo digitale

Alessandro Picchiarelli » 89

Il *capitalismo dell'intelligenza artificiale* (IA)

Antonio Sacco » 107

Lavorare e scrivere con le proprie mani: tecnica e tecnologia al servizio della missione paolina	
<i>Gian Luca Carrega</i>	» 129
I padri della Chiesa e la «tecnologia»: fra giudizio (<i>krisis</i>) e buon uso (<i>chrēsis</i>)	
<i>Alberto Nigra</i>	» 145
Dalla soggettività all'oggettività: la filosofia di Bernard Lonergan come fondamento per il design sensibile ai valori	
<i>Steven Umbrello</i>	» 161
Intelligenza artificiale e medicina: sfide tecniche ed etiche	
<i>Alessandro Mantini</i>	» 173
Teologia dell'educazione. Come educare al tempo dell'IA, come insegnare teologia al tempo dell'IA	
<i>Marco Sanavio</i>	» 199

RECENSIONI

M. FERRARIS – G. SARACCO, <i>Tecnosofia. Tecnologia e umanesimo per una scienza nuova</i> (O. Aime)	» 217
L. PEYRON, <i>Incarnazione digitale. Custodire l'umano nell'infosfera</i> (C. Corbella)	» 220
Y. BERIO RAPETTI, <i>La società senza sguardo. Divinizzazione della tecnica nell'era della teocnocrazia</i> (M. Grossi)	» 222
P. BENANTI <i>Human in the Loop. Decisioni umane e intelligenze artificiali</i> (P. Simonini)	» 226
J.C. DE MARTIN, <i>Contro lo smartphone. Per una tecnologia più democratica</i> (P. Simonini)	» 230
L. FLORIDI, <i>Etica dell'intelligenza artificiale. Sviluppi, opportunità, sfide</i> (G. Zeppegno)	» 233
M. PRIOTTO, <i>L'itinerario geografico-teologico dei patriarchi di Israele (Gen 11–50)</i> (G. Galvagno)	» 236

B. KOWALCZYK, <i>La «Vetus Syra» del vangelo di Marco. Commento e traduzione</i> (G.L. Carrega)	»	238
T. HALÍK, <i>Pomeriggio del cristianesimo. Il coraggio di cambiare</i> (O. Aime)	»	242
E. IULA, <i>La pazienza del vasaio. La riparazione a confronto con la modernità</i> (P. Mirabella)	»	245
H. DE LUBAC – H.U. von BALTHASAR, <i>Conversazioni sulla Chiesa. Interviste di Angelo Scola</i> , a cura di J.-R. ARMOGATHE (L. Casto)	»	248
M.V. CERUTTI (a cura di), <i>Allo specchio dell'altro. Strategie di resilienza di «pagani» e gnostici tra II e IV secolo d.C.</i> (A. Nigra)	»	254
L. BERZANO, <i>Senza più la domenica. Viaggio nella spiritualità secolarizzata</i> (O. Aime)	»	260
M. CONDÉ, <i>Il vangelo del nuovo mondo</i> (M. Nisii)	»	263

SCHEDE

G. PALESTRO – M. ROSSINO – G. ZEPPEGNO, <i>Uomo e ambiente. Movimenti ambientalisti e proposta cristiana a confronto</i> (F. Casazza)	»	269
S. RONDINARA (a cura di), <i>Metodo</i> (A. Piola)	»	270

Intelligenza artificiale e medicina: sfide tecniche ed etiche

Alessandro Mantini

Introduzione... a largo spettro

Intelligenza artificiale (IA) è un'espressione che si è cominciata ad affermare a partire dal 1956 in vari ambiti scientifici come una nuova sfida al pensiero, all'agire e alla razionalità umana. Tradizionalmente si parla infatti di IA in relazione allo sviluppo di algoritmi e strumenti software «intelligenti» (implementabili poi in strutture robotiche), nella prospettiva di realizzare processi quali: pensare umanamente (*thinking humanly*), pensare razionalmente (*thinking rationally*), agire umanamente (*acting humanly*) e agire razionalmente (*acting rationally*).¹

L'uso della parola «intelligenza», troppo audacemente accostato a quella, assai diversa, di «artificiale», ci invita a porre in evidenza i contorni del suo inquadramento semantico proprio per evitare confusioni o fraintendimenti. L'intelligenza umana si colloca nel dinamismo dell'immagine e della somiglianza con Dio (*imago Dei*), secondo i tratti dell'*imago Christi*, ove Gesù il Cristo è la Rivelazione e l'Incarnazione del *Verbo-Logos*, Persona reale, accessibile e celeste, secondo la logica della *ratio* e del *verbum*.² Dall'*imago Trinitatis*, dunque, in cui si inscrivono gli orizzonti dell'intelligenza e del realismo delle scienze proprio alla luce dell'incarnazione nella fruttuosa potenza dello Spirito Santo, l'uomo trae origine, significato e vita, per esprimersi, in libertà e in responsabilità, nel continuo collegamento tra il particolare e l'universale, il parziale e il tutto, in una visione sempre nuova e superiore, «complessa», creativa e partecipativa. Pertanto, tre tratti che caratterizzano l'intelligenza umana sono: l'*intus-legere* (leg-

¹ Cf. S.J. RUSSELL – P. NORVIG, *Artificial intelligence. A Modern Approach*, Pearson, Harlow 2016.

² Cf. G. TANZELLA-NITTI, *Gesù Cristo, Rivelazione e Incarnazione del Logos*, in G. TANZELLA-NITTI – A. STRUMIA (a cura di), *Dizionario Interdisciplinare di Scienza e Fede*, vol. 1, Città Nuova, Roma 2002, 695.

gere e scrutare in profondità, tra fisica e meta-fisica); il *cum-prehendere* (afferrare con la mente e collegare insieme – partecipare); il *cum-gnoscere* (conoscere, riconoscere e valutare – percezione e pensiero).

Ricollocare l'intelligenza nell'ordine della creazione, con tutta la fiducia che la santissima Trinità, nell'atto di donarla, conferisce all'uomo, permette dunque alla persona umana di realizzare quel duplice movimento estatico dell'amore di Dio, rappresentato rispettivamente:

- dalla *partecipazione creaturale* all'intelligenza creatrice (cf. Gen 1,26);
- dalla *responsabilità creaturale* nel contesto di una intelligenza obbediente e illuminata (cf. Gen 2,16-17).

In questo orizzonte, l'intelligenza dapprima viene liberata dall'autosufficienza che, essendo una menzogna, la costringe e la limita nella sua espressività, e successivamente viene spinta verso le alte vette della Sapienza, perché sostenuta dallo zampillare dello Spirito Santo che la rende capace di contemplare (cf. 1Gv 1,1), da *cum-templum*. Quest'ultimo è il quarto, decisivo e omnicomprensivo significato di cui l'intelligenza umana naturalmente vive e al quale costantemente tende e si riferisce. Essa, infatti, partecipa responsabilmente dello «stare» dinamico e comunionale con Dio, nel suo Tempio, ove la santissima Trinità coinvolge l'uomo in una vitalità adorante, riconoscente, generante e quindi sapiente (i sette doni dello Spirito Santo mostrano proprio questa ampiezza di prospettive). Proprio l'incarnazione di Gesù il Cristo rende partecipe l'uomo dell'intelligibilità e del mistero, nello stupore, nella meraviglia e nell'umiltà (cf. Sal 26,4).

Da questo sintetico inquadramento, conseguie che l'espressione «intelligenza artificiale» è in realtà un ossimoro nelle due parole che la compongono: l'artificiale è un arte-fatto, un «prodotto statico» (perché non creaturale e non relazionale) delle mani dell'uomo, l'intelligenza invece è un «frutto dinamico» della relazione con Dio e partecipa della sua immagine e somiglianza dalle quali riceve, senza mai poterla comunicare ad alcunché di artificiale, la luce della sua espressività. L'intelligenza, infatti è per sua natura «incomunicabile», in quanto strettamente legata al mondo spirituale e interiore dell'anima e della coscienza, quest'ultima mirabilmente definita come «il nucleo più segreto e il sacrario dell'uomo, dove egli è solo con Dio, la cui voce risuona nell'intimità».³

Credo pertanto che si debba più propriamente parlare di *elaborazione artificiale*, ma, pur intendendo questo, per semplicità in quanto segue manterrò comunque la dicitura classica, aggiungendo però il corsi-

³ CONCILIO VATICANO II, *Gaudium et spes*, 7 dicembre 1965, n. 16.

vo, *intelligenza* artificiale (o in forma abbreviata scrivendo *IA*). In questo modo si vuole richiamare una comprensione meno incantata o fatalista, ma più antropocentrica e ragionata, per indicare proprio quello che essa realmente è, cioè una elaborazione complessa di dati attraverso algoritmi matematici e statistici molto avanzati. L'*IA* quindi deriva e rimanda all'intelligenza umana perché quest'ultima si sviluppi in modo più esteso e ampio, creativo e profondo, mantenendo chiara la sua distinzione e la sua centralità.

