

Intelligenza Artificiale: come funziona, cosa cambia

Francesco Masera

Il Medico Veterinario nella società delle mangrovie — Varese, 18 aprile 2026

1. Introduzione: il paradosso dell'adozione

Iniziamo con un dato. Uno studio pubblicato nel 2026 su *Frontiers in Veterinary Science* (Li & Lai, 2026) ha analizzato l'uso dell'intelligenza artificiale tra i medici veterinari cinesi. Il risultato è sorprendente per due motivi opposti: il 71% degli intervistati dichiara di utilizzare già strumenti di AI nel proprio lavoro quotidiano. Ma il 44,6% di questi stessi utilizzatori ammette di avere scarsa familiarità con la tecnologia che sta usando.

Sette veterinari su dieci usano l'intelligenza artificiale. Quasi la metà di loro non sa bene come funziona. I ricercatori lo hanno chiamato il “paradosso dell'adozione”, e questa espressione cattura perfettamente una situazione che non riguarda solo la Cina. Una survey 2026 sugli specialisti ACVIM e ECVIM-CA mostra che il 42% degli internisti nordamericani e il 33% degli europei usa l'AI nella pratica clinica, ma la maggioranza di questi utenti riferisce conoscenze solo moderate o basse (Bertin et al., 2026). Il fenomeno è globale: viviamo immersi in tecnologie che utilizziamo quotidianamente senza comprenderne i meccanismi.

Nella maggior parte dei casi, questa ignoranza è innocua. Non è necessario capire come funziona un motore a combustione interna per guidare un'automobile in sicurezza. Ma la medicina — umana o veterinaria — è diversa. Quando uno strumento influenza una diagnosi, una prognosi o una scelta terapeutica, la mancata comprensione del suo funzionamento non è più innocua. Diventa un rischio. Chi non capisce come funziona uno strumento non sa nemmeno quando quello strumento può sbagliare. E non saperlo è il primo passo verso errori che si sarebbero potuti evitare.

Questo documento nasce per offrire ai medici veterinari una mappa del territorio. Non un manuale tecnico: non serve sapere programmare per capire l'AI, così come non serve essere ingegneri per capire un'ecografia. Serve però conoscere il principio di funzionamento, i punti di forza e soprattutto i limiti. Per chi desiderasse un percorso più approfondito, dai fondamenti matematici alle competenze operative, rimando al mio volume dedicato (Masera, 2026). In queste pagine, invece, mi concentro sugli aspetti più rilevanti per il professionista veterinario che vuole capire cosa sta usando e quando può fidarsi.

Nelle pagine che seguono vedremo prima come le macchine imparano dai dati — il meccanismo alla base di quasi tutte le applicazioni AI in ambito medico. Poi affronteremo l'intelligenza artificiale generativa, quella dei ChatGPT e degli assistenti testuali, che funziona in modo diverso e porta con sé rischi diversi. Infine apriremo alcune finestre sulle applicazioni già operative in medicina veterinaria — non come catalogo di prodotti, ma come illustrazione concreta dei principi descritti — e chiuderemo con i rischi che ogni professionista dovrebbe conoscere.

2. Come le macchine imparano dai dati

2.1 L'analogia del bambino

Quando diciamo “intelligenza artificiale”, nella grande maggioranza dei casi parliamo di una cosa sola: una macchina che impara dai dati. Questa espressione — machine learning, apprendimento automatico — può sembrare astratta, ma il principio che la governa è sorprendentemente intuitivo.

Pensate a come un bambino impara a riconoscere un gatto. Nessuno gli fornisce una definizione formale: “mammifero domestico del genere Felis, con quattro zampe, orecchie a punta, coda lunga e pelo variabile”. Il bambino vede gatti. Ne vede decine, centinaia, in contesti diversi: gatti grandi, gatti piccoli, gatti neri, gatti a pelo lungo, gatti fotografati dall'alto o di profilo. E a un certo punto, senza saper formulare una regola esplicita, riconosce un gatto anche se non lo ha mai visto prima. Il suo cervello ha estratto dall'esperienza un pattern — un insieme di regolarità ricorrenti — che gli permette di generalizzare.

Il machine learning funziona esattamente allo stesso modo. La differenza è di scala: invece di centinaia di esempi, se ne usano centinaia di migliaia o milioni. E invece di immagini viste con gli occhi, si usano numeri: valori di un esame del sangue, pixel di una radiografia, temperature rilevate da un sensore, parole contenute in una cartella clinica. A partire da questi dati, il software cerca regolarità statistiche — pattern — che gli permettano di classificare, predire o riconoscere qualcosa.

2.2 Apprendimento supervisionato: imparare dagli esempi

Il tipo di apprendimento più comune nelle applicazioni mediche è quello chiamato supervisionato. Il concetto è semplice: si fornisce alla macchina un grande numero di esempi già classificati, e le si chiede di trovare le regole che distinguono una categoria dall'altra.

