



## ***SULLA RESPONSABILITÀ PER DANNI CAUSATI DA DISPOSITIVI MOSSI DALL'INTELLIGENZA ARTIFICIALE: L'AMBITO SANITARIO.***

*Sommario: I. Premessa: il quadro normativo europeo e nazionale. – II. I profili di irresponsabilità. – III. Mappatura degli attuali usi diagnostici dell'IA nella pratica clinica (Sintesi dell'Appendice). – IV. Lettura della mappatura in chiave di responsabilità – V. Tendenze future nello sviluppo degli usi clinici diagnostici della IA generativa. – Appendice al documento.*

### **I. Premessa: il quadro normativo europeo e nazionale.**

Intelligenza artificiale: minaccia o risorsa? Strumento al servizio dell'uomo, o espediente per sostituirlo?

Il prepotente ingresso dell'IA nella vita umana e la connessa fusione in un contesto ibrido fatto di tecnologia e di realtà materiale apre scenari e schiude interrogativi un tempo inimmaginabili, rendendo possibile un'interazione (e non una simbiosi) uomo/macchina che, se edificata attraverso un approccio sobrio e vigile, potrebbe esprimere l'accettabile punto di equilibrio, costituito dal potenziamento delle capacità umane che dia vita a un'intelligenza “aumentata” piuttosto che aridamente e inquietantemente “artificiale”.

Una prospettiva, questa, foriera di nuove strade per il bene comune di una *magnifica humanitas* che sappia rinvenire nella tecnica non una forza antagonista rispetto alla persona, bensì lo strumento per un significativo miglioramento delle sue condizioni di vita. La sfida dell'esonante paradigma tecnocratico nel mondo globalizzato è ormai avvertita in tutta la sua potenza: è la corsa all'algoritmo più performante e alla banca dati più vasta. Non basta perciò regolarla: va disarmata e resa ospitale – è scritto nella recentissima enciclica leonina – e disarmare «*non significa rinunciare alla tecnologia ma impedirle di governare l'umano*».

Chi è chiamato a governare non può tuttavia che prendere le mosse dalla regolamentazione. Si spiega così che dopo una lunga stagione in cui le applicazioni iniziali avevano riguardato sistemi chiusi o autosufficienti – il gioco della dama, quello degli scacchi, i primi ed embrionali *robot* – così apparendo mezzo per il conseguimento di risultati scarsamente applicabili a problemi reali, il progredire di studi e ricerche ha posto i legislatori dinanzi alla necessità imperiosa di individuare regole e principi: la guida di un veicolo senza alcun intervento umano, un'ampia

gamma di applicazioni finanziarie, commerciali e scientifiche, le criticità e i rischi ad esse ancorati hanno imposto interventi multilivello dai quali è ormai impossibile prescindere.

**Sul piano eurolunionale**, fondamentale è il Regolamento 2024/1689 sull'IA (noto come *AI Act*) approvato nel giugno 2024 dal Parlamento europeo e dal Consiglio, dopo un *iter* legislativo protrattosi per oltre un triennio, e pubblicato sulla Gazzetta Ufficiale dell'UE il successivo 12 luglio.

Poiché i regolamenti sono applicabili in via immediata in tutti gli Stati membri, la scelta di ricorrere a questo strumento in luogo di una semplice direttiva (di per sé suscettibile di una più o meno sollecita e puntuale traduzione in leggi nazionali nei singoli Paesi) ha avuto l'effetto di determinare vincoli uniformi e automaticamente applicabili sull'intero territorio dell'Unione.

Alla base di questo importante atto normativo è l'intento di modulare la disciplina dell'IA alla luce del grado di rischio di lesione dei diritti fondamentali e delle libertà dei cittadini. I sistemi di IA sono stati pertanto classificati attraverso una quadripartizione, a seconda che comportino:

1. un rischio *inaccettabile* (sistemi vietati – Capo II);
2. un rischio *alto* (sistemi ammessi solo dietro garanzia di ossequio a precisi requisiti di trasparenza e di non esclusività algoritmica – Capo III);
3. un rischio *limitato* (sistemi sottoposti a obblighi di trasparenza- Capo IV);
4. un rischio *non valutabile* preventivamente (stante l'eterogeneità dei compiti dei modelli di IA per finalità generali - Capo V).

I Capi VI-IX investono plurimi profili (misure a sostegno dell'innovazione; governance; banca dati dell'UE per i sistemi ad alto rischio; monitoraggio successivo all'immissione sul mercato).

Il Capo X prevede per i sistemi caratterizzati da un rischio non significativo (sistemi volontariamente subordinabili alle prescrizioni di auspicabili codici di condotta) l'incoraggiamento dei singoli Stati membri all'elaborazione di tali codici, da adottarsi su base volontaria.

Seguono altri Capi di vario contenuto (deleghe di potere; sanzioni; disposizioni finali).

In questo quadro, meritevole di peculiare menzione per l'ambito sanitario si rivela l'art. 57, contemplante le c.d. *sandbox*, costituite da ambienti controllati in cui le strutture possono testare nuove soluzioni tecnologiche salvaguardando la sicurezza

dei pazienti in attività di telemedicina o di diagnostica predittiva, ovvero utilizzando dati clinici anonimi per addestrare modelli di IA senza ledere la *privacy* dei pazienti. Innovazione ed esigenze di *compliance* sono così armonicamente governate, raggiungendo insieme obiettivi di innovazione tecnologica atti a consentire la valutazione di rischi e benefici, e di protezione dei dati personali.

In particolare, le *regulatory sandbox* consentono di superare le rigidità dettate dalla necessità di cautela dell'evoluzione tecnologica rendendo possibile una sperimentazione regolata e controllata. Il menzionato art. 57 dell'*AI Act* impone per l'appunto spazi di sperimentazione normativa che ogni Stato è chiamato ad adottare entro il prossimo 2 agosto, in una logica di innovazione responsabile che convogli esigenze di sperimentazione e protezione dei diritti fondamentali dei singoli pazienti in un quadro proteso alla massima realizzazione della salute pubblica.

L'emanando decreto delegato deliberato dal Governo nel giugno scorso (ma ancora in testo provvisorio e non definitivamente varato) dedica grande attenzione all'attuazione di tale disposizione giacché introduce nel Capo II (artt. 2-34) un'intera Sezione (la sesta, artt. 26-31) riservata agli "Spazi di sperimentazione normativa", così istituendo, sotto la vigilanza diretta delle Autorità nazionali per l'IA, lo "Spazio di sperimentazione italiano" ai sensi del detto art. 57. In particolare: l'art. 26 prevede l'istituzione di tale "spazio di sperimentazione"; l'art. 27 ne disciplina la gestione strategica; l'art. 28 ne regola la gestione operativa e l'attuazione tecnica; l'art. 29 prevede ai fini dell'ottimizzazione delle risorse il coinvolgimento di altri attori competenti nell'ambito IA; l'art. 30 prevede ulteriori spazi di sperimentazione normativa per l'IA nei campi della difesa, della giustizia, delle attività bancarie, finanziarie e assicurative.

**Sul piano nazionale**, l'IA è disciplinata dalla legge 23 settembre 2025 n. 132, recante "*Disposizioni e deleghe al Governo in materia di intelligenza artificiale*", pubblicata in G.U. 25 settembre 2025 n. 223, ed entrata in vigore il 10 ottobre 2025.

L'art. 1 manifesta la volontà legislativa di promuovere «*un utilizzo corretto, trasparente e responsabile, in una dimensione antropocentrica, dell'intelligenza artificiale, volto a cogliere le opportunità*» e di garantirne la vigilanza «*sui rischi economici e sociali e sull'impatto sui diritti fondamentali*».

L'art. 3 sancisce «*il primato della persona sulla macchina*», e si traduce «*in tre corollari fondamentali*»:

- il principio di non esclusività della decisione algoritmica, secondo cui nessuna decisione che incida significativamente sui diritti e le libertà delle persone può

basarsi esclusivamente su un processo decisionale autorizzato, essendo imprescindibile un sostanziale contributo umano;

- il principio di conoscibilità degli algoritmi, in forza del quale i sistemi di IA devono essere trasparenti e comprensibili, tali da consentire agli interessati l'accesso alle informazioni significative sulla logica utilizzata;
- il principio di non discriminazione algoritmica, secondo cui le procedure debbono essere progettate in modo da prevenire effetti discriminatori, specie nei confronti delle categorie protette o vulnerabili.

Con specifico riguardo all'attività sanitaria, si trascrivono qui di seguito gli articoli da 7 a 10:

« Art. 7

*Uso dell'intelligenza artificiale in ambito sanitario e di disabilità*

1. *L'utilizzo di sistemi di intelligenza artificiale contribuisce al miglioramento del sistema sanitario, alla prevenzione, alla diagnosi e alla cura delle malattie, nel rispetto dei diritti, delle libertà e degli interessi della persona, anche in materia di protezione dei dati personali.*
2. *L'introduzione di sistemi di intelligenza artificiale nel sistema sanitario non può selezionare e condizionare l'accesso alle prestazioni sanitarie secondo criteri discriminatori.*
3. *L'interessato ha diritto di essere informato sull'impiego di tecnologie di intelligenza artificiale.*
4. *La presente legge promuove lo sviluppo, lo studio e la diffusione di sistemi di intelligenza artificiale che migliorano le condizioni di vita delle persone con disabilità, agevolano l'accessibilità, la mobilità indipendente e l'autonomia, la sicurezza e i processi di inclusione sociale delle medesime persone anche ai fini dell'elaborazione del progetto di vita di cui all'articolo 2, comma 1, lettera n), del decreto legislativo 3 maggio 2024, n. 62.*
5. *I sistemi di intelligenza artificiale in ambito sanitario costituiscono un supporto nei processi di prevenzione, diagnosi, cura e scelta terapeutica, lasciando impregiudicata la decisione, che è sempre rimessa agli esercenti la professione medica.*
6. *I sistemi di intelligenza artificiale utilizzati in ambito sanitario e i relativi dati impiegati devono essere affidabili, periodicamente verificati e aggiornati al fine di minimizzare il rischio di errori e migliorare la sicurezza dei pazienti.*

Art. 8

*Ricerca e sperimentazione scientifica nella realizzazione di sistemi di intelligenza artificiale in ambito sanitario*

1. *I trattamenti di dati, anche personali, eseguiti da soggetti pubblici e privati senza scopo di lucro, dagli Istituti di ricovero e cura a carattere scientifico (IRCCS), di cui al decreto legislativo 16 ottobre 2003, n. 288, nonché da soggetti privati operanti nel settore sanitario nell'ambito di progetti di ricerca a cui partecipano soggetti pubblici e privati senza scopo di lucro o IRCCS, per la ricerca e la sperimentazione scientifica nella realizzazione di sistemi di intelligenza artificiale per finalità di prevenzione, diagnosi e cura di malattie, sviluppo di farmaci, terapie*

