



Mediterranean Insecurity

L'era dell'AI e l'evoluzione della criminalità: continuità e innovazione tra crimine tradizionale e criminalità digitale.

**PARTE SECONDA – Reti criminali e reti tecnologiche:
l'evoluzione delle organizzazioni criminali nell'era
dell'intelligenza artificiale tra Europa e contesto globale**

Walter Floris

Dopo aver delineato, nella Parte Prima, il quadro storico e geografico relativo alla nascita della criminalità organizzata “vecchia scuola”¹, si procede all’analisi dei principali ambiti nei quali essa ha sviluppato la propria attività.

Che cosa è l'intelligenza artificiale?

Questo interrogativo, pur sembrando apparentemente semplice, si rivela in realtà assai più complesso di quanto possa sembrare a prima vista.

¹ FLORIS WALTER, *Cenni storici e geografici sulla nascita della criminalità organizzata “vecchia scuola”*, in *Mediterranean Insecurity*, agosto 2026.





Figura 1: fonte: <https://www.cybersecurity360.it/news/europol-lia-e-catalizzatore-del-crimine-organizzato-nuove-sfide-per-la-sicurezza-europea/> Pubblicato il 21 mar 2025

L'obiettivo consiste, dunque, nel verificare se sia possibile individuare una definizione condivisa di IA, aderente allo stato attuale della tecnica e non influenzata da prospettive eccessivamente speculative o futuristiche. Prima di affrontare la complessità definitoria propria del concetto di IA, che negli ultimi anni ha attirato l'attenzione di legislatori, dottrina, giurisprudenza e, più in generale, anche dei "non addetti ai lavori", appare tuttavia opportuno procedere per gradi, introducendo alcuni concetti preliminari. L'analisi prenderà pertanto avvio dal concetto di dato, per giungere successivamente a quello di algoritmo.

In una prima e necessariamente sommaria prospettiva, l'IA può essere definita come un "insieme di tecnologie che combina dati, algoritmi e potenza di calcolo"². L'intelligenza artificiale non costituisce un'entità materiale percepibile in senso fisico, ma si manifesta attraverso sistemi computazionali fondati sull'elaborazione di dati mediante algoritmi e modelli matematici. Sebbene tali sistemi siano integrati all'interno di infrastrutture hardware e software, essi non possono essere identificati con il supporto materiale che ne consente il funzionamento. A caratterizzare l'IA sono, infatti, i processi computazionali e i modelli attraverso i quali il sistema acquisisce ed elabora informazioni, producendo determinati risultati e, in alcuni casi, modificando le proprie prestazioni sulla base dei dati e dell'esperienza acquisita.

In questa prospettiva, l'IA può essere considerata un sistema computazionale complesso, generalmente costituito da componenti software ed eseguito mediante un'infrastruttura hardware, vale a dire l'insieme degli elementi fisici che compongono un dispositivo elettronico. In ambito informatico, tale

² RUFFOLO U., *Intelligenza Artificiale, machine learning e responsabilità da algoritmo*, in Giur. it., luglio 2019.

infrastruttura coincide con l'elaboratore, mentre nel settore della robotica può essere rappresentata dal robot. Pur trattandosi di una definizione volutamente ampia, essa consente di individuare alcuni elementi essenziali per una corretta comprensione del fenomeno.

Anzitutto, questa impostazione evidenzia i principali fattori che hanno contribuito alla rapida evoluzione dell'IA negli ultimi vent'anni, tra cui il significativo incremento delle capacità computazionali e la crescente disponibilità di dati. All'interno di tale prospettiva assumono particolare rilievo proprio i dati e gli algoritmi, che costituiscono componenti fondamentali dei sistemi di IA. I dati rappresentano una risorsa indispensabile per il loro sviluppo e funzionamento³, mentre gli algoritmi consistono nelle procedure attraverso cui tali dati vengono elaborati al fine di ottenere informazioni e risultati. In particolare, nei sistemi di apprendimento automatico, gli algoritmi permettono di analizzare i dati, individuare schemi e correlazioni e formulare nuove inferenze.

Tuttavia, le difficoltà definitorie già riscontrate con riferimento alla nozione di IA si ripropongono anche rispetto ai concetti di dati, Big Data e algoritmi. Si tratta, infatti, di espressioni ormai ampiamente diffuse nel linguaggio comune, il cui utilizzo può talvolta risultare improprio o eccessivamente generico, anche a causa di una conoscenza non sempre approfondita delle caratteristiche e delle implicazioni proprie di tali concetti.

AI e valutazione della pericolosità criminale: gli algoritmi predittivi.

Negli ultimi anni si è registrato un aumento dei casi di criminalità informatica nei quali l'intelligenza artificiale assume un ruolo centrale, evidenziando uno dei principali rischi derivanti dall'utilizzo di questa tecnologia. In tale prospettiva, è possibile individuare quattro diverse modalità di impiego degli algoritmi: *l'algoritmo delinquente*, *l'algoritmo investigante*, *l'algoritmo consulente* e *l'algoritmo giudicante*.

Con l'espressione "algoritmo delinquente" si fa riferimento all'utilizzo dell'IA per la realizzazione di attività illecite o per il perseguimento di scopi

³ Di VIZIO F., *Prevenzione e investigazioni: l'uso di IA, big data e soluzioni tecnologiche in ambito finanziario e nel contrasto al riciclaggio (AML) e al finanziamento del terrorismo (CFT)*, in DisCrimen, 11 gennaio 2024.

criminali. Le potenzialità della tecnologia possono, infatti, essere sfruttate per rendere le condotte criminose maggiormente organizzate, efficaci e complesse da rilevare.

L'intelligenza artificiale può, però, essere utilizzata anche per finalità preventive e di contrasto alla criminalità, soprattutto nell'ambito della polizia predittiva e della giustizia predittiva. In ambito investigativo, l'algoritmo investigante consente di elaborare e mettere in relazione grandi quantità di dati al fine di individuare possibili situazioni di rischio o di fornire indicazioni utili all'individuazione dei responsabili di un reato già commesso.

Tali strumenti rientrano nel più ampio settore del *law enforcement* e possono essere impiegati dalle forze dell'ordine per analizzare informazioni e individuare eventuali ricorrenze nei fenomeni criminali. Un'applicazione significativa consiste nell'identificazione degli *hot spot*, ossia delle zone territoriali caratterizzate, sulla base dei dati disponibili, da una maggiore probabilità di verifica di determinati reati.

L'utilizzo dell'IA in questo settore presenta, tuttavia, profili problematici. Tra i principali si collocano i *bias* algoritmici, che possono influenzare i risultati e produrre effetti discriminatori, oltre alle questioni relative alla protezione dei dati personali e alla privacy. Anche in Italia sono stati sperimentati sistemi di questo genere, tra cui “*KeyCrime*”, utilizzato dalla Questura di Milano⁴. Nel campo della giustizia predittiva assumono invece rilievo l'algoritmo consulente e l'algoritmo giudicante. Lo sviluppo di queste tecnologie ha portato a ipotizzare scenari nei quali l'IA possa essere utilizzata anche per formulare previsioni sull'esito di un procedimento giudiziario⁵.

L'algoritmo consulente ha una funzione essenzialmente ausiliaria: attraverso l'elaborazione dei dati disponibili, fornisce al giudice informazioni e valutazioni che possono contribuire alla formazione della decisione, lasciando comunque a quest'ultimo la responsabilità dell'interpretazione e della scelta finale. L'algoritmo giudicante, invece, indica sistemi caratterizzati da un

⁴ NAST C., “Il software italiano che ha cambiato il mondo della polizia predittiva”, Wired Italia, maggio 2019.

⁵ COSTA C., “Intelligenza artificiale e Giustizia: tempi ancora prematuri”, AI4Business, maggio 2019.

coinvolgimento più diretto nel processo decisionale, ad esempio nelle valutazioni relative al *pretrial release* o nella predisposizione di decisioni giudiziarie.

L'impiego di strumenti algoritmici in ambito giudiziario deve essere accompagnato da adeguate garanzie, dal momento che le valutazioni prodotte possono incidere significativamente sulla libertà e sui diritti delle persone. Al tempo stesso, se sottoposti a controlli adeguati e utilizzati secondo criteri di trasparenza, verificabilità e affidabilità, tali sistemi possono rappresentare un utile strumento di supporto, soprattutto per la loro capacità di elaborare una quantità di informazioni difficilmente analizzabile mediante le sole capacità umane.

Un'applicazione particolarmente significativa della giustizia predittiva riguarda la valutazione del rischio di recidiva. Gli algoritmi possono essere utilizzati per stimare la probabilità che un soggetto reiteri determinate condotte criminose e fornire, di conseguenza, elementi utili alla valutazione giudiziaria, anche in relazione alle misure cautelari. La valutazione algoritmica dovrebbe, in ogni caso, mantenere una funzione meramente strumentale, senza sostituire il giudizio del magistrato.

Le applicazioni dell'IA non si limitano, tuttavia, al settore penale. La tecnologia trova crescente spazio anche in altri ambiti del diritto, tra cui quello contrattuale e della risoluzione delle controversie. In Paesi come i Paesi Bassi e il Regno Unito sono stati sviluppati sistemi capaci di gestire controversie sorte online e, in determinate circostanze, di contribuire alla loro risoluzione automatizzata. In ambito stragiudiziale, tali strumenti possono essere utilizzati nell'arbitrato e nella conciliazione, offrendo soluzioni che, nei casi previsti, possono assumere rilievo anche nell'eventuale successiva fase giudiziale⁶.

È indubbio che l'intelligenza artificiale stia trasformando e continuerà a trasformare le pratiche legali e che il diritto sia chiamato ad affrontare e regolamentare le nuove tecnologie, nonché le molteplici questioni giuridiche ad esse connesse.

L'IA è intelligente e pensante?

⁶ Si veda “Online Dispute Resolution e giustizia digitale,” Altalex, 23 febbraio 2021.

La tendenza a descrivere l'intelligenza artificiale come un “computer o robot capace di pensare autonomamente, apprendere e svolgere attività proprie dell'uomo” riflette una visione fortemente antropomorfica del fenomeno, nella quale il funzionamento dei sistemi di IA viene assimilato, almeno in parte, a quello dell'intelligenza umana. Tale impostazione può risultare fuorviante, soprattutto quando si ricorre a espressioni come “apprendimento”, “pensiero”, “ragionamento” e “comprensione”, termini tradizionalmente riferiti alle capacità cognitive dell'essere umano⁷. Il loro utilizzo, infatti, rischia di attribuire ai sistemi di IA caratteristiche e capacità che non possiedono realmente, o comunque di rappresentarne le potenzialità in maniera eccessivamente ampia.

Un simile linguaggio non produce conseguenze soltanto sul piano terminologico, ma può influenzare anche il modo in cui l'intelligenza artificiale viene percepita e discussa nella società. La rappresentazione dell'IA come una forma di intelligenza simile a quella umana può alimentare, da un lato, l'idea che tali sistemi siano destinati a sostituire progressivamente l'uomo in numerose attività e, dall'altro, scenari più estremi nei quali le macchine arriverebbero a sfuggire al controllo umano, fino a mettere in discussione la stessa sopravvivenza dell'umanità. Ne deriva l'esigenza di utilizzare una terminologia più precisa e consapevole, evitando di confondere le effettive capacità dei sistemi artificiali con quelle proprie dell'intelligenza umana.

D'altra parte, un simile equivoco non appare del tutto sorprendente, considerando che già la scelta di accostare il termine “intelligenza” all'aggettivo “artificiale” ha contribuito a creare l'idea che le macchine possano essere “intelligenti” secondo modalità comparabili a quelle dell'essere umano.