Immaginate di avere 100.000 radiografie toraciche di cani, ciascuna accompagnata dal referto di un radiologo che indica se c'è un versamento pleurico oppure no. L'algoritmo esamina tutte queste immagini — i “sì” e i “no” — e impara a riconoscere le caratteristiche visive che distinguono le due categorie. Non gli viene spiegato cosa cercare: è lui a trovarlo, attraverso un processo iterativo di prova, errore e correzione.

Questo processo si chiama addestramento, o training. Durante il training, il modello esamina ogni esempio, formula una previsione (“questa radiografia mostra un versamento” oppure “no”), confronta la previsione con la risposta corretta, e aggiusta i propri parametri interni per sbagliare un po' meno la volta successiva. Ripetuto milioni di volte, questo ciclo produce un modello che ha imparato a riconoscere pattern che l'occhio umano potrebbe non cogliere — o che coglierebbe, ma con minor velocità e costanza.

Un aspetto cruciale: la qualità del risultato dipende interamente dalla qualità e dalla quantità dei dati di addestramento. Se le 100.000 radiografie provengono tutte da cani di taglia grande, il modello potrebbe non funzionare bene sui gatti o sui cani toy. Se i referti contengono errori sistematici, il modello imparerà quegli errori. In informatica si usa un'espressione efficace: *garbage in, garbage out* — se entrano dati spazzatura, escono risultati spazzatura. Questo vale per qualsiasi sistema di AI, e vale la pena tenerlo sempre a mente.

2.3 Le reti neurali: strati su strati

Nella maggior parte delle applicazioni moderne, lo strumento matematico che effettua l'apprendimento è una rete neurale artificiale. Il nome è evocativo — e in parte fuorviante. Le reti neurali sono state ispirate, molto alla lontana, dall'architettura del cervello biologico, ma in realtà sono oggetti matematici: sequenze di operazioni numeriche organizzate in strati successivi.

Ogni strato riceve dei numeri in ingresso, li trasforma applicando pesi e funzioni matematiche, e passa il risultato allo strato successivo. Il primo strato riceve i dati grezzi — i pixel dell'immagine, i valori del sangue. Gli strati intermedi estraggono caratteristiche sempre più astratte: nel caso di un'immagine, il primo strato potrebbe riconoscere bordi e contrasti, il secondo forme geometriche, il terzo strutture anatomiche, e così via. L'ultimo strato produce l'output: “versamento pleurico sì/no”, oppure un valore numerico come “probabilità di sviluppare la malattia: 87%”.

I pesi — i numeri che governano ogni trasformazione — sono i parametri che il modello impara durante il training. Un modello moderno può avere milioni o miliardi di parametri. Quando diciamo che un modello “sa fare” qualcosa, intendiamo che quei milioni di numeri sono stati calibrati attraverso il training in modo da produrre risposte accurate. Non c'è nessuna comprensione in senso umano: c'è una configurazione numerica che produce risultati utili.

La profondità della rete — il numero di strati — è ciò che dà il nome al deep learning, l'apprendimento profondo. Le reti profonde riescono a catturare pattern più complessi rispetto a quelle con pochi strati, ma richiedono più dati di addestramento e più potenza di calcolo.

2.4 Validazione: come si misura se funziona

Un punto fondamentale che ogni professionista dovrebbe conoscere: il fatto che un modello funzioni bene sui dati di addestramento non garantisce che funzionerà bene nella pratica. Se un bambino ha visto solo gatti siamesi, potrebbe non riconoscere un persiano. Allo stesso modo, un modello addestrato su radiografie di una sola marca di apparecchiatura potrebbe produrre risultati diversi con un'altra.

Per questo motivo, lo sviluppo di un sistema AI prevede sempre una fase di validazione: si testa il modello su dati che non ha mai visto durante il training. Se funziona bene anche su questi dati nuovi, si può essere ragionevolmente fiduciosi che abbia davvero imparato il pattern sottostante, e non si sia limitato a memorizzare gli esempi. In gergo tecnico, si dice che il modello generalizza bene.

Il rischio opposto si chiama *overfitting*: il modello ha imparato i dati di addestramento troppo bene, inclusi i loro rumori e le loro particolarità casuali, e quando incontra dati nuovi sbaglia. È come uno studente che ha memorizzato le risposte di un compito in classe specifico, ma non ha capito la materia e non sa risolvere un esercizio diverso.

Nella pratica clinica, questo ha un'implicazione importante: quando un produttore di software AI dichiara un'accuratezza del 95%, la domanda giusta da porsi non è “è tanto o poco?”, ma “su quali dati è stata misurata?”. Se quei dati erano molto simili a quelli di addestramento, la cifra potrebbe essere gonfiata. Se provenivano da cliniche diverse, con apparecchiature diverse, su popolazioni di pazienti diverse, allora la cifra è molto più attendibile.