*e tecnologie riabilitative, realizzazione di apparati medicali, incluse protesi e interfacce fra il corpo e strumenti di sostegno alle condizioni del paziente, salute pubblica, incolumità della persona, salute e sicurezza sanitaria nonché studio della fisiologia, della biomeccanica e della biologia umana anche in ambito non sanitario, in quanto necessari ai fini della realizzazione e dell'utilizzazione di banche di dati e modelli di base, sono dichiarati di rilevante interesse pubblico in attuazione degli articoli 32 e 33 della Costituzione e nel rispetto di quanto previsto dall'articolo 9, paragrafo 2, lettera g), del regolamento (UE) 2016/679 del Parlamento europeo e del Consiglio, del 27 aprile 2016.*

- 2. Ai medesimi fini, fermo restando l'obbligo di informativa in favore dell'interessato, che può essere assolto anche mediante un'informativa generale messa a disposizione nel sito web del titolare del trattamento e senza ulteriore consenso dell'interessato ove inizialmente previsto dalla legge, è sempre autorizzato l'uso secondario di dati personali privi degli elementi identificativi diretti, anche appartenenti alle categorie indicate all'articolo 9 del regolamento (UE) 2016/679, da parte dei soggetti di cui al comma 1, salvi i casi nei quali la conoscenza dell'identità degli interessati sia inevitabile o necessaria al fine della tutela della loro salute.*
- 3. Negli ambiti di cui al comma 1 o per le finalità di cui all'articolo 2-sexies, comma 2, lettera v), del codice in materia di protezione dei dati personali, di cui al decreto legislativo 30 giugno 2003, n. 196, è sempre consentito, previa informativa all'interessato ai sensi dell'articolo 13 del regolamento (UE) 2016/679, il trattamento per finalità di anonimizzazione, pseudonimizzazione o sintetizzazione dei dati personali, anche appartenenti alle categorie particolari di cui all'articolo 9, paragrafo 1, del medesimo regolamento (UE) 2016/679. È consentito altresì il predetto trattamento finalizzato allo studio e alla ricerca sui gesti atletici, sui movimenti e sulle prestazioni nell'attività sportiva in tutte le sue forme, nel rispetto dei principi generali di cui alla presente legge e dei diritti di sfruttamento economico dei dati relativi alle attività agonistiche che spettano a chi le organizza.*
- 4. L'Agenzia nazionale per i servizi sanitari regionali (AGENAS), previo parere del Garante per la protezione dei dati personali, tenendo conto di standard internazionali e dello stato dell'arte e della tecnica, può stabilire e aggiornare linee guida per le procedure di anonimizzazione di dati personali, di cui al comma 3, e per la creazione di dati sintetici, anche per categorie di dati e finalità di trattamento.*
- 5. I trattamenti di dati di cui ai commi 1 e 2 devono essere comunicati al Garante per la protezione dei dati personali con l'indicazione di tutte le informazioni previste dagli articoli 24, 25, 32 e 35 del regolamento (UE) 2016/679, nonché con l'indicazione espressa, ove presenti, dei soggetti individuati ai sensi dell'articolo 28 del medesimo regolamento (UE) 2016/679, e possono essere avviati decorsi trenta giorni dalla predetta comunicazione se non sono stati oggetto di provvedimento di blocco disposto dal Garante per la protezione dei dati personali.*
- 6. Restano fermi i poteri ispettivi, interdittivi e sanzionatori del Garante per la protezione dei dati personali.*

#### Art. 9

##### *Disposizioni in materia di trattamento di dati personali*

- 1. Il trattamento dei dati personali anche particolari come definiti dall'articolo 9 del regolamento (UE) 2016/679, con il massimo delle modalità semplificate consentite dal predetto regolamento per finalità di ricerca e sperimentazione anche tramite sistemi di intelligenza artificiale e machine learning, inclusi la costituzione e*

*l'utilizzo di spazi speciali di sperimentazione a fini di ricerca, anche mediante l'uso secondario dei dati personali, è disciplinato con decreto del Ministro della salute da emanare entro centoventi giorni dalla data di entrata in vigore della presente legge, sentiti il Garante per la protezione dei dati personali, gli enti di ricerca, i presidi sanitari nonché le autorità e gli operatori del settore.*

Art. 10

*Disposizioni in materia di fascicolo sanitario elettronico, sistemi di sorveglianza nel settore sanitario e governo della sanità digitale*

*1. Al decreto-legge 18 ottobre 2012, n. 179, convertito, con modificazioni, dalla legge 17 dicembre 2012, n. 221, dopo l'articolo 12 è inserito il seguente:*

*«Art. 12-bis (Intelligenza artificiale nel settore sanitario). - 1. Al fine di garantire strumenti e tecnologie avanzate nel settore sanitario, con uno o più decreti del Ministro della salute, di concerto con l'Autorità politica delegata in materia di innovazione tecnologica e transizione digitale e con l'Autorità delegata per la sicurezza della Repubblica e per la cybersicurezza e sentita la Conferenza permanente per i rapporti tra lo Stato, le regioni e le province autonome di Trento e di Bolzano, sono disciplinate le soluzioni di intelligenza artificiale aventi funzione di supporto alle finalità di cui all'articolo 12, comma 2. Con i decreti di cui al primo periodo sono individuati i soggetti che, nell'esercizio delle proprie funzioni, accedono alle soluzioni di intelligenza artificiale secondo le modalità ivi definite.*

*2. Per il supporto alle finalità di cura, e in particolare per l'assistenza territoriale, è istituita una piattaforma di intelligenza artificiale. La progettazione, la realizzazione, la messa in servizio e la titolarità della piattaforma di cui al primo periodo sono attribuite all'AGENAS in qualità di Agenzia nazionale per la sanità digitale. La piattaforma di cui al primo periodo eroga servizi di supporto:*

- a) ai professionisti sanitari per la presa in carico della popolazione assistita con suggerimenti non vincolanti;*
- b) ai medici nella pratica clinica quotidiana con suggerimenti non vincolanti;*
- c) agli utenti per l'accesso ai servizi sanitari delle Case della comunità.*

*3. Previa intesa in sede di Conferenza permanente per i rapporti tra lo Stato, le regioni e le province autonome di Trento e di Bolzano, l'AGENAS, con proprio provvedimento, esplicita i servizi di supporto di cui al comma 2.*

*4. La piattaforma di cui al comma 2 è alimentata con i dati strettamente necessari per l'erogazione dei servizi di cui al medesimo comma 2, trasmessi dai relativi titolari del trattamento.*

*L'AGENAS è titolare del trattamento dei dati raccolti e generati all'interno della piattaforma.*

*5. Previo parere del Ministero della salute, del Garante per la protezione dei dati personali e dell'Agenzia per la cybersicurezza nazionale, l'AGENAS, con proprio provvedimento, valutato l'impatto del trattamento, specifica i tipi di dati trattati e le operazioni eseguite all'interno della piattaforma, nonché le misure tecniche e organizzative per garantire un livello di sicurezza adeguato al rischio e per tutelare i diritti fondamentali e gli interessi dell'interessato, in coerenza con le disposizioni del regolamento (UE) 2016/679».*

*2. Dall'attuazione del presente articolo non devono derivare nuovi o maggiori oneri a carico della finanza pubblica. L'AGENAS provvede alle attività di cui al presente articolo con le risorse umane, strumentali e finanziarie disponibili a legislazione vigente».*

Con riferimento all'intero contenuto della legge in parola, deve sottolinearsi che essa è eminentemente rivolta all'enunciazione di principi di ordine generale ovvero riferiti a singoli settori (informazione e riservatezza dei dati personali; sviluppo economico; sicurezza e difesa nazionale; sanità; lavoro; professioni intellettuali; pubblica amministrazione; attività giuridica; ecc.), e che la concreta disciplina applicativa è invece affidata a uno o più decreti che il Governo è chiamato ad emanare entro un anno dall'entrata in vigore della legge (e quindi entro il 10 ottobre 2026). Si tratta infatti di una "legge-delega", ovverossia di una legge contenente le linee di insieme, da riempire di regole applicative attraverso i decreti delegati al potere esecutivo. Ed infatti:

- l'art. 16 delega il Governo ad adottare uno o più decreti legislativi *«per definire una disciplina organica relativa all'utilizzo di dati, algoritmi o metodi matematici per l'addestramento di sistemi di intelligenza artificiale...»*;
- l'art. 24 c. 1 delega il Governo ad adottare uno o più decreti legislativi *«per l'adeguamento della normativa nazionale al Regolamento UE 2024/1689 (il c.d. AI Act)»*;
- l'art. 24 c. 3 delega il Governo ad adottare uno o più decreti legislativi *«per adeguare e specificare la disciplina dei casi di realizzazione e di impiego illecito di sistemi di intelligenza artificiale»*.

In attuazione di queste deleghe, e più precisamente in diretta relazione con quella contemplata all'art. 24 c. 1, nelle passate settimane il Governo ha esaminato e approvato due decreti legislativi (non ancora pubblicati in G.U.):

- uno di adeguamento della normativa nazionale al Regolamento UE 2024/1689, in materia di poteri delle Autorità nazionali e di utilizzo dell'intelligenza artificiale nella formazione in ambito scolastico e universitario, nella pubblica amministrazione, **nonché nel campo sanitario**, delle professioni, della tutela dei lavoratori e della formazione dei magistrati;
- un altro di adeguamento della normativa nazionale al Regolamento UE 2024/1689 in materia di utilizzo dei sistemi di AI per l'attività di polizia e di responsabilità civile e penale.

Diretto rilievo ai presenti fini presenta soprattutto il primo di essi: tale decreto (che in realtà, come sopra chiarito, è ad oggi un mero testo provvisorio), investe eminentemente profili organizzativi di ordine generale: l'istituzione dell'"Autorità nazionale di notifica" (prevista dall'art. 70 del Regolamento UE<sup>1</sup>), l'individuazione

---

<sup>1</sup> «Ciascuno Stato membro istituisce o designa come autorità nazionale competente ai fini del presente regolamento almeno un'autorità di notifica e almeno un'autorità di vigilanza del mercato. Tali

delle autorità nazionali di vigilanza del mercato, il coordinamento con le Autorità settoriali di vigilanza del mercato stesso, i poteri dell'Agenzia per l'Italia Digitale, i poteri sanzionatori delle varie Autorità coinvolte, ecc.

Di particolare rilievo ai fini della sanità è il Capo III (artt. 35-50), che detta regole in materia di IA per i percorsi formativi, le professioni, il lavoro, **la sanità** e la pubblica amministrazione.

