

Intelligenza artificiale e novità normative in materia penale: dalla legge n. 132/2025 ai decreti attuativi

Clementina Salvi¹

1. Dalla regolazione dell'intelligenza artificiale alla responsabilità penale

Con la legge 23 settembre 2025, n. 132, recante disposizioni e deleghe al Governo in materia di intelligenza artificiale, il legislatore italiano ha introdotto una prima disciplina nazionale organica dell'IA nel quadro tracciato dal Regolamento (UE) 2024/1689 (AI Act). La legge interviene in vari settori e contiene già alcune disposizioni di rilievo penalistico, tra cui l'introduzione dell'art. 612-quater c.p. in materia di illecita diffusione di contenuti generati o manipolati con sistemi di intelligenza artificiale.

La legge n. 132/2025 costituisce, tuttavia, soltanto il primo passaggio dell'intervento nazionale. L'art. 24 delega il Governo ad adottare la normativa necessaria ad adeguare l'ordinamento all'AI Act e, più specificamente, a intervenire sulla realizzazione e sull'impiego illeciti dei sistemi di IA. In attuazione della delega sono stati predisposti due decreti legislativi. Il primo assume particolare rilievo sul versante penalistico: disciplina l'utilizzo dei sistemi di IA nell'attività di polizia e introduce nuove disposizioni di diritto penale sostanziale e processuale, oltre ad ampliare il perimetro della responsabilità amministrativa degli enti ai sensi del d.lgs. 8 giugno 2001, n. 231.

La relazione illustrativa chiarisce la logica complessiva dell'intervento: il diritto penale è chiamato a presidiare l'effettività delle prescrizioni regolatorie nei segmenti nei quali vengono in gioco beni di rango primario, quali la vita, l'incolumità individuale e pubblica e la sicurezza dello Stato. Il dato più significativo non risiede quindi semplicemente nell'introduzione di nuovi reati "commessi con l'IA", ma nel passaggio dalla regolazione dell'intelligenza artificiale alla rilevanza penale della sua governance. Obblighi tecnici e organizzativi propri del modello europeo di regolazione del rischio – sicurezza del sistema, sorveglianza umana e gestione delle diverse fasi del ciclo di vita – possono diventare parametri rilevanti ai fini della responsabilità penale.

2. Quando la mancata AI compliance assume rilevanza penale: il nuovo art. 437-bis c.p.

Il fulcro dell'intervento è rappresentato dall'introduzione dell'art. 437-bis c.p., rubricato "Omessa adozione di misure di sicurezza nei sistemi di intelligenza artificiale e alterazione illecita degli stessi". La disposizione ruota attorno a tre condotte: l'omessa adozione delle misure tecniche di sicurezza previste per la progettazione, l'addestramento, la produzione, l'immissione sul mercato o l'utilizzo professionale di sistemi di IA ad alto rischio; l'omessa adozione di misure

¹ Senior Advisor, Pollicino Advisory.

di sorveglianza umana; e, su un piano distinto, l'alterazione illecita del funzionamento di tali sistemi.

La collocazione sistematica è significativa. Il nuovo delitto viene inserito tra i delitti di comune pericolo mediante violenza e presenta una evidente affinità strutturale con l'art. 437 c.p., relativo all'omissione o rimozione dolosa di cautele contro gli infortuni sul lavoro. La scelta riflette i beni giuridici tutelati e la logica preventiva della fattispecie: non viene incriminata la mera non conformità regolatoria, ma una condotta che determini un pericolo concreto per la vita o l'incolumità individuale, ovvero, nelle ipotesi più gravi, per l'incolumità pubblica o la sicurezza dello Stato.

La fattispecie è quindi costruita come reato di pericolo concreto. La punibilità delle condotte omissive presuppone che dall'omessa predisposizione delle misure tecniche o di human oversight derivi un pericolo effettivamente accertabile per i beni protetti. Il trattamento sanzionatorio è graduato in ragione dell'interesse esposto a pericolo: la pena è più severa quando la condotta mette a rischio l'incolumità pubblica o la sicurezza dello Stato. Una logica analoga caratterizza la condotta attiva di alterazione del sistema, per la quale il legislatore prevede una cornice edittale più elevata, coerentemente con la maggiore offensività della manipolazione esterna e intenzionale del sistema.

Particolarmente rilevante è il rapporto tra la nuova fattispecie e il ciclo di vita dei sistemi ad alto rischio. La disposizione richiama espressamente progettazione, addestramento, produzione, immissione sul mercato e utilizzo professionale. Secondo la relazione illustrativa, tale elencazione è funzionale a coprire le diverse posizioni di garanzia riconducibili agli operatori individuati dall'AI Act, dal fornitore al deployer, passando per importatori e distributori. Il rischio penale non si concentra quindi esclusivamente sul momento dell'utilizzo finale del sistema, ma può emergere lungo l'intera catena di sviluppo e commercializzazione, a seconda degli obblighi concretamente gravanti sul soggetto.

