22/10/2024

IA e indagini penali: nuove prospettive e vecchie soluzioni?

La recente entrata in vigore dell'AI ACT e la possibile imminente approvazione del ddl destinato a integrare a livello nazionale la normativa nel settore suggeriscono di affrontare alcune tematiche di carattere generali legate al concetto di sistema di intelligenza artificiale, alla natura degli accertamenti su tali sistemi in ambito penale e alla rilevanza di specifici aspetti legati a tali accertamenti.

di Cesare Parodi

SOMMARIO

Premessa: la definizione di sistema di IA

Gli accertamenti tecnici ex art. 359 c.p.p. e la necessità di procedere a incidente probatorio

La necessità e i limiti dell'accertamento

Oggetto e strumenti di accertamento: i diritti della difesa

Infine: da dove deriva il “learning”?

In conclusione

Premessa: la definizione di sistema di IA

Per molto, moltissimo tempo dovremo dedicarci ad analizzare e interpretare le nuove disposizioni in tema di intelligenza artificiale (d'ora in poi IA) derivante dall'entrata in vigore del Regolamento (UE) 2024/1689 del Parlamento europeo e del Consiglio del 13 giugno 2014 che stabilisce regole armonizzate sull'intelligenza artificiale e modifica i regolamenti (CE) n. 300/2008, (UE) n. 167/2013, (UE) n. 168/2013, (UE) 2018/858, (UE) 2018/1139 e (UE) 2019/2144 e le direttive 2014/90/UE, (UE) 2016/797 e (UE) 2020/1828 (regolamento sull'intelligenza artificiale) – nonché – ovviamente – della normativa italiana destinata a dare concreta attuazione allo stesso (si rimanda sul punto a C. Parodi, *Sistema penale e ddl sull'intelligenza artificiale: prospettive e criticità*, in *IUS Penale*, 13 Maggio 2024).

Nondimeno, senza entrare nel dettaglio dei singoli – innumerevoli – aspetti che dovranno essere affrontati con riguardo all'utilizzo di programmi di IA nell'indagine penale e al contrasto di forme di IA per porre in essere, in vario modo, condotte illecite – possono essere enucleate alcune tematiche di generale – verrebbe da dire generalissima – natura sottese in via prioritaria alle due "facce" del problema.

Si tratta, prima di tutto, della definizione del **concetto di sistema di intelligenza artificiale** che il ddl ha ripreso integralmente dal testo del regolamento: «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».

Abbiamo già manifestato (C. Parodi, *Sistema penale e ddl sull'intelligenza artificiale: prospettive e criticità*) quale perplessità su una formula assolutamente "centrale" e indispensabile per delineare il quadro oggetto di analisi. In realtà, confrontando la definizione con il considerando n. 12 del Regolamento menzionato, il quadro appare più chiaro: «…la definizione dovrebbe essere basata sulle principali caratteristiche dei sistemi di IA, che la distinguono dai tradizionali sistemi software o dagli approcci di programmazione più semplici, e non dovrebbe riguardare i sistemi basati sulle regole definite unicamente da persone fisiche per eseguire operazioni in modo automatico. Una caratteristica fondamentale dei sistemi di IA è la loro capacità inferenziale. Tale capacità inferenziale si riferisce al processo di ottenimento degli output, quali previsioni, contenuti, raccomandazioni o decisioni, che possono influenzare gli ambienti fisici e virtuali e alla capacità dei sistemi di IA di ricavare modelli o algoritmi, o entrambi, da input o dati. Le tecniche che consentono l'inferenza nella costruzione di un sistema di IA comprendono approcci di apprendimento automatico che imparano dai dati come conseguire determinati obiettivi e approcci basati sulla logica e sulla conoscenza che traggono inferenze dalla conoscenza codificata o dalla rappresentazione simbolica del compito da risolvere. La capacità inferenziale di un sistema di IA trascende l'elaborazione di base dei dati consentendo l'apprendimento, il ragionamento o la modellizzazione. Il termine «automatizzato» si riferisce al fatto che il funzionamento dei sistemi di IA prevede l'uso di macchine…. I sistemi di IA sono progettati per funzionare con livelli di autonomia variabili, il che significa che dispongono di un certo grado di autonomia di azione rispetto al coinvolgimento umano e di capacità di funzionare senza l'intervento umano. L'adattabilità che un sistema di IA potrebbe presentare dopo la diffusione si riferisce alle capacità di autoapprendimento, che consentono al sistema di cambiare durante l'uso…».

Sul punto si è soffermata di recente la Circolare n. 14 del 26 luglio 2024 di ASSONIME "Il regolamento europeo sull'intelligenza artificiale: analisi ragionata della nuova disciplina e prospettive di policy per le imprese". Il documento di ASSONIME contribuisce indubbiamente a delineare la portata della definizione, laddove precisa che «Il tratto distintivo di un sistema di intelligenza artificiale, nell'accezione fatta propria dal regolamento, è la capacità inferenziale, di generare risultati che possono influenzare gli ambienti fisici o virtuali in cui insistono attraverso un'interazione dinamica e con livelli di autonomia variabili. Il regolamento non riguarda i software tradizionali, programmati per

regolamento non riguarda i software tradizionali, programmati per svolgere un compito in una logica di automazione più o meno evoluta ('if-then')».

In tale prospettiva le "condizioni" per il riconoscimento di un sistema di IA sono state così individuate; un sistema è tale se (art. 3 n. 1 e considerando 12):

a) automatizzato (nella versione inglese 'machine-based' in quanto 'run(s) on machines');

b) progettato per funzionare con livelli di autonomia variabile rispetto all'intervento/coinvolgimento umano;

c) variamente adattabile dopo la diffusione/messa in funzione, ossia capace di adeguarsi al contesto dinamico in cui si inserisce durante l'uso (come accade ad esempio per i sistemi che generano raccomandazioni personalizzate sulla base delle preferenze);

d) caratterizzato da capacità inferenziale, ossia di generare, a partire dagli input che riceve, per obiettivi impliciti o espliciti, output – quali contenuti, previsioni, raccomandazioni o decisioni – in grado di influenzare ambienti fisici o virtuali.

Le inferenze sono strettamente legate all'impiego, nella costruzione del sistema, di tecniche di machine learning (che imparano dai dati come conseguire determinati obiettivi) e "approcci basati sulla logica e sulla conoscenza" (che traggono inferenze dalla conoscenza codificata o dalla rappresentazione simbolica del compito da risolvere), che consentono l'apprendimento, il ragionamento o la modellizzazione.

Inoltre «secondo quanto chiarito nei considerando, il regolamento non riguarda i software tradizionali, gli approcci di programmazione più semplici, i sistemi che eseguono automaticamente operazioni in base a regole predefinite dall'uomo. In sostanza, è esclusa la programmazione statica o deterministica (riconducibile all'approccio 'if-then'), contrapposta a quella statistico-probabilistica».

Una quadro definitorio certamente molto ampio, non solo sul piano della indicazione strettamente "tecnologica" ma anche su quello della prospettiva di utilizzo dei sistemi di IA; come chiarisce sempre il considerando n. 12 del Regolamento "I sistemi di IA possono essere utilizzati come elementi indipendenti (stand-alone) o come componenti di un prodotto, a prescindere dal fatto che il sistema sia fisicamente incorporato nel prodotto (integrato) o assista la funzionalità del prodotto senza esservi incorporato (non integrato)". Pertanto, come precisato dalla Circolare ASSONIME «Le applicazioni interessate sono comunque molteplici: dal riconoscimento vocale e facciale alla computer vision, all'elaborazione del linguaggio naturale (natural language processing-NLP), ai decision support system, ai sistemi robotici intelligenti».

Il punto che dovrebbe "differenziare" un (si fa per dire) banale programma da un sistema di IA parrebbe, pertanto, individuabile non tanto nella capacità di determinare "output" (che, indubbiamente, sono determinati anche da "sistemi" basati su una logica di automazione più o meno evoluta) quanto dalla capacità inferenziale, ossia di "alimentarsi" in base agli elementi – di varia natura – immessi o considerati dal sistema. Una logica, pertanto, induttiva e non deduttiva.

Tale aspetto – che deriva in termini inequivoci dalla lettura congiunta della definizione e nel considerando del Regolamento sopra riportati – ha dei riflessi sugli accertamenti che, nell'ambito di un procedimento penale, devono (e sempre più dovranno) essere svolti, specie dal momento in cui il ddl sull'intelligenza artificiale introdurrà nel sistema nuove aggravanti, nuove ipotesi di reato e modificherà alcune fattispecie, sul presupposto della "presenza" tra gli elementi della condotta dell'utilizzo di sistemi di IA? Indubbiamente si: vediamo per quale ragione.

Gli accertamenti tecnici ex art. 359 c.p.p. e la necessità di procedere a incidente probatorio

Come è noto, in base all'art. 359, comma 1, c.p.p. «Il pubblico

Come è noto, in base all'art. 359, comma 1, c.p.p. «Il pubblico ministero, quando procede ad accertamenti, rilievi segnaletici, descrittivi o fotografici e ad ogni altra operazione tecnica per cui sono necessarie specifiche competenze, può nominare e avvalersi di consulenti, che non possono rifiutare la loro opera».

Non possono esservi dubbi sul fatto che il P.M. deve avvalersi di un c.t. – nel momento in cui si trova nella necessità di:

- individuare la natura di un sistema informatico
- verificarne le modalità di funzionamento e le ricadute sull'attività per il quale lo stesso è utilizzato
- accertare (in casi di "machine learning") natura e provenienza delle informazioni attraverso le quali il sistema è "alimentato".

Indubbiamente, la P.G. specializzata – in alcuni casi – potrebbe assumere una forma di "supplenza" dell'attività del c.t., ma si deve ritenere che a fronte di sistemi di particolare complessità e delicatezza la scelta di procedere in esito alla nomina di un "esperto" potrà presentarsi come sostanzialmente obbligata.

Nel momento in cui il c.t. dovrà esaminare un programma "tradizionale" (per capirci, quelli che "eseguono automaticamente operazioni in base a regole predefinite dall'uomo") si tratterà di un'analisi "statica" con oggetto (in certi casi, indubbiamente complesso) un algoritmo non solo sviluppato dal suo autore, ma le cui funzioni di elaborazione si presentano in termini "statici": l'analisi di tali programmi non è – pertanto – correlata a un dato temporale.

Non così, evidentemente, per i sistemi di intelligenza artificiale che fanno della potenzialità di funzionare in base a un'interazione dinamica e con livelli di autonomia variabili. Tali sistemi – ontologicamente – non potranno essere costanti e identici nel tempo, ma sono destinati a una naturale implementazione, che potrà avvenire con tempi ed impatti differenti, ma connaturata alla natura del sistema.

Se è così, occorre considerare la possibilità (*rectius*, la necessità) di considerare che oggetto dell'accertamento tecnico è un "fenomeno" dinamico e non statico. Un po' come il fiume di Eraclito, un sistema di IA non è mai – e non sarà mai – uguale a sé stesso: dunque, chi si "cala" in una analisi prodromica a una valutazione, deve tenere conto di tale aspetto.

Ciò non significa che – in tutti i casi – si dovrà fare ricorso alla procedura di cui all'art 360 c.p.p. e – ove necessario – alla richiesta di incidente probatorio ex art. 392 c.p.p., ma che tale prospettiva deve essere considerata, laddove «gli accertamenti previsti dall'articolo 359 riguardano persone, cose o luoghi **il cui stato è soggetto a modificazione…**». Si tenga conto, in tale prospettiva, delle conseguenze di una valutazione errata del P.M., al riguardo, di cui all'art. 392 c.p.p., a fronte delle richieste della difesa: «Qualora, prima del conferimento dell'incarico, la persona sottoposta alle indagini formuli riserva di promuovere incidente probatorio, il pubblico ministero dispone che non si proceda agli accertamenti salvo che questi, se differiti, non possano più essere utilmente compiuti…. se il pubblico ministero, malgrado l'espressa riserva formulata dalla persona sottoposta alle indagini e pur non sussistendo le condizioni indicate nell'ultima parte del comma 4, ha ugualmente disposto di procedere agli accertamenti, i relativi risultati non possono essere utilizzati nel dibattimento».

Se i criteri del percorso ermeneutico sono chiari, resta da sviscerare il significato dell'espressione "persone, cose o luoghi il cui stato è soggetto a modificazione". È solo una questione di strumenti a disposizione dell'osservatore, indispensabili per cogliere e eventualmente "misurare" tali modificazioni, atteso che proprio il tempo è spesso la variabile che ci consente di individuare il "quantum" della modificazione. Il punto è che la formula della norma non può essere intesa in senso puramente meccanicistico o filosofico, in quanto, in entrambi i casi, per difetto o per eccesso, la stessa non risulterebbe funzionale alle finalità che la stessa è chiamata a realizzare.

Può essere utile un confronto rispetto alla realtà imprenditoriale in generale. Laddove la rispondenza di un impianto o stabilimento sottoposto a sequestro ad una serie di disposizioni in materia

ambientale possa essere svolto solo non intervenendo sul medesimo, occorre porsi il problema se – di nuovo – sia ipotizzabile una stasi della fruizione dinamica del complesso di beni sino al momento di un – ipotetico – dibattimento. È verosimile ritenere che proprio la natura necessariamente dinamica di un complesso di beni finalizzati all'esercizio di impresa debbano essere considerati necessariamente sottoposti a modifica – connaturata alla fruizione dei medesimi e che pertanto un accertamento sugli stessi possa essere svolte con le forme dell'art.360 c.p.p. o – più frequentemente – dell'accertamento peritale ex art. 392 c.p.p.

Proviamo a trasporre nel settore di interesse tali principi. L'accertamento su un sistema di IA può essere finalizzato alla ricostruzione di responsabilità su un evento specifico e determinato ovvero in generale sulle funzionalità del programma. Nel primo caso sarà necessario un sequestro informatico in grado di "cristallizzare" il sistema, di modo che l'accertamento tecnico sullo stesso potrà anche svolgersi con le forme dell'art. 359 c.p.p. (questo non esclude altre problematiche altrettanto serie: le vedremo al punto successivo).

Nel caso, al contrario, di un accertamento sulla funzionalità o comunque finalizzato ad analizzare il sistema su un arco temporale, indubbiamente si pone il problema della modificabilità dello stesso.

La possibilità per il pubblico ministero di disporre accertamenti ex art. 360 c.p.p. su cose il cui stato sia soggetto a modificazione – o eventualmente di formulare istanza di incidente probatorio – non può essere contestata con giudizio *ex post,* sulla base del concreto risultato delle indagini, ma va valutata *ex ante* sulla base di una razionale previsione effettuata nel momento in cui l'accertamento stesso viene disposto. Ciò che conta non è la modificazione in sé, quanto la modificazione che può pregiudicare l'esito dell'accertamento che su "persone, cose o luoghi" deve essere svolto: tanto maggiore è la rilevanza della modificazione quanto maggiore è l'incidenza sul potenziale esito della valutazione.

In questa prospettiva, il concetto di modificazione deve essere inteso in via prioritaria in chiave "naturalistica", del medesimo devono essere valutato altre e differenti prospettive ermeneutiche. È modificabile anche un bene che, pur non essendo direttamente ed immediatamente mutevole sul piano astrattamente naturalistico, prevede **una fruizione tale da alterarne, almeno in parte, la natura**. Nel caso di specie, proprio la costante implementazione del programma tramite il meccanismo del *machine learning* impone di considerare l'accertamento "dinamico" sul sistema tale da suggerire le forme dell'incidente probatorio.

La necessità e i limiti dell'accertamento

In realtà, preso atto della potenziale necessità di richiedere un approfondimento tecnico con le forme dell'incidente probatorio, si pone un problema di ancora maggiore complessità e – soprattutto – che impone un'analisi in una duplice prospettiva. Sia il P.M., per l'accertamento delle responsabilità penali, sia la difesa, per la verifica sugli strumenti di indagini utilizzati dalla P.G. come dal P.M. con riguardo a un sistema di IA hanno una concreta e ineludibile esigenza di conoscere (attraverso i propri consulenti) nel dettaglio il "meccanismo di funzionamento del sistema" (lo stesso problema potrà averlo, nel caso, il perito nominato dal giudice).

Si può trattare di sistemi sostanzialmente di elevata complessità, verosimilmente frutto di un'attività di ricerca e sviluppo durata anni e tale da avere determinato elevati costi, la cui conoscenza nel dettaglio deve essere considerata oggetto di tutela. I produttori di tali sistemi potrebbero legittimamente avere dei timori a "disvelare" la struttura degli stressi e potrebbero rifiutare forme di collaborazione con le parti interessate.

Sul piano giuridico, il sistema di IA deve essere ricondotto al concetto di *software* in quanto tale è tutelabile come opera d'ingegno, "protetta" ai sensi della l. n. 633/1941, il cui art. 2 comma 8 prevede che sono tutelati «i programmi per elaboratore in qualsiasi forma espressi, purché originali quale

risultato di creazione intellettuale dell'autore». Pur non potendo entrare nel dettaglio del tema nella presente sede, non sono mancate prospettive ermeneutiche dirette a riconoscere tutela ai sensi della disciplina dei brevetti laddove il software sia inserito nell'ambito di una invenzione, ossia laddove si faccia uso di un elaboratore programmato in modo da ottenere un risultato tecnico nuovo e originale tramite il coordinamento di elementi e mezzi già conosciuti. Una situazione – quest'ultima – certamente ravvisabile anche attraverso sistemi di IA "applicati" a contesti produttivi o di servizi. In questo senso, sia l'art. 45 del codice della proprietà industriale italiano (decreto legislativo 10 febbraio 2005, n. 30) che l'art. 52 (Invenzioni brevettabili), paragrafo 2, punto c), della Convenzione sul brevetto europeo escludono che il software possa essere brevettato; possono essere verosimilmente ammesse alcune eccezioni, potendo risultare brevettabile le c.d. *computer implemented inventions* (software capaci di produrre un effetto tecnico ulteriore quali ad es. come un robot aspirapolvere) laddove non sono brevettabili software gestionali (es. motori di ricerca).

Il significato è il valore dei sistemi di AI sono elementi assolutamente presenti – in termini indiretti ma inequivoci – nello scenario delineato dal Regolamento sopra menzionato. In questo senso vale la pena di richiamare, prima di tutto, tre considerando:

88. Lungo la catena del valore dell'IA, spesso più parti forniscono sistemi di IA, strumenti e servizi, ma anche componenti o processi, che sono integrati dal fornitore nel sistema di IA con varie finalità, inclusi l'addestramento dei modelli, la riqualificazione dei modelli, la prova e la valutazione dei modelli, l'integrazione nel software o altri aspetti dello sviluppo dei modelli. Tali parti svolgono un ruolo importante nella catena del valore nei confronti del fornitore del sistema di IA ad alto rischio in cui i loro sistemi di IA, strumenti, servizi, componenti o processi sono integrati e dovrebbero fornire a tale fornitore mediante accordo scritto le informazioni, le capacità, l'accesso tecnico e qualsiasi altra forma di assistenza necessari sulla base dello stato dell'arte generalmente riconosciuto, **al fine di consentire al fornitore di adempiere pienamente gli obblighi di cui al presente regolamento, senza compromettere i propri diritti di proprietà intellettuale o segreti commerciali**.

107. Al fine di aumentare la trasparenza sui dati utilizzati nelle fasi di pre-addestramento e addestramento dei modelli di IA per finalità generali, compresi testo e dati protetti dalla normativa sul diritto d'autore, è opportuno che i fornitori di tali modelli elaborino e mettano a disposizione del pubblico una sintesi sufficientemente dettagliata dei contenuti utilizzati per l'addestramento del modello di IA per finalità generali. **Pur tenendo debitamente conto della necessità di proteggere i segreti commerciali e le informazioni commerciali riservate**, la presente sintesi dovrebbe essere di respiro ampio e generale, anziché dettagliata sotto il profilo tecnico, al fine di agevolare le parti con interessi legittimi, compresi i titolari dei diritti d'autore, nell'esercitare e far rispettare i loro diritti ai sensi del diritto dell'Unione, ad esempio elencando le principali raccolte o serie di dati che sono state inserite nell'addestramento del modello, quali grandi banche dati o archivi di dati privati o pubblici, e fornendo una descrizione delle altre fonti di dati utilizzate.

Particolarmente significativo risulta il considerando 167, che espressamente affronta (sebbene non in chiave penale) il tema della cooperazione con le autorità preposte ai controlli del settore:

167. Al fine di garantire una cooperazione affidabile e costruttiva delle autorità competenti a livello dell'Unione e nazionale, è opportuno che tutte le parti coinvolte nell'applicazione del presente regolamento rispettino la riservatezza delle informazioni e dei dati ottenuti nell'assolvimento dei loro compiti, in conformità del diritto dell'Unione o nazionale. Dovrebbero svolgere i loro compiti e le loro attività in modo **da proteggere, in particolare, i diritti di proprietà intellettuale, le informazioni commerciali riservate e i segreti commerciali**, l'efficace attuazione del presente regolamento, gli interessi pubblici e di sicurezza nazionale, l'integrità del procedimento penale o amministrativo e l'integrità delle informazioni classificate.

Principi trasfusi in singoli articoli del Regolamento, laddove all'art.

25 si precisa «…la necessità di rispettare e proteggere i diritti di proprietà intellettuale, le informazioni commerciali riservate e i segreti commerciali conformemente al diritto dell'Unione e nazionale», nonché in particolare all'art. 78 (Riservatezza) per il quale «in conformità del diritto dell'Unione o nazionale, la Commissione, le autorità di vigilanza del mercato e gli organismi notificati, nonché le altre persone fisiche o giuridiche che partecipano all'applicazione del presente regolamento, garantiscono la riservatezza delle informazioni e dei dati ottenuti nello svolgimento dei loro compiti e delle loro attività in modo da tutelare…».

Il quadro è chiaro, anche se, in estrema sintesi, si deve rilevare che quanto maggiori saranno sul piano della potenziale invasività, della complessità e del contenuto innovativo i sistemi di IA (oggetto di indagine come strumento di indagine) tanto maggiore potrà essere la difficoltà di effettuare analisi e verifiche funzionali ad accertarne le effettive potenzialità di utilizzo e le condotte concretamente realizzate attraverso gli stessi.

Oggetto e strumenti di accertamento: i diritti della difesa

Come sopra abbiamo specificato, la questione deve essere esaminato in una differente prospettiva. Non è questa la sede per analizzare in quale misura programmi di IA potranno essere concretamente utilizzati nello svolgimento delle indagini: si tratta di tematica di estremo rilievo, sulla quale molte riflessioni dovranno essere effettuate, atteso che verosimilmente un approccio di taglio “manicheo” (qualsiasi o nessuna) non risponde alla prospettiva più verosimile. Al proposito si è osservato che programmi di IA potrebbero essere utilizzati, ad es., per «valutare meglio il grado di attendibilità dei testi oculari, essendo in grado di fornire più precise indicazioni su condizioni atmosferiche, visibilità, distanza del teste dal luogo di svolgimento dei fatti, o stabilire con maggior precisione l'autenticità o la provenienza di un documento» (così R.E., Kostoris, *Intelligenza artificiale, strumenti predittivi e processo penale*, in *Cass. pen*. 2024)

Nei casi nei quali si riterrà possibile (e lecito e, prima o poi, doveroso) utilizzare tali strumenti, è evidente che si porrà il problema di assicurare alla difesa la possibilità di **conoscere l'algoritmo** utilizzato per giungere ai risultati presentati dall'accusa. In tale ottica – a differenza di quanto vedremo *infra* – potrebbe essere di minore rilievo la conoscenza delle informazioni sulla base della quali l'algoritmo ha “lavorato”, atteso che si tratterà di elementi probatori comunque riversati nel fascicolo del P.M. o comunque provenienti da banche dati o raccolte di informazioni pubbliche, come tali, accessibili e verificabili.

Il nodo centrale del problema, per altro, è quello della necessità o meno di **ostensione dei codici sorgente del programma**. In questo senso «partendo dalla premessa che i sistemi di intelligenza artificiale funzionano stabilendo correlazioni tra enormi masse di dati, si è ritenuto che il primo problema sia quello di garantire a tutti gli attori processuali la conoscenza dei dati di partenza inseriti nell'algoritmo» si deve ritenere che «quando un siffatto strumento sia impiegato da autorità giudiziarie in un processo penale, non dovrebbe essere mai consentito opporre il segreto industriale sui dati, che integrano il c.d. codice sorgente», trattandosi di una condizione necessaria per poter esperire un controllo indipendente sui dati forniti dal sistema (cfr. A. Balsamo, *L'impatto dell'intelligenza artificiale nel settore della giustizia*, in *sistemapenale.it*).

La questione è chiarissima; molto meno la risposta: «L'incapacità di stabilire collegamenti causali tra i fatti e la incapacità di processare i dati in funzione semantica, unite al carattere opaco dei sistemi di autoapprendimento (che non consentono neppure di controllare, il procedimento attraverso il quale sono pervenuti a un certo risultato), impediscono di contestare con argomenti razionali il risultato ottenuto, incidendo quindi sul diritto di difesa e sul contraddittorio nella formazione della prova, come pure sulla possibilità del giudice di poter comprendere l'affidabilità dei dati che dovrebbe porre a fondamento del suo giudizio e di poter di conseguenza motivare adeguatamente» (così A. Balsamo, *op.cit.*).

Se la rilevanza del problema pare oggettiva, non scontata la

Se la rilevanza del problema pare oggettiva, non scontata la risposta che la giurisprudenza potrà fornire. Indubbiamente non si tratta di una questione del tutto nuova, trattandosi di una criticità potenziale del rapporto tra accusa e difesa presente da tempo, seppure in forme differenti e con differente oggetto, che frequentemente si è presentata – potremmo dire: ovviamente – proprio nel settore delle indagini per reati informatici o commessi attraverso l'utilizzo di strumenti informatici.

Emblematica, in questo senso la notissima vicenda Sky Ecc – con oggetto la diffusione e l'utilizzo di criptocellulari ritenuti sostanzialmente non intercettabili – in relazione alla "ricaduta" in termini di utilizzabilità negli ordinamenti diversi da quello francese nel momento in cui la autorità di tale paese non hanno ritenuto di "disvelare" sostanzialmente le metodiche tecniche utilizzate per analizzare la RAM, cioè la memoria dei server, su cui erano immagazzinate tutte le informazioni necessarie al sistema per funzionare il sistema.

Situazioni nelle quali il diritto di difesa – costituzionalmente garantito – si scontra con l'efficacia dell'indagine penale (che frequentemente ha per oggetto la tutela di altri interessi di pare rilievo costituzionale). Proprio nel settore informatico la costante "rincorsa" tra lo sviluppo di nuovi strumenti e programmi (siano essi a finalità illecita o a finalità investigativa) e la predisposizione di adeguate contromisure impone – almeno sul piano logico – di "colmare" rapidamente il gap che si apre a fronte di ogni nuovo sviluppo, spesso con sforzi rilevanti e investimenti significativi. *Nihil sub sole novi,* anche se dobbiamo aspettarci che la tematica- in relazione agli accertamenti sui sistemi di IA, sarò amplificata se non elevata all'ennesima potenza. Confidiamo sul fatto che la S.C. saprà fornire indicazioni precise sul punto.

In definitiva, per quanto riguarda l'analisi dei sistemi di IA utilizzati in chiave penale le problematiche che la pubblica accusa dovrò affrontare non si discostano- se non in termini quantitativi- da quelle che si sono presentate in contesti analoghi: pensiamo, allora, alla necessità di fruire di strumenti di collaborazione giudiziaria internazionale, laddove i produttori dei programmi abbiano sede all'estero (e magari non nell'ambito europeo) e alla – in alcuni casi – indispensabile collaborazione della autorità di tali paesi (un nome a caso: USA) per "convincere" i produttori a rendere disponibili le informazioni sui sistemi. Anche in questo caso – come è già accaduto per la possibilità di "lettura" degli I-phone anni orsono – la disponibilità di imprese straniere può essere direttamente proporzionale alla "bontà" del rapporto con l'a.g. del paese ove la società ha sede e inversamente proporzionale all'interesse economico-commerciale a una – anche solo parziale – "*discovery*" degli algoritmi.

Per quanto riguarda, al contrario, all'uso di sistemi di IA in fase di indagine, la categoria con la quale si rende necessario un confronto è quella della **prova atipica**, di cui all'art. 189 c.p.p., trattandosi di prova digitale e scientifica non specificamente disciplinata dalla legge. Una tipologia di prova che deve rispondere a specifici requisiti, in quanto deve risultare:

> autonomamente compatibile con i principi costituzionale
> oggettivamente idonea ad assicurare l'accertamento dei fatti per i quali viene disposta

Sul punto, la S.C., già in epoca remota (Cass. pen., sez. V, 26 novembre 1998, n. 1858, Rv. 212468 – 01) aveva stabilito che nell'assumere prove non disciplinate dalla legge a norma dell'art.189 c.p.p., il giudice è tenuto ad applicare i criteri legali stabiliti per gli analoghi mezzi di prova tipici ovvero a ricorrere a consolidate massime di esperienza o regole di inferenza secondo una disciplina scientifica.

Proprio la "scientificità" di uno strumento quale un sistema di IA impone di trasporre nella valutazione sulla idoneità di uno sistema di IA applicato all'indagine penale principi enucleati dalla S.C. in una notissima decisione (Cass. pen., sez. IV, 13 dicembre 2010, n. 43786, Rv. 248943-4): «Per valutare l'attendibilità di una teoria occorre esaminare gli studi che la sorreggono; le basi fattuali sulle quali essi sono condotti; l'ampiezza, la rigorosità, l'oggettività della ricerca; il grado di sostegno che i fatti accordano alla tesi; la discussione critica che ha accompagnato l'elaborazione dello

studio, focalizzata sia sui fatti che mettono in discussione l'ipotesi, sia sulle diverse opinioni che nel corso della discussione si sono formate; l'attitudine esplicativa dell'elaborazione teorica. Ancora, rileva il grado di consenso che la tesi raccoglie nella comunità scientifica. Infine, dal punto di vista del giudice, è di preminente rilievo l'identità, l'autorità indiscussa, l'indipendenza del soggetto che gestisce la ricerca, le finalità per le quali si muove. Dopo aver valutato l'affidabilità metodologica e l'integrità delle intenzioni, occorre infine valutare se esista una teoria sufficientemente affidabile ed in grado di fornire concrete, significative ed attendibili informazioni idonee a sorreggere l'argomentazione probatoria inerente allo specifico caso esaminato; deve trattarsi, cioè, di una teoria sulla quale si registra un preponderante, condiviso consenso….

Di tale complessa indagine il giudice è infine chiamato a dar conto in motivazione, esplicitando le informazioni scientifiche disponibili e fornendo razionale spiegazione, in modo completo e comprensibile a tutti, dell'apprezzamento compiuto».

Proviamo a ipotizzare a un sistema di analisi comparativa tra dati in termini massivi o di riconoscimento di forme di contraffazione di dati personali (immagine, voce o video) analizzati a confrontati con sistemi di IA. Spetterà, pertanto, alla giurisprudenza – a fronte di specifici “stimoli” da parte della P.G. – e della pubblica accusa – verificare, nell'ambito del principio del contraddittorio – l'ammissibilità, l'affidabilità e l'autorevolezza degli elementi probatori derivanti dall'utilizzo di sistemi di IA in fase di indagine.

Infine: da dove deriva il “learning”?

Vi è ancora un aspetto che merita di essere brevemente affrontato, in quanto apparentemente meno rilevante di altri: ma solo apparentemente. Vediamo perché.

Le valutazioni – in chiave di penale rilevanza – di condotto derivanti da sistemi di IA – è destinata indubbiamente a concentrarsi sull'algoritmo sul quale si fonda la funzionalità del sistema, e quindi – comi abbiamo visto – sui codici sorgente, sugli sviluppi successivi e – una differente prospettiva- sulla individuazione dei soggetti che tale sistema hanno ipotizzato, realizzato, implementato e – infine – applicato. È giusto e logico che sia così.

Nondimeno, siccome ci stiamo occupando di modelli di IA di natura induttiva, destinati a operare con un “meccanismo” inferenziale fondato sul machine learning, non possiamo trascurare la rilevanza – assoluta – che è rappresentata dai dati e dalle informazioni che di tale apprendimento costituiranno il “propellente”.

In questa prospettiva emerge il problema della documentazione del dataset, ossia dei database utilizzati per addestrare le IA; un problema serio, in *quanto* «…non sapere la provenienza dei dati da cui le IA imparano a dialogare, risolvere problemi, creare immagini, scrivere testi e molto, molto altro ancora, espone a rischi di discriminazione, abusi da parte dei governi, mancato rispetto delle minoranze e violazione di copyright…». Si tratta di «un aspetto molto delicato dello sviluppo delle IA…: la provenienza dei dati che servono per allenarle. Va ricordato che questa mole enorme di informazioni viene concessa agli sviluppatori in modo gratuito, con l'idea che ne facciano un uso non commerciale. Che non ci guadagnino su, insomma...» (così E. Capone, *L'intelligenza artificiale oltre quella umana e gli umani come formiche: OpenAI e i rischi della IA Forte*, *repubblica.it*).

E allora, appare evidente che un'indagine su un sistema di IA deve tenere conto non solo del meccanismo tecnico “astratto” che costituisce la struttura del sistema, ma della provenienza, della autenticità, della completezza e della pertinenza dei dati e delle informazioni che – si può immaginare – gli ideatori del programma avevano ipotizzato. È certo che un “coefficiente” di potenziale illiceità nella “predisposizione” e nell'utilizzo di un sistema può derivare proprio dalla scelta di “alimentare” l'apprendimento con dati e informazioni non rispondenti alle caratteristiche astratte in relazione alle quali il sistema stesso era stato immaginato.

E' una differente prospettiva di valutazione della possibile penale rilevanza di una condotta che può presentarsi come alternativa a

rilevanza di una condotta, che può presentarsi come alternativa o cumulativa rispetto alle valutazioni del sistema in sé, e che può richiedere forme di indagine non necessariamente ed esclusivamente tecnico- scientifiche, quanto anche “storiche”, laddove si presenti l'esigenza di comprendere la provenienza di dati e informazioni, la sussistenza di forme di “selezione” preventiva prima dell''utilizzo per la implementazione del sistema – e ovviamente – la individuazione dei soggetti ai quali tali ultimi aspetti sono direttamente riferibili.

Insomma: ci sono ottime ragioni per pensare che non ci annoieremo.

In conclusione

Per affrontare le questioni interpretative legate all'entrata in vigore dell'AI Act, è importante puntualizzare il concetto di sistema di intelligenza artificiale contenuto nel Regolamento (UE) 2024/1689 Regolamento, che è destinata a essere recepita dal legislatore italiano.
I sistemi di intelligenza artificiale fondati sul meccanismo del machine learning devono essere considerati strumenti dinamici in costante modifica e pertanto gli accertamenti tecnici sugli stessi potrebbero doversi svolgere con le forme dell'accertamento irripetibile o dell'incidente probatorio.
Gli accertamenti tecnici menzionati si possono presentare di particolare delicatezza e complessità, in quanto evidentemente i sistemi in oggetto sono frutto di investimenti – dal punto di vista tecnologico ed economico di grande rilievo; nondimeno, non vi sono ragioni per ritenere che l'approccio a tali approfondimenti debba essere differente rispetto a quello già maturato in analoghi campi nell'ambito delle indagini penali.
Non si può trascurare la rilevanza – in relazione alle indagini su sistemi di IA – della valutazione sulla scelta dei dati e delle informazioni utilizzati per implementare il sistema, in quanto forme di illiceità possono derivare anche, evidentemente, sotto tale profilo.

FONTE: IUS/Penale