

QUESITI

MATTIA DI FLORIO

**Il diritto penale che verrà.
Brevi considerazioni sul possibile
impiego dell'IA per prevenire il rischio di disastri
colposi**

L'articolo si sofferma sull'impatto delle nuove tecnologie, in particolare dell'intelligenza artificiale, in una prospettiva per certi versi originale che non attiene alla nota tematica della giustizia predittiva, bensì ai futuri "riverberi" sulla colpa penale. L'autore trae spunto da alcuni (più o meno) recenti casi di disastri colposi per soffermarsi, in adesione alla concezione *debole* di IA, sui possibili concreti effetti di quest'ultima sulla colpa penale. È auspicabile che il ricorso, nell'immediato futuro, all'IA consenta di ridurre l'incidenza della condotta umana nella verifica dei disastri colposi.

The criminal law to come. Brief considerations on the possible use of AI (and Big Data) to prevent the risk of negligent disasters

The article focuses on the impact of new technologies, in particular of artificial intelligence, in a somewhat original perspective that does not relate to the thematic note of predictive justice, but rather the future "reverberations" on criminal guilt. The author draws inspiration from some (more or less) recent cases of negligent disasters to focus, in adherence to the "weak" conception of AI, on the possible concrete effects the latter on criminal guilt. It is desirable that the recourse, in the immediate future, to the AI consent to reduce the incidence of human conduct in the verification of negligent disasters.

SOMMARIO: 1. Il recente disastro della funivia di Stresa: il caso. - 2. (segue) La rilevanza della condotta umana in materia di disastro ed omicidio colposo. - 3. *Uno sguardo tecnologico: l'IA per rendere le infrastrutture smart.* - 4. (segue) La proposta di "Regolamento IA" dell'UE - 5. *Brevi considerazioni critiche di diritto penale sulla concezione forte dell'IA.* - 6. (segue) *La concezione debole dell'IA e le implicazioni penalistiche per la colpa.* - 7. Conclusioni.

1. Il recente disastro della funivia di Stresa: il caso. L'incidente della funivia Stresa-Alpino-Mottarone, in Piemonte, del 23 maggio 2021, in cui sono morte quattordici persone (tra cui due bambini di due e cinque anni), mentre una sola è rimasta gravemente ferita (un bambino di cinque anni) ha costituito un tragico evento che ha "segnato" il graduale "ritorno alla normalità" *post* emergenza sanitaria da SARS-CoV-2: l'incidente si è, infatti, verificato il giorno immediatamente successivo alla riapertura degli impianti di risalita in monta-

gna, nel rispetto delle linee guida di settore, come previsto dal D.L. 18 maggio 2021, n. 65, pubblicato nella G.U. 18 maggio 2021, n. 117.

In breve i fatti. Poco dopo mezzogiorno del 23 maggio 2021 una delle cabine dell'impianto di funivia che collega Stresa alla sommità del Mottarone, con a bordo quindici persone (il numero dei passeggeri era ridotto a causa delle misure di distanziamento anti-SARS-CoV-2), mentre stava percorrendo in risalita la tratta che dall'Alpino sale al Mottarone, ormai pressoché raggiunta la stazione di arrivo, repentinamente ed improvvisamente invertiva il senso di marcia a causa della rottura della fune traente, e finiva per retrocedere lungo il cavo portante, a velocità sempre maggiore rispetto a quella normale. Giunta in prossimità di uno dei piloni del tracciato, a causa del brusco cambio di pendenza del cavo portante che regge il carico, la cabina si sganciava dal cavo stesso e precipitava nel vuoto fino a schiantarsi al suolo dopo una caduta di oltre venti metri.

All'arrivo dei soccorsi, tra i 15 passeggeri soltanto due bambini erano ancora in vita dopo il violentissimo impatto: trasportati tramite elisoccorso all'ospedale, il più piccolo decedeva dopo poche ore a causa delle gravissime lesioni riportate, mentre l'altro bambino, che veniva sottoposto ad intervento chirurgico, nonostante le plurime fratture riportate, sopravviveva.

La Procura di Verbania, che ha avviato le indagini per disastro ed omicidio colposo nei confronti dell'amministratore unico, del direttore di esercizio e del capo servizio della funivia, ha richiesto la convalida del fermo e di applicazione della misura coercitiva della custodia cautelare in carcere per i tre indagati; detta richiesta non è stata, tuttavia, convalidata dal GIP che ha applicato, invece, la misura cautelare degli arresti domiciliari solo nei confronti del capo servizio della funivia¹.

Uno dei nodi dell'indagine attiene all'apposizione da parte del capo servizio di dispositivi blocca freni che hanno inibito il normale funzionamento del sistema frenante di emergenza della cabina, collocato sulla fune più alta (portante). Come si legge nel decreto di fermo, il capo servizio della funivia avrebbe ammesso di avere «deliberatamente e ripetutamente inserito i dispositivi blocca freni (forchettoni), disattivando il sistema frenante di emergenza», una condotta «di cui erano stati ripetutamente informati» anche gli altri due indagati, i quali «avvallavano tale scelta e non si attivavano per consentire i necessari interventi di manutenzione che avrebbero richiesto il fermo

¹ L'ordinanza del GIP di Verbania del 29 maggio 2021 è reperibile in www.giurisprudenzapenale.com, 31 maggio 2021.

dell'impianto, con ripercussioni di carattere economico»².

Nelle more delle indagini la società responsabile della manutenzione ordinaria e straordinaria dell'impianto di Stresa ha reso noti i controlli recenti, eseguiti su richiesta del gestore, senza riscontrare problemi tecnici; essa ha, altresì, annunciato di costituirsi parte civile nel procedimento giudiziario, dal momento che «la manomissione degli impianti di sicurezza che ha portato alla tragica morte di 14 persone è un atto gravissimo. L'utilizzo dei cosiddetti forchettoni è espressamente vietato con persone a bordo. Eventuali risarcimenti verranno devoluti alle famiglie delle vittime»³.

