CATEGORIE TRADIZIONALI DEL DIRITTO PENALE E INTELLIGENZA ARTIFICIALE: CRISI O PALINGENESI? (*)

Le raccomandazioni dell'*Association Internationale de Droit Pénal* e la rilevanza del recente regolamento europeo sull'intelligenza artificiale

di Lorenzo Picotti

Il testo, rielaborato in italiano con alcune integrazioni e corredato di note bibliografiche, riprende la relazione, tenuta in francese, nel primo panel del XXI Congresso internazionale di Diritto penale, organizzato dall'AIDP a Parigi il 25-28 giugno scorso. Vengono in sintesi richiamate le raccomandazioni approvate all'esito dei lavori della prima Sezione, avente per tema di parte generale la domanda: "crisi o palingenesi" delle categorie penalistiche tradizionali di fronte all'emergere dell'intelligenza artificiale? Esclusa l'opportunità di un'imputazione diretta della responsabilità penale ai sistemi "intelligenti" in quanto tali, le raccomandazioni esposte, relative alla causalità, ai reati dolosi, ai reati colposi, alle posizioni di garanzia ed ai modelli di responsabilità da reato delle persone giuridiche, sono aggiornate ed integrate alla luce di significative disposizioni contenute nel regolamento dell'Unione europea sull'intelligenza artificiale (AI Act), recentemente approvato.

(*) Relazione tenuta il 26 giugno 2024, nel primo panel del XXI Congresso internazionale di diritto penale organizzato dall'*Association Internationale de Droit Pénal* (AIDP), Parigi, 25-28.6.2024 (Congresso del Centenario dalla fondazione).

Editore Associazione "Progetto giustizia penale" | via Festa del Perdono 7, 20122 Milano
c/o Università degli Studi di Milano, Dipartimento di Scienze Giuridiche "Cesare Beccaria"
redazione@sistemapenale.it

1. Introduzione.

Ringrazio Katalin Ligeti, che presiede questo primo panel e mi ha presentato in modo lusinghiero, dandomi la parola[1]; ringrazio il prof. Didier Rebut, che è il responsabile dell'organizzazione di questa straordinaria celebrazione parigina del Centenario della fondazione dell'AIDP[2]; ringrazio i colleghi ed amici del Consiglio direttivo, primo fra tutti il presidente John Vervaele, per avermi dato (ancora una volta) fiducia, assegnandomi - cinque anni orsono - il ruolo e la responsabilità di relatore generale di un'importante sezione del Congresso internazionale, dedicata al diritto penale generale, come già era stato per il Congresso di Istanbul del 2009.

Parlo in francese, non solo per ragioni di territorialità e di ospitalità. Il francese è la lingua con cui la nostra Associazione è stata fondata e battezzata a Parigi un secolo fa, e la preferisco, come bella lingua latina, che ho appreso in gioventù, prima del tedesco e dell'inglese, quando era considerata la lingua internazionale per eccellenza, ed era una delle lingue prescelte nelle scuole italiane.

Ricordo che ai primi incontri dell'AIDP cui ho partecipato, a partire da un colloquio preparatorio tenutosi a Friburgo in Brisgovia, in Germania, all'inizio degli anni '80 sulla criminalità economica, diretto dai professori Hans-Heinrich Jescheck e Klaus Tiedemann, la lingua parlata era soprattutto il francese. Ed avevo quindi potuto seguire bene le relazioni e la discussione in quella lingua, che loro stessi conoscevano perfettamente, al pari di autorevoli professori, che – come il prof. Cesare Pedrazzi – intervennero in quell'occasione.

È, quindi, un grande onore per me, intervenire oggi come relatore generale sul tema oggetto della prima Sezione del Congresso del Centenario, dedicato alla Giustizia penale di fronte all'intelligenza artificiale, ed in particolare sulle sfide che questa pone, oltre che alla società contemporanea, agli studiosi ed ai penalisti, che devono mettere alla prova le loro tradizionali categorie dogmatiche con le nuove realtà.
È una sfida emblematica, oggi forse la più forte, ma non certo la prima, che l'emergere di nuove tecnologie pone al sistema penale.

Inizierò, quindi, con una breve premessa di carattere metodologico sul suo significato paradigmatico (par. 2), per poi presentare una sintesi schematica di quelli che considero i punti più rilevanti delle raccomandazioni approvate al Colloquio internazionale svoltosi a Siracusa nel settembre 2022, a conclusione dei lavori della prima Sezione, intitolata per l'appunto alle sfide poste dall'intelligenza artificiale alle "categorie tradizionali del diritto penale: crisi o palingenesi?"[3] (par. 3), raccomandazioni

---

[1] La prof.ssa Katalin Ligeti, Direttrice della Facoltà di Giurisprudenza dell'Università di Lussemburgo, è stata eletta nuova Presidente dell'AIDP dall'assemblea generale dei soci, tenutasi a Parigi il 28.6.2024 al termine del Congresso, subentrando così al prof. John Vervaele, già presidente per due mandati dal 2014.

[2] Il prof. Didier Rebut è il Direttore dell'Istituto di Criminologia e di diritto penale dell'Università Panthéon-Assas di Parigi, che ha ospitato l'evento.

[3] Il rapporto generale, il testo della risoluzione ivi approvata ed una selezione dei piú significativi rapporti nazionali si possono leggere nel fascicolo monografico, a cura di L. PICOTTI e B. PANATTONI, *Traditional Criminal Law Categories and AI: Crisis or Palingenesis?*, in *Revue Internationale de Droit Pénal* (d'ora in poi: *RIDP*), 1/2023.

fatte ufficialmente proprie dall'AIDP in questo Congresso. Ma aggiungerò - visto il tempo trascorso – anche qualche considerazione ulteriore e, soprattutto, taluni riferimenti a disposizioni del regolamento europeo sull'intelligenza artificiale, recentemente approvato dopo un lungo *iter* e numerosi emendamenti[4], in quanto tali norme – come già era stato evidenziato nella risoluzione di Siracusa con riferimento alla proposta della Commissione[5], su cui iniziava il dibattito - possono avere un significativo impatto per una più precisa individuazione dei presupposti e dei limiti delle responsabilità penali configurabili in relazione alla produzione, alla diffusione ed all'utilizzo di questi sistemi.

2. Premessa generale e metodologica.

La rapidità dell'evoluzione sociale, scientifica e tecnologica ha sempre rappresentato una sfida rispetto alla lentezza con cui la legislazione penale, la giurisprudenza e, in generale, le categorie giuridiche dei sistemi penali si adattano, elaborando le necessarie risposte.

Molte nozioni tradizionali, come la causalità, il dolo, la colpa, non sono state mai improvvisamente accantonate, perché potevano sembrare obsolete di fronte a nuove conoscenze scientifiche, conquiste tecnologiche od a fenomeni nuovi che si manifestavano. Il tempo più lungo, di cui hanno sempre avuto bisogno i diversi formanti del diritto, per adeguarsi alle nuove realtà e fornire idonee risposte, non è necessariamente da vedere quale fattore negativo, quale freno che le categorie tradizionali rappresenterebbero rispetto all'evoluzione tempestiva del sistema giuridico, limitandone o ritardandone il cambiamento, quasi fossero solo funzionali a conservare lo *status quo*.

Piuttosto, si è trattato – e si tratta – di tempi utili, se non indispensabili, per consentire uno sviluppo ponderato e graduale, che partendo dai concetti acquisiti, di cui deve essere sempre riconosciuto e mantenuto il valore dogmatico e sistematico, anche per la pratica legislativa e giurisprudenziale, deve però portare all'elaborazione di contenuti innovativi, più adatti alla diversa realtà da regolamentare, man mano che questa emerge.

Certo, restano dei limiti invalicabili da salvaguardare, espressi dai principi di legalità, di colpevolezza, di *extrema ratio*, di proporzionalità, per citare solo i più importanti,

[4] Se veda il testo del regolamento (UE) 2024/1689 del Parlamento europeo e del Consiglio del 13 giugno 2024 che stabilisce regole armonizzate sull'intelligenza artificiale e modifica i regolamenti (CE) n, 300/2008, (UE) n, 167/2013, (UE) n, 168/2013, (UE) 2018/858, (UE) 2018/1139 e (UE) 2019/2144 e le direttive 2014/90/UE, (UE) 2016/797 e (UE) 2020/1828 (regolamento sull'intelligenza artificiale), pubblicato in GUUE del 12.07.2024, serie L, It. La prima proposta della Commissione europea risale al 21.4.2021 (COM (2021) (206) final. Le citazioni di articoli, non seguite nel presente testo da una diversa specificazione, si intendono riferite al regolamento predetto (c.d. *AI Act*).

[5] Cfr. in particolare il par. 4, punto 14, delle Raccomandazioni contenute nella *Resolution on Traditional Criminal Law Categories and AI*, in *RIDP*, cit., p. 57.

enunciati ormai solennemente anche dalle Carte dei diritti fondamentali e dalle Costituzioni degli Stati di diritto.

Ma i dogmi concettuali, in quanto tali, non devono essere considerati immutabili, perché sono il risultato di un processo di formazione storicamente determinato, frutto dello sviluppo della dottrina e della giurisprudenza, oltre che della legislazione, affermatesi in concrete condizioni storiche, culturali, sociali, politiche e filosofiche, che quindi con il loro mutare non possono che modificarsi, come del resto è già stato in passato.

Avvicinandomi al tema odierno, si può considerare paradigmaticamente il concetto penalistico di "azione", che si è evoluto da un'originaria nozione naturalistica e causale, legata al movimento corporeo dell'uomo, capace di modificare il mondo esterno (in contrapposizione laica ad una concezione moralistica interiore, derivata dall'idea di peccato), che rifletteva le acquisizioni del positivismo dell'Ottocento, verso una nozione finalista e personalistica, basata sul ruolo caratterizzante dello scopo perseguito dal soggetto agente, in grado di guidare consapevolmente la propria condotta, in conformità alla rivendicata centralità dell'individuo, fino alla nozione c.d. sociale, che ravvisa le caratteristiche rilevanti per il diritto penale nell'incidenza che ha la condotta umana (non solo attiva, ma anche omissiva) nelle relazioni con i consociati e con la collettività, evidenziandone l'incidenza sugli interessi e diritti da proteggere, in un'acquisita prospettiva solidaristica.

Ebbene, si tratta di una nozione che oggi può e anzi deve essere ulteriormente sviluppata, fino ad abbracciare l'apporto delle nuove tecnologie, che in una parte sempre più rilevante giungono a supportare ed anche a sostituire atti significativi dell'uomo, determinando effetti giuridicamente rilevanti nelle relazioni con gli altri, con il mondo esterno, con l'ambiente, compresa l'eventuale causazione di lesioni a beni giuridicamente protetti dal diritto penale, che non si presenta però quale diretta estrinsecazione di movimenti né di concrete decisioni di una determinata persona umana, data l'interazione che si ha con gli strumenti tecnologici, dotati di un crescente grado di autonomia. E nondimeno tali offese richiedono – come è emerso nei lavori di cui esporrò i risultati - un'imputazione alla volontà e responsabilità di persone umane, perché si continui a garantire la protezione dei medesimi beni ed interessi.

Cito, ad esempio, la recente normativa francese sulla guida dei veicoli automatici, che fa riferimento alla possibile "delega di guida" dal conducente umano al mezzo che può operare in modalità autonoma, non escludendo, ma determinando un possibile trasferimento di responsabilità al produttore, su cui il collega francese, Maître Julien Walther, presente nel nostro panel, sarà certamente più preciso, come lo è stato nel rapporto nazionale redatto con la sua collega Marion Lacaze[6].

Anche la nozione di "colpa" si è da tempo allontanata da una concezione naturalistica, che la riconduceva all'elemento psicologico del reato, come ancora si legge nella rubrica dell'art. 43 del codice penale italiano, secondo cui la negligenza,

[6] Rinvio a M. LACAZE, J. WALTHER, *French Report on Traditional Criminal Law Categories and AI*, in *RIDP*, 1/2023, p. 153 s.

l'imprudenza e perfino l'imperizia erano concepite come mancanze psicologiche di attenzione e di cautela, per giungere ad una concezione strettamente normativa, incentrata sulla violazione di oggettive regole cautelari, scritte o ricavate dall'esperienza, a fronte della quale il contenuto personalistico di colpevolezza è garantito dalla c.d. misura soggettiva, che deve accompagnare quella oggettiva, vale a dire dal rimprovero personalistico per "non aver agito in modo diverso" da quello che sarebbe stato non solo legalmente richiesto, ma anche personalmente possibile e, dunque, esigibile, senza però che un effettivo contenuto psichico debba venire in rilievo.

Richiamo questi paradigmi concettuali e normativi, perché sono la base da cui muovere per poterli adattare alle sfide poste dai sistemi di intelligenza artificiale.

Ed un ultimo esempio è utile alla nostra argomentazione, e riguarda la responsabilità penale delle persone giuridiche, che era un tabù nell'Europa continentale fino a pochi decenni orsono, in cui l'adagio "*societas delinquere non potest*" era considerato alla stregua di un dogma insormontabile, collegato alla nozione personalistica di azione penalmente rilevante, ma che, tuttavia, per la necessità di un più efficace contrasto alla criminalità economica e delle società commerciali in genere, vere attrici dei processi sociali, ha portato ad accogliere - sotto l'impulso determinante dell'Unione europea - l'esperienza dei Paesi di *common law*, e si è rapidamente diffusa nei nostri ordinamenti giuridici, dalla Francia all'Olanda, dal Belgio alla Spagna, fino ai più riluttanti, come l'Italia e la Germania, rimasti maggiormente condizionati dal menzionato dogma: per cui l'hanno recepita qualificando nominalmente come "amministrative" le relative sanzioni, pur aventi un espresso contenuto punitivo, correlato e proporzionato ai reati ad esse imputabili.

Questo modello di responsabilità si sta rivelando sempre più importante nei settori in cui operano organizzazioni complesse, necessarie per gestire le forme più avanzate di produzione e di attività, non solo economiche, nelle quali l'applicazione delle nuove tecnologie gioca un ruolo crescente e crea anche nuove tipologie di rischi: per cui dovrà essere opportunamente valorizzato nell'ambito che vogliamo trattare (cfr. *infra*, par. 3.5).

Si pensi al campo della tutela dell'ambiente, od a quello della protezione della salute e sicurezza sui luoghi di lavoro, od ancora a quello delle transazioni di borsa ormai sempre più gestite da algoritmi, fino all'intero settore *Fin-tech*, e tanti altri ancora, nei quali l'Unione europea ha introdotto nuovi obblighi di criminalizzazione, richiedendo sempre che accanto alla responsabilità penale delle persone fisiche, cui sono imputabili i reati, si preveda anche la responsabilità delle persone giuridiche, nell'interesse delle quali esse agiscono.

Questo nuovo modello di responsabilità, pur essendo costruito con riferimento ad entità collettive costituite comunque da esseri umani, prefigura uno schema di responsabilità per fatti che non sono materialmente commessi dalle persone giuridiche, vale a dire dai soggetti cui si applicano le corrispondenti sanzioni, bensì dalle persone fisiche che "vi stanno dentro" agendo per loro conto e nel loro interesse. E può dunque rilevare quale schema di riferimento per l'imputazione di fatti materialmente realizzati tramite sistemi di intelligenza artificiale, che operano nell'interesse di persone od enti che "vi stanno dietro".

3. I punti rilevanti delle raccomandazioni approvate nel colloquio internazionale di Siracusa conclusivo dei lavori della prima Sezione.

Sulla base di tali premesse generali, passo ora ad offrire un riassunto schematico dei punti più rilevanti, contenuti nelle raccomandazioni approvate all'esito del colloquio internazionale conclusivo dei lavori della prima Sezione del nostro Congresso, svoltosi il 15 e 16 settembre 2022 presso il "*Siracusa International Institute for Criminal Justice and Human Rights*". In tale esposizione richiamerò – rinviando a necessari futuri approfondimenti – talune disposizioni del sopra menzionato regolamento europeo sull'intelligenza artificiale, recentemente approvato, che seppur non ancora direttamente applicabile[7], possono rilevare per definire i presupposti e limiti della responsabilità penale che può nascere in relazione ai nuovi sistemi.

3.1. Il punto di partenza del questionario e la risposta negativa sull'ipotetica capacità d'agire penalmente rilevante dei sistemi di intelligenza artificiale.

Il questionario, piuttosto articolato, distribuito ai gruppi nazionali, che è servito come punto di partenza per i lavori della prima Sezione[8], aveva, innanzitutto, messo in evidenza la caratteristica peculiare di crescente autonomia che distingue i sistemi di intelligenza artificiale da altri sistemi informatici, i quali, anche se molto sviluppati dal punto di vista tecnologico, si basano pur sempre sull'esecuzione di programmi predefiniti dall'uomo e destinati ad elaborare un *set* predeterminato di dati ed informazioni.

Viceversa, ciò che qualifica i sistemi di intelligenza artificiale, ponendo nuove sfide al diritto non solo penale, specie in tema di responsabilità correlate alla loro produzione ed al loro utilizzo, è l'autonomia di apprendimento (attraverso le varie tecniche di *machine learning*) e l'autonomia decisionale, ad essa correlata, che si sviluppa sulla base dell'"esperienza" via via acquisita dai sistemi stessi, in grado di elaborare in tempi minimi quantità enormi di dati ed informazioni, sempre aggiornate, da essi ricercate nel *web* e nel mondo esterno (grazie alla potenza delle connessioni, delle memorie e dei processori, oltre alla qualità dei sensori visivi, acustici, termici, ecc. di cui possono avvalersi): tanto che gli stessi algoritmi sono in grado di adattarsi ed evolversi, con risultati finali (*out put*) che possono essere imprevedibili (*unpredictable*) anche per i programmatori ed i produttori, oltre che per gli utilizzatori finali.

---

[7] Come stabilisce l'art. 113 del regolamento, entrato formalmente in vigore il ventesimo giorno successivo alla sua pubblicazione (vale a dire il 2.7.2024), esso si applica però solo dal 2.8.2026, per dare tempo ad autorità, aziende, cittadini di adeguarvisi, salvo alcune importanti norme che si applicano già dal 2.2.2025 (par. 3, lett. a) e dal 2.8.2025 (par. 3, lett. b), mentre l'art. 6 par. 1 si applicherà solo dal 2.8.2027 (par. 3, lett. c). Nondimeno, ai fini delle valutazioni penalistiche sui vari punti trattati, il regolamento ha fin d'ora un importante valore giuridico e concettuale, cui fare riferimento.

[8] Si può leggere nel sito dell'AIDP: https://www.penal.org/sites/default/files/Questionnaires%20EN.pdf

Se l'autonomia nascente dalle capacità cognitive e decisionali dei sistemi di intelligenza artificiale ne rappresenta il valore aggiunto, non essendo l'uomo in grado di garantire analoghe prestazioni, tantomeno nei medesimi tempi, essa pone però la questione della c.d. *gap responsibility*, trattandosi di stabilire chi risponda per il loro operato, dal quale possono scaturire nuovi rischi e danni, non prodotti direttamente da atti dell'uomo, né da lui singolarmente decisi o controllati, essendo del resto un controllo completo confliggente con l'utilità stessa dell'automazione. Ma in gioco vengono beni giuridici e diritti fondamentali meritevoli di protezione penale, anzi già oggetto di tutela anche penale, se le offese provengono direttamente dall'azione di un uomo (si pensi a morti o lesioni cagionate da una manovra difettosa di un'auto a guida autonoma, o da erronei interventi diagnostici o chirurgici realizzati da robot, o ad abusi di mercato realizzati dagli algoritmi ad alta frequenza, che dominano oggi le transazioni di borsa, od ancora ad uccisioni o danneggiamenti posti in essere da droni ed armi c.d. intelligenti, ecc.).

La risoluzione approvata a Siracusa, sulla base della grande maggioranza delle posizioni espresse anche nei rapporti nazionali[9], ha escluso che - allo stato attuale dello sviluppo tecnologico - si possa o, comunque, sia utile riconoscere una vera e propria capacità di agire, giuridicamente rilevante, in capo ai sistemi di intelligenza artificiale in quanto tali, che possa essere la base per una loro diretta responsabilità penale, come pur suggerito da una minoritaria parte della dottrina americana, e prospettato in alcuni paesi come la Cina, su cui la collega Wang Xumei, partecipante al presente panel, ha avuto modo di soffermarsi in modo dettagliato nel relativo rapporto nazionale[10].

Nella risoluzione citata è stato sottolineato che i sistemi di intelligenza artificiale restano ontologicamente diversi dalle persone umane, perché ad essi manca la libertà morale di autodeterminazione e la capacità di considerare e valutare, con la necessaria flessibilità, anche il concreto contesto personale, etico e sociale, rispetto a cui va decisa la soluzione migliore di possibili problemi o dilemmi, oltre al fatto che sono privi di coscienza di sé nel passato, nel presente e, quindi, anche proiettata nel futuro[11].

La sanzione penale, pur se ad essi adattata, non potrebbe, quindi, svolgere una funzione di prevenzione speciale, e neppure di prevenzione generale, né tantomeno una funzione retributiva, in assenza di qualsiasi possibilità di rimprovero etico-giuridico per decisioni e comportamenti realizzati dagli algoritmi[12], e non dalle persone che pur li hanno concepiti, prodotti, messi a disposizione od utilizzati.

---

[9] Per riferimenti rinvio a L. PICOTTI, *Traditional Criminal Law Categories and AI: Crisis or Palingenesis? General report*, in *RIDP*, 1/2023, p. 11 s., in specie p. 16 s.

[10] Cfr. X. WANG, X. ZHANG, *Chinese Report on Traditional Criminal Law Categories and AI*, in *RIDP*, 1/2023, p. 111 s.

[11] Cfr. in particolare il par. 1, punti 4 e 5 delle Raccomandazioni di cui alla *Risolution*, cit., p. 53 s., in specie 55.

[12] *Ibidem.*

3.2. Se il diritto penale debba o meno fornire risposte per prevenire e punire i reati commessi da, attraverso o contro i sistemi di intelligenza artificiale.

La successiva domanda fondamentale, a cui è stata dedicata la prima Sezione, è stata quindi se il diritto penale debba o meno fornire un'adeguata risposta, in grado di prevenire e punire i reati commessi da, attraverso o contro i sistemi di intelligenza artificiale, indirizzandosi alle persone umane (fisiche e giuridiche) che "stanno dietro" il loro operato; ed, in caso affermativo, se per tale risposta siano applicabili le tradizionali categorie su cui si fonda la responsabilità penale o se, al contrario, ne sia necessaria una "palingenesi".

La risoluzione ha dato una risposta sostanzialmente affermativa ad entrambe le questioni, muovendo dalla riconosciuta necessità di proteggere, anche con sanzioni penali da minacciare ed applicare a persone umane (fisiche o anche giuridiche), i diritti fondamentali ed i beni giuridici che vengano offesi tramite sistemi di intelligenza artificiale, o con aggressioni a loro danno, visto che tali beni giuridici già beneficiano di una tutela penale che, per la loro meritevolezza, è ritenuta necessaria dagli ordinamenti vigenti. Per cui non è possibile accettare, da un punto di vista logico-giuridico, oltre che politico criminale, che fatti che costituirebbero reato, se commessi da un essere umano, restino o diventino non punibili, solo perché è coinvolto nella loro commissione o, comunque, riguardano un sistema di intelligenza artificiale[13]. Tanto più che una tale area di impunità sarebbe destinata ad espandersi in maniera esponenziale ed intollerabile nel prossimo futuro.

*3.3. Sui criteri di attribuzione della responsabilità penale alle persone che "stanno dietro" ai sistemi di intelligenza artificiale***.**

Il passo successivo è stato quello di identificare le persone e le entità giuridiche che – stando "dietro" ai sistemi di intelligenza artificiale – possano essere ritenute penalmente responsabili, sulla base della preliminare constatazione empirica che, se un sistema di intelligenza artificiale viene progettato, prodotto, commercializzato, reso disponibile ed utilizzato, tutto ciò avviene nell'interesse od a vantaggio di una persona fisica o giuridica che ne può fruire e disporre, nelle sue varie fasi di vita[14].

Questa osservazione preliminare non è ovviamente sufficiente, per individuare anche i criteri di attribuzione della responsabilità penale, che devono soddisfare i requisiti di personalità e di colpevolezza per il fatto offensivo da cui dipende la pena o, comunque, la sanzione punitiva minacciata dall'ordinamento. Per cui è apparsa necessaria una parziale revisione o quantomeno adattamento delle categorie tradizionali che vengono in rilievo.

---

[13] Par. 3, punto 9 delle Raccomandazioni di cui alla *Risolution,* cit., p. 55.

[14] Par. 3, punti 10 ed 11 delle Raccomandazioni di cui alla *Risolution,* cit., p. 56.

3.3.1. Sul nesso di causalità.

In primo luogo, va sottolineata la necessità di verificare l'esistenza di un nesso di causalità tra la condotta od il comportamento attribuibile ad un'azione o ad un'omissione umana, ed il fatto o l'evento che si è poi verificato, a seguito dell'intervento finale e decisivo di un sistema di intelligenza artificiale.

A questo proposito, sembrano adottabili i comuni criteri di imputazione oggettiva, che muovono dalla preliminare formula della *conditio sine qua non*, secondo la quale il nesso causale penalmente rilevante deve essere affermato, a seguito di un ragionamento ipotetico, con riferimento a qualsiasi contributo – attivo od omissivo – senza il quale l'evento od il fatto non si sarebbe verificato. Ed è indubbio che alla base della progettazione, produzione, distribuzione, messa a disposizione ed utilizzazione di sistemi di intelligenza artificiale vi è sempre un volontario atto di soggetti umani, senza il quale la catena causale non verrebbe neppure attivata. A tale prima verifica deve però aggiungersi anche un criterio di selezione normativo, sia per evitare un *regressum ad infinitum*, sia perché devono escludersi dall'imputazione oggettiva i fatti od eventi del tutto anomali, che siano cioè completamente estranei alla relazione (o nesso) di rischio che deve comunque sussistere, secondo criteri scientifici, tecnici o statistici, tra l'intervento umano da considerare e l'evento (o fatto) che è stato "deciso" od anche direttamente realizzato, nel suo anello finale, dal sistema di intelligenza artificiale[15].

Criteri normativi che dovranno orientare anche quella preventiva valutazione dei rischi (*risk assessment*) che ora, con vincolante precetto di diritto positivo, è imposta ai titolari ed operatori dei sistemi classificati "*ad alto rischio*"[16] dal menzionato

[15] Il problema nasce soprattutto con riferimento al fenomeno delle c.d. *black box*, vale a dire a di quei processi o fasi che portano ad *output* di cui non è possibile ricostruire, neppure da parte dei tecnici che hanno progettato i sistemi stessi, tutti i singoli anelli della catena che ha portato in concreto a quell'esito, che potrebbe non essere ripetibile. Problema che se può avere un rilievo in sede di accertamento (rispetto a cui, peraltro, la memorizzazione automatica e le tracce elettroniche delle varie procedure ed attività di tali sistemi possono fornire un fondamentale supporto) non preclude di per sé la possibilità d'imputazione oggettiva, secondo criteri normativi che sono - almeno metodologicamente - da tempo condivisi, non occorrendo mai la spiegazione di ogni singolo passaggio della catena causale da cui si produce l'evento finale, potendo soccorrere anche correlazioni statistiche od ipotetiche ragionevolmente plausibili.

[16] La relativa classificazione si basa sulle caratteristiche di cui all'art. 6, par. 1, del regolamento (la cui effettiva applicazione è però posticipata al 2.8.2027: cfr. *supra*, nota 7) e rinvia, inoltre, all'elenco di settori in cui operano, contenuto nell'Allegato III al regolamento stesso, peraltro modificabile, al pari di altre norme, dalla Commissione tramite atti delegati (in forza dell'art. 7), nel rispetto delle condizioni date. Sono esclusi da tale classificazione (e dai conseguenti obblighi), per cui si possono delimitare anche *a contrario* le caratteristiche dei "sistemi ad alto rischio" (che sono certamente i più importanti da considerare, non solo per il gran numero riconducibile a tale categoria, ma soprattutto ai fini dell'imputazione della responsabilità penale, perché sono quelli che coinvolgono per definizione i più rilevanti interessi e beni giuridici, nonché diritti fondamentali da proteggere), i sistemi che invece non presentino "*un rischio significativo di danno per la salute, la sicurezza o i diritti fondamentali delle persone fisiche, anche nel senso di non influenzare materialmente il risultato del processo decisionale*" (art. 6, par. 3), per i quali valgono solo alcune norme di carattere generale.

regolamento dell'Unione europea sull'intelligenza artificiale, il quale fornisce al contempo elementi rilevanti per delineare siffatti criteri[17].

Dunque non si tratta certo di abbandonare la fondamentale garanzia dell'oggettivo nesso di causalità, che fonda l'imputazione oggettiva, e che deve sussistere tra la condotta umana punibile e l'evento o fatto complessivamente imputato. Ma i relativi criteri normativi di delimitazione ed accertamento dovranno essere adeguati al portato ed alle conoscenze delle nuove tecnologie, oltre che alla pertinente disciplina extrapenale, per stabilire la sussistenza in concreto dell'oggettivo "nesso di rischio" realizzatosi (o meno) nell'evento o fatto penalmente rilevante.

3.3.2. Sull'elemento soggettivo o psicologico del reato.

Il dibattito sul contenuto della risoluzione approvata si è concentrato soprattutto sul tema dell'elemento soggettivo o psicologico del reato, al quale sono state dedicate una gran parte delle raccomandazioni, distinguendo preliminarmente far reati commissibili nell'ambito di attività *ab origine* illecite e quelli commissibili invece nell'ambito di attività in sé lecite di base. Nel primo caso l'attenzione si è ovviamente appuntata sui reati dolosi, nel secondo sui reati colposi.

3.3.2.1. Sui reati dolosi.

La risoluzione muove dalla constatazione che se esiste un'intenzione che abbraccia anche l'operato del sistema di intelligenza artificiale, prescelto come *strumento* (sotto questo profilo non differente da altri possibili) per raggiungere il fine prefissato o, comunque, per realizzare la "volontà consapevole" dell'agente (si pensi all'uccisione di una persona tramite un drone intelligente in grado di ricercarla e riconoscerla; o ad un abuso di mercato realizzato tramite algoritmi ad alta frequenza, impostati per sfruttare al meglio un rialzo di valori, previamente indotto tramite una moltiplicazione artificiosa di domande d'acquisto), non si ravvisano ostacoli particolari per imputare e rimproverare il fatto all'agente umano che "sta dietro" all'utilizzo di detti sistemi di intelligenza artificiale, bastando richiamare i comuni principi in materia di dolo, comprese le regole da applicare nei casi di *aberratio ictus* e di *aberratio delicti,* se i risultati siano diversi da quelli che l'agente voleva e si era rappresentato[18].

Tuttavia, è stata sottolineata la necessità di introdurre nuovi reati preparatori, od a consumazione anticipata, per punire le violazioni dolose di specifici precetti, che

---

[17] Si veda in particolare l'art. 9 rubricato "*Sistema di gestione dei rischi*" che – fra le altre cose - impone di considerare sia quelli "*che possono emergere quando il sistema è utilizzato in conformità alla sua finalità prevista*" sia quando lo è "*in condizioni di uso improprio ragionevolmente prevedibile*" (par. 2, lett. b). Sull'importanza basilare di tale disciplina, dettagliatamente articolata, anche per definire i limiti dei "*rischi accettabili*", si avrà modo di ritornare *infra*, par. 3.3.2.2.

[18] Cfr. par. 3, punto 12, *sub* a.1, delle Raccomandazioni di cui alla *Risolution,* cit., p. 56.

riguardino singoli operatori od anelli dell'intera catena che precede l'operato concreto di questi sistemi, che va dalla progettazione, allo sviluppo, alla produzione e commercializzazione, fino alla fornitura e messa a disposizione per l'utilizzo finale, in modo che si possano più efficacemente prevenire offese e violazioni più gravi.

Al riguardo la risoluzione approvata a Siracusa ha rinviato ai limiti ed alle condizioni generali che legittimano questo tipo di incriminazioni di atti meramente preparatori di più gravi delitti, contenuti nella risoluzione della prima Sezione del XVIII Congresso AIDP tenutosi ad Istanbul nel 2009[19], secondo cui occorre, fra l'altro, che vi sia un pericolo chiaro ed attuale di offesa alla vita, all'integrità o alla libertà di altri esseri umani[20].

Bisogna ora aggiungere che il regolamento europeo sull'intelligenza artificiale contiene un nucleo molto importante di precetti, obbligazioni e standards da rispettare, fin dalle diverse fasi prodromiche rispetto all'effettivo utilizzo, che possono costituire quell'indispensabile premessa normativa dell'intervento penale, già richiamata, in termini generali, nelle raccomandazioni approvate a Siracusa[21].

In particolare, il regolamento europeo vieta fin dall'origine la creazione di determinati sistemi di intelligenza artificiale che comportino "*rischi inaccettabili*"[22] e pone poi un reticolo di norme da rispettare, da parte dei diversi attori, od "*operatori*"[23], per la progettazione, lo sviluppo, la produzione, la messa in circolazione nonché a disposizione dei "*sistemi ad alto rischio*".

Tali divieti e precetti possono, dunque, essere alla base di autonome fattispecie penalmente sanzionate, aventi un momento consumativo anticipato rispetto al verificarsi di eventi dannosi o, comunque, alla compiuta offesa dei beni giuridici di elevato valore da proteggere, quali la vita, l'integrità personale e la libertà, o più ingenerale i diritti fondamentali, più volte richiamati dal regolamento quali limiti da salvaguardare in ogni attività che coinvolga i sistemi di intelligenza artificiale.

Ad esempio, si possono citare le norme che vietano in modo assoluto le pratiche elencate nell'art. 5 del regolamento, come quelle dell'intelligenza artificiale "*manipolativa*" (lett. a), o che comportano lo sfruttamento delle vulnerabilità della vittima per ragioni dovute all'età, alla disabilità o ad una particolare situazione sociale o

---

[19] Si veda il rapporto generale curato dal sottoscritto, presentato e discusso nel Colloquio preparatorio di La Coruña (5-8.9.2007), e la proposta di risoluzione ivi approvata, in *RIDP*, 3-4/2007, rispettivamente p. 355 s. e p. 509 s.

[20] Il testo della risoluzione definitivamente approvata, dopo l'ampia discussione nel XVIII Congresso di Istanbul del 20-27.9.2009 (che ha apportato alcuni emendamenti, come era possibile secondo le regole interne allora vigenti), si può leggere sempre in *RIDP*, 1-2/2015, p. 421 s. (sul punto cfr. in specie lettera A, par. II. punto 2).

[21] Cfr. l'intero par. 2 delle Raccomandazioni di cui alla *Risolution,* cit., p. 55, che precede tutti quelli poi dedicati ai profili specificamente penali, ed è intitolato: "*Sulla necessità di regole, standards e obbligazioni extrapenali*".

[22] Il Capo II del regolamento, costituito dal solo art. 5, composto peraltro da 8 articolati paragrafi, contiene un elenco di "*pratiche di intelligenza artificiale vietate*", per cui sono previste le pesanti sanzioni amministrative stabilite dall'art. 99, par. 3 (su tale sistema sanzionatorio si tornerà nel testo).

[23] Ai sensi dell'art. 3, n. 8 del regolamento, è considerato "*operatore*" "*un fornitore, un fabbricante del prodotto, un deployer, un rappresentante autorizzato, un importatore o un distributore*".

economica (lett. b); od ancora le pratiche che permettono di valutare o classificare le persone sulla base del loro comportamento sociale o delle loro caratteristiche personali, anche se note, o dedotte o previste (lett. c); i sistemi di *predictive policing* fondati unicamente sulla profilazione di una persona fisica o sulla valutazione di tratti della personalità (lett. d), come pure certi sistemi di riconoscimento facciale (lett. e) o che permettono di dedurre le emozioni di una persona sul posto di lavoro e per l'insegnamento negli istituti scolastici (lett. f); oppure che permettono la categorizzazione biometrica, traendo deduzioni o inferenze in merito a razza, opinioni politiche, appartenenza a sindacati, convinzioni religiose o filosofiche, vita od orientamento sessuale, ecc.

La gravità di simili pratiche, se intenzionalmente o comunque consapevolmente poste in essere, potrebbe legittimare, a prescindere dalla concreta offesa che si cagioni ai beni giuridici da proteggere e dall'applicazione di fattispecie penali già vigenti, un'incriminazione penale per la violazione dolosa dei relativi divieti, pur tenendo conto che già lo stesso regolamento, all'art. 99, par. 3, prevede per esse "*sanzioni amministrative pecuniarie fino a 35 000 000 Euro o, se l'autore del reato è un'impresa, fino al 7 % del fatturato mondiale totale annuo dell'esercizio precedente, se superiore*"[24].

Ma anche la violazione di altri singoli obblighi o divieti, che sono imposti a produttori, sviluppatori, fornitori, importatori, distributori, deployer[25] dagli artt. 8 e seguenti, in relazione ai sistemi classificati "*ad alto rischio*", cui come detto è dedicato l'intero Capo III, potrebbero essere oggetto di autonome incriminazioni penali, da affiancare alle meno gravi sanzioni amministrative previste invece dal successivo par. 4 del citato art. 99.

---

[24] Tale sistema di sanzioni amministrative - che comprende l'indicazione dettagliata dei criteri di commisurazione (art. 99, par. 7) e ne estende la possibilità di applicazione anche ad organi ed istituzioni dell'Unione (art. 100), oltre che ai "fornitori di modelli di intelligenza artificiale a finalità generale" (art. 101), di cui non ci si può occupare in questa sede - non si pone in alternativa, quanto piuttosto in parallelo rispetto alle eventuali sanzioni penali, che restano di competenza degli Stati, e che non avrebbero del resto potuto essere oggetto del regolamento dell'Unione europea, che ha una competenza solo concorrente in materia penale, da esercitare tramite direttive (*ex* art. 83 TFUE). Infatti il par. 1 del citato art. 99 del regolamento si limita a prevedere che "*gli Stati membri stabiliscano norme relative alle sanzioni*" (ed altre misure di esecuzione), che devono essere "*effettive, proporzionate e dissuasive*". Sulla compatibilità di un sistema di doppio binario con il divieto di *bis in idem*, alle condizioni elaborate dalla sofferta giurisprudenza europea e nazionale, sia consentito rinviare, per brevità, a L. PICOTTI, *Doppio binario sanzionatorio* e ne bis in idem*: verso un accettabile epilogo del lungo dialogo fra le corti?*, in A. CADOPPI, P. VENEZIANI, P. ALDROVANDI (a cura di), *Legalità e diritto penale dell'economia. Studi in onore di Alessio Lanzi*, Roma, 2020, p. 510 s.

[25] Si tratta di una figura inserita nel corso dei lavori parlamentari, che è così definita dall'art. 3, n. 4: "*una persona fisica o giuridica, un'autorità pubblica, un'agenzia o un altro organismo che utilizza un sistema di IA sotto la propria autorità, tranne nel caso in cui il sistema di IA sia utilizzato nel corso di un'attività personale non professionale*". Il termine non è tradotto in italiano nel testo pubblicato nella Gazzetta Ufficiale UE, ma potrebbe rendersi con una circonlocuzione che indichi la persona o l'ente - proprietario, possessore o con altro titolo che lo legittimi in termini giuridici - che concretamente "*mette a disposizione*" dell'utilizzatore finale il sistema d'intelligenza artificiale, interponendosi, quindi, rispetto al "*distributore*" o "*fornitore*" ed alle altre figure, definite rispettivamente ai numeri 7), 3) ecc. del medesimo art. 3, dedicato ad un lungo elenco di ben 69 definizioni.

3.3.2.2. Sui reati colposi.

Più complesse si sono dimostrate le questioni relative ai reati colposi, che possono essere commessi nell'esercizio di molteplici e tendenzialmente sempre più numerose attività lecite di base (quali la circolazione di veicoli a guida autonoma o l'uso di sistemi intelligenti e robotici nella diagnosi e nel trattamento medico-chirurgico, per restare ai settori già menzionati)[26]. In quest'ambito, in conformità al principio dell'*ultima ratio* che deve governare la previsione ed applicazione delle sanzioni penali, è anche concettualmente necessario che, prima che esse intervengano, vi sia o sia possibile ricostruire una regolamentazione extra-penale, che stabilisca gli standard tecnici da rispettare, le situazioni e condizioni di rischio da prevenire, le misure cautelari da adottare, in particolare in caso di *red flags*, e quant'altro va previsto e messo in opera nell'esercizio di dette attività per circoscrivere i rischi: dunque, fin dalle fasi della progettazione e sviluppo, e quindi della produzione, commercializzazione, vendita, messa a disposizione ed infine utilizzo concreto dei predetti sistemi[27].

A questo proposito, si possono richiamare i molteplici obblighi e divieti introdotti dal citato regolamento europeo sull'intelligenza artificiale in relazione ai sistemi "*ad alto rischio*".

Si tratta sia di obblighi generali, relativi, ad esempio, alla "qualità" dei *data set* su cui deve svolgersi il necessario allenamento (*training*) dei sistemi stessi (art. 10, che regola dettagliatamente quali debbano essere le "*pratiche di governance e gestione dei dati adeguate alla finalità prevista del sistema*", da adottare per i "*set di dati di addestramento, convalida e prova*"), sia di obblighi più specifici previsti per singoli operatori, ad esempio in capo al produttore ed allo sviluppatore, che devono garantire un adeguato livello di trasparenza e di informazione al deployer (art. 13), ovvero di accuratezza, di robustezza e di cybersicurezza dei sistemi stessi (art. 15).

Ma vengono in rilievo anche l'art. 42, che pone le condizioni perché i dati di addestramento e prova si possano presumere "*conformi*" ai requisiti di qualità menzionati, e l'art. 44, che regola le condizioni e procedure di valutazione della "*conformità*" ai requisiti stabiliti per tutti i sistemi "*ad alto rischio*" dalla sezione 2 del Capo III.

Certamente, dall'insieme delle norme in esame, si possono desumere specifiche regole cautelari, la cui violazione può essere alla base di un addebito per colpa. Ma nella definizione dei limiti della responsabilità penale per eventi e fatti avversi, che non sono voluti né concretamente previsti dalla persona umana, essendo riferibili all'impiego di

---

[26] Cfr. in specie il punto 12.b del par. 3 delle Raccomandazioni di cui alla *Risolution,* cit., p. 56-57, in cui si evidenziano i più importanti punti di attrito fra criteri di attribuzione della responsabilità penale e caratteristiche tecniche dei sistemi di intelligenza artificiale: la loro autonomia (1), la concreta imprevedibilità delle loro decisioni e del loro funzionamento (2), l'opacità dei meccanismi che li regolano (3), la complessità del loro processo di programmazione, sviluppo, produzione, aggiornamento e manutenzione (4).

[27] Cfr. il par. 2, punti 6 e 7 delle Raccomandazioni di cui alla *Risolution,* cit., p. 55.

sistemi di intelligenza artificiale nell'esercizio di attività di base lecite, la questione più delicata riguarda la previa delimitazione del perimetro del c.d. "rischio consentito"[28], non bastando la sola violazione di una cautela, pur doverosa, a fondare la responsabilità penale per il fatto o l'evento offensivo che si sia poi in concreto verificato.

Di grande rilievo appare a tal fine l'art. 9 del regolamento (già richiamato *supra* al par. 3.3.1. a proposito dell'individuazione dei parametri d'imputazione oggettiva), che stabilisce l'obbligo di istituire un "*sistema di gestione dei rischi*", quale requisito essenziale dei sistemi di intelligenza artificiale "*ad alto rischio*". Esso va "*inteso come un processo iterativo continuo pianificato ed eseguito nel corso dell'intero ciclo di vita* […], *che richiede un riesame e un aggiornamento costanti e sistematici*", da documentare e mantenere nel tempo. Rientrano nelle varie fasi in cui si deve articolare la gestione dei rischi, l'identificazione, l'analisi e la valutazione di quelli "*noti e* ragionevolmente prevedibili *che il sistema* [...] *può porre per la salute, la sicurezza e i diritti fondamentali*", sia quando è usato "*conformemente alla sua finalità prevista*" sia quando lo sia "*in condizioni di uso improprio* ragionevolmente prevedibile" (lettere a) e b), evidenziazioni aggiunte), mentre si devono considerare anche quelli "*derivanti dall'analisi dei dati raccolti dal sistema di monitoraggio successivo all'immissione sul mercato*" (lett. c).

L'ultima fase è quella dell'adozione di "*misure di gestione dei rischi opportune e mirate*" (lett. d), che per il par. 5 sono quelle per cui "*i pertinenti* rischi residui *associati a ciascun pericolo nonché il* rischio residuo complessivo *dei sistemi sono considerati* accettabili" (evidenziazioni aggiunte).

Il legislatore europeo riconosce quindi, espressamente, un'area di "rischi consentiti" o "*accettabili*"; e l'adeguatezza delle misure che devono (se non eliminarli) "*ridurli*", è da commisurare a "*le conoscenze tecniche, l'esperienza, l'istruzione e la formazione che ci si può aspettare dal deployer e* [ad] *il contesto presumibile in cui il sistema è destinato ad essere usato*" (art. 9, par. 5, comma 3). Fra le numerose altre prescrizioni da seguire nella gestione dei rischi[29], vanno segnalate quelle per cui tali sistemi devono essere "*sottoposti a prova al fine di individuare le misure di gestione dei rischi più appropriate e mirate*" garantendo che "*funzionino in modo coerente per la finalità prevista e che siano conformi ai requisiti*" stabiliti dal regolamento stesso (par. 6).

Di grande rilievo è, infine, l'"*obbligo di sorveglianza umana*" (art. 14), sui cui presupposti e limiti si tornerà a proposito delle posizioni di garanzia (*infra* par. 3.4).

Questa fitta rete di regole extrapenali, da applicare ai sistemi "*ad alto rischio*" – che come detto sono quelli più importanti da considerare (cfr. *supra*, nota 16), non solo per il gran numero riconducibile a tale categoria, ma soprattutto ai fini dell'imputazione della responsabilità penale, coinvolgendo per definizione i più rilevanti beni giuridici e diritti fondamentali da proteggere – si indirizza ai diversi "*operatori*" coinvolti nella loro catena di vita, secondo i diversi ruoli da essi svolti. Regole che costituiscono, quindi, una

---

[28] Cfr. il par. 3, punto 12.b delle Raccomandazioni di cui alla *Risolution,* cit., p. 56, e par. 4, punto 17, *ivi*, p. 57.

[29] Ad esempio relative alla documentazione tecnica e alla conservazione delle registrazioni (artt. 11 e 12), oltre ad altre più specifiche per i fornitori (artt. da 16 a 22), per gli importatori (art. 23) per i distributori (art. 24) e per i diployer (artt. 26 e 27).

solida base positiva, non solo per circoscrivere il perimetro del rischio consentito (od "*accettabile*") nella loro implementazione e nel loro uso, ma anche - per converso – per desumere le concrete regole cautelari da osservare da parte dei rispettivi destinatari, la cui violazione può essere fonte di responsabilità penale "*per colpa*", una volta accertato il nesso di causalità (o, meglio, di imputazione oggettiva) con il fatto o l'evento offensivo, che pur essendo realizzato, nel suo ultimo anello, dalla "decisione" finale o dal comportamento (*output*) del sistema di intelligenza artificiale, rientri nel nesso di rischio come sopra individuabile[30].

In ogni caso, alla misura oggettiva della colpa, così ricostruita, deve aggiungersi il momento soggettivo di riprovazione personale, che in conformità al principio di colpevolezza deve basarsi sulla concreta possibilità di agire diversamente, da parte del soggetto umano, in conformità ai precetti cautelari come sopra ricostruiti.

A tale riguardo, la risoluzione approvata a Siracusa ha evidenziato che non è necessario che il soggetto umano preveda in concreto l'evento od il fatto effettivamente realizzatosi, essendo sufficiente, perché si configuri la sua responsabilità per colpa, muovendo dai principi generali, la "prevedibilità", cioè la mera "possibilità di prevedere" che l'*out put* prodotto dal sistema di intelligenza artificiale rientri nella correlata area di rischio. Prevedibilità che può, dunque, essere estesa – in sintonia con il modello della "colpa di organizzazione", su cui si tornerà nel par. 3.5., relativo alla responsabilità delle persone giuridiche – ai fatti od eventi riconducibili alla tipologia di quelli oggetto della doverosa valutazione *ex ante* dei rischi correlati al sistema d'intelligenza artificiale, per cui deve includere anche la doverosa consapevolezza di *output* singolarmente "imprevedibili", per evitare i quali il sistema di gestione dei rischi e, più in specifico, le misure anche organizzative di prevenzione e contenimento da rispettare sono state poste e sono dunque individuabili, quali regole cautelari che già scontano la "prevedibilità" di tali categorie di fatti od eventi[31].

3.4. Sulle posizioni di garanzia.

Le richiamate disposizioni del regolamento europeo possono contribuire a definire più concretamente anche le posizioni di garanzia penalmente rilevanti, delle quali la risoluzione adottata a Siracusa ha sottolineato l'importanza, per responsabilizzare penalmente soggetti chiamati ad impedire fatti od eventi costituenti reato, in forza delle funzioni da essi svolte, in specie "di controllo" sulle diverse fasi della catena che va dalla progettazione e sviluppo, alla produzione, distribuzione, messa a

[30] Cfr. il par. 4 delle Raccomandazioni di cui alla *Risolution,* cit., p. 57 s., in cui si evidenzia la necessità di "*adattamento dei criteri di attribuzione della responsabilità penale alle caratteristiche dei sistemi di intelligenza artificiale ed, in particolare, al loro grado di autonomia*"; ed ivi, in specie, i punti 13, 14, 15, 17, 18 e 19.a, secondo periodo, p. 57-58.

[31] Cfr. il punto 19.a del par. 4 (in specie seconda parte) delle Raccomandazioni di cui alla *Risolution,* cit., p. 58.

disposizione ed infine utilizzazione di tali sistemi, in conformità a chiari obblighi giuridici "*di natura tecnica, organizzativa e di controllo*"[32].

Come già esposto, non solo l'art. 9 del regolamento richiede un "*sistema di gestione dei rischi*" che si articola in una dettagliata serie di obblighi e prescrizioni, facenti capo ai vari operatori coinvolti nelle diverse fasi, i quali devono garantire l'adozione ed aggiornamento costante di specifiche misure di prevenzione ed intervento dirette a ridurli, se non eliminarli.

Di grande rilievo è soprattutto il menzionato art. 14, che impone un "*obbligo di sorveglianza umana*", fin dalla fase di progettazione e sviluppo dei sistemi, da realizzare "*anche con strumenti di interfaccia uomo-macchina adeguati*", in modo tale che essi possano "*essere efficacemente supervisionati da persone fisiche durante il periodo in cui sono in uso*" (par. 1). La sorveglianza umana non deve né può certamente entrare nel contenuto di ogni singola fase e modalità operativa dei sistemi di intelligenza artificiale, perché ne intralcerebbe od impedirebbe il funzionamento, contraddicendo l'utilità della loro autonomia. Piuttosto deve mirare a "*prevenire o ridurre al minimo i rischi per la salute, la sicurezza o i diritti fondamentali che possono emergere quando un sistema di intelligenza artificiale ad alto rischio è utilizzato conformemente alla sua finalità prevista o in condizioni di uso improprio ragionevolmente prevedibile, in particolare qualora tali rischi persistano nonostante l'applicazione di altri requisiti di cui alla presente sezione*" (par. 2 del citato art. 14 del regolamento).

Ebbene, i soggetti destinatari di tali obblighi, che – secondo i criteri espressi nelle raccomandazioni approvate nel Colloquio di Siracusa - li rendono "*garanti*" anche ai fini penali, con corrispondenti doveri d'impedimento dei reati in cui si possano concretizzare i predetti rischi, avendo una espressa posizione "di controllo" su di essi, si possono considerare espressamente individuati dal citato art. 14, che oltre alle "*persone fisiche alle quali è affidata la sorveglianza umana*" (par. 4) [33], menziona i "*fornitori*" ed i "*disployer*", che possono invece essere anche persone giuridiche, i primi dei quali devono "*individuare e, ove tecnicamente possibile, integrare nel sistema, fin da prima della sua immissione sul mercato o messa in servizio*" le prescritte misure di sorveglianza umana (par. 3, lett. a), mentre i secondi devono in ogni caso "*attuarle*" (par. 3, lett. b)[34].

Dagli obblighi d'identificazione sempre aggiornata e di valutazione adeguata dei rischi derivanti dalle attività dei sistemi di intelligenza artificiale, cui deve conseguire l'adozione, attuazione ed aggiornamento delle corrispondenti misure di contenimento, facenti espressamente capo alle persone umane ed, eventualmente, agli enti che stanno dietro ad essi, discende dunque la configurabilità di posizioni di garanzia penalmente

---

[32] Cfr. il punto 19.a del par. 4 (in specie prima parte) delle Raccomandazioni di cui alla *Risolution,* cit., p. 58.

[33] *Ibidem*. Sulla molteplicità di possibili figure di garanti, si veda il "rapporto speciale" presentato nell'ambito dei lavori della prima Sezione da C. GRANDI, *Positive obligations (Garantestellung) grounding Criminal Responsibility for not having avoided an illegal Result connected to the AI Functioning*, in *RIDP*, 1/2023, p. 67 s., che raccomanda di formalizzarle e sistematizzarle, in relazione alla struttura organizzative che venga in questione, distinguendo preliminarmente fra "garanti tecnici" e "garanti decisionali" (ivi, p. 74 s.)

[34] La norma dettaglia nei paragrafi seguenti, quali siano i compiti di dette "*persone fisiche*" che devono essere di regola almeno due, e devono essere messe, dal disployer, in condizioni di adempierli, sulla base di debite informazioni e misure organizzative.

rilevanti, che possono fondare una responsabilità penale per omesso impedimento di eventi o fatti costitutivi di reati, cagionati nell'ultimo anello dai sistemi stessi, in conformità ai canoni generali elaborati in relazione all'art. 40, capoverso, c.p., da adeguare però alla nuova specifica disciplina normativa per applicarli efficacemente alla realtà tecnologica che abbiamo di fronte.

3.5. Sulla responsabilità da reato delle persone giuridiche.

Le raccomandazioni contenute nella risoluzione approvata a Siracusa pongono in primo piano il modello di responsabilità da reato delle persone giuridiche[35], nei confronti delle quali devono essere previste specifiche sanzioni di natura punitiva, proporzionate ai reati commessi da, attraverso o contro i sistemi di intelligenza artificiale[36].

A tal fine, si suggerisce di introdurre o adeguare il modello di responsabilità "da reato", in modo che quella della persona giuridica non dipenda da una presupposta responsabilità penale di una determinata persona fisica, dovendo piuttosto dipendere direttamente ed autonomamnete dall'operato del sistema di intelligenza artificiale prodotto, sviluppato, fornito, messo a disposizione od utilizzato nell'interesse o a vantaggio della persona giuridica[37].

Tale modello di responsabilità potrebbe sviluppare quel principio di "autonomia" della responsabilità della persona giuridica, rispetto alla responsabilità penale della persona fisica, enunciato nell'art. 8 del nostro d.lgs. 231/2001, e basarsi quindi soltanto sulla "colpa d'organizzazione" dell'ente stesso, con esclusione in ogni caso di una mera responsabilità oggettiva. "Colpa d'organizzazione" ravvisabile

---

[35] Al par. 4, punto 16, delle Raccomandazioni di cui alla *Risolution,* cit., p. 57, si rimanda, seppur genericamente, anche al modello della responsabilità da prodotti difettosi, le cui specifiche regole di imputazione sono intese ad assicurare un maggior grado di protezione della salute e dei beni dei consumatori, come già previsto dalla direttiva 85/354/CEE, che verrà a breve sostituita dalla più incisiva proposta della Commissione europea presentata il 28.9.2022 (2022/232 COD) ed approvata dal Parlamento europeo nel marzo 2024. Con questa nuova direttiva, nel concetto di "*prodotto*" sono espressamente inclusi anche "*i file per la fabbricazione digitale...e il software*" (art. 4, par. 1, *sub* n. 1 della proposta di nuova direttiva). Mentre altri elementi possono essere tratti dalla disciplina della responsabilità per la salute e sicurezza sul luogo di lavoro, parimenti menzionata dalle Raccomandazioni al punto citato, oggetto delle innumerevoli direttive dell'Unione europea (elencate, ad esempio, nella premessa del d.lgs. 9.4.2008, n. 81 e successive modifche), dirette a regolare dettagliatamente anche le prescrizioni e le responsabilità relative all'uso di macchinari e prodotti ad alta tecnologia.

[36] Cfr. in specie il punto 19.b del par. 4 delle Raccomandazioni di cui alla *Risolution,* cit., p. 58; e quanto alle relative sanzioni "*punitive*", il par. 5, punto 20, p. 59, in cui si sottolinea che accanto alle sanzioni pecuniarie, si dovrebbe prevedere l'ingiunzione di modificare il sistema di conformità e controllo interno dell'ente, con possibilità di ordinare un periodo di monitoraggio pubblico, per garantire la conformazione agli standard imposti.

[37] Per un'attenta valutazione delle possibili tecniche normative e degli ostacoli riferibili al riconoscimento di una responsabilità delle persone giuridiche per reati correlati a sistemi di intelligenza artificiale, si veda il "rapporto speciale" presentato nell'ambito dei lavori della prima Sezione da V. MONGILLO, *Corporate Criminal Liability for AI-realted Crimes: Possible Legal Techniques and Obstacles*, in *RIDP*, 1/2023, p. 77 s.

nell'assenza, carenza od inadeguatezza delle misure organizzative e di prevenzione prescritte dalle norme pertinenti (a partire da quelle del regolamento europeo già sopra diffusamente richiamate), che devono essere finalizzate ed idonee a contenere e gestire gli specifici rischi derivanti dai sistemi di intelligenza artificiale, che operano o sono comunque riconducibili all'interesse o vantaggio della stessa persona giuridica.

Al riguardo, è da richiamare la necessità di efficaci meccanismi di vigilanza, che garantiscano l'implementazione e l'aggiornamento costante di dette misure, eventualmente supportati - a loro volta - da appositi sistemi di intelligenza artificiale[38], ferma la sorveglianza umana imposta nei termini sopra esaminati dal regolamento europeo.

Negli ordinamenti giuridici nazionali che prevedono che la responsabilità delle persone giuridiche sia limitata ad una lista chiusa di reati, come accade in Italia, la raccomandazione formulata a Siracusa è che tale elenco sia esteso a tutti i reati, anche colposi, che possano essere commessi per mezzo di, attraverso o contro sistemi di intelligenza artificiale[39].

Si aggiunga che il regolamento europeo, come detto, all'art. 99 introduce un ampio sistema di sanzioni amministrative pecuniarie, per vero anche molto incisive[40], applicabili direttamente alle persone giuridiche (che potrebbero essere addirittura maggiormente afflittive rispetto a quelle penali), per la violazione di specifici precetti del regolamento stesso e con riferimento ai diversi operatori e soggetti, siano persone fisiche o giuridiche, che ne sono destinatari. Per cui si può già parlare di un modello di responsabilità autonoma delle persone giuridiche per l'operato dei sistemi di intelligenza artificale.

3.6. Sulla formulazione di nuove fattispecie di reato.

Per individuare eventualmente nuove fattispecie di reato, che siano specificamente riferite all'utilizzo o coinvolgimento di sistemi di intelligenza artificiale, come suggerito da alcune risposte al questionario, e come sta emergendo in taluni ordinamenti[41], compresa oggi l'Italia[42], il legislatore europeo o nazionale potrebbe utilizzare un criterio di "analogia" rispetto a reati comuni esistenti, come già è avvenuto

---

[38] Un sistema di intelligenza artificiale è per ciò stesso considerato "*ad alto rischio*" se è "*destinato a essere utilizzato come componente di sicurezza di un prodotto*" ed è "*soggetto a una valutazione della conformità da parte di terzi ai fini dell'immissione sul mercato o della messa in servizio*" (cfr. art. 6, par. 1, lettere a) e b) del regolamento).

[39] Cfr. il punto 19.b del par. 4 delle Raccomandazioni di cui alla *Risolution,* cit., p. 59.

[40] Cfr. *supra,* par. 3.3.2.1 e nota 24.

[41] Per alcuni riferimenti rinvio a L. PICOTTI, *Traditional Criminal Law Categories* (*General report*), cit., in specie p. 23 s.

[42] Si veda il Capo V, intitolato "Disposizioni penali", contenente l'art. 25 del disegno di legge governativo presentato il 14.4.2024, che prevede – oltre all'introduzione di molteplici circostanze aggravanti per reati commessi coinvolgendo sistemi di intelligenza artificiale – anche l'introduzione di un nuovo delitto rubricato: "Art. 612-*quater* c.p. - *Illecita diffusione di contenuti generati o manipolati con sistemi di intelligenza artificiale*", peraltro ancora incerto nella sua formulazione, presentata con due possibili testi alternativi.

per contrastare la criminalità informatica (si pensi alla frode informatica, affiancata alla truffa comune; od alla falsità informatica, aggiunta alla falsità documentale). Ma sulle possibili riforme di parte speciale, fermo quanto già esposto circa l'opportunità di introdurre specifici reati preparatori (cfr. *supra*, par. 3.3.2.1), rimando ai lavori della seconda Sezione ed alla relativa risoluzione, approvata nel Colloquio internazionale di Bucarest del 14-16.6.2023, per cui è stato relatore generale l'amico prof. Fernando Miró-Llinares[43], che interverrà nel prossimo panel.

3.7. Misure complementari e civili.

Infine, occorre evidenziare che il paragrafo 6 delle raccomandazioni approvate a Siracusa, muovendo dal principio di *ultima ratio* da rispettare nel ricorso alle sanzioni penali, ed evidenziando le difficoltà (o ritardi) che possono riscontrarsi nell'implementare un efficace sistema di responsabilità penale nei confronti delle persone fisiche e giuridiche che "stanno dietro" i sistemi di intelligenza artificiale, suggerisce di ricorrere anche a misure complementari di *enforcement*, oltre che a rimedi civilistici[44]: quali sistemi di autorizzazioni e di certificazioni amministrative, modelli di *compliance*, interventi di giustizia riparativa, accordi con vittime e con pubbliche autorità.

4. Osservazioni conclusive.

Il valore del lavoro della nostra Associazione internazionale di penalisti, che si confronta da oltre un secolo con i fenomeni in forte evoluzione della società in cui viviamo, è data dalla capacità di rispondere alle sfide che essa pone, stimolando l'evoluzione e l'adeguamento, di pari passo, della legislazione, della giurisprudenza, della dottrina, al fine di garantire la razionalità e l'efficienza del sistema penale, che deve continuare ad adempiere nel modo migliore alla sua funzione fondamentale di tutela dei beni giuridici e dei diritti fondamentali della persona e della collettività, offesi da reati che si manifestano in sempre nuove forme e modalità.

La loro repressione e, per quanto possibile, prevenzione deve però coniugare sempre la necessità di cambiamento e di adattamento, con quella imperativa di salvaguardia dei principi basilari del diritto penale, che sono garanzie inalienabili dello Stato di diritto.

Come è emerso dai lavori congressuali, per far fronte alle sfide poste dalle nuove conquiste tecnologiche, rappresentate dall'introduzione e sempre più estesa applicazione dei sistemi di intelligenza artificiale, che determinano corrispondenti rivolgimenti nei rapporti sociali, economici, politici ed interpersonali, è apparso però

---

[43] Il rapporto generale, il testo della risoluzione approvata ed una selezione dei più significativi rapporti nazionali si possono leggere nel fascicolo monografico a cura di F. MIRÓ-LLINARES, C. DUVAC, T. TOADER, M.S. GALARZA, *Criminalisation of AI-related Offences*, in *RIDP*, 1/2024.

[44] Cfr. par. 6, punti 23 e 24 Raccomandazioni di cui alla *Risolution*, cit., p. 59.

necessario suggerire una riconsiderazione di alcune categorie dogmatiche tradizionali, in specie quelle della causalità, della colpa, delle posizioni di garanzia, della responsabilità da reato delle persone giuridiche, per rinnovarne i contenuti.

Il lavoro non può certamente dirsi compiuto, essendo stato anzi da poco intrapreso, ed occorrendo un tempo adeguato di ricerche e di confronto, per l'adattamento ponderato e graduale dei concetti tradizionali, alla luce di quanto emergerà anche dalla prassi applicativa e dall'implementazione delle nuove norme europee appena emanate.

Dovrà di certo svilupparsi ulteriormente un'elaborazione teorica capace di rinnovarli, pur mantenendone e riconoscendone il valore dogmatico e la funzione sistematica, per rispondere nel modo migliore alla nuova realtà da regolamentare, affinché non sia essa che finisca per sopraffarci.