Con questa introduzione terminologica e semantica, entrerò ora direttamente nel contesto dell'*IA* applicata allo specifico settore della medicina e chirurgia per evidenziarne rispettivamente alcuni tratti tecnici, applicativi ed etici, utili per futuri approfondimenti.

1. Aspetti tecnici dell'*intelligenza artificiale* in medicina e chirurgia

Tutti gli aspetti relativi alla gestione dei «dati» provenienti dagli esami clinici, tecnici e anamnestici richiedono per loro stessa natura (cioè, per il fatto che si riferiscono a persone umane nella loro «complessità» vitale) un approccio trans-disciplinare, per non ridurre la persona ad una somma di dati, per quanto elaborati. I dati ed il loro contenuto informativo vanno considerati piuttosto all'interno di un contesto nel quale le discipline che aiutano a decifrarli possano interagire nell'ottica di un servizio alla persona integrale (cioè considerata nei suoi aspetti umano-personale, corporeo-clinico, interiore-spirituale, relazionale-familiare). La sinergia complessa e trans-disciplinare tra algoritmo ed etica è, come vedremo, proprio un primo esempio di quanto detto. L'etica è strettamente connessa infatti con l'antropologia e la sua fondazione, ed entrambe poggianno sulla teologia rivelata che ne illumina non solo l'approfondimento⁴ ma anche l'orizzonte vitale, secondo una visione prospettica integrale. La nuova «rivoluzione scientifica» in atto richiama dunque le discipline a interagire sapientemente avendo la persona umana come centro «centrato», cioè come primo destinatario nell'orizzonte del «servizio» che a sua volta si colloca in una prospettiva interamente cristologica. Questo è l'auspicio che la teologia cattolica può formulare, offrendo il suo contributo sapienziale per lo svi-

⁴ Cf. *Gaudium et spes*, n. 22: «Cristo, che è il nuovo Adamo, proprio rivelando il mistero del Padre e del suo amore svela anche pienamente l'uomo a se stesso e gli manifesta la sua altissima vocazione».

luppo di una «medicina personalizzata», secondo il più che mai necessario trinomio:

(tecnologia ↔ medicina ↔ teologia) → servizio alla persona umana

Quando ci si riferisce alla persona umana, e quando la si riconosce «centrata» in Cristo, entrano in gioco con maggiore evidenza tutte le altre prospettive oltre a quella biologica e corporea, quali l'interiorità e la coscienza, quindi l'anima, nonché la relazionalità e la sensibilità, tutte uniche e irripetibili. Questi ultimi ed essenziali aspetti, riferiti alla persona umana come unitotalità unificata di anima e corpo, rendono ancora più complessa la gestione dei «dati» in medicina, trattandosi ogni volta di casi unici, per quanto pur sempre analizzabili anche al livello di clusterizzazioni diagnostiche e cliniche. Non va mai dimenticato, infatti, nell'uso dell'IA e quindi dei *Big Data* in medicina, il doppio dinamismo dell'unicità della singola persona (dimensione verticale) e della conformità dei riscontri clinici tra le persone (dimensione orizzontale), che solo potrà favorire un uso efficace della tecnologia, per il bene ultimo proprio della persona:

unicità (↓) + conformità (↔) → servizio al bene della persona

1.1. Il «pasto» dell'IA: ovvero come gestire i Big Data nella medicina

Affacciandoci ora sul fronte tecnico, per prima cosa incontriamo la «materia prima» dell'IA, che poi, a ben guardare, è una «immateria prima» in quanto si tratta solo di dati estratti dal mondo reale con procedimenti di campionatura, e connotati appunto dalla loro immaterialità. Questo è a buon titolo il «pasto» dell'IA, cioè quanto è necessario affinché gli algoritmi possano sviluppare le loro elaborazioni e produrre i loro risultati.

I *Big Data* che abbiamo a disposizione in medicina sono ormai dell'ordine di grandezza dei ZB (Zettabyte = 10^{21} byte). Tale abbondanza non basta che sia raccolta, ma occorre che sia anche «organizzata» e «classificata» per poter essere a sua volta «gestita» dall'IA, e per poter poi essere trattata al fine di estrarne il contenuto informativo. La gestione e l'utilizzo dei dati clinici, infatti, sono aspetti estremamente delicati e sensibili.

La comunità internazionale ha dunque messo a disposizione diversi strumenti per regolamentare, organizzare e gestire i dati nel contesto della *Health-Care*,⁵ distinguendo dapprima tra «dati strutturati» e dati «non strutturati»:

⁵ Per quanto segue, cf. A.C. CHANG, *Intelligent-Based Medicine. Artificial Intelligence and Human Cognition in clinical Medicine and Healthcare*, Academic Press, London 2020, 47-65.

- *Electronic Medical Records* (EMRs): tipicamente dati strutturati dal punto di vista elettronico, attualmente presentati in tabelle o in *database* relazionali;
- *Advanced Imaging Studies* (AIS): dati non strutturati (risonanze magnetiche, immagini radiografiche, scansioni di tomografia computerizzata, ecocardiogrammi, angiogrammi, ecc.);
- altri dati non strutturati: tutti quei dati cartacei, in numero largamente maggioritario, attualmente disponibili.

Una volta raccolti e categorizzati in opportuni *database*, una seconda esigenza è quella di uniformare tali dati ai fini di una loro standardizzazione che li renda più fruibili per essere confrontati tra loro e quindi utilizzati e valutati in rete. Sono stati dunque elaborati alcuni standard internazionali per gli *Health-Care Data* per consentire una efficace interoperabilità:

- *International Statistical Classification of the Diseases and Related Health Problems* (ICD), che offre tutta una serie di classificazioni per le varie diagnosi (corredate di corrispondenti segni e sintomi) con codici specifici che permettono di identificarle;
- *Current Procedural Terminology* (CPT) *code*, che identifica i tipi di servizi connessi e conseguenti alla diagnosi;
- *SNOMED Clinical Terms* (CT), che contiene una raccolta di insiemi di codici di concetti, descrizioni, relazioni e riferimenti clinici;
- *Unified Medical Language System* (UMLS), che fornisce un sistema di interoperabilità per informazioni biomediche, tra cui gli *Electronic Health Records* (EHR);
- *Logical Observation Identifiers Names and Codes* (LOINC): una sorta di dizionario per le osservazioni e i dati di laboratorio;
- *Digital Imaging and Communications in Medicine* (DICOM): utile per la comunicazione e la gestione dei dati provenienti da immagini digitalizzate;
- *Picture Archive and Communication System* (PACS), che rappresenta una *computer network* di informazioni legate alle immagini digitalizzate, importante per la loro trasmissione, revisione, archiviazione e conservazione.

I *Big Data*, come è noto, si descrivono con le famose quattro V che li caratterizzano: *Volume*, *Variety*, *Velocity* e *Veracity* (cui possono aggiungersi anche: *Value*, *Visualization* e *Variability*), non essendo infatti solo puntuali, ma anche raccolti nel tempo (per cui diventano dati estremamente abbondanti e complessi). In questi casi (*holter* pressorio o cardiaco, tracciati elettrocardiografici, la respirazione di un paziente intubato, la valutazione del rischio di sepsi, ecc.) si tende a valutare non solo l'intera «storia» raccolta, ma in modo particolare le osservazioni immediatamente

precedenti al momento di volta in volta analizzato (si utilizzano allora i cosiddetti *Hidden Markov Models* – HMM). Questi ultimi stimano variabili non sempre evidenti ma spesso nascoste o sconosciute, che abbiano un effetto sull'osservazione in corso, da valutare sulla base di modelli probabilistici in stretta relazione tra loro.

I concetti fondamentali in gioco sono di conseguenza: il *data size*, cioè le dimensioni di questi dati; il *data location*, quindi il loro formato e la loro raccolta; la *data structure*, quindi la qualità strutturata o non strutturata dei dati (l'80% attualmente è non strutturata); la *data integrity*, che rappresenta un problema enorme, legato a dati eventualmente mancanti, oppure non appropriatamente segnalati; la *data consistency* in riferimento a quei dati che hanno bisogno di essere registrati e ridefiniti in maniera adeguata a causa della loro iniziale inconsistenza.

Dopo aver affrontato il problema dell'*Health-Care Database* subentra la necessità dell'*Health-Care Data Management*, cioè della loro effettiva gestione. Si usa parlare tipicamente, a tal proposito, di ETL (*Extract – Transform – Load*),⁶ per indicare, cioè, l'importanza dell'estrazione, della trasformazione e della conservazione dei dati, al fine di poterne disporre efficacemente.

Uno dei progetti attualmente in fase di sviluppo è quello di creare una specie di «mondo virtuale» (su server o su strutture distribuite chiamati *DataBase Management System* – DBMS), nel quale tutti possano poi accedere a questi dati (una volta standardizzati) nella logica dell'interoperabilità. Quest'ultimo concetto oggi non solo è estremamente interessante dal punto di vista della possibilità di crescita nella ricerca scientifica e nell'apprendimento dei sistemi esperti basati sull'IA, ma è anche necessario, proprio nell'orizzonte di una ricerca che non può più essere mono-disciplinare ma sempre più trans-disciplinare e quindi interattiva e relazionale.