2. (segue) La rilevanza della condotta umana in materia di disastro ed omicidio colposo. Il tragico incidente della funivia di Stresa induce ad interrogarsi su di un profilo preliminare, che attiene alla distinzione concettuale tra “caso” e rilevanza della condotta umana in materia di disastro colposo.

Ci sono due fattori compresenti suscettibili di portare ad un fatto simile: la rottura (o il distacco) anomalo della fune traente ed il mancato funzionamento del freno sulla fune portante.

Potrebbe apparire “strana” la rottura improvvisa della fune traente che costituisce un fenomeno “lento”: le funi in acciaio vengono periodicamente controllate, e la loro rottura non è certo un fenomeno inatteso, ma è un fenomeno lento di usura, corrosione, che deve essere previsto dal tecnico che monitora, a meno non ci sia stato qualcosa di stravolgente, come un'azione meccanica forte, un fulmine, un oggetto che cade in modo molto violento ed anche pesante (che almeno dalle prime fasi dell'indagine non sembrerebbe essere emerso).

Tra i controlli periodici delle funi, vanno richiamati, per la loro importanza, almeno i controlli magnetoinduttivi, che consistono nella sottoposizione delle funi (portanti o traenti) ad un campo magnetico che le controlla per la loro estensione; in tal modo, la fune localmente magnetizzata risponde con un segnale che è proporzionale, pur con margini di errore, all'area sottoposta al campo, per cui il tecnico è in grado di capire la situazione interna al cavo, anche se non vede fili rotti. Questi controlli sono impiegati da decenni e costituiscono la tecnica più consolidata; sebbene ci siano possibilità di ulteriori approfondimenti (come, ad es., l'apertura della fune o un'azione meccanica lunga), i controlli magnetoinduttivo già danno un segnale della situazione.

² Lo stralcio del decreto di fermo disposto dalla Procura di Verbania è riportato da www.ansa.it, 28 maggio 2021.

³ V. ancora www.ansa.it, 28 maggio 2021.

Da uno sguardo sommario all'elenco dei controlli eseguiti dalla società responsabile della manutenzione dell'impianto sembrerebbe evincersi che il controllo magnetoinduttivo delle funi traenti (e di tutte le funi dell'impianto) era stato eseguito con esito positivo in data 5 novembre 2020 (secondo la periodicità annuale imposta dalla legge).

Tra i precedenti dell'incidente della funivia di Stresa può ricordarsi, ad esempio, quello del Monte Bianco, dove il 29 agosto del 1961 il cavo portante della funivia tra *Punta Helbronner* e *l'Aiguille du Midi* venne tranciato da un cacciabombardiere francese in volo sopra il Monte Bianco: precipitarono tre cabine e sei occupanti morirono; i restanti passeggeri trascorsero la notte sospesi nel vuoto prima di essere salvati. Quell'incidente si verificò a distanza di 24 anni e mezzo dall'installazione della corda che doveva essere sostituita entro 25 anni: da allora l'obbligo della sostituzione della corda è passato a intervalli inferiori a venti anni.

Uno dei possibili problemi che incontrano gli impianti di alta quota è l'escursione termica: nelle giornate estive la corda è investita dai raggi solari ed aumenta di molto la temperatura anche intorno ai 70-80°; di notte, alle altezze più elevate la temperatura scende di parecchi gradi e l'umidità all'interno delle corde condensa e le gocce di acqua si radunano verso il fondo, e quindi è possibile che nel periodo di non utilizzo della funivia ci sia stata un accumulo di umidità e di formazione di gocce d'acqua che si sono concentrate verso la fine. Questo fenomeno, che prende il nome di *fretting fatigue* (ben noto ai tecnici del settore), comporta che negli strati interni della corda traente potrebbe esserci stata una rottura dei vari fili c.d. a becco di flauto.

Queste prime (sommario) osservazioni sembrerebbero escludere il rilievo del caso e far propendere per l'incidenza della condotta umana, dal momento che il caso fortuito deve costituire un avvenimento imprevedibile ed eccezionale che si inserisce d'improvviso nell'azione del soggetto (*ex art. 45 c.p.*).

Ulteriori conferme (ove mai ce ne fosse bisogno) in tal senso potrebbero trarsi da pregresse tristi vicende di disastri colposi che hanno segnato la più recente storia del nostro Paese.

Ci si riferisce, ad esempio, al disastro ferroviario del 12 luglio 2016, quando uno scontro frontale fra due convogli della tratta Andria-Corato della Ferrotramviaria provocò la morte di 23 persone ed il ferimento di altri 51 passeggeri; nel processo davanti al Tribunale di Trani sono imputate la Ferrotram-

viaria s.p.a.⁴ ed altre 17 persone fisiche (dipendenti e dirigenti dell'azienda pugliese di trasporti e del Ministero delle Infrastrutture accusate a vario titolo di disastro ferroviario, omicidio colposo e lesioni colpose gravi, omissione dolosa di cautele, violazione delle norme sulla sicurezza e falso. Durante un'udienza del processo - come si apprende dalle cronache giudiziarie - sarebbe stato oggetto di contraddittorio l'obsolescenza del regime del c.d. blocco telefonico, cioè il sistema di telefonate per far partire i treni (risalente alla metà dell'Ottocento), che regolava la sicurezza della circolazione ferroviaria nella tratta dove era avvenuto l'incidente⁵. Il mancato adeguamento tecnologico della linea - secondo i consulenti tecnici della Procura - avrebbe, inoltre, comportato che per ben 146 volte prima del disastro si erano venuti a creare i presupposti potenziali per uno scontro frontale tra due convogli.