In questo ambito, si dispone che le istituzioni scolastiche promuovano attività formative sull'intelligenza artificiale, nonché moduli volti al potenziamento delle competenze scientifiche e tecnologiche mediante l'utilizzo degli strumenti dell'IA (art. 35); che dev'essere promossa la formazione dei docenti, oltre che degli studenti, e ciò anche per favorire un uso etico, responsabile, trasparente e critico delle tecnologie IA nel rispetto dei diritti fondamentali della persona, della protezione dei dati personali e dei principi di equità e non discriminazione (art. 36); che il Ministero del Lavoro e delle politiche sociali è chiamato a coordinare le misure volte a sostenere l'adattamento del capitale umano all'introduzione e alla diffusione dei sistemi di intelligenza artificiale, con l'obiettivo di favorire la transizione professionale, l'aggiornamento delle competenze e la prevenzione dei fenomeni di obsolescenza professionale, specie nei settori maggiormente esposti agli effetti della trasformazione tecnologica (art. 39); che il Ministero per la pubblica amministrazione indirizza e coordina le politiche di reclutamento, formazione e riqualificazione professionale dei dipendenti pubblici per consentire l'introduzione, l'utilizzo e la diffusione dei sistemi di intelligenza artificiale (art. 40); che il Ministero del Lavoro e delle politiche sociali assicura l'introduzione e l'utilizzo dei sistemi di IA nei contesti lavorativi nel rispetto dei principi di tutela della salute e della sicurezza dei lavoratori (art. 42); che le università e gli enti pubblici di ricerca debbano mirare alle attività di alfabetizzazione, formazione e divulgazione degli impieghi e dei limiti delle tecnologie di IA, compiti questi, cui sono anche chiamati ordini professionali e associazioni di categoria (artt. 43-45 e 47). Con specifico riguardo alle professioni sanitarie, l'art. 47 c. 6 stabilisce che per esse *«l'inclusione dei percorsi di formazione all'uso dei sistemi di intelligenza artificiale nei programmi di educazione continua in medicina è disposta con accordo in sede di Conferenza permanente per i rapporti tra lo Stato, le regioni e le province autonome di Trento e di Bolzano, da adottare entro sei mesi dalla data di entrata in vigore del presente decreto, ...»*. Inoltre l'art. 50 detta le regole sulla "Formazione in

---

*autorità nazionali competenti esercitano i loro poteri in modo indipendente, imparziale e senza pregiudizi...».*

materia di intelligenza artificiale nell'ambito della formazione continua in medicina – ECM – e della formazione manageriale in sanità”.

Occorre tuttavia anche rilevare che la bozza del decreto legislativo in parola si conclude con un'allarmante “Clausola di invarianza finanziaria” (art. 53), ove è sancito che le amministrazioni interessate debbano provvedere alle attività di cui al decreto stesso *«mediante l'utilizzo delle risorse umane, strumentali e finanziarie disponibili a legislazione vigente e, comunque, senza nuovi o maggiori oneri a carico della finanza pubblica»*. Un sano realismo giustifica il timore che la mancata previsione di risorse aggiuntive possa costituire un ostacolo al più razionale ed efficace perseguimento degli obiettivi.

Accanto a questi provvedimenti in vista dei quali il Parlamento, attraverso la legge in parola, ha delegato i poteri legislativi al Governo, devono inoltre collocarsi le normative secondarie cui la stessa legge n. 132/2025 fa riferimento. Così, con specifico riguardo al settore sanitario:

- in forza dell'art. 8 c. 4, l'Agenzia nazionale per i servizi sanitari regionali (AGENAS) *«può stabilire e aggiornare linee guida per le procedure di anonimizzazione di dati personali... e per la creazione di dati sintetici...»* (art. 8, c. 4); altre funzioni sono attribuite ad AGENAS dall'art. 10 commi 1 e 2;
- in forza dell'art. 9 il Ministero della salute è stato chiamato ad emanare un decreto volto alla concreta regolamentazione del trattamento dei dati personali. Ciò sarebbe dovuto avvenire *«entro centoventi giorni dalla data di entrata in vigore della presente legge»*, e quindi entro il 7 febbraio 2026. Ad oggi non consta che sia avvenuto. Il che si spiega, anche in ragione del fatto che tale decreto ministeriale dev'essere emanato *«sentiti il Garante per la protezione dei dati personali, gli enti di ricerca, i presidi sanitari nonché le autorità e gli operatori del settore»*. La genericità della dizione è tale per cui appare ben comprensibile che i tempi subiscano qualche dilatazione;
- in forza dell'art. 10 c. 1, che ha disposto l'inserimento dell'art. 12*bis* all'interno del d.l. 18 ottobre 2012 n. 179 (trattasi di un provvedimento recante molteplici misure *«per la crescita del Paese»*), il Ministero della salute è chiamato a disciplinare le soluzioni di intelligenza artificiale aventi funzione di supporto alle finalità di diagnosi, cura e riabilitazione, prevenzione, profilassi internazionale, studio e ricerca scientifica in campo medico, biomedico ed epidemiologico, programmazione sanitaria, ecc.

## **II. I profili di responsabilità.**

Chi risponde dei danni provocati da dispositivi deputati al compimento di operazioni in tutto o in parte sottratte al controllo diretto dall'uomo e affidate a sistemi di calcolo e a "decisioni" la cui gestione è demandata a un algoritmo?

Fino a pochi anni orsono non sorgeva alcun dubbio sul fatto che "l'agire" di una "macchina" chiamasse direttamente in causa l'operato di chi la aveva ideata, progettata, sviluppata, programmata. Il problema si è concretamente affacciato al cospetto di situazioni in cui la misura di automazione è divenuta talmente elevata da rendere la "macchina" stessa a sé stante rispetto a chi ha avuto parte nella sua realizzazione.

A chi ascrivere, in questo frangente, le conseguenze dell'azione posta in essere dal sistema di intelligenza artificiale? A chi lo ha configurato? A chi se ne avvantaggia? O direttamente al sistema medesimo, da qualificarsi "soggetto artificialmente intelligente"?

Muovendo dall'ultima ipotesi qui prospettata, può subito osservarsi che riconoscere soggettività a tale sistema non sarebbe in sé e per sé né impossibile, né stravagante, ove si consideri che sul piano del diritto da tempo immemorabile si riconosce autonomia soggettiva a entità diverse dalla persona fisica: il legislatore attribuisce infatti la capacità, intesa quale idoneità ad operare nel mondo delle relazioni giuridiche, anche ad enti astratti, nettamente distinti dagli esseri umani. Si tratta delle *persone giuridiche*, entità che stipulano contratti, assumono impegni giuridicamente vincolanti, compiono atti leciti e illeciti, rispondono dei danni eventualmente causati ad altri.

Esse sono, cioè soggetti di diritto, ancorché prive della corporeità e della carnalità. Sovente, è vero, sottendono una collettività di persone fisiche riunite tra loro per un certo scopo, quasi un'*universitas*, un corpo sociale costituente un macroantropo (si pensi a un'associazione, a un comitato, a una società commerciale). Ma non sempre è così, e ci si trova al cospetto di entità cui è riconosciuta la soggettività, quantunque non costituite da una pluralità di persone fisiche: è il caso della *fondazione*, soggetto di diritto costituito da un patrimonio vincolato a una determinata finalità. È vero che per operare nella realtà giuridica anch'essa ha bisogno dell'uomo, ma solo quale braccio operativo, strumento, organo agente non per sé, bensì in suo nome e conto.

Finzione giuridica? Struttura artificialmente creata dal legislatore? Centro di imputazione di diritti, di doveri, di responsabilità, di obblighi facenti capo direttamente ad esso? Mero espediente semantico evocativo di interessi superindividuali?

Non è questa la sede per discuterne, e si vuole qui soltanto sottolineare che, in via di principio, la diretta responsabilità del sistema di intelligenza artificiale non necessariamente costituirebbe un assurdo, purché accompagnata da precise regole e da articolate cautele.

Non sarebbe del tutto irragionevole, in altre parole, l'individuazione sul piano dello stretto diritto di una "personalità elettronica", sempreché ad essa si ricollegli – così come accade per i soggetti diversi dalle persone fisiche – non solo un'astratta responsabilità, ma anche una concreta capacità patrimoniale atta a fronteggiare le eventuali rivendicazioni di ordine risarcitorio provenienti dai terzi danneggiati. Anche se, a ben vedere, il problema finirebbe per spostarsi sull'identificazione dei soggetti chiamati a conferire le occorrente risorse, sì che in realtà la responsabilità finirebbe comunque per gravare su altri (vuoi persone fisiche, vuoi persone giuridiche).

L'agire del meccanismo algoritmico che procuri un danno finirebbe dunque per implicare sempre un'imputazione della responsabilità e conseguenti implicazioni patrimoniali, dirette o indirette, in capo a uno o più soggetti.

Il problema è che un esasperato sviluppo dell'intelligenza artificiale, fondato sull'autoalimentazione attraverso algoritmi di autoapprendimento potrebbe precludere l'identificazione di un responsabile: se la macchina si emancipa da chi l'ha ideata, realizzata, istruita, fino a divenire capace di decisioni autonome, come queste ultime potrebbero essere imputate, se dannose, al suo "genitore"? Senza considerare, poi, che i padri possono essere molteplici: il realizzatore dell'*hardware*, l'autore del *software*, il programmatore, l'addestratore, l'addetto alla manutenzione atta a prevenire malfunzionamenti.

È questo, in estrema e semplificatrice sintesi, il tema della responsabilità dell'agire di un sistema "intelligente".

Il tema presenta particolare delicatezza e complessità in ambito sanitario, campo peculiarmente delicato ed eticamente e bioeticamente sensibile. All'astratta responsabilità della struttura che si avvale del sistema e del medico cui il paziente è affidato potrebbero aggiungersi le plurime responsabilità (individuali? solidali?) di coloro che hanno contribuito alla creazione del sistema, al suo funzionamento, ai controlli manutentivi, e ad ogni altro apporto di attività e di competenze: questioni affrontate in chiave preventiva dal Regolamento UE 2024/1689 e nel prisma risarcitorio dalla Direttiva UE 2024/2853 concernente la responsabilità da prodotto.

In via del tutto astratta e teorica, ci si potrebbe anche interrogare sulla possibilità di ricorrere per via analogica alle disposizioni di legge che disciplinano le

responsabilità di genitori e rappresentanti legali, ravvisando così in determinate persone fisiche (fabbricanti, programmatori, addestratori, controllori) una responsabilità “educativa” di istruzione e/o di controllo e/o di vigilanza sull’operato del sistema affidato alle loro cure, ovvero alle previsioni normative che addebitano automaticamente a “padroni e committenti” i danni cagionati dai loro ausiliari nell’esercizio delle incombenze cui sono addetti, o ancora a quelle che sanciscono una legale presunzione di responsabilità (salva prova contraria) a carico di chi svolge un’attività di per sé pericolosa, sì da riguardarlo sempre quale responsabile a meno che non sia egli stesso a provare l’assenza di qualsivoglia sua colpa. Si tratterebbe tuttavia di forzature che non risolverebbero i problemi.