Per raccogliere, strutturare, rappresentare e memorizzare i *Big Data* in medicina ci si riferisce dunque ai DBMS come, ad esempio, l'*Health Level-7* (HL7), la cui idea è quella di costruire un mondo nel quale ognuno possa avere un accesso sicuro e giusto per l'utilizzo dei dati clinici, ovunque e sempre questo fosse necessario. L'*Interoperability* degli EHRs nell'ottica di *Multi-institutional Collaborations* richiede infatti la «capacità di diversi sistemi informativi, dispositivi o applicazioni di connettersi in modo coordinato, all'interno e al di là dei confini organizzativi, per accede-

⁶ Vale la pena menzionare i *Data Warehouse* per dati essenzialmente strutturati, i *Data Lake* per dati grezzi e non strutturati, ed i *Data Reservoir*, per raccogliere entrambe le tipologie precedenti.

re, scambiare e utilizzare in modo cooperativo i dati tra le parti interessate con l'obiettivo di ottimizzare la salute degli individui e delle popolazioni».⁷

I DBMS possono essere a loro volta strutturati come «modelli classici relazionali» che si compongono essenzialmente di singole entità, o come modelli *object oriented* i cui elementi sono più complessi comprendendo diverse specifiche delle singole entità. Questi ultimi sono molto più performanti in quanto offrono la possibilità di far interagire, e quindi di memorizzare, anche le relazioni tra i dati (non solo quindi tabelle *Excel* ma «oggetti» composti connessi tra loro). Infine, ci sono i «modelli a grafo», che sono ancora più performanti, perché prevedono l'ulteriore distinzione tra nodi e relazioni, e contengono più informazioni e più possibilità di esplorazione. I modelli a grafo sono particolarmente indicati per *database* in costante evoluzione quali quelli propri della medicina. Attualmente la maggior parte dei *database* sono di tipo relazionale, mentre la loro conversione in modelli a grafo sarebbe auspicabile, proprio in vista del loro utilizzo per i nuovi algoritmi di *Deep Learning* e delle architetture «cognitive».

Se è indubbio il fascino immediato dell'IA nel suo manifestarsi al grande pubblico, in questo momento di forte ascesa e pervasività del suo influsso nell'attuale contesto tecnologico, è anche vero che essa non può in alcun modo prescindere da questa preziosa premessa che riguarda la raccolta e la gestione dei dati, i quali appunto sono proprio il «pasto» dell'IA. Questo pasto va «servito cotto», il che vuol dire che i ricercatori devono spendere molto tempo per trasformare i dati in formati realmente utilizzabili dagli algoritmi di IA in vista di un loro utilizzo significativo e quanto più affidabile possibile.

1.2. Le «abilità» dell'IA: ovvero come potenzia la medicina?

Le strutture tradizionali di IA per le applicazioni in medicina e chirurgia utilizzano la tecnica del *Machine Learning* (ML), nelle sue distinzioni in operazioni di «classificazione» (che assegnano un dato a una classe o categoria di dati) o di «regressione» (in cui i risultati sono gruppi continui, come per esempio la stima delle dimensioni di un tumore). In entrambi i casi si ricorre a strumenti di algebra lineare e non lineare.

Per consentire l'«apprendimento», i dati raccolti vengono suddivisi in dati di *training* (usati per addestrare l'algoritmo), dati di *test* (usati appunto per verificare) e dati di misura (usati finalmente per la valutazione dei casi). È inoltre necessario determinare il tipo di implementazione, gli iper-parametri in gioco (tasso di apprendimento, numero di neuroni artifi-

⁷ Cf. CHANG, *Intelligent-Based Medicine*, 54 (mia traduzione).

ciali, ecc.), il pre-processamento dei dati richiesto prima delle misure vere e proprie, ecc.

Accanto al ML tradizionale vi sono poi altre strutture di apprendimento (tipo *Deep Learning* – DL) che fanno uso delle *Neural Networks* (NN), e che meglio si applicano al caso di dati con uno sviluppo temporale dinamico e continuo (quali ad esempio i filmati intraoperatori). Si tratta di reti interconnesse di neuroni artificiali composte da un *layer* di neuroni di ingresso, uno o più *layer* intermedi (o nascosti – *hidden layers*) e un *layer* di uscita, tutti di numerosità variabile a seconda dei casi e della potenza di calcolo disponibile. Particolarmente interessanti sono nello specifico, le *Deep Neural Networks*, tra cui le *Recurrent Neural Networks* (RNN) e le *Convolutional Neural Networks* (CNN).

Un problema attualmente non trascurabile è quello della stesura dei *reports* degli interventi chirurgici, i quali, in vista del loro utilizzo con il *Deep Learning*, richiedono ai chirurghi molto tempo affinché presentino una adeguata qualità e uniformità, senza la quale spesso non sono utilizzabili con l'IA.

Prima di entrare in maggiore dettaglio nell'analisi delle potenzialità dell'IA applicata alla medicina, vale la pena accennare al problema della *privacy* e della *security* per le persone coinvolte. Infatti, occorre fare ricorso a modelli capaci di analizzare i dati raccolti in formati tali da preservare l'identità dei pazienti (ad esempio si ricorre a tecniche di DL per il riconoscimento facciale al fine di eliminare i volti nelle riprese delle sale operatorie). Occorre fare inoltre sempre molta attenzione ai dati provenienti da fonti diverse, al fine di evitare o ridurre possibili *biases* (condizionamenti dei dati stessi a opera dell'operatore umano che li raccoglie e li assegna all'algoritmo). Per tale motivo, una volta ottenuti dei risultati è sempre importante riportare una descrizione puntuale del *set* dei dati in ingresso, della loro classificazione e dei procedimenti applicati ad essi prima dell'addestramento. Questi aspetti aggiungono naturalmente ulteriori e non trascurabili tratti alla complessità della gestione dei dati clinici.

Vista allora la grande quantità di dati e il conseguente impegno necessario per il loro impiego nell'IA, sempre di più si tende ad utilizzare il *Distributed Learning* che consente di allenare e validare gli algoritmi che analizzano i dati provenienti da diverse istituzioni, evitando la centralizzazione dei dati stessi e quindi senza la necessità di «muoverli». Si tratta quindi di sfruttare la rete internet per «distribuire» i dati, in modo che il singolo algoritmo possa attingere a diversi *database*.

I *Big Data*, una volta raccolti in modo adeguato, offrono una duplice potenzialità: quella della numerosità campionaria (cui consegue una statistica capace di offrire parametri valutativi con alte probabilità di efficacia) e quella della qualità (nella misura in cui i dati sono ben raccolti). Entrambi questi aspetti contribuiscono in modo significativo al passaggio dall'uso

delle Reti neurali artificiali «pre-addestrate» e già molto diffuse in vari ambiti, all'uso delle *Deep Neural Network* (DNN) specifiche, per applicazioni nel mondo della medicina e chirurgia. Queste ultime offrono risultati molto migliori, al prezzo naturalmente di una maggiore complessità di calcolo.

Un aspetto interessante applicabile anche in ambito clinico, riguarda l'applicazione dell'IA al NLP (*Natural Language Processing*), che potrebbe aiutare i medici nella *review* di articoli o di casi chirurgici o clinici, favorendo una maggiore numerosità di valutazioni, rispetto a una valutazione solo personale. L'IA è utilizzata per la «lettura» dei diversi articoli scientifici, fornendone una classificazione per argomenti o per risultati, come aiuto al medico o al chirurgo.

Attualmente, nella chirurgia in modo particolare, accanto allo sviluppo di una ingegnerizzazione di alta qualità e alla produzione di robot specializzati, si sta diffondendo in modo crescente il ML nella *Computer Vision* (CV), che fa uso delle CNN per identificare strumenti operatori e tratti anatomici, e per valutare le procedure chirurgiche a partire dai video degli interventi sia in differita che in diretta. In questi casi, trattandosi di filmati è utile l'uso di «modelli ricorsivi» che facilitano l'acquisizione delle informazioni temporali legate allo sviluppo dei singoli *frames* dei video (si usano a questo scopo le RNN). Le immagini endoscopiche degli interventi in laparoscopia certamente forniscono una quantità enorme di dati, ma non sempre presentano una accettabile uniformità delle riprese a parità di passaggi operatori. Infatti, in alcuni casi, ad esempio, l'immagine prodotta è piena (riempie l'intero *frame*), in altri è presente un *black frame* che circonda l'inquadratura «tonda» della telecamera. Se ne deduce la necessità di un «non indolore» processamento previo delle immagini raccolte che le renda confrontabili tra loro e quindi gestibili in modo efficace dagli algoritmi, i quali come si sa non hanno la capacità di astrazione. Le DNN e il *Deep Learning*, necessari per la «preparazione» o «allenamento» delle NN, sono infatti vincolati da procedimenti, normalmente abbastanza complessi, di *pre-processing* dei dati (immagini e filmati).

Tenendo conto che gli elementi in gioco sono l'acquisizione delle immagini in movimento, la definizione della geometria tridimensionale, la risoluzione, il cosiddetto *frame rate*, i colori, ecc., il loro processamento non è affatto banale e richiede NN ben addestrate (con i problemi che ne conseguono) e molto consistenti in termini di neuroni e di livelli neurali profondi (*hidden layers* nelle DNN).