Un altro disastro ferroviario, prevedibile ed evitabile, secondo la Procura di Milano che ha concluso le indagini (come si apprende dalla cronaca giudiziaria), è stato quello di Pioltello sulla ferrovia Milano-Venezia dove, il 25 gennaio 2018, il deragliamento di un treno passeggeri causò tre morti e 46 feriti; l'incidente che, secondo una relazione dei consulenti della Procura milanese, sarebbe stato causato dall'ormai noto "spezzone di rotaia" di 23 centimetri,

⁴ Il Tribunale di Trani, con ordinanza del 7 maggio 2019, ha revocato la precedente ordinanza emessa dal GUP del medesimo Tribunale, ed ha ammesso la costituzione di parte civile nei confronti dell'ente imputato (entrambe le ordinanze sono reperibili in www.giurisprudenzapenale.com, 8 maggio 2019).

⁵ Appare utile richiamare integralmente le dichiarazioni dell'allora Ministro delle infrastrutture e dei trasporti, Graziano Del Rio, rilasciate alla Camera per le comunicazioni del Governo in relazione al tragico incidente ferroviario sulla tratta Andria-Corato: «la sicurezza della tratta coinvolta dall'incidente è regolata tramite consenso telefonico che lascia interamente all'uomo la gestione ed è tra i sistemi meno evoluti e più rischiosi di regolazione della circolazione ferroviaria. Purtroppo un sistema come quello del consenso telefonico che lascia interamente all'uomo la possibilità di intervento è oggi considerato maggiormente a rischio anche se utilizzato su tratte di esercizio compatibili con il sistema. La sicurezza della circolazione ferroviaria nella tratta dove è avvenuto l'incidente, è regolata tramite il meccanismo del consenso telefonico: nel regime del blocco telefonico il capostazione non può inviare un treno alla stazione successiva se non ha domandato e ottenuto dal capostazione della predetta il consenso a inviare quel determinato treno. La sezione di linea è dunque considerata normalmente bloccata, e viene liberata per la circolazione di volta in volta mediante il consenso dell'inoltro del treno; con tale procedura sulla sezione di linea può essere presente un solo treno per volta. Il sistema di segnalamento con consenso telefonico, pur essendo sicuro, è certamente un sistema tra i meno evoluti rispetto alle tecnologie disponibili per la regolazione della circolazione ferroviaria: infatti il sistema si affida interamente all'uomo, nella fattispecie all'operatività dei capistazione, come sopra descritto. Le tecnologie oggi disponibili sono molteplici, e si adattano ai diversi regimi di esercizio in relazione alle caratteristiche della rete, alla frequenza dei convogli e alla velocità di esercizio. Nel caso di specie, sulla tratta a binario semplice in esame, il sistema di consenso telefonico è in uso da oltre sessant'anni: l'attuale frequenza dei convogli è praticamente inalterata da circa dieci anni, durante i quali non si sono evidenziati inconvenienti all'applicazione del sistema. Il sistema, ripeto, è di completa responsabilità della Ferrotramviaria, della società di gestione» (il testo è reperibile in www.agi.it).

fratturatosi nel c.d. “punto zero”, non si sarebbe verificato se solo si fossero prese in seria considerazione le segnalazioni degli addetti alla manutenzione, e se non vi fosse stata una lunga serie di omissioni nella sicurezza, al solo fine di risparmiare i costi di manutenzione per la società RFI⁶.

Ulteriore esempio di disastro colposo è rappresentato dal crollo del viadotto Polcevera di Genova, noto anche come ponte “Morandi”, che il 14 agosto del 2018 causò la morte di 43 persone. La Procura di Genova, come si apprende dalle cronache giudiziarie, ha chiuso di recente le indagini a carico dei vari soggetti (71 persone più le due società Autostrade Aspi e Spea) coinvolti a vario titolo per disastro ed omicidio colposo, oltre che per attentato alla sicurezza dei trasporti, rimozione dolosa di cautele contro infortuni sul lavoro ed omicidio stradale. Dalle indagini emergerebbe un dato inquietante, ossia che «nonostante numerosi segni premonitori, nessuno ha preso decisioni per la messa in sicurezza degli stralli, le parti più critiche del viadotto»; dall’inaugurazione del viadotto, risalente al 1967, «i cavi della pila collassata non sono stati oggetto di alcun sostanziale intervento di manutenzione»⁷.

Nell’incidente della funivia del Mottarone – come negli altri disastri colposi precedentemente accennati – appare pacifico che non vi fosse alcun coinvolgimento dell’intelligenza artificiale (in inglese *artificial intelligence*, nota con l’acronimo *AI*), nel monitoraggio delle possibili criticità delle infrastrutture.

Vero ciò, sembra lecito chiedersi, su di un piano generale, che cosa accadrebbe, invece, se nel prossimo futuro si assistesse ad una concreta “transizione” verso le nuove promettenti tecnologie, come, ad es. l’*IA*.

Il quesito, cui si cercherà di rispondere nel prosieguo, è se un giorno non (molto) lontano l’*IA* potrebbe quanto meno “ridurre” l’incidenza della condotta umana in relazione a disastri colposi come, ad esempio, quelli precedentemente passati in rassegna. E, forse, quel giorno, come vedremo meglio

⁶ Occorre aggiungere che il GUP del Tribunale di Milano, con ordinanza depositata il 2 febbraio 2021, ha dichiarato inammissibile la costituzione di parte civile nei confronti della società imputata (l’ordinanza è reperibile in www.giurisprudenzapenale.com, 3 febbraio 2021).

Più di recente, lo stesso Gup di Milano ha disposto il processo, fissato per il 12 ottobre 2021, oltre che per la società RFI (imputata sulla base del d.lgs. n. 231/2001), per l’ex ad di RFI ed altre 7 persone tra dirigenti, dipendenti e tecnici di RFI, imputati per disastro ferroviario colposo, omicidio colposo plurimo, lesioni colpose e violazione delle normative sulla sicurezza (www.ansa.it, 21 giugno 2021).

⁷ Si veda nella cronaca giudiziaria *Ponte Morandi, chiusa l’inchiesta: 71 indagati. “Manutenzioni non adeguate. Decisioni operative sulla sicurezza della struttura mai prese*, in www.illattoquotidiano.it, 22 aprile 2021.

