

TECNOLOGIE DIGITALI ED ETICA

DIGITAL TECHNOLOGIES AND ETHICS

VINCENZO PAGLIA ¹

DOSSIER

Introduzione

In diverse occasioni papa Francesco ha sottolineato come il tempo in cui oggi il mondo si trova non è tanto un'epoca di cambiamenti, ma quello di un cambiamento d'epoca.² Nella profondità della trasformazione che stiamo vivendo un ruolo centrale è svolto dalla scienza e dalla tecnica, certo in stretto collegamento con l'organizzazione economico-finanziaria. Se esaminiamo tale ruolo in modo complessivo, vediamo che coinvolge e sovrverte tutte le dimensioni dell'esistenza umana e delle relazioni sociali. Jacques Ellul, già nel 1954, ci aveva preavvertito identificando chiaramente la questione della tecnica come *l'enjeu du siècle* («la questione cruciale del secolo»)³ Ma non smise di ripeterlo nel corso della sua riflessione: «sappiamo bene che tutto dipenderà finalmente dall'esito dell'avventura tecnica» (*Nous savons tous que tout dépendra finalement de l'issue de l'aventure technique*).⁴ Questo contributo svolgerà l'argo-

mento seguendo un percorso in tre passi: 1) quali sono le caratteristiche e le implicazioni della tecnologia con cui oggi ci troviamo a vivere; 2) quale etica ci viene chiesto di elaborare perché possa avere un impatto effettivo sulla tecnologia e quindi 3) alcuni criteri secondo cui configurare la tecnologia che stiamo realizzando per il futuro o, per riprendere la domanda che già negli anni '90 Dietmar Mieth formulava come titolo del suo libro: *Cosa vogliamo potere?*⁵

In conclusione, sarà presentato anche un progetto concreto che si va realizzando, su iniziativa della Pontificia Accademia per la Vita (PAV): la *Rome Call for AI ethics*. Essa intende dare concretezza anche operativa alla riflessione teorica, in modo da promuovere principi etici fondamentali per una crescita di responsabilità comune nell'ambito della cosiddetta "intelligenza artificiale" (IA).

Ma prima di entrare nell'argomento, occorre premettere un breve chiarimento lessicale. Senza dilungarci in una più elaborata distinzione tra "tec-

nica” e “tecnologia”, precisiamo l’uso che faremo dei termini. Ci baseremo sull’analogia di quanto avviene per altre simili coppie di concetti, tipo “psiche” e “psicologia”: si designerà con “tecnica” un fenomeno reale (che include strumenti e processi) e con “tecnologia” il sapere o la scienza relativa a tale ambito di realtà. L’uso al plurale del termine, cioè “tecnologie”, indicherà invece gli specifici artefatti che grazie a questi saperi e procedure vengono costruiti⁶. Impiegheremo cioè il significato che il termine ha assunto nell’uso corrente, in conformità al quale alcuni parlano anche di “intelligenze artificiali”⁷, con riferimento ai molteplici dispositivi a cui gli algoritmi possono essere congiunti.

1. La tecnologia oggi

Sappiamo come nel suo cammino l’umanità abbia vissuto molti cambiamenti sconvolgenti che, nel modificare le conoscenze sulla realtà, hanno anche dislocato la sua posizione nell’universo.

L’uomo si è trovato sospinto sempre più verso regioni periferiche: con Copernico la terra ha perso la sua collocazione centrale nel sistema planetario; con Darwin la specie umana è stata iscritta nella concatenazione evolutiva delle altre specie animali; con Freud l’Io ha scoperto che quel controllo totale che credeva di avere sulle proprie facoltà era illusorio, trovandosi sottomesso a spinte inconscie che influiscono su decisioni e comportamenti. Ognuno di questi passaggi ha comportato

non solo trasformazioni sul modo di comprendere la realtà fisica, biologica o psichica, ma anche una nuova auto-comprensione.

Del resto, la storia della tecnica ha parimenti conosciuto vere e proprie rivoluzioni. In particolare, nel campo dell’industria, si parla di una prima rivoluzione in cui la macchina a vapore ha sostituito la forza umana. Solo per segnalare la rilevanza sociale delle invenzioni tecnologiche, notiamo quanto tale sviluppo sia stato intrecciato con il graduale superamento della schiavitù, almeno nella sua versione classica.⁸ Una seconda svolta si è realizzata con l’avvento dell’energia elettrica, dell’acciaio e della chimica industriale (all’inizio soprattutto rilevante per l’agricoltura).

Il terzo passaggio è stato promosso dalla microelettronica e dall’automazione nelle manifatture.⁹ Di questa rivoluzione ancora si discute la durata: se oggi la “meccatronica” (intesa come somma di meccanica, elettronica e informatica) non sia che una manifestazione più ampia ed esplicita di quanto era già in essere, oppure se l’IA e i robot ci abbiano introdotti in una nuova era. Novità tanto più difendibile quanto più i dispositivi digitali escono dagli spazi circoscritti della produzione industriale (o dell’organizzazione militare) ed entrano nella sfera sociale e nella vita domestica, divenendo possibili funzioni sempre più flessibili.