Un compito dell'IA usata per il processamento delle immagini è la determinazione dei contorni e la loro individuazione/riconoscimento specifica. Questo avviene tramite l'uso di opportuni «filtri» (ottenuti sempre con elaborazioni matematiche e statistiche a partire dal contenuto informativo delle immagini). In questo caso il DL aiuta nel riconoscimento dei bordi e quindi degli «oggetti», cioè di parti anatomiche o strumenti chirurgici. Tec-

niche simili si usano per le immagini e per i video, nel cui caso le sequenze temporali e il movimento richiedono una certa velocità di lettura e una più attenta fase di «addestramento». Si parla in proposito di *image classification* per individuare la presenza o meno di determinati oggetti nell'immagine; di *object detection* per localizzare gli oggetti all'interno dell'immagine; di *semantic segmentation* per l'individuazione effettiva delle immagini riconosciute, contornandole; di *instance segmentation* per l'individuazione con contorni precisi di ogni singolo oggetto nelle loro distinzioni numeriche. Tutto questo viene applicato anche per il riconoscimento automatico dei passi operatori che caratterizzano un singolo intervento.

Un altro aspetto estremamente interessante e sfidante dell'IA applicata alla medicina è quello delle «immagini» in 3D, ottenute utilizzando riprese nella forma di *multiple-views*, che introducono diverse prospettive sul campo operatorio. Questo richiede un'ulteriore complessità legata alla gestione di un unico sistema di riferimento (*world-coordinate frame of reference*) cui si rapportano e si coordinano le immagini elaborate dalle diverse prospettive delle telecamere. L'utilità di questo approccio è quella di poter gestire forme e dimensioni degli oggetti sotto investigazione.

In tutto questo orizzonte si pone poi il problema opposto a quello dei *Big Data*, cioè quello della «compressione» delle immagini, secondo diverse tecniche ad essa associate. Sembra un paradosso, ma la necessità di compattare i dati o, meglio, di ottenerne una sintesi che sia comunque significativa, è richiesta proprio dalla loro abbondanza e dalla loro qualità. Si tratta allora di trovare il giusto equilibrio tra le due esigenze, vista la particolare operazione alla base dell'IA. Essa infatti «digitalizza» il reale, il quale, pur essendo «disponibile», sempre preserva il suo mistero⁸ chiamando il ricercatore all'umiltà e alla consapevolezza di avere a disposizione un'«immagine» parziale, campionata, digitale, limitata, di una realtà molto più complessa, profonda e «continua».

L'IA in chirurgia inoltre è molto utile in modo particolare per la formazione dei futuri chirurghi, grazie all'uso di tecniche assistite da robot (*Robotic Process Automation* – RPA). L'utilizzo congiunto di algoritmi e di innovazioni tecnologiche ingegnerizzate, ma anche della Realtà virtuale (*Virtual Reality* – VR), permette di dotare i cammini formativi di quella che viene chiamata Realtà aumentata (*Augmented Reality* – AR). Si tratta cioè di ambienti in cui la realtà «reale» e la realtà «virtuale» interagiscono, offrendo quindi agli studenti e/o specializzandi di avvicinarsi all'anatomia

⁸ Il mistero «crea la possibilità di un allargamento concettuale [...] nella scoperta di una teoria di ordine superiore [...] il mistero invita a impegnarsi, ma resiste a ogni chiusura», A.E. MCGRATH, *The Territories of Human Reason. Science and Theology in an Age of Multiple Rationalities*, Oxford University Press, Oxford 2019, 186-201 (mia traduzione).

e alle tecniche operatorie con l'ausilio delle indicazioni «virtuali» riconosciute ed elaborate dall'IA, come in un campo operatorio dilatato. Nello studio dell'anatomia, lo studente può vedere ad esempio un organo o un apparato reale, cui si aggiungono virtualmente evidenziazioni/riconoscimenti di alcuni tratti anatomici (individuati dall'IA) che si «sovrappongono», virtualmente appunto, man mano che procede l'esplorazione; oppure se ne può vedere l'intera struttura in modo completamente virtuale permettendo allo studente un «viaggio» al suo interno, con la possibilità di effettuare «esplosioni» parziali delle strutture interne, in modo da acquisirne (sempre virtualmente) familiarità, anche laddove sarebbe molto difficile la stessa operazione dal punto di vista reale (cf. strutture cellulari).

È opportuno fare attenzione, anche in questo caso, all'uso dell'altro ossimoro proposto e largamente diffuso: Realtà virtuale. Infatti, ciò che è virtuale di per sé non può essere reale, pertanto si accostano anche in questo caso due parole poste su differenti piani epistemologici. È allora opportuno, per non cadere in confusione e quindi per un più responsabile e libero uso delle tecnologie associate, distinguere chiaramente le parole specificando che il «virtuale» elaborato rappresenta appunto una pseudo-realtà, disponibile all'apprendimento, ma diversa dalla realtà cui facciamo riferimento come punto di partenza e base della nostra conoscenza. Questo, in analogia con l'altro ossimoro già evidenziato rappresentato dall'*intelligenza artificiale*.

L'IA può aiutare, insieme alla VR, anche nell'integrazione della dimensione visiva con quella «sensitivo-tattile», producendo artificialmente realistiche «sensazioni» tattili che si ripercuotono sullo strumento utilizzato dal chirurgo. L'idea è quella di offrire diversi «apparati virtuali» disponibili all'apprendimento, quali: il *thinking* attraverso il processamento dei dati, l'*acting* tramite attuatori che aiutano o sostituiscono il chirurgo, e il *sensing* che fornisce un livello percettivo artificiale. Gli ultimi due passaggi interagiscono direttamente con il campo operatorio. Nella chirurgia robotica, il chirurgo lavora infatti su una *console*, distinta dal campo operatorio nel quale invece «lavorano» bracci robotici, i quali devono poter fornire all'operatore anche le «sensazioni tattili».

La novità di questi metodi di addestramento, che richiedono una formazione integrata, dinamica e tecnologicamente avanzata, può essere applicata ai casi di procedure *robot-assisted* e a quelle di semplice simulazione. In generale, comunque, l'utilizzo dell'IA in chirurgia può fornire, tra gli altri, i seguenti ausili.⁹

⁹ B. NAMAZI – V. DEVERAJAN – G. SANKARANARAYANAN, *Artificial Intelligence for Surgical Education and Intraoperative Analysis*, in D.A. HASHIMOTO (a cura di), *Artificial Intelligence in*

- tracciamento degli strumenti chirurgici identificandone la presenza, la posizione o la posa da immagini fisse, flussi video e dati cinematici;
- organizzazione dei video o dei dati cinematici in diverse fasi, attività/sotto-attività e gesti, e classificazione delle parti chirurgiche;
- segmentazione della scena chirurgica localizzando le strutture anatomiche, gli strumenti e le parti chirurgiche;
- valutazione basata sulle prestazioni per brevi frammenti di attività o per l'intera durata dell'operazione.

Prima dell'intervento chirurgico, è inoltre possibile ricorrere all'aiuto dell'IA al fine di sviluppare modelli di valutazione o stratificazione dei rischi pre-operatori, basandosi sulle cartelle cliniche poste online che permettono di conseguenza un progressivo aumento di accuratezza, senza mai dimenticare il ruolo centrale e imprescindibile di sintesi e di verifica del medico. Se da un lato, infatti, l'ausilio della tecnologia (sia algoritmica che robotica) può alleggerire alcuni tratti delle procedure clinico/chirurgiche, dall'altro è richiesta una sempre maggiore preparazione e attenzione da parte della persona umana, nella valutazione del peso da dare ai risultati provenienti da questi strumenti. Tale valutazione inizia dalla presa di coscienza di quale sia la natura strutturale dell'IA in modo da non sopravvalutarne o idealizzarne i tratti in riferimento alle sue applicazioni e ai suoi risultati. Si potrà così sviluppare un senso critico capace di valORIZZARE invece al massimo le potenzialità umane arricchite dai dati (perché tali sono e non la realtà), raccolti dagli strumenti tecnologici.

Un'altra prospettiva molto importante per l'IA in medicina è quella di utilizzare gli algoritmi delle *Real-Time Strategies* (RTS), cioè dei giochi in tempo reale che si basano su informazioni imperfette, pianificazioni a lungo termine, ampi spazi di azione. Queste «tecnologie» saranno molto utili per affrontare questioni cliniche, vista la loro applicabilità a circostanze in continuo e imprevedibile sviluppo.

Infine, è bene segnalare la possibilità di:

- realizzare interventi chirurgici in remoto: il chirurgo in un luogo al comando di una *console* e una *IA*-robotica in un altro luogo, che materialmente «opera» sul paziente;
- *IA* che sappiano interagire in modo diretto e dialogico con i pazienti in remoto per fornire un ausilio clinico a largo spettro, recependo e registrando misurazioni cliniche tramite opportuni sensori, per monitorare i decorsi clinici o per sostenere e ricordare la somministrazione

Surgery. Understanding the Role of AI in Surgical Practice, McGraw Hill, New York 2021, 141-171, qui 164 (mia traduzione).

di terapie, «dialogando» e «rassicurando» i pazienti durante l'evolversi dei *check-up*, ecc.