2.5 Quando il modello sbaglia: i tipi di errore

Qualsiasi modello di AI, per quanto accurato, commetterà errori. Per un professionista è fondamentale capire non solo che il modello sbaglia, ma come sbaglia, perché i diversi tipi di errore hanno conseguenze cliniche molto diverse.

Nella classificazione medica, gli errori si dividono in due categorie fondamentali. Un falso negativo è un caso in cui il modello dice “normale” quando in realtà c'è una patologia: il paziente malato viene classificato come sano. Un falso positivo è il contrario: il modello segnala un'anomalia che non c'è, generando un allarme ingiustificato. In medicina, i falsi negativi sono generalmente più pericolosi — un tumore non visto è peggio di un allarme in più — ma i falsi positivi non sono innocui: generano ansia, esami aggiuntivi inutili, costi e, nel caso degli animali da reddito, possono portare a trattamenti o abbattimenti ingiustificati.

Due metriche catturano questa distinzione. La sensibilità misura la capacità del modello di trovare i veri positivi: su tutti i pazienti malati, quanti ne identifica correttamente? La specificità misura la capacità di escludere i sani: su tutti i pazienti sani, quanti vengono correttamente classificati come tali? Un modello può essere molto sensibile ma poco specifico (trova quasi tutti i malati, ma genera molti falsi allarmi) o molto specifico ma poco sensibile (raramente dà falsi allarmi, ma perde qualche caso).

Nella pratica, non esiste un modello perfetto. Ogni sistema rappresenta un compromesso tra sensibilità e specificità, e la scelta del punto di equilibrio dipende dal contesto clinico. In un programma di screening, dove il costo di perdere un caso è alto, si preferisce un'alta sensibilità anche a prezzo di più falsi positivi. In un contesto di conferma diagnostica, si

preferisce un'alta specificità. Quando valutate un software AI, chiedete sempre: quali sono la sensibilità e la specificità? Su quale popolazione sono state misurate? Sono adeguate al mio contesto d'uso?

Una precisazione utile: in medicina veterinaria, dove le prevalenze possono essere basse e i dataset sbilanciati, l'accuratezza complessiva da sola dice poco. Un modello che classifica come sano il 98% dei pazienti sarà accurato al 98% in una popolazione dove solo il 2% è malato — ma non avrà individuato nessun caso. Metriche più informative includono, oltre a sensibilità e specificità, l'AUC (area sotto la curva ROC, che misura la capacità discriminante complessiva del modello), il valore predittivo positivo (PPV: se il modello dice “malato”, con quale probabilità ha ragione?) e il valore predittivo negativo (NPV: se dice “sano”, quanto è affidabile?). Questi valori cambiano con la prevalenza della malattia nella popolazione esaminata, ed è per questo che i risultati ottenuti in un contesto non sono automaticamente trasferibili a un altro.

2.6 Apprendimento non supervisionato e rinforzo: altre strade

L'apprendimento supervisionato non è l'unico modo in cui una macchina può imparare. Vale la pena conoscere almeno altri due approcci, perché sono alla base di applicazioni che incontrerete sempre più spesso.

Il primo si chiama apprendimento non supervisionato. In questo caso, alla macchina non vengono forniti esempi già classificati: le si danno i dati e le si chiede di trovare da sola raggruppamenti, anomalie o strutture. Immaginate di avere migliaia di cartelle cliniche senza etichette diagnostiche: un algoritmo non supervisionato potrebbe scoprire che certi gruppi di pazienti condividono pattern simili nei parametri ematici, suggerendo sottotipi di una malattia che non erano stati identificati. Non classifica — esplora. È uno strumento di scoperta, non di conferma.

Il secondo è l'apprendimento per rinforzo, che funziona in modo ancora diverso: la macchina impara per tentativi, ricevendo una ricompensa quando fa bene e una penalità quando sbaglia. È il principio con cui i programmi di AI hanno imparato a giocare a scacchi o a Go meglio di qualsiasi essere umano, e in medicina si sta esplorando per l'ottimizzazione dei protocolli terapeutici: data una serie di opzioni di trattamento e una misura di risultato, il sistema impara quale sequenza di decisioni massimizza le probabilità di successo.

Per la pratica quotidiana, l'apprendimento supervisionato resta di gran lunga il più utilizzato. Ma conoscere l'esistenza di questi altri approcci aiuta a comprendere un punto importante: non esiste “una” intelligenza artificiale. Esistono famiglie diverse di tecniche, ciascuna con punti di forza e limiti propri. La domanda giusta non è mai “l'AI funziona?”, ma “quale tipo di AI, addestrato su quali dati, per quale compito specifico?”.

2.7 Un riassunto in tre concetti

Prima di proseguire, fissiamo tre concetti che torneranno utili nel resto del documento:

Pattern. L'AI trova regolarità nei dati che l'occhio umano non può cogliere, non per mancanza di competenza, ma per limiti cognitivi naturali: nessun cervello umano può incrociare sette variabili su centomila pazienti contemporaneamente.