Fuori causa è senz’altro la tecnologia robotica chirurgica, se ed in quanto caratterizzata da apparecchiature prive di autonomia e costantemente seguite dal sanitario. Il problema si presenta invece con riguardo a quei sistemi che grazie ad algoritmi classificano e valutano le condizioni del paziente, effettuano diagnosi e suggeriscono terapie. Fino a qual punto il sanitario, per quanto competente, esperto e vigile, è in grado di controllarne (ed eventualmente rilevarne e correggerne) un imprevedibile malfunzionamento? In disparte le situazioni di clamorosa deviazione della “macchina” dai canoni di ragionevole risposta, quale sanitario sarebbe disposto ad assumere la responsabilità di disattendere le indicazioni? E come potrebbe essergli addebitata la mancata dissociazione da queste, se il sistema è costruito e programmato per dare autonomamente risposte di generale quasi assoluta affidabilità? Semmai, potrebbe in questi casi limite individuarsi una qualche responsabilità in chi ha rivestito funzioni di “addestramento” del sistema di intelligenza artificiale ovvero ruoli di preventivo controllo della sua efficienza nel duplice versante dell’*hardware* e del *software*.

Va da sé che nel caso limite di errore del sistema (paziente ammalato, ma dato per sano), il paziente che è stato privato delle cure (o il suo familiare che ne lamenta la morte per assenza di esse) agirebbe nei confronti della struttura sanitaria e del medico secondo i criteri di cui alla legge 8 marzo 2017 n. 24 (la c.d. legge Gelli-Bianco): responsabilità contrattuale della prima (con conseguente presunzione di responsabilità a suo carico, salva prova contraria) e responsabilità extracontrattuale del sanitario (con conseguente onere, per il paziente o per il congiunto che ne piange il decesso, di fornire la prova delle sue dirette responsabilità). È di tutta evidenza, d’altro canto, che la parte danneggiata si guarderà bene dal tentare di dimostrare il malfunzionamento della tecnologia, e saranno semmai la struttura sanitaria e/o il

medico a chiamare in garanzia altri soggetti (fabbricante, addestratore, manutentore, controllore, ecc.).

È tuttavia anche da mettere in conto che la complessità dei meccanismi di intelligenza artificiale possa, in fatto, rendere indimostrabile la colpa di uno o più dei soggetti dell'astratta catena di responsabilità. Dovrà dunque in questo caso concludersi che il paziente non curato o mal curato, ovvero il congiunto che ne lamenta il decesso, subirà il rigetto della domanda risarcitoria, e sarà magari anche condannato al pagamento delle spese di lite, delle complesse ed infruttuose consulenze tecniche e di quant'altro connesso alla sfortunata iniziativa giudiziaria? Ne scaturirebbe un evidente problema etico e bioetico: un malfunzionamento del sistema che esuli dalla sfera di controllo del medico e della struttura sanitaria, ma rispetto al quale non sia risultato possibile individuare, tra i vari soggetti potenzialmente responsabili o corresponsabili (produttore, commercializzatore, programmatore, addetto all'aggiornamento del sistema, addetto alla manutenzione, ecc.) chi effettivamente sia in colpa, può lasciare privo di ristoro il paziente danneggiato (o chi per lui)?

Parrebbe doversi concludere che il cattivo funzionamento del macchinario, fonte del pregiudizio arrecato al paziente, non possa in alcun caso escludere la responsabilità della struttura sanitaria quand'anche il mancato o difettoso funzionamento non sia dipeso da fatto ad essa imputabile. In altre parole, la disfunzione incolpevole, figlia di causa ignota o non accertabile, non può portare a concludere che il paziente è spogliato del risarcimento: la struttura sanitaria, in quanto chiamata all'erogazione di una prestazione assicurata attraverso l'utilizzo di un macchinario dotato di intelligenza artificiale, diverrebbe ineluttabilmente garante del suo funzionamento, sì da doversi fare carico della disfunzione pur in totale assenza di una propria colpa. Sembrerebbe quindi prospettabile, in siffatta ipotesi, una sua responsabilità di natura oggettiva, pur nell'incolpevolezza di quella disfunzione, **non potendo in nessun caso essere eluso il tema etico e bioetico della necessità di assicurare adeguato ristoro a chi abbia subito danni procuratigli da un sistema di intelligenza artificiale.** Va da sé, del resto, che l'accorta organizzazione sanitaria si sarà ragionevolmente premunita attraverso un adeguato strumento assicurativo.

Naturalmente, quando si parla di responsabilità per danni cagionati al paziente ci si riferisce alla responsabilità civile e quindi ai connessi profili risarcitori. Altro e distinto tema è quello della responsabilità penale, per sua natura sempre e soltanto personale. Mentre l'obbligazione risarcitoria può transitare da un soggetto ad un altro (si pensi al ruolo delle compagnie di assicurazioni, alla posizione del terzo che assuma

un debito altrui, al datore di lavoro che risponde dei danni cagionati dai suoi subordinati nell'esercizio delle incombenze), nulla di simile può accadere con riguardo all'imputabilità del reato, sia esso doloso o anche semplicemente colposo.

Già si è segnalato in precedenza (par. 1) che la legge nazionale 132/2025 ha delegato al Governo l'emanazione delle norme concernenti i casi di realizzazione e di impiego illeciti di sistemi di intelligenza artificiale (art. 24, c. 3). In particolare, il Governo è chiamato a prevedere *«strumenti, anche cautelari, finalizzati a inibire la diffusione e a rimuovere contenuti generati illecitamente anche con sistemi di intelligenza artificiale, assistiti da un sistema di sanzioni effettive, proporzionate e dissuasive»*, nonché a introdurre *«autonome fattispecie di reato, punite a titolo di dolo o di colpa, incentrate sull'omessa adozione o sull'omesso adeguamento di misure di sicurezza per la produzione, la messa in circolazione e l'utilizzo professionale di sistemi di intelligenza artificiale, quando da tali omissioni deriva pericolo concreto per la vita o l'incolumità pubblica o individuale o per la sicurezza dello Stato, e infine a precisare i criteri di imputazione della responsabilità penale delle persone fisiche e amministrative degli enti...»* (art. 24, c. 5).

Nell'accennato contesto, un ruolo di grande rilievo compete alla punta avanzata dell'IA, costituita dai c.d. "sistemi esperti", applicazioni tecnologiche costituite da *software* che riproducono artificialmente le prestazioni di una o più persone esperte in un determinato campo di attività.

Il "sistema esperto", adeguatamente istruito, è in grado di dedurre informazioni da un insieme di dati e da ragguagli di partenza; può mettere in atto, cioè, procedure di logica, induttiva o deduttiva, che lo conducono a una conclusione frutto dell'analisi dei fatti e delle circostanze di cui ha avuto notizia.

Mentre i programmi tradizionali presentano una struttura tipicamente algoritmica, e quindi operano tramite serie ordinate di istruzioni miranti a risolvere un determinato problema pervenendo a un risultato unico e certo, i sistemi esperti si basano sull'esperienza, e posti dinanzi a un problema prospettano un arco di risposte alternative, offrendo una visione trasparente delle decisioni assunte e giustificando le conclusioni raggiunte, così che l'utente possa sempre avere il controllo sul funzionamento interno del sistema.

Tali sistemi possono quindi assistere l'uomo nel processo decisionale fornendogli consulenza, formulando diagnosi, offrendo spiegazioni, suggerendo soluzioni alternative. Ci si riferisce qui, in particolare, alle *Chat GPT (Chat Generative Pre-trained Transformer)*, sistemi di comunicazione tra l'uomo e la

macchina (*chatbot*): un software che simula la conversazione umana e consente all'utente di interagire con un dispositivo digitale ponendo domande e ricevendo risposte.

La peculiarità è che questi sistemi possiedono la capacità di generare (di qui l'aggettivo *Generative*) in maniera testuale con linguaggio umano risposte scritte o sonore attingendo in maniera creativa alle informazioni acquisite nel *training* e nel corso delle successive elaborazioni: non si limitano quindi a un lavoro di analisi o di classificazione, essendo addestrati a fornire risposte sulla base di un'elevatissima quantità di dati forniti al sistema o da questo acquisiti (P sta per *Pre-trained*, mentre la T di GPT sta per *Transformer*, perché le risposte vengono convertite in linguaggio naturale).

*Chat GPT* non è dunque un semplice motore di ricerca, ma qualcosa di più, giacché può muoversi tra migliaia di siti diversi cercando con successivi aggiustamenti la risposta più completa: in quanto sistema generativo preaddestrato, è in grado di svolgere ruoli di assistente personale per quesiti legati alle esigenze quotidiane, di strumento didattico per approfondimenti nei più disparati ambiti, di collaboratore professionale nel settore sanitario. Può infine precisarsi, ad ogni buon conto, che queste specialissime attitudini non si sottraggono al quadro di riferimento giuridico sopra sommariamente tracciato.

### **III. Mappatura degli attuali usi diagnostici dell'IA nella pratica clinica (Sintesi dell'Appendice)<sup>2</sup>.**

Oggi l'uso diagnostico degli algoritmi di intelligenza artificiale è reale, ma non omogeneo: è concentrato soprattutto nelle specialità ad alta densità di immagini o di segnali standardizzati, in particolare oftalmologia, endoscopia digestiva, radiologia senologica e toracica, neuroradiologia d'urgenza, anatomia patologica digitale e, in misura più selettiva, cardiologia e dermatologia.

Nella pratica, la gran parte dei prodotti regolati non formula una diagnosi "autonoma" completa: funge invece da *concurrent reader*, *triage/prioritization tool*, *second read* o sistema di rilevazione di lesioni/anomalie.

L'eccezione più netta, oggi, è lo screening autonomo della retinopatia diabetica; un altro caso vicino all'autonomia, ma in un contesto di screening ben

---

<sup>2</sup> Contributo fornito dal prof. Giuseppe Remuzzi con la precisazione che si tratta di documento preparato da Alberto Arrigoni e Anna Caroli del laboratorio di Artificial Intelligence for Medicine dell'Istituto Mario Negri.

definito, è la Computer-Aided Detection (CAD) per tubercolosi su radiografia del torace, raccomandata dall'OMS<sup>3</sup>.

I casi d'uso con la migliore combinazione di regolazione, studi prospettici e impatto operativo documentato sono<sup>4</sup> : retinopatia diabetica su fotografia del fundus; CAD in colonscopia per polipi/adenomi; supporto alla lettura della mammografia/tomosintesi digitale mammaria; CAD su radiografia del torace per lo screening della tubercolosi; triage neuroradiologico per Ostruzione di grandi vasi / Emorragia intracranica; seconda lettura in anatomia patologica digitale per biopsie prostatiche.