Per questo, sono allo studio piattaforme, strumentazioni e IA di *Health Care Remote Monitoring* (HCRM), per il quale sono necessari, soprattutto nelle aree più disagiate e povere, (1) piani per l'integrazione, l'archiviazione e la condivisione dei dati; (2) programmi di educazione e addestramento, per sviluppare, implementare e mantenere questi sistemi; (3) reti e interfacce per l'uso di smartphone con collegamento internet in stretto contatto con strutture cliniche; (4) sviluppo di dispositivi «indossabili» dai pazienti che consentano l'acquisizione dei dati in tempo reale; (5) sviluppo e utilizzo di un collegamento diretto con medici o specialisti disponibili, oppure, solo come interfaccia intermedia e non sostitutiva, l'utilizzo di della *Conversational Artificial Intelligence Chatbots*.

2. Etica della tecnologia e dell'intelligenza artificiale

Quale etica, e quindi come gestire, in contesti così complessi e tecnologici, l'essenziale interazione persona-persona (medico-paziente e paziente-équipe di cura), che sola, può offrire realmente un *surplus* alla cura clinica, favorendone l'umanizzazione?

2.1. La «bravura» dell'IA: ovvero l'etica della tecnologia

L'IA e le tecnologie hardware e software ad essa connesse si stanno sviluppando a un ritmo sempre più elevato e con risultati sempre più affascinanti che però pongono interrogativi importanti. Il problema oggi è proprio quello del non altrettanto rapido sviluppo di regolamentazioni, di un'etica e di leggi che possano aiutare a gestire, a livello internazionale, le implicazioni spesso estremamente delicate dell'uso dell'IA nella sua attualità e nelle sue prospettive.

«Ascoltare» e scoprire, attraverso un sintetico *excursus* storico, alcuni risultati emersi lungo il cammino fino ad oggi intrapreso dalla comunità internazionale in merito alla riflessione etica sulla tecnologia, sulla robotica e sull'IA, potrà aiutare lo sviluppo di una futura e robusta etica della tecnologia. Il materiale attualmente disponibile in merito è abbastanza ampio e ricco, vista la consapevolezza della sua importanza,¹⁰ anche se

¹⁰ Cf. P. BENANTI, *La condizione tecno-umana. Domande di senso nell'era della tecnologia*, EDB, Bologna 2016.

rimane ancora per lo più al livello di *soft regulation* (cioè, senza pieno valore vincolante). D'altra parte, l'interconnessione delle proposte e il loro fondamento appaiono comunque strutturati intorno al criterio di fondo che è quello della dignità dell'uomo e del bene dell'umanità.

In riferimento all'etica applicata nella medicina, ambito nel quale è particolarmente evidente come la tecnologia debba essere posta al servizio del bene, l'orizzonte etico classico e primario si fonda sui seguenti principi morali¹¹ da tutti riconosciuti, che rappresentano la base operativa per ogni successivo sviluppo:

- 1) rispetto per l'autonomia;
- 2) non maleficenza;
- 3) beneficenza;
- 4) giustizia.

In particolare, la beneficenza «è necessariamente il principio prevalente in medicina e tutti gli altri principi sono sussidiari ad essa [...] l'attenzione alla beneficenza [...] li contestualizza e rende l'intera impresa medica più coerente».¹² La stessa considerazione è di conseguenza applicabile *in toto* all'intera impresa tecnologica che oggi coinvolge l'umanità, fondandone auspicabilmente il necessario orientamento al vero bene.

Un importante e storico riferimento originario dello sviluppo etico della tecnologia lo troviamo in Isaac Asimov (1920-1992) il quale nel 1942 ha definito e proposto le tre leggi fondamentali della robotica:¹³

- 1) un robot non può ferire un essere umano o, attraverso l'inazione, consentire a un essere umano di subire danni;
- 2) un robot deve obbedire agli ordini impartiti dagli esseri umani, a meno che tali ordini non siano in conflitto con la Prima Legge;
- 3) un robot deve proteggere la propria esistenza purché tale protezione non sia in contrasto con la Prima o la Seconda Legge.

A partire da queste «basi etiche» fondative, mentre sarà necessario identificare quale sia il vero bene, inteso in senso assoluto, cioè valido per ogni persona umana, sono poi stati definiti diversi valori etici applicabili alla tecnologia, che in campo medico assumono una importanza decisiva.

Passi concreti nascono solo nel 2002 con la proposta di Gianmarco Veruggio, della Scuola di robotica presso la Scuola S. Anna di Pisa, che coniò

¹¹ Cf. T.L. BEAUCHAMP – J.F. CHILDRESS, *Principles of Biomedical Ethics*, Oxford University Press, Oxford 2009.

¹² R.M. TAYLOR, *Ethical principles and concepts in medicine*, in J.L. BERNAT – R. BERESFORD (a cura di), *Handbook of Clinical Neurology, 118 (3rd series): Ethical and Legal Issues in Neurology*, Elsevier, Amsterdam 2013, 7 (mia traduzione).

¹³ Cf. I. ASIMOV, *Runaround*, in *Astounding Science-Fiction*, Street & Smith Publications, New York 1942, 94-103.

e sviluppò una *roboetica* «per indicare il rapporto positivo che dovrebbe intercorrere tra progettista/produttore/utente di robot e queste macchine intelligenti. Non solo norme negative, dunque, ma la complessa relazione che collega gli umani ai loro artefatti intelligenti e autonomi».¹⁴ Successivamente dal 30 al 31 gennaio 2004 si è tenuto a Sanremo il primo Simposio internazionale sulla *roboetica* con il titolo *The ethics, social, humanitarian and ecological aspects of Robotics*, sempre organizzato in collaborazione con la Scuola di robotica. Lo stesso Veruggio pose in evidenza la necessità di approfondire primariamente l'etica umana per poi riflettere su quella da applicare ai robot: «prima ancora di affrontare il problema dell'etica artificiale di cui saranno dotate le nostre macchine intelligenti, ci sta a cuore l'etica umana di chi queste macchine progetta, produce e impiega».¹⁵

L'importanza del tema è anche sottolineata dal riconoscimento della necessità di coinvolgere diverse discipline per la sua elaborazione, come: «campi principali da coinvolgere nella *roboetica*: robotica, informatica, intelligenza artificiale, filosofia, etica, teologia, biologia, fisiologia, scienze cognitive, neuroscienze, giurisprudenza, sociologia, psicologia, disegno industriale».¹⁶ Nel 2018 si è poi passati dalla *roboetica* all'*algoretica*, per indicare la necessità di accostare una riflessione etica anche allo sviluppo degli algoritmi.¹⁷

I tratti salienti del dibattito in corso sui singoli valori etici della tecnologia in generale e declinabili in modo particolare nel contesto dell'IA e della robotica in medicina, da un lato sono la testimonianza di una sensibilità etica crescente, dall'altro evidenziano la necessità di una chiamata alla strutturazione etica normativa (vincolante) che sia completamente *Human Person centered*. Tali valori etici, unitamente a specifiche linee programmatiche, emergono in corrispondenza di diverse e strettamente correlate raccolte di studi, di seguito riportate in modo sintetico e schematico:

2015 *Handbook of Ethics, Values, and Technological Design. Sources, Theory, Values and Application Domains*: responsabilità + trasparenza; democrazia + giustizia; benessere umano; inclusività; presenza; privacy; regolamentazione; sicurezza; sostenibilità; fiducia.¹⁸

¹⁴ Cf. F. OPERTO – G. VERUGGIO, *A dieci anni dalla nascita della Roboetica*, in *Mondo Digitale*, ottobre 2014, 3; G. VERUGGIO, *Il cammino della Roboetica*, in *Le Scienze* 40 (2007), n. 461, 34-35.

¹⁵ Cf. G. VERUGGIO – F. OPERTO, *Roboethics: Social and Ethical Implications of Robotics*, in B. SICILIANO – O. KHATIB (a cura di), *Springer Handbook of Robotics*, Springer, New York 2016, 990-998; cf. OPERTO – VERUGGIO, *A dieci anni dalla nascita della Roboetica*, 3.

¹⁶ G. VERUGGIO, *The EURON Roboethics Roadmap*, in *Proc. Humanoids '06: 6th IEEE-RAS Int. Conf. Humanoid Robots*, 2006, 612-617, qui 613.

¹⁷ Cf. P. BENANTI, *Oracoli. Tra algoretica e algocrazia*, Luca Sossella Editore, Roma 2018.

¹⁸ Cf. J. VAN DEN HOVEN – P.E. VERMAAS – I. VAN DE POEL (a cura di), *Handbook of Ethics, Values, and Technological Design. Sources, Theory, Values and Application Domains*, Springer, New York 2015.