Dati. La qualità del risultato dipende dalla qualità e dalla rappresentatività dei dati di addestramento. Un modello non è mai migliore dei dati su cui è stato costruito.

Validazione. Nessun numero di accuratezza ha significato senza sapere come e su quali dati è stato misurato. La domanda critica è sempre: questi risultati valgono anche nel mio contesto clinico?

3. L'intelligenza artificiale generativa

3.1 Un meccanismo fondamentalmente diverso

Tutto ciò che abbiamo descritto finora — il riconoscimento di pattern nelle immagini, la predizione a partire da valori ematici, la classificazione di dati clinici — rientra in ciò che si chiama AI analitica o predittiva. Questo tipo di AI riceve dati e produce una classificazione o una stima: “questa radiografia mostra un versamento”, “il rischio di malattia renale è elevato”, “questa vacca potrebbe avere una mastite entro 48 ore”.

Ma c'è un altro tipo di intelligenza artificiale che avete quasi certamente incontrato, e che funziona in modo profondamente diverso. È quella che tutti conosciamo come ChatGPT, oppure come Gemini, Claude, Copilot — i cosiddetti modelli linguistici di grandi dimensioni, o Large Language Models. Si chiama AI generativa perché, invece di classificare dati esistenti, genera contenuti nuovi: testi, immagini, codice informatico, e in prospettiva audio e video.

La differenza non è solo nel risultato ma nel meccanismo. L'AI predittiva risponde a domande chiuse: sì o no, alto o basso, presente o assente. L'AI generativa produce risposte aperte: frasi, paragrafi, ragionamenti articolati. E il modo in cui lo fa è al tempo stesso elegante e profondamente controintuitivo.

3.2 Come funziona un modello linguistico

Un modello linguistico come ChatGPT è stato addestrato su una quantità quasi inconcepibile di testo: miliardi di pagine provenienti da libri, articoli scientifici, siti web, manuali tecnici, forum, enciclopedie, documenti legali, e molto altro. Da tutta questa lettura, il modello ha imparato i pattern statistici del linguaggio: quali parole tendono a seguire quali altre parole, in quali contesti, con quale probabilità.

Quando ponete una domanda a ChatGPT, il modello non “cerca” la risposta in un database. Non consulta un archivio. Fa una cosa diversa: genera la sequenza di parole statisticamente più plausibile dato il contesto della vostra domanda. Una parola dopo l'altra, ogni volta scegliendo la più probabile. È come un completamento automatico straordinariamente sofisticato: dato quello che avete scritto voi, il modello produce ciò che — secondo i pattern appresi — dovrebbe venire dopo.

Questo produce risultati spesso impressionanti. Se chiedete a un modello linguistico di spiegarvi la fisiopatologia della malattia renale cronica nel gatto, otterrete probabilmente una risposta chiara, ben strutturata, con terminologia corretta. Non perché il modello “sappia” di nefrologia felina in senso umano: ma perché, nel suo addestramento, ha elaborato così tanti testi di nefrologia veterinaria che ha imparato quali parole e concetti tendono ad accompagnarsi in quel contesto. La risposta è il riflesso statistico di un'enorme quantità di letteratura.

3.3 Il problema fondamentale: le allucinazioni

Ed è qui che emerge il limite strutturale più importante dell'AI generativa: poiché il modello genera parole probabili e non fatti verificati, può produrre risposte che sono perfettamente formulate ma completamente false. Lo fa con la stessa sicurezza con cui genera quelle corrette, perché non ha un meccanismo interno che distingue il vero dal falso. Non verifica ciò che scrive. Non sa di non sapere.

In gergo tecnico, questo fenomeno si chiama allucinazione, ed è un termine appropriato: il sistema vede cose che non ci sono e le racconta con assoluta convinzione. Può inventare citazioni bibliografiche che non esistono, attribuire risultati a studi mai condotti, descrivere farmaci con dosaggi plausibili ma sbagliati, fornire diagnosi differenziali ragionevoli ma basate su premesse errate.

Questo non è un bug: è una conseguenza diretta del meccanismo. Il modello non distingue tra “produrre la sequenza di parole più probabile su un argomento in cui esiste molta letteratura attendibile” e “produrre la sequenza di parole più probabile su un argomento in cui le informazioni disponibili sono scarse o contraddittorie”. In entrambi i casi, genera con la stessa fluidità e la stessa apparente sicurezza.

Per la pratica veterinaria, l'implicazione è immediata: se chiedete a un modello linguistico un parere su un caso clinico, potreste ricevere una risposta articolata, con terminologia corretta, struttura logica impeccabile e riferimenti a studi — e scoprire che gli studi non esistono, che il dosaggio è sbagliato, o che il ragionamento diagnostico, pur sembrando coerente, parte da un'associazione statistica che non corrisponde alla realtà clinica.