Questa panoramica ci è sufficiente per cogliere quanto profondi siano

RIASSUNTO

Nell'attuale «cambiamento d'epoca», le tecnologie digitali svolgono un ruolo determinante. L'etica (teologica) è sollecitata non solo a elaborare una riflessione sempre più transdisciplinare, ma anche a individuare passi operativi che coinvolgano altri soggetti attivi in questo ambito. La *Rome Call for AI Ethics*, promossa dalla Pontificia Accademia per la Vita, si colloca in tale prospettiva: promuove una regolamentazione etica, frutto di una comune elaborazione, e l'assunzione di responsabilità da parte dei firmatari. I rappresentanti delle imprese e delle istituzioni, internazionali ed ecclesiali, che hanno firmato il documento si impegnano anche ad ampliare la rete degli aderenti.

Parole chiave

Teologia morale, tecnologie digitali, intelligenza artificiale, umanesimo.

SUMMARY

In the current “change of epoch,” digital technologies play a defining role. Theological ethics is under pressure not only to articulate increasingly trans-disciplinary reflections, but also to identify operational steps that involve other actors who are active in the field. The Rome Call for AI Ethics, promoted by the Pontifical Academy for Life, is situated in such a perspective, promoting ethical regulation produced in collaboration with and the assumption of responsibility by signatories.

The representatives of international and ecclesial companies and institutions who signed the document also commit themselves to expanding the network of adherents.

Keywords

Moral theology, digital technologies, artificial intelligence, humanism.

stati i cambiamenti che nella storia hanno inciso sul modo di vivere e di lavorare dell'intera umanità. Però, per rendersi conto della specificità del cambiamento che stiamo vivendo, occorre aggiungere che le tecnologie dell'informazione e della comunicazione (TIC) operano in stretta connessione con quelle che vengono definite «tecnologie emergenti e convergenti». ¹⁰ Con tale denominazione si indicano le nanotecnologie, ¹¹ le

biotecnologie e le scienze cognitive, oltre a quelle digitali. Questa sinergia amplifica l'impatto di ciascuna.

Se per un verso, computer e algoritmi consentono di elaborare e di progettare interventi finora non immaginabili sul patrimonio genetico, per altro verso, la ricerca sui sistemi computazionali si ispira ai risultati delle neuroscienze: pensiamo alle “reti neurali”, al *machine learning* o all'idea di impiegare modelli biologici, come il

RESUMEN

Las tecnologías digitales juegan un papel decisivo en el actual «cambio de época».

Se solicita a la ética (teológica) no solo desarrollar una reflexión cada vez más transdisciplinar, sino también identificar pasos operativos que involucren a otros sujetos activos en esta área.

La *Rome Call for AI Ethics*, promovida por la Pontificia Academia para la Vida, se sitúa en esta perspectiva: promueve una reglamentación ética, fruto de una elaboración común, y la asunción de responsabilidad por parte de los firmantes.

Los representantes de las empresas e instituciones, tanto internacionales como eclesiales, que firmaron el documento se comprometen también a ampliar la red de miembros.

Palabras clave

Teología moral, tecnologías digitales, inteligencia artificial, humanismo.

DNA, per conservare dati. L'esito complessivo è tale da indurre trasformazioni riguardo a dimensioni che risultavano evidenti nello scorrere abituale della vita, ma che ora vengono messi in questione.

Ne possiamo evidenziare quattro principali, a livello di percezione diffusa: a) la crescente difficoltà a distinguere tra ambiti reali e ambiti virtuali; b) lo sfumarsi dei confini tra quanto è umano, quanto è meccanico (o artefatto)

e quanto è naturale; c) il passaggio da una situazione di scarsità di informazione a una disponibilità crescente e (sovr)abbondante di dati; d) lo slittamento da una priorità attribuita alla sussistenza isolata delle cose e alle relazioni binarie, verso la maggiore rilevanza conferita alle interazioni, ai processi e alle reti.¹²

Quindi possiamo osservare come questa nuova situazione richieda un ripensamento della concezione di noi stessi (chi siamo); delle nostre reciproche relazioni (come socializziamo); della nostra concezione della realtà (la nostra metafisica); delle nostre interazioni con la realtà (la nostra capacità di agire).¹³

In particolare, la comprensione di noi stessi che va affermandosi pone un accento prioritario sulla centralità dell'informazione: l'essere umano è anche un organismo informazionale (*inforg*) reciprocamente connesso con altri organismi simili e parte di un ambiente informazionale (*l'infosfera*), che condividiamo con altri agenti, naturali e artificiali, che processano informazioni in modo logico e autonomo. Nella realtà "iperconnessa" in cui siamo immersi non ha più neanche senso porre la domanda se si è *online* o *offline*, perché attraverso la molteplicità di dispositivi con cui costantemente interagiamo e che comunicano tra loro, spesso a nostra insaputa (come avviene per l'Internet delle cose), sarebbe più corretto dire che siamo *onlife*: un neologismo che alcuni pensatori usano per dire l'inestricabile intreccio tra vita e universo digitale.¹⁴