Più di recente, la Procura di Genova, dopo aver stralciato dieci posizioni, ha chiesto il rinvio a giudizio, oltre che per le due società Aspi e Spea (imputate in base al d.lgs. n. 231/2001), per 59 imputati, gli ex vertici, i dirigenti e tecnici delle due società, nonché funzionari del Ministero dei trasporti e del Provveditorato alle opere pubbliche (www.illattoquotidiano.it, 24 giugno 2021).

in seguito, potrebbe segnare, per così dire, un “nuovo inizio” per la colpa penale, dove la condotta umana potrebbe avere come “parametro” il *super agente modello dell’IA*.

3. Uno sguardo tecnologico: l’IA per infrastrutture smart. L’IA, intesa come l’impiego dei *computer* per simulare le funzioni cognitive degli esseri umani⁸, si fonda su algoritmi che sono una serie di istruzioni precise ed espressioni matematiche che affrontano un dato problema e cercano di risolverlo⁹; con una metafora, potrebbe dirsi che essi esprimono in «formule la complessità del comportamento umano»¹⁰.

Basti considerare gli algoritmi di apprendimento automatico che «identificano relazioni precise nei dati osservati, senza avere regole e modelli espliciti pre-programmati» (c.d. *machine learning*), o gli algoritmi basati su un modello semplificato del cervello umano, le c.d. reti neurali (*neural network*) che di recente, grazie a *computer* più potenti, sono capaci di un apprendimento “profondo” (*deep learning*) che si riferisce «ai molti strati di unità coinvolti nell’algoritmo»¹¹. Quest’ultimo tipo di apprendimento non si limita alle immagini, ma trova applicazione «in casi più impegnativi, come le predizioni»¹². Con una certa enfasi, può affermarsi che disporre di grandi quantità di dati da cui estrarre informazioni ed algoritmi significa oggi avere un dono simile a quello dell’oracolo delfico nell’antica società Grecia: «sono loro gli indovini digitali del nostro presente. Come ai tempi antichi, sono al centro di enormi interessi economici e di potere, e sono diventati oggetto di adorazione o paura. Sempre misteriosi, sia quando risolvono i nostri problemi sia quando indicano futuri scomodi. Terrorizzanti quando ci fanno intuire un mondo dove l’algoritmo non solo legge il futuro, ma lo definisce»¹³.

Non meno importanti per gli algoritmi dell’IA sono i *Big Data*¹⁴, fondamentali

⁸ Le basi concettuali dell’AI furono gettate dal famoso matematico britannico Alan Turing, in un suo scritto intitolato *Computing Machinery and Intelligence*, in *Mind*, 1950, 49, 433 ss.

Sul concetto di intelligenza artificiale, v. BODEN, *L’intelligenza artificiale*, trad. it., Bologna, 2016, 7 ss.

⁹ Per una definizione di algoritmo, v. TOFFALORI, *Algoritmi*, Bologna, 2015, 7.

¹⁰ VESPIGNANI, *L’algoritmo e l’oracolo. Come la scienza predice il futuro e ci aiuta a cambiarlo*, Milano, 2019, 65.

¹¹ VESPIGNANI, *L’algoritmo e l’oracolo*, cit., 66-67; su *machine learning*, *reti neurali* e *deep learning*, v. anche CRESCENZI, PAGLI, *Problemi, algoritmi e coding*, Bologna, 2017. Sugli algoritmi di *machine learning*, v. *funditus* DOMINGOS, *L’algoritmo definitivo. La macchina che impara da sola e il futuro del nostro mondo*, trad. it., Torino, 2016.

¹² VESPIGNANI, *L’algoritmo e l’oracolo*, cit., 76.

¹³ VESPIGNANI, *L’algoritmo e l’oracolo*, cit., 19.

¹⁴ DELMASTRO, NICITA, *Big data*, Bologna, 2019, cui si rinvia anche per gli ulteriori approfondimenti bibliografici.

beni economici nell'era del capitalismo digitale, i quali possono essere intesi nell'accezione di *megadati* (accolta dall'UE), vale a dire di «grandi quantità di tipi diversi di dati prodotti da varie fonti, fra cui persone, macchine e sensori»¹⁵. Il volume di questa enorme massa di dati raccolti e processati nelle piattaforme digitali globali (*over the top*)¹⁶, serve a migliorare l'efficienza degli algoritmi e, «a sua volta, l'uso dell'algoritmo da parte di ciascuno di noi genera nuovi dati, e così via, insegnando all'algoritmo come migliorare e, persino, come imparare ad imparare meglio»¹⁷.

L'efficienza algoritmica dell'*IA* integrata dal grande volume dei *Big Data* – oltre alle tecnologie di corollario come “l'*Internet* delle cose” (che estende la connessione virtualmente a tutte le cose) e la robotica avanzata – sembrerebbero aver avviato la “quarta rivoluzione industriale”, termine coniato dall'economista Klaus Schwab¹⁸, fondatore del *World Economic Forum*.

Tra le svariate applicazioni dell'*IA* nel futuro prevedibile¹⁹ – per quel che rileva maggiormente in questa sede – v'è anche il monitoraggio delle infrastrutture: l'*IA* potrebbe, infatti, “intercettare” quegli “indicatori di sofferenza” delle infrastrutture inizialmente non riconoscibili, ma che hanno una tendenza a crescere a dismisura, mettendo a rischio l'integrità di una struttura.

Per raggiungere elevati *standard* di sicurezza, grazie alle nuove tecnologie, la parola chiave dell'immediato futuro sembra essere la manutenzione “predittiva”, in luogo della manutenzione “correttiva”: ogni infrastruttura potrebbe essere dotata di sensori adatti al monitoraggio “intelligente” (*smart*). L'impiego dell'*IA* potrebbe analizzare i parametri rilevati in modo automatico e continuativo sulle strutture, per elaborare previsioni sui comportamenti dell'opera e indicazioni sugli interventi da effettuare; in tal modo, l'*IA* sarebbe di assai prezioso ausilio ai tecnici del settore per l'individuazione e la soluzione di problemi complessi.