In questi ambiti l'evidenza è più forte per l'accuratezza diagnostica, l'aumento del tasso di rilevazione e i miglioramenti dei workflow; è generalmente più debole per gli endpoint longitudinali come la mortalità, la riduzione dei tumori intervallari o gli esiti funzionali paziente-centrati.

Negli Stati Uniti, il principale riferimento regolatorio è l'autorizzazione della FDA tramite *De Novo* o *510(k)*. Nell'Unione europea, il percorso ordinario per i dispositivi medici, inclusi molti software diagnostici, è la marcatura CE secondo MDR/IVDR tramite organismi notificati e autorità nazionali competenti. L'EMA non svolge un ruolo generale equivalente a quello della FDA per i dispositivi medici, ma interviene in ambiti specifici, ad esempio fornendo supporto ai panel di esperti per dispositivi ad alto rischio o combinazioni farmaco-dispositivo.

Inoltre, il quadro regolatorio conferma la rapida crescita dell'offerta, ma anche la necessità di distinguere tra autorizzazione, adozione e impatto clinico. Secondo l'AI Index Report 2026 (Stanford University – Human-Centered Artificial Intelligence), entro il 2025 la FDA aveva autorizzato 1.357 dispositivi AI/ML-enabled, di cui 258 nel solo 2025, distribuiti in 17 specialità cliniche e provenienti da 693 aziende. La radiologia domina nettamente il panorama, con 1.039 dispositivi, pari al 76,6% del totale, confermando che l'IA diagnostica più matura resta concentrata nelle aree ad alta densità di immagini e di segnali standardizzati. Tuttavia, la stessa fonte evidenzia che l'evidenza clinica continua a inseguire la crescita regolatoria, con una quota molto ridotta di dispositivi supportata da trial randomizzati. In altre parole, la clearance regolatoria non equivale automaticamente a una prova del beneficio clinico nel mondo reale.<sup>5</sup>

---

<sup>3</sup> <https://www.who.int/publications/i/item/9789240110373>

<sup>4</sup> Per maggiori dettagli e bibliografia di riferimento si veda l'Appendice a questo Documento

<sup>5</sup> «The 2026 AI Index Report | Stanford HAI».

Le principali vulnerabilità dei sistemi di IA diagnostica sono trasversali: bias di selezione, spectrum bias (una rappresentazione parziale del fenomeno da parte della coorte dello studio), drift tra centri, scanner o dispositivi, gestione degli input non gradabili, scarsa trasparenza architetture, reporting non standardizzato e sorveglianza post-market disomogenea. Le nuove linee guida di reporting, come STARD-AI, insieme a CONSORT-AI, SPIRIT-AI e TRIPOD+AI, sono rilevanti proprio perché la qualità della letteratura implementativa è ancora eterogenea.

#### **IV. Lettura della mappatura in chiave di responsabilità<sup>6</sup>.**

Questa lettura collega la mappatura alla questione della responsabilità per i danni causati da sistemi diagnostici basati sull'intelligenza artificiale.

Il punto di partenza è che l'IA diagnostica non costituisce una categoria unica: varia notevolmente in base al livello di autonomia del sistema, al momento del workflow in cui interviene, alla possibilità di controllo umano e alla tracciabilità dell'errore. Questa distinzione potrebbe incidere sul modo in cui vengono attribuite le responsabilità.

La mappatura mostra infatti che, nella pratica clinica attuale, la maggior parte dei sistemi di IA non sostituisce integralmente il medico, ma si inserisce nel percorso diagnostico come strumento di supporto, secondo lettore, sistema di allerta, filtro preliminare o ausilio alla prioritizzazione. Solo in pochi casi l'IA si avvicina a una funzione diagnostica relativamente autonoma. In molti altri ambiti l'algoritmo contribuisce alla decisione, ma non l'esaurisce.

Questa funzione ausiliaria, tuttavia, non riduce la rilevanza del problema della responsabilità. Anche quando non decide formalmente al posto del medico, l'IA può orientare in modo significativo l'attenzione, i tempi, le priorità, le soglie di approfondimento e il grado di fiducia dell'operatore. Un sistema che segnala una lesione, classifica un caso come urgente, rassicura sull'assenza di reperti patologici o suggerisce la necessità di un referral non è un semplice strumento neutro: entra nella dinamica della decisione clinica e può modificarne concretamente l'esito.

Da questo punto di vista, il danno eventualmente causato dall'IA diagnostica difficilmente può essere ricondotto a un'unica causa. Non sempre si tratterà semplicemente di "errore dell'algoritmo" o di "errore del medico". Più spesso il danno potrà derivare dall'interazione tra più elementi: qualità dell'immagine o del segnale, popolazione su cui il sistema è utilizzato, limiti del modello, interfaccia con cui

---

<sup>6</sup> Contributo del prof. Giuseppe Remuzzi.

l'output viene presentato, integrazione nel workflow, formazione degli operatori, manutenzione, aggiornamenti, monitoraggio post-market e possibilità di ricostruire ex post ciò che è accaduto.

In questo senso, l'IA diagnostica rende più visibile il carattere già strutturalmente organizzativo della medicina contemporanea. La decisione clinica non nasce più soltanto dall'incontro tra medico e paziente, ma da un ecosistema composto da professionisti, dispositivi, software, dati, protocolli, infrastrutture digitali e assetti organizzativi.

L'errore, quindi, può collocarsi in punti diversi della catena: nell'acquisizione del dato, nella sua elaborazione, nella segnalazione prodotta dal sistema, nell'interpretazione dell'output, nella decisione finale o nella mancata sorveglianza del funzionamento nel tempo.

Fino a che punto il medico può essere ritenuto responsabile per non aver corretto o disatteso l'esito prodotto da un sistema progettato proprio per aumentare l'accuratezza diagnostica?

Il problema diventa particolarmente evidente quando l'algoritmo opera su pattern difficilmente verificabili dall'occhio umano o quando la sua affidabilità dichiarata induce il professionista a considerarlo una fonte particolarmente autorevole. In questi casi, il controllo umano rischia di essere più formale che sostanziale: il medico rimane teoricamente al centro della decisione, ma concretamente può non disporre degli strumenti per comprendere perché il sistema abbia prodotto un determinato output o per riconoscere tempestivamente un suo malfunzionamento.

Al tempo stesso, non si può neppure concludere che la presenza dell'IA esoneri il professionista da ogni responsabilità. Soprattutto nei casi in cui l'output sia visibile, verificabile e inserito in un quadro clinico più ampio, il giudizio medico mantiene un ruolo essenziale.

La difficoltà sta allora nel comprendere quale grado di vigilanza sia ragionevolmente esigibile nel caso concreto, tenendo conto della funzione effettiva del sistema, del suo grado di autonomia, della chiarezza dei suoi limiti, della qualità dell'input disponibile e del modo in cui la struttura sanitaria ne ha organizzato l'utilizzo.

Anche la struttura sanitaria assume infatti un ruolo centrale: è il soggetto che decide di integrare il sistema nella prestazione diagnostica, ne definisce il contesto

d'uso, forma gli operatori, organizza i flussi, predispone i controlli, gestisce gli aggiornamenti e conserva la tracciabilità delle decisioni.

Un ulteriore aspetto di riflessione riguarda la qualità informativa dell'output dell'IA: in pratica clinica non basta chiedersi se il sistema sia accurato in termini generali, ma occorre capire se il suo risultato sia comprensibile, interpretabile e utilizzabile nel caso concreto. In questa prospettiva, la spiegabilità dei modelli, la quantificazione dell'incertezza, la calibrazione e l'utilità clinica diventano elementi rilevanti anche nella riflessione sulla responsabilità. Un sistema trasparente, che segnala i propri limiti, distingue i casi incerti, comunica il grado di confidenza, mantiene una buona calibrazione nel contesto d'uso e offre elementi esplicativi dell'output, consente al medico non solo di integrarlo più ragionevolmente nel percorso diagnostico, ma anche di esercitare una forma effettiva di sorveglianza critica sull'IA, confrontandone il risultato con il quadro clinico e con gli altri dati disponibili.

Al contrario, un sistema chiuso, opaco, non calibrato sulla popolazione locale, privo di indicazioni sull'incertezza o validato solo su endpoint intermedi rischia di produrre una falsa impressione di oggettività e sicurezza, riducendo la capacità del medico di intercettare anomalie o incongruenze e rendendo più difficile riconoscere l'errore, prevenirlo e ricostruirne la responsabilità.

\* \* \* \*

Oggi l'IA diagnostica è clinicamente più solida quando il problema è circoscritto, l'input è standardizzato, il workflow è ben definito e la decisione finale resta umana oppure è confinata a uno screening stretto. In altre parole, l'IA è già pratica clinica corrente soprattutto come “occhio aggiuntivo” o “corsia preferenziale”, meno spesso come sostituto del medico. Le eccezioni principali sono poche, precisamente delimitate e legate a contesti di screening o di supporto altamente strutturati.

Il passaggio alla pratica clinica richiede inoltre di distinguere tra la prestazione tecnica del modello e l'utilità reale del sistema implementato. Il report AI Index 2026 mostra che gli esempi più convincenti di adozione su scala aziendale riguardano strumenti integrati nei workflow clinici — come la documentazione ambientale, la predizione della sepsi e gli strumenti generativi collegati alla cartella clinica — più che sistemi diagnostici pienamente autonomi. Questo suggerisce che il valore dell'IA in sanità dipende non solo dalla performance algoritmica, ma anche dalla capacità di essere incorporata in processi clinici governati, tracciabili, accettabili da medici e pazienti e in grado di produrre benefici misurabili.

## V. Tendenze future nello sviluppo degli usi clinici diagnostici della IA generativa<sup>7</sup>.

In questi sistemi diagnostici autonomi basati sull'intelligenza artificiale, in cui i Large Language Models sono addestrati su singoli compiti ben definiti, l'interfaccia degli algoritmi resta comunque il medico. Nei più recenti sviluppi, una nuova generazione di algoritmi si interfaccia direttamente con il consumatore o il paziente (Direct-to-Consumer tools), senza alcuna mediazione con un medico<sup>8</sup>.

Un esempio recente di questo tipo di sistemi è ChatGPHealth di Open AI<sup>9</sup>, rilasciato al pubblico (non ancora in Italia) nel Gennaio 2026, un chatbot progettato per stabilire un dialogo e fornire una prima linea di triage medico sulla base di domande mediche specifiche e sintomi riportati, suggerendo follow-up clinici senza alcuna intermediazione con personale medico.