2. Quale etica?

All'interno di questo scenario, che mette in questione coordinate antropologiche fondamentali, siamo sollecitati ad andare oltre un'etica che consideri gli atti umani avulsi dai contesti e dai soggetti agenti, isolandoli dal piano relazionale e sociale. Si tratta qui di non limitare la nostra attenzione alla sola punta dell'*iceberg*, ma di prendere in considerazione per quanto possibile anche la sua base e, ancora più ampiamente, l'oceano in cui tutto l'*iceberg* si trova (e che gli consente di galleggiare). In questa luce, possiamo evidenziare in particolare due ambiti cruciali.

2.1. Linguaggio, comprensione, responsabilità

Nel discorso che si tiene a proposito dei nuovi dispositivi si nota un consistente - e in gran parte inavvertito - impiego di un lessico che viene abitualmente riferito a esseri umani. Così i termini passano a designare entità o procedimenti dotati di una certa automazione:¹⁵ si parla di robot che hanno esperienze percettive o cognitive, intelligenze collettive condivise nel *cloud*, capacità di prendere decisioni. Ma l'etica non può importare con disinvoltura nel proprio discorso termini che veicolano nozioni che distorcono la realtà. La stessa espressione "intelligenza artificiale" è pericolosamente equivoca: anche se non facile da correggere o da sostituire, occorre almeno esercitare un controllo critico, che è una delle forme irrinunciabili del discernimento.

Papa Francesco ha sottolineato la necessità di non rimanere invischiati in questo equivoco: «la denominazione di "intelligenza artificiale", pur certamente di effetto, può rischiare di essere fuorviante. I termini occultano il fatto che - a dispetto dell'utile assolvimento di compiti servili (è il significato originario del termine "robot") -, gli automatismi funzionali rimangono qualitativamente distanti dalle prerogative umane del sapere e dell'agire. E pertanto possono diventare socialmente pericolosi. È del resto già reale il rischio che l'uomo venga tecnologizzato, invece che la tecnica umanizzata: a cosiddette "macchine intelligenti" vengono frettolosamente attribuite capacità che sono propriamente umane. Dobbiamo comprendere meglio che cosa significano, in questo contesto, l'intelligenza, la coscienza, l'emotività, l'intenzionalità affettiva e l'autonomia dell'agire morale. I dispositivi artificiali che simulano capacità umane, in realtà, sono privi di qualità umana. Occorre tenerne conto per orientare la regolamentazione del loro impiego, e la ricerca stessa, verso una interazione costruttiva ed equa tra gli esseri umani e le più recenti versioni di macchine».¹⁶ Il punto nevralgico su cui attrarre l'attenzione è che questi dispositivi sono privi di corpo; sono fondamentalmente delle macchine computazionali che possono elaborare solo flussi astratti di dati. Il fatto che noi percepiamo comportamenti o effetti di processi dotati di automazione, che noi abitualmente interpretiamo

come l'esito di capacità intellettive, ci induce a trascurare che le macchine vi giungono attraverso processi molto diversi. Essi sono una imitazione delle apparenze.¹⁷ Anche quando sono muniti di sensori e attuatori (come avviene nei robot), gli algoritmi lavorano riducendo certi aspetti del reale a codici numerici, escludendo un'infinità di dimensioni che invece la nostra sensibilità coglie e che sfuggono ai principi di una modellizzazione matematica. Quindi la trasposizione acritica sulle macchine di un linguaggio che è adeguato per gli esseri umani amputa e distorce ciò che il processo dell'intelligenza presuppone. Quest'ultima implica infatti un intreccio indisciungibile con il desiderio che si radica nella corporeità con i suoi relativi vissuti e con la sua tensione a un apprendimento multisensoriale non sistematizzabile dell'ambiente esterno: «Per dirla semplicemente, cervello e corpo sono nella stessa barca e insieme rendono possibile la mente».¹⁸

2.2. *Far maturare l'ethos condiviso*

Un secondo ambito riguarda la riflessione sugli esiti che le nuove tecnologie nel loro complesso esercitano sull'ethos condiviso. Non è saggio lasciarsi sedurre dallo svolgimento più preciso o più efficiente o meno costoso di alcuni compiti, concentrandosi solo sulla singola prestazione e perdendo di vista l'insieme. Così come non è sufficiente limitarsi a una, pur necessaria, educazione per sensibilizzare e istruire

gli "utilizzatori finali" sul modo di impiegare gli strumenti disponibili. Infatti, come abbiamo visto, questi dispositivi sono molto più che strumenti: essi plasmano il mondo e le coscienze (anche attraverso il linguaggio). Occorre assumere responsabilmente questo tema, che riguarda l'educazione umana alle qualità non tecniche della coscienza e l'amore per il bene della comunità anche quando non se ne ricava un vantaggio. Esiste, dunque, una questione politica relativa al potere di legittimazione comunitaria dell'ethos dell'IA, che non va confuso con il potere di distribuzione dei suoi vantaggi individuali e astrattamente funzionali. In altri termini: non basta semplicemente educare le singole persone a un «consumo critico» dei prodotti tecnologici, né limitarsi a ragionare sulla sensibilità morale di chi fa ricerca e progetta sistemi e algoritmi, affidandoci ad essa.