3.4 Comprendere non è generare

Un equivoco molto diffuso è che, poiché i modelli linguistici producono risposte che sembrano intelligenti, allora “capiscano” ciò di cui parlano. Vale la pena essere chiari: non è così, o quantomeno non nel senso in cui usiamo la parola “capire” per un essere umano.

Un modello linguistico non ha esperienza clinica. Non ha mai visitato un paziente. Non sa cosa significhi palpare un addome o interpretare il comportamento di un animale sofferente. Quello che ha è un'enorme conoscenza statistica di come le parole si combinano nei testi medici. Può generare un testo che assomiglia perfettamente a quello che scriverebbe un esperto, senza avere nulla dell'esperienza che rende quell'esperto tale.

Questo non significa che sia inutile. Tutt'altro. Un modello linguistico può essere un assistente prezioso per redigere documentazione clinica, sintetizzare letteratura, formulare ipotesi da verificare, comunicare con i proprietari in modo chiaro. Ma è esattamente questo: un assistente. Non un consulente. La differenza è che il consulente si assume la responsabilità di ciò che dice; l'assistente produce un output che deve essere verificato da chi ha la competenza per farlo.

3.5 AI generativa multimodale

Un'evoluzione recente ha ampliato le capacità dei modelli generativi oltre il testo. I modelli cosiddetti multimodali possono elaborare e generare non solo parole, ma anche immagini, audio e video. Potete, per esempio, mostrare a un modello multimodale la fotografia di una lesione cutanea e chiedergli una descrizione, oppure fornirgli un referto scritto e chiedergli di generare un riassunto per il proprietario dell'animale.

Anche in questo caso, il principio fondamentale non cambia: il modello lavora su pattern statistici, non su comprensione. Un modello multimodale che descrive una lesione cutanea lo fa perché ha visto milioni di immagini associate a descrizioni testuali, non perché “vede” la lesione come la vede un dermatologo. Questo amplia le possibilità d'uso, ma i rischi di allucinazione e di errore restano identici — anzi, possono essere amplificati quando il modello lavora su dati visivi ambigui o poco rappresentati nel suo addestramento.

3.6 Un dato clinico concreto: LLM in teriogenologia

Per passare dalla teoria alla pratica, vale la pena esaminare uno un recente studio sull'uso dei modelli linguistici in ambito veterinario. Okur et al. (2026), in un lavoro pubblicato su JAVMA a febbraio 2026, hanno sottoposto 15 scenari ostetrici e ginecologici standardizzati a quattro “valutatori”: due clinici esperti in teriogenologia, due veterinari novizi, ChatGPT-5 e ChatGPT-5 Thinking (la versione con ragionamento a catena). Le risposte sono state giudicate da un pannello di esperti in cieco, su una scala di qualità a 5 punti.

Il risultato è significativo: ChatGPT-5 Thinking ha ottenuto i punteggi qualitativi complessivi più alti, seguito da ChatGPT-5, poi dai clinici esperti, e infine dai novizi. Le risposte generate dai modelli erano generalmente più consistenti e complete di quelle umane. La conclusione degli autori è che i modelli linguistici, in particolare quelli con capacità di ragionamento strutturato, possono funzionare come strumenti di supporto decisionale in scenari clinici complessi e time-sensitive.

Ma attenzione: lo studio è stato condotto su scenari simulati, non su pazienti reali, e in condizioni cold-start (senza contesto clinico pregresso). Nel mondo reale, il clinico esperto ha accesso a informazioni che nessun modello linguistico possiede: la storia del paziente, l'esame fisico, l'intuizione costruita su anni di pratica. Il dato non dice "l'AI è più brava del veterinario"; dice che, in un compito specifico e ben definito, un modello con ragionamento strutturato produce risposte di alta qualità che possono integrare — non sostituire — il giudizio clinico. La distinzione tra assistente e consulente resta intatta.

3.7 AI generativa nella pratica quotidiana: usi ragionevoli

Dopo aver descritto rischi e potenzialità, è importante non lasciare l'impressione che l'AI generativa sia solo un pericolo da evitare. Al contrario, ci sono ambiti in cui può essere uno strumento genuinamente utile nella pratica veterinaria — a patto di usarla per ciò che sa fare bene, e non per ciò che non sa fare.

La redazione è forse l'ambito più naturale. Scrivere note cliniche strutturate, tradurre un referto tecnico in un linguaggio comprensibile per il proprietario, redigere lettere di riferimento a colleghi specialisti, preparare comunicazioni per il personale: sono tutti compiti in cui il modello linguistico eccelle, perché il suo punto di forza — generare testo coerente e ben strutturato — coincide con ciò che serve. L'errore fattuale è ancora possibile, ma se il veterinario fornisce al modello i dati corretti del caso e poi rilegge l'output, il rischio si riduce a un livello gestibile.