Mancano ancora sufficienti dati indipendenti per la verifica di sicurezza ed affidabilità di questo strumento<sup>10</sup>, come preconditione essenziale, prima di procedere ad altre necessarie valutazioni di natura medica (importanza del rapporto medico-paziente, dell'esame obiettivo etc.), legale, etica, politico-sanitaria (conseguenze degli errori di triage<sup>11</sup>).

### *Conversational medical artificial intelligence*

L'Intelligenza Artificiale Diagnostica Conversazionale rappresenta una delle frontiere odierne degli sviluppi tecnologici in sanità digitale. I sistemi di Conversational Artificial Intelligence (CAI) non si limitano a rispondere a domande mediche fisse, ma conducono una vera e propria raccolta anamnestica strutturata attraverso un dialogo interattivo con il paziente. Questi si basano su Large Language Models generalisti su larga scala (c.d. Foundation Models o modelli di fondazione)

---

<sup>7</sup> Paragrafo elaborato in bozza dal compianto prof. Antonino Cattaneo.

<sup>8</sup> Per una tassonomia dei vari sistemi di IA generativa in campo sanitario, ed una discussione generale, si veda Angus et al (2025) AI, Health and Health Care Today and Tomorrow. The JAMA Summit Report on Artificial Intelligence. *Journal of American Medical Association* 334: 1650-1664.

<sup>9</sup> <https://openai.com/it-IT/index/introducing-chatgpt-health/>

<sup>10</sup> Il sistema Chat GPHealth è stato sottoposto ad un primo stress test strutturato indipendente, nel quale sono stati dimostrati numerosi errori di sotto- o sovra-valutazione in casi clinici estremi (Ramaswamy et al (2026) Chat GPHealth performance in a structured test of triage recommendations. *Nature Medicine* <https://doi.org/10.1038/s41591-026-04297-7>)

<sup>11</sup> Le conseguenze derivanti da errori nel triage sono asimmetriche: una sottovalutazione dell'esito del triage preclude accesso a trattamenti potenzialmente salva vita, mentre al contrario una sopravvalutazione del rischio del paziente determina un aumento di costi per il sistema sanitario.

che codificano in modo “enciclopedico” le conoscenze biomediche (“*LLM encode clinical knowledge*”, sic) (Generalist Medical AI)<sup>12 13</sup>

Il punto di riferimento globale in questo ambito è il sistema AMIE (Articulate Medical Intelligence Explorer)<sup>14</sup>, un LLM (sviluppato da Google Research) ottimizzato per il ragionamento diagnostico e la diagnosi differenziale, in modo autonomo o come ausilio interattivo al medico. L’applicazione di AMIE ad un dialogo clinico con pazienti simulati ha dimostrato in studi clinici simulati una precisione diagnostica e una qualità di gestione pari o superiore a quella dei medici di base<sup>15</sup>. Una prima valutazione di fattibilità clinica reale di questo sistema, come prima linea di interazione con pazienti reali è in corso<sup>16</sup>. I modelli più recenti integrano la conversazione testuale con capacità multimodali per analizzare contemporaneamente audio, video ed elettrocardiogrammi (Multimodal AMIE)<sup>17</sup>.

Un’area in espansione, di notevole rilevanza etica, e che richiede pertanto attente riflessioni, riguarda l’uso degli algoritmi di CAI come presidio per la salute mentale e la psicoterapia, senza intermediazione medica o psicologica<sup>18</sup>. La natura non diagnostica, ma direttamente interattiva col soggetto (sia esso “paziente” o meno) è di particolare rilevanza, per soggetti a rischio e vulnerabili.

Negli articoli pubblicati di recente (NB: pubblicati da parte delle aziende private produttrici delle tecnologie), la motivazione dichiarata sottostante allo sviluppo tecnologico della Conversational Multimodal Diagnostic Artificial Intelligence è collegata alla crisi dei sistemi sanitari, alla carenza di personale medico, alla necessità di liberare i medici da compiti “gravosi”, e di ottimizzare l’organizzazione delle strutture sanitarie. Da questo, il lancio della iniziativa AI co-

---

<sup>12</sup> Moor et al (2023) Foundation models for generalist medical artificial intelligence. *Nature* 616: 259–265

<sup>13</sup> Singhal Karan et al (2023) Large language models encode clinical knowledge. *Nature* 620: 172-180; Singhal et al (2025) Toward expert-level medical question answering with large language models. *Nature Medicine* 31, 953-950 descrivono il sistema MedPaLM messo a punto da Google DeepMind.

<sup>14</sup> McDuff et al (2025) Towards accurate differential diagnosis with large language models. *Nature* 642: 451–457 (2025).

<sup>15</sup> Tu Tao et al (2025) Towards conversational diagnostic artificial intelligence- *Nature* 642: 442-450.

<sup>16</sup> Brodeur et al (2026) A prospective clinical feasibility study of a conversational diagnostic AI in an ambulatory primary care clinic. ArXiv 2603.08448v3.

<sup>17</sup> Saab et al (2026) Advancing Conversational diagnostic AI with multimodal reasoning *Nature Medicine* 32: 1726–1736.

Shah Meet et al (2026) Towards Conversational Medical AI with eyes, ears and a voice. Submitted May 10th 2026 in ArXiv <https://doi.org/10.48550/arXiv.2605.09272>.

<sup>18</sup> Meadi MR et al (2025) Exploring the Ethical Challenges of Conversational AI in Mental Health Care: Scoping Review *JMIR Ment Health* 12:e60432 doi: [10.2196/60432](https://doi.org/10.2196/60432).

Jana Sedlakova & Manuel Trachsel (2023) Conversational Artificial Intelligence in Psychotherapy: A New Therapeutic Tool or Agent?, *The American Journal of Bioethics*, 23:5, 4-13, DOI: 10.1080/15265161.2022.2048739.

clinician research initiative<sup>19</sup>. La visione di medio periodo è che il AI Co-Clinician sia il terzo “soggetto (sic)” di una triade (Triadic care, che include il paziente, il medico e il AI co-clinician, e già questo linguaggio di equiparazione tra soggetti fa riflettere)<sup>20</sup>.

In realtà, al di là delle parole di cautela e necessità di maggiori approfondimenti, la visione tecnocratica e l’obiettivo, non esplicitamente scritti, è che algoritmi di Conversational and Multimodal Diagnostic Artificial Intelligence, possano essere proposti, nel prossimo futuro, come prima linea di visita medica dei pazienti, con il medico di base “sostituito” e trasformato ad AI Primary Care Physician<sup>21</sup>.

Si può anche facilmente prevedere che gli algoritmi di Conversational Diagnostic AI ad uso sanitario possano essere integrati nel prossimo futuro con algoritmi di coinvolgimento emotivo ed empatico, per simulare un rapporto medico-paziente positivo e creare una relazione di piena fiducia. Questi metodi sono oggi largamente implementati nei cosiddetti “emotional empathic AI companions” e resi popolari da piattaforme quali Character AI e Replika. Si tratta di chatbot basati sulla personalizzazione su larga scala di LLM che costituiscono avatar virtuali capaci di ingaggiare rapporti “personali” emulando apparenza emotiva (emotional

---

<sup>19</sup> Dal sito Google DeepMind (April 30<sup>th</sup> 2026) <https://deepmind.google/blog/ai-co-clinician/>  
“Health systems worldwide are striving for better outcomes, lower costs, and an improved experience for both patients and clinicians. However, progress is constrained by a global shortage of clinical experts, with the World Health Organization predicting a shortfall of more than 10 million health workers by 2030.

While AI is often seen as the key to bridging this gap, it has not yet been able to fully meet the needs of clinicians and patients. That's why, today, we are announcing our AI co-clinician research initiative, to explore how AI could better amplify doctors' expertise and deliver higher quality care to patients.”  
Si noti, per inciso, che la “figura” (ovvero l’algoritmo) dell’AI co-clinician fa il paio con algoritmi di AI-co-scientist, che vengono implementati come generatori di ipotesi (e di scoperte scientifiche (sic!)) in ausilio ai ricercatori, un argomento che meriterebbe una riflessione di per sé, anche se di natura differente...  
David Adam (2026) The AI Co-Scientist is here. *Nature Medicine* 32: 772-775.

Gottweiss et al (2026) Accelerating scientific discovery with Co-Scientist. *Nature*  
<https://doi.org/10.1038/s41586-026-10644-y>

C Lu et al (2024) The ai scientist: Towards fully automated open-ended scientific discovery  
arXiv preprint arXiv:2408.06292, 2024•arxiv.org

Sull’AI Co-scientist, dal sito di Google DeepMind

<https://deepmind.google/blog/co-scientist-a-multi-agent-ai-partner-to-accelerate-research/>

<sup>20</sup> “We hypothesize that the next evolution of healthcare delivery will entail “triadic care” where AI agents can help patients in their care journeys under the clinical authority of their physician. Medicine has always been a team sport, and AI agents can bring more teammates onto the field: extending clinicians' reach while ensuring they retain judgment and control.

This serves as the foundation of our AI co-clinician research initiative: AI designed to function as a collaborative member of the care team that interacts with patients under expert clinical supervision. We designed and evaluated AI co-clinician in both clinician and patient-facing settings. Addressing both perspectives is key for AI to enhance the quality, cost, availability and experience of care delivery.”  
Tratto dal sito Google DeepMind <https://deepmind.google/blog/ai-co-clinician/>

<sup>21</sup> Vivek Natarajan (Google DeepMind), comunicazione orale (April 13-17 Spring School Synthetic Biology, Scuola Normale Superiore(Pisa).

engagement), creando attaccamento, legami e dipendenze disfunzionali. La crescente diffusione di questi sistemi pone già oggi, di per sé, rilevanti problemi etici e regolatori che sono oggetto di attiva discussione<sup>22</sup>.

La incorporazione di AI companion chatbots (compagno IA) in sistemi conversazionali per la salute, sia salute fisica che mentale, porrà ulteriori quesiti sia etici che regolatori e legislativi. L'integrazione della IA direttamente interattiva con il consumatore/utente/paziente nei domini della salute fisica e mentale si sta sviluppando con enorme velocità e gli aspetti regolatori e l'individuazione dei profili di responsabilità legale non hanno tenuto il passo.

Uno studio del Parlamento Europeo, pubblicato il 19 Maggio 2026, richiama l'attenzione sulle sfide che questi sistemi interattivi pongono<sup>23</sup> e la necessità urgente di una loro regolamentazione, per le conseguenze sui soggetti più vulnerabili ed osserva che pochissimi Paesi hanno approvato leggi per la loro regolamentazione. Neppure la EU, si osserva nel rapporto, ha leggi specifiche per gli AI companions, anche se il quadro legislativo esistente (ivi incluso il AI Act ed il Digital Services Act) potrebbe fornire una cornice di riferimento.

Non possiamo che richiamarci alla urgente necessità di “custodia della persona umana nel tempo della intelligenza artificiale” alla quale ci ha richiamato il Pontefice Leone XIV con la *Magnifica humanitas*.

---

<sup>22</sup> Editorial (2024) *Nat. Mach. Intell.* **6**, 495 (2024). Empathic AI can't get under the skin.  
Editorial (2025): Emotional risks of AI companions demand attention. *Nature Machine Intelligence* **7**, 981–982 .  
Kirk, H. et al (2024) The benefits, risks and bounds of personalizing the alignments of large language models to individuals *Nat. Mach. Intell.* **6**, 383–392 (2024)  
De Freitas & Cohen (2025) Unregulated emotional risks of AI wellness apps *Nat. Mach. Intell.* **7**, 813–815.

---

<sup>23</sup> Mar Negreiro and Oyku Dilara Anac (2026) The spread of AI companions and the challenges they generate. PE 789.299 [https://www.europarl.europa.eu/thinktank/it/document/EPRS\\_BRI\(2026\)789299](https://www.europarl.europa.eu/thinktank/it/document/EPRS_BRI(2026)789299)

## **APPENDICE AL DOCUMENTO** “*Sulla responsabilità per danni causati da dispositivi mossi dall’intelligenza artificiale: l’ambito sanitario*”<sup>24</sup>

### **Mappatura dell’uso diagnostico dell’IA nella pratica clinica**

#### **Indicazioni alla lettura**

In questo report di analisi, “uso diagnostico” include quattro famiglie di impiego: screening autonomo, rilevazione di lesioni/anomalie durante l’esame, triage/prioritizzazione dei casi sospetti, e seconda lettura o supporto al professionista.

*Sono esclusi gli usi puramente amministrativi, gli strumenti di refertazione generativa non regolati come dispositivi diagnostici e gli algoritmi esclusivamente predittivi/prognostici non direttamente inseriti in un atto diagnostico.*

Questa mappatura è selettiva, clinicamente orientata ai casi con la migliore documentazione regolatoria e bibliografica.<sup>25</sup>

Il livello di evidenza è riportato in termini pragmatici:

- alto quando esistono studi prospettici/pivotal e/o più validazioni esterne con sintesi autorevoli e, idealmente, trial randomizzati o raccomandazioni di società scientifiche;
- intermedio quando ci sono approvazioni regolatorie e validazioni esterne o reader studies robusti, ma pochi trial pragmatici;
- limitato quando un prodotto è autorizzato, ma i dati indipendenti post-market o gli studi comparativi restano pochi.

Va ricordato che le revisioni recenti sui trial randomizzati (RCT) di IA in pratica clinica mostrano una crescita, ma una base ancora relativamente piccola rispetto all’ampiezza dell’offerta commerciale.

Un’ultima nota metodologica: la documentazione regolatoria pubblica spesso non esplicita con precisione l’architettura del modello. In molti prodotti è quindi corretto parlare di deep learning proprietario<sup>26</sup>.

---

<sup>24</sup> Per la paternità, si rinvia alla nota n. 2.

<sup>25</sup> Wenderott et al., «Effects of Artificial Intelligence Implementation on Efficiency in Medical Imaging—a Systematic Literature Review and Meta-Analysis».

<sup>26</sup> Han et al., «Randomised Controlled Trials Evaluating Artificial Intelligence in Clinical Practice».

## Casi d'uso e maturità clinica

### Oftalmologia

Il caso d'uso più maturo e più vicino all'autonomia diagnostica reale è lo screening della retinopatia diabetica, riferibile su fotografie del fundus, in persone adulte con diabete. Gli input sono in genere due immagini per occhio, una centrata sulla papilla e l'altra sulla macula; i modelli commerciali impiegano il deep learning su immagini retiniche, talora in ensemble con diversi algoritmi.

*I prodotti emblematici sono IDx-DR / LumineticsCore (FDA DEN180001; successive 510(k) per aggiornamenti, ad esempio K213037), EyeArt (FDA K200667, aggiornato con K223357) e AEYE-DS (FDA K221183). Nei documenti FDA, IDx-DR è definito come dispositivo di screening diagnostico autonomo per identificare “more than mild diabetic retinopathy”; EyeArt è descritto come ensemble di algoritmi di machine learning/deep learning; AEYE-DS come software diagnostico retinico per mtmDR.*

[https://www.accessdata.fda.gov/cdrh\\_docs/reviews/DEN180001.pdf](https://www.accessdata.fda.gov/cdrh_docs/reviews/DEN180001.pdf)

Studi recenti suggeriscono che l'adozione nelle cure primarie può migliorare l'aderenza allo screening e l'accesso all'oculista, soprattutto nelle popolazioni svantaggiate.<sup>27</sup>

I benefici dimostrati sono quindi: portare lo screening nel luogo di cura, ridurre la dipendenza dagli specialisti per la prima valutazione, aumentare il completamento degli esami e, in molti contesti, mantenere un'accuratezza comparabile a quella della valutazione effettuata dagli esperti.

I principali limiti restano rilevanti. Questi sistemi possono essere usati solo su popolazioni specifiche e con dispositivi già validati. Inoltre, non sempre le immagini ottenute sono di qualità sufficiente per essere analizzate correttamente. L'IA non può sostituire una visita oculistica completa, perché valuta solo alcuni aspetti dell'occhio.

In generale, le sue prestazioni possono variare a seconda del contesto in cui viene applicata, ad esempio con pazienti di diverse etnie, in strutture diverse o con dispositivi diversi. Infine, la diagnosi è limitata ad alcune patologie retiniche e non copre tutte le malattie oculari.

### Gastroenterologia ed endoscopia

Nell'endoscopia digestiva, il caso più consolidato è la rilevazione in tempo reale di polipi e adenomi durante la colonscopia. Qui l'IA è soprattutto Computer-Aided Detection (CAD) - diagnosi assistita da computer: non formula di per sé una diagnosi istologica, ma segnala regioni sospette nel video endoscopico e aumenta l'attenzione dell'operatore.

*I prodotti meglio documentati sono GI Genius (FDA DEN200055), EndoScreener (FDA K211326), SKOUT (FDA K213686) e CADDIE (FDA K240044). Nel De Novo di GI Genius l'FDA chiarisce che il dispositivo è un “computer-assisted reading tool” per lesioni mucose coliche, non è uno stand-alone diagnostic device, non caratterizza la lesione in luogo della biopsia e non sostituisce il giudizio clinico; l'architettura include esplicitamente una convolutional neural network.*

[https://www.accessdata.fda.gov/cdrh\\_docs/reviews/DEN200055.pdf](https://www.accessdata.fda.gov/cdrh_docs/reviews/DEN200055.pdf)

Il razionale clinico è forte: la colonscopia è il gold standard per rilevare e rimuovere lesioni precancerose, ma l'adenoma detection rate varia tra gli endoscopisti ed è inversamente associato al rischio di cancro coloretale intervallare.

La posizione dell'European Society of Gastrointestinal Endoscopy, cioè la Società Europea di Endoscopia Gastrointestinale, del 2025, basata su revisioni sistematiche della letteratura, rileva oltre 40 studi clinici randomizzati e controllati sull'uso della rilevazione assistita da computer. Nella sintesi

---

<sup>27</sup> Wang et al., «Systematic Review and Meta-Analysis of Regulator-Approved Deep Learning Systems for Fundus Diabetic Retinopathy Detections».

della Società Europea di Endoscopia Gastrointestinale, l'uso della CAD aumenta il tasso di rilevazione degli adenomi, con un rischio relativo di 1,22 sulla base di 40 studi e 30.674 partecipanti.<sup>28</sup>

Tuttavia, la raccomandazione è debole, perché l'effetto sugli esiti clinici più importanti, come l'incidenza e la mortalità per cancro coloretale, resta incerto. Inoltre, il beneficio è controbilanciato da una maggiore rilevazione di polipi non neoplastici, dal potenziale rischio di sovradiagnosi e da un aumento dei controlli di sorveglianza.

*Il percorso autorizzativo FDA De Novo relativo a GI Genius conclude che il dispositivo migliora il tasso di rilevazione degli adenomi e il numero di adenomi rilevati per colonscopia rispetto alla colonscopia standard, ma riconosce l'incertezza sulla generalizzabilità dei risultati e sul legame con gli esiti oncologici finali.*

In termini pratici, questo è quindi un caso d'uso molto diffuso, ma, concettualmente, va classificato come supporto diagnostico intraprocedurale, non come "diagnosi automatica". I vantaggi più solidi sono l'aumento dei tassi di rilevazione di adenomi/polipi e la standardizzazione della vigilanza visiva. Le criticità principali sono l'impatto potenziale sui falsi positivi, la minore chiarezza sui benefici clinici a lungo termine, la dipendenza dalla qualità della preparazione, dell'endoscopio e della tecnica dell'operatore, e il fatto che la rilevazione non equivale alla caratterizzazione istologica.

[https://www.accessdata.fda.gov/cdrh\\_docs/reviews/DEN200055.pdf](https://www.accessdata.fda.gov/cdrh_docs/reviews/DEN200055.pdf)

## Radiologia senologica

Nel breast imaging, l'intelligenza artificiale è oggi utilizzata soprattutto come supporto alla lettura concomitante delle immagini, oppure come alternativa parziale alla doppia lettura umana nella mammografia e nella tomosintesi digitale mammaria. Gli input utilizzati dai sistemi sono immagini di mammografia digitale a pieno campo (Full-Field Digital Mammography, FFDM) e/o immagini di tomosintesi digitale mammaria (Digital Breast Tomosynthesis, DBT); gli algoritmi sono in genere sistemi di deep learning per la rilevazione e la valutazione delle lesioni.

*I prodotti con chiara documentazione regolatoria includono ProFound AI di iCAD (FDA K191994, K203822, aggiornamenti come K240417), MammoScreen di Therapixel (FDA K192854, K211541, K240301, K243679) e Transpara di ScreenPoint Medical, che il fabbricante dichiara CE-marked e FDA-cleared per 2D/3D mammography, con riferimenti K241831 e K232096. La documentazione FDA su ProFound AI esplicita l'uso di deep learning e pattern recognition; la documentazione di Transpara e iCAD dichiara marcatura CE e impiego clinico concorrente.*

[https://www.accessdata.fda.gov/cdrh\\_docs/pdf20/K203822.pdf](https://www.accessdata.fda.gov/cdrh_docs/pdf20/K203822.pdf)

L'evidenza qui è fra le più interessanti tra le applicazioni non autonome. Nel 2023, il trial MASAI ha mostrato una riduzione del carico di lavoro del 44,3% grazie allo screening supportato dall'AI, con un tasso di rilevazione del cancro simile a quello della doppia lettura standard. I risultati del 2025 invece, riportano anche un aumento del cancer detection rate, 6,4 per 1000 nel gruppo AI contro 5,0 per 1000 nel controllo, con rapporto 1,29, e suggeriscono che l'IA contribuisce alla rilevazione precoce di tumori clinicamente rilevanti e riduce il carico di lettura senza aumentare i richiami in modo penalizzante.<sup>29</sup>

Altri studi su Lancet Digital Health hanno indicato che i modelli con un lettore umano più IA possono aumentare la rilevazione rispetto alla doppia lettura interamente umana. La lettura complessiva, però, resta prudente: è un'area molto promettente e già usata, ma i risultati dipendono fortemente dal protocollo di screening, dalla soglia operativa, dall'integrazione dello storico e dalla composizione della popolazione.<sup>30</sup>

I benefici dimostrati riguardano soprattutto l'efficienza, la priorità dei casi ad alta probabilità e, in alcuni contesti, l'aumento della rilevazione. I limiti riguardano la trasferibilità tra vendor e programmi

---

<sup>28</sup> Bretthauer et al., «Use of Computer-Assisted Detection (CADe) Colonoscopy in Colorectal Cancer Screening and Surveillance».

