Lorenzo Picotti
(Professore ordinario di Diritto penale e Diritto penale dell'Informatica – Università di Verona)

Profili penali dell'utilizzo di sistemi di intelligenza artificiale e robotici

Sommario:

1. ***Introduzione: dall'anno scorso….***
2. ***Il Colloquio Internazionale dell'AIDP e le raccomandazioni su intelligenza artificiale e responsabilità penali***
3. ***Per un adeguamento dei modelli d'imputazione all'uomo od all'ente che "sta dietro" i sistemi di I.A.***

1. Introduzione: dall'anno scorso….

Nel ringraziare Giuseppe Cassano e Francesco Di Ciommo per aver organizzato la nuova edizione degli "Stati generali del Diritto di Internet" e per il nuovo invito, non posso non muovere dalle considerazioni espresse lo scorso anno nella prima edizione.

Tracciavo un rapido excursus storico del diritto penale dell'informatica, dai computer crime ai cybercrime, sottolineando l'impatto del cyberspace (quale infosfera connotata da un'iperconnettività permanente e da una dimensione davvero imponente di acquisizione e gestione quotidiana di dati, provenienti da tutti gli utenti: i dati quali nuovo "petrolio" della moderna economia e società, come è stato detto) per sottolineare le novità degli ulteriori crimini informatici o, meglio, cibernetici nascenti di fronte alle nuove manifestazioni della criminalità.

Ed affacciavo l'insufficienza - di fronte all'emergere dell'I.A. - della definizione di "sistema informatico" fornito dalla pur basilare Convenzione Cybercrime del 2002, di cui all'art. 1, lettera a), che si limita alla nozione di "elaborazione automatica di dati in esecuzione di un programma"

Infatti, dalla mera "automazione", consistente nell'esecuzione automatizzata del trattamento di dati secondo un programma, dipendente dall'input dei dati e dei programmi posti dall'uomo, si profila oggi la crescente "autonomia" dei nuovi sistemi di I.A., a partire dall'autonomia dei processi di apprendimento, delle logiche di conoscenza, degli approcci statistici e degli incroci di queste ed altre fonti di esperienza, attivate e sviluppate dai sistemi stessi, nei vari settori in cui possono operare e sempre più operano ed opereranno, per cui procedono alla stessa ricerca ed acquisizione diretta di dati, non solo nel cyberspace, ma anche nel mondo esterno, tramite sempre più sofisticati sensori di ogni genere, ottici, sonori, termici, tattili e quant'altro.

"Autonomia" che si esprime parallelamente nello sviluppo degli algoritmi di nuova generazione, in specie i c.d. algoritmi adattivi, che possono evolversi e modificarsi senza intervento dei programmatori, sulla base delle esperienze ed informazioni accumulate. Tanto che si prospetta oggi una c.d. I.A. "forte" (strong) che sempre più si distingue da quella c.d. "debole" per il crescente livello di autonomia che raggiunge.

A fronte di questi sviluppi, vi è la basilare "fame di dati" dell'intero sistema globale, che necessita di un'accumulazione (non impropriamente definibile come "originaria") di dati ed informazioni, da acquisire con ogni modalità ed in ogni occasione, non sempre legalmente, ma spesso con tecniche e mezzi anche "predatori", necessaria allo sviluppo ed espansione della società, dell'economia, dell'amministrazione

digitali, come l'accumulazione originaria di capitali (ampiamente predatoria ed illegale) fu necessaria per la nascita ed evoluzione del sistema capitalistico.

E riproponevo – come si ripropone - l'annosa questione se ci si dovesse rassegnare al dominio del "codice tecnico" (*code is law*), ovvero, oggi, degli algoritmi di nuova generazione, in sintesi: al dominio dei sistemi di I.A. sul diritto stesso, destinato a perdere spazio e possibilità di regolazione di fronte a fenomeni strutturalmente troppo rapidi ed estesi per essere riconducibili al suo controllo; oppure se la regola giuridica avesse ed abbia necessità e possibilità di prevalere.

Al riguardo, segnalavo il ruolo di enforcement che gli stessi sistemi di I.A. possono al contempo svolgere e in parte già svolgono nel preservare i diritti ed interessi giuridici meritevoli di tutela, sia nella fase di raccolta e gestione dei dati, sia nella fase della loro diffusione o messa a disposizione: ad es. tramite un adeguato intervento di prevenzione delle attività e dei comportamenti illeciti che gli stessi più importanti ISP o "signori del web" attivano mediante un'applicazione automatizzata di filtri e regole interne, o policy per gli utenti e i partner commerciali, capaci di riconoscere, bloccare ed/od eliminare automaticamente contenuti illeciti o non graditi, attivando anche organi di controllo e di valutazione interna.

Ma un altrettanto rilevante apporto possono già dare i sistemi I.A. alla giustizia penale, a partire dalle indagini sui reati e dagli interventi di intelligence, fino alla gestione dei procedimenti e dello stesso processo penale, pur rifuggendo dall'introduzione di sistemi di giustizia o polizia predittive o di decisioni totalmente affidate a tali sistemi autonomi (come indica anche il molto citato art. 22 GDPR, e già la giurisprudenza del nostro Consiglio di Stato).

La risposta era ed è che il diritto, anche penale, non può abdicare dalla sua funzione regolatrice e di garanzia dei diritti fondamentali e degli interessi o beni giuridici meritevoli di tutela: funzione che richiede il ricorso alla sanzione penale quale *ratio extrema*, ma indispensabile, come dimostra la sua stessa storia, perché come già insegnava il grande utilitarista inglese Jeremy Bentham, alle origini del diritto penale moderno, di fronte alle offese più gravi non basta contrastare il c.d. "male immediato" materiale che riguarda la singola vittima lesa nei suoi personali interessi o diritti, e che può conseguire ristoro risarcitorio anche dal diritto civile. Vi sono infatti offese più gravi, che producono un c.d. "male mediato" (o di second'ordine, per Bentham), che deriva dal primo, ma si espande sull'intera società, quale allarme e timore di ripetizione, minando la sicurezza di tutti, fino a creare il pericolo di una disgregazione dei rapporti sociali e della stessa società, un ritorno all'*homo homini lupus* di Hobbes, che invocava per questo un Leviatano.

Il tema era ed è allora ancora oggi di definire i presupposti ed i limiti della responsabilità penale di fronte alle nuove manifestazioni criminose, caratterizzate ormai dall'impiego od anche semplicemente dall'affidamento ai sistemi di I.A., in sempre più numerose ed importanti attività umane, non solo illecite, ma anche e soprattutto lecite.

E non può non rinviarsi al compito dei giuristi, e dei penalisti in specie, per definire nuovi modelli ed adeguate categorie di imputazione dei fatti offensivi, che possano integrare reati già esistenti o da introdurre specificamente nell'ordinamento giuridico, tramite cui sia riconosciuto il ruolo crescente che i sistemi d'I.A. e robotici hanno assunto e sono sempre più destinati ad assumere.

2. Il Colloquio Internazionale dell'AIDP e le raccomandazioni su intelligenza artificiale e responsabilità penali

2.1 Questo compito è stato assunto ed intrapreso, non certo del tutto adempiuto, nell'anno trascorso.

In particolare, richiamo i lavori della I sezione del Congresso AIDP (Association Internationale de Droit Pénal) sul tema di diritto penale generale "*Traditional law categories and AI: crisis or palingenesis?*" sfociati nel colloquio internazionale svoltosi a Siracusa presso il prestigioso Syracuse Institute il 15 e 16 settembre scorso, dopo che sono stati raccolti, sulla base di un questionario, 22 rapporti nazionali e 3 rapporti speciali, sulla cui base è stato possibile redigere il mio rapporto generale, con l'insostituibile supporto di Beatrice Panattoni, che ringrazio (anche di essere oggi presente), e giungere a formulare una bozza di risoluzione, contenente le raccomandazioni da sottoporre infine all'approvazione del XXI Congresso Internazionale dell'AIDP che si terrà a Parigi nel 2024, interamente dedicato, nelle sue quattro sezioni, al tema "*Artificial Intelligence and Criminal Justice*".

Nel corso del Colloquio, cui hanno partecipato, oltre al Presidente dell'AIDP ed ai rapporteurs speciali, numerosi esponenti dei gruppi nazionali, sono stati apportati numerosi emendamenti, integrazioni, aggiunte alla bozza di risoluzione, che infine è stata approvata all'unanimità e verrà pubblicata nel prossimo numero (primo del 2023) della *Revue Internationale de Droit Pénal*, unitamente al rapporto generale, ai rapporti speciali ed a una selezione di rapporti nazionali.

Da questi è emerso un quadro variegato di (scarse) previsioni normative (in specie in Germania e Francia, in materia di veicoli a guida autonoma), singole pronunce giurisprudenziali e, per ora, soprattutto articolati contributi dottrinali, compresi alcuni importanti progetti di riforma in materia, che vanno dalla Cina all'Ucraina.

2.2 Dopo il vivace e costruttivo dibattito svoltosi nel citato Colloquio, le raccomandazioni approvate con la risoluzione finale sottolineano, in primo luogo, la **necessità della protezione penale** di fronte alle offese più gravi, realizzate da o tramite o a danno di sistemi di I.A. e robotici, non essendo accettabile che restino impunite, mentre sarebbero sanzionate penalmente se commesse da persone fisiche.

Il secondo punto caratterizzante della risoluzione approvata è che occorre una specifica regolamentazione, che sappia coniugare le menzionate ed emergenti esigenze di tutela penale, con le garanzie fondamentali del diritto penale, rappresentate in primis dai principi di legalità e di personalità della responsabilità penale, vale a dire di colpevolezza.

Al riguardo è stata respinta l'ipotesi di una responsabilità penale degli stessi sistemi di I.A. come tali (insuscettibili di effettiva capacità di intendere e di volere, ed insensibili all'effetto "rieducativo" della sanzione penale, che fosse minacciata ed applicata direttamente ad essi, pur con i debiti adattamenti) e si è raccomandato di prevedere invece specifiche discipline per i singoli settori di attività in cui operano, muovendo dalla fissazione normativa di regole di sicurezza, standard di produzione, cautele operative e quant'altro possa garantire, entro livelli accettabili, il controllo e la limitazione dei rischi specifici connessi al loro utilizzo.

In effetti ritengo che il termine I.A. sia una suggestiva metafora, che non identifica un fatto, un soggetto o un "ente" unitario, che si possa regolare in termini generali.

I rapporti nazionali concordano sul punto, sottolineando altresì che a livello normativo e concettuale i diversi sistemi di I.A. non possono – allo stato, e salvi futuri sviluppi della c.d. I.A. "forte" – essere direttamente destinatari di precetti (e sanzioni) penali, nonostante il livello di autonomia raggiunto, dovendosi piuttosto risalire all'uomo od all'ente che vi sta dietro e che vi si affida.

2.3. In particolare, ai fini penali va operata una **distinzione basilare**, espressa nelle raccomandazioni, fra **l'utilizzo doloso,** in attività criminose od illecite, dei predetti sistemi, e **l'utilizzo** ben più diffuso **in attività lecite**.

Nel primo caso ne è stata sottolineata la maggior pericolosità e capacità offensiva, in particolare in ipotesi di reati realizzati con sistemi robotici (si pensi a droni che seguano e colpiscano il bersaglio o i bersagli da

individuare e prescegliere, o più in generale all'uso di armi letali autonome, vietate dal diritto internazionale penale; o di algoritmi finalizzati alla realizzazione sistematica di frodi, anche tributarie, abusi di mercato, riciclaggio, e quant'altro). A tali ipotesi possono applicarsi le regole dell'*aberratio ictus* e dell'*aberratio delicti*, nonché prevedere, se del caso, specifici reati preparatori.

Nei casi più rilevanti del ricorso o dell'affidamento, sempre più diffuso, ai sistemi di I.A. in attività lecite, emergono invece i maggiori problemi che attengono ai presupposti e limiti di un'imputazione a titolo di colpa.

Al riguardo sono emersi approcci diversi, come ad esempio in Germania ed in Francia, con riferimento alla **disciplina anche penale delle self driving cars**, su cui sono intervenuti di recente i rispettivi legislatori.

In Germania si persegue la responsabilizzazione, comunque, del conducente (*Fahrer*), che deve mantenere sempre una posizione di controllo e garanzia sulla circolazione del veicolo: esso deve rispettare tutte le regole della circolazione ed il conducente deve essere sempre in grado di disattivare in tutto od in parte la modalità di guida autonoma (distinta in "elevata" e "piena" automazione) per assumere la diretta guida del veicolo (§§ 1 a ed 1 b StrassenVerkehrG – modificati con la legge del 7.5.2021). Ma anche il produttore deve, a monte, rispettare precisi standard di sicurezza, conformi a quelli stabiliti a livello europeo ed internazionale e normativamente definiti, per poter mettere in commercio veicoli a guida autonoma o automatizzata ad elevato livello, che devono previamente conseguire una specifica licenza dall'autorità pubblica di controllo, spettando poi al Ministero federale della circolazione operare una periodica valutazione da riferire al Governo ed al Bundestag.

In Francia la modifica del codice della strada ha invece introdotto l'interessante concetto di "delegazione della guida" da parte del conducente al veicolo che disponga di un sistema automatizzato di guida, che può essere di tre livelli: parziale, elevato o pieno, di cui mantiene la responsabilità nei limiti della "delega".

Un interessante caso emblematico è stato di recente deciso in Finlandia, con condanna del conducente che non si è reso conto di un errore del sistema di navigazione, che non aveva riconosciuto che l'auto si trovava in una rampa d'uscita dell'autostrada, in cui avrebbe dovuto moderare la velocità, per cui avvea tamponato l'auto che la precedeva causando lesioni ai loro passeggeri.

2.4. Occorre dunque procedere per settori considerando le specifiche regole in ciascuno esistenti o da introdurre o adattare.

Uno dei più interessanti, settori da considerare, accanto a quello della circolazione stradale e dei trasporti in genere, è rappresentato dal **ricorso a sistemi di I.A. e robotici nel campo medico**, dalla formulazione di diagnosi e terapia, all'esecuzione di interventi chirurgici, alla c.d. telemedicina.

Non si tratta di meri strumenti passivi, ma di sistemi che sono oggetto di un vero e proprio affidamento, se non anche di una vera delega, quantomeno parziale, di compiti ed attività specialistiche da parte di un operatore, al fine di ottenere prestazioni più sicure e precise, che muovano da dati più ampi ed aggiornati, in tempi immediati. Per cui vengono in rilievo, da un lato, le capacità "cognitive" e "decisionali" di questi sistemi, dall'altro i beni giuridici fondamentali che possono essere offesi da errori o malfunzionamento dei sistemi stessi, quali la vita, l'incolumità, l'autodeterminazione informata dei pazienti.

Diverso è il rapporto che con loro si instaura, rispetto a quello che si ha con chi è offeso da sinistri nella circolazione stradale, poiché vi è alla base un voluto od accettato affidamento degli stessi pazienti nel ricorso a tali sistemi, all'interno del rapporto di cura con gli operatori sanitari.

In comune è che si tratta di possibili reati colposi di evento (morte, lesioni personali), per cui ogni modalità tramite cui si realizza l'offesa può integrare la fattispecie penale incriminatrice.

2.4.1 Però si pone innanzitutto un problema di accertamento della **causalità**, fra condotta ed evento, che già pervade la casistica medica: rispetto alla patologia in atto, quale può esser il peso del contributo causale pur concorrente con altre condizioni, che però non escludano il nesso di causalità (ex art. 41 c.p.), portato dall'intervento del sistema di I.A:? se si considerasse quest'ultimo quale causa sopravvenuta "*da sola sufficiente a determinare l'evento*" – in quanto "scelta" operata autonomamente dal sistema - sarebbe interrotto il legame causale con la condotta (attiva od omissiva) dell'operatore sanitario che vi ha fatto affidamento. Ma così si creerebbero inaccettabili sfere di impunità, destinate ad espandersi, quanto più si estenda il ricorso a detti sistemi.

E fino a che punto si dovrebbe o potrebbe risalire nella catena causale, fino all'ideatore e/o produttore del sistema o del robot, per un suo malfunzionamento?

Problema acuito dalla possibile causalità c.d. normativa "per omissione", *ex* art. 40 capoverso, c.p., per il cui accertamento quale "equivalente" della causalità naturale (la patologia che ha in concreto causato la morte) si deve poter stabilire se sia configurabile e fin dove si spinga il "dovere giuridico di impedimento" dell'esito avverso in capo all'agente umano che sta "dietro" ai sistemi di I.A. cui si è affidato.

2.4.2 Il problema confina così con quello della **colpa**. Per imputare una responsabilità penale ai soggetti umani che "stanno dietro" i sistemi di I.A. occorre accertare anche la loro colpevolezza soggettiva, non essendo ammissibile (ex art. 27 Cost.) una mera responsabilità oggettiva, che può eventualmente fondare una responsabilità civile, contrattuale od extracontrattuale.

Ma per aversi colpa, occorre preliminarmente individuare le "regole cautelari" riconoscibili, il cui rispetto avrebbe consentito di evitare l'evento offensivo.

Potrebbero così emergere profili di colpa specifica, per la violazione, ad esempio, delle istruzioni fornite dal produttore o dal manutentore; oppure anche profili di colpa generica, per cui però non è facile individuare – data la novità della materia e l'imprevedibilità, per molti aspetti, delle scelte ed attività concrete del sistema di I.A. - quel modello dell' *homo eiusdem condicioni s et professionis* da cui si dovrebbero dedurre le regole non scritte di comportamento diligente, prudente e perito.

A proposito dell'imperizia soccorre solo in parte la disciplina dell'art. 590 *sexies* c.p. introdotto dalla c.d. legge Gelli Bianco, che esclude la responsabilità colposa degli operatori sanitari che rispettino le Linee Guida e le Buone pratiche clinico assistenziali, adeguate al caso concreto, quali riconosciute formalmente secondo la relativa disciplina.

Ma esse includono, od in che misura possono estendersi anche all'operato dei sistemi di I.A. e robotici? La valutazione di "adeguatezza" alla specificità del caso concreto abbraccia il ricorso ad essi, al controllo dei requisiti di sicurezza ed aggiornamento? In altri termini: la perizia medica abbraccia la conoscenza specialistica e la abilità nell'uso dei robot?

Certamente essi aumentano la precisione e l'accuratezza degli interventi sanitari, innalzando, nel contempo, lo standard di sicurezza e di precisione da osservare, rispetto a quello esigibile da un soggetto umano.

2.4.3 Vengono, infine, in rilievo le **"posizioni di garanzia"** del personale medico sanitario nei confronti dei pazienti, ai sensi dell'art. 40, cpv., c.p. (corrispondente al § 13 StGB), rispetto alla loro vita e salute, oltre che al loro diritto all'autodeterminazione in materia di trattamenti sanitari. Al riguardo può subentrare una responsabilità per omissione, rispetto agli obblighi giuridici incombenti sul personale sanitario, da estendere al controllo del corretto funzionamento dei sistemi di I.A. e robotici.

Mentre possono rilevare, rispetto al necessario consenso informato - che deve includere espressamente il ricorso a detti sistemi – deviazioni impreviste nel loro funzionamento od output, dipendente dal margine di autonomia ed operatività immediata dei sistemi predetti.

2.4.4 La complessità della catena di produzione, commercializzazione, applicazione e controllo dei sistemi in esame, fino al loro impiego concreto, suggerisce di ricorrere ad alcuni schemi collaudati di imputazione della responsabilità nel campo della disciplina dei prodotti, della salute e sicurezza sui luoghi di lavoro, della stessa **responsabilità "da reato" degli enti** ai sensi del d.lgs. 231/2001.

Infatti, emerge da queste discipline il rilievo giuridico, anche penale, degli obblighi di previa valutazione dei rischi connessi all'attività esercitata e delle conseguenti misure di contenimento e controllo da adottare, mediante adeguati "modelli di organizzazione", la cui violazione fonda la c.d. colpa di organizzazione.

È infatti attraverso una corretta strutturazione delle diverse competenze, da ripartire fra soggetti c.d. apicali e subordinati, che si può articolare un efficiente sistema di prevenzione dei reati e degli illeciti che possono essere commessi nell'esercizio delle in sé altrimenti lecite attività dell'ente.

Ed è interessante sottolineare che, in base al principio di "autonomia" della responsabilità dell'ente, rispetto a quella dell'autore del reato, di cui può mancare l'identificazione specifica o che potrebbe risultare non punibile (art. 8 d.lgs. 231/2001), l'ente potrebbe essere comunque chiamato a rispondere del reato ascrivibile, nell'ultimo anello della catena causale, al sistema di I.A., in forza dei criteri di imputazione stabiliti dalla stessa normativa, anche se non viene identificato o non è personalmente punibile un singolo soggetto persona fisica.

Il problema qui nasce però nel nostro ordinamento anche dalla limitazione del novero dei soggetti collettivi che possono rispondere dei reati commessi nel loro interesse o vantaggio, ai sensi d.lgs. 231/2001, in quanto in quanto l'art. 1 del detto decreto non include gli "enti pubblici non economici": e tali vengono considerate le pur denominate "aziende" sanitarie pubbliche.

Inoltre, lo stesso decreto legislativo non prevede l'omicidio colposo e le lesioni personali colpose (neppure se gravi o gravissime), fra i reati presupposto della responsabilità dell'ente, menzionando solo (nella vincolante lista contenuta in questo corpo normativo, pour via via estesa nel tempo), i delitti colposi predetti commessi con violazione delle norme a tutela della salute e sicurezza sui luoghi di lavoro (come recita l'art. 25 septies).

La proposta di riforma potrebbe essere quindi quella di estendere anche alle aziende sanitarie pubbliche questa forma di responsabilità, includendo i reati colposi di cui agli art. 589 e 590 c.p., con i limiti di cui all'art. 590 sexies c.p., nell'elenco dei reati presupposto.

2.5. Altro settore di grande interesse per le possibili responsabilità penali derivanti dall'uso di sistemi di I.A. è quello del c.d. trading algoritmico, che può portare a configurare abusi di mercato (ad es. abuso di informazioni privilegiate ex art. 184 TUF e manipolazioni del mercato ex art. 185 TUF) realizzati tramite calcoli probabilistici e strategie operative automatizzate.

3. Per un adeguamento dei modelli d'imputazione all'uomo od all'ente che "sta dietro" i sistemi di I.A.

Ma al centro deve sempre riconoscersi l'interazione uomo-macchina e, come suggerisce anche la citata risoluzione dell'AIDP, lungi dal lasciare spazio a zone di impunità che diverrebbero sempre più ampie ed inaccettabili, con il crescente ricorso a detti sistemi algoritmici, occorre piuttosto rielaborare ed adattare le categorie dell'imputazione penale ai nuovi fenomeni, muovendo dai "modelli" già esistenti di responsabilità dell'ente, di responsabilità da prodotto, di responsabilità per la salute e sicurezza sui luoghi di lavoro, che offrono paradigmi sperimentati per ascrivere la responsabilità "da reato" a soggetti umani che pur non

siano direttamente gli autori materiali dell'ultimo anello della catena che determina l'offesa dei beni giuridici meritevoli e bisognosi della tutela penale.

In questa rielaborazione, resta essenziale l'apporto della dottrina e l'interdisciplinarietà degli approcci, che devono integrare le conoscenze giuridiche fra i vari ambiti del diritto, anche in una prospettiva comparata e sovranazionale, con quelle scientifiche e tecniche, indispensabili per comprendere i meccanismi che presiedono al funzionamento ed all'utilizzazione dei sistemi in esame, risalendo all'uomo "che sta dietro" di essi.