Un secondo ambito è l'esplorazione intellettuale. Chiedere a un modello linguistico di elencare le possibili diagnosi differenziali per una presentazione clinica insolita, o di riassumere lo stato dell'arte su un trattamento, può essere un modo efficace per stimolare il ragionamento — a patto di trattare la risposta come un punto di partenza, non come una conclusione. È come consultare un collega molto colto ma non sempre affidabile: le sue idee possono essere preziose, ma vanno sempre verificate.

Un terzo ambito, in crescita rapida, è la formazione. I modelli linguistici possono generare casi clinici simulati per l'addestramento degli studenti, creare quiz di autovalutazione, spiegare concetti complessi con analogie personalizzate. In questo contesto, il rischio di errore è meno critico perché l'obiettivo non è ottenere una verità clinica ma stimolare l'apprendimento.

La regola pratica è semplice: usate l'AI generativa per compiti in cui potete verificare l'output con la vostra competenza. Non usatela come fonte autorevole per decisioni che non siete in grado di valutare indipendentemente.

Una direzione promettente è quella dei “copiloti” specializzati: modelli addestrati esclusivamente su corpora di linee guida validate, non su internet nel suo complesso. Un esempio significativo, dalla medicina umana, è ESC Chat, lanciato nel 2025 dalla Società Europea di Cardiologia: un assistente AI addestrato unicamente sulle linee guida cliniche ESC, che fornisce risposte con tracciabilità della fonte e link diretti alla sezione pertinente della linea guida. Non è un chatbot generico: è uno strumento vincolato a un corpus autorevole. In medicina veterinaria non esistono ancora equivalenti, ma la direzione è questa: meno ChatGPT generico, più strumenti verticali addestrati su fonti controllate e verificabili.

3.8 I limiti attuali e il ritmo del cambiamento

Un aspetto che è importante considerare è la velocità con cui questa tecnologia si evolve. I modelli linguistici disponibili oggi sono significativamente più capaci di quelli di due anni fa, e quelli di due anni fa erano già enormemente più avanzati rispetto a quelli di cinque anni fa. Questo ritmo di sviluppo rende difficile fare affermazioni definitive: un limite che oggi è reale potrebbe essere parzialmente superato tra sei mesi.

Ciò detto, alcuni limiti sono strutturali e non sembrano destinati a scomparire nel breve termine. I modelli linguistici non hanno accesso a informazioni in tempo reale (a meno che non siano collegati a strumenti di ricerca esterni). Non possono verificare autonomamente la correttezza di ciò che generano. Non hanno memoria persistente delle conversazioni passate, se non all'interno di una singola sessione. E soprattutto, non hanno esperienza del mondo fisico: non sanno cosa significa toccare, osservare, annusare, percepire un paziente. Tutta la loro conoscenza è mediata dal linguaggio.

Per il professionista veterinario, questo suggerisce un atteggiamento pragmatico: esplorare con curiosità le possibilità che questi strumenti offrono, testarli nel proprio contesto, ma mantenere sempre la consapevolezza che si tratta di strumenti in rapida evoluzione, con capacità impressionanti e limiti altrettanto reali. La tecnologia migliorerà; la necessità di spirito critico, no.

3.9 L'AI agentica: quando la macchina agisce

Un'evoluzione più recente, e ancora in fase iniziale, merita attenzione perché potrebbe cambiare radicalmente il modo in cui l'AI interagisce con la pratica professionale. Si chiama AI agentica, o agentic AI, e rappresenta un salto concettuale rispetto a tutto ciò che abbiamo descritto finora.

L'AI predittiva analizza dati e produce una classificazione. L'AI generativa produce testo o immagini in risposta a una domanda. L'AI agentica fa qualcosa di diverso: riceve un obiettivo e agisce autonomamente per raggiungerlo, prendendo decisioni intermedie, utilizzando strumenti esterni e adattando il proprio comportamento in base ai risultati ottenuti.

Per capire la differenza, pensate a tre livelli di interazione. Al primo livello, chiedete a un sistema AI: “Questa radiografia mostra un versamento pleurico?”. Il sistema risponde sì o no. È AI predittiva. Al secondo livello, chiedete: “Descrivi questa radiografia e suggerisci un piano diagnostico”. Il sistema genera un testo articolato. È AI generativa. Al terzo livello, dite: “Questo paziente ha una dispnea acuta. Gestisci il workup diagnostico”. Il sistema legge la radiografia, identifica un versamento, consulta la cartella clinica per verificare la storia del paziente, ordina un'ecografia toracica, suggerisce un prelievo per citologia, e prepara una bozza di comunicazione per il proprietario. È AI agentica: non risponde a una domanda, ma esegue una sequenza di azioni orientate a un obiettivo.

I sistemi agentici sono oggi in fase sperimentale nella medicina umana e non ancora operativi in veterinaria. Ma la direzione è chiara: le piattaforme di gestione clinica stanno integrando capacità sempre più autonome, e la transizione da “assistente che risponde” a “agente che agisce” è già in corso. Per il professionista, questo alza ulteriormente la posta in gioco della comprensione: un sistema che non solo suggerisce ma agisce richiede un livello di supervisione e di comprensione dei meccanismi ancora più elevato.