Occorre invece lavorare per creare corpi sociali intermedi che assicurino rappresentanza alla sensibilità etica degli utilizzatori, in tutte le fasi del processo. Si tratta insomma di evitare di assegnare un ruolo dogmatico e dirigistico sia al totalitarismo politico sia al liberismo tecnocratico. Stiamo iniziando a intravedere una nuova disciplina che riguarda "lo sviluppo etico degli algoritmi" o più semplicemente "algor-etica".¹⁹ Rimane aperta naturalmente la ricerca di come realizzare tali obiettivi.²⁰ E questo ci conduce al nostro passo successivo.

3. Raccordare tecnologia, antropologia ed etica

Quindi, come nel momento analitico occorre favorire uno sguardo ampio e complessivo, così sul piano etico non è sufficiente circoscrivere l'attenzione al controllo dei singoli dispositivi. Una visione ristretta e parziale ci porterebbe a un richiamo astratto e generico al rispetto dei diritti soggettivi, della dignità della persona e di quei principi che la Dottrina sociale della Chiesa molto opportunamente ci mette a disposizione (come la giustizia, la solidarietà e la sussidiarietà). Questi principi mantengono certo tutta la loro validità, ma la complessità del mondo tecnologico contemporaneo ci chiede di elaborarne una interpretazione che, da una parte, sia sufficientemente precisa - in modo da evitare che siano strumentalizzati da una *business ethics* centrata sull'interesse economico - e, dall'altra, possa renderli incisivi nella concretezza delle situazioni.

3.1. Dignità e libertà

Abbiamo visto che il nuovo livello dell'intermediazione tecnologica cosiddetta "intelligente" tende a comprimere, se non a escludere, le valutazioni del singolo, cortocircuitando così questa fondamentale espressione della sua dignità. In questa fase, appare ormai chiaro che "l'umano" è condizionato in modo da "assecondare" il dispositivo IA, molto più che il contrario: è il dispositivo stesso a plasmare l'utente come "degnò" e "libero" di farne uso. Il *feed-back* ca-

ratteristico dell'IA è appunto l'alimentazione e la sofisticazione di questo meccanismo selettivo di conformità sociale al dispositivo. Si mantiene così un circolo costrittivo che, con lo scopo di condizionare e dirigere i comportamenti, trae vantaggio proprio dall'inconsapevole assoggettamento dell'utente che lo alimenta.²¹ Questa problematica esige un nuovo paradigma, per l'inquadramento della eticità praticabile in tale ambito. Tra l'altro osserviamo come giunga qui a un paradosso il mito neoliberale che, in nome della priorità assegnata all'autodeterminazione, conduce a espropriare la altrui (e alla fine anche la propria) capacità di libera decisione.²²

Il compito che si profila è di individuare modalità di monitoraggio che articolino diverse discipline per la ricerca condivisa di un'etica che sia effettivamente transdisciplinare.²³ Occorre accompagnare l'intero percorso in cui intervengono le diverse competenze nell'elaborazione degli apparati tecnologici (ricerca, progettazione, produzione, distribuzione, utilizzo individuale e collettivo). È una mediazione ormai indispensabile, vista la capacità della strumentazione IA di determinare vere e proprie forme di controllo e orientamento delle abitudini mentali e relazionali, e non solo di potenziamento delle funzioni cognitive e operative. Vanno esaminate dai diversi punti di vista le ricadute prevedibili dei singoli momenti del processo in corso. L'obiettivo è quello di assicurare una verifica competente e

condivisa dei processi secondo cui si integrano i rapporti tra gli esseri umani e le macchine nella nuova era aperta dalla IA. È questo un compito che richiede disponibilità al dialogo e alla collaborazione. Non mancano segni di attenzione e di impegno su questo fronte. Un esempio interessante in questa linea ci è dato da *Ingegneri senza frontiere*. L'organizzazione, sorella minore di *Medici senza frontiere*, ha proposto un manifesto perché gli ingegneri, dato il ruolo strategico che svolgono nel mondo delle tecnologie, assumano responsabilità anche per le dimensioni sociali del loro lavoro. Essi propongono tra l'altro di integrare nel percorso universitario discipline che introducano diverse prospettive e aiutino a comprendere le implicazioni degli artefatti tecnici nelle relazioni sociali e nella sfera politico-culturale.²⁴ Sarà così possibile anche riflettere sugli strumenti disponibili per valutare l'innovazione tecnologica in vista di scelte politiche eticamente consapevoli. Si tratta quindi di garantire un effettivo esercizio della libertà e il riconoscimento della dignità. Per esempio, come sarà tutelato chi decide di non usare certi dispositivi di assistenza, che considera non consoni alla propria concezione della persona e delle relazioni umane? Questa libertà cresce simmetricamente con la disponibilità dei vantaggi potenziali del suo uso oppure il suo vincolo costrittivo cresce fatalmente con la sua potenza tecnologica? Si potrà scegliere una infermiera umana, anche rinunciando