2017 *I 23 Principi di Asilomar*: suddivisi in 3 categorie:

- *ricerca*: intelligenza benefica; uso vantaggioso; collegamento scienza-politica; cultura della ricerca: cooperazione – fiducia – trasparenza; evitare la gara;
- *etica e valori*: sicurezza; trasparenza del fallimento; trasparenza giudiziaria; responsabilità; allineamento del valore con i valori umani; valori umani; libertà e privacy personale; beneficio condiviso; prosperità condivisa; controllo umano; non sovversione; evitare la corsa agli armamenti con l'IA;
- *lungo termine*: attenzione alle capacità; importanza + rischi; auto-miglioramento ricorsivo; bene comune.¹⁹

2018 *Dichiarazione di Montréal*: benessere; autonomia; privacy – intimità; solidarietà; partecipazione democratica; equità; diversità e inclusione; attenzione; responsabilità; sostenibilità.²⁰

2018 «EGE – *Ethical Principles and democratic prerequisites*»: dignità umana; autonomia; responsabilità; giustizia – equità – solidarietà; democrazia; responsabilità di regole e leggi; sicurezza – integrità corporea e mentale; protezione dei dati e privacy; sostenibilità.²¹

2018 *AI Code*: bene comune – benefici per l'umanità; intelligibilità – equità; non riduzione dei diritti sui dati o sulla privacy; fiorire mentalmente, emotivamente ed economicamente; non ferire, distruggere o ingannare l'essere umano.²²

2018 *The Tenets of the Partnership on AI*: robusto – affidabile – sicuro; comprensibile – interpretabile; avvantaggiare e responsabilizzare le persone; educare, ascoltare, coinvolgere, cercare, informare, rivolgersi alle parti interessate; ricerca aperta – implicazioni del dialogo dell'IA; impegno e responsabilità nei confronti di un'ampia gamma di parti interessate; garantire che le preoccupazioni e le opportunità specifiche del dominio siano comprese e affrontate; proteggere la privacy e la sicurezza; comprendere e rispettare; responsabilità; nessuna violazione dei diritti umani – nessun danno; cooperazione – fiducia – apertura.²³

2018 *AI4People*: non maleficenza: privacy – sicurezza – capacità di attenzione; beneficenza: benessere – dignità – sostenibilità; giusti-

¹⁹ *Principles developed in conjunction with the 2017 Asilomar conference* [Benevolent AI 2017], in <https://futureoflife.org/ai-principles> (accesso: 1° marzo 2022).

²⁰ Cf. *Montréal Declaration for a Responsible Development of Artificial Intelligence*, 2018.

²¹ EUROPEAN GROUP ON ETHICS IN SCIENCE AND NEW TECHNOLOGIES (EGE), *Statement on Artificial Intelligence, Robotics and "Autonomous" Systems*, 2018.

²² *Five overarching principles for an AI code*, 2018, par. 417.

²³ *The Tenets of the Partnership on AI*, 2018, in <https://partnershiponai.org/about/#tenets> (accesso: 1° marzo 2022).

zia: prosperità – solidarietà; spiegabilità: intelligibilità – responsabilità; consentire l'autorealizzazione umana, senza svalutare le capacità umane; potenziare l'agire umano, senza rimuovere la responsabilità umana; aumentare le capacità della società, senza ridurre il controllo umano; coltivare la coesione sociale, senza erodere l'autodeterminazione umana; autonomia: potere di decidere.²⁴

2019 «*IEEE Ethically Aligned Design (EAD)*»: suddivisi in quattro categorie:

- *EAD Principi generali*: diritti umani; benessere; azione dei dati; efficacia; trasparenza; responsabilità; consapevolezza dell'uso improprio; competenza;
- *EAD Pilastri*: valori umani universali; autodeterminazione politica e azione dei dati; affidabilità tecnica;
- *EAD Capitoli: computing affettivo*: sistemi tra culture – cura dei sistemi – manipolazione/spostamento/inganno del sistema – sistemi a supporto del potenziale umano – sistemi con emozioni sintetiche; dati personali e azione individuale: creazione – cura – controllo – problema dei dati per i minori; metodi per guidare la ricerca e la progettazione etica: istruzione e ricerca interdisciplinare – pratiche aziendali su A/IS (*Autonomous and Intelligent Systems*) – responsabilità e valutazione; A/IS per lo sviluppo sostenibile: A/IS al servizio dello sviluppo sostenibile per tutti – pari disponibilità – A/IS e occupazione – educazione per l'età A/IS – A/IS e azione umanitaria; diritto: benessere – ostacoli alla fiducia informata – efficacia – competenza – responsabilità – trasparenza.²⁵

2019 *Orientamenti etici per un'IA affidabile*:

- *IA Affidabile*: legalità; eticità; robustezza;
- *4 principi etici*: rispetto dell'autonomia umana; prevenzione dei danni; equità; esplicabilità;
- *7 requisiti fondamentali*: intervento e sorveglianza umani; robustezza tecnica e sicurezza; riservatezza e *governance* dei dati;

²⁴ L. FLORIDI – J. COWLS – M. BELTRAMETTI ET AL., *AI4People – An Ethical Framework for a Good AI Society: Opportunities, Risks, Principles, and Recommendations*, in *Minds and Machines* 28 (2018), 689–707.

²⁵ THE IEEE GLOBAL INITIATIVE ON ETHICS OF AUTONOMOUS AND INTELLIGENT SYSTEMS, *Ethically Aligned Design: A Vision for Prioritizing Human Well-being with Autonomous and Intelligent Systems*, 2019. Cf. <https://standards.ieee.org/content/ieee-standards/en/industry-connections/ec/autonomous-systems.html> (accesso: 1° marzo 2022).

trasparenza; diversità, non discriminazione ed equità; benessere sociale e ambientale; *accountability*.²⁶

- 2019 *20 principi guida*: dignità; divieto di discriminazione; trasparenza; *accountability*; libertà di espressione; privacy e protezione dei dati personali; giustizia; pubblica amministrazione; regolamentazione; soggettività; *smart contracts* e tecnologia *blockchain*; effettività; granularità; mercati; responsabilità civile; colpevolezza; *product liability*; veicoli a guida autonoma; processo; sindacabilità.²⁷
- 2020 *Designing for human rights in AI*: dignità; libertà; uguaglianza; solidarietà; diritto alla vita; non discriminazione; privacy; trasparenza; sicurezza.²⁸
- 2021 *WHO*: proteggere l'autonomia: assistere gli esseri umani – efficienza – trasparenza; promuovere il benessere umano: sicurezza – interesse pubblico (nessun danno fisico o mentale); trasparenza – spiegabilità – intelligibilità; responsabilità: garanzia umana; inclusività – equità; reattività – sostenibilità.²⁹
- 2023 *UE AI Act*: propone una regolazione dell'uso dell'IA con un approccio prevalentemente basato sulla valutazione dei diversi livelli di rischio: consentiti senza restrizioni; consentiti ma soggetti a obblighi di informazione/trasparenza; consentiti previo rispetto dei requisiti IA e valutazione della conformità ex ante; vietati.³⁰

L'Italia, attraverso il Ministero dello sviluppo economico, nel 2020 ha elaborato il documento *Proposte per una strategia italiana per l'Intelligenza artificiale*, ove si sottolinea che «un approccio antropocentrico all'intelligenza artificiale richiede che l'essere umano venga valorizzato e supportato dalla tecnologia nel suo realizzarsi come individuo, nel suo agire

²⁶ GRUPPO INDEPENDENTE DI ESPERTI AD ALTO LIVELLO SULL'INTELLIGENZA ARTIFICIALE, istituito dalla Commissione europea nel giugno 2018, *Orientamenti etici per un'IA affidabile*, Commissione europea, Bruxelles 2019.

²⁷ G. GUAITA, *I pericoli dell'Intelligenza Artificiale: i fenomeni da tenere d'occhio per governare il processo*, 2019, in <https://aemmedi.it/wp-content/uploads/2020/02/G-Guaita.pdf> (accesso 1° marzo 2022).

²⁸ E. AIZENBERG – J. VAN DEN HOVEN, *Designing for human rights in AI*, in *Big Data & Society* 7/2 (2020), 1-14.

²⁹ WORLD HEALTH ORGANIZATION, *Ethics and governance of Artificial Intelligence for health: WHO guidance*, Geneva 2021.

³⁰ Cf. COMMISSIONE EUROPEA, *Un approccio europeo all'intelligenza artificiale*, in <https://digital-strategy.ec.europa.eu/it/policies/european-approach-artificial-intelligence> (accesso: 5 febbraio 2024).

come cittadino, nelle sue mansioni e competenze come lavoratore e nelle sue decisioni come consumatore».³¹

L'IA, in questo orizzonte etico, può essere allora definita come HAI (*Human-centered AI*), con «l'impegno a studiare, guidare e sviluppare tecnologie e applicazioni di Intelligenza artificiale incentrate sull'uomo».³²

Come si vede, lo sviluppo tecnologico «chiama» in modo pressante alla comprensione, attenzione, centratura e valorizzazione della dignità della persona umana, per cercarne il fondamento sostanziale. La persona umana può elevarsi, infatti, dilatando gli orizzonti della ricerca, alla trascendenza e finalmente approdare alla rivelazione di Dio in Gesù Cristo che, come abbiamo visto, svela pienamente l'uomo all'uomo. Quest'ultimo, rivestito di una dignità altissima che risuona all'interno della dimensione creaturale, si comprende come figlio nel Figlio, centrato nel seno delle tre Persone divine. Evidentemente tali conclusioni fondative non sono riconosciute nell'ambito internazionale della discussione etica dell'IA, ma non possiamo non considerarne l'importanza risolutiva proprio in vista di un consapevole contributo «scientifico», strettamente legato al mistero dell'incarnazione, che la teologia cattolica e l'antropologia teologica in particolare possono offrire al dibattito in corso.