3.10 Un riassunto in tre concetti

Anche qui, fissiamo i punti chiave:

Generazione, non conoscenza. L'AI generativa produce testo (o immagini, o audio) statisticamente plausibile, non fatti verificati. Il modello non garantisce verità fattuale né ragionamento clinico affidabile: genera output in base a pattern statistici del linguaggio, non a un processo di verifica dei fatti.

Allucinazioni. Il modello può generare contenuti falsi con la stessa fluidità e sicurezza con cui genera quelli corretti. Non sa di non sapere.

Assistente, non consulente. L'AI generativa può essere uno strumento prezioso per la redazione, la sintesi e la comunicazione. Ma ogni suo output deve essere verificato da chi ha la competenza clinica per farlo.

4. L'AI nella pratica veterinaria: sei finestre

I principi descritti nei capitoli precedenti non sono astrazioni teoriche: sono già operativi nella medicina veterinaria. In questa sezione apriamo sei finestre — brevi, concrete — su altrettante aree in cui l'intelligenza artificiale sta cambiando la pratica professionale. Ogni

finestra parte dal meccanismo per arrivare all'applicazione, con un occhio sempre aperto sui limiti.

Le sei finestre che seguono non pretendono di essere un catalogo esaustivo: le applicazioni dell'AI in veterinaria sono già più numerose e si moltiplicano rapidamente. Ho scelto questi sei ambiti perché ciascuno illustra un meccanismo diverso dell'intelligenza artificiale — riconoscimento di immagini, predizione da dati numerici, personalizzazione terapeutica, monitoraggio da sensori, elaborazione del linguaggio, integrazione di dati eterogenei. L'obiettivo non è mappare tutti i prodotti disponibili, ma mostrare come i principi descritti nei capitoli precedenti si traducono in strumenti concreti, con i loro punti di forza e i loro limiti.

4.1 Pattern nei pixel — Diagnostica per immagini

La diagnostica per immagini è il campo in cui l'AI veterinaria è più avanzata, e il motivo è semplice: è il tipo di problema che le reti neurali risolvono meglio. Quando un modello viene addestrato su centinaia di migliaia di radiografie, ciascuna accompagnata dalla diagnosi di uno specialista, impara a riconoscere anomalie con una velocità e una costanza impossibili per l'occhio umano. Oggi almeno otto aziende nel mondo offrono software di questo tipo per la veterinaria, con la più diffusa utilizzata in oltre 2.300 cliniche e capace di analizzare circa 50.000 radiografie alla settimana (Ndiaye et al., 2025).

Uno studio del 2025 ha messo a confronto uno di questi sistemi con undici specialisti certificati in radiologia veterinaria: l'AI ha eguagliato il migliore di loro in accuratezza descrittiva, risultando più precisa nel confermare la normalità, ma meno efficace nell'individuare le anomalie (Ndiaye et al., 2025). In termini tecnici: alta specificità, sensibilità inferiore. Per la pratica clinica, questo significa che il sistema è un buon secondo lettore — rinforza la fiducia quando tutto è normale — ma non può sostituire l'occhio esperto nella ricerca di ciò che non va.

Questo studio merita però una lettura critica. Una commentary (Joslyn et al., 2025) ha sollevato un punto metodologico importante che illustra un concetto valido per qualsiasi applicazione AI: il bias di prevalenza. Nel dataset utilizzato, la percentuale di casi patologici era così bassa che una strategia banale — classificare tutto come “normale” senza nemmeno guardare le immagini — avrebbe raggiunto un'accuratezza dell'84%. Questo non invalida i risultati, ma ci ricorda una regola fondamentale: quando i casi positivi sono rari, l'accuratezza complessiva diventa una metrica fuorviante, e può dare un'impressione troppo positiva delle prestazioni di qualsiasi sistema.

4.2 Pattern nei numeri — Medicina predittiva

Qui il principio è lo stesso — trovare regolarità nei dati — ma applicato a numeri invece che a immagini. L'esempio più significativo viene dalla nefrologia del gatto: un modello di

machine learning addestrato su oltre 106.000 pazienti che analizza esami del sangue e delle urine di routine per predire la malattia renale prima che si manifesti clinicamente (Bradley et al., 2019).

Il dato che rende questo esempio illuminante dal punto di vista dell'AI è questo: tutti quei parametri erano già nelle mani dei veterinari. Erano gli stessi esami di sempre. Ciò che il modello ha aggiunto è la capacità di riconoscere micro-variazioni combinate — un pattern troppo sottile e troppo diverso da paziente a paziente per essere colto dall'osservazione umana. Vicino al momento della diagnosi, il modello individua oltre il 90% dei casi. Ma la sfida vera è la predizione anticipata: a un anno di distanza la sensibilità scende al 63%, a due anni al 44%, pur mantenendo pochissimi falsi allarmi.

