

Integrare la bioetica con l'etica della tecnologia per formare i professionisti del futuro

di Viola Schiaffonati

Introduzione

È indubbio che scienza e tecnologia abbiano sempre giocato un ruolo fondamentale nell'aprire nuove possibilità di cambiamento per la vita umana e, di conseguenza, abbiano profondamente influenzato anche la pratica medica e le riflessioni bioetiche a essa connesse. Recentemente alcune tecnologie informatiche, come l'intelligenza artificiale (IA)¹, hanno guadagnato una certa centralità sia nel dibattito accademico sia in quello pubblico, promettendo di rivoluzionare la medicina e di superare alcuni dei problemi che l'affliggono, come per esempio l'alto tasso di errori nella diagnosi. Accanto alla fiducia in tali promesse – spesso roboanti e non fondate su solide evidenze – sono emerse anche diverse preoccupazioni, come quelle relative ai rischi che comporta automatizzare decisioni complesse e l'impatto che tale automatizzazione potrebbe avere su diverse categorie di persone. Il tema della cosiddetta medicina digitale e delle sue ricadute etiche è già stato ampiamente affrontato. A questo proposito per esempio (Blasimme 2021), oltre a mostrare la problematicità della definizione stessa di medicina digitale, ha rilevato come i temi più ampiamente dibattuti nell'ambito delle questioni etiche siano al momento tre: la privacy e la protezione dei dati personali; le questioni relative all'apprendimento automatico; la trasparenza degli algoritmi di IA.

¹ La disciplina intelligenza artificiale può essere definita in vari modi e nel corso della sua pur breve storia sono state proposte diverse definizioni. In questo capitolo mi attengo a una definizione classica che considera l'intelligenza artificiale la disciplina, appartenente all'area dell'informatica, che sviluppa sistemi artificiali in grado di svolgere attività complesse, che nelle persone richiedono intelligenza (percezione, ragionamento, apprendimento, comunicazione o azione), senza l'intervento diretto dell'essere umano (Nilsson 2010).

L'obiettivo di questo capitolo non è di trattare per esteso le questioni etiche della medicina digitale, ma di mostrare come alcuni aspetti della bioetica tradizionale debbano essere rivisti alla luce dell'attuale dibattito nell'etica della tecnologia. In altre parole, la dimensione tecnologica che caratterizza oggi la pratica medica, e più in generale l'ambito della salute, deve informare anche la riflessione etica ad essa connessa. In questo capitolo mi concentro quindi su due specifici aspetti che emergono, in particolare, con l'avvento delle tecnologie di IA: l'approccio incrementale all'etica e le sfide connesse alla moralizzazione delle tecnologie. Questi elementi non sono specifici della medicina; tuttavia, entrambi possono avere un impatto decisivo sulla pratica medica e sulle professionalità delle persone coinvolte. Questo impatto è tanto più decisivo in un momento come quello che stiamo vivendo, in cui la rivoluzione digitale sembra essere a un punto di svolta in grado di modificare le vite dei singoli e delle società.

Questo capitolo è profondamente informato dall'esperienza dell'insegnamento del modulo *Bioethics* nel corso *Bioethics, Professionalism and History of Medicine* nel corso di Laurea in Bioingegneria e Medicina (*MedTech*) nato dalla collaborazione fra Politecnico di Milano e Humanitas University. La progettazione e l'insegnamento di questo modulo mi hanno reso possibile riflettere più nello specifico su quali siano le domande etiche che lo sviluppo della tecnologia in medicina chiama in causa e quale possa essere il modo migliore per sostenere la riflessione critica delle studentesse e degli studenti di fronte a questi temi.

Dai quattro principi della bioetica alle sedici condizioni per le tecnologie sperimentali

La bioetica, intesa come il contesto di analisi delle questioni etiche connesse alla salute e al benessere umani, si fonda tradizionalmente su quattro principi (Beauchamp & Childress 2012): rispetto per l'*autonomia* umana, *non maleficenza*, *beneficenza* e *giustizia*. Il rispetto per l'autonomia deve essere inteso non semplicemente come attitudine alla tolleranza delle scelte altrui, ma come una forma di azione tesa ad assicurare che sia difesa e promossa la capacità altrui di decidere. I principi di *non maleficenza* e di *beneficenza* traggono origine dal noto motto latino "*Primum non nocere*" e si applicano non solo all'etica medica, nella misura in cui la salute di un paziente deve sempre essere prioritaria, ma anche più in generale alla salute umana, come per esempio alla possibilità di evitare l'inquinamento o il riscaldamento globale. Tali principi prescrivono quindi non solo di non fare del male agli altri ma anche di cercare di prevenire questo male, suggerendo che i possibili danni debbano essere controbilanciati dai benefici. Infine, il principio di *giustizia* prescrive che i costi e i benefici, per esempio di una cura o di un trattamento, debbano essere equamente distribuiti.