4. (segue) *La proposta di “Regolamento IA” dell'UE*. In ogni caso va osserva-

¹⁵ DELMASTRO, NICITA, *Big data*, cit., 10 ss.

¹⁶ L'espressione inglese “*over the top*” sta a significare che «le piattaforme digitali sviluppano servizi che si trovano gerarchicamente al di sopra delle infrastrutture fisiche di telecomunicazione fisse e mobili grazie alle quali accediamo alla rete» (DELMASTRO, NICITA, *Big data*, cit., 48).

¹⁷ DELMASTRO, NICITA, *Big data*, cit., 11.

¹⁸ SCHWAB, *La quarta rivoluzione industriale*, trad. it., Milano, 2019, secondo cui la quarta rivoluzione industriale sarebbe iniziata nel 2016 (dopo la terza rivoluzione industriale, iniziata nel 1960, nota anche come rivoluzione digitale), grazie in particolare agli algoritmi di *machine learning* dell'*AI* ed alla crescita dei *Big Data*.

¹⁹ Sulle principali applicazioni dell'*AI*, che riguardano i più svariati settori della società v. in generale LONGO, SCORZA, *Intelligenza artificiale*, Milano, 2020, 60-61.

to che il Legislatore europeo con la proposta di “Regolamento *IA*” pubblicata dalla Commissione Europea il 21 aprile 2021, ha introdotto per la prima volta un quadro normativo sistematico cui occorre brevemente accennare²⁰.

Viene, innanzitutto, fornita una pionieristica definizione di sistema di *IA*, che ricomprende i *software* realizzati con tecniche od approcci tra quelli specificati nel relativo Allegato I, tra i quali rientrano gli algoritmi di *machine learning*. Ancor più interessante è la classificazione delle pratiche di *IA* consentite o vietate a seconda dei rischi per i diritti fondamentali e la sicurezza: per quel che maggiormente interessa in questa sede, i sistemi di *IA* relativi a infrastrutture pubbliche essenziali sono qualificate “ad alto rischio”, dal momento che il loro impiego è consentito solo in presenza di specifici controlli di sicurezza, secondo un modello di classificazione incentrato sul rischio associato al prodotto stesso.

In particolare, una prima categoria comprende i sistemi di *IA*, elencati nell’Allegato II del Regolamento, i quali troveranno impiego come componenti di sicurezza dei prodotti (o che sono essi stessi un prodotto). Una seconda categoria riguarda, invece, i sistemi di *IA* a sé stanti, elencati nell’Allegato III, il cui impiego può incidere sui diritti fondamentali.

Il Regolamento *IA* prevede una serie di controlli da effettuare sui sistemi di *IA* ad alto rischio, al fine di garantire elevati *standard* di trasparenza e di sicurezza. A corollario di detta articolata “rete” di controlli sono previsti importanti obblighi di responsabilizzazione del fornitore (come, ad es., la tracciabilità e verificabilità degli *output* di *IA* ad alto rischio per l’intero ciclo di vita), un apposito sistema di gestione del rischio, pratiche di *governance* dei dati che presentino, tra l’altro, proprietà statistiche adeguate a supportare l’uso del sistema di *IA*, e, *last but not least*, un’efficace supervisione umana.

Il Regolamento prevede anche l’istituzione di un “Comitato” europeo con funzioni di consulenza e di assistenza alla Commissione europea e sanzioni amministrative pecuniarie in caso di inosservanza dell’articolata disciplina.

5. *Brevi considerazioni critiche di diritto penale sulla concezione forte dell’IA.* A fronte del futuro scenario di infrastrutture *smart* che potrebbero aiutare a prevedere rischi di disastri colposi (e conseguenti vittime), lo studioso del diritto penale deve preliminarmente soffermarsi sulla contrapposizione tra la concezione *forte* di *IA*, intesa come autonomo soggetto di diritto, e la

²⁰ Per un maggior grado di approfondimento sull’articolata disciplina della proposta di Regolamento *IA*, si rinvia a LUSARDI, *Regolamento UE sull’Intelligenza Artificiale: uno strumento articolato per gestire il rischio*, in *Quot. giur.*, 3 giugno 2021.

concezione *debole*, che afferma, invece, il carattere strumentale dell'*IA* rispetto alla condotta umana²¹.

A differenza dei *software* tradizionali, un sistema di *IA* non si basa sulla programmazione (cioè sul lavoro di sviluppatori che scrivono il codice di funzionamento del sistema), ma su tecniche di apprendimento: vengono dunque definiti algoritmi che elaborano un'enorme quantità di dati dai quali è il sistema stesso che deve derivare le proprie capacità di comprensione e di ragionamento.

Le decisioni dell'*IA* si fondano, pertanto, su processi decisionali che non possono essere compresi da osservatori esterni: è questo il noto problema c.d. di *black box* che si riferisce al fatto che «molti algoritmi prendono un dato in entrata e ne producono uno in uscita, passando attraverso un processo di apprendimento che è una scatola nera non interpretabile dall'esterno»²².

Il *machine learning* ha, inoltre, permesso, agli algoritmi di aggirare il c.d. paradosso di Polanyi che può riassumersi nella frase di buon senso «conosciamo di più di quello che possiamo spiegare», per cui «implicitamente, sappiamo molto sul modo in cui funziona il mondo intorno a noi, ma non siamo in grado di esplicitare questa conoscenza [...]; gli algoritmi di *machine learning*, e in particolare le reti neurali, sono proprio lo strumento con cui i computer acquisiscono un tipo di conoscenza implicita attraverso input esemplificativi, senza però essere in grado di esplicitare il perché dei loro risultati»²³.

In molti casi, inoltre, l'*IA* può apprendere non solo dalle proprie esperienze, ma anche dai suoi “simili”, mediante le tecnologie di *cloud computing* che consentono di incrementare l'apprendimento delle macchine intelligenti che possono sommare le loro “esperienze” operative nei più svariati ambiti applicativi²⁴.

Orbene, la predetta concezione *forte* richiederebbe un diritto penale “robotico”²⁵, in cui l'*IA* verrebbe considerata un soggetto suscettibile di un giudizio di colpevolezza per disastri colposi che si sarebbe potuto prevedere ed evitare, alla luce di un agente “modello”.

Si tratta, è bene rilevarlo, di una concezione suggestiva ma (assai) problemati-

²¹ PIERGALLINI, *Intelligenza Artificiale: da “mezzo” ad “autore” del reato?*, in *Riv. it. dir. proc. pen.*, 2020, 1759.

²² VESPIGNANI, *L'algoritmo e l'oracolo*, cit., 74.

²³ *Ibidem*.

²⁴ Sul punto, v. CAPPELLINI, *Machina delinquere non potest? Brevi appunti su intelligenza artificiale e responsabilità penale*, in *Criminalia*, 27 marzo 2019, 6, cui si rinvia anche per i riferimenti bibliografici.

²⁵ PIERGALLINI, *Intelligenza Artificiale*, cit., 1763 ss.

ca²⁶, poiché l'*IA* non ha una “realtà interiore”, anche solo lontanamente “avvicinabile” alla coscienza (umana), né può comprendere, ad es., il “disvalore” di una condotta illecita, che costituisce una “costruzione” sociale, ancor prima che giuridica. Almeno fino a quando una macchina non saprà pensare, sembra “premature” interrogarsi sulla problematica della “macchina-uomo”, vale a dire il tentativo di descrivere le macchine come esseri umani: secondo il famoso *test* di Turing una macchina saprà pensare solo quando saprà parlare con un uomo e l'uomo non saprà decidere se a parlare con lui è una macchina autonoma o un programmatore “nascosto”²⁷.

Allo stesso modo, sembra destinata a fallire la “prospettiva”, altrettanto insidiosa, dell’“uomo-macchina” (risalente al filosofo illuminista La Mettrie), ossia il tentativo di descrivere gli uomini come se fossero macchine²⁸.

Occorre, al riguardo, evidenziare che l'intelligenza umana non è un «cervello nella vasca», come immaginato dal filosofo Putnam negli anni 80 del secolo scorso²⁹, bensì un'intelligenza “incorporata”, non solo razionale (fredda, distaccata), ma anche “emotiva” (corporea, istintiva); come ipotizzato dal famoso studioso di psicologia comportamentale, Kahneman, sono configurabili due “sistemi” di pensiero, integrati tra loro, che connotano l'intelligenza umana, «pensieri veloci» corrispondenti al «sistema 1» (“automatico”), e «pensieri lenti» che attengono al «sistema 2» (“razionale”)³⁰. Quindi, per avere un sistema di *IA* che riesca ad imitare l'uomo occorrerebbe sviluppare non solo la parte “mentale”, ma anche la parte “corporea” della macchina.

Come ha scritto il noto scienziato Federico Faggin, «oggi leggiamo che l'intelligenza artificiale, basata sulle reti neurali artificiali, potrà presto sorpassare le capacità umane. Queste previsioni sono sostenute da risultati impressionanti, come la sconfitta da parte di un computer anni fa del campione mondiale di scacchi e, più recentemente, del campione mondiale di go, un gioco cinese ritenuto ancora più immaginativo e quindi più “umano” degli scacchi. Alcuni scienziati prevedono addirittura che in meno di quarant'anni il computer diventerà cosciente. Questi risultati e previsioni di esperti vanno contro il buon senso di molte persone e sfidano la nostra comprensione intuitiva della vita, della consapevolezza e della natura della realtà. La scienza deve urgentemente affrontare queste tematiche, che nel passato sono state studiate

²⁶ In questo senso, cfr. PIERGALLINI, *Intelligenza Artificiale*, cit.

²⁷ Alan Turing, descrive il suo famoso *test* in *Computing Machinery and Intelligence*, cit.

²⁸ LA METTRIE, *L'uomo macchina*, trad. it., Milano, 2015.

²⁹ Per la discussione filosofica dello scenario del “cervello nella vasca”, v. PUTNAM, *Ragione, verità e storia*, trad. it., Milano, 1985.

³⁰ V. funditus, KAHNEMAN, *Pensieri lenti e veloci*, trad. it., Milano, 2012.

quasi esclusivamente dalla filosofia. Che cos'è che ci distingue dalle macchine? Siamo semplicemente dei computer biologici? Secondo la maggior parte degli scienziati, sì. Basandomi sulle mie esperienze dirette e sulle mie riflessioni, io penso di no»³¹.

Il passo richiamato è significativo perché esprime il punto di vista di uno scienziato “in minoranza” rispetto alla concezione “materialista” che sembra, invece, equiparare *tout-court* il cervello umano ad un computer.

A parere di chi scrive, proprio la mancanza di una forma di “coscienza” dell'IA, oltre che di un'intelligenza “corporea”, milita contro la concezione forte, con la conseguenza che la suggestione del codice penale “robotico” sembra destinata a rimanere tale.

Per richiamare uno stimolante contributo del Burchard, professore di Diritto penale dell'Università *Goethe* di Francoforte sul Meno, è possibile rispondere negativamente al quesito se l'IA e la conseguente trasformazione algoritmica della società, implichi la fine del diritto penale³².

6. (segue) *La concezione debole dell'IA e le implicazioni penalistiche per la colpa.* Maggiormente “fruibile” per il diritto penale appare la concezione debole, secondo cui l'IA opererebbe come un sistema in ausilio dell'uomo, senza avere la capacità di delinquere o di essere punita (*machina delinquere et puniri non potest*)³³.

Questo non significa che debbano essere minimizzate le promesse dell'IA che si tradurrebbero in una “maggiore” garanzia del bene giuridico della sicurezza ed integrità pubblica, oltre che della efficienza dei controlli (con particolare riguardo alle infrastrutture).

Si tratta, più realisticamente, di prendere atto che l'IA potrebbe rappresentare uno strumento per ridurre il rischio di pericoli alla sicurezza ed incolumità degli utenti delle infrastrutture che in Italia scontano, in media, ritardi “atavici” nei controlli (come dimostrato, ad es. dal crollo del Ponte “Morandi”).

È di palmare evidenza che una riduzione complessiva del rischio, grazie anche all'impiego dell'IA, comporterebbe, potenzialmente, una riduzione dell'incidenza della condotta umana colposa (commissiva od omissiva) nella manutenzione delle infrastrutture. Con questo non si vuol sostenere, ovvia-

³¹ FAGGIN, *Silicio. Dall'invenzione del microprocessore alla nuova scienza della consapevolezza*, Milano, 2019, 187-188.

³² BURCHARD, *L'intelligenza artificiale come fine del diritto penale? Sulla trasformazione algoritmica della società*, in *Riv. it. dir. proc. pen.*, 2019, 1909 ss.

³³ Cfr. CAPPELLINI, *Machina delinquere non potest?*, cit.,

mente, che non si celebrerebbero più processi per disastro ed omicidio colposo, ma piuttosto, che potrebbe assistersi ad una diminuzione della probabilità di verificazione di simili reati che potrebbe tradursi, in termini assoluti, in una “contrazione” rispetto alla frequenza con la quale si verificano.

Ciò potrebbe contribuire anche a “rassenerare” la comunità dei consociati che troppo spesso vede inascoltate le proprie richieste di controlli sulle criticità di infrastrutture, a causa della carenza di manutenzione da parte delle Autorità preposte.

Basti citare, ad esempio, il crollo di un ponte sul Magra nella Lunigiana vicino ad Aulla, che ha destato l’interesse anche dei *mass media*; il crollo si verificò l’8 aprile del 2020 e coinvolse due autisti, uno illeso e l’altro ricoverato a causa di lesioni, ma non in pericolo di vita. La tragedia poteva assumere proporzioni di gran lunga maggiori se in quel periodo non avesse operato il blocco degli spostamenti per effetto del *lockdown* da emergenza sanitaria SARS-CoV-2.

Al di là di questo ulteriore aspetto, il dato inquietante per lo studioso del diritto penale è che, nonostante le plurime segnalazioni di cittadini, e del sindaco di Aulla, sulla doverosità di ulteriori controlli da parte dell’Anas, considerate le criticità del ponte, quest’ultima sostenne, con un comunicato dell’agosto 2019 (oggetto peraltro di interrogazione parlamentare), che il viadotto in questione già attenzionato e sorvegliato da personale Anas non presentasse al momento criticità tali da compromettere la sua funzionalità statica, e che, sulla base di ciò, non fossero giustificati provvedimenti emergenziali per il viadotto stesso³⁴.

L’IA potrebbe integrare (*rectius*: sostituirebbe) il parametro dell’agente modello, con la conseguenza che la condotta dell’agente reale “conforme” ad esso renderebbe inesigibile una condotta alternativa, a meno che il modello previsionale delineato dall’IA sulla sicurezza dell’infrastruttura non si riveli inattendibile, a causa di errori della macchina nella fase di apprendimento (tramite algoritmi di *machine learning*).

A fronte di una simile evenienza, la colpa dell’uomo potrebbe comunque ritenersi “lieve”, se non del tutto esclusa, alla luce dell’affidamento riposto nelle maggiori capacità di carattere tecnico e calcolatorio della macchina³⁵, in quali-

³⁴ Si veda *Camera dei Deputati - 4 - 05188 - Interrogazione sulla grave vicenda della caduta del ponte sul fiume Magra nonostante precedenti segnalazioni di evidenti criticità all’Anas*, in www.infoparlamento.it, 10 aprile 2020.

³⁵ Su di un piano puramente filosofico, potrebbe forse congetturarsi che la calcolabilità degli algoritmi di *machine learning* appare, per certi versi, in grado di realizzare pienamente quell’idea di calcolo inteso da Max Weber come strumento di precisione della razionalità economica e del capitalismo occidentale

tà di “nuovo” agente modello; residuerebbe, tuttavia, un margine di colpa “grave” quando l’errore dell’IA sia stato determinato da imperizia umana (ad es., per scarsa conoscenza di statistica) nell’interpretazione dei modelli previsionali.

Questi possibili “riverberi” sulla colpa penale, potrebbero prestarsi a rilievi critici relativi al fatto che la predetta concezione *debole* non sarebbe in grado di giustificare una “qualificazione” in termini di (*super*) agente modello dell’IA, poiché il parametro di riferimento dell’agente reale deve essere pur sempre “umano” – anche se, in certi ambiti, “idealizzato” dalla stessa giurisprudenza, a tal punto da apparire “inavvicinabile” all’uomo³⁶. Si tratta, in ogni caso, di rilievi che sembrano muovere più da un argomento formale, che sostanziale ed, in definitiva, non appaiono insormontabili per *il diritto penale che verrà*, dove la condotta umana dovrà comunque rapportarsi alle maggiori capacità tecniche dell’IA. La cooperazione, per così dire, tra l’agire (creativo) umano e le “macchine intelligenti”³⁷, pare ragionevolmente destinato ad incrementare lo *standard* di cautela nella gestione e nei controlli delle infrastrutture. L’esito di una simile cooperazione appare orientata a beneficio esclusivo degli uomini, come già intuito dal famoso scrittore di fantascienza Isaac Asimov nelle sue tre celebri “leggi” della robotica³⁸.

7. *Conclusioni* Avviandoci alle conclusioni di queste brevi osservazioni che “guardano” all’IA da una prospettiva “differente” da quella della (ormai) nota ed ampia problematica della giustizia predittiva³⁹, non sembra fuor di luogo

(sulla “calcolabilità weberiana”, v. *funditus* IRTI, *Un diritto incalcolabile*, Torino, 2016, spec. 33 ss.