È proprio questo collegamento a sollevare uno dei profili più delicati della nuova disciplina. Il precetto penale si innesta su un apparato regolatorio prevalentemente tecnico e in continua evoluzione. Stabilire quali misure di sicurezza o di sorveglianza umana fossero dovute nel caso concreto richiederà necessariamente di confrontarsi non solo con l'AI Act, con la normativa nazionale di attuazione, ma anche con gli standard tecnici applicabili. Si crea così un punto di contatto particolarmente stretto tra risk-based regulation e principio di legalità penale, con inevitabili questioni di determinatezza, prevedibilità e individuazione della regola cautelare violata.

In questa prospettiva assume rilievo la scelta di circoscrivere la responsabilità colposa alle ipotesi di colpa grave. La relazione illustrativa collega espressamente tale opzione all'esigenza di proporzionalità e alla necessità di evitare la criminalizzazione di ogni scostamento dallo standard tecnico in un settore caratterizzato da rapida evoluzione e dalla compresenza di standard nazionali, europei e internazionali. Il diritto penale interviene dunque come presidio ulteriore della sicurezza dei sistemi ad alto rischio, ma la soglia di criminalizzazione resta

ancorata alla gravità della violazione e alla concreta esposizione a pericolo dei beni tutelati.

3. Dalla responsabilità individuale alla responsabilità dell'ente

Il secondo passaggio di particolare rilievo per le imprese riguarda la responsabilità amministrativa degli enti. Il decreto introduce nel d.lgs. n. 231/2001 il nuovo art. 25-vicies, rubricato “Reati commessi con l’uso di sistemi di intelligenza artificiale”, inserendo nel catalogo dei reati presupposto sia il nuovo art. 437-bis c.p., sia l’art. 612-quater c.p. sui deepfake, introdotto dalla legge n. 132/2025.

Per il delitto di cui all’art. 437-bis è prevista una sanzione pecuniaria da seicento a mille quote; per il delitto di cui all’art. 612-quater una sanzione da duecento a settecento quote. In entrambe le ipotesi trovano inoltre applicazione specifiche sanzioni interdittive, tra cui la sospensione o revoca di autorizzazioni, il divieto di contrattare con la pubblica amministrazione, l’esclusione da agevolazioni e il divieto di pubblicizzare beni o servizi. La relazione illustrativa evidenzia come l’integrazione del catalogo dei reati presupposto sia volta a preservare la coerenza sistematica del d.lgs. n. 231/2001, mantenendo ferma la tipicità penale del fatto presupposto.

L’inclusione dell’art. 437-bis rende particolarmente evidente il passaggio dalla *AI compliance* alla responsabilità penale. La corretta individuazione delle responsabilità, l’adozione delle misure tecniche di sicurezza e l’effettività della sorveglianza umana non rilevano più soltanto ai fini della conformità al quadro regolatorio europeo, ma possono incidere sulla prevenzione di un reato presupposto della responsabilità dell’ente. Per le organizzazioni che sviluppano o utilizzano sistemi di IA ad alto rischio, ciò rende necessario verificare se la mappatura dei rischi e i modelli organizzativi esistenti siano adeguati a intercettare anche le nuove condotte penalmente rilevanti.

La scelta di includere anche l’art. 612-quater amplia ulteriormente questa prospettiva. Il legislatore collega infatti il sistema 231 non soltanto ai rischi derivanti dal funzionamento di sistemi ad alto rischio, ma anche agli usi illeciti dell’IA generativa e dei contenuti sintetici. Ne deriva l’esigenza di considerare, nella governance aziendale, non soltanto i rischi tecnici connessi ai sistemi impiegati, ma anche le modalità attraverso cui strumenti di generazione o manipolazione di contenuti possono essere utilizzati nell’interesse o a vantaggio dell’ente.

4. L’IA nelle indagini: identificazione biometrica remota e garanzie processuali

Il decreto interviene anche sul versante opposto: non soltanto sull’IA come possibile fonte di rischio penalmente rilevante, ma sull’IA come strumento dell’investigazione penale. Il nuovo art. 359-ter c.p.p. disciplina l’impiego, nelle indagini preliminari, di sistemi di intelligenza artificiale per l’identificazione biometrica remota in tempo reale, collocandosi nel perimetro dell’art. 5 dell’AI Act, che pone un divieto generale all’utilizzo di tali sistemi negli spazi accessibili al pubblico per finalità di law enforcement, salvo ipotesi tassativamente individuate e nel rispetto di stringenti garanzie.

La disciplina italiana individua specifiche ipotesi di impiego, tra cui la conferma dell'identificazione e la ricerca mirata di un indiziato, la localizzazione di un sospetto non identificato e la ricerca di un latitante destinatario di misura cautelare o di ordine di esecuzione. L'utilizzo è ulteriormente circoscritto attraverso il rinvio ai reati già contemplati dall'AI Act e mediante soglie di gravità previste dalla disciplina nazionale. È inoltre contemplata la ricerca di vittime di sottrazione, tratta o sfruttamento sessuale.

Alla particolare invasività dello strumento corrisponde un sistema rafforzato di garanzie. In via ordinaria, l'impiego è subordinato alla richiesta del pubblico ministero e all'autorizzazione del giudice per le indagini preliminari. Il provvedimento autorizzativo deve essere motivato e delimitare l'area geografica, le persone ricercate e il periodo di utilizzo, che non può eccedere quindici giorni, salva proroga con decreto motivato. Il richiamo ai requisiti europei di necessità e proporzionalità rafforza la natura eccezionale dello strumento, mentre la riserva di giurisdizione costituisce il principale presidio rispetto alla sua elevata intrusività. Nei casi di urgenza opera un meccanismo di attivazione e successiva convalida scandito da termini rigorosi. Quando non sia possibile attendere il provvedimento del pubblico ministero, gli ufficiali di polizia giudiziaria possono attivare il sistema, ma devono trasmettere la richiesta entro dodici ore; seguono la richiesta di convalida del pubblico ministero e la decisione del giudice entro i termini previsti. A chiusura del sistema, i risultati non possono essere utilizzati quando il sistema sia stato impiegato fuori dai casi consentiti dalla legge o in violazione delle principali garanzie procedurali. In tali ipotesi, dati personali, risultati e output devono essere cancellati, salvo che costituiscano corpo del reato.

5. Contenuti generati con IA e strumenti cautelari

Un ulteriore tassello riguarda il sequestro preventivo dei contenuti online. Il decreto modifica l'art. 104 delle disposizioni di attuazione del codice di procedura penale, ampliando l'ambito dei contenuti suscettibili di sequestro mediante ordine rivolto ai prestatori di servizi intermediari e includendo espressamente quelli generati con sistemi di intelligenza artificiale. La relazione illustrativa collega l'intervento alla necessità di adattare gli strumenti giudiziari alla circolazione di contenuti online sottoposti a interventi dell'IA, consentendo l'inibizione dell'accesso al contenuto e, secondo le modalità già previste dall'ordinamento, l'oscuramento delle relative risorse online.

La modifica si inserisce in una tendenza più ampia, già visibile nella legge n. 132/2025, a considerare il contenuto sintetico non soltanto come oggetto di obblighi di trasparenza o di regolazione digitale, ma anche come possibile oggetto di interventi penalistici e cautelari. Si rafforza così il raccordo tra diritto penale, disciplina processuale e governance delle piattaforme nella gestione dei rischi derivanti dall'uso illecito dell'IA generativa.

6. Conclusioni

Letti nel loro insieme, gli interventi della legge n. 132/2025 e della disciplina attuativa mostrano un'evoluzione significativa dell'approccio italiano all'intelligenza artificiale. Il diritto penale non interviene soltanto a valle, per sanzionare specifici usi dannosi della tecnologia, come nel caso dei deepfake. Con il nuovo art. 437-bis c.p. entra direttamente nella governance del sistema, attribuendo rilevanza alla mancata adozione di cautele tecniche e organizzative quando questa determini un pericolo concreto per beni fondamentali.

Per imprese e organizzazioni, questo passaggio rende sempre meno sostenibile una separazione netta tra AI Act compliance, cybersecurity, governance e criminal risk. L'individuazione delle responsabilità lungo il ciclo di vita del sistema, la documentazione delle misure di sicurezza, l'effettività della human oversight e l'aggiornamento della mappatura dei rischi 231 diventano parti di uno stesso processo di gestione preventiva. La compliance all'AI Act non coincide naturalmente con l'assenza di responsabilità penale, ma costituisce sempre più chiaramente uno degli elementi attraverso i quali il rischio penale connesso all'IA deve essere individuato, governato e documentato.

È questo il dato di maggiore interesse della nuova disciplina: l'AI governance cessa di essere soltanto una materia regolatoria e diventa, in determinate condizioni, anche una questione di prevenzione del rischio penale.