Questi principi aspirano a esprimere una sorta di moralità comune a tutte le diverse culture che, almeno in linea di principio, possa essere sottoscritta da tutti. Sebbene questa idea sia stata criticata da più parti, si deve comunque riconoscere

che essa funge da agevole punto di partenza in molti contesti dell'etica medica e biomedica, soprattutto quando utilizzata in modo costruttivo e flessibile, e non come soluzione a tutti i problemi morali.

L'avvento di una pratica medica sempre più intrecciata all'impiego di tecnologie complesse ha fatto emergere una vasta riflessione sulla necessità di rivedere, ampliare o sostituire i quattro tradizionali principi della bioetica. L'IA in particolare sembra offrire, da un lato, uno stimolo a riconcettualizzare alcuni temi, dall'altro, mette in discussione alcune delle categorie tradizionali. Per questo motivo, negli ultimi anni è emersa una massa di linee guida e di codici etici e di comportamento tesi a indirizzare lo sviluppo e l'uso di strumenti di IA in modo etico (Jobin et al. 2018). Sebbene le diverse proposte abbiano spesso punti di partenza diversi, è significativo notare che spesso i quattro principi fondamentali della bioetica vengono mantenuti, anche se arricchiti da principi come la spiegabilità. Floridi & Cowls 2019 sostengono per esempio che il principio della spiegabilità deve essere aggiunto ai quattro principi della bioetica nei casi in cui occorra regolare eticamente l'uso dell'IA. Per spiegabilità i due autori intendono sia l'intelligibilità dell'IA, ossia la possibilità di comprendere come essa funzioni, sia la responsabilità, ossia la capacità di capire chi è responsabile del modo in cui l'IA funziona.

Tra le diverse proposte avanzate in tal senso, in questo capitolo mi concentro su un *framework* etico che, partendo dai quattro principi della bioetica, propone la loro articolazione in sedici diverse condizioni. Queste condizioni servono ad applicare i principi generali al caso specifico delle cosiddette tecnologie sperimentali (van de Poel 2016). La ragione della mia scelta risiede principalmente nel carattere incrementale di questa proposta, ossia nel riconoscere che *le tecnologie sperimentali* (ossia quelle tecnologie per le quali la conoscenza operativa è limitata ed è quindi difficile valutarne a priori rischi e benefici sociali) debbano essere gradualmente introdotte nel loro contesto d'uso. La ragione di ciò sta nella possibilità di monitorare in maniera costante i loro effetti, che non possono essere del tutto anticipati nella fase di progettazione. Questo *framework* non si concentra su nessun contesto d'uso in particolare, ma si può applicare a diverse situazioni. L'ambito medico assume senza dubbio una rilevanza centrale data la particolare vulnerabilità delle persone coinvolte quando vengono impiegate tecnologie sperimentali e quindi la necessità di una gradualità nella loro introduzione.

Un elemento centrale delle tecnologie sperimentali, di cui l'IA fa a buon diritto parte, è la loro *opacità epistemica*: tali sistemi sono estremamente complessi sia come artefatti tecnici (Vermaas et al. 2011) sia nelle loro interazioni con l'ambiente circostante, soprattutto quando sono coinvolti esseri umani. Si può quindi affermare che essi presentano diversi gradi di *incertezza* che rendono difficile prevedere il comportamento che avranno nel loro contesto d'uso e le possibili conseguenze. **Pensiamo per esempio al caso di un robot autonomo che opera nel contesto ospedaliero a supporto di diverse attività, come l'assistenza**

alle persone ricoverate. Oppure ai sistemi di IA in ambito diagnostico, e in particolare agli algoritmi di apprendimento automatico (*machine learning*) utilizzati nell'analisi delle immagini cliniche e radiologiche oppure per elaborare previsioni circa il decorso di una malattia sulla base della storia clinica dei pazienti.

A causa di questa incertezza i rischi connessi all'impiego di tecnologie sperimentali sono difficili, e in alcuni casi impossibili, da prevedere. Date queste difficoltà, può accadere che una volta che si siano materializzate le conseguenze negative, sia ormai troppo tardi per modificare lo sviluppo e l'impatto delle tecnologie sperimentali. L'approccio incrementale insiste quindi proprio sul fatto che, sebbene possano sempre accadere eventi inaspettati, le tecnologie sperimentali, proprio per l'altro grado di incertezza che le contraddistingue, dovrebbero essere introdotte gradualmente nei loro contesti d'uso. L'idea è di poter monitorare gli effetti sociali che esse producono ed essere così in grado di interrompere il loro utilizzo e di modificare le loro funzionalità in modo da evitare il più possibile tali effetti negativi. Gli effetti negativi possono essere legati alla mancanza di sicurezza e affidabilità delle applicazioni mediche legate all'IA, ma anche alla mancanza di trasparenza degli algoritmi utilizzati in questo ambito, allo scarso livello di rappresentatività dei dati che vengono utilizzati per addestrare gli algoritmi di apprendimento automatico, nonché ai *bias* che essi contengono e che impattano sui loro risultati. Per tutte queste ragioni l'approccio incrementale concettualizza l'introduzione delle tecnologie sperimentali nella società come un *esperimento sociale* (Krohn & Weyer 1994), nel quale la società possa fungere da laboratorio per testare le tecnologie sperimentali.

Consideriamo ora più nel dettaglio il *framework* etico proposto per gli esperimenti che coinvolgono tecnologie sperimentali (van de Poel 2016). Le sedici condizioni proposte possono essere suddivise in tre gruppi corrispondenti ai principi bioetici di non maleficenza, di beneficenza, e di rispetto per l'autonomia e giustizia. Vale la pena ricordare che queste condizioni non sono rigide ma costantemente aperte al miglioramento sulla base dell'esperienza accumulata dalla loro applicazione a situazioni concrete.

Le condizioni relative al principio morale di *non maleficenza* sono:

1. *Assenza di altri mezzi ragionevoli per ottenere informazioni sui rischi e sui benefici.* Questa condizione richiede che, prima che una tecnologia sperimentale, come per esempio un robot autonomo per l'assistenza agli anziani, sia impiegata in un determinato contesto, siano stati esplorati tutti gli altri possibili modi per acquisire conoscenza sul possibile comportamento del sistema e sui rischi a esso associati. Adottando una prospettiva incrementale, prima di rilasciare i robot nel loro reale contesto d'uso, occorre che siano fatti esperimenti in laboratorio o in ambienti controllati per valutare le loro prestazioni in varie condizioni. L'obiettivo finale è di avere un sistema robotico che non causi danni, o perlomeno minimizzi questa possibilità, una volta che sia stato adottato.

2. *Monitoraggio dei rischi e dei danni.* Questa condizione richiede che le operazioni di una tecnologia sperimentale siano monitorate per evitare danni alle

persone e per proteggere la loro privacy.

3. *Possibilità e disponibilità a modificare o a fermare l'esperimento.* Questa condizione è collegata alla condizione 2 e richiede che le attività di una tecnologia sperimentale possano sempre essere fermate o modificate in caso di possibili danni alle persone.

4. *Contenimento del rischio.* Questa condizione richiede che, almeno in linea di principio, il rischio sia il più possibile contenuto. Tuttavia è evidente che una situazione completamente priva di rischio non è mai possibile. Più realisticamente questa condizione richiede, quindi, che siano state prese tutte le misure per contenere il rischio.

5. *Ridimensionamento volontario per evitare un danno su larga scala e migliorare quanto appreso.* Questa condizione richiede che una tecnologia sperimentale sia introdotta gradualmente nell'ambiente in cui deve operare. In questo modo si può incrementare il grado di conoscenza del funzionamento del sistema in modo graduale e cercare così di ridurre il rischio su larga scala.

6. *Organizzazione flessibile dell'esperimento e tentativo di evitare di rimanere bloccati su una determinata tecnologia.* Questa condizione richiede maggiore flessibilità nelle diverse fasi di realizzazione di un esperimento per incrementare via via la sua complessità. In assenza di flessibilità, un errore possibile è quello di rimanere bloccati su una specifica opzione tecnologica senza considerare altre possibilità. È infatti piuttosto comune ignorare valide alternative tecnologiche una volta che si sia fatta una determinata scelta, senza considerare che scelte diverse potrebbero portare a dei benefici.

7. *Evitare esperimenti che possano minare la resilienza.* Questa condizione esprime l'idea che il parametro di minimizzazione del rischio non possa essere misurato in termini assoluti, ma dipenda invece dalla resilienza di ciascuno degli individui coinvolti in una situazione di rischio non atteso. Per esempio, minimizzare il rischio per robot che operano per l'assistenza agli anziani in un ambiente ospedaliero è ben diverso dal doverlo fare nel caso in cui le persone assistite non siano anziate e vivano nelle loro case.

Le condizioni relative al principio morale di *beneficenza* sono:

8. *Aspettativa di ricevere benefici sociali dall'esecuzione dell'esperimento.* Questa condizione suggerisce che l'introduzione di una tecnologia sperimentale in uno specifico contesto dovrebbe fornire benefici tangibili per la società, per esempio aumentando la possibilità di supporto per i pazienti o la qualità della diagnosi e delle terapie. Dato che l'uso di una tecnologia sperimentale non è privo di rischi, i vantaggi per la società dovrebbero compensare i rischi.

9. *Distribuzione chiara delle responsabilità nel progettare, eseguire, monitorare, valutare, adattare e terminare l'esperimento.* Questa condizione è particolarmente critica, dato che richiede una netta distinzione fra i ricercatori che hanno una specifica formazione in merito al processo sperimentale e i soggetti dell'esperimento, ossia le persone che sono coinvolte nell'esperimento e che invece solitamente non hanno alcuna specifica formazione in merito.

Le condizioni relative ai principi morali del *rispetto dell'autonomia* e della *giustizia* sono:

10. *I soggetti sperimentali devono essere informati.* Questa condizione si applica alle situazioni in cui dei soggetti umani sono coinvolti negli esperimenti. Naturalmente in questi casi è necessario il consenso informato. Tuttavia, dato che spesso le tecnologie sperimentali non solo impiegano dati, ma anche i risultati di complesse inferenze da questi dati, il tradizionale strumento del consenso informato dovrebbe essere rivisto, dato che non sempre è chiaro in anticipo come i dati raccolti saranno utilizzati.

11. *L'esperimento è approvato da organi legittimati democraticamente.* Questa condizione si applica alle situazioni in cui il consenso non è fornito individualmente, ma deriva da un organo legittimato attraverso un processo democratico.

12. *I soggetti sperimentali possono influenzare la progettazione, l'esecuzione, il monitoraggio, la valutazione, l'adattamento e la fine dell'esperimento.* Questa condizione garantisce che la condizione 11 non possa essere oggetto di abuso da parte di un gruppo che decida di imporre sacrifici a un singolo individuo. I soggetti coinvolti dovrebbero poter sempre avere un ruolo nel definire gli esperimenti. Il rispetto di questa condizione è tutt'altro che semplice nell'ambito medico. Tuttavia, occorre trovare modi per poter tenere conto di questa condizione coinvolgendo per esempio le associazioni di pazienti nella definizione dei processi sperimentali.

13. *I soggetti sperimentali possono ritirarsi dall'esperimento.* Questa condizione è simile alla 12 e richiede che un soggetto possa ritirarsi in un qualsiasi momento da un esperimento in una qualsiasi delle sue fasi.

14. *I soggetti sperimentali più vulnerabili non dovrebbero essere soggetti di questi esperimenti oppure dovrebbero essere ulteriormente protetti o avvantaggiarsi in modo particolare dei risultati degli esperimenti.* Nel contesto generale degli esperimenti con tecnologie sperimentali questa condizione richiede che le persone più vulnerabili siano particolarmente protette, o addirittura non coinvolte, negli esperimenti che si occupano di valutare l'introduzione di nuove tecnologie nella società.

15. *Giusta distribuzione dei potenziali pericoli e benefici.* I pericoli e i benefici dovrebbero essere distribuiti in maniera equa tra i soggetti coinvolti, in modo da prevenire la situazione in cui i costi ricadano solo su alcuni e i benefici solo su altri.

16. *Reversibilità del danno o, se impossibile, compensazione del danno.* Questa condizione richiede che, se il danno è inevitabile, devono esserci dei meccanismi di compensazione. Una possibilità potrebbe essere quella di intendere la compensazione dal punto di vista monetario sulla scorta dei meccanismi usati nell'ambito assicurativo.

Queste sedici condizioni elaborate nel *framework* dei quattro principi della

bioetica mostrano l'importanza di un approccio incrementale e graduale nell'introduzione di tecnologie sperimentali nel loro contesto d'uso. Inoltre, mettono in luce come questa gradualità consenta di acquisire preziose informazioni alla cui luce può e deve essere rivista la progettazione delle stesse tecnologie sperimentali con l'obiettivo di evitare alcuni rischi. Nel prossimo paragrafo mi concentro quindi sul tema della progettazione (*design*), mostrando come accanto alle molte potenzialità che sussistono nel progettare una tecnologia che promuova effetti benefici, ci siano delle enormi criticità di cui occorre tenere conto.

La moralizzazione delle tecnologie: benefici e rischi

La responsabilità, intesa come la capacità di rispondere delle proprie azioni e dei loro effetti, è uno dei concetti chiave dell'etica della tecnologia. La responsabilità entra in gioco quando accade qualcosa di inaspettato e di negativo e il compito dell'etica è di cercare di stabilire di chi sia la responsabilità morale di quanto accaduto. Un esempio classico è il disastro del Challenger, la navicella spaziale progettata dalla NASA, avvenuto nel 1986 quando, a causa di alcuni problemi dei razzi che dovevano portare in orbita il Challenger, la navicella spaziale esplose dopo circa un minuto dal lancio. Questo caso viene discusso perché i problemi che causarono il disastro erano in parte noti e si può dire che non furono presi in seria considerazione né durante la fase di progettazione né durante quella di lancio. Dopo il disastro furono fatti degli sforzi ingenti per capire di chi fosse la responsabilità di quanto accaduto, ripercorrendo a ritroso la catena degli eventi che portarono all'evento negativo finale. L'idea di rivolgere l'attenzione a quanto è avvenuto prima di un evento negativo è tipica della cosiddetta *responsabilità passiva* (van de Poel & Royakkers 2011). Quando però entrano in gioco tecnologie potenti, e in parte imprevedibili, come le tecnologie sperimentali occorre riconsiderare l'idea della responsabilità passiva. Infatti il *design*, inteso come il progetto che porta alla realizzazione di una nuova tecnologia, può giocare un ruolo determinante nella gestione di alcuni problemi etici della tecnologia. Questo tipo di responsabilità è chiamata *responsabilità attiva*. La responsabilità attiva consiste nel progettare le tecnologie in modo tale che non solo sia possibile evitare effetti negativi, ma che le tecnologie stesse, per il modo in cui sono fatte, siano anche in grado di promuovere effetti positivi. Un esempio assai noto di visione attiva della responsabilità è il cosiddetto *Value-sensitive design* che, mediante un approccio multidisciplinare che opera fin dall'inizio delle fasi di progettazione di una tecnologia, considera i valori morali come requisiti della progettazione (Van den Hoven 2007). Per esempio, quando si parla oggi di *privacy by design*, ci si riferisce proprio all'idea che il valore della privacy sia incorporato nella progettazione, e che quindi un sistema informatico sia progettato in modo tale da garantirla a priori.

L'idea di incorporare valori morali in una tecnologia in fase di progettazione

prende anche il nome di *moralizzazione della tecnologia* (Verbeek 2011). Secondo questa idea, la moralità non riguarda solo gli esseri umani ma anche come questi interagiscono con le tecnologie che li circondano attraverso un processo di mediazione tecnologica che influenza sia la percezione sia l'azione. Per esempio, i tornelli per accedere alla metropolitana incorporano nella loro stessa struttura il fatto che si debba pagare un biglietto per utilizzare la metropolitana. In altre parole, sono gli esseri umani ma anche l'ambiente circostante – tecnologie incluse – a dare forma all'azione morale.

Il tema della moralizzazione delle tecnologie, e in particolare di come progettare una tecnologia in modo che la sua adozione promuova effetti positivi, è quanto mai attuale nel caso dell'IA. Per indicare questo indirizzo oggi vengono usate diverse etichette, come per esempio *Beneficial AI* oppure *AI for Good*. Certamente l'idea di poter progettare tecnologie che, oltre ad essere molto potenti, sono anche pensate per 'fare del bene' è stimolante e rassicurante, e sembra spazzare via definitivamente le perplessità in merito all'impatto dell'IA nelle nostre società. Non sarebbe 'benefico' progettare automobili che abbiano un misuratore del livello di alcol incorporato nel sistema di accensione, così da non poter essere guidate se il livello dell'alcol nel nostro sangue è sopra i limiti di legge? Forse si potrebbero evitare i molti decessi per guida in stato di ubriachezza che ancora oggi avvengono sulle nostre strade e che coinvolgono spesso le persone più giovani. E non sarebbe ancora più 'benefico' se questo rilevatore fosse obbligatorio per tutte le automobili e progettato in modo tale da riconoscere la tolleranza individuale all'alcol, attraverso un sofisticato sistema di IA in grado di valutare, in ogni momento e sulla base delle specifiche condizioni al contorno, la nostra capacità individuale di tolleranza all'alcol? In questo caso non dovremmo neppure più attenerci ai limiti di legge che sono uguali per tutti, ma potremmo avere un rilevatore personalizzato, visto che è noto che la tolleranza all'alcol varia in maniera determinante da persona a persona.

Questi temi divengono ancora più pressanti – e in certo senso ancora più critici – quando si parla dell'impiego di strumenti di IA in medicina. In questo contesto sembrano aprirsi possibilità concrete di risolvere tanti problemi per i quali educazione, buone pratiche e suggerimenti non riescono a essere cogenti. Anche pensando a casi semplici, come l'importanza di promuovere l'attività fisica o di alimentarsi in una certa maniera per preservare la nostra salute, appare immediatamente chiaro quanto ampio possa essere lo spazio di azione per tecnologie moralizzanti progettate per stimolare e supportare in maniera personalizzata determinati comportamenti e stili di vita. Non sarebbe 'benefico' progettare una applicazione che, sulla base delle caratteristiche individuali e delle attività quotidiane, consenta di migliorare l'aderenza a una terapia che magari è complessa e si prolunga nel tempo? Del resto sono ormai largamente diffusi i dispositivi commerciali indossabili (*wearables*) per il monitoraggio dei nostri stili di vita (come per esempio l'attività fisica o la qualità del sonno) che sulla base di questo monitoraggio ci 'consigliano' di muoverci di più o di andare a

riposare prima. Negli ultimi anni tali dispositivi indossabili sono notevolmente migliorati nelle loro prestazioni e sono in grado di misurare, con una precisione che in alcuni casi è paragonabile a quella dei dispositivi biomedici certificati, la pressione sanguigna o il livello di ossigeno nel sangue. Tali dispositivi non sono in sé tecnologie moralizzanti, ma hanno le potenzialità per diventare tali, soprattutto quando al monitoraggio, e sulla base di questo, si aggiungono le funzioni di diagnosi e intervento clinico, nonché di stimolo a determinati comportamenti. In tutti questi casi, gli algoritmi predittivi basati su tecniche di IA sono in grado di testare migliaia di parametri per scoprire i più statisticamente significativi dal punto di vista clinico. In altri termini, la tecnologia ‘apprende da sola’ quali parametri biologici, fenotipici e clinici sono correlati a una determinata condizione patologica (Blasimme 2021). Non sarebbe ancora più ‘benefico’ quindi progettare tecnologie per la salute che, sulla base dell’analisi di dati fenotipici digitali e dati genomici, nonché di quelli relativi a determinanti ambientali e socioeconomici della salute, siano in grado di personalizzare sempre più le pratiche diagnostiche e i trattamenti rispetto alle caratteristiche del singolo paziente? Un paziente che potrebbe essere stimolato a compiere certe azioni per prevenire o curare determinate condizioni e a evitarne altre con il supporto di tecnologie di IA che offrono previsioni personalizzate.

Grande attenzione è rivolta oggi ai numerosi problemi della progettazione e dell’uso delle tecnologie sperimentali in contesti sensibili come quello della medicina e della salute: dai problemi di privacy fino a quelli di giustizia. Problemi che sono ormai all’ordine del giorno nel dibattito attuale. Tuttavia, quando ci concentra sulle tecnologie sperimentali progettate per incorporare valori e per promuovere effetti benefici – le cosiddette tecnologie moralizzanti – sorgono ulteriori problemi. La moralizzazione della tecnologia solleva, infatti, domande importanti relative alla libertà umana e a come essa sia indissolubilmente legata alla nostra idea di dignità, al rischio che siano le tecnologie e non gli esseri umani a decidere, alla possibilità di rendere queste scelte frutto di processi democratici e non risultato delle decisioni poco trasparenti di un ristretto gruppo di persone o organizzazioni.

Torniamo per un momento all’esempio del sistema di accensione dell’auto che non consente di guidare al conducente che abbia un tasso di alcol sopra i livelli di legge. Accanto agli indubbi vantaggi che una soluzione di questo tipo potrebbe offrire se applicata su larga scala, non dobbiamo dimenticare le critiche di chi sostiene che ciò costituirebbe una delega della responsabilità alle tecnologie, senza che esse abbiano la flessibilità di gestire situazioni di emergenza. Come gestire quei casi in cui, anche a fronte di un livello di alcol troppo elevato, poter guidare un’auto costituisce letteralmente una questione di vita o di morte? È vero che le leggi limitano già in maniera significativa la nostra libertà, ma occorre riconoscere che i limiti imposti dalla legge (per esempio quella che indica che è vietato guidare in certe condizioni) sono ben diversi da quelli imposti da una tecnologia: infatti siamo sempre liberi di trasgredire una legge, ma è assai

difficile superare un limite imposto da una tecnologia.

Un altro problema riguarda la possibilità di delegare progressivamente la gestione dei nostri comportamenti e delle nostre scelte a tecnologie che, se certamente molto affidabili sotto diversi punti di vista, potrebbero farci correre il rischio di diventare sempre meno capaci di decidere e scegliere. La moralizzazione delle tecnologie può rischiare di diventare un modo per delegare la responsabilità per cui, invece di educare le persone a comportamenti più corretti nell'ambito della salute, queste attività sono 'appaltate' a tecnologie potenti che offrono soluzioni personalizzate.

Queste criticità devono essere prese sul serio e sono in parte già entrate nella riflessione accademica e nel dibattito pubblico. Ciò che ancora manca è l'attenzione a un altro tema, altrettanto serio e su cui le nostre società dovranno confrontarsi a breve: è possibile moralizzare le tecnologie in maniera democratica? In altri termini chi decide e in che modo quali sono i valori da incorporare in una determinata tecnologia? Come funziona questo processo?

Sulla base di quali criteri vengono fatte certe scelte invece che altre? I valori e le scelte, prima di essere incorporati in una tecnologia, sono condivisi e pubblicamente discussi?

Sono domande aperte, a cui questo capitolo non intende certo fornire una risposta semplificata. Per costruire una risposta plausibile, tuttavia, abbiamo bisogno di un radicale cambio di paradigma, alla costruzione del quale siamo tutti chiamati a contribuire. In primo luogo i progettisti delle tecnologie che devono essere consapevoli di questi problemi e di come, per citare uno dei fondatori della filosofia della tecnologia, le tecnologie hanno una politica (Winner 1980). Ma anche gli utenti di queste tecnologie che devono essere consapevoli della posta in gioco. E soprattutto tutti coloro che sono e saranno chiamati a utilizzarle in contesti altamente complessi e in cui l'impiego delle tecnologie sperimentali ha degli effetti potenzialmente dirompenti, come la medicina. Tutto ciò richiede molti cambiamenti, uno dei quali è la formazione dei futuri professionisti che deve tenere conto di come includere una consapevolezza critica su questi temi.

Conclusioni

La confluenza fra la dimensione umana e quella tecnologica che caratterizza oggi le nostre società in generale si rivela in tutta la sua importanza nei contesti come la medicina e la salute, che tradizionalmente hanno le persone al centro della loro pratica e riflessione. Proprio per il ruolo sempre più centrale che le tecnologie, e quelle di IA in particolare, hanno nelle professioni mediche, la riflessione critica sulla dimensione tecnologica è necessaria. I futuri professionisti devono essere capaci non solo di impiegare al meglio le diverse tecnologie, ma anche di tenere conto del loro impatto e della intrinseca dimensione socio-tecnica che le caratterizza. Ciò significa riconoscere che le tecnologie danno forma

alla nostra società che, a sua volta, dà forma alle tecnologie (Johnson 2009). Perché questo avvenga deve essere riconosciuta una sempre maggiore attenzione all'interdisciplinarietà nella formazione dei professionisti. Interdisciplinarietà naturalmente non deve significare essere esperti di molti domini, quanto piuttosto saper integrare nella propria professionalità alcune prospettive diverse che la possano arricchire. I medici del futuro non devono diventare esperti *anche* di bioetica e di etica della tecnologia; piuttosto devono essere consapevoli dei problemi che la progettazione e l'impiego di tecnologie sempre più potenti comporta nella loro professione. E per farlo devono essere in grado di dialogare con gli esperti di bioetica e di etica della tecnologia.

In questo capitolo ho mostrato come la bioetica, lungi dall'essere superata nella formazione dei futuri professionisti della salute, assume oggi una dimensione ancora più centrale. All'insegnamento della bioetica tradizionale deve essere affiancata però anche l'etica della tecnologia, affinché gli strumenti critici della dimensione tecnologica, con cui i futuri professionisti si troveranno ad avere sempre di più a che fare, siano oggetto di riflessione. Ciò non significa che il semplice fatto di insegnare questi temi sia sufficiente a sviluppare una riflessione critica su di essi. Certamente il modo di insegnare gioca un ruolo fondamentale, anche se purtroppo non ci sono risposte definitive che garantiscano che questa capacità critica sia stata acquisita. Un aspetto importante è senza dubbio mostrare la complessità di questi temi per evidenziare come la riflessione etica in generale sia un processo che non garantisce il risultato. L'etica non è un manuale di istruzioni e neppure una serie di regole su come dovremmo comportarci nelle nostre vite. L'etica è un processo complesso che richiede ogni volta informazioni contestuali, dipende dalle nostre attitudini morali ed è profondamente imbevuta della storia individuale e sociale di ciascuno di noi. Occorre in primo luogo riconoscere questa complessità per affrontare le sfide future che richiedono un cambio di paradigma a partire certamente dalla formazione delle future generazioni fino ad arrivare alla dimensione politica, intesa nella sua accezione più ampia (Crawford 2021).

Take-home messages

- La bioetica dovrebbe essere integrata con l'etica della tecnologia per includere nella sua riflessione la dimensione tecnologica che caratterizza oggi la medicina e i diversi ambiti della salute. Questa integrazione deve cominciare dalla formazione dei futuri professionisti che devono essere accompagnati ad affrontare in maniera critica i temi che la dimensione tecnologica fa emergere in questi contesti.
- In questo capitolo si discute di due aspetti che emergono, in particolare, con l'avvento delle tecnologie di intelligenza artificiale (IA): la necessità di un approccio incrementale nell'adozione di tecnologie di IA e le sfide connesse alla loro moralizzazione, ossia l'incorporare valori positivi nel loro *design*. Sebbene questi

aspetti non siano specifici della medicina, entrambi possono avere un impatto decisivo sulla pratica medica e sulle professionalità delle persone coinvolte.

Riferimenti

- Beauchamp, T. and Childress, J. (2012), *Principles of biomedical ethics* (7 ed.), Oxford University Press, New York.
- Blasimme, A. (2021), “Medicina Digitale e IA: profili etici”, in F. Fossa, V. Schiaffonati, G. Tamburrini (a cura di). *Automi e persone. Introduzione all’etica dell’intelligenza artificiale e della robotica*, Carocci editore, Roma, 55-69.
- Crawford, C. (2021), *Atlas of AI*, Yale University Press.
- Floridi, Luciano and Cowls, Josh (2019), “A Unified Framework of Five Principles for AI in Society”, Available at SSRN: <https://ssrn.com/abstract=3831321> or <http://dx.doi.org/10.2139/ssrn.383132>.
- Krohn, W., Weyer, J. (1994), “Society as a Laboratory”, *Science and Public Policy*, 21, 3, 173-183.
- Jobin, A., Ienca, M. and Vayena, E. (2019), “The global landscape of AI ethics guidelines”, *Nat Mach Intell* 1, 389–399.
- Johnson, D. (2009), *Computer Ethics*, Pearson.
- Nilsson, N. (2010), *The Quest for Artificial Intelligence: A History of Ideas and Achievements*, Cambridge University Press, Cambridge-New York.
- Van den Hoven, J. (2007) “ICT and value sensitive design”, *The information society: Innovation, legitimacy, ethics and democracy in honor of Professor Jacques Berleur SJ*, IFIPAICT, volume 233, 67-72.
- van de Poel, I. (2016), “An Ethical Framework for Evaluating Experimental Technology”, *Science and Engineering Ethics*, 22, 667-686.
- van de Poel, I. and Royakkers, L. (2011), *Ethics, Technology and Engineering*, Wiley-Blackwell.
- Verbeek, P.P. (2011), *Moralizing Technologies*, The University of Chicago Press.
- Vermaas, P., Kroes, P., van de Poel, I., Franssen, M. and Houkes, W. (2011), *A Philosophy of Technology: From Technical Artefacts to Sociotechnical Systems*, Morgan & Claypool Publishers.
- Winner, L. (1980), “Do artifacts have politics?”, *Daedalus*, 109, 121-136