Una nota di cautela che vale per tutta la medicina predittiva: predire non è curare. La predizione precoce è clinicamente utile solo se cambia la gestione del paziente. Se non esistono ancora terapie efficaci in quella fase così precoce, sapere in anticipo che un animale si ammalerà potrebbe non modificare sostanzialmente l'esito. È un potenziale promettente, non ancora una soluzione definitiva.

4.3 Pattern nel singolo paziente — Medicina personalizzata

La medicina personalizzata rappresenta forse l'applicazione più ambiziosa dell'AI in ambito clinico: non più classificare popolazioni, ma predire come risponderà il singolo individuo a una specifica terapia. Il principio è quello del machine learning predittivo, ma applicato a una scala diversa: i dati provengono da un singolo paziente — le sue cellule, il suo profilo molecolare — e il modello li confronta con migliaia di casi precedenti per predire la risposta individuale.

Un esempio concreto viene dall'oncologia veterinaria. La piattaforma Imprimed testa farmaci antitumorali sulle cellule vive prelevate dal tumore del paziente, poi usa modelli AI per predire quale terapia funzionerà meglio per quell'individuo specifico. Uno studio su cani con linfoma (Callegari et al., 2024) ha riportato una concordanza elevata tra le previsioni del modello e la risposta effettiva alla terapia. I risultati sono promettenti, anche se i dati peer-reviewed sono ancora limitati e vanno interpretati con cautela. Caso raro e significativo: la piattaforma si sta espandendo alla medicina umana — una traslazione dalla veterinaria verso l'umana.

Un approccio più recente e meno invasivo arriva dalla genomica: analizzare il sangue periferico del paziente, senza bisogno di prelievi dal tumore, per individuare firme genetiche che predicono la risposta al trattamento. Rao et al. (2026) hanno dimostrato che tre geni rilevabili già al settimo giorno dopo il trattamento predicono l'esito a lungo termine nei cani con linfoma. Il principio AI è lo stesso — riconoscimento di pattern in dati molecolari — ma la barriera d'ingresso per il clinico si abbassa.

Un terzo ambito in cui l'AI sta avendo un impatto concreto è la patologia digitale. In oncologia, la classificazione dei tumori dipende spesso dal conteggio manuale di strutture cellulari al microscopio — un compito tedioso e notoriamente soggetto a variabilità tra un patologo e l'altro: due esperti possono guardare lo stesso vetrino e arrivare a conteggi diversi, con conseguenze dirette sulla prognosi. Algoritmi di deep learning addestrati su decine di migliaia di immagini annotate (Aubreville et al., 2022) hanno dimostrato di ridurre drasticamente questa variabilità, portando la concordanza tra patologi assistiti da AI a livelli significativamente superiori rispetto al conteggio manuale. È un caso esemplare del ruolo che l'AI dovrebbe avere: non sostituire l'esperto, ma renderlo più accurato e riproducibile.

4.4 Pattern nella mandria — Monitoraggio intelligente degli allevamenti

Cambiando scala completamente, l'AI si applica anche al monitoraggio continuo degli animali da reddito — un campo noto come Precision Livestock Farming. Il meccanismo è quello del machine learning alimentato da sensori: collari, ear tag elettronici, biosensori che rilevano in tempo reale temperatura, frequenza cardiaca, attività e posizione di ogni singolo animale.

Il principio è quello delle anomalie: il modello impara qual è il comportamento normale di un animale e segnala le deviazioni, spesso prima che siano visibili all'occhio dell'allevatore. Una review recente (Ghavi Hossein-Zadeh, 2025) documenta la rapida espansione di queste applicazioni, con modelli di deep learning e computer vision che raggiungono prestazioni promettenti nella sorveglianza sanitaria automatizzata, pur sottolineando le sfide legate alla qualità dei dati, alla variabilità tra un allevamento e l'altro, e alla necessità di validazione in condizioni reali. Per la veterinaria pubblica, è il passaggio dalla sorveglianza reattiva — intervenire quando il problema è conclamato — a quella predittiva: intervenire quando il modello segnala che qualcosa sta per accadere.

4.5 Pattern nelle parole — Documentazione clinica

Qui entra in gioco l'AI generativa applicata al lavoro quotidiano del veterinario: dettatura automatica di note cliniche, sintesi delle cartelle, comunicazioni ai proprietari, triage telefonico. La survey sugli specialisti ACVIM/ECVIM-CA (Bertin et al., 2026) indica che i cosiddetti “scribi AI” — sistemi che trascrivono e strutturano le note cliniche durante la visita — sono già lo strumento AI più adottato nella pratica internistica.

Il meccanismo è quello del modello linguistico applicato a un compito specifico: il sistema ascolta la conversazione tra veterinario e proprietario, la trascrive, e la riorganizza in un formato clinico strutturato. Il vantaggio è evidente: risparmio di tempo, documentazione più completa, minor rischio di dimenticare dettagli. Ma il rischio è altrettanto evidente: il modello può fraