

SP

SISTEMA
PENALE

FASCICOLO

12/2024

COMITATO EDITORIALE Giuseppe Amarelli, Roberto Bartoli, Hervè Belluta, Michele Caianiello, Massimo Ceresa-Gastaldo, Adolfo Ceretti, Cristiano Cupelli, Francesco D'Alessandro, Angela Della Bella, Gian Paolo Demuro, Emilio Dolcini, Novella Galantini, Mitja Gialuz, Glauco Giostra, Antonio Gullo, Stefano Manacorda, Vittorio Manes, Luca Masera, Anna Maria Maugeri, Melissa Miedico, Vincenzo Mongillo, Francesco Mucciarelli, Claudia Pecorella, Marco Pelissero, Lucia Riscato, Marco Scoletta, Carlo Sotis, Costantino Visconti.

COMITATO SCIENTIFICO (REVISORI) Andrea Abbagnano Trione, Alberto Alessandri, Silvia Allegrezza, Chiara Amalfitano, Enrico Mario Ambrosetti, Ennio Amodio, Gastone Andreazza, Ercole Aprile, Giuliano Balbi, Marta Bargis, Enrico Basile, Fabio Basile, Alessandra Bassi, Teresa Bene, Carlo Benussi, Alessandro Bernardi, Marta Bertolino, Francesca Biondi, Rocco Blaiotta, Manfredi Bontempelli, Carlo Bonzano, Matilde Brancaccio, Carlo Bray, Renato Bricchetti, David Brunelli, Carlo Brusco, Silvia Buzzelli, Alberto Cadoppi, Lucio Camaldo, Gaia Caneschi, Stefano Canestrari, Giovanni Canzio, Francesco Caprioli, Matteo Caputo, Fabio Salvatore Cassibba, Donato Castronuovo, Elena Maria Catalano, Mauro Catenacci, Antonio Cavaliere, Francesco Centonze, Federico Consulich, Carlotta Conti, Stefano Corbetta, Roberto Cornelli, Fabrizio D'Arcangelo, Marcello Daniele, Gaetano De Amicis, Cristina De Maglie, Alberto De Vita, Jacopo Della Torre, Ombretta Di Giovine, Gabriella Di Paolo, Giandomenico Dodaro, Massimo Donini, Salvatore Dovere, Tomaso Emilio Epidendio, Luciano Eusebi, Riccardo Ferrante, Giovanni Fiandaca, Giorgio Fidelbo, Stefano Finocchiario, Carlo Fiorio, Roberto Flor, Luigi Foffani, Désirée Fondaroli, Gabriele Fornasari, Gabrio Forti, Piero Gaeta, Alessandra Galluccio, Marco Gambardella, Alberto Gargani, Loredana Garlati, Giovanni Grasso, Giulio Illuminati, Gaetano Insolera, Roberto E. Kistoris, Giorgio Lattanzi, Sergio Lorusso, Ernesto Lupo, Raffaello Magi, Vincenzo Maiello, Adelmo Manna, Grazia Mannozi, Marco Mantovani, Luca Marafioti, Enrico Marzaduri, Maria Novella Masullo, Oliviero Mazza, Francesco Mazzacuva, Claudia Mazzucato, Alessandro Melchionda, Chantal Meloni, Vincenzo Militello, Andrea Montagni, Gaetana Morgante, Lorenzo Natali, Renzo Orlandi, Luigi Orsi, Francesco Palazzo, Carlo Enrico Paliero, Lucia Parlato, Annamaria Peccioli, Chiara Perini, Lorenzo Picotti, Carlo Piergallini, Paolo Pisa, Luca Pistorelli, Daniele Piva, Oreste Pollicino, Domenico Pulitanò, Serena Quattrocchio, Tommaso Rafaraci, Paolo Renon, Maurizio Romanelli, Bartolomeo Romano, Gioacchino Romeo, Alessandra Rossi, Carlo Ruga Riva, Francesca Ruggieri, Elisa Scaroina, Laura Scomparin, Nicola Selvaggi, Sergio Seminara, Paola Severino, Rosaria Sicurella, Piero Silvestri, Fabrizio Siracusano, Nicola Triggiani, Andrea Francesco Tripodi, Giulio Ubertis, Maria Chiara Ubiali, Antonio Vallini, Gianluca Varraso, Vito Velluzzi, Paolo Veneziani, Francesco Viganò, Daniela Vigoni, Francesco Zaccà, Stefano Zirulia.

REDAZIONE Francesco Lazzeri, Giulia Mentasti (coordinatori), Enrico Andolfatto, Silvia Bernardi, Pietro Chiaraviglio, Beatrice Fragasso, Cecilia Pagella, Emmanuele Penco, Gabriele Pontepino, Tommaso Trinchera.

Sistema penale (SP) è una rivista *online*, aggiornata quotidianamente e fascicolata mensilmente, ad accesso libero, pubblicata dal 18 novembre 2019.

La *Rivista*, realizzata con la collaborazione scientifica dell'Università degli Studi di Milano e dell'Università Bocconi di Milano, è edita da Progetto giustizia penale, associazione senza fine di lucro con sede presso il Dipartimento di Scienze Giuridiche "C. Beccaria" dell'Università degli Studi di Milano, dove pure hanno sede la direzione e la redazione centrale. Tutte le collaborazioni organizzative ed editoriali sono a titolo gratuito e agli autori non sono imposti costi di elaborazione e pubblicazione.

La *Rivista* si uniforma agli standard internazionali definiti dal *Committee on Publication Ethics* (COPE) e fa proprie le relative linee guida.

I materiali pubblicati su *Sistema Penale* sono oggetto di licenza CC BY-NC-ND 4.00 International. Il lettore può riprodurli e condividerli, in tutto o in parte, con ogni mezzo di comunicazione e segnalazione anche tramite collegamento ipertestuale, con qualsiasi mezzo, supporto e formato, per qualsiasi scopo lecito e non commerciale, conservando l'indicazione del nome dell'autore, del titolo del contributo, della fonte, del logo e del formato grafico originale (salvo le modifiche tecnicamente indispensabili). La licenza è consultabile su <https://creativecommons.org/licenses/by-nc-nd/4.0/>.

Peer review I contributi che la direzione ritiene di destinare alla sezione "Articoli" del fascicolo mensile sono inviati a un revisore, individuato secondo criteri di rotazione tra i membri del Comitato scientifico, composto da esperti esterni alla direzione e al comitato editoriale. La scelta del revisore è effettuata garantendo l'assenza di conflitti di interesse. I contributi sono inviati ai revisori in forma anonima. La direzione, tramite la redazione, comunica all'autore l'esito della valutazione, garantendo l'anonimato dei revisori. Se la valutazione è positiva, il contributo è pubblicato. Se il revisore raccomanda modifiche, il contributo è pubblicato previa revisione dell'autore, in base ai commenti ricevuti, e verifica del loro accoglimento da parte della direzione. Il contributo non è pubblicato se il revisore esprime parere negativo alla pubblicazione. La direzione si riserva la facoltà di pubblicare nella sezione "Altri contributi" una selezione di contributi diversi dagli articoli, non previamente sottoposti alla procedura di *peer review*. Di ciò è data notizia nella prima pagina della relativa sezione.

Di tutte le operazioni compiute nella procedura di *peer review* è conservata idonea documentazione presso la redazione.

Modalità di citazione Per la citazione dei contributi presenti nei fascicoli di *Sistema penale*, si consiglia di utilizzare la forma di seguito esemplificata: N. COGNOME, *Titolo del contributo*, in *Sist. pen.* (o *SP*), 1/2023, p. 5 ss.

INDICE DEI CONTRIBUTI

PAPER

E. CARLONI, <i>Conflitto di interessi e abuso di ufficio: problemi aperti tra amministrazione e costituzione</i>	5
E. RECCIA, <i>Populismo e diritto penale del nemico nella legislazione del migrante</i>	21
O. LOMBARDI, <i>Responsabilità penale dell'uomo per il danno cagionato attraverso condotte dolose e colpose nell'impiego dei sistemi di intelligenza artificiale</i>	57

SCHUDE

G.L. GATTA, <i>"Asfissia applicativa" e "ineffettività" del riformato reato di traffico di influenze illecite: la Procura di Roma chiede di sollevare questione di legittimità costituzionale (processo sulle mascherine – covid)</i>	75
S. ZIRULIA, <i>Il "decreto Piantedosi" sulle operazioni di soccorso in mare approda alla Consulta</i>	79
M.C. UBIALI, <i>Concorso pubblico truccato anticipando i temi delle prove: non potendo più ricorrere alla turbativa d'asta e all'abuso d'ufficio, il Tribunale di Milano condanna per rivelazione di segreti d'ufficio</i>	83
F. VITARELLI, <i>Lo sfruttamento del lavoro intellettuale è fuori dal tipo descritto dall'art. 603-bis c.p.?</i>	89
C. MISTRORIGO, <i>Una fotografia delle indegne condizioni presenti nelle carceri italiane in una ordinanza del Magistrato di Sorveglianza di Firenze a seguito di reclamo ex art. 35 bis o.p.</i>	99
C. PASINI, <i>L'impatto del c.d. pacchetto sicurezza sulle persone straniere in Italia e sul fenomeno dell'immigrazione</i>	107

RESPONSABILITÀ PENALE DELL'UOMO PER IL DANNO CAGIONATO ATTRAVERSO CONDOTTE DOLOSE E COLPOSE NELL'IMPIEGO DEI SISTEMI DI INTELLIGENZA ARTIFICIALE

di Oriana Lombardi

Il tentativo del legislatore europeo di positivizzare la nozione di Artificial Intelligence nell'A.I. Act ha il pregio di fornire una nozione in nuce di intelligenza artificiale dalla quale gli operatori del diritto possono e devono iniziare a trarne delle conseguenze. I sistemi di intelligenza artificiale (id est i loro meccanismi di funzionamento) costituiscono una rivoluzione copernicana che chiama tutti gli studiosi ad interrogarsi. Sovviene la considerazione che l'i.a., come ogni strumento, potrebbe essere utilizzata non solo a vantaggio della collettività, ma al contrario per realizzare finalità criminali. Il penalista è spinto a domandarsi in che modo l'uso di sistemi di intelligenza artificiale possa essere ricondotto nel sistema del diritto penale, nel rispetto del principio di legalità e dei suoi corollari e in armonia con le categorie tradizionali della materia. In particolare, il suo utilizzo potrebbe creare incertezze in termini di imputabilità, in relazione al nesso di causalità (che deve legare l'azione/omissione e la relativa modificazione della realtà materiale) nonché rispetto alla nozione di colpa e di negligenza. Di qui, questo articolo si ripropone di sviluppare i seguenti interrogativi: in primo luogo, sul versante delle condotte assistite da dolo, come viene punito un soggetto che abbia realizzato un reato mediante l'aiuto di sistemi di i.a.? Ancora, si possono imputare ad un soggetto che abbia agito con l'ausilio dell'i.a. tutte le conseguenze dannose derivanti dalla collaborazione uomo-macchina? In secondo luogo, sul versante delle condotte colpose, come impatta la collaborazione uomo-macchina sull'imputabilità e sul nesso causale e quali sono gli obblighi incombenti sull'uomo? Vedremo che anche nel diritto penale si rende necessario un bilanciamento tra i benefici dello sviluppo tecnologico e la tutela dei diritti fondamentali della persona, il cui sacrificio sull'altare del progresso non può essere giustificato.

SOMMARIO: 1. Introduzione. – 2. Condotte dolose. – 2.1. Elemento materiale: riconducibilità delle condotte realizzate con l'intelligenza artificiale a reati preesistenti. – 2.2. Elemento soggettivo: qual è l'oggetto del dolo? – 3. Condotte colpose. – 3.1. Premessa. – 3.2. Gli obblighi positivi introdotti dall'A.I. Act. – 3.4. Obbligo di *human oversight*. – 4. L'accertamento causale. – 5. Un possibile scenario: fra colpa dell'utilizzatore e *responsibility gap*. – 6. Considerazioni conclusive.

1. Introduzione.

L’A.I. Act adottato il 13 giugno 2024 dal Parlamento europeo e dal Consiglio reca la prima definizione vincolante di intelligenza artificiale¹. In buona sostanza, con la locuzione intelligenza artificiale (d’ora in poi i.a.) ci si riferisce alla capacità di una macchina di operare, quasi di ragionare, in modo non dissimile dalla mente umana elaborando grandi quantità di dati sulla base di istruzioni fornite da algoritmi capaci di aggiornarsi in modo automatico, cd. *machine learning*². Tali sistemi sono impiegati ormai da decenni in numerosi settori per delegare alle macchine compiti specifici affinché vengano svolti più velocemente e in modo più accurato (auspicabilmente) rispetto a quanto l’uomo potrebbe fare, consentendogli nel frattempo di concentrarsi su quanto invece non è delegabile³. Del resto, quando si parla di progresso l’obiettivo è sempre lo stesso: permettere all’uomo una vita più agevole. Tuttavia, ogni attività umana, specialmente se svolta con l’ausilio di tecnologie avanzate e come tali molto spesso poco intellegibili, espone la collettività a dei rischi più o meno elevati rendendoci tutti beneficiari di un’evoluzione che ci mette inevitabilmente in pericolo. Per le attività pericolose, ma consentite e dunque regolamentate poiché socialmente utili, il legislatore e gli amministratori sono sempre alla ricerca di una normazione che individui un punto di equilibrio fra massimizzazione dei vantaggi e minimizzazione dei rischi. A fronte di un rischio incerto, questo bilanciamento avviene tradizionalmente in base al principio di precauzione in virtù del quale la fissazione del livello di protezione prescelto incombe sui responsabili politici e deve tener conto dei valori che prevalgono in quella società e in un determinato momento storico⁴.

¹ Ai sensi dell’art. 3, co.1, 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) 2019/2144 e le direttive 2014/90/UE, (UE) 2016/797 e (UE) 2020/1828 (regolamento sull’intelligenza artificiale) per sistema di intelligenza artificiale si intende «un sistema automatizzato progettato per funzionare con livelli di autonomia variabili e che può presentare adattabilità dopo la diffusione e che, per obiettivi espliciti o impliciti, deduce dall’input che riceve come generare output quali previsioni, contenuti, raccomandazioni o decisioni che possono influenzare ambienti fisici o virtuali».

² Secondo la Commissione europea: «Machine learning (ML) is a subset of AI, where algorithms are trained to infer certain patterns based on a set of data in order to determine the actions needed to achieve a given goal» (Si veda il Libro Bianco “Libro bianco sull’intelligenza artificiale – Un approccio europeo all’eccellenza e alla fiducia”), Commissione europea del 19 febbraio 2020. Per un ulteriore approfondimento sul tema, si vedano S. RUSSELL - P. NORVIG (a cura di), *Artificial Intelligence: A Modern Approach*, Prentice Hall, IV ed., New York, 2020, pp. 634 ss.; e H. SURDEN, *Machine Learning and Law*, in *Washington Law Review*, vol. 89, n. 1, 2014, pp.87 ss.

³ Quanto ai settori in cui ha trovato larga applicazione, si pensi al campo delle *self-driving cars*, in cui l’impiego dell’i.a. potrebbe ridurre notevolmente il rischio di sinistri o al suo impiego in ambito medico, dove minimizzato il rischio di *bias* si potrebbe effettuare uno screening oncologico più accurato grazie alla diagnostica per immagini dotata sistemi di i.a., utilizzare tecniche di prevenzione di malattie cardiovascolari e ricorrere ai cd. *surgical robotics* in ambito chirurgico. Su quest’ultimo aspetto, si veda E. SANTORO, *Intelligenza artificiale in medicina: qual è il suo impatto?*, in Istituto di ricerche farmacologiche Mario Negri IRCCS, 28 giugno 2023, disponibile al seguente link: <https://www.marionegri.it/magazine/intelligenza-artificiale-medicina>.

⁴ Si vedano D. FRANZONE, *Il principio di precauzione in diritto comunitario*, in *Il principio precauzionale nel diritto*

Il legislatore europeo ha compiuto questa scelta graduando gli obblighi incombenti sui soggetti coinvolti nella produzione e nell'utilizzo di sistemi di i.a. proprio in base al livello di rischio che la tecnologia prodotta o impiegata è in grado di ingenerare, introducendo in particolare delle regole specifiche per i cd. sistemi *ad alto rischio*⁵. Tali previsioni, nonostante siano state introdotte con la finalità di armonizzare la materia nei diversi Stati membri a tutela del mercato interno, devono essere prese in considerazione dal penalista che è chiamato a misurarsi con la prima regolamentazione della materia con efficacia giuridica vincolante⁶.

Segnatamente, quanto alla definizione di i.a. recata dall'A.I. *Act*, ci si domanda se a questa possa essere attribuito un ruolo descrittivo all'interno delle preesistenti fattispecie incriminatrici tenuto conto del divieto di analogia *in malam partem*. Inoltre, ci si chiede se la violazione degli obblighi introdotti dal regolamento per i cd. sistemi *ad alto rischio* possa integrare l'elemento normativo della colpa. E ancora, si cerca di capire come far fronte ai problemi che la collaborazione uomo-macchina è in grado di porre in relazione all'accertamento della causalità materiale.

Pertanto, in attesa che il legislatore interno decida eventualmente di innovare l'attuale quadro giuridico di riferimento onde evitare possibili vuoti di tutela, questo contributo intende interrogarsi sulla qualificazione dei fatti storici che, nell'imminenza, potrebbero essere commessi mediante l'utilizzo di sistemi di i.a.

In questa analisi verranno presi in considerazione, separatamente, i fatti storici realizzati con dolo e quelli realizzati con colpa dal momento che, in base al diverso elemento soggettivo, si pongono interrogativi differenti e dotati di un diverso grado di problematicità, più spiccato nel caso delle condotte colpose.

internazionale e comunitario, a cura di A. BIANCHI - M. GESTRI, Giuffrè, Milano, 2006, pp. 8 ss.; A. MASSARO, [Principio di precauzione e diritto penale: nihil novi sub sole?](#), in *Dir. Pen. Con.*, 9 maggio 2011; F. GIUNTA, *Il diritto penale e le suggestioni del principio di precauzione*, in *disCrimen*, 19 dicembre 2018, consultabile al seguente link <https://discrimen.it/il-diritto-penale-e-le-suggestioni-del-principio-di-precauzione/>; G. FORTI, "Accesso" alle informazioni sul rischio e responsabilità: una lettura del principio di precauzione, in *disCrimen*, 18 dicembre 2018, p. 163. Per un inquadramento sociologico, si veda per tutti U. BECK, *La società del rischio. Verso una seconda modernità*, Carocci editore, Roma, 2013. Infine, sempre sul punto, si consideri anche la comunicazione della Commissione del 2 febbraio 2000, sul principio di precauzione, (COM 2000) 1 final.

⁵ Per le regole di classificazione dei sistemi ad alto rischio, si veda l'art. 6 dell'A.I. *Act*.

⁶ Di ciò se ne ritrae conferma dal fatto che la base giuridica del Regolamento è l'art. 114 TFUE (v. considerando 3), anche se con tale atto il legislatore europeo si ripropone di tutelare contestualmente altri interessi pubblici quali la salute, la sicurezza e la protezione dei diritti fondamentali come riconosciuti e tutelati dal diritto dell'Unione.

2. Condotte dolose.

2.1. Elemento materiale: riconducibilità delle condotte realizzate con l'intelligenza artificiale a reati preesistenti

Gli *Artificial Intelligence Crimes*, vale a dire le condotte criminali realizzate con dolo attraverso l'utilizzo di sistemi di i.a., non appaiono di difficile qualificazione nel quadro giuridico di riferimento potendo, in linea di massima, essere agevolmente ricondotte alle fattispecie incriminatrici attualmente esistenti.

La questione può essere affrontata analizzando partitamente le ipotesi in cui il fatto posto in essere con l'uso di sistemi di i.a. sia riconducibile ad un reato a forma libera oppure ad uno a forma vincolata.

Nel primo caso, con i reati a forma libera il legislatore ha attribuito rilevanza ad ogni comportamento umano che sia causa, con qualsiasi modalità, di un determinato evento⁷. Si pensi esemplarmente all'uccisione di una persona tramite un drone intelligente capace di ricercarla e riconoscerla⁸. In un caso del genere, verificata la consapevole causazione dell'evento da parte del soggetto agente, l'*human-behind-the-machine* che ha dato la fatale istruzione al drone risponde di omicidio volontario ai sensi dell'art. 575 c.p.

Diversamente, nei reati a forma vincolata le modalità della condotta sono descritte in modo dettagliato perché l'ordinamento ha scelto di proteggere il bene giuridico da alcune classi di condotte, e non da altre. L'azione concreta sarà penalmente rilevante, cioè tipica, solo se corrispondente allo specifico modello di comportamento descritto dalla norma. Rispetto a questa categoria di reati, data l'introduzione della definizione di i.a. ad opera dell'*A.I. Act*, l'interprete potrebbe domandarsi se sia necessario inserire nella descrizione della condotta un riferimento esplicito alle nuove tecnologie onde evitare frizioni con il principio di legalità, e in modo particolare con i suoi corollari del principio di tassatività e con il divieto di interpretazione analogica *in malam partem*. «Si pensi esemplarmente ad un abuso di mercato realizzato tramite algoritmi ad alta frequenza⁹ impostati per sfruttare al meglio un rialzo di valori,

⁷ Per un approfondimento sulla distinzione fra reati a forma libera e reati a forma vincolata, si veda, per tutti G. MARINUCCI - DOLCINI - G.L. GATTA, *Manuale di Diritto Penale, Parte generale*, IX ed., Milano, Giuffrè, pp. 234 e ss.

⁸ L. PICOTTI, *Categorie tradizionali del diritto penale e intelligenza artificiale: crisi o palingenesi?*, in questa Rivista, 31 luglio 2024.

⁹ «Per attività di *trading* ad alta frequenza, o *high frequency trading*, si intende una modalità operativa basata sull'impiego di algoritmi che consentono di acquisire, elaborare e reagire alle informazioni di mercato con una velocità elevata», così V. CAIVANO - S. CICCARELLI - G. DI STEFANO - M. FRATINI - G. GASPARRI - M. GILIBERTI - N. LINCiano - I. TAROLA, *Il trading ad alta frequenza, Caratteristiche, effetti, questioni di policy*, in *Discussion papers*, CONSOB, 5 dicembre 2012. Per dare una misura della diffusione dell'uso del *trading* ad alta frequenza, si tenga in considerazione che in base a studi condotti dall'ESMA, dal 2018 è stato evidenziato un incremento del *trading* algoritmico con riferimento al mercato azionario europeo del 50/70%. A tal riguardo, si veda ESMA, *European Securities and Markets Authority, Consultation Paper.MiFID II/MiFIR, review report on Algorithmic Trading*, 18 dicembre 2020.

previamente indotto tramite una moltiplicazione artificiosa di domande d'acquisto»¹⁰. Il reato di manipolazione del mercato previsto dall'art. 185 del D.Lgs. 58/1998 (TUF) oltre alle condotte di diffusione di notizie false e di realizzazione di operazioni simulate, incrimina anche l'alterazione del mercato realizzata con "altri raggiri" quale modalità ulteriore di realizzazione del reato.¹¹ Con questa locuzione sembra che il legislatore abbia inteso far riferimento ad altre condotte, diverse dalla diffusione di notizie false e dalla realizzazione di operazioni simulate, ma comunque connotate da una certa capacità decettiva e perciò in grado di realizzare il pericolo concreto della sensibile alterazione del prezzo¹². In questo caso, si potrebbe sostenere che l'interprete possa ricondurre la condotta consistita nell'uso di algoritmi ad alta frequenza alla locuzione "altri artifici" rimanendo comunque nel perimetro tracciato dalla fattispecie incriminatrice dal momento che l'espressione avrebbe la funzione di sanzionare delle condotte che "non sono illecite in sé", ma che pur tuttavia si sostanziano in comportamenti manipolativi del mercato¹³.

Un ragionamento simile è spendibile anche in riferimento al settore dei cd. *cybercrimes*¹⁴, rispetto al quale può ragionevolmente presumersi che l'offensività delle condotte sia destinata ad incrementare proprio in virtù dell'impiego di strumenti di i.a., con la messa in pericolo di beni giuridici quali il diritto alla riservatezza dei propri dati e il diritto all'integrità.

La materia è stata innovata di recente dalla legge n. 90 del 28 giugno 2024. Sotto un punto di vista lessicale, giova osservare come al netto di queste modifiche, il riferimento alle tecnologie in essi contenuto avviene con le immutate espressioni quali, per citarne alcune, "apparecchiature, dispositivi, programmi informatici" o "sistemi informatici o telematici"¹⁵ oppure più genericamente "apparecchiature e altri mezzi"¹⁶.

¹⁰ L. PICOTTI, *Categorie tradizionali del diritto penale*, cit.

¹¹ In questo senso, Cass. Pen., Sez. V, 2 ottobre 2008, n. 2063.

¹² Proprio in virtù di questa locuzione il reato di manipolazione del mercato è stato definito da autorevole dottrina come "una truffa in *incertam personam*": «la patrimonialità degli interessi in gioco e la connotazione oggettivamente soggettivamente fraudolenta autorizzano l'immagine di una truffa in *incertam personam*, ove all'anticipazione dell'intervento repressivo (tipica delle tutele a largo raggio) fa riscontro la severità del trattamento punitivo», si veda C. PEDRAZZI, *Società commerciali (disciplina penale)*, in *Digesto delle discipline penalistiche*, XIII ed., Torino, 1998, p. 305.

¹³ M. GAMBARDELLA, *Condotte economiche e responsabilità penale*, II ed., Giappichelli, Torino, pp. 412 e ss. Inoltre, si consideri che il ricorso a questa interpretazione potrebbe non rendersi necessario ove l'art. 185 TUF dovesse essere modificato nel senso prospettato dal d.d.l. n. 1146 sull'intelligenza artificiale il cui art. 25, co. 4, prevede l'introduzione nella lettera dell'art. 185 TUF di un espresso riferimento all'ipotesi in cui la manipolazione del mercato sia commessa mediante l'impiego di sistemi di intelligenza artificiale, qualificando tuttavia l'impiego di i.a. quale circostanza aggravante. Allo stesso modo, l'art. 25, co. 2, del d.d.l. prevede l'introduzione di un'aggravante per il reato di aggrottaggio di cui all'art. 2637 c.c. se il fatto è commesso con l'impiego di i.a.

¹⁴ Con questa espressione si fa riferimento alla commissione di comportamenti illeciti commessi sul web, si pensi esemplarmente alle condotte di *phishing*, attacchi mediante *malware*, diffamazioni e truffe online.

¹⁵ Espressioni che nell'art. 635-*quater*, co. 1, c.p., introdotto dall'art.16, co. 1, lettera q) della L. 28 giugno 2024 n. 90, si riferiscono rispettivamente agli strumenti di realizzazione del reato e all'oggetto materiale.

¹⁶ Espressioni presenti nell'art. 617-*bis* c.p.

Insomma, il legislatore è intervenuto sui reati informatici dopo l'adozione dell'A.I. Act e non ha sentito l'esigenza di introdurre un esplicito riferimento agli strumenti di i.a. Piuttosto, la tendenza sembra essere quella di una tecnica legislativa che privilegia l'uso di formule ampie per sanzionare l'impiego criminale di strumenti informatici genericamente intesi, al di là di specifiche definizioni, le quali talvolta possono rivelarsi un'arma a doppio taglio. Le espressioni conservate dal legislatore sembrano potersi indirizzare, indistintamente, a computer, *software* e alle nuove forme di intelligenza artificiale e chissà, ad ogni altro dispositivo *hardware* e *software* che l'ingegneria informatica sarà in grado di offrirci nel prossimo futuro. Infatti, proprio con riguardo alle fattispecie che fanno riferimento all'uso di strumenti tecnologici, va da sé che i mezzi cui queste si riferiscono al momento della loro entrata in vigore, con ogni probabilità, verranno superati nei decenni a seguire. Si pensi al noto *revirement* giurisprudenziale sull'applicabilità dell'art. 57 c.p. al direttore del periodico *online*, occasione in cui le Sezioni Unite hanno constatato la necessità di interpretare estensivamente il termine "stampa" di cui all'art. 1 della l. n. 47 del 1948, in modo da ricomprendervi anche la testata giornalistica telematica, abbandonando un'interpretazione legata alle tecnologie dell'epoca al fine di evitare disparità di trattamento fra condotte egualmente meritevoli di pena¹⁷.

Di conseguenza, non appare difficile ritenere che una condotta criminale tradizionalmente realizzata attraverso un computer o un *software*, oggi continui ad essere tipica ove commessa con l'i.a. in virtù di un'interpretazione evolutiva della norma che suggerisce all'operatore del diritto di interpretare una regola alla luce del mutamento sociale¹⁸.

Del resto, anche l'art. 1, lettera a) della Convenzione in materia di *Cybercrime* include nella definizione di "*computer system*" un sistema capace di processare dati in via autonoma¹⁹. Dunque, nient'altro che *old wine in new bottles*²⁰?

A questo approccio, certamente conciliante e risolutivo, si potrebbe obiettare che un conto è, in realtà, parlare di un computer o di un *software*, i quali per quanto all'avanguardia possano essere, restano inequivocabilmente meri strumenti nelle mani degli uomini²¹; un conto è riferirsi all'intelligenza artificiale. È davvero corretto

¹⁷ Cass. Pen., SS.UU., 17 luglio 2015, n. 31022; Cass. Pen., Sez. V, 23 ottobre 2018, n. 1275.

¹⁸ E. CANNIZZARO, *Interpretazione conforme fra tecniche ed effetti normativi*, in *L'interpretazione conforme al diritto dell'Unione europea, Profili e limiti di un vincolo problematico*, a cura di A. BERNARDI, Jovene, Napoli, 2015, precisamente p. 7. Per tutti, si veda, E. BETTI, *Interpretazione della legge e sua efficienza evolutiva*, in *JUS, Rivista di scienze giuridiche*, n. 2, 1959.

¹⁹ « "*Computer system*" means any device or a group of interconnected or related devices, one or more of which, pursuant to a program, performs automatic processing of data».

²⁰ «To be sure, some of the manifestations are new. But a great deal of crime committed with or against computers differs only in terms of the medium. While the technology of implementation, and particularly its efficiency, may be without precedent, the crime is fundamentally familiar», così P. N. GRABOSKY, *Virtual criminality: old wine in new bottle?*, in *SOCIAL & LEGAL STUDIES*, 2001, vol. 10, n. 2, pp. 243-249.

²¹ «Compared to other computer software, ML algorithms do not require explicit instructions from humans but rather extract patterns and learn implicit rules from a considerable number of examples», in *trend micro research, Malicious Use and Abuses of Artificial Intelligence*, 2020, consultabile sul sito <https://www.europol.europa.eu/cms/sites/default/files/documents/ma->

ricondurre i sistemi di i.a., il cui tratto distintivo è l'essere dotati di un variabile grado di autonomia, alla generica nozione di sistema informatico? Oppure sarebbe preferibile tenerne in considerazione le specificità, cioè la concreta eventualità che attraverso il loro impiego si giunga a risultati che vadano al di là delle indicazioni ricevute dall'uomo? Secondo questa diversa ricostruzione, tenere distinti gli *Artificial Intelligence Crimes* dai *Cybercrimes* potrebbe rivelarsi un approccio più cauto. In questo caso, l'incriminazione di un fatto penalmente rilevante realizzato con l'uso di i.a. necessiterebbe dell'*interpositio legislatoris*.

Secondo questa diversa ricostruzione, perciò, la prossima mossa spetterebbe al legislatore il quale, in un'ottica *de iure condendo*, potrebbe compiere nuove scelte di politica criminale. Presumibilmente, questo accadrà in relazione al reato di *deepfake*²². Infatti, il Consiglio dei ministri lo scorso 23 aprile 2024 ha approvato il disegno di legge n. 1146 che prevede l'inserimento di una nuova fattispecie incriminatrice, il nuovo art. 612-*quater* c.p., volto a sanzionare l'«*Illecita diffusione di contenuti generati o alterati con sistemi di intelligenza artificiale*», il reato di *deepfake* per l'appunto²³. Si tratta di un'ipotesi di reato che presenta potenzialità offensive nei confronti di una pluralità di beni giuridici meritevoli di tutela. In primo luogo, questa è una fattispecie capace di ledere la reputazione, la dignità e la libertà morale della persona protagonista del contenuto artefatto. Di ciò se ne ritrae conferma dal fatto che l'art. 612-*quater* c.p. sarebbe destinato a collocarsi nel Libro II, Titolo XII, Capo III, Sez. III del codice penale, vale a dire fra le fattispecie incriminatrici poste a tutela della libertà morale dell'individuo. Questa collocazione sistematica non è casuale dal momento che, molto spesso, il materiale prodotto mediante sistemi di i.a. generativa tende a colpire reputazione, dignità e libertà morale, nonché la libertà sessuale dell'individuo, con contenuti che non di rado sono sessualmente espliciti e peraltro usualmente *gender oriented*, a discapito delle donne. L'Unione europea, cogliendo questa tendenza, ha adottato la Direttiva (UE) 2024/1385

licitious uses and abuses of artificial intelligence europol.pdf. Si tenga inoltre presente, in un'ottica di comparazione, che la legislazione di alcuni Paesi europei si sta mostrando sensibile a questa distinzione, esemplarmente nei Paesi Bassi è usata la dicitura “*automated device*” invece di “*computer system*” all'art. 80-*sexies* c.c.

²² Etimologicamente, “*deepfake*” deriva dall'unione dell'espressione “*deep learning*” e l'aggettivo “*fake*” e sta ad indicare una tipologia di algoritmi che sfrutta l'intelligenza artificiale per generare dei falsi, “*fake*” appunto, per creare un'immagine, un audio voce e/o video umano sintetico, F. ARRUZZOLI, *Deepfake-Significato, Storia, Evoluzione*, in *ICT Security Magazine*, 17 ottobre 2022. Per un approfondimento sull'intelligenza artificiale generativa e sul *deepfake* si veda il Libro Bianco della Commissione Europea, *Generative AI and Disinformation: Recent Advances, Challenges, and Opportunities*, Febbraio 2024, disponibile al seguente link <http://edmo.eu/wp-content/uploads/2023/12/Generative-AI-and-Disinformation-White-Paper-v8.pdf>.

²³ Il prospettato art. 612-*quater* c.p. recita: «*Chiunque cagiona un danno ingiusto ad una persona, cedendo, pubblicando o altrimenti diffondendo, senza il suo consenso, immagini, video o voci falsificati o alterati mediante l'impiego di sistemi di intelligenza artificiale e idonei a indurre in inganno sulla loro genuinità, è punito con la reclusione da uno a cinque anni. Il delitto è punibile a querela della persona offesa. Si procede tuttavia d'ufficio se il fatto è connesso con altro delitto per il quale si deve procedere d'ufficio ovvero se è commesso nei confronti di persona incapace, per età o per infermità, o di una pubblica autorità a causa delle funzioni esercitate*». Per un approfondimento sul d.d.l. n. 1146/2024 si veda S. DE FLAMMINEIS, [Fattispecie penali nel contesto dell'intelligenza artificiale. Lo spunto del d.d.l. 1146/2024](#), in questa Rivista, fasc. 9/2024.

del Parlamento europeo e del Consiglio, del 14 maggio 2024, sulla lotta alla violenza contro le donne e alla violenza domestica, il cui art. 5, par. 1, lett. b) rivolge agli Stati membri un obbligo di incriminazione proprio in materia di *deepfake* a sfondo sessuale²⁴. Il nuovo reato contenuto nel d.d.l. n. 1146 costituirebbe un adempimento tempestivo da parte dell'Italia a quest'obbligo di tutela, ma non solo. Infatti, data la formulazione dell'art. 612-*quater* c.p., il presidio penale non si rivolgerebbe in via esclusiva al solo bene giuridico della libertà morale dell'individuo (nonostante la sua collocazione sistematica), piuttosto la norma risulta finalizzata a sanzionare qualsiasi condotta di *deepfake* capace di ingenerare un danno ingiusto. Tra queste, si pensi esemplarmente all'ipotesi di diffusione di un contenuto foto/video/audio falsificato riguardante una pubblica autorità in virtù delle funzioni che questa esercita con lo scopo di influenzare l'opinione pubblica di un Paese, e in particolare l'opinione politica²⁵. Questa eventualità è espressamente tenuta in considerazione dal prospettato inserimento dell'art. 612-*quater* nel codice penale, disposizione che nel suo secondo comma prevede che il reato diventi procedibile d'ufficio ove il contenuto manipolato diffuso abbia ad oggetto una pubblica autorità per il ruolo che questa ricopre, emergendo anche attraverso il piano della procedibilità le potenzialità lesive di questa condotta, spiccate in effetti, allorquando il *deepfake* sia finalizzato a realizzare disinformazione e propaganda.

In virtù di queste preoccupazioni, l'art. 50 dell'A.I. Act ha previsto degli obblighi di trasparenza rivolti a fornitori e *deployers* di i.a. generativa, con lo scopo di rendere nota la non genuinità del contenuto prodotto. In particolare, l'art. 50, par. 4, dell'A.I. Act prevede che i *deployers* «di un sistema di IA che genera o manipola immagini o contenuti audio o video che costituiscono un «deep fake» rendono noto che il contenuto è stato generato o manipolato artificialmente». Attraverso quest'obbligo di trasparenza (e forse, anche con il nuovo reato di *deepfake* che ne sanziona penalmente la violazione ove questa sfoci in un danno ingiusto alla persona) si persegue l'obiettivo di arginare i *cons*, tutt'altro che trascurabili, derivanti da questo particolare impiego delle nuove tecnologie.

Inoltre, è opportuno evidenziare che il ddl. n. 1146 sull'intelligenza artificiale prevede altresì l'introduzione nel codice penale di un certo numero di aggravanti, di cui una comune e le restanti speciali. Quanto all'aggravante comune, all'art. 61, co. 1, c.p. si andrebbe ad aggiungere il numero 11-*decies*²⁶, mentre le aggravanti speciali si

²⁴ L'art. 5 della direttiva, rubricato «Condivisione non consensuale del materiale intimo o manipolato», al suo par. 1, lett. b) dispone: «Gli Stati membri provvedono a che siano punite come reato le condotte intenzionali seguenti: produrre, manipolare o alterare e successivamente rendere accessibile al pubblico, tramite TIC, immagini, video o analogo materiale in modo da far credere che una persona partecipi ad atti sessualmente espliciti, senza il consenso della persona interessata, qualora tali condotte possano arrecare un danno grave a tale persona».

²⁵ È il caso della diffusione di contenuti falsificati con finalità di disinformazione e propaganda politica, fra i quali rientra il celeberrimo l'episodio risalente al febbraio del 2022 in cui è stato diffuso un video *fake* del Presidente ucraino Volodymyr Zelensky con il quale questi avrebbe chiesto al popolo ucraino di deporre le armi.

²⁶ In base al quale costituirebbe aggravante comune «l'aver commesso il fatto mediante l'impiego di sistemi di intelligenza artificiale, quando gli stessi, per la loro natura o per le modalità di utilizzo, abbiano costituito mezzo insidioso, ovvero quando il loro impiego abbia comunque ostacolato la pubblica o la privata difesa, ovvero aggravato le conseguenze del reato».

rivolgerebbero a fattispecie determinate fra cui, a titolo esemplificativo, la truffa, la frode informatica e i reati in senso lato di riciclaggio.

Rispetto alla prospettata introduzione della circostanza comune, è apprezzabile che ad aggravare il fatto non sia il mero utilizzo di un sistema di i.a., cosa che in effetti non conferisce necessariamente alla condotta un *quid pluris* in termini di offensività, bensì è richiesto l'accertamento in concreto del fatto che la condotta si sia rivelata effettivamente più riprovevole proprio perché realizzata impiegando l'intelligenza artificiale.

Al contrario, la formulazione delle aggravanti speciali è caratterizzata da un riferimento al mero all'impiego dell'intelligenza artificiale, senza che questo venga ulteriormente connotato²⁷.

Dunque, in relazione alle aggravanti speciali, ci si domanda (tornando così al punto di partenza di questo esame delle condotte dolose) se l'aggiunta del riferimento espresso all'impiego dell'intelligenza artificiale sia stata mossa dalla preoccupazione che le condotte sarebbero altrimenti rimaste impunte perché non riconducibili ai reati preesistenti, oppure, se ci sia stato in realtà il timore che queste non venissero punite abbastanza severamente.

Forse, richiedere anche per l'applicabilità delle circostanze speciali, oltre che per l'aggravante comune, l'accertamento in concreto che l'impiego dell'i.a. si sia rivelato maggiormente offensivo, eviterebbe quella che a tutti gli effetti sembra una sorta di presunzione di maggior gravità del fatto.

2.2. Elemento soggettivo: qual è l'oggetto del dolo?

In relazione all'elemento soggettivo degli A.I. *crimes*, appunto il dolo, vale la pena porsi due interrogativi.

In primo luogo, qual è l'oggetto del dolo? Per attribuire al soggetto agente i risultati della propria condotta è necessario che questi, avvalsi di sistemi di i.a., ne comprenda in modo piuttosto approfondito i sistemi di funzionamento (cosa che presupporrebbe una conoscenza tecnica-informatica superiore alla media) oppure è sufficiente che sia conscio di essersi avvalso di sistemi di i.a.²⁸?

²⁷ Si consideri, a titolo esemplificativo, l'art. 25, co. 1, lett. f) del menzionato d.d.l. che, per quanto riguarda il reato di truffa, aggiungerebbe all'art. 640, co. 2, il numero 2-ter introducendo una circostanza aggravante «se il fatto è commesso mediante l'impiego di sistemi di intelligenza artificiale»; lo stesso avverrebbe in relazione al reato di frode informatica, ex art. 640-ter c.p., che subirebbe la stessa modifica al suo terzo comma ai sensi della lett. g) del d.d.l.; similmente è previsto in relazione ai reati di riciclaggio (art. 648-bis c.p.), impiego di denaro, beni o utilità di provenienza illecita (art. 648-ter c.p.) e autoriciclaggio (art. 648-ter., co. 1, c.p.) che dovrebbero essere innovati rispettivamente dalle lett. h), i) ed l) del d.d.l. sull'intelligenza artificiale proprio con un riferimento al solo impiego dell'i.a.

²⁸ La questione viene sollevata da Beck e Gerndt, nei seguenti termini: «it has to be noted that understanding the workings of the AI system or its susceptibility to errors is necessary for the commission of an intentional crime in order to be able to draw conclusions about the causality from the perspective of the perpetrator and to hold the perpetrator responsible for intentional crime», si veda S. BECK - S. GERNDT, *German report on traditional criminal*

In secondo luogo, sono attribuibili al soggetto agente le conseguenze ulteriori della propria condotta a questa legate dal nesso causale ma non rientranti nel fuoco del dolo? Nei sistemi dotati di un tasso più elevato di autonomia, può accadere che manchi una diretta corrispondenza fra *input* fornito dal soggetto agente e *output* ottenuto, benché l'*input* fornito all'i.a. fosse comunque diretto alla commissione di un reato. Autorevole dottrina reputa applicabili i tradizionali istituti dell'*aberratio ictus* e dell'*aberratio delicti*, sostenendo che «*the mere material diversity of the harmed object must not represent an excuse if its characteristics are not relevant for the configuration of the criminal offence (e.g. killing one person instead of another is not relevant for the realization of the crime of murder, when it is intended by the agent)*»²⁹.

Le questioni restano aperte.

3. Condotte colpose.

3.1. Premessa.

Visto il largo impiego dei sistemi intelligenti nelle attività lecite svolte dall'uomo, è molto probabile che di qui a breve nelle aule dei tribunali si dovrà stabilire se, e in che misura, possa essere mosso un addebito colposo nei confronti di taluno per un evento nefasto realizzatosi in occasione della collaborazione-uomo macchina.

Si pensi, all'ipotesi paradigmatica delle *self-driving cars*, visti i precedenti³⁰, o ai casi in cui la tecnologia offra ai medici, e soprattutto alla collettività, strumenti innovativi per migliorare le capacità di diagnosi e di cura³¹. Ma si pensi anche all'attività di produzione e messa in commercio degli stessi sistemi di i.a. In queste circostanze potrebbe verificarsi un evento avverso non voluto dal soggetto agente. Ma allora perché è accaduto? E soprattutto, per colpa di chi?³²

law categories and AI, in *Traditional Criminal Law Categories and AI: Crisis or Palingenesi?*, a cura di L. PICOTTI - B. PANATTONI, International Colloquium Section I, Siracusa, 15-16 settembre 2022, p. 211.

²⁹ L. PICOTTI, *Resolution on traditional criminal law categories and AI*, in L. PICOTTI - B. PANATTONI (a cura di), cit., p. 54.

³⁰ In Florida una vettura Tesla modello S si infilava sotto ad un camion bianco, non riuscendo a distinguerlo dal cielo luminoso, distruggendo completamente l'abitacolo e provocando così la morte del conducente, 7 maggio 2016; in Arizona una vettura Uber completamente autonoma investiva un pedone, provocandone la morte, il 18 marzo 2018; in California, una donna che spingeva una bicicletta in una strada a quattro corsie è stata colpita e uccisa da un veicolo di prova Uber, il 23 marzo 2018.

³¹ Merita menzione il *software* italiano "THAITI", l'innovativa applicazione dell'i.a. nell'ambito dell'*imaging* medico relativo alla risonanza magnetica cardiaca. Il *software* è in grado di calcolare il valore ottimale del cosiddetto "tempo di inversione", un parametro necessario per l'acquisizione di immagini mirate a identificare l'eventuale presenza di tessuto cicatriziale nel cuore a seguito di somministrazione di un mezzo di contrasto, per un approfondimento: <https://www.unimib.it/comunicati/inventato-italia-software-ottimizzare-qualita-delle-immagini-risonanza-magnetica-cardiaca-1>.

³² Per un inquadramento di teoria generale sulle fattispecie colpose si vedano, *ex multis*, G. MARINUCCI, *La colpa*, Milano, 2013, e G. FORTI, *Colpa ed evento nel diritto penale*, Milano, 1990.

Ci si chiede come impatterà l'impiego dell'i.a. nell'accertamento della responsabilità colposa. Segnatamente, quali sono le regole cautelari da rispettare per un prudente impiego dell'i.a.? Quali quelle rivolte ai produttori di i.a.?

In questa parte, che affronta le questioni più spinose in tema di responsabilità penale dell'uomo connessa all'uso di A.I. *systems*, il proposito dello scritto è quello di individuare i nodi da sciogliere³³ e tentarne una cauta ricostruzione partendo dalla lettera del recente Regolamento (UE) 2024/1689 del Parlamento europeo e del Consiglio, il quale ha introdotto degli obblighi specifici e un generale obbligo di *human oversight*.

3.2. Gli obblighi positivi introdotti dall'A.I. Act.

Per stare al passo con i tempi, il legislatore europeo si è premurato di introdurre delle regole armonizzate non solo a tutela del mercato interno, ma anche a presidio dei diritti fondamentali dell'uomo³⁴.

Tra gli obblighi positivi introdotti figurano obblighi specifici (artt. 8 e ss. dell'A.I. Act) indirizzati agli operatori economici coinvolti nella produzione e nello sviluppo dei sistemi cd. ad *alto rischio*. In un'ottica di massima riduzione del rischio, ove tali obblighi positivi specifici non dovessero bastare ad evitare eventi avversi, per ragioni di prossimità rispetto all'eventuale causazione di un danno, è rivolto un generale obbligo di *human oversight* all'utilizzatore (art. 14 A.I. Act).

Quanto agli obblighi specifici che si rivolgono esclusivamente agli operatori economici, ci si domanda se si tratta di regole di condotta la cui violazione è idonea ad integrare l'elemento normativo della colpa specifica, se siano cioè delle regole cautelari. Più precisamente, sono tutte regole connotate dalla finalità preventiva di riduzione del rischio o molte di esse hanno una finalità diversa, quella appunto di dettare regole uniformi per il commercio dell'i.a. nel mercato interno?

Forse, sarebbe prudente affermare che ove dovesse accadere un evento dannoso in occasione della violazione di uno degli obblighi specifici dettati dal regolamento agli artt. 8 e ss., andrebbe verificato caso per caso se la regola violata fosse genuinamente connotata da finalità preventiva, cioè che questa fosse finalizzata a ridurre il rischio di verificazione di quello specifico evento e che non fosse, al più, una prescrizione indirizzata agli operatori economici per garantire un *fair play* nel mercato interno. Certo, non è da escludere che alcune regole di condotta possano perseguire contestualmente più di una finalità, fra qui quella propria delle regole cautelari.

Inoltre, da alcune recenti pronunce di legittimità emerge la tendenza all'abbandono della teoria del cd. "rischio consentito" in base alla quale la previsione di

³³ Nodi che la dottrina penalistica internazionale più esperta in materia ha individuato e sta sciogliendo in questi anni. Il riferimento, rispetto alle questioni di diritto penale sostanziale, è agli interventi tenuti da K. LIGETI - L. PICOTTI - J. WALTHER - X. WANG - S. BACK - S. AIRES DE SOUSA - C.E. JAPIASSÙ - F. MIRÒ LINARÈS - M. NORD-WAGNER - A. GULLO - N. GROFU - A. KHALIFA, il 26 giugno 2024, nelle Sezioni I e II, del XXI Congresso internazionale di diritto penale organizzato dall' *Association Internationale de Droit Pénal* (AIDP).

³⁴ Cons. 3 dell' A.I. Act.

regole di condotta scritte costituisce un limite alla possibilità di muovere un rimprovero per colpa generica, a favore dell'esclusivo ricorso alle categorie generali che informano la colpa.

Infatti, rimanendo nel campo delle attività pericolose ma lecite e regolamentate perché socialmente utili, la Suprema Corte di Cassazione in una recente pronuncia in materia di lesioni colpose cagionate nell'esercizio di un'attività sportiva ha reputato non soddisfacente l'impostazione in base alla quale il limite fra illecito sportivo e illecito penale sia rappresentato, appunto, dall'area di rischio consentito. Piuttosto, ha ritenuto che lo sportivo di una pratica "di contatto", accanto al dovere di rispettare le regole dello sport intrapreso, sia tenuto al più generale onere di osservanza delle regole cautelari, prudenziali e di diligenza³⁵.

Applicando questo principio all'attività di produzione e messa in circolazione dell'i.a., attività ora regolamentata dall'A.I. *Act*, non è improbabile che anche in questo settore il rapporto fra colpa specifica e colpa generica sia visto in un'ottica di integrazione piuttosto che di esclusione, andando a sommarsi per i produttori l'ipotesi di una responsabilità per colpa generica a quella per colpa specifica³⁶.

A ciò si aggiunga che proprio l'A.I. *Act* offre, in parte, una base normativa per questo tipo di interpretazione in relazione alla posizione dei soggetti coinvolti nella produzione e nella messa in circolazione dei sistemi di i.a. cd. *ad alto rischio*. Infatti, l'art. 9, rubricato "sistema di gestione dei rischi" nel suo par. 4, lett. a), impone ai soggetti produttori di individuare le misure di gestione del rischio garantendo «l'eliminazione o la riduzione dei rischi per quanto possibile attraverso un'adeguata progettazione e fabbricazione»³⁷.

Pertanto, sarebbe tutt'altro che privo di fondamento ritenere che, quanto meno nel momento in cui questi sono chiamati ad individuare le misure di gestione del rischio, la condotta dei soggetti che producono e commercializzano i sistemi di i.a. sia raffrontabile al parametro dell'*homo eiusdem condicionis et professionis*, rendendo ancora

³⁵ Cass. Pen., Sez. IV, 21 ottobre 2021, n. 3284, oltre alla tematica del rischio consentito, la pronuncia si addentra anche sulla dicotomia regole di condotta e regole cautelari; dello stesso segno si veda anche, sulla questione del rischio consentito, Cass. Pen., Sez. IV, 28 ottobre 2021, n. 8609.

³⁶ Per un approfondimento del rapporto fra colpa generica e colpa specifica si veda G. CIVELLO, *Colpa stradale, affidamento e autoresponsabilità: una sentenza esemplare*, Nota a Cass., sez. IV, sent. 14 dicembre 2023 (dep. 20 dicembre 2023), n. 50816, pres. Piccialli, rel. Mari, in questa [Rivista](#), 2/2024, p. 42; con riferimento al rapporto tra colpa generica e colpa specifica nella responsabilità per colpa medica, si veda C. CUPELLI, [Lo statuto penale della colpa medica e le incerte novità della legge Gelli-Bianco](#), in *Dir. Pen. Cont. - Riv. trim.*, 4/2017, p. 206, secondo cui: «Più esauriente – rispetto alle attività rischiose, giuridicamente autorizzate – è la moderna concezione della cd. *duplice dimensione* o *misura dell'imputazione colposa*: a) *la dimensione oggettiva*, consistendo il primo elemento essenziale nella condotta violatrice della preesistente regola cautelare di condotta, volta a salvaguardare i beni giuridici; b) *la dimensione soggettiva*, consistendo il secondo elemento essenziale nell'esigibilità dell'osservanza di tale regola da parte dell'agente (e, quindi, nell'attribuibilità al medesimo dell'inosservanza)».

³⁷ (Corsivo aggiunto).

più concreta l'ipotesi che il rispetto delle regole cautelari scritte non escluda nei loro confronti un addebito per colpa generica³⁸.

3.4. Obbligo di *human oversight*.

L'art. 9 dell'A.I. Act, al suo par. 4, definisce accettabili i rischi che, nonostante l'adozione delle misure di gestione di cui all'art. 9, par. 2, lettera d), residuano per la salute, la sicurezza e per i diritti fondamentali e prevede che questa circostanza sia comunicata all'utente. Insomma, nel passaggio dell'i.a. dal soggetto che l'ha prodotta al soggetto che ne fa uso, su quest'ultimo è allocato il rischio (di danno) residuo³⁹.

In effetti, in base all'art. 14, par. 2, del regolamento, è l'utilizzatore ad essere tenuto ad un obbligo di *human oversight* «in particolare quando tali rischi persistono nonostante l'applicazione di altri requisiti di cui al presente capo»⁴⁰.

Ma che ruolo svolge, *rectius* potrebbe svolgere, questo obbligo sul terreno della responsabilità penale? Più schiettamente, l'art. 14 dell'A.I. Act fonda una posizione di garanzia in capo al soggetto utilizzatore oppure rivolge a questi una regola cautelare?

Nel primo caso, in virtù dell'art. 40, co. 2, c.p., l'uomo-supervisore risponderebbe dell'evento dannoso non solo quando lo ha cagionato, ma anche quando non lo abbia impedito⁴¹, cioè ove non sia intervenuto su una catena causale innescata da altri interrompendola prima della verifica del danno;

ove invece si intenda la norma quale fonte di una regola cautelare, *l'human in the loop*⁴² sarebbe responsabile se è parte della catena causale che ha portato alla verifica del danno, cioè se ha fornito un contributo materiale alla causazione dell'evento con una condotta commissiva colposa.

Potremmo sostenere che il riportato inciso contenuto nell'art. 14, par. 2, del regolamento lasci intendere che ove diligenza, prudenza e perizia possano ridurre la possibilità di verifica di rischi ulteriori, queste accortezze debbono essere prestate dal soggetto utilizzatore secondo quanto ci si può ragionevolmente aspettare dall'*homo eiusdem condicionis et professionis*. Dunque, per le ragioni che seguono, si potrebbe qualificare l'obbligo di *human oversight* quale fonte di una regola cautelare.

In primo luogo, dato che l'i.a. può presentare diversi gradi di automazione, è condivisibile l'opinione in base alla quale si potrebbe rinvenire nell'obbligo di

³⁸ In questa direzione anche B. FRAGASSO, [La responsabilità penale del produttore di sistemi di intelligenza artificiale](#), in questa Rivista, 13 giugno 2023, p.15.

³⁹ Infatti, in base all'art. 14, par. 1, del regolamento «I sistemi di IA ad alto rischio sono progettati e sviluppati, anche con strumenti di interfaccia uomo-macchina adeguati, in modo tale da poter essere efficacemente supervisionati da persone fisiche durante il periodo in cui sono in uso».

⁴⁰ Art. 14, par. 2, del regolamento.

⁴¹ Propende per questa ricostruzione in L. PICOTTI, *Categorie tradizionali del diritto penale e intelligenza artificiale: crisi o palingenesi?*, cit.

⁴² Espressione traducibile «l'uomo al centro dell'algoritmo». Per un approfondimento sul tema, si vedano, A. GIANNINI, *Intelligenza artificiale, human oversight e responsabilità penale*, in *disCrimen*, 21 novembre 2022, p. 16; B. GREEN - A. KAK, *The false Comfort of human oversight as an antidote to A.I.*, in *Slate*, 15 giugno 2021.

sorveglianza umana la fonte di una posizione di garanzia soltanto quando la macchina è del tutto automatica e non quando il contributo umano resta assolutamente necessario⁴³. Infatti, solo se il sistema intelligente fosse del tutto autonomo all'uomo potrebbe essere attribuito anche un ruolo "esterno" sulla catena causale degli eventi, quale quello di garante dell'operato altrui (dell'i.a. in questo caso). Tuttavia, allo stato attuale, il contributo umano è ancora necessario nonché espressamente richiesto⁴⁴.

V'è da aggiungere che, in relazione alle fonti delle posizioni di garanzia secondo quella parte della dottrina che sostiene la concezione formale, affinché una certa norma possa essere reputata fonte di una posizione di garanzia, fra gli altri requisiti, deve essere dotata di un sufficiente grado di specificità sì da rendere prevedibile per il soggetto agente che il non fare, in un certo contesto, viene equiparato al fare⁴⁵. In questo caso, la dicitura *human oversight* non sembra presentare i necessari caratteri di specificità.

Ma in cosa si sostanzia quindi la sorveglianza umana? La sola richiesta di *human oversight* suggerisce poco. Sulla scorta di quanto è già accaduto per gli *autonomous vehicles* in Germania e in Francia⁴⁶, si potrebbe ragionevolmente ritenere che in questa parte il regolamento abbia lasciato agli Stati Membri la possibilità di disciplinare gli aspetti riguardanti la posizione dell'utilizzatore. Il che è apprezzabile, dal momento che il legislatore europeo, giova ribadirlo, è intervenuto sulla base dell'art. 114 TFUE a tutela del mercato interno. Con tale intervento, realizzato con una fonte normativa direttamente applicabile, si può dire che questi abbia legiferato nel pieno rispetto dei principi di sussidiarietà e di proporzionalità⁴⁷ concentrandosi sulla posizione degli operatori economici ma senza ignorare il ruolo svolto poi in concreto dagli utilizzatori degli a.i. *systems*, obbligando anche questi ultimi quanto meno a sorvegliare il sistema di i.a. quando questo è in uso.

In definitiva, si potrebbe affermare che la prescrizione di *human oversight* abbia la natura giuridica di norma cautelare, pur nella sua generalità. Questa fungerebbe da *memorandum* del fatto che lo svolgimento di attività lecite con l'aiuto dell'i.a. debba essere condotto secondo la perizia, la prudenza e la diligenza che si possono ragionevolmente attendere dall'*homo eiusdem condicionis et professionis*, ferma restando la

⁴³ In base al più diffuso standard J3016 della *SAE International* (*Society of Automotive Engineers*), è possibile distinguere sei livelli di automazione complessiva dell'auto automatica da 0 a 5. Per un approfondimento sul tema, si veda A. CAPPELLINI, [Profili penalistici delle self-driving cars](#), in *Dir. Pen. Con., Riv. trim.*, 2/2019, p. 325.

⁴⁴ Questo è espressamente richiesto dall'A.I. *Act* all'art. 14, par. 1, in base al quale: «*I sistemi di IA ad alto rischio sono progettati e sviluppati, anche con strumenti di interfaccia uomo-macchina adeguati, in modo tale da poter essere efficacemente supervisionati da persone fisiche durante il periodo in cui sono in uso*».

⁴⁵ Nella manualistica tradizionale, si veda, G. MARINUCCI-DOLCINI - G.L. GATTA, cit., pp. 270 e ss. Sui requisiti della fonte della posizione di garanzia, si veda, F. MANTOVANI - G. FLORA, *Diritto Penale, Parte Generale*, XII ed., Wolters Kluwer CEDAM, Padova, p. 160.

⁴⁶ In Germania è stata approvata una legge che è intervenuta sul codice della strada per disciplinare le ipotesi in cui il conducente si avvalga di un veicolo a guida semi-autonoma o completamente autonoma; allo stesso modo, nel 2021 con una *ordonnance* il legislatore francese ha modificato il *Code de la route* disciplinando l'ipotesi di responsabilità penale derivante dall'uso di auto dotate di i.a.

⁴⁷ Principi previsti dall'art. 5 TUE rispettivamente ai paragrafi 3 e 4 al fine di regolamentare l'esercizio delle competenze da parte dell'Unione europea.

possibilità di una futura introduzione di obblighi positivi che specifichino in che cosa consista la sorveglianza dell'utilizzatore nel singolo settore di riferimento⁴⁸. Interventi di questo tipo aprirebbero la strada ad un possibile rimprovero per colpa specifica nei riguardi dell'utente-sorvegliante riproponendosi anche rispetto a quest'ultimo le tradizionali incertezze applicative relative al rapporto fra colpa generica e colpa specifica esaminate in riferimento ai soggetti produttori.

4. L'accertamento causale.

In linea di massima, il tradizionale procedimento di accertamento della causalità materiale⁴⁹ può ritenersi applicabile anche in relazione ad un fatto storico integrante reato commesso con l'uso dell'i. a.

L'accertamento del nesso causale, come accade rispetto ad ogni accertamento in concreto, dovrà essere condotto tenendo conto delle specificità della materia, in questo caso cioè applicando le regole di spiegazione scientifica relative al funzionamento dell'i.a.⁵⁰. Tuttavia, tale adattamento non è privo di incertezze dal momento che i sistemi di i.a. possono essere dotati di un più o meno spiccato livello di autonomia nel *decision making process*. Questo significa che l'evento dannoso verificatosi in occasione della collaborazione uomo-macchina possa essere causalmente legato ad un *output* del sistema non totalmente deducibile dall'*input* che la macchina ha ricevuto da parte dall'utilizzatore-sorvegliante⁵¹. In questi casi, dunque, non è possibile attribuire con certezza l'evento verificatosi alla condotta del soggetto-utilizzatore. In altre parole, si deve tenere in considerazione il fenomeno della *black box*, cioè il fatto che la ricostruzione dei passaggi compiuti dal sistema possa non avvenire in maniera del tutto spiegabile. Infatti, può accadere che questi possano non essere totalmente ripercorribili attraverso il giudizio controfattuale compiuto *ex post* proprio in virtù del grado di autonomia della macchina. Tuttavia, ciò «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».⁵²

V'è da aggiungere che l'A.I. Act all'art. 12, par. 1, rivolge ai fornitori di sistemi di intelligenza artificiale un obbligo di «registrazione automatica degli eventi "log" per la durata

⁴⁸ Risvolto auspicato da S. BECK - S. GERNDT, *German report on traditional criminal law categories and AI*, cit.

⁴⁹ In base all'art. 27, co. 1, della Costituzione, il principio di personalità della responsabilità penale richiede che l'offesa prodottasi sul piano fenomenico, in termini di danno o di messa in pericolo, sia conseguenza dell'azione o della mancata azione del soggetto agente. Il contegno del soggetto agente deve costituire, secondo la celebre formula del criminalista tedesco Von Buri, la *conditio sine qua non* della verifica dell'offesa. La condotta del soggetto è *condizione senza la quale* l'evento dannoso non si sarebbe verificato ogni qual volta, attraverso un giudizio ipotetico controfattuale condotto *ex post*, alla luce di leggi di spiegazioni scientifica o massime di esperienza quell'evento non si sarebbe accaduto.

⁵⁰ in L. PICOTTI, *Categorie tradizionali del diritto penale e intelligenza artificiale*, cit., p. 9.

⁵¹ Il riferimento è al *machine learning*, per il quale si rimanda alla nota n. 2.

⁵² in L. PICOTTI, *Categorie tradizionali del diritto penale e intelligenza artificiale*, cit., p. 9, si veda *supra* nota n. 15.

del ciclo di vita del sistema». Tale registrazione è preordinata all'assolvimento dell'ulteriore obbligo di conservazione della traccia degli eventi, obbligo la cui portata è ulteriormente specificata all'art. 19 del regolamento⁵³. Queste previsioni sono infatti espressamente finalizzate a «garantire un livello di tracciabilità del funzionamento del sistema di intelligenza artificiale ad *alto rischio*»⁵⁴, tracciabilità di cui ci si potrà servire ove debba essere accertata la responsabilità penale dell'utilizzatore. Infatti, adempimenti quali la registrazione e la conservazione di determinati dati quali, a titolo esemplificativo, «il periodo di ciascun utilizzo del sistema (data e ora di inizio e di fine di ciascun utilizzo), la banca dati di riferimento utilizzata dal sistema per verificare i dati di input, i dati di input per i quali la ricerca ha portato a una corrispondenza»⁵⁵ giocheranno indubbiamente un ruolo importante allorquando il giudice penale sarà chiamato ad accertare la sussistenza del rapporto causa-effetto intercorrente fra una certa condotta e un dato evento. Tuttavia, bisogna ammettere che la strada per la completa *explainability* dei sistemi di i.a. particolarmente autonomi è ancora tutta in salita⁵⁶.

5. Un possibile scenario: fra colpa dell'utilizzatore e *responsibility gap*.

Quando dall'uso di una macchina che non presenta malfunzionamenti deriva un evento dannoso, questo potrebbe essere riconducibile ad una insufficiente sorveglianza da parte del soggetto utilizzatore⁵⁷. Se si volesse portare a conseguenze la ricostruzione dell'obbligo di *human oversight* di cui all'art. 14. del regolamento quale norma cautelare, intanto il danno sarà attribuibile al soggetto sorvegliante in quanto, al posto suo, l'agente modello lo avrebbe previsto ed evitato. Infatti, si potrebbe osservare che molto spesso ciò che sembra in linea di principio meglio conoscibile da chi programma e costruisce una certa cosa, sia in realtà più familiare a chi si cimenta quotidianamente nel suo impiego. Questa riflessione può valere tanto per il *driver* mediamente esperto che in certe condizioni di tempo e di luogo sa (o dovrebbe sapere) quali imprevisti possono verificarsi lungo il percorso anche quando la guida dell'auto è in modalità semi-autonoma, quanto per il medico mediamente esperto che abbia alle spalle una casistica

⁵³ Quanto ai log generati automaticamente, l'art. 19, par. 1, del regolamento prevede che «i fornitori di sistemi di IA ad alto rischio conservano i log di cui all'articolo 12, paragrafo 1, generati automaticamente dai loro sistemi di IA ad alto rischio, nella misura in cui tali log sono sotto il loro controllo. Fatto salvo il diritto dell'Unione o nazionale applicabile, i log sono conservati per un periodo adeguato alla finalità prevista del sistema di IA ad alto rischio, della durata di almeno sei mesi, salvo diversamente disposto dal diritto dell'Unione o nazionale applicabile, in particolare dal diritto dell'Unione in materia di protezione dei dati personali».

⁵⁴ Art. 12, par. 2, del regolamento.

⁵⁵ Art. 12, par. 3, lett. a), b), c).

⁵⁶ Con «*Explainable Artificial Intelligence (xAI)*» si intende «l'insieme delle conoscenze scientifiche finalizzate alla spiegazione della logica interna che sta dietro all'algoritmo utilizzato dal sistema intelligente per fare chiarezza su quanto succede nella «*black box*» dei dati e degli algoritmi che portano ad un certo *output* a fronte di un dato *input*».

⁵⁷ Anche l'A.I. Act all'art. 14, par. 4, lett. b) prende in considerazione il rischio di *overreliance* da parte dell'utilizzatore sull'*output* prodotto dal sistema di i.a.

tale da permettergli di rendersi conto che l'*output* prodotto da un sistema diagnostico basato sull'i.a., in quel caso specifico non è calzante.

In sede di accertamento della responsabilità colposa per non aver vigilato, bisogna verificare che il *driver*-agente modello o il medico-agente modello, al posto delle persone concretamente coinvolte, avrebbero previsto ed evitato l'evento in base al grado di competenza ed esperienza che ci si può ragionevolmente attendere.

Se così non fosse, l'evento dannoso non sarebbe attribuibile né alla catena di sviluppo del sistema intelligente, essendo lo stesso perfettamente funzionante (nonché rispondente alle previsioni dell'A.I. *Act*), né al sorvegliante, avendo egli prestato l'accortezza possibile. Questa è l'ipotesi in cui un fatto dannoso penalmente rilevante non è ascrivibile a nessuna persona fisica. Allo stato dell'arte, il cd. *responsibility gap*⁵⁸ è un rischio congenito all'impiego dell'i.a. nelle attività lecite con cui il legislatore e i tribunali non hanno ancora fatto i conti.

6. Considerazioni conclusive.

Forse un giorno l'ipotesi di una responsabilità diretta dell'i.a.⁵⁹ diventerà realtà, ma fino ad allora il problema del *responsibility gap*, nutrito da una difficile *explainability* del modo in cui operano i sistemi di i.a. caratterizzati dal *machine learnig*, continuerà a rappresentare un *vulnus* nel giudizio di attribuzione della responsabilità penale.

L'augurio è che la scienza nel prossimo futuro sia sempre più in grado di approfondire e chiarire le complesse dinamiche che governano l'aggiornamento automatico degli algoritmi dell'i.a., fornendo agli studiosi delle indicazioni chiare sui legami intercorrenti fra certi impieghi dell'i.a. e il verificarsi di determinate conseguenze. Ove maggior chiarezza dovesse essere fatta, ciò potrebbe comportare una nuova valutazione dell'area del rischio consentito e dunque una nuova individuazione del punto di equilibrio fra vantaggi e svantaggi derivanti dall'impiego dell'intelligenza artificiale, o nel senso di un più esteso ambito di applicazione, o al contrario, di una sua limitazione ove certi *outcomes* dovessero rimanere inspiegabili, e dunque, inevitabili.

Fino ad allora, si auspica che la presa sul principio di personalità della responsabilità penale resti ben salda e che nonostante la tendenza dell'animo umano sia, per natura, quella di cercare qualcuno cui attribuire la responsabilità di un evento nefasto, il legislatore sappia regolamentare l'impiego dell'i.a. senza trasmodare dal principio di precauzione al "precauzionismo"⁶⁰.

⁵⁸ Per un approfondimento sul tema, si veda, per tutti, F. SANTONI DE SIO, *Four responsibility gaps with Artificial Intelligence, Why they matter and how to address them*, disponibile al seguente link: <https://link.springer.com/article/10.1007/s13347-021-00450-x>.

⁵⁹ Si veda, R. ABBOTT - A. SARCH, *Punishing Artificial Intelligence: Legal Fiction or Science Fiction?*, pp. 323 e ss., disponibile in open access al link: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3327485; si veda inoltre M.E. FLORIO, [Il dibattito sulla responsabilità penale diretta delle IA: "molto rumore per nulla"?](#), in questa *Rivista*, fasc. 2/2024.

⁶⁰ Per un'analisi dei pericoli connessi al "precauzionismo", si veda, A. MASSARO, *Principio di precauzione e diritto penale*, cit.

Editore

ASSOCIAZIONE
**"PROGETTO GIUSTIZIA
PENALE"**