Il dischiudersi di questa fioritura di valori etici offre certamente una rosa dinamica di indirizzi volti a favorire la scelta e l'implementazione di tecnologie e algoritmi che siano chiaramente centrati sulla persona umana; d'altra parte, però, questa abbondanza emergente nel panorama internazionale

rischia di diventare soverchiante e fuorviante sollevando due potenziali problemi. O i vari insiemi di principi etici per l'IA sono simili, portando inutili ripetizioni e ridondanze, oppure, se differiscono in modo significativo, sono suscettibili di ingenerare confusione e ambiguità. Il peggior risultato sarebbe quello di creare un «mercato dei principi» in cui le parti interessati potrebbero essere tentate di «acquistare» quelle più allettanti.³³

Questa è la conseguenza di una «sensibilità» che, mentre si mostra attenta alla dimensione etica, allo stesso tempo, mantenendola sul piano più o meno esortativo di *soft law*, finisce per renderla ininfluente.

³¹ GRUPPO DI ESPERTI MISE SULL'INTELLIGENZA ARTIFICIALE, *Proposte per una strategia italiana per l'Intelligenza artificiale*, MISE, Roma 2020, 87-102.

³² Cf. G. COVINO, *Etica e Intelligenza Artificiale: i Principi di Asilomar e la Human-centered AI*, in *Briciole Filosofiche*, 8 giugno 2020, in <https://bricioledifilosofia.com/2020/06/08/etica-e-intelligenza-artificiale-i-principi-di-asilomar-e-la-human-centered-ai/> (accesso 1° marzo 2022). Cf. WEI XU, *Toward Human-Centred AI: A Perspective from Human-Computer Interaction*, in *Interactions* 26/4 (2019), 42-46, qui 44.

³³ FLORIDI, *Etica dell'intelligenza artificiale*, 92-93. Per una trattazione sinottica più completa dei principi etici cf. *ivi*, 91-105.

L'esame dei principi etici proposti, secondo l'*AI Ethics Guidelines Global Inventory* (rassegna globale delle linee guida etiche per l'IA) di *Algorithm Watch*,³⁴ offre una sintesi di 5 principi fondamentali, aggiungendone solo uno ai 4 iniziali: beneficenza, non maleficenza, autonomia, giustizia ed esplicabilità. Quest'ultimo tratto permette all'interfaccia persona – macchina/algoritmo di essere sempre «spiegabile» e quindi controllabile da parte nostra, anche nel suo orientamento etico e quindi di servizio. Infatti, uno dei maggiori problemi che si incontrano nell'uso dell'IA è il passaggio dalla «comprendere» alla «efficienza», che sposta il livello epistemologico sul piano del «funziona» (come fosse una *Black Box*), e non su quello del «perché funziona», ponendo di conseguenza seri rischi di derive, appunto o non controllate o pilotate.

In conclusione, è importante citare la *Rome Call for AI Ethics: «RenAIssance. A human-centric Artificial Intelligence»* tenutasi a Roma al termine della XXVI assemblea generale dei membri della Pontificia accademia per la vita dal 26 al 28 febbraio 2020. In quell'occasione Brad Smith (presidente di Microsoft), John Kelly III (vice-presidente esecutivo di IBM), Qu Dongyu (direttore generale della FAO), Paola Pisano (rappresentante del ministro dell'innovazione tecnologica del governo italiano), Vincenzo Puglia (presidente della Pontificia accademia per la vita) hanno siglato un accordo storico sull'etica del «buon algoritmo»:

Ora più che mai, dobbiamo garantire una visione in cui l'IA sia sviluppata non in funzione della tecnologia, ma piuttosto per il bene dell'umanità e dell'ambiente, della nostra casa comune e condivisa e dei suoi abitanti umani, che sono inestricabilmente connessi. In altre parole, una visione in cui gli esseri umani e la natura siano al centro dello sviluppo dell'innovazione digitale, supportati e non gradualmente sostituiti da tecnologie che si comportano come attori razionali, ma non sono in alcun modo umani. È ora di iniziare a prepararsi per un futuro più tecnologico in cui le macchine avranno un ruolo più importante nella vita degli esseri umani, ma anche un futuro in cui sia chiaro che il progresso tecnologico afferma la genialità della razza umana e rimane dipendente dalla sua integrità etica.³⁵

Una importante convergenza è stata ratificata tra i partecipanti in corrispondenza dei seguenti «principi etici fondamentali per una buona innovazione»:

1) *trasparenza*: i sistemi di IA devono essere spiegabili;

³⁴ ALGORITHM WATCH, *AI Ethics Guidelines Global Inventory*, 9 aprile 2019, in <https://algorithmwatch.org/en/hashtag/ethics-guidelines/> (accesso: 7 marzo 2022).

³⁵ Cf. V. PAGLIA – R. PEGORARO (a cura di), *The «Good» Algorithm? Artificial Intelligence: Ethics. Law, Health*, Pontifical Academy for Life, Roma 2021, 239-240 (mia traduzione).

- 2) *inclusione*: i bisogni di tutti gli esseri umani devono essere presi in considerazione affinché tutti possano trarne beneficio, e a tutti gli individui possano essere offerte le migliori condizioni possibili per esprimersi e svilupparsi;
- 3) *responsabilità*: coloro che progettano e implementano l'uso dell'IA devono procedere con responsabilità e trasparenza;
- 4) *imparzialità*: non creare o agire in base a pregiudizi, salvaguardando così l'equità e la dignità umana;
- 5) *affidabilità*: i sistemi di IA devono essere in grado di funzionare in modo affidabile;
- 6) *sicurezza e privacy*: i sistemi di IA devono funzionare in modo sicuro e rispettare la privacy degli utenti.

In riferimento alla *Caritas in veritate* di Benedetto XVI (cf. nn. 5-6.68-77), si è infine evidenziata la necessità di una vera e propria *educatio vitae* per una *Person-centered Ethics* nell'era della *IA*, la cui semantica è più completa rispetto all'espressione *Human-centered*. Questo sarà possibile quanto più si prenderà coscienza dell'esigenza, accanto all'impiego dell'IA in ambito clinico, dell'insostituibile importanza del «tocco umano».³⁶

Il passaggio da un'etica dei principi a un'etica centrata sulla persona umana potrebbe ulteriormente maturare, particolarmente a partire dall'ambito medico, nella direzione di un'*etica della relazione*. Si tratta di un auspicio che, attingendo alla struttura trinitaria della persona umana, appunto costitutivamente "dalla, per e nella" relazione, permetterebbe lo sviluppo di una medicina che più sapientemente sappia gestire la tecnologia per favorire l'emergere di un nuova e vera umanità: nella carità.

3. L'intelligenza artificiale per le cure radio e chemio-terapiche: un'applicazione pratica

Presso il Policlinico universitario «A. Gemelli» di Roma, un gruppo di ricerca trans-disciplinare nato intorno al reparto Art4ART (Arte per la *Advanced Radio Therapy*) sta portando avanti un progetto innovativo di applicazione dell'IA all'accompagnamento integrale della persona in cura oncologica. Questo avveniristico reparto declina l'arte, nelle sue molteplici forme espressive, lungo e accanto al percorso umano e clinico delle per-

³⁶ Cf. PAGLIA – PEGORARO, *The «Good» Algorithm?*, 85-89.113.243.

sone in cura radio- e chemio-terapica. Ogni singola persona che accede al reparto si trova a vivere il tempo dell'accoglienza e delle terapie in un contesto di bellezza immersiva e di bellezza relazionale con il personale, grazie al contributo di numerosi artisti e alla formazione specifica di tutti coloro che si pongono al servizio (personale sanitario e volontari). È stato realizzato infatti un ambiente atto a offrire una vera e propria occasione di cura integrale che coinvolga clinica (altissima professionalità congiunta alle più moderne tecnologie), umanità, spiritualità e relazionalità, a servizio della persona umana come unitotalità unificata di corpo e anima, rendendo così l'approccio terapeutico estremamente più efficace e completo. Il reparto stesso è ricco di opere artistiche, giardini verticali, colori abbinati, grandi schermi per la proiezione di filmati, ecc.

L'IA utilizzata in questo contesto è previsto che sia di aiuto almeno su due livelli: il primo è quello della sua stessa progettazione, sviluppata in modo strutturalmente etico, ponendo al centro la persona umana nel suo orizzonte teologico (basandosi sulle linee guida delle beatitudini);³⁷ il secondo è quello del suo impiego al servizio della clinica, per potenziare la personalizzazione della cura.