a qualche efficienza, in cambio di umanità, oppure il sistema sanitario fornirà obbligatoriamente agenti robotici da prendere o lasciare? Per chi viene introdotto in protocolli sanitari, anche diagnostici, saranno garantiti criteri di imparzialità o i *bias* che affliggono gli algoritmi avranno effetti discriminatori?²⁵ Analogamente potremmo argomentare riguardo ai *Big data*: sarà possibile per tutti e per le differenti comunità una qualche modalità intuitiva e praticabile di accesso ai *Big data*, più o meno analoga a quella che oggi consente anche agli anziani di accedere a Internet, per comporsi il loro quadro di informazione, ricerca, contatti? Oppure con lo sviluppo dell'IA, l'accesso e l'elaborazione saranno selettivamente riservate alle grandi *holding* economiche, ai sistemi di pubblica sicurezza, agli attori della *governance* politica? In altri termini, si potranno svolgere delle ricerche di informazioni in modo equo oppure la sofisticazione dei servizi sarà automaticamente sottratta a chi vive ordinariamente nella società e alla comunità cui appartiene, aumentando le già molteplici forme di divario?

3.2. Differenze e culture

Considerando poi la famiglia umana nel suo complesso, dobbiamo avere cura di come possono essere rispettate le differenze tra comunità. Occorre da una parte evitare che l'IA ne tenga conto solo in funzione dell'affinamento dei suoi dispositivi regolativi, e dall'altra insistere perché sia garantito lo spazio di intervento, per-

ché il soggetto comunitario possa determinare gli effetti collettivi di un'automazione sofisticata e autogestita dei prodotti, delle funzioni, dei servizi. Certo, "di diritto" è giusto proclamare l'uguaglianza, la razionalità, la coscienza, la libertà, come dimensioni paritetiche e universali della dignità e della legalità, che riguardano ogni singolo senza distinzioni. "Di fatto" però, le diversità non sono accessori dell'umano, ma costituiscono componenti della sua esistenza reale. Ci sono diversità indesiderabili (deficit intellettivi, deprivazioni sociali, vuoti educativi, vulnerabilità sistemiche). Ma ci sono diversità delle quali viviamo (la nostra rete familiare, la nostra dedizione religiosa, il nostro lavoro, il nostro ruolo, il nostro *habitat*). Confondere queste diversità è sbagliato, come è ugualmente erroneo ignorare il loro handicap permanente o, rispettivamente, la loro funzione antropologica. Tale possibile confusione espone il discorso etico, giuridico, politico a una pericolosa astrattezza.

Va favorito l'impegno a costruire modi di vita che non puntino all'eliminazione di ogni vulnerabilità, con l'intento immaginario di liberarci dalle nostre affezioni a vantaggio di un'organizzazione ideale delle cose, inseguendo una forma compiuta e perpetua di (illusoria) perfezione. È invece importante garantire lo spazio per il riconoscimento della fondamentale imperfezione propria dell'esistenza e per una sua assunzione riconciliata. Al contempo la diversità degli individui

va celebrata, come del resto anche l'autonomia responsabile della volontà, la nostra capacità di apprendimento multisensoriale della realtà, edificando così modalità di stare insieme nella reciproca ospitalità.²⁶ Qui potrà avere un ruolo importante la teologia, intesa come sapere critico dell'esperienza della fede cristiana. Essa potrà introdurre, anche nel dibattito pubblico, una riflessione che eviti l'appiattimento sui criteri operativi e funzionali che tendono a prendere il sopravvento: occorre mantenere viva la ricerca del pensiero per una elaborazione antropologica umanistica, che faccia valere il richiamo alla dignità umana in quanto istanza trascendente, cioè da una parte non determinata aprioristicamente e dall'altra effettivamente implicata nelle forme concrete dell'agire tecnico.²⁷ Come del resto la riflessione teologica può contribuire a una interpretazione del limite e della passività non solo come ostacoli da abbattere, ma come dimensioni costitutive della relazione interpersonale e pertanto da includere in ogni attività umana, anche tecnico-scientifica.²⁸

4. Alla ricerca di passi concreti e condivisi

L'ampiezza e la profondità dei cambiamenti in atto non permettono di farsi illusioni sulla oggettiva complessità del compito che ci attende. Nessun singolo soggetto o nessuna realtà organizzativa può presumere di poter affrontare isolatamente le questioni in cui ci troviamo immersi.

