

*ABSTRACT*

**Biorobotica, robotica e diritto penale**

**1. Genetica comportamentale, neuroscienze e diritto penale**

Il tema della genetica comportamentale e delle neuroscienze, nell'ambito della più ampia questione dei rapporti tra diritto e tecnologia, in ambito giuridico, pone sullo sfondo la questione della concezione della corporeità e della antropologia sottesa alle concezioni giuridiche. Sul piano penalistico molte sono le ripercussioni dei recenti studi di genetica comportamentale, di biologia molecolare e le neuroscienze, studi che ci consentono di comprendere meglio il processo decisionale e deliberativo umano, individuando una base biologico-chimica che impone una ridefinizione della nozione di libero arbitrio. L'impatto di studi di filosofia della mente, biologici e neurologici si riverbera sulle categorie importanti del diritto penale e del processo penale; viene in rilievo la categoria della imputabilità (in cui capacità di intendere comprende: empatia, pensiero morale e ragionamento controfattuale, e la capacità di volere: la capacità di controllare l'impulso motorio), della capacità a stare in giudizio, della valutazione della prova dichiarativa, rileva nell'ambito dei reati in materia di abuso di sostanze stupefacenti, con finalità di prevenzione del reato, nella fase esecutiva della pena, laddove la concezione preventiva della pena esprime quella funzione inibitoria, di orientamento culturale e di veto che agisce sulle funzioni di controllo

**2. Biorobotica, interfacce cervello-macchina e potenziamento umano: filosofia precauzionale e euristica di avversione al rischio.**

Il campo della robotica e della biorobotica è per eccellenza interdisciplinare. L'aspetto più attuale di questa tematica della robotica riguarda la biorobotica, ovvero la combinazione o ibridazione uomo-macchina, mediante innesto di parti robotiche nel corpo umano (computer indossabili, impianti di computer nel corpo, arti artificiali) con finalità terapeutiche o di potenziamento fisico. Il dato sorprendente, che distingue l'ibridazione uomo-macchina dalle altre protesi non cibernetiche, risulta proprio nell'interazione tra sistema nervoso, impulsi cerebrali e animazione della protesi, al fine di restituire al soggetto la percezione del movimento. La biorobotica si avvale delle c.d. interfacce cervello-macchina: canali che offrono la possibilità di influenzare gli stati mentali/di coscienza di una persona e che permettono di trasmettere direttamente al cervello dei segnali elettrici esterni. In tal modo l'arto bionico è in grado di riconoscere la volontà del soggetto ed eseguire gli ordini motori del cervello in tempo reale. Le interfacce tra il cervello umano e macchina consentono di leggere e utilizzare i segnali neurali associati all'attività cognitiva per controllare un arto artificiale o la traiettoria di una piattaforma robotica mobile.

Sorgono importanti questioni etiche: sul piano bioetico e filosofico l'evoluzione tecnologica mette in discussione il concetto di "umano" e di "umanità", di identità personale e perfino, sul piano giuridico, il concetto di responsabilità. La reale presenza dei fenomeni ibridativi uomo-macchina (cyborg) deve farci riflettere anche sui problemi giuridici connessi al funzionamento o al disfunzionamento di queste nuove tecnologie ed al regime giuridico di responsabilità per danni o reati commessi dai soggetti ibridati, ai prossimi congiunti ed a soggetti terzi. Il profilo di responsabilità penale, qualora siano stati realizzati reati, solleva i seguenti interrogativi: è possibile

affermare se l'incidente sia stato causato da un problema di controllo degli arti artificiali? Quindi il funzionamento degli arti bionici risponde a leggi scientifiche? A chi attribuire la responsabilità qualora venga cagionata la morte di un terzo, se non è possibile accertare un vizio di costruzione o funzionamento dell'arto biologico? L' "azione" compiuta con un arto artificiale può essere considerata "cosciente e volontaria"? Più a monte, si collocano i soliti profili relativi ai limiti alla disponibilità del corpo e della integrità fisica, ove il tema si intreccia con la distinzione tra intervento terapeutico (e relativo statuto epistemologico: consenso libero consapevole e informato) e intervento migliorativo o di potenziamento a scopo non terapeutico. Il punto è che è assai difficile concettualmente tracciare una distinzione tra terapia e potenziamento, e si potrebbe perfino sostenere che essa manca di significato pratico. In gioco è la stessa nozione di salute che se intesa in senso soggettivo, così come recepita dalla Organizzazione Mondiale della Sanità, ovvero come sinonimo di completo (ed utopico) stato di benessere, comprensivo non solo di una dimensione fisica ma anche di quella psichica che va al di là di quanto necessario per ristabilire uno stato di alterazione, sembrerebbe annullare la differenziazione concettuale tra terapia e potenziamento puro non terapeutico.