Prima di analizzare il rapporto tra intelligenza umana e intelligenza artificiale, è tuttavia opportuno svolgere una breve considerazione preliminare di carattere metodologico. L'utilizzo del termine “intelligenza”, infatti, non solo può rendere meno immediata la comprensione della reale natura dell'IA, ma presenta anche significative difficoltà sul piano definitorio. Il concetto di intelligenza, infatti, è tutt'altro che univoco e continua a essere oggetto di discussione e approfondimento nell'ambito delle scienze cognitive, della psicologia e delle

⁷ L'ambiziosa associazione tra macchine e intelligenza umana può infatti rendere più suggestiva, ma al contempo meno accurata, la comprensione di una tecnologia complessa e di particolare rilevanza.

neuroscienze. La domanda relativa a cosa debba intendersi per “intelligenza” rimane, pertanto, ancora oggi priva di una risposta universalmente condivisa.

La difficoltà di definire concetti quali “intelligenza” e “pensiero” era stata già riconosciuta da Alan Turing, il quale adottò un approccio prevalentemente operativo. Attraverso il celebre test di Turing, il matematico inglese propose di valutare non tanto l’esistenza di una vera e propria intelligenza nella macchina, quanto la sua capacità di manifestare, in determinate condizioni, un comportamento non distinguibile da quello di un essere umano. In tal modo, Turing spostò il problema dal piano puramente teorico a quello dell’osservazione concreta del comportamento della macchina.

Tra le definizioni maggiormente condivise, l’intelligenza viene descritta come l’insieme delle capacità cognitive fondamentali che permettono di affrontare e risolvere problemi in contesti differenti”⁸. In questa prospettiva, l’essere umano può essere definito intelligente proprio in virtù della capacità di utilizzare conoscenze, esperienze, ragionamento e immaginazione per affrontare situazioni nuove e adattarsi alle diverse circostanze.

Una concezione in parte analoga è quella elaborata dallo psicologo statunitense Howard Gardner, autore della teoria delle intelligenze multiple. Secondo tale impostazione, l’intelligenza può essere intesa come la capacità di utilizzare abilità cognitive e conoscenze per apprendere, affrontare problemi e perseguire obiettivi che assumono valore per il singolo individuo o per la società e la cultura di appartenenza.

Ne consegue che il semplice utilizzo del termine “intelligenza” in riferimento ai sistemi artificiali non dovrebbe condurre automaticamente a considerarli equivalenti all’intelligenza umana. Le capacità proprie dell’IA devono, infatti, essere analizzate tenendo conto delle modalità attraverso cui tali sistemi elaborano informazioni, apprendono dai dati e producono determinati risultati, evitando di sovrapporre a questi processi caratteristiche cognitive tipicamente umane”⁹.

Tuttavia, studi più recenti hanno evidenziato come una concezione dell’intelligenza formulata in questi termini consenta di affermare che un

⁸ KAPLAN J., *Intelligenza artificiale*, LUISS University press, pp. 43 ss., Roma, 2018.

⁹ GARDNER H., *Intelligenze multiple*, Anabasi, Milano, 1994.

individuo sia in grado di risolvere determinati problemi, senza necessariamente chiarire attraverso quali processi o modalità tale risultato venga raggiunto. Per tali ragioni, può risultare preferibile fare riferimento al concetto di abilità piuttosto che a quello, più ampio e problematico, di intelligenza.

Non è questa la sede per approfondire ulteriormente il tema e i suoi risvolti filosofici. Poiché l'obiettivo dell'analisi è formulare alcune considerazioni critiche sul rapporto tra intelligenza biologica e intelligenza artificiale, si assumerà, quale premessa dell'analisi, che l'intelligenza, così come altri attributi ad essa strettamente connessi, quali il pensiero e la coscienza, costituisca una caratteristica propria dell'esperienza umana.

AI e le investigazioni. La tecnologia come strumento nella lotta alla criminalità.

Tra i settori nei quali l'utilizzo di strumenti basati sull'intelligenza artificiale sta assumendo crescente importanza rientra quello investigativo. In questo ambito, l'IA può rappresentare un significativo strumento di supporto per gli investigatori, sia sul piano giudiziario, contribuendo alla ricostruzione dei fatti e all'individuazione dei possibili responsabili, sia sul piano della pubblica sicurezza, attraverso applicazioni orientate alla prevenzione della criminalità.

Questa evoluzione trova conferma anche nel documento elaborato in occasione dell'incontro annuale degli esperti di polizia promosso dall'*Organization for Security and Co-operation in Europe* (OSCE). Il documento sottolinea come la necessità di rendere le attività di polizia sempre più efficaci ed efficienti, anche alla luce del rapido sviluppo tecnologico, abbia reso indispensabile l'adozione di strumenti avanzati per la gestione e l'analisi dell'enorme quantità di dati raccolti dalle forze dell'ordine. Attraverso tali tecnologie è possibile, ad esempio, individuare ricorrenze e schemi nei dati, effettuare valutazioni del rischio e orientare la distribuzione delle risorse umane ed economiche verso i settori nei quali risultano maggiormente necessarie.

Sebbene si tratti di un fenomeno relativamente recente, diversi sistemi di IA sono già stati sperimentati e, in alcuni Paesi, vengono concretamente utilizzati dalle forze di polizia. Tra le principali applicazioni rientrano i software destinati all'analisi di immagini e filmati, i sistemi di riconoscimento facciale e

identificazione biometrica, i droni autonomi, le tecnologie robotiche e gli strumenti di analisi predittiva. Questi ultimi possono essere impiegati per individuare le cosiddette “*zone calde*” della criminalità, ossia le aree caratterizzate da una maggiore incidenza di determinati fenomeni criminosi, oppure per stimare la probabilità che possano verificarsi specifiche condotte illecite, soprattutto quando presentano un elevato livello di pericolosità¹⁰.

In tale prospettiva, l'intelligenza artificiale offre agli investigatori possibilità particolarmente significative. Da un lato, consente di elaborare valutazioni orientate al futuro e quindi funzionali alla prevenzione dei reati; dall'altro, permette di ottenere una conoscenza più approfondita di eventi già verificatisi, contribuendo, ad esempio, alla ricostruzione della dinamica e della scena del crimine.

In base alla finalità perseguita, gli strumenti di IA impiegati in ambito investigativo possono essere ricondotti, in termini generali, a due categorie: quelli caratterizzati da una funzione predittiva e quelli dotati di una funzione conoscitiva. Il principale elemento di innovazione risiede nella capacità di questi sistemi di elaborare informazioni già disponibili e, attraverso la loro combinazione e analisi, produrre nuove conoscenze utili all'attività investigativa.

Sicurezza e criminalità. Il software made in Italy “XLAW”.

Nel contesto della polizia predittiva applicata alla sicurezza urbana, anche l'Italia ha intrapreso diverse iniziative finalizzate a potenziare le attività di prevenzione e contrasto dei fenomeni criminali. Tra le principali sperimentazioni realizzate nel nostro Paese si colloca XLAW, un sistema già adottato in diverse città italiane, tra cui Napoli, Prato, Salerno, Venezia, Modena e Parma.

La sperimentazione di XLAW è stata avviata dal Dipartimento della Pubblica Sicurezza – Direzione Centrale Anticrimine della Polizia di Stato (DAC), in collaborazione con la Questura di Napoli¹¹. L'efficacia e l'affidabilità del sistema sono state inoltre oggetto di valutazione da parte di soggetti esterni,

¹⁰ 2019 OSCE Annual Police Experts Meeting *Artificial Intelligence and Law Enforcement: An Ally or an Adversary?*, 23-24 September 2019, Vienna, Hofburg.

¹¹ Si veda il sito <https://www.xlaw.it/presentazione/>

tra cui ricercatori dell'Università degli Studi di Napoli Federico II e dell'Università degli Studi di Napoli "Parthenope".



Figura 2: Polizia Predittiva Per La Sicurezza Urbana: Sistema Xlaw Made In Italy fonte <https://www.piusicurezza.com/2022/11/07/polizia-predittiva-xlaw/>

L'elemento distintivo di XLAW consiste nell'elaborazione di un modello concettuale di natura criminologica, definito "riserve di caccia"¹², sul quale si fonda il modello statistico impiegato per l'analisi dei fenomeni criminali. Il sistema concentra la propria attenzione principalmente sul rapporto tra il reato e il luogo in cui viene commesso. A intervalli regolari, l'algoritmo analizza i dati disponibili e genera previsioni e avvisi relativi alle ore successive, indicando le aree territoriali nelle quali risulta più elevata la probabilità che si verifichino determinate tipologie di reato.

XLAW si basa su un algoritmo euristico finalizzato soprattutto alla previsione dei cosiddetti reati predatori. Il sistema utilizza dati storici geolocalizzati relativi agli illeciti già verificatisi e, attraverso la loro elaborazione, individua le aree nelle quali determinate condotte criminali tendono a concentrarsi. Questi luoghi, definiti appunto "riserve di caccia", possono comprendere, ad esempio, zone turistiche, aree circostanti stazioni ferroviarie o metropolitane e mercati, ossia contesti nei quali alcuni fenomeni criminali possono manifestare specifiche ricorrenze sia dal punto di vista spaziale sia temporale.

Attraverso questa metodologia, il sistema è in grado di individuare gli *hot spot*, riconoscere schemi ricorrenti (*pattern*) associati ai reati predatori e realizzare una vera e propria mappatura della criminalità (*crime mapping*). In questo modo vengono evidenziate le zone nelle quali determinati reati presentano una maggiore

¹² LOMBARDO E., *Sicurezza 4P. Lo studio alla base del software XLAW per prevedere e prevenire i crimini*, Mazzanti Libri, 2019.

L'analisi non si limita, tuttavia, alla localizzazione geografica dei reati. XLAW prende in considerazione anche diversi elementi del fenomeno criminale, tra cui le modalità operative degli autori, le caratteristiche delle vittime, le peculiarità del territorio e la conformazione dell'ambiente urbano. Vengono valutati, ad esempio, gli eventuali percorsi di fuga e la presenza di ostacoli o barriere, naturali o artificiali, che possono incidere sulle modalità di commissione del reato e sulla possibilità di allontanamento.

Tale metodologia può contribuire a contrastare l'attività di soggetti che commettono reiteratamente reati quali furti, rapine e scippi, soprattutto quando le condotte poste in essere presentano caratteristiche e modalità operative ricorrenti. Il sistema, infatti, è orientato a rilevare schemi e ricorrenze spazio-temporali nella distribuzione dei fenomeni criminali, tenendo conto della possibilità che determinati autori adottino nel tempo modalità di azione consolidate e privilegino specifici contesti territoriali considerati particolarmente favorevoli alla commissione dei reati.



Gli alert generati da XLAW permettono alle forze dell'ordine di indirizzare le attività di pattugliamento verso le aree geografiche che il sistema individua

come maggiormente esposte al rischio di determinati reati. Prima dell'utilizzo del software, una pattuglia percorreva mediamente circa 125 km al giorno; grazie all'analisi dei dati e alla conseguente distribuzione mirata delle risorse, il percorso medio giornaliero sarebbe sceso a circa 23 km. L'impiego della tecnologia consente quindi di concentrare la presenza degli operatori nelle zone considerate più vulnerabili alla criminalità predatoria, rendendo il controllo del territorio maggiormente selettivo ed efficiente.

I risultati attribuiti alla sperimentazione di XLAW sembrerebbero indicare una maggiore efficacia rispetto alle modalità tradizionali di sorveglianza. Secondo i dati diffusi nelle fonti relative alla sperimentazione, il sistema avrebbe raggiunto un'accuratezza superiore all'80% nell'individuazione delle aree caratterizzate da un maggiore rischio di specifici reati. Nelle città coinvolte nella sperimentazione sarebbe stata inoltre registrata una riduzione dei reati predatori, accompagnata da un incremento del 20% degli arresti in flagranza¹³. I possibili benefici riguarderebbero anche l'aspetto economico, grazie a una gestione più razionale delle risorse disponibili e agli effetti derivanti dalla prevenzione dei fenomeni criminali.

Secondo i risultati riportati, l'utilizzo del sistema avrebbe inoltre contribuito a rafforzare la percezione di sicurezza dei cittadini e la fiducia nei confronti delle forze dell'ordine, favorendo al contempo una maggiore efficacia nelle attività decisionali e operative degli agenti. La riduzione dei reati potrebbe inoltre tradursi in una diminuzione dei costi complessivamente sostenuti dalla collettività.