Alle molteplici provocazioni che si presentano, possiamo aggiungere anche quella di promuovere pratiche di collaborazione. La Pontificia Accademia per la Vita, a partire da molteplici sollecitazioni che le sono state rivolte da diversi soggetti operanti nell'ambito della ricerca e della produzione di sistemi digitali, ha lanciato una iniziativa che intende mettere in opera quanto si va elaborando teoricamente. Il progetto ha preso consistenza nella *Rome Call for AI Ethics*, presentata pubblicamente il 28 febbraio 2020.²⁹

4.1. Significato della "Call"

Il testo su cui la *Call* si basa è il risultato di una comune elaborazione in cui i partecipanti hanno individuato alcuni riferimenti condivisi, anche nella ricerca di un linguaggio comprensibile per tutti. I diritti umani sono stati di aiuto, sia quanto a contenuti, sia come esempio di convergenza possibile a un livello intermedio. Essi infatti, pur non essendo indipendenti da presupposti fondativi sul piano antropologico, consentono tuttavia di trovare un terreno di possibile convergenza di diverse visioni del mondo, a livello sia culturale sia religioso. E, d'altra parte, non entrano nella formulazione di norme giuridiche troppo dettagliate e particolari.

Il documento di sviluppa secondo tre grandi direttrici. La prima è quella etica, che richiama il quadro di valori fondamentali sotteso alla *Dichiarazione universale dei diritti umani*. Viene così affermato un contesto di

riferimento necessario per ogni sviluppo tecnologico eticamente valido. Sono in particolare sottolineati: l'inclusione, la contemporanea attenzione al bene dell'intera umanità e di ogni essere umano, il rispetto e la salvaguardia del pianeta, "nostra casa comune e condivisa". La seconda direttrice riguarda l'educazione per le giovani generazioni, che sarà segnata profondamente dalla disponibilità delle nuove risorse tecnologiche e a cui va garantito un accesso senza disuguaglianze. Data la velocità della trasformazione, occorrerà anche un'educazione permanente, soprattutto per chi rischia di essere lasciato indietro. La formazione dovrà includere pure l'attenzione alla coscienza e alle motivazioni che la rendono capace di impegnarsi per il bene della comunità, anche a scapito del proprio interesse. L'ultima direttrice è infine quella del diritto. Qui emerge l'esigenza di tradurre i principi enunciati in regolamentazioni effettive e di renderli incisivi attraverso un approccio etico *by design*, cioè che accompagni ogni passaggio del ciclo produttivo delle tecnologie, fin dall'inizio.

4.2. Principi di riferimento

Per realizzare gli obiettivi formulati e dare qualche indicazione più precisa su come operare in modo etico nell'ambito dell'IA, sono stati messi a fuoco alcuni principi: 1) *Trasparenza*: in linea generale, i sistemi di intelligenza artificiale devono essere spiegabili; 2) *Inclusione*: i bisogni di tutti gli esseri umani devono essere presi

in considerazione in modo che ognuno possa trarne beneficio e a tutti gli individui possano essere offerte le migliori condizioni possibili per esprimersi e svilupparsi; 3) *Responsabilità*: coloro che progettano e mettono in opera queste tecnologie devono procedere con responsabilità e trasparenza; 4) *Imparzialità*: evitare di creare o di agire secondo pregiudizi, salvaguardando così l'equità e la dignità umana; 5) *Affidabilità*: i sistemi di intelligenza artificiale devono essere in grado di funzionare in modo attendibile; 6) *Sicurezza e privacy*: i sistemi di intelligenza artificiale devono funzionare in modo sicuro e rispettare la riservatezza degli utenti. Questi principi sono riferimenti fondamentali per una buona innovazione, anche se naturalmente la loro declinazione operativa non è automatica e richiede una continua attenzione. Sottolineiamo come essi siano in sintonia con i documenti elaborati da vari organismi dell'Unione Europea:³⁰ è molto interessante che siano risultati condivisibili anche da grandi imprese del mondo statunitense, che anzi ne hanno elogiato la validità in diversi contesti.³¹ La presentazione del documento si è conclusa con la sua firma da parte del presidente di Microsoft, Bradford Lee Smith, del vicepresidente esecutivo IBM, John Kelly III, del direttore generale della FAO, Qu Dongyu e del ministro del governo italiano per l'innovazione tecnologica e la digitalizzazione, Paola Pisano. Era presente anche il presidente del Parlamento

europeo David Sassoli, che ha espresso l'interesse delle istituzioni europee per il processo inaugurato, sottolineando l'importanza di una collaborazione internazionale degli sponsor della *Call*. Come si vede dalla varietà dei ruoli e dei profili dei firmatari, sono coinvolti differenti portatori di interessi, provenienti dal mondo produttivo, istituzionale, politico, scientifico e accademico. Ognuno nel suo specifico ruolo, come è emerso dai loro rispettivi interventi, si impegna a diffondere la *Call* in modo che anche altri ne divengano firmatari, condividendo così sia la ricerca per una comprensione sempre più approfondita e condivisa dei cambiamenti che stiamo vivendo, sia l'assunzione delle responsabilità che ne derivano.