<sup>29</sup> Hernström et al., «Screening Performance and Characteristics of Breast Cancer Detected in the Mammography Screening with Artificial Intelligence Trial (MASAI)».

<sup>30</sup> Lång et al., «Artificial Intelligence-Supported Screen Reading versus Standard Double Reading in the Mammography Screening with Artificial Intelligence Trial (MASAI)».

nazionali, l'effetto su recall e intervallari, la calibrazione nei seni densi e l'assenza, per ora, di una prova definitiva di beneficio sugli endpoint finali come la mortalità. In più, gran parte della documentazione pubblica riguarda le performance e i workflow, mentre la sorveglianza post-market comparativa è meno rappresentata.

## Radiologia toracica per screening TB e reperti selezionati

Nel torace la situazione va distinta. L'uso generalista dell'IA per "abnormal chest X-ray" è ampio, ma più eterogeneo; il caso più robustamente normato e raccomandato è la lettura CAD delle radiografie del torace per screening della tubercolosi negli adulti/adolescenti  $\geq 15$  anni.

*I prodotti FDA comprendono, per esempi rilevanti, Lunit INSIGHT CXR Triage (FDA K211733) per pleural effusion/pneumothorax, qXR-LN (FDA K231805) per noduli polmonari incidentali, qXR-CTR (FDA K231149) per cardiothoracic ratio, e più recentemente qXR-Detect (FDA K251934) come CAD concorrente per ROI sospette in sei categorie. La documentazione FDA parla esplicitamente di deep learning nel caso di qXR-LN e di algoritmi AI-based per Lunit; i prodotti sono però etichettati come ausili di rilevazione/triage, non come sostituti del radiologo.*

Sul versante TB, l'OMS raccomanda dal 2021 l'uso del CAD per la lettura automatizzata della radiografia del torace nello screening e nel triage della TB, e nel 2025 ha dichiarato che sei software hanno raggiunto gli standard dell'OMS per l'uso nello screening di persone di 15 anni o più. Una review RSNA del 2025 sottolinea che gli strumenti clinicamente rilevanti includono CAD per CXR e triage tools per varie urgenze toraciche, e richiama numerosi studi recenti, inclusi RCT e studi multicentrici, su noduli, pneumotorace e tubercolosi.<sup>31</sup> Questo rende il CAD-TB uno dei pochi contesti in cui esiste una vera policy recommendation globale per l'uso diagnostico e di screening dell'IA.<sup>32</sup>

I benefici documentati sono particolarmente evidenti nei contesti a bassa disponibilità di radiologi: aumento della fattibilità dello screening, standardizzazione della selezione dei test microbiologici successivi e accelerazione del flusso. I limiti sono altrettanto importanti: la raccomandazione dell'OMS è contestuale allo screening della TB, non va estesa indebitamente alla diagnosi generale di tutte le patologie toraciche; inoltre, la taratura della soglia conta molto; l'uso in età pediatrica non è ancora raccomandato; e gli esiti possono variare in presenza di storia di TB pregressa, comorbidità, hardware e qualità dell'immagine.

## Casi d'uso diffusi come supporto o triage

### Neuroradiologia d'urgenza e stroke

Nell'ictus e nelle urgenze neuroradiologiche, l'intelligenza artificiale è oggi ampiamente diffusa come software per il triage, la priorità e la notifica. Questi sistemi aiutano a identificare rapidamente casi urgenti, come l'occlusione di un grosso vaso cerebrale (Large Vessel Occlusion, LVO), l'emorragia intracranica (Intracranial Hemorrhage, ICH) e altre condizioni acute rilevabili alla tomografia computerizzata (TC) e all'angiografia con tomografia computerizzata (CTA).

*I prodotti esemplari sono Viz LVO / ContaCT (FDA DEN170073, poi K223042) e varie versioni di Aidoc BriefCase per ICH, PE, LVO (ad esempio K180647, K221240, K192383, K213721). La documentazione FDA è molto chiara su un punto: questi sistemi non sono dispositivi diagnostici autonomi, ma strumenti di parallel workflow o di computer-aided triage and notification a supporto della prioritizzazione delle immagini. Per i prodotti specifici, l'architettura dettagliata non è sempre divulgata pubblicamente; la letteratura di settore parla tipicamente di deep learning/CNN per imaging volumetrico e segmentazione/riconoscimento vascolare, ma per molti 510(k) la descrizione resta ad alto livello.*

L'evidenza clinica qui è soprattutto di tipo implementativo. I software per la rilevazione di occlusioni vascolari sono ormai ampiamente disponibili per il triage dell'ictus; la letteratura recente si concentra soprattutto sulla riduzione dei tempi di percorso e sul coordinamento interospedaliero, più che su una superiorità diagnostica autonoma.<sup>33</sup>

---

<sup>31</sup> «AI Applications for Thoracic Imaging: Considerations for Best Practice | Radiology».

<sup>32</sup> «WHO Consolidated Guidelines on Tuberculosis».

<sup>33</sup> «Automated Vessel Occlusion Software in Acute Ischemic Stroke: Pearls and Pitfalls | Stroke».

La criticità principale è concettuale: l'IA è spesso percepita come “diagnostic tool”, ma i documenti regolatori la qualificano più precisamente come strumento volto a non perdere o a non ritardare i casi sospetti. I rischi includono falsi negativi nelle lesioni piccole o nelle anatomie complesse, falsi positivi che disturbano il workflow, differenze di prestazioni tra scanner e protocolli e la dipendenza dall'integrazione PACS-messaging.

[https://www.accessdata.fda.gov/cdrh\\_docs/pdf22/K223042.pdf](https://www.accessdata.fda.gov/cdrh_docs/pdf22/K223042.pdf)

## Anatomia patologica digitale

In anatomia patologica, l'ambito d'uso dell'intelligenza artificiale più chiaramente regolato riguarda la seconda lettura di biopsie prostatiche digitalizzate.

*Il riferimento principale è Paige Prostate (FDA DEN200080); come esempio europeo si può affiancare Ibox Prostate, che il fabbricante dichiara CE marked e, per alcune versioni, IVDR. Paige è indicato per assistere il patologo nel rilevare foci sospetti di carcinoma nelle whole-slide images di biopsie prostatiche, con output binario “suspicious / not suspicious” e le coordinate del punto più sospetto; il documento FDA definisce esplicitamente il sistema come un deterministic deep learning system e ne chiarisce il ruolo adjunctive, non sostitutivo della diagnosi del patologo. [23]*

I benefici documentati sono: riduzione del rischio di piccoli foci mancati, standardizzazione del second-read e possibile ottimizzazione del tempo del patologo. I limiti sono: infrastruttura digitale necessaria, dipendenza da scanner/viewer specifici, rischio di errori dovuti a una bassa qualità dello scan, differenza tra le performance per-biopsy e l'utilità per-patient, e trasparenza pubblica limitata per molti prodotti CE.

[https://www.accessdata.fda.gov/cdrh\\_docs/reviews/DEN200080.pdf](https://www.accessdata.fda.gov/cdrh_docs/reviews/DEN200080.pdf)

## Cardiologia basata su segnali

In cardiologia, l'uso, dal punto di vista diagnostico, è oggi meno uniforme rispetto all'imaging, ma esistono prodotti regolati per l'analisi dei segnali fisiologici.

*Un esempio ben documentato è Eko Analysis Software (FDA K192004), che analizza fonocardiogramma + ECG simultaneo per rilevare murmurs sospetti e fibrillazione atriale/ritmo sinusale. La documentazione FDA lo presenta come supporto alla valutazione clinica di suoni cardiaci ed ECG; l'architettura dettagliata del prodotto non è completamente esplicitata nel 510(k), ma la letteratura sul sistema Eko parla di algoritmi di deep learning.*

<https://www.accessdata.fda.gov/CDRH510K/K192004.pdf>

L'evidenza disponibile è promettente, ma nel complesso meno solida rispetto a quella presente in ambiti come l'oftalmologia e l'endoscopia. Uno studio del 2023 su algoritmi autorizzati dalla FDA per il rilevamento dei murmurs cardiaci associati a cardiopatie strutturali ha riportato una sensibilità dell'85,6% e una specificità dell'84,4%. L'American Heart Association, in un documento scientifico del 2024 sull'uso dell'intelligenza artificiale in cardiologia, considera questi strumenti parte di una trasformazione reale della pratica clinica, ma sottolinea anche la necessità di un'implementazione rigorosa, di una validazione esterna continua e di un monitoraggio clinico nel tempo.<sup>34</sup>

Il vantaggio clinico potenziale è notevole in medicina generale, nel triage e in ambienti con accesso limitato all'ecocardiografia. I limiti sono il rumore del segnale, la variabilità tra i dispositivi, il fatto che il tool resta un filtro e non sostituisce gli esami di conferma, e una base di studi pragmatici meno robusta.

## Dermatologia

*In dermatologia, il prodotto regolato più chiaro è DermaSensor (FDA DEN230008), indicato per aiutare medici non dermatologi a decidere il referral di lesioni sospette per melanoma, BCC o SCC in pazienti di età ≥40 anni. È importante notare che non è uno strumento di screening massivo e non si basa principalmente sulla fotografia clinica standard, ma su analisi ottiche/spettroscopiche della*

---

<sup>34</sup> Armoundas et al., «Use of Artificial Intelligence in Improving Outcomes in Heart Disease».

*lesione; l'esatta architettura algoritmica pubblica non è dettagliata nel De Novo e va trattata come proprietaria/non pienamente specificata.*

Sul piano delle prove, la dermatologia AI dispone di meta-analisi recenti che mostrano come i clinici possano beneficiare dell'assistenza AI nella diagnosi di tumori cutanei e, in alcune sintesi, i meno esperti sembrano trarre più beneficio dal supporto algoritmico.

15 ottobre 2026

Commissione lineea di Bioetica