Nello specifico, e in sintesi, ad ogni singola persona in cura presso il reparto, durante il tempo della terapia (ma anche in quello dell'accoglienza che lo precede) viene consegnato un tablet nel quale è stata sviluppata una applicazione che consente un vero e proprio «viaggio» tra espressioni della bellezza (musica, poesia e letteratura, cinema, natura, pittura, scultura e monumenti, fotografia, mestieri) e sentimenti (amicizia, amore, attenzione, coraggio, cura di sé, entusiasmo, passione, spiritualità). Un vasto numero di contenuti multimediali è stato donato e messo a disposizione da artisti e istituzioni (ad esempio i Musei Vaticani), per il bene delle persone in questo delicato tratto di cura. In tale matrice bidimensionale l'utente può selezionare di volta in volta i due aspetti che preferisce abbinare, tra le combinazioni possibili di semi di bellezza e di sentimenti, secondo lo specifico «momento» vissuto, per percorrere tratti di «immersione» in se stesso e nelle relazioni con il personale, accompagnato dal dinamismo della bellezza. Questi percorsi, scelti e dinamici, offrono altresì l'invito e la possibilità di una interazione diretta, ove poter esprimere il tesoro delle proprie «risonanze» personali secondo diverse modalità di partecipazione.

Il compito dell'algoritmo di IA che ho avuto il compito di sviluppare, sotteso a questa matrice, è quello di costruire, con delicatezza e con dinamicità, prendendo in considerazione alcuni parametri dell'interazione

³⁷ Cf. A. MANTINI, *Technological Sustainability and Artificial Intelligence Algor-ethics*, in *MDPI – Sustainability* 14/6 (2022), n. 3215. Esteso e tradotto in: P.A. BONINI – A. MANTINI, *Ricercare nella Verità*, CISU, Roma 2023, 161-191.

dell'utente con tale offerta di bellezza che potremmo chiamare «riflessiva» e «proiettiva», una sorta di «immagine» tridimensionale in movimento del cammino interiore vissuto dalla persona in cura durante le fasi della terapia. Questo «cammino» si estende all'intero percorso del trattamento, individuando alcuni tratti della personale dinamica dell'interiorità (da considerarsi comunque e sempre come una possibile, timida e umile indicazione del vissuto della persona). Tali immagini dinamiche tengono conto dei diversi parametri legati all'interazione dell'utente con i semi di bellezza che gli sono offerti, nell'alternarsi degli stati emotivi e sentimentali, sia relazionali che personali. Esse vengono poi lette e utilizzate dall'intera équipe di cura, al fine di calibrare, passo dopo passo, non solo la qualità della cura stessa, ma anche la sua efficacia, che quindi non è più solo farmacologico-clinica. La cura integrale della persona, in questo modo, è resa possibile dall'interazione, di stampo trans-disciplinare, tra clinica, tecnologia, interiorità, relazioni e spiritualità. La bellezza, offerta attraverso diverse espressioni donate dalla carità di altre persone, diventa il fermento, il lievito, che aiuta ad armonizzare e potenziare il percorso stesso, realmente personalizzato e centrato sulla singola persona in cura.

Conclusione

Di fronte all'importanza delle relazioni persona-persona e delle interazioni persona-macchina e macchina-macchina, che sono coinvolte nel complesso contesto dell'uso dell'IA in medicina e chirurgia, non è realmente possibile, ritengo, costruire una convoluzione³⁸ tra «intelligenze», poiché si tratta di realtà diverse non solo strutturalmente ma anche sostanzialmente. Quest'ultimo termine è infatti da rivalutare nell'oggi delle interconnessioni, in cui si parla di *internet of everything* e di *internet of things*, riducendo indebitamente la realtà a «dati» (*dataismo*) e le relazioni a «correlazioni» (*infosfera*), pensando di superare il problema metafisico e teologico della sostanza, e quindi del fondamento. Si può certamente parlare di sinergia, di collaborazione, nell'ottica di un servizio della tecnologia orientato all'uomo, ma non di convoluzione, sia da un punto di vista morale, sia appunto sostanziale, in quanto quella umana è intelligenza, l'altra no.

Alcuni ricercatori parlano anche di «*Medical Intelligence*»³⁹ come nuova frontiera e quasi nuova realtà legata alla «convoluzione» di intelligenza umana e *intelligenza artificiale*. A mio avviso bisogna fare molta attenzio-

³⁸ Contrariamente all'opinione di CHANG, *Intelligent-Based Medicine*, 445-446.

³⁹ Cf. *ivi*, 428.

ne in questo processo di «ibridazione», perché tale prospettiva può ingenerare l'idea transumanista del superamento quasi evolutivo della persona umana. Invece ritengo più esatto parlare di «intelligenza medica» riferita alla sola persona umana, seppur potenziata dall'ausilio di risultati e opportunità tecniche (non quindi «intelligenti») forniti degli algoritmi di IA, con tutte le loro frontiere «cognitive», senza lasciare intendere l'insorgere di una nuova realtà ibrida. L'intelligenza umana può davvero sviluppare maggiormente le sue potenzialità, proprio coinvolgendosi in avventure di ricerca estese che, lungi dal sostituire la persona umana, invece ne stimolano ulteriormente la professionalità e creatività clinica, potendosi dedicare con molta più attenzione e tempo alla determinante dimensione di incontro personale del medico con la persona in cura!

Questa nuova frontiera del progresso umano ci chiama pertanto a una più luminosa e fondata responsabilità anche nell'imparare a lavorare in equipe di cura. Tale collaborazione serve a in-formare i medici delle potenzialità, dei rischi e del grado di affidabilità dell'IA, mentre i tecnici hanno bisogno di conoscere l'approccio dei medici per meglio sviluppare una tecnologia al servizio della persona, per una ri-umanizzazione della medicina, dove i medici avranno più tempo per rilevare tutti quei «dati» non contenuti nei freddi *database*. Solo questi «dati» permettono una valutazione integrale con un livello di attenzione e di profondità che, oltre a stimolare il medico nella sua capacità di visione olistica, favoriscono la relazione come parte integrante della cura stessa!

La sfida dell'*intelligenza* artificiale pone a medici e tecnici l'urgenza di cercare la verità dell'uomo e di sviluppare l'intelligenza reale (IR), il pensiero, il ragionare, il prendersi cura, l'amore. Un appello alla fondazione di una antropologia che esca dalle strettoie del riduzionismo e delle differenze culturali, per osare ascoltare umilmente una Rivelazione che è lì disponibile per la regione. La medicina è l'ambito privilegiato per questo dinamismo, offrendo (1) l'occasione preziosa per convergere in unità di fronte alla riscoperta della persona umana nella sua identità unica e irripetibile creata ad immagine e somiglianza di Dio; (2) l'occasione sublime della riscoperta dell'intelligenza umana (IR) come riflesso del Cristo Logos; (3) l'occasione vitale per alimentare quella identità relazionale della persona umana che esiste dalla e nella Trinità.

Persona, ragione, relazione, sono dunque i tre pilastri che ora siamo chiamati a potenziare perché ogni tecnologia, opera delle mani dell'uomo, sia una nuova opportunità per cantare la lode e la meraviglia di essere noi tutti opera delle mani di Dio, Padre, Figlio e Spirito Santo.

PS Non sembrino fuori luogo i riferimenti diretti al prezioso contributo della teologia cattolica, che sono volutamente esplicitati proprio al fine

di riportare alla nostra consapevolezza quanto essa possa porsi in diretto e qualificato dialogo con la riflessione in corso. La rinnovata e stupita riscoperta delle «potenzialità» e della «forza» dell'incarnazione di Gesù il Cristo, con tutta la rivelazione che ne consegue e la teo-logia che si sviluppa attraverso il trinomio sacra Scrittura, tradizione e magistero ecclesiastico, potrà favorire lo sviluppo autentico della scienza e della medicina in particolare, come servizio alla persona umana nell'orizzonte di un vero progresso dell'umanità verso il suo compimento: «La creazione stessa attende con impazienza la rivelazione dei figli di Dio; essa infatti è stata sottomessa alla caducità – non per suo volere, ma per volere di colui che l'ha sottomessa – e nutre la speranza di essere lei pure liberata dalla schiavitù della corruzione, per entrare nella libertà della gloria dei figli di Dio» (Rm 8,19-21).

Alessandro Mantini
Università Cattolica del Sacro Cuore
Largo F. Vito, 1
00168 Roma
alessandro.mantini@unicatt.it

Sommario

È proposta una sintesi ragionata e trans-disciplinare relativa all'uso dell'*intelligenza artificiale* (che consiste di algoritmi matematici e statistici molto avanzati) nella medicina e chirurgia dal punto di vista tecnico, applicativo ed etico. La medicina, potenziata dall'*IA* che manca della profondità della coscienza umana, è chiamata a progredire centrando ancora più decisamente sulla dignità della persona umana e il suo vero bene.

Summary – Artificial *Intelligence* and Medicine: Technical and Ethical Challenges

A reasoned and trans-disciplinary synthesis about the use of Artificial *Intelligence* (consisting of highly advanced mathematical and statistical algorithms) in Medicine and Surgery from a technical, applicative, and ethical perspective is proposed. Medicine, empowered by *AI* that lacks the depth of human consciousness, is called to progress by focusing even more decisively on the Dignity of the Human Person and his True Good.