Conclusione

In questa linea possiamo cogliere, anche se fra molte contraddizioni e tensioni, la graduale maturazione di un comune impegno per un autentico sviluppo, che intende il progresso in termini non solo economici o quantitativi. Il pur indispensabile contributo del sapere tecnico-scientifico viene infatti articolato con altre prospettive, incluse quelle della teologia e dell'etica, in modo da orientarci, con la collaborazione di tutte le persone di buona volontà, verso uno sviluppo umano effettivamente integrale.

NOTE

¹ Mons. Vincenzo Paglia è Presidente della Pontificia Accademia per la Vita e Gran Cancelliere del Pontificio Istituto Teologico Giovanni Paolo II per le Scienze del Matrimonio e della Famiglia di Roma.

² Cf FRANCESCO, *Il "Buon" Algoritmo? Intelligenza Artificiale, Etica, Diritto, Salute*. Discorso alla XXVI Assemblea Generale della Pontificia Accademia per la Vita, Roma 28 febbraio 2020, in http://www.vatican.va/content/francesco/it/speeches/2020/february/documents/papa-francesco_20200228_accademia-perla-vita.html (28-12-2020).

³ Cf ELLUL Jaques, *La Technique: L'Enjeu du siècle*, Paris, Armand Colin 1954.

⁴ ELLUL Jacques, *Le bluff technologique*, Paris, Hachette 2012 (edizione originale 1988), 92. Per una sintetica presentazione critica del pensiero dell'autore cf CHIODI Maurizio - REICHLIN Massimo, *Morale della vita. Bioetica in prospettiva filosofica e teologica*, Brescia, Queriniana 2017, 62-66.

⁵ MIETH Dietmar, *Cosa vogliamo potere? Etica nell'epoca della biotecnica*, Brescia, Queriniana 2003.

⁶ Per un più ampio inquadramento delle questioni sottese a questa scelta lessicale cf BENANTI Paolo, *La condizione techno-umana*, Bologna, EDB 2016, 11-22.

⁷ Id., *Le macchine sapienti. Intelligenze artificiali e decisioni umane*, Bologna, Marietti 2018. A questo uso al plurale si oppongono coloro che ritengono l'IA una disciplina collegata alla *computer science*: «the branch of computer science that deals with creating intelligent machines capable of performing new tasks and resolving new problems, of adapting to the environment and understanding it, and of understanding natural language» (CIFERRI Ludovico - BELTRAMETTI Monica - FLORIDI Luciano - TRIVELLATO Piero, *Artificial Intelligence as a new growth factor*, Aspen Institute 2017, in https://www.myecole.it/biblio/wp-content/uploads/2020/11/3DK_4AI_Artificial-Intelligence-ENG.pdf 3) [28-12-2020].

⁸ Sulla circolarità dei rapporti tra invenzione tecnologica, strutture sociali e impatto politico cf WINNER Langdon, *Do Artifacts Have Politics?*, in *Daedalus* 109(1980)1, 121-136.

⁹ Cf CARROZZA Maria Chiara, *I Robot e noi*, Bologna, Il Mulino 2017, 16-17.

¹⁰ Papa Francesco ha sollecitato la PAV ad approfondire queste tematiche nella Lettera che ha inviato per il 25° anniversario della sua fondazione: cf FRANCESCO, *Humana communitas*. Lettera del Santo Padre Francesco al Presidente della Pontificia Accademia per la Vita in occasione del XXV anniversario della sua istituzione, 6 gennaio 2019, n. 12, in http://www.vatican.va/content/francesco/it/letters/2019/documents/papa-francesco_20190106_lettera-accademia-vita.html (20-12-2020). Per una presentazione di questo testo cf CASALONE Carlo, «*Humana communitas*». *La vita nella trama delle relazioni*, in *La Civiltà Cattolica* 1(2019)2, 209-222.

¹¹ Il prefisso "nano" indica una misura corrispondente a 10⁹ metri, cioè un milionesimo di millimetro.

¹² Cf FLORIDI Luciano (ed.), *The Onlife Manifesto. Being Human in a Hyperconnected Era*, Heidelberg - New York - Dordrecht - London, SpringerOpen 2015, ed. Kindle, pos. 243.

¹³ Il compito che ci attende è quindi quello di ricomprendere come categorie che finora hanno avuto un rilievo nell'interpretazione di noi stessi, del mondo e delle nostre relazioni interpersonali e comunitarie richiedano una profonda revisione (o, come direbbero i filosofi, di essere "risemantizzate"). Si tratta, come si vedrà, di concetti chiave come attenzione, proprietà-titolarietà (*ownership*), privacy, agire responsabile (*agency*).

¹⁴ Cf FLORIDI Luciano, *La quarta rivoluzione. Come l'infosfera sta trasformando il mondo*, Milano, Cortina 2017.

¹⁵ L'automazione indica quella caratteristica per cui processi o funzioni risultano «indipendenti dall'azione umana, ma al tempo stesso controllabili, ripetibili ed efficienti» (CARROZZA, *I Robot e noi* 21).

¹⁶ FRANCESCO, Discorso alla XXV Assemblea Generale della PAV, 25 febbraio 2019, in PAGLIA Vincenzo - PEGORARO Renzo (edd.), *Robo-ethics. Humans, Machines and Health*, Roma, PAV 2020, 18-19.