³⁶ Emblematico il riferimento, ad esempio, al campo all’attività medica-chirurgica, dove un orientamento dottrinale ha evidenziato il rischio che, sotto l’agente modello invocato dalla giurisprudenza, si nasconderebbe «una pretesa di diligenza forgiata sul parametro estremo del medico *Gottsähnlich*, iperbolico, e non avvicinabile dal medico concreto», per cui l’agente modello dell’*homo eiusdem professionis et conditionis* «finirebbe per ripeterne i limiti di inarrivabilità» (CAPUTO, *Colpa penale del medico e sicurezza delle cure*, Torino, 2017, 92).

³⁷ Sui rapporti in generale tra l’intelligenza umana e l’IA, cfr. MCAFEE, BRYNJOLFSSON, *La nuova rivoluzione delle macchine*, Milano, 2015, spec. 90 ss.

³⁸ ASIMOV, *Trilogia della fondazione*, trad. it., Milano, 2004, dove sono affermate le tre famose “leggi” della robotica: 1. Un robot non può recare danno a un essere umano o permettere con l’inazione che un essere umano possa essere danneggiato. 2. Un robot deve ubbidire agli ordini degli esseri umani, tranne quando tali ordini entrano in conflitto con la Prima legge. 3. Un robot deve proteggere la propria esistenza finché tale protezione non entri in conflitto con la Prima e la Seconda legge.

³⁹ Per una disamina dei profili penalistici dell’IA, con particolare riferimento alla differente problematica della giustizia predittiva, v., *ex multis*, F. BASILE, *Intelligenza artificiale e diritto penale: quattro possibili percorsi di indagine*, in *Riv. Dir. pen. uomo*, 10, 2019, 44 ss.; DI GIOVINE, *Il judge-Bot e le sequenze giuridiche in materia penale (intelligenza artificiale e stabilizzazione giurisprudenziale)*, in *Cass. pen.*, 2020, 951 ss.; MAGRO, *Decisione umana e decisione robotica. Un’ipotesi di responsabilità da procrea-*

richiamare alcune interessanti considerazioni di un matematico francese, il quale, in un noto testo di divulgazione scientifica, scriveva che: «i computer, muniti di algoritmi sempre più complessi e performanti, sembrano oggi in grado di superare gli esseri umani nella maggior parte delle loro competenze. Guidano automobili, partecipano a operazioni chirurgiche, possono creare musica o dipingere quadri originali. Difficile immaginare un'attività umana che, da un punto di vista tecnico, non possa essere realizzata da una macchina pilotata da un algoritmo adeguato. [...]. Difficile, oggi, pronosticare di quali prodezze saranno artefici le macchine di domani. Sarebbe solo sorprendente il fatto che non ci sorprendano»⁴⁰.

Non v'è dubbio che, tra le attività tecniche, i controlli delle infrastrutture, nell'immediato futuro, potrebbero costituire un importante campo di applicazione dell'*IA*, con conseguente probabile “riduzione” dell'incidenza della condotta umana colposa, come accennato in precedenza.

Non è da escludere, inoltre, che *il diritto penale che verrà* potrebbe essere meno “totale” di quanto appaia oggi allo studioso del diritto penale⁴¹, dal momento che l'impiego della potente capacità di calcolo degli algoritmi nel monitorare le infrastrutture potrebbe, per certi versi, “attenuare” la convinzione che nel diritto penale si possa trovare (necessariamente) la “soluzione”, a fronte di un prospettato disastro colposo. In altri termini, lo stigma penale potrebbe non configurarsi in relazione ad una condotta umana “conforme” al *super agente* modello.

Eppure, se diamo un'occhiata al recente Piano nazionale di ripresa e resilienza (PNRR) trasmesso dal Governo all'UE, in attuazione al programma *Next Generation EU*, non sembra che il nostro Paese voglia investire seriamente sull'*IA*⁴², considerata sia la scarsità delle risorse, sia l'assenza di un istituto di ricerca *ad hoc*, sul modello dell'*Alan Turing Institute* britannico.

zione robotica, in www.lalegislazionepenale.eu, 10 maggio 2020; CANZIO, *Intelligenza artificiale, algoritmi e giustizia penale*, in www.sistemapenale.it, 8 gennaio 2021; *funditus* MANES, *L'oracolo algoritmico e la giustizia penale: al bivio tra tecnologia e tecnocrazia*, in RUFFOLO (a cura di), *Intelligenza artificiale - il diritto, i diritti, l'etica*, Milano, 2020; GARAPON, LASSEGUE, *La giustizia digitale. Determinismo tecnologico e libertà*, trad. it., Bologna, 2021; più di recente, MAUGERI, *L'uso di algoritmi predittivi per accertare la pericolosità sociale: una sfida tra evidence based practices e tutela dei diritti fondamentali*, in *Arch. pen.* (web), 1, 2021. Con riferimento al diritto processuale, volendo, v. CASTELLI, PIANA, *Giusto processo e intelligenza artificiale*, Rimini, 2019; NIEVA-FENOLL, *Intelligenza artificiale e processo*, trad. it., Torino, 2019.

⁴⁰ LAUNAY, *Il grande romanzo della matematica*, trad. it., Milano, 2019, 150.

⁴¹ V. SGUBBI, *Il diritto penale totale*, Bologna, 2019.

⁴² In questo senso, si veda DELLA PORTA, *PNRR, l'intelligenza artificiale dov'è? Una dimenticanza pericolosa*, in www.agendadigitale.eu, 11 febbraio 2021.

Al momento, si può solo sperare (*sic*) che il D.L. 73/2021 *Sostegni-bis*⁴³, recante misure urgenti connesse all'emergenza da SARS-CoV-2, per le imprese, il lavoro, i giovani, riesca a stanziare i fondi necessari per attivare l'Istituto Italiano per l'intelligenza artificiale (*I3A*).

⁴³ Il citato Decreto *Sostegni bis* è stato pubblicato sulla G.U. del 25 maggio del 2021.