Nel complesso, gli esiti delle sperimentazioni condotte in diverse città italiane sembrerebbero sostenere l'utilità di un modello algoritmico fondato sull'analisi congiunta del luogo e della tipologia di reato. In tale prospettiva, il modello delle "riserve di caccia" potrebbe rappresentare un interessante strumento di analisi per ulteriori studi in ambito sociologico e criminologico, oltre che un possibile supporto all'elaborazione di strategie più mirate di prevenzione e contrasto della criminalità.

Esistono punti d'incontro tra AI e giustizia?

¹³ Lombardo E., Sicurezza 4P, p. 84, op. cit.

L'impiego dell'intelligenza artificiale nel settore giudiziario incontra ancora numerosi limiti di natura giuridica. Allo stato attuale, l'ipotesi di affidare la funzione giudiziaria a un sistema di IA appare difficilmente compatibile con il quadro costituzionale italiano. In particolare, l'art. 102 Cost. stabilisce che la funzione giurisdizionale sia esercitata da magistrati ordinari istituiti e disciplinati secondo le norme dell'ordinamento giudiziario. A tale previsione si affiancano le garanzie previste dagli artt. 25 e 101 Cost., che rendono problematica sia la sostituzione del giudice con un sistema automatizzato sia l'eventualità che il magistrato debba necessariamente conformarsi alle conclusioni elaborate dall'intelligenza artificiale. Un sistema nel quale le decisioni automatizzate assumessero carattere vincolante solleverebbe, inoltre, rilevanti problematiche in relazione ai principi del contraddittorio e del giusto processo.

Anche la disciplina europea in materia di protezione dei dati personali pone specifiche garanzie. L'art. 22 GDPR riconosce, infatti, il diritto dell'interessato a non essere sottoposto a una decisione fondata esclusivamente su un trattamento automatizzato, mentre l'art. 10 GDPR disciplina in maniera specifica il trattamento dei dati personali relativi a condanne penali e reati, prevedendo particolari condizioni e garanzie.

In ambito giudiziario, pertanto, l'intelligenza artificiale può essere utilizzata come strumento di supporto all'attività del magistrato, fornendo informazioni e analisi utili alla valutazione del caso. Non può, invece, sostituirsi integralmente al giudice nell'esercizio della funzione giurisdizionale, né l'esito di un procedimento automatizzato può costituire, da solo, il fondamento della decisione giudiziaria.

I sistemi di giustizia predittiva e, più in generale, gli *automated decision-making systems*, sono stati sperimentati soprattutto in ambito civile e amministrativo. Nel settore penale, il loro utilizzo ha invece assunto particolare rilievo con riferimento agli strumenti destinati alla valutazione del rischio. Le informazioni elaborate da tali sistemi possono essere sottoposte al giudice come elementi aggiuntivi per valutare la posizione dell'imputato e orientare la scelta delle misure da adottare, senza tuttavia sostituire la valutazione giudiziale.

Applicazioni di questo tipo sono state oggetto di sperimentazione e dibattito anche negli Stati Uniti, in particolare in Stati quali Colorado, Arizona, Kentucky,

Delaware e Louisiana, evidenziando le potenzialità ma anche le numerose questioni giuridiche connesse all'utilizzo dell'IA nei procedimenti giudiziari.

In Europa, nel 2018, la Commissione europea per l'efficacia della giustizia (CEPEJ) del Consiglio d'Europa ha adottato la “Carta etica europea sull'uso dell'intelligenza artificiale nei sistemi giudiziari e negli ambiti connessi”, nella quale sono individuati alcuni principi fondamentali che devono orientare lo sviluppo e l'impiego dei sistemi di IA in ambito giudiziario:

- principio di trasparenza, neutralità ed equità;
- principio del controllo da parte di chi l'utilizza;
- principio di non-discriminazione;
- principio del rispetto dei Diritti Fondamentali;
- principio di qualità e sicurezza.

Il giudice è chiamato a decidere sulle controversie sottoposte alla sua valutazione e, in tale prospettiva, alcune delle attività che oggi rientrano nelle sue competenze potrebbero, almeno teoricamente, essere supportate da sistemi di intelligenza artificiale. Un simile impiego potrebbe contribuire a ridurre il carico di lavoro degli uffici giudiziari, spesso caratterizzati da un elevato numero di procedimenti e da risorse umane e temporali limitate.

In quest'ottica, si potrebbe prevedere una graduale individuazione delle attività maggiormente adatte a essere affidate, almeno in parte, a strumenti automatizzati, privilegiando quelle più semplici e standardizzabili, che non richiedano un significativo esercizio dell'interpretazione o della valutazione discrezionale del giudice. L'utilizzo dell'IA in questi settori potrebbe consentire di rendere più rapida la gestione delle controversie e di ridurre i tempi necessari per la loro definizione.

Non si può inoltre escludere che, in prospettiva futura, sistemi adeguatamente sviluppati e addestrati possano essere utilizzati come strumenti di ausilio alla decisione, soprattutto nei casi caratterizzati da una struttura semplice e nei quali l'applicazione della norma giuridica non presenti particolari difficoltà interpretative. In linea teorica, l'utilizzo di tali strumenti potrebbe anche contribuire a limitare alcune forme di arbitrarietà o discriminazione legate ai bias umani. Occorre, tuttavia, considerare che anche gli algoritmi possono essere

condizionati da pregiudizi presenti nei dati utilizzati per il loro addestramento, arrivando persino a riprodurli o accentuarli.

L'automazione potrebbe inoltre favorire una maggiore coerenza nell'applicazione delle norme e nel trattamento di casi caratterizzati da elementi analoghi. Ciò non dovrebbe però comportare una meccanica uniformazione dell'attività giudiziaria. La decisione deve continuare a tenere conto delle caratteristiche concrete di ciascuna controversia e delle peculiarità delle persone coinvolte, preservando così la necessaria dimensione individuale e valutativa della funzione giudiziaria.



Figura 4: fonte <https://aniasafe.it/product/problematiche-giuridiche-dellintelligenza-artificiale/>

“Machina delinquere non potest”? La responsabilità penale dell'intelligenza artificiale.

Un ulteriore profilo problematico riguarda la capacità dei sistemi di IA di classificare e profilare gli individui attraverso l'elaborazione di quantità molto elevate di dati. Da tale attività possono derivare vere e proprie identità algoritmiche, che non necessariamente coincidono né con l'immagine che una persona ha di sé né con la sua qualificazione giuridica formale. Gli algoritmi di *machine learning*, infatti, non si limitano all'analisi dei dati biometrici, ma possono individuare correlazioni statistiche riguardanti comportamenti, preferenze, abitudini e relazioni sociali, contribuendo così alla costruzione di un articolato profilo digitale dell'individuo.

A ciò si aggiunge un problema significativo legato alla trasparenza. L'utente, infatti, spesso non dispone di informazioni sufficienti sui criteri utilizzati dal sistema, sui parametri presi in considerazione o sulle modalità attraverso cui vengono elaborate le valutazioni. Si determina così una marcata

asimmetria informativa tra l'individuo e i soggetti che gestiscono le piattaforme e le infrastrutture tecnologiche.

La questione assume particolare rilevanza quando il sistema commette errori di classificazione o riproduce pregiudizi e stereotipi discriminatori. In tali circostanze, si pone il problema di stabilire a chi debba essere attribuita la responsabilità per le conseguenze prodotte: al soggetto che ha progettato l'algoritmo, all'impresa che lo ha sviluppato o commercializzato oppure a chi ne ha disposto l'utilizzo concreto?

È proprio in questo contesto che la filosofia del diritto è chiamata a interrogarsi sull'opportunità di elaborare nuovi criteri di imputazione, capaci di considerare la pluralità dei soggetti coinvolti nelle diverse fasi di progettazione, sviluppo, addestramento e utilizzo dei sistemi di IA. La riflessione deve inoltre confrontarsi con il rapporto tra *ius conditum*, ossia il diritto vigente, e *ius condendum*, vale a dire il diritto ancora da elaborare, con l'obiettivo di individuare soluzioni adeguate alle trasformazioni determinate dalla società digitale.

Particolarmente dibattuta è anche la possibilità di riconoscere una forma di soggettività giuridica ai sistemi artificiali autonomi. Tale ipotesi nasce dall'esigenza di individuare un centro di imputazione della responsabilità, ma si scontra con la natura complessa e distribuita dei processi algoritmici, nei quali intervengono diversi soggetti e componenti tecnologiche.

Le difficoltà aumentano in ragione del crescente livello di autonomia e imprevedibilità raggiunto dai sistemi di IA¹⁴. I meccanismi di *machine learning* e *deep learning*, infatti, consentono alle macchine di elaborare dati e produrre risultati attraverso processi che, in determinate circostanze, possono non essere pienamente prevedibili neppure da coloro che li hanno progettati. Ne deriva una progressiva difficoltà nel determinare in anticipo tutte le possibili risposte o comportamenti del sistema.

Da qui nasce il dibattito sull'eventuale necessità di superare il tradizionale modello vicario di responsabilità, fondato sul principio secondo cui *machina delinquere (et puniri) non potest*: la macchina, non essendo considerata soggetto

¹⁴ Il termine “attanti” viene utilizzato da parte della dottrina per riferirsi ai sistemi artificiali senza assimilarli agli attori, in quanto privi di soggettività, ma distinguendoli al contempo dai meri oggetti.

di diritto, non può essere direttamente destinataria della responsabilità penale, che deve pertanto essere ricondotta a un soggetto umano.

La rapida diffusione dell'intelligenza artificiale e la sua progressiva integrazione nelle attività quotidiane hanno quindi aperto nuovi interrogativi anche per il diritto penale. Diventa sempre più importante comprendere se gli attuali criteri di imputazione della responsabilità penale, tradizionalmente costruiti intorno alla condotta e alla volontà dell'essere umano, siano sufficienti a disciplinare situazioni nelle quali un comportamento potenzialmente criminoso venga realizzato attraverso sistemi di IA.

La crescente complessità tecnologica potrebbe dunque rendere necessario un ripensamento degli strumenti giuridici tradizionali, soprattutto nei casi in cui risulti difficile individuare il rapporto tra l'azione umana, il funzionamento autonomo del sistema e l'evento prodotto. La realtà tecnologica, infatti, sta progressivamente proponendo situazioni concrete che impongono al diritto di confrontarsi con scenari un tempo difficilmente prevedibili¹⁵.



Figura 5, “Nuovi reati per chi usa l’Intelligenza Artificiale”, fonte https://www.laleggepertutti.it/748181_nuovi-reati-per-chi-usa-lintelligenza-artificiale

Tra le diverse problematiche connesse all'impiego di sistemi artificiali dotati di intelligenza, una merita particolare attenzione per le sue rilevanti implicazioni, sia attuali sia future: quella relativa ai reati colposi contro la vita e l'incolumità delle persone. La questione non riguarda soltanto le molteplici modalità attraverso cui i sistemi di IA possono essere utilizzati nella commissione di reati, ma investe più specificamente le ipotesi in cui un comportamento lesivo

¹⁵ CALDWELL A., GRIFFIN T., *AI-enabled future crime*, in Crime Science, 2020.

sia materialmente determinato dall'azione di un sistema artificiale dotato di una certa autonomia.

La diffusione sempre più ampia di questi prodotti intelligenti solleva, infatti, importanti interrogativi in materia di sicurezza degli utenti e dei terzi. Il problema si collega, almeno in parte, alla più generale disciplina della responsabilità per danno da prodotto, poiché anche l'adozione di elevati standard di sicurezza non consente di escludere completamente la possibilità che tali tecnologie provochino eventi dannosi o mettano in pericolo l'incolumità delle persone.

Non si tratta, peraltro, di una questione puramente teorica. Alcuni episodi si sono già verificati, come dimostrano gli incidenti che hanno coinvolto veicoli a guida autonoma, talvolta con conseguenze particolarmente gravi. La rilevanza del problema emerge soprattutto considerando la struttura dei reati colposi, nell'ambito dei quali l'utilizzo di sistemi capaci di assumere decisioni in modo autonomo impone di individuare non soltanto le probabilità che si verifichi un evento dannoso, ma anche il soggetto o i soggetti ai quali possa essere attribuita la relativa responsabilità penale.

Naturalmente, nei reati d'evento è necessario che si verifichi l'evento lesivo previsto dalla fattispecie incriminatrice affinché il reato possa configurarsi. Tuttavia, le regole cautelari svolgono una funzione preventiva fondamentale, poiché stabiliscono anticipatamente quali comportamenti siano consentiti e quali, invece, debbano essere evitati proprio per ridurre il rischio di eventi dannosi.

Di conseguenza, la necessità di disciplinare adeguatamente scenari particolarmente gravi, come un incidente mortale o un evento catastrofico, produce effetti anche sulle attività ordinarie nelle quali vengono utilizzate tecnologie intelligenti. La valutazione delle possibili conseguenze estreme contribuisce, infatti, a determinare le condizioni e i limiti entro cui tali attività possono essere legittimamente svolte.

Resta quindi centrale il problema di stabilire quali cautele debbano essere adottate per consentire l'utilizzo di questi sistemi e come debbano essere ripartiti gli eventuali profili di responsabilità tra i diversi soggetti coinvolti nella loro progettazione, produzione, distribuzione e impiego. Si tratta di una valutazione che richiede un necessario bilanciamento tra interessi contrapposti, nel quale i benefici derivanti dall'innovazione devono essere confrontati con i rischi che essa

comporta, considerando anche la possibilità, seppur non necessariamente frequente, che si verifichino gli eventi più gravi e dannosi¹⁶.

Nell'ambito delle tecnologie di intelligenza artificiale, il tema della colpa giuridica ripropone alcune problematiche di fondo, di natura non soltanto giuridica, ma anche sociopolitica e antropologica. Tali problematiche assumono una particolare rilevanza nell'ipotesi in cui il contesto di realizzazione del reato colposo coinvolga un sistema dotato di intelligenza artificiale, rendendo ancora più complessa l'individuazione dei presupposti e dei criteri di imputazione della responsabilità.

L'incremento esponenziale delle tecnologie emergenti cui si è assistito negli ultimi anni ha determinato un profondo mutamento della società, producendo effetti su molteplici settori economici e ponendo le basi per una progressiva trasformazione degli stili di vita e delle modalità di lavoro. Tra le innovazioni di maggiore rilievo, l'intelligenza artificiale, e in particolare quella generativa, costituisce un fattore di sviluppo destinato a incidere su numerosi ambiti, influenzando non soltanto il modo di essere, ma anche quello di agire delle persone.

L'evolversi delle tecnologie e l'adeguamento degli ordinamenti.

Un esame completo dell'intelligenza artificiale non può prescindere dalla sua dimensione giuridica. La diffusione sempre più ampia di queste tecnologie all'interno delle diverse sfere della vita quotidiana, con conseguenze significative sia sul piano sociale sia su quello economico, ha infatti favorito un intenso confronto tra istituzioni, imprese e altri soggetti interessati sulle principali questioni giuridiche derivanti dal loro utilizzo¹⁷.

Il dibattito sulla regolamentazione dell'IA può essere ricondotto a tre interrogativi fondamentali: se sia effettivamente necessario introdurre una disciplina specifica; a quale soggetto debba essere attribuito il compito di regolamentare il fenomeno; e, infine, attraverso quali modalità e strumenti tale regolamentazione debba essere realizzata e quali obiettivi debba perseguire.

¹⁶ BECH J., *La società del rischio. Verso una seconda modernità* (1986), Roma, 2000.

¹⁷ SEVERINO P., *Intelligenza artificiale e diritto penale, Intelligenza artificiale, il diritto, i diritti, l'etica*, Giuffrè, Milano, 2020.

Queste questioni hanno assunto particolare rilevanza negli ultimi anni, dando origine a posizioni dottrinali differenti e, talvolta, profondamente contrapposte.

Secondo una prima impostazione, l'intervento legislativo dovrebbe essere limitato alle situazioni in cui risulti concretamente indispensabile. In questa prospettiva, gli strumenti giuridici già esistenti, accompagnati da un'adeguata interpretazione e applicazione della normativa vigente, potrebbero essere sufficienti a disciplinare gran parte delle problematiche poste dall'intelligenza artificiale.

Un diverso orientamento sostiene, invece, l'opportunità di predisporre una disciplina specificamente dedicata all'IA, in considerazione della crescente diffusione di queste tecnologie e delle loro caratteristiche peculiari, che possono determinare situazioni e problematiche non sempre adeguatamente affrontabili attraverso le categorie giuridiche tradizionali.

Quest'ultima posizione appare oggi sempre più affermata e trova un consenso crescente anche nel settore economico. Gli stessi operatori del mercato, inizialmente più cauti rispetto all'introduzione di nuove regole, sembrano riconoscere con maggiore chiarezza l'esigenza di disporre di un quadro normativo specifico e organico, capace di accompagnare lo sviluppo dell'intelligenza artificiale e, al contempo, di disciplinarne i rischi e le possibili conseguenze¹⁸.

Se la questione relativa alla necessità di regolamentare l'intelligenza artificiale sembra ormai aver trovato una risposta in senso affermativo, restano ancora aperti gli interrogativi riguardanti chi debba intervenire e con quali modalità. In particolare, ci si interroga sull'opportunità di adottare una disciplina più stringente o, al contrario, caratterizzata da una minore invasività, nonché sulla possibilità di affidare la regolamentazione prevalentemente alle autorità pubbliche oppure di valorizzare forme di autoregolamentazione privata. Ulteriore questione riguarda la scelta tra un approccio generale, applicabile trasversalmente ai diversi settori, e una disciplina settoriale, calibrata sulle specifiche caratteristiche dei singoli ambiti di utilizzo.

Questi interrogativi assumono particolare rilevanza perché la diffusione dell'IA, al pari delle precedenti trasformazioni tecnologiche, ha posto i legislatori

¹⁸ SCHERER M. U., *Regulating Artificial Intelligence Systems: Risks, Challenges, Competencies, and Strategies*, in Harv. JL & Tech., 2016.



di fronte alla necessità di individuare un delicato punto di equilibrio tra innovazione e tutela. Da un lato, è necessario favorire la ricerca, lo sviluppo e l'impiego delle tecnologie di IA, affinché possano contribuire alla crescita economica e al progresso sociale; dall'altro, occorre predisporre principi e regole idonei a garantirne un utilizzo sicuro, responsabile e rispettoso dei diritti e delle garanzie previste dall'ordinamento.

Il problema nasce proprio dalla possibile tensione tra queste due esigenze. Una regolamentazione particolarmente rigida potrebbe rallentare la ricerca e limitare lo sviluppo di nuove applicazioni, riducendo i benefici potenzialmente derivanti dall'innovazione. Al contrario, un quadro normativo eccessivamente permissivo potrebbe risultare insufficiente a prevenire o contenere i rischi associati all'utilizzo dell'intelligenza artificiale. La sfida regolatoria consiste, pertanto, nell'individuare un modello capace di accompagnare l'innovazione senza rinunciare alla necessaria tutela degli individui e della collettività.



Figura 6: fonte <https://www.peridirittiumani.com/2019/07/23/il-diritto-internazionale-umanitario-e-dei-diritti-umani-un-approccio-variegato/>

Cooperazione nazionale e internazionale. Alla ricerca di un delicato equilibrio.

Nell'ambito della regolamentazione dell'intelligenza artificiale, il coinvolgimento delle organizzazioni internazionali assume particolare importanza, poiché consente di affrontare in maniera coordinata le questioni giuridiche, etiche e sociali derivanti dalla diffusione di queste tecnologie. La dimensione globale dell'IA rende infatti necessario un confronto tra i diversi Stati e gli altri soggetti coinvolti, al fine di promuovere forme condivise di governance.

Le organizzazioni internazionali possono svolgere diverse funzioni a questo riguardo. Innanzitutto, costituiscono una sede neutrale nella quale gli Stati

e gli altri attori della comunità internazionale possono confrontarsi e cooperare, favorendo la definizione di principi, standard e regole comuni.

In questa prospettiva si inserisce il progetto relativo alla creazione di un Osservatorio Globale sull'IA (GAIO), concepito come uno strumento destinato a raccogliere e analizzare informazioni sullo sviluppo e sull'utilizzo dell'intelligenza artificiale. Tra le attività previste rientrano anche la realizzazione e il costante aggiornamento di un registro degli incidenti e delle applicazioni emergenti, nonché la pubblicazione di rapporti periodici sullo stato dell'IA a livello globale.

Un'iniziativa di questo tipo mira a favorire uno sviluppo dell'intelligenza artificiale coerente con i principi e gli standard internazionali, rafforzando la trasparenza e la responsabilità degli operatori. Al tempo stesso, il monitoraggio sistematico delle tecnologie e dei relativi rischi può contribuire a individuare tempestivamente le criticità e ad adottare misure idonee a limitarne le possibili conseguenze negative.

La *Roadmap for Digital Cooperation*¹⁹ è il documento pubblicato nel 2020 dalle Nazioni Unite con l'obiettivo di delineare una serie di raccomandazioni rivolte alla comunità internazionale, volte a promuovere un ambiente digitale inclusivo, sicuro e rispettoso dei diritti umani. Prendendo atto dell'importanza centrale assunta dalle tecnologie digitali nel panorama globale²⁰, la Roadmap individua alcune linee d'azione strategiche finalizzate a favorire una più ampia diffusione dei benefici offerti dalle ICT e, al contempo, a ridurre i potenziali rischi connessi al loro utilizzo. Il documento affronta, tra le altre, questioni sempre più complesse quali l'intelligenza artificiale, la protezione dei dati, la sorveglianza, la privacy, l'ecosistema di Internet, la responsabilità online e la sicurezza digitale, sottolineando la necessità di uno sforzo multilaterale volto a rafforzare la cooperazione digitale internazionale e a fare della governance digitale una priorità dell'agenda globale delle Nazioni Unite²¹.

¹⁹ United Nation, General Assembly, *Road map for digital cooperation*, Report of the Secretary-General, 22 May 2020.

²⁰ Resolution adopted by the General Assembly on 21 September 2020 - Declaration on the commemoration of the seventy-fifth anniversary of the United Nations, United Nations.

²¹ ABBA L., articolo in Rivista italiana di informatica e diritto, *Carenze attuali e soluzioni future nei meccanismi per la cooperazione digitale*, ISSN 2704-7318, Fascicolo 1-2021.

Le Nazioni Unite e l'Unione europea svolgono un ruolo centrale nel promuovere la cooperazione internazionale in materia di intelligenza artificiale e nel favorire l'elaborazione di principi e regole diretti a garantirne un utilizzo sicuro, affidabile e rispettoso dei diritti fondamentali. In particolare, le Nazioni Unite hanno inserito la cooperazione globale sull'IA tra le priorità della propria agenda sulla trasformazione digitale, sostenendo lo sviluppo di sistemi di intelligenza artificiale che siano affidabili, sicuri, sostenibili e fondati sulla tutela dei diritti umani.

Anche l'Unione Europea, attraverso l'*AI Act* e le proprie iniziative di cooperazione internazionale, persegue l'obiettivo di promuovere un modello di IA antropocentrico e affidabile, contribuendo allo stesso tempo alla costruzione di una più ampia governance internazionale della tecnologia.

La cooperazione tra organizzazioni quali ITU, UNESCO e OHCHR favorisce, inoltre, l'integrazione dei diritti umani, dei principi etici e degli standard internazionali nei processi di sviluppo e utilizzo delle tecnologie digitali e dell'intelligenza artificiale. Tale collaborazione non sembra necessariamente orientata alla creazione di un unico sistema normativo mondiale, quanto piuttosto alla progressiva definizione di principi, standard e strumenti di governance condivisi, capaci di rafforzare il coordinamento tra gli Stati e gli altri soggetti interessati. Un simile approccio risulta fondamentale per consolidare la governance dell'IA e favorire una distribuzione il più possibile equa dei benefici derivanti dal progresso tecnologico.

In questa prospettiva, le Nazioni Unite considerano prioritario anche il rafforzamento della fiducia e della sicurezza digitale attraverso un confronto inclusivo a livello internazionale. La *Roadmap for Digital Cooperation* evidenzia, infatti, l'esigenza di proteggere le tecnologie digitali che supportano servizi e infrastrutture essenziali per la collettività, tra cui alimentazione, acqua, abitazioni, energia, sanità e trasporti.

L'obiettivo è quello di limitare le vulnerabilità e orientare lo sviluppo tecnologico verso finalità di interesse generale, favorendo una crescita economica e sociale equa e sostenibile. Tale esigenza assume particolare rilevanza alla luce dell'aumento degli attacchi informatici e della diffusione della disinformazione online, fenomeni che possono compromettere il funzionamento di servizi e infrastrutture fondamentali.

Il rapporto tra intelligenza artificiale e relazioni internazionali può essere esaminato anche con riferimento al possibile contributo dell'IA ai processi decisionali strategici. Nell'ambito della teoria dei giochi e dell'analisi delle strategie, ad esempio, i sistemi di IA possono elaborare grandi quantità di dati, costruire differenti scenari e confrontare possibili strategie. Ciò potrebbe contribuire a ridurre l'incidenza di alcune componenti emotive che caratterizzano le decisioni umane, quali ansia, diffidenza o sospetto. Tale funzione potrebbe risultare particolarmente utile nelle situazioni in cui gli interessi personali dei decisori possono entrare in contrasto con gli obiettivi dell'istituzione di appartenenza.

Tuttavia, l'utilizzo dell'IA non garantisce automaticamente maggiore obiettività o razionalità. I risultati prodotti dai sistemi possono essere condizionati dai *bias* contenuti nei dati di addestramento, dalle scelte effettuate nella progettazione dei modelli e dalle istruzioni impartite dagli operatori. In questo modo, l'algoritmo può riprodurre o addirittura accentuare pregiudizi e distorsioni già presenti nel processo decisionale umano.

Ne consegue che la presunta neutralità dell'intelligenza artificiale può generare una falsa percezione di oggettività. Per questo motivo, l'utilizzo di tali strumenti dovrebbe essere accompagnato da adeguati meccanismi di supervisione e controllo umano, indispensabili per valutare criticamente i risultati e le raccomandazioni formulate dai sistemi automatizzati.

Per questo motivo, il principale motivo di preoccupazione non dovrebbe essere tanto l'ipotesi di un'IA capace di ragionare come un decisore politico, quanto il rischio che un sistema orientato secondo logiche simili a quelle di un decisore umano venga considerato, proprio perché intangibile e apparentemente estraneo a interessi personali ed emozioni, come necessariamente più neutrale e affidabile.

Un ulteriore settore nel quale l'intelligenza artificiale può produrre effetti significativi è quello della diplomazia, e in particolare dell'attività svolta dalle ambasciate. Le rappresentanze diplomatiche stanno progressivamente digitalizzando le proprie funzioni e attività, determinando una trasformazione delle modalità tradizionali attraverso cui viene esercitata la diplomazia. Questo processo non incide soltanto sull'organizzazione e sul funzionamento delle strutture diplomatiche, ma amplia anche il numero e la varietà degli interlocutori

con i quali esse possono entrare in contatto, includendo soggetti esterni alle istituzioni governative.

Tale evoluzione assume particolare rilievo nell'ambito della *public diplomacy*, attraverso la quale gli Stati comunicano con le opinioni pubbliche e con gli attori stranieri allo scopo di promuovere la propria immagine, i propri valori e i propri interessi, nonché di costruire e consolidare relazioni internazionali nel lungo periodo.

La digitalizzazione rappresenta, dunque, uno degli elementi centrali della trasformazione della diplomazia contemporanea. L'introduzione delle nuove tecnologie richiede agli Stati di adeguare rapidamente strutture, strumenti e modalità operative, favorendo così la progressiva affermazione della *Digital Public Diplomacy*²². In questo contesto, una delle principali questioni da affrontare riguarda la gestione dei dati, soprattutto qualora il ricorso alle tecnologie digitali e ai sistemi di intelligenza artificiale venga progressivamente integrato nei processi decisionali. Dati e informazioni rappresentano infatti una risorsa di rilevante valore strategico ed economico per associazioni, organizzazioni, istituzioni e imprese, ma sono al contempo caratterizzati da un elevato grado di sensibilità, in particolare quando contengono informazioni personali o riservate. La loro importanza aumenta ulteriormente quando vengono utilizzati nell'interazione tra esseri umani e sistemi di intelligenza artificiale, le cui prestazioni dipendono in larga misura dalla qualità, dall'accuratezza e dalla completezza dei dati impiegati. Dati inesatti, incompleti, non aggiornati o non adeguatamente strutturati possono infatti determinare risultati analitici distorti e, conseguentemente, condurre a decisioni non corrette o inefficaci.

Carta di Nizza e CEDU alla prova dell'AI.

I principali strumenti sovranazionali deputati alla tutela dei diritti fondamentali nello spazio europeo sono la Carta dei diritti fondamentali dell'Unione europea, comunemente denominata Carta di Nizza²³ e la

²² BAGNULO E., "Come cambia la diplomazia con l'intelligenza artificiale: vantaggi e incognite", Network Digital 360, 31 Marzo 2023.

²³ La Carta dei diritti fondamentali dell'Unione Europea, proclamata a Nizza il 7 dicembre 2000, ha acquisito, con l'entrata in vigore del Trattato di Lisbona e in virtù dell'art. 6 TUE, lo stesso valore giuridico dei Trattati dell'Unione Europea.

Convenzione europea dei diritti dell'uomo (CEDU)²⁴. Sebbene appartengano a sistemi giuridici distinti, entrambe rappresentano riferimenti fondamentali per la protezione dei diritti della persona: la prima opera nell'ambito dell'ordinamento dell'Unione Europea, mentre la seconda costituisce il principale strumento convenzionale del sistema del Consiglio d'Europa. La crescente diffusione delle tecnologie digitali impone, tuttavia, di verificare se tali strumenti siano sufficientemente adeguati a fronteggiare le nuove problematiche derivanti dall'evoluzione dell'intelligenza artificiale.

Per quanto riguarda i sistemi algoritmici capaci di elaborare grandi quantità di dati, l'interpretazione dei principi posti a tutela dei diritti fondamentali evidenzia innanzitutto la necessità di garantire il rispetto della dignità e dell'autonomia della persona. Questi valori costituiscono il presupposto per assicurare una tutela effettiva degli altri diritti fondamentali riconosciuti dalla Carta di Nizza e dalla CEDU.

A tali principi si affiancano la non discriminazione, la sicurezza, nonché i criteri di imparzialità, equità e trasparenza. L'obiettivo è fare in modo che i processi attraverso i quali i dati vengono raccolti, trattati e trasformati in informazioni siano, per quanto possibile, comprensibili, controllabili e verificabili.

Un elemento particolarmente significativo è rappresentato dalla qualità dei dati sui quali i sistemi di IA vengono sviluppati e addestrati. Le caratteristiche dei dati possono infatti incidere direttamente sui risultati prodotti dagli algoritmi. Errori nella fase di progettazione, addestramento o implementazione possono determinare l'esclusione, la sottovalutazione o la rappresentazione non corretta di determinate informazioni, portando il sistema a riprodurre o addirittura accentuare distorsioni già presenti nei dati o nel contesto sociale. In questo senso, anche l'UNESCO ha evidenziato il rischio che l'IA possa incorporare e amplificare bias, contribuendo a fenomeni di discriminazione, disuguaglianza ed esclusione.

I sistemi di IA possono, infatti, riprodurre pregiudizi e stereotipi legati, ad esempio, al genere, all'età, all'origine etnica o ad altre caratteristiche personali, quando tali distorsioni siano già presenti nei dati utilizzati oppure derivino dalle modalità di progettazione e funzionamento del sistema. Si parla in questo caso di

²⁴ Convenzione europea per la salvaguardia dei diritti dell'uomo e delle libertà fondamentali (CEDU), sottoscritta a Roma il 4 novembre 1950 ed entrata in vigore il 3 settembre 1953.

bias algoritmici²⁵, ossia di alterazioni sistematiche capaci di condizionare i risultati generati dagli strumenti automatizzati e di produrre conseguenze discriminatorie.

Tali distorsioni possono dipendere da diversi fattori, tra cui la qualità e la rappresentatività dei dati, i criteri utilizzati per selezionarli e organizzarli e le scelte effettuate durante la progettazione e l'addestramento degli algoritmi. Ne deriva che una corretta gestione e governance dei dati costituisce una condizione essenziale per ridurre i rischi discriminatori e migliorare l'affidabilità e l'equità dei risultati.

I rischi possono diventare particolarmente significativi quando sistemi algoritmici caratterizzati da errori progettuali o alimentati da dati incompleti e non rappresentativi vengono utilizzati in ambiti delicati. È il caso, ad esempio, delle procedure di selezione nel mercato del lavoro o delle valutazioni relative all'accesso ai servizi finanziari, nelle quali un algoritmo potrebbe contribuire a rafforzare discriminazioni fondate sull'età, sul genere o sull'origine etnica. Ancora più delicata è la situazione nel settore penale, qualora sistemi automatizzati vengano utilizzati per supportare decisioni capaci di incidere sulla libertà personale.

Anche il diritto alla riservatezza e alla protezione dei dati personali può essere esposto a rischi derivanti dall'utilizzo dell'IA. Un esempio è costituito dai sistemi di riconoscimento facciale, che possono consentire la raccolta e l'analisi di informazioni personali senza che l'interessato sia adeguatamente informato o, nei casi in cui sia necessario, abbia prestato il proprio consenso. Un ulteriore problema è rappresentato dai deepfake, attraverso i quali immagini, video o altri contenuti possono essere artificialmente creati o modificati fino a risultare difficilmente distinguibili dalla realtà. Questi strumenti possono essere utilizzati per alterare la percezione degli eventi, compromettere la reputazione delle persone o condizionare le decisioni di individui e istituzioni.

Ciò non significa, tuttavia, che tali problematiche siano intrinseche all'intelligenza artificiale. Fenomeni analoghi possono infatti derivare anche dall'impiego di tecnologie e strumenti differenti. Ciò che caratterizza l'IA è piuttosto la capacità di elaborare enormi quantità di informazioni, automatizzare

²⁵ VANLEENHOVE C., *Artificial Intelligence and the Law*, Cambridge, p. 123 ss, 2022.

determinati processi e operare su larga scala, caratteristiche che possono aumentare significativamente la rapidità, la diffusione e l'impatto di eventuali errori, discriminazioni o comportamenti dannosi.

Ne consegue che, affinché l'impiego dell'intelligenza artificiale risulti compatibile con la tutela dei diritti fondamentali, è necessario garantire tale tutela fin dalle prime fasi di progettazione e sviluppo dei sistemi, promuovendo al contempo un approccio responsabile alla loro implementazione e al loro utilizzo. Sebbene si sia ormai sviluppata una crescente consapevolezza circa le possibili ripercussioni dei sistemi di IA sulla privacy e sulla protezione dei dati personali²⁶, soprattutto in relazione all'elaborazione di grandi quantità di dati (Big Data), appare ancora più complessa l'analisi degli effetti che tali sistemi possono produrre sugli altri diritti fondamentali, soprattutto quando vengono integrati in processi decisionali o in sistemi di governance automatizzata.



Figura 7 : fonte <https://www.certifico.com/news/carta-dei-diritti-fondamentali-dell-unione-europea>

La CEDU e le linee guida del Consiglio d'Europa sull'IA.

L'intervento del Consiglio d'Europa si è sviluppato in risposta alla rapida evoluzione delle tecnologie basate sull'intelligenza artificiale e ai nuovi rischi che il loro impiego può comportare per la protezione dei diritti umani, molti dei quali sono tutelati dalla Convenzione europea dei diritti dell'uomo (CEDU). In questo contesto, nel 2018 il Consiglio d'Europa ha pubblicato uno studio di particolare rilevanza, realizzato da un gruppo di esperti nell'ambito dello *Steering Committee on Media and Information Society* (CDMSI), dal titolo *Algorithms and Human*

²⁶ CAGGIANO G., CONTALDI G, MANZINI P., Verso una legislazione europea su mercati e servizi digitali, Bari, 2021.

*Rights: Study on the Human Rights Dimensions of Automated Data Processing Techniques and Possible Regulatory Implications*²⁷.

Anziché puntare sull'introduzione immediata di un quadro normativo uniforme, il Comitato sottolinea l'esigenza di approfondire preliminarmente le conseguenze sociali e giuridiche derivanti dall'utilizzo degli algoritmi. Tale cautela appare particolarmente significativa considerando che, al momento dell'elaborazione dello studio, molte delle tecnologie analizzate si trovavano ancora in una fase iniziale di sviluppo. Viene inoltre evidenziata la necessità di un impegno comune tra i diversi soggetti interessati, attraverso un approccio multistakeholder fondato sulla ricerca transdisciplinare, sull'analisi di dati ed evidenze concrete e sulla condivisione di conoscenze ed esperienze.

In tale ottica, il Comitato ritiene che il Consiglio d'Europa, in quanto organizzazione paneuropea impegnata nella salvaguardia dei diritti umani, possa assumere un ruolo centrale nello studio delle conseguenze dell'intelligenza artificiale sull'effettivo esercizio dei diritti fondamentali, sia nei rapporti tra individuo e autorità pubbliche sia nell'ambito delle relazioni private.

Un ulteriore riferimento significativo è costituito dalla Carta etica europea sull'uso dell'intelligenza artificiale nei sistemi giudiziari e negli ambiti connessi, adottata dalla Commissione europea per l'efficienza della giustizia (CEPEJ)²⁸ del Consiglio d'Europa. Il documento esamina le principali problematiche etiche derivanti dall'utilizzo dell'IA nel settore giudiziario e individua una serie di principi destinati a garantire la compatibilità di tali tecnologie con i diritti fondamentali e con i principi dello Stato di diritto.

La Carta riguarda specificamente l'impiego dell'intelligenza artificiale nell'ambito della giustizia e individua principi sia sostanziali sia metodologici volti a orientare la progettazione, lo sviluppo e l'utilizzo dei sistemi algoritmici. I suoi destinatari sono molteplici e comprendono non solo gli operatori e gli sviluppatori coinvolti nella realizzazione e nell'utilizzo di tali strumenti, ma anche

²⁷ Committee of expert of Internet Intermediaries, Algorithms and Human Rights. <https://edoc.coe.int/en/internet/7589-algorithms-and-human-rights-study-on-the-human-rightsdimensions-of-automated-data-processing-techniques-and-possible-regulatory-implications.html>.

²⁸ Commissione istituita nel 2002 con l'obiettivo di monitorare e misurare la qualità dei sistemi giudiziari dei Paesi membri.

le autorità pubbliche e i legislatori nazionali, chiamati a predisporre un adeguato quadro normativo.

La CEPEJ riconosce il potenziale delle applicazioni di IA nel migliorare l'efficienza e la qualità della giustizia, purché il loro impiego sia accompagnato da adeguate garanzie per la tutela dei diritti fondamentali previsti dalla CEDU e dalla Convenzione n. 108 del Consiglio d'Europa in materia di protezione dei dati personali. A tal fine, viene sottolineata l'importanza di integrare la tutela dei diritti fondamentali già nelle prime fasi di progettazione e sviluppo dei sistemi di IA.

Questa impostazione si riflette nei concetti di “*ethical-by-design*” e, in particolare, di “*human-rights-by-design*”, secondo i quali i principi etici e le garanzie dei diritti fondamentali devono essere incorporati nella struttura stessa della tecnologia sin dalla fase progettuale, anziché essere introdotti soltanto successivamente, quando il sistema è ormai stato sviluppato²⁹.

Il CAHAI (*Ad hoc Committee on Artificial Intelligence*) ha contribuito in modo significativo ai lavori del Consiglio d'Europa, concentrandosi sul rapporto tra lo sviluppo e l'utilizzo dell'intelligenza artificiale e la protezione dei diritti umani, della democrazia e dello Stato di diritto. Nel corso della propria attività, il Comitato ha esaminato sia le opportunità offerte dall'IA sia i rischi connessi al suo impiego, valutando al contempo l'opportunità di predisporre un quadro giuridico specifico per disciplinarne lo sviluppo e l'utilizzazione.

Secondo l'impostazione elaborata dal Comitato, l'impiego delle tecnologie di IA dovrebbe essere accompagnato da adeguate garanzie di trasparenza e affidabilità, nonché da elevati standard nella qualità dei dati. Particolare attenzione dovrebbe essere inoltre riservata alla tutela dell'autonomia umana, alla possibilità di esercitare un controllo effettivo sui sistemi e alla presenza di adeguati meccanismi di supervisione. Gli individui dovrebbero, inoltre, poter contestare le decisioni adottate mediante sistemi automatizzati e disporre di strumenti di tutela effettivi qualora tali decisioni producano conseguenze sui loro diritti.

In quest'ottica, il CAHAI ha evidenziato la necessità di orientare lo sviluppo e l'utilizzo dell'intelligenza artificiale verso modelli compatibili con la

²⁹ CEPEJ, *European Ethical Charter on the Use of Artificial Intelligence in Judicial Systems and their environment*, cit., p. 8. Il Comitato (cd. T-PD) della Convenzione 108.

tutela dei diritti fondamentali. L'obiettivo è fare in modo che tali tecnologie possano sostenere e rafforzare i valori della democrazia e dello Stato di diritto, anziché determinare effetti in grado di comprometterne i principi essenziali.



Figura 8 : fonte <https://www.altalex.com/documents/news/2019/04/08/giustizia-predittiva>

La proposta di un regolamento europeo sull'Intelligenza Artificiale: “AI Act”.

La proposta di regolamento sull'intelligenza artificiale presentata dalla Commissione Europea nel 2021 ha rappresentato un passaggio fondamentale nell'elaborazione di un quadro normativo europeo in materia di IA, in risposta alle sollecitazioni provenienti dalle istituzioni dell'Unione Europea e, in particolare, dal Parlamento Europeo e dal Consiglio³⁰.

L'iniziativa aveva un duplice obiettivo: da un lato, garantire il corretto funzionamento del mercato interno; dall'altro, favorire lo sviluppo e l'utilizzo di sistemi di IA sicuri, affidabili e compatibili con la tutela dei diritti fondamentali. Al contempo, la proposta mirava a individuare e contenere i principali rischi associati all'impiego di tali tecnologie.

Uno degli elementi caratterizzanti della proposta era rappresentato dall'adozione di un approccio basato sul rischio, attraverso la classificazione dei sistemi di IA in diverse categorie, alle quali corrispondevano specifici obblighi e requisiti in funzione del livello di rischio. La disciplina prevedeva inoltre regole relative alla progettazione, alla commercializzazione e all'utilizzo dei sistemi, accompagnate da strumenti di vigilanza, controllo e sanzione destinati a garantirne l'effettiva applicazione.

³⁰ BRADFORD A., *Effetto Bruxelles. Come l'Unione europea regola il mondo*, Milano 2021.

Il processo legislativo si è successivamente concluso con l'adozione del Regolamento (UE) 2024/1689, noto come *AI Act*³¹, che rappresenta il primo intervento normativo organico dell'Unione europea specificamente dedicato all'intelligenza artificiale. Il regolamento introduce un insieme armonizzato di disposizioni finalizzate a favorire lo sviluppo e l'impiego di sistemi di IA sicuri e affidabili, assicurando al contempo un elevato livello di protezione della salute, della sicurezza e dei diritti fondamentali.



Figura 9: fonte <https://www.iconsulentiprivacy.it/approvato-il-regolamento-sullintelligenza-artificiale-ai-act-punti-fondamentali/>

L'*AI Act* rappresenta il primo tentativo dell'Unione Europea di predisporre una disciplina organica e sistematica dell'intelligenza artificiale, con l'obiettivo di conciliare il sostegno all'innovazione tecnologica con la salvaguardia dei diritti fondamentali. Attraverso un modello regolatorio fondato sulla valutazione dei rischi, l'Unione mira a favorire lo sviluppo e l'impiego di sistemi di IA che siano sicuri, affidabili e compatibili con i principi e i valori propri dell'ordinamento europeo, assumendo al contempo una posizione di particolare rilievo nel dibattito internazionale sulla regolamentazione e sulla governance dell'intelligenza artificiale.

L'iter legislativo ha preso avvio con la proposta di regolamento presentata dalla Commissione europea il 21 aprile 2021, denominata *Artificial Intelligence Act*. L'iniziativa mirava a introdurre una disciplina uniforme dell'intelligenza artificiale all'interno dell'Unione, attraverso l'armonizzazione delle norme applicabili e la predisposizione di un sistema comune per l'individuazione e la gestione dei rischi connessi all'utilizzo di tali tecnologie. La proposta teneva

³¹ Proposta di regolamento del Parlamento Europeo e del Consiglio (legge sull'intelligenza artificiale) e modifica alcuni atti legislativi dell'Unione (*AI Act*), 21 aprile 2021, COM/2021/206 final.

conto, inoltre, delle conseguenze economiche e sociali derivanti dalla diffusione dell'IA, ponendo particolare attenzione alla protezione della salute, della sicurezza e dei diritti fondamentali.

L'adozione dell'AI Act non costituisce, tuttavia, un intervento isolato, ma si inserisce in un percorso più ampio attraverso il quale le istituzioni europee hanno progressivamente elaborato strategie e strumenti destinati a disciplinare l'intelligenza artificiale. Tale processo ha acquisito particolare rilevanza a partire dal 2017, nell'ambito della più generale strategia europea per la trasformazione digitale e delle politiche riguardanti i dati, i servizi digitali, la robotica e le tecnologie emergenti.

Un momento particolarmente significativo di questo percorso è rappresentato dalla Comunicazione della Commissione europea del 25 aprile 2018, "L'intelligenza artificiale per l'Europa". Con tale documento è stata definita una strategia finalizzata a sostenere la ricerca e gli investimenti nel settore dell'IA, favorendone lo sviluppo e la diffusione e accompagnando, al contempo, i processi di trasformazione economica e sociale determinati dall'innovazione tecnologica.

La strategia delineata nel 2018 ha inoltre evidenziato la necessità di affiancare alla dimensione tecnologica ed economica un adeguato quadro etico e giuridico, fondato sui valori dell'Unione europea e sul rispetto dei diritti fondamentali. In questo modo, l'approccio europeo all'intelligenza artificiale è stato progressivamente orientato non soltanto verso il rafforzamento dell'innovazione e della competitività, ma anche verso la costruzione di un sistema basato sulla fiducia, sulla sicurezza, sulla responsabilità e sulla tutela della persona³².

La Commissione Europea ha inoltre promosso la costituzione di un Gruppo indipendente di esperti di alto livello sull'intelligenza artificiale (*High-Level Expert Group on Artificial Intelligence – HLEG AI*), incaricato di approfondire le principali questioni connesse allo sviluppo e all'impiego di tali tecnologie. Nell'aprile 2019 il Gruppo ha pubblicato le *Ethics Guidelines for Trustworthy AI*, nelle quali sono stati individuati tre requisiti essenziali affinché un sistema di intelligenza artificiale possa essere qualificato come affidabile (*trustworthy*): il

³² L'Intelligenza artificiale per l'Europa, 25 aprile 2018, COM(2018) 237 final, pp. 3-4. Oltre alla Carta dei diritti, la Commissione si riferisce all'articolo 2 del TUE per la definizione dei valori fondamentali dell'Unione.

rispetto della legge, la conformità ai principi etici e la robustezza, intesa sia sotto il profilo tecnico sia in relazione alle sue implicazioni sociali.

Le indicazioni elaborate dal Gruppo di esperti hanno contribuito alla progressiva definizione del modello europeo di governance dell'IA, caratterizzato dalla ricerca di un equilibrio tra la promozione dell'innovazione tecnologica e della competitività e la necessità di garantire un'adeguata protezione dei diritti fondamentali. Questa impostazione ha successivamente trovato espressione anche nel processo legislativo culminato nell'adozione dell'AI Act, il cui sistema regolatorio si fonda principalmente sulla classificazione delle applicazioni di IA in relazione al livello di rischio che esse comportano.

Il regolamento prevede, infatti, una distinzione tra pratiche di intelligenza artificiale vietate, sistemi ad alto rischio, sistemi sottoposti a specifici obblighi di trasparenza e applicazioni caratterizzate da un livello di rischio minimo o trascurabile. Una particolare attenzione è dedicata alle pratiche ritenute incompatibili con i principi e i valori fondamentali dell'Unione europea. Tra queste sono ricomprese, in determinate circostanze, forme di manipolazione o di sfruttamento del comportamento umano suscettibili di arrecare danni rilevanti alle persone, oltre alle altre pratiche espressamente individuate dal regolamento come vietate.

L'*AI Act* deve inoltre essere considerato in rapporto alla disciplina europea relativa alla protezione dei dati personali, e in particolare al GDPR. I due strumenti normativi, pur presentando profili di interconnessione, perseguono finalità e operano secondo ambiti differenti. Il GDPR è principalmente diretto a disciplinare il trattamento dei dati personali e a tutelare i diritti degli interessati, mentre l'*AI Act* introduce specifici requisiti e obblighi riguardanti la progettazione, lo sviluppo, l'immissione sul mercato e l'utilizzo dei sistemi di intelligenza artificiale. Ne consegue che l'*AI Act* non sostituisce la disciplina in materia di protezione dei dati personali e non introduce, in termini generali, un divieto del processo decisionale automatizzato, ma si inserisce nel quadro normativo europeo come disciplina complementare volta a regolare i rischi propri delle tecnologie di IA.

È dunque sulla base di un approccio graduato secondo il livello di rischio che si struttura l'*AI Act*, adottato definitivamente dall'Unione europea nel 2024. Il Regolamento (UE) 2024/1689 è stato approvato dal Parlamento Europeo

e successivamente adottato dal Consiglio dell'Unione Europea il 21 maggio 2024, divenendo il primo quadro normativo organico e completo al mondo volto a disciplinare l'intelligenza artificiale secondo un approccio basato sul rischio.

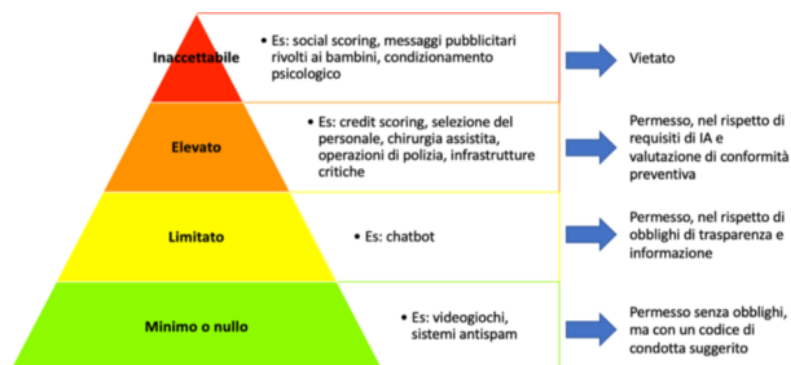


Figura 10: fonte <https://www.ispionline.it/it/pubblicazione/ai-act-la-sfida-europea-136317> pubblicato il 21 luglio 2023

Uno degli aspetti che maggiormente distingue l'AI Act dalle precedenti forme di autoregolamentazione e dagli strumenti di *soft law* in materia di IA è il suo carattere giuridicamente vincolante. Il regolamento adotta un modello basato sul grado di rischio: all'aumentare del rischio connesso a un sistema di IA corrispondono requisiti, obblighi e responsabilità più rigorosi per i soggetti coinvolti nel suo sviluppo, nella commercializzazione e nell'utilizzo.

Particolare rilievo assumono i sistemi di IA ad alto rischio, sottoposti a specifiche condizioni volte a garantirne sicurezza, affidabilità e rispetto dei diritti fondamentali prima della loro immissione sul mercato o della loro messa in servizio. La qualificazione come sistema ad alto rischio dipende soprattutto dalla finalità e dal contesto di utilizzo, non dalla tecnologia considerata in sé, tenendo conto anche dei possibili effetti sulla salute, sulla sicurezza e sui diritti delle persone.

Per tali sistemi sono previsti diversi obblighi, tra cui, nei casi stabiliti, la valutazione di conformità e la marcatura CE, oltre a prescrizioni relative alla gestione dei rischi, alla qualità dei dati, alla documentazione, alla trasparenza, alla sorveglianza umana, all'accuratezza, alla robustezza e alla sicurezza. Distinta dalla valutazione di conformità è la valutazione d'impatto sui diritti fondamentali, richiesta in determinate circostanze e finalizzata a individuare e analizzare le

possibili conseguenze dell'utilizzo del sistema sui diritti delle persone interessate³³.

Il Comitato per l'intelligenza artificiale (CAI) del Consiglio d'Europa ha avuto un ruolo determinante nella predisposizione della Convenzione quadro sull'intelligenza artificiale e i diritti umani, la democrazia e lo Stato di diritto, adottata il 17 maggio 2024 e aperta alla firma il 5 settembre dello stesso anno. La Convenzione rappresenta il primo trattato internazionale vincolante specificamente dedicato all'IA e mira a garantire che lo sviluppo e l'utilizzo dei sistemi di intelligenza artificiale siano compatibili con i diritti umani, la democrazia e lo Stato di diritto.

Sul piano europeo, il Libro bianco sull'intelligenza artificiale³⁴, pubblicato dalla Commissione nel febbraio 2020 ha consolidato l'indirizzo volto a conciliare innovazione tecnologica e tutela dei valori fondamentali dell'Unione, ponendo l'attenzione sull'adeguatezza del quadro normativo, sulla protezione dei diritti, sulla governance e qualità dei dati, sulla sicurezza e sulla responsabilità. Anche il Parlamento europeo, con la Risoluzione del 20 ottobre 2020 sulla responsabilità civile connessa all'IA, ha sottolineato l'esigenza di adeguare i tradizionali modelli di responsabilità alle peculiarità delle nuove tecnologie.

Questo percorso è confluito nella proposta di *AI Act*, che ha introdotto un sistema di obblighi e controlli proporzionati al livello di rischio dei diversi sistemi di IA, prevedendo misure applicabili sia prima sia dopo la loro immissione sul mercato e perseguendo finalità quali la tutela della sicurezza e della salute, dei diritti fondamentali e dei consumatori³⁵.

Per quanto riguarda la responsabilità civile, occorre distinguere la disciplina prevista dall'*AI Act*. A livello europeo, il tema dei danni derivanti dall'utilizzo dell'IA ha suscitato particolare attenzione, soprattutto a causa delle difficoltà probatorie legate alla complessità e all'autonomia di alcuni sistemi, tanto da ipotizzare strumenti volti ad agevolare la prova del danneggiato e, in determinati casi, l'introduzione di presunzioni sul nesso causale.

³³ CONTALDI G., *Intelligenza artificiale e dati personali*, in *Ordine internazionale e diritti umani*, 2021, pp. 1193-1213.

³⁴ Libro Bianco sull'Intelligenza artificiale, 19 febbraio 2020, COM/2020/65 final.

³⁵ CONTALDI G., *La proposta di regolamento sull'intelligenza artificiale e la protezione dei dati personali*, op. Cit.

La responsabilità, tuttavia, non può essere attribuita automaticamente al produttore, ma deve essere valutata in base al regime applicabile e al ruolo concretamente svolto da ciascun soggetto coinvolto, distinguendo tra produttore, operatore e altri eventuali responsabili.

L'impegno per un IA affidabile dell'Unione Europea.

La crescente diffusione dell'intelligenza artificiale interessa inevitabilmente anche il settore giudiziario, nel quale i sistemi algoritmici possono incidere sia sull'organizzazione degli uffici sia sui processi decisionali. In tale ambito, la CEPEJ del Consiglio d'Europa ha analizzato l'impiego delle tecnologie digitali e dell'IA nella giustizia, elaborando principi in materia di *cybergiustizia* e adottando la Carta etica europea sull'uso dell'intelligenza artificiale nei sistemi giudiziari.

Anche con riferimento alla regolamentazione dell'IA, i rapporti tra competenze dell'Unione europea e degli Stati membri devono essere interpretati alla luce del principio di leale cooperazione sancito dall'art. 4, par. 3, TUE, che impone alle parti di collaborare nell'esercizio delle rispettive competenze e di evitare condotte che possano ostacolare il perseguimento degli obiettivi dell'Unione³⁶.

Tale principio impone agli Stati membri e all'Unione di cooperare reciprocamente nell'adempimento dei rispettivi obblighi e nell'esercizio delle proprie competenze, evitando comportamenti suscettibili di compromettere il conseguimento degli obiettivi dell'Unione.

Nei rapporti con le organizzazioni internazionali, il principio di leale cooperazione assume particolare rilievo, poiché gli Stati membri possono essere soggetti a obblighi internazionali già assunti e devono coordinarli con quelli derivanti dall'ordinamento dell'Unione, nel rispetto del riparto delle competenze previsto dai Trattati. In presenza delle condizioni di cui all'art. 3, par. 2, TFUE, l'Unione può disporre di una competenza esterna esclusiva, impedendo agli Stati membri di assumere autonomamente impegni internazionali suscettibili di incidere sulle norme comuni dell'Unione.

³⁶ HYETT S., *The Duty of Cooperation: A Flexible Concept, the General Law of EC External Relations*, London, 2000.

Tale questione assume particolare importanza con riferimento alla Convenzione del Consiglio d'Europa sull'intelligenza artificiale, la cui partecipazione da parte dell'Unione consente di coordinare gli obblighi internazionali con il quadro normativo europeo in materia di IA, riducendo il rischio di sovrapposizioni o conflitti. La Convenzione, incentrata sulla tutela dei diritti umani, della democrazia e dello Stato di diritto, presenta significativi elementi di convergenza con i valori dell'Unione e con l'impostazione dell'AI Act, contribuendo così a rafforzare il coordinamento tra la dimensione europea e quella internazionale della governance dell'intelligenza artificiale.

In tale prospettiva, il coordinamento dell'azione esterna dell'Unione nella negoziazione e nella conclusione della Convenzione risponde all'esigenza di tutelare non soltanto il corretto funzionamento del mercato interno, ma anche l'autonomia normativa dell'Unione e il suo quadro costituzionale e valoriale, sul quale si fonda la disciplina europea dell'intelligenza artificiale.

La partecipazione dell'Unione europea alla Convenzione quadro del Consiglio d'Europa sull'intelligenza artificiale.

Il Consiglio d'Europa (CdE), attraverso l'elaborazione della Convenzione quadro sull'intelligenza artificiale, ha trasformato in obblighi giuridicamente vincolanti alcuni dei principi già affermati nei precedenti strumenti e documenti internazionali in materia di IA. La Convenzione disciplina le diverse fasi del ciclo di vita dei sistemi di intelligenza artificiale, introducendo specifiche garanzie e obblighi destinati a trovare applicazione sia nel settore pubblico sia, entro i limiti stabiliti dalla stessa Convenzione, in quello privato.

La rilevanza della Convenzione per l'Unione Europea riguarda non solo il suo contenuto e i soggetti interessati, ma anche l'esercizio della competenza esterna dell'UE, ossia la capacità di concludere accordi internazionali. In tale prospettiva, il Consiglio dell'Unione, ai sensi dell'art. 218 TFUE e su proposta della Commissione, ha autorizzato l'avvio dei negoziati per la partecipazione dell'UE alla Convenzione del Consiglio d'Europa sull'IA, individuando nell'art. 114 TFUE la relativa base giuridica. Con la decisione (UE) 2024/2218 del 28 agosto 2024, il Consiglio ha inoltre autorizzato la firma della Convenzione, poi sottoscritta dall'Unione il 5 settembre 2024.