¹⁷ «On the one hand, there is a reduction of intelligence to only behaviour. On the other hand, 'behaving machines' are shaped by

and contributing to sociocultural context, thus enforcing the connection with human beings» (SINIBALDI Edoardo et al., *Contributions from the Catholic Church to ethical reflections in the digital era*, in *Nature Machine Intelligence* 2[2020], 242-244, in <https://doi.org/10.1038/s42256-020-0175-4> [27-12-2020]).

¹⁸ DAMASIO Antonio, *Lo strano ordine delle cose*, Milano, Adelphi 2018, 274; cf anche LEXCELLENT Christian, *Pourra-t-on pirater l'être humain? IA versus intelligence humaine*, Toulouse, Cépaduès 2019, in particolare l'ultimo capitolo il cui titolo (tradotto) suona «La verità dell'IA: la sparizione del corpo»; BONACCORSO Giorgio, *Critica della ragione impura. Per un confronto tra teologia e scienza*, Assisi, Cittadella 2016, 69-101 e 191-210; VANTINI Lucia, *Il sé esposto. Teologia e neuroscienze in chiave fenomenologica*, Assisi, Cittadella 2017, 23-104.

¹⁹ Cf BENANTI Paolo, *Oracoli. Tra algoretica e algocrazia*, Roma, Luca Sossella editore 2018.

²⁰ Il tema è sentito e vari contributi forniscono indicazioni interessanti: cf ad es. DIGNUM Virginia, *Responsible Artificial Intelligence. How to Develop and Use AI in a Responsible Way*, Cham (Switzerland), Springer 2019; JOBIN Anna - IENCA Marcello - VAYENA Effy, *Artificial Intelligence: the global landscape of ethics guidelines*, in *Nature Machine Intelligence* 1(2019) 389-399, in <https://doi.org/10.1038/s42256-019-0088-2> (27-12-2020); KAEBNICK Gregory E. - GUSMANO Michael K. (a cura di), *Governance of Emerging Technologies*, New York, The Hastings Center 2018; BRUNDAGE Miles, *Artificial Intelligence and Responsible Innovation*, in MÜLLER Vincent C. (a cura di), *Fundamental Issues of Artificial Intelligence*, Switzerland, Springer 2016, 541-552.

²¹ Cf SADIN Eric, *Critica della ragione artificiale: una sfida dell'umanità. Una difesa dell'umanità*, Roma, Luiss University Press 2019 (ed. Kindle), pos. 165.

²² Cfr ZUBOFF Shoshana, *Il capitalismo della sorveglianza. Il futuro dell'umanità nell'era dei nuovi poteri*, Roma, LUISS 2019.

²³ FRANCESCO, *Costituzione apostolica Veritatis gaudium* circa le Università e le Facoltà ecclesiarie, 8 dicembre 2017, 4c.

²⁴ Cfr INGÉNIEURS SANS FRONTIÈRES, *Manifeste pour une formation citoyenne des ingénieur-e-s*,

in <https://www.isf-france.org/node/1211> (27-12-2020); cf anche KANT Vivek - KERR Eric, *Taking Stock of Engineering Epistemology: Multidisciplinary Perspectives*, in *Philosophy & Technology* 32(2019)4, 685-726.

²⁵ Cf RIGOLI Felix H., *Artificial Intelligence in the road of Health for All. Perils and Hope*, in PAGLIA Vincenzo - PEGORARO Renzo (a cura di), *The Good Algorithm? Artificial Intelligence, Ethics, Law, Health*, PAV, Roma [in via di pubblicazione].

²⁶ Per una suggestiva elaborazione letteraria di questa prospettiva cf MCEWAN Ian, *Macchine come me* [Machines Like Me and People Like You, Jonathan Cape 2019], Torino, Einaudi 2019.

²⁷ Cf MANZONE Gianni, *Morale artificiale. Nanotecnologie, intelligenza artificiale, robot. Sfide e promesse*, Bologna, EDB 2020, 8 e 195-233.

²⁸ Cf SEQUERI Pierangelo, *La fede e la giustizia degli affetti. Teologia fondamentale della forma cristiana*, Siena, Cantagalli 2019, 162-171.

²⁹ Il testo e i diversi interventi che hanno avuto luogo durante la presentazione sono disponibili, in https://romecall.org/_/27-12-2020.

³⁰ Cf GRUPPO DI ESPERTI DELLA COMMISSIONE EUROPEA, *Orientamenti etici per una IA affidabile*, European Union 2018; EUROPEAN COMMISSION, *Artificial Intelligence. A European Perspective*, European Union 2018.

³¹ Cf in particolare SMITH Brad, *Tools and Weapons, The Promise and the Peril of the Digital Age*, New York (NY), Penguin Press 2019, 131-150, dove il presidente di Microsoft esamina la *General Data Protection Regulation*, adottata dell'Unione Europea nel 2016, e ne valuta molto positivamente gli effetti, anche per la conversione tecnologica cui ha condotto le aziende.