La questione della differenza morale e giuridica tra intervento a scopo terapeutico e intervento a scopo migliorativo interferisce con il tema di fondo: quali gli effetti delle tecnologie di potenziamento, benefici o dannosi? Tali effetti (benefici o dannosi) sono isolati all'individuo o si estendono all'ambiente e alle generazioni future? Il tema del potenziamento apre una nuova frontiera del dibattito bioetico, destando una serie di preoccupazioni etiche dovute alla situazione di doppia incertezza epistemica che lo connota. Il primo livello di incertezza epistemica è di tipo scientifico. Infatti, da un lato esse costituiscono delle tecniche relativamente nuove, alcune futuribili, altre allo stadio di prima sperimentazione, pertanto non vi è certezza o meglio, non vi è alcuna prova statistico scientifica sui loro possibili utilizzi in termini di efficacia, sicurezza, sia nel breve che nel lungo periodo. Il secondo livello di incertezza concerne la valutazione sul piano morale ed etico degli effetti sull'individuo, sulle generazioni future e sull'ambiente. Si pone il problema del come valutare, sotto il profilo etico, l'affinarsi di tecniche di potenziamento in termini di "autenticità" (è più autentica la vita di un individuo che ha migliorato le proprie potenzialità o quella di un individuo che vive nelle foreste?), di dignità (ma quale concetto di dignità umana assumere, come obiettivo di realizzazione o come naturalità?).

Il dibattito attuale vede fronteggiare sostenitori che articolano teorie e argomentazioni a difesa del potenziamento (teoria libertaria, teoria utilitarista, dei c.d. "tecnofili" Nicholas Agar, Allen Buchanan, Nick Bostrom, John Harris, Julian Savulescu) e d'altro lato detrattori o c.d. "bioconservatori" (Francis Fukuyama, Jurgen Habermas, Leon Kass, Michael Sandel) che analizzano le possibili minacce aperte all'uomo e alle generazioni future. In questa prospettiva il problema del potenziamento viene affrontato nell'ambito delle teorie della giustizia, con specifico riferimento al problema della disuguaglianza (potenziati/non potenziati), delle possibili ripercussioni dell'accesso al potenziamento sulla non accettazione della disabilità (dato il divario sempre crescente tra dis-abili, abili, super-abili o potenziati).

Si pone quindi un problema di individuare un modello di regolazione giuridica in situazioni di incertezza scientifica. Viene quindi in rilievo l'ipotesi del divieto ultraprudenziale di comportamenti i cui effetti sono benefici. Se gli effetti ipotizzabili delle tecnologie di potenziamento bionico e biorobotico sono dannosi (per l'individuo, per le generazioni future o per l'ambiente) o in parte benefici (per l'individuo) ma dannosi non le generazioni future e per l'ambiente, l'eventuale penalizzazione delle tecnologie di potenziamento umano di biorobotica trova fondamento nel paradigma classico di legittimazione del divieto di cagionare danno a terzi o a se stessi. Tuttavia, se come si diceva, vige una totale incertezza scientifica sulla dannosità di tali interventi, e di essi si accerta solo la beneficialità sotto il profilo fisico-psichico, sia pure sul solo singolo individuo su cui vengono praticate, si pone il problema sotto il profilo etico se il biopotenziamento non debba costituire un dovere morale e sociale (se poi gli effetti benefici per

tutti, sia a livello personale che generale). La ipotetica e futuribile penalizzazione di queste tecnologie in questo caso rifletterebbe una inversione del paradigma paternalistico forte e debole: si vieta ciò che è bene per l'uomo sul presupposto che il potenziamento è espressione di motivi futili. Si va oltre il divieto di cagionare danno a se stessi o ad altri: si intravede una norma di divieto ultraprudenziale che va oltre al modello di giustificazione del paternalismo forte e che recepisce l'euristica di precauzione forte.

### **3. La robotica, i droni e le Intelligenze Artificiali.**

La robotica comprende lo studio delle Intelligenze Artificiali, ossia la costruzione di macchine capaci di sentire, adattarsi all'ambiente, di imparare, dotate di capacità evolutiva e persino di "capaci di empatia". Queste macchine presentano le seguenti caratteristiche: sono interattive e reattive all'ambiente, agiscono in modo autonomo e imprevedibile e non determinato, flessibile e influenzabile; sono dotate di un sistema di apprendimento automatico e quindi dotate di autonomia, interattività, adattabilità, e in grado di migliorare le loro performance. Le organizzazioni complesse si avvalgono di sistemi informatici e tecnologici intelligenti in grado di prendere delle decisioni autonome (decisioni che non prevedono alcuna esplicita autorizzazione da parte di un essere umano) e che operano come supporto a manager soprattutto nella gestione di infrastrutture tecnologiche ad alto rischio. Le organizzazioni artificiali sono strutture mutate da quelle umane, all'interno delle quali ogni agente intelligente, denominato personoide occupa un ruolo ben preciso (che comprende accesso a informazioni, doveri responsabilità) e producono diversi processi decisionali.

Gli ambiti di applicazione delle Intelligenze Artificiali sono i più disparati: si va dagli agenti artificiali e reti neurali artificiali in ambito economico (i robot agiscono come agenti economici intelligenti privi di emozioni, e quindi realizzano le condizioni di equilibrio competitivo nella domanda e nell'offerta, ovvero il modello di efficienza distributiva perfetta), creazione di robot di servizio e assistenziali (badanti robot). In medicina, con l'avvento della chirurgia mini-invasiva, i robot sono utilizzati nella c.d. telechirurgia. Applicazioni assai vaste sono in ambito militare, con l'uso di droni come armi e di robot soldati. In proposito si richiama il monito di molti scienziati a vietare o a usare cautele nell'uso di robot autonomi letali, come sistemi intelligenti in ambito militare in cui la decisione è autonoma, le cui conseguenze possono essere assai devastanti. Questi robot armi possono innescare armi letali senza un intervento umano nel processo decisionale. La Corte penale internazionale ha dichiarato l'illegittimità dell'uso di armi robot e droni.

Il comportamento dei robot può essere imprevedibile, in quanto le Intelligenze Artificiali hanno capacità di apprendimento automatico, di adattamento all'ambiente, capacità decisionale e creatività rispetto una rosa di scelte. I metodi di apprendimento automatico non sempre consentono di appurare se il robot abbia veramente imparato (o anche solo approssimato in modo soddisfacente) ciò che vogliamo insegnargli. Si pongono quindi problemi di valutazione dei risultati ottenuti nei processi di apprendimento automatico da parte di robot. Inoltre, l'interazione tra un sistema aperto di apprendimento condizionato proprio delle Intelligenze artificiali e il suo ambiente (comunicazione asincrona) è spesso soggetta a vincoli temporali che non consentono di assicurare la presenza di un essere umano nei cicli di controllo. Quindi certamente non sono completamente controllabili dall'uomo.

Gli studi dei metodi formali di verifica del software nascono dall'esigenza di verificare se ogni esecuzione di un determinato programma per un calcolatore soddisfa alcuni requisiti fondamentali. Più recentemente, queste metodologie sono state estese al problema di specificare e verificare le proprietà di sistemi, generalmente detti "ibridi", che comprendono varie tipologie di sistemi robotici. Alcuni risultati limitativi che sono stati ottenuti a proposito dei sistemi ibridi indicano che,

in generale, non è possibile verificare con queste metodologie se un sistema robotico soddisfi o non soddisfi determinati vincoli spaziali o temporali nell'esecuzione di un dato compito.

Quali sono le implicazioni pratiche normative delle riflessioni epistemologiche sulla nostra limitata capacità di prevedere il comportamento delle macchine? Il produttore o il programmatore di un sistema intelligente che apprende dall'esperienza è in grado di prevederne precisamente il comportamento, anche nelle condizioni normali d'uso, ed è quindi responsabile? Si profilano due ordini di problemi. Il primo teorico filosofico, ossia la questione dello statuto ontologico di macchine particolarmente evolute. Sono semplici oggetti, oppure travalicano la soglia dei requisiti minimi per il riconoscimento di un livello di soggettività? Il secondo è pratico-funzionale-giuridico, ossia, la questione se sia opportuno e utile ai fini giuridici conferire a simili artefatti un livello di soggettività. Il problema quindi è stabilire se la tecnologia moderna ha creato una nuova tipologia di soggetto, il soggetto non-umano, l'agente non-umano. Queste Intelligenze artificiali, in quanto dotate di autonomia, agenti imprevedibili sono soggetti realmente capaci di manifestare una "propria" intenzionalità e quindi un proprio agire? Sono meri strumenti, o soggetti, anche mediati rispetto ad un altro autore?

Secondo la corrente di pensiero detta Intelligenza Artificiale Forte i calcolatori sono macchine capaci di esprimere autentico pensiero e in grado di produrre processi intellettuali identici a quelli umani. Non vi sarebbe alcuna differenza ontologico-qualitativa tra cervello umano e cervello elettronico e tra intelligenza umana e intelligenza artificiale, la sola differenza sarebbe la sede o il supporto fisico, la testa umana fatta di carne, ossa e altri materiali biologici e la struttura di un calcolatore fatto di metallo e energia. Non conta la struttura ma la funzione. Esistono intelligenze in sé indipendentemente dalla sede fisica in cui risiedono (tesi funzionalista). L'altra tesi, quella contraria detta della Intelligenza Artificiale Debole, ossia della non autenticità del pensiero meccanico e della diversità ontologica tra intelligenza artificiale e intelligenza naturale sostiene che le macchine simulano e riproducono soltanto i processi intellettuali umani e ne rappresentano delle copie; conta la struttura e non la funzione ossia la sede dove si svolge l'attività non come si svolge. Da queste due visioni derivano due opposte concezioni a proposito dello statuto ontologico delle macchine.

A) I profili di responsabilità penale: l'ipotesi fantastica del robot come agente non-umano responsabile penalmente, nuovo soggetto di diritto penale. Questa ipotesi solleva problemi di definizione della capacità soggettiva, della nozione di azione, della nozione di colpevolezza: i robot sono rimproverabili e motivabili? Esprimono una loro volontà, possono essere "autori" di reati? Il tema presenta un'analogia con quello della responsabilità penale-amministrativa degli enti collettivi e cozza con una concezione antropomorfa del diritto penale; gli enti collettivi non hanno corpo né anima, ma sono soggetti alla legge penale; i robot hanno un "corpo", una materia su cui far ricadere la sanzione penale (ad esempio, la disattivazione della macchina o la sua distruzione), ma non hanno un'anima, pur essendo dotati di autonoma capacità decisionale ed essendo in grado di interagire con l'ambiente. Ma l'agire del robot è un "vero" agire rilevante penalmente, cioè si può definire cosciente e volontario?

B) Ipotesi della responsabilità penale della persona fisica: se i robot sono mezzi e non agenti -non umani. Se si ritiene che i robot non siano agenti in senso stretto, si pone quindi il problema della responsabilità esclusivamente umana del programmatore o utilizzatore. A che titolo può essere imputata la responsabilità all'uomo, se si assume che il sistema informatico intelligente è un decision maker? I robot intelligenti agiscono in modo non programmato e imprevedibile. Questa imprevedibilità pone problemi di attribuzione della responsabilità penale a carico dei programmatori, dei costruttori e degli utilizzatori. Problema inverso di imputazione agente umano-ente collettivo: l'agente umano agisce, l'ente ne risponde. Qui l'agente non-umano agisce, la persona fisica ne risponde.

La responsabilità penale del programmatore o utilizzatore: che modello di responsabilità penale a carico dell'agente umano? Responsabilità diretta o indiretta, colposa o dolosa? Si profilano due soluzioni: A) a titolo di dolo, perché l'azione "volontaria" dell'agente non-umano incardina responsabilità sull'uomo e perché l'azione del robot si indentifica e rappresenta una longa manus della persona umana (teoria dell'immedesimazione organica: il robot). B) a titolo di colpa, individuata nel difetto di programmazione, costruzione, scelta, utilizzo, manutenzione funzionamento di robot intelligenti.

La colpa del programmatore: occorre elaborare un nuovo concetto normativo di colpevolezza che consente di elaborare un concetto di "colpa da programmazione", ove si può ipotizzare di prospettare una causa di esclusione della colpevolezza consistente nella predisposizione di misure di sicurezza che prevengano la realizzazione di reati da parte di robot intelligenti (in pratica che mettano in atto le tre leggi della robotica di Asinov).

Inoltre, si individueranno le regole cautelari relative alla programmazione, scelta, uso, controllo e manutenzione di robot. Limiti alla responsabilità del programmatore, costruttore, utilizzatore sono da individuare nel caso di uso improprio del robot, o nel caso fortuito o nella forza maggiore.

Quando l'agente artificiale è evoluto e gode di margine di creatività e decisionalità (è un decision maker) e non vi è un visibile difetto nel controllo, uso, manutenzione dei robot, a che titolo sarà l'imputazione soggettiva del reato a carico della persona fisica? Nella letteratura anglosassone si parla di prevedibilità e di logico sviluppo prevedibile, similmente al 116 c.p. in tema di responsabilità del concorrente per l'evento diverso più grave non voluto, purchè prevedibile, e anche se non sono perfettamente noti tutti gli anelli della catena causale di produzione dell'evento. La differenza è che il programmatore o l'utilizzatore non concorrono nella realizzazione di un reato di base e quindi la regola del versari in re illicita non vale, ma creano una condizione di rischio lecito tollerato. La prevedibilità come componente della colpa, sulla base di leggi scientifiche, anche se non corroborate da studi e non consolidate, o nella totale ignoranza di leggi scientifiche di spiegazione causale. In definitiva, è colposo il comportamento de programmatore che non preveda l'imprevedibilità del robot intelligente! Anche la prevedibilità astratta e generica di futuri danni ancora non ben identificati incardina la responsabilità per colpa, anche quando non sono noti tutti gli anelli del processo causale: rischio di identificare la colpa con la precauzione e il prudenzialismo.