La partecipazione dell'UE trova fondamento anche nell'art. 3, par. 2, TFUE, che riconosce una competenza esterna esclusiva quando la conclusione di un accordo internazionale può incidere sulle norme comuni dell'Unione o modificarne la portata. Ciò assume particolare rilievo alla luce del quadro normativo europeo sull'IA e, in particolare, dell'AI Act. La partecipazione dell'Unione mira dunque a garantire la coerenza tra gli obblighi internazionali e l'acquis dell'UE, evitando possibili sovrapposizioni o interferenze con la normativa europea.

La Convenzione quadro del Consiglio d'Europa sull'intelligenza artificiale mira a conciliare innovazione tecnologica e tutela dei diritti umani, della democrazia e dello Stato di diritto, attraverso un approccio basato sulla valutazione dei rischi connessi all'IA. A tal fine, introduce principi quali dignità e autonomia della persona, trasparenza, supervisione, responsabilità, non discriminazione, tutela della vita privata e sicurezza.

Le disposizioni della Convenzione presentano numerose analogie con il Regolamento (UE) 2024/1689 (*AI Act*); per questo, la sua attuazione nell'Unione è ricondotta all'AI Act e alle altre pertinenti norme dell'acquis europeo. La partecipazione autonoma dell'UE alla Convenzione risponde quindi all'esigenza di garantire coerenza tra gli obblighi internazionali e il diritto dell'Unione, evitando possibili interferenze o modifiche delle norme comuni già vigenti in materia di IA³⁷.

È significativo, in questo senso, il parere dell'EDPS del 2022, secondo cui il mandato del Consiglio d'Europa presenta una portata più ampia rispetto all'impostazione prevalentemente orientata al mercato interno dell'AI Act. La Convenzione assume infatti una prospettiva più generale, volta a tutelare diritti umani, democrazia e Stato di diritto, promuovendo anche il confronto pubblico, la partecipazione degli stakeholder e l'alfabetizzazione digitale.

L'*AI Act*, pur perseguendo anch'esso obiettivi di tutela della salute, della sicurezza e dei diritti fondamentali, è principalmente finalizzato alla creazione di un quadro normativo armonizzato per il mercato interno dell'Unione. La differenza tra i due strumenti risiede quindi soprattutto nella diversa ampiezza delle finalità: mentre l'*AI Act* privilegia l'armonizzazione della disciplina

³⁷ Corte di giust. parere 19 marzo 1993, 2/91, punto 25; 5 novembre 2002, C467/98, Commissione/Danimarca, punto 82

europea, la Convenzione del Consiglio d'Europa adotta una dimensione internazionale maggiormente incentrata sulla protezione dei diritti umani, della democrazia e dello Stato di diritto³⁸.

Nel quadro normativo vigente, tuttora in continua evoluzione, l'assenza di una delimitazione sempre chiara dell'ambito delle competenze dell'Unione rispetto a quello degli Stati membri rende complesso determinare, caso per caso, se questi ultimi debbano astenersi dall'assumere autonomamente impegni nei confronti di soggetti terzi, al fine di evitare possibili interferenze con il diritto dell'Unione, oppure se conservino una competenza esterna propria in relazione alla conclusione di accordi internazionali in materia di intelligenza artificiale.

La questione trova ulteriore conferma nella decisione del Consiglio che autorizza la firma della Convenzione quadro del Consiglio d'Europa sull'IA. I relativi considerando chiariscono infatti il rapporto tra le competenze dell'Unione e quelle degli Stati membri, fornendo indicazioni utili anche per definire l'ambito delle rispettive competenze esterne. In particolare, la decisione precisa che, nel settore dell'intelligenza artificiale, tale ripartizione deve essere valutata alla luce delle caratteristiche e della natura dei singoli strumenti giuridici interessati³⁹.

Questi aspetti richiamano il principio di leale cooperazione, che impone all'Unione di considerare le competenze e gli obblighi internazionali degli Stati membri, garantendo al contempo coerenza tra diritto europeo e internazionale. La Convenzione del Consiglio d'Europa rafforza, attraverso obblighi giuridicamente vincolanti, principi comuni in materia di IA, diritti umani, democrazia e Stato di diritto, in larga parte coerenti con i valori dell'Unione. La partecipazione dell'UE contribuisce così a coordinare la dimensione europea e internazionale della governance dell'IA, preservando l'autonomia e la coerenza dell'ordinamento europeo.

Conclusioni

Particolare attenzione è stata dedicata all'uso improprio dell'intelligenza artificiale da parte della criminalità, nonché alle opportunità e alle criticità

³⁸ WEATHERILL S., *The Fundamental Question of Minimum or Maximum Harmonisation, The Internal Market 2.0*, London, 2020.

³⁹ Decisione 2024/2218, considerando 8.

connesse all'impiego dell'IA come strumento di predizione, valutazione e supporto decisionale nella lotta al crimine. In tale contesto, è stata analizzata anche la risposta dell'Unione Europea, impegnata nella definizione di un quadro normativo armonizzato tra gli Stati membri, finalizzato a garantire un ambiente digitale sicuro e affidabile, nel rispetto dei diritti fondamentali.

L'avvento dell'intelligenza artificiale ha profondamente trasformato numerosi ambiti della vita quotidiana, influenzando anche il settore della giustizia. Da tempo, infatti, si discute della possibilità di introdurre il cosiddetto “giudice-macchina”, espressione di una concezione fortemente tecnicistica del diritto e della funzione giurisdizionale. Tuttavia, allo stato attuale, tale prospettiva non appare compatibile con i principi che governano l'esercizio della funzione giurisdizionale.

L'attività del giudice, infatti, non si esaurisce in un processo meramente logico-meccanico di applicazione delle norme, ma richiede interpretazione, valutazione delle circostanze del caso concreto e assunzione di responsabilità nella decisione. L'intelligenza artificiale può pertanto svolgere una funzione di supporto, ma non può sostituire il giudice nella decisione finale. Tale impostazione trova conferma anche nel quadro europeo: l'AI Act considera ad alto rischio determinati sistemi destinati ad assistere le autorità giudiziarie nell'interpretazione dei fatti e del diritto e nell'applicazione della legge al caso concreto, precisando che l'utilizzo dell'IA non deve sostituire il potere decisionale del giudice e che la decisione finale deve rimanere un'attività guidata dall'essere umano.

A ciò si aggiunge il riferimento all'art. 102 della Costituzione italiana, secondo cui la funzione giurisdizionale è esercitata da magistrati ordinari istituiti e regolati dalle norme sull'ordinamento giudiziario. Ne consegue che la sostituzione integrale del giudice umano con un sistema di intelligenza artificiale non risulterebbe compatibile con l'attuale assetto costituzionale della funzione giurisdizionale.

Diverso è il discorso relativo all'impiego dell'intelligenza artificiale nella fase investigativa, ambito nel quale si registra una crescente diffusione di strumenti tecnologici destinati a supportare l'attività delle forze di polizia e degli organi investigativi. Tali strumenti possono contribuire alle attività di prevenzione, alla sicurezza pubblica, all'analisi delle informazioni e

all'individuazione di possibili correlazioni utili alle indagini. In questo contesto si collocano, ad esempio, sistemi di analisi predittiva come XLAW, sviluppato e utilizzato dalla Polizia di Stato italiana, che ha impiegato gli alert prodotti dal software come supporto all'attività operativa.

L'impiego di tali strumenti deve tuttavia essere accompagnato da particolari garanzie, considerato che l'utilizzo dell'IA nel settore della *law enforcement* può incidere significativamente sui diritti fondamentali, sulla presunzione di innocenza, sui diritti della difesa e sul diritto a un ricorso effettivo. L'*AI Act* qualifica, infatti, come ad alto rischio diverse applicazioni destinate alle attività di contrasto al crimine, tra cui alcuni sistemi utilizzati per la valutazione del rischio di commissione o reiterazione di reati, per la valutazione dell'affidabilità delle prove e per il profiling nell'ambito delle attività di prevenzione, indagine e perseguimento dei reati.

Da quanto emerso, appare evidente che l'integrazione dell'intelligenza artificiale nel sistema della giustizia non rappresenti una mera innovazione tecnologica, bensì l'espressione di un nuovo rapporto tra tecnologia e diritto. La governance dell'IA richiede infatti un approccio multidisciplinare e una stretta cooperazione tra Stati, istituzioni e organizzazioni internazionali, chiamati ad affrontare in maniera coordinata le sfide giuridiche, etiche e sociali derivanti dall'utilizzo di tali tecnologie.

Questo nuovo rapporto tra diritto e innovazione deve fondarsi sulla tutela della persona e dei suoi diritti fondamentali, garantendo che la giustizia continui a essere orientata alle esigenze dell'essere umano. Ciò implica non solo un costante aggiornamento delle tecnologie impiegate, ma anche dei modelli organizzativi, dei processi di governance e delle competenze professionali necessarie per governare efficacemente la trasformazione digitale.

Bibliografia

ABBA L., articolo in Rivista italiana di informatica e diritto, *Carenze attuali e soluzioni future nei meccanismi per la cooperazione digitale*, ISSN 2704-7318, Fascicolo 1-2021.

BAGNULO E., "Come cambia la diplomazia con l'intelligenza artificiale: vantaggi e incognite", Network Digital 360, 31 Marzo 2023. BECH J., *La società del rischio. Verso una seconda modernità (1986)*, Roma, 2000.

- BRADFORD A., *Effetto Bruxelles. Come l'Unione europea regola il mondo*, Milano 2021.
- CAGGIANO G., CONTALDI G., MANZINI P., *Verso una legislazione europea su mercati e servizi digitali*, Bari, 2021.
- CALDWELL A., GRIFFIN T., *AI-enabled future crime*, in *Crime Science*, 2020.
- COSTA C., "Intelligenza artificiale e Giustizia: tempi ancora prematuri", AI4Business, maggio 2019.
- Di VIZIO F., *Prevenzione e investigazioni: l'uso di IA, big data e soluzioni tecnologiche in ambito finanziario e nel contrasto al riciclaggio (AML) e al finanziamento del terrorismo (CFT)*, in *DisCrimen*, 11 gennaio 2024.
- FINOCCHIARO G., *Intelligenza artificiale e responsabilità*, in *Contratto e impresa*, febbraio 2020.
- GARDNER H., *Intelligenze multiple*, Anabasi, Milano, 1994.
- GOODFELLOW I., Bengio Y., and Courville A., *Deep learning*. MIT Press, 2017.
- KAPLAN J., *Intelligenza artificiale*, LUISS University press, pp. 43 ss., Roma, 2018.
- HYETT S., *The Duty of Cooperation: A Flexible Concept, the General Law of EC External Relations*, London, 2000
- LOMBARDO E., *Sicurezza 4P. Lo studio alla base del software XLAW per prevedere e prevenire i crimini*, Mazzanti Libri, 2019.
- NAST C., "Il software italiano che ha cambiato il mondo della polizia predittiva", *Wired Italia*, maggio 2019.
- PIZZETTI F., *Intelligenza artificiale, protezione dei dati personali e regolazione*, pp. 191 ss, Giappichelli, Milano, 2018.
- RUFFOLO U., *Intelligenza Artificiale, machine learning e responsabilità da algoritmo*, in *Giur. it.*, luglio 2019.
- SCHERER M. U., *Regulating Artificial Intelligence Systems: Risks, Challenges, Competencies, and Strategies*, in *Harv. JL & Tech.*, 2016.
- SEVERINO P., *Intelligenza artificiale e diritto penale, Intelligenza artificiale, il diritto, i diritti, l'etica* Giuffrè, Milano, 2020.
- VANLEENHOVE C., *Artificial Intelligence and the Law*, Cambridge, p. 123 ss, 2022.
- WEATHERILL S., *The Fundamental Question of Minimum or Maximum Harmonisation, The Internal Market 2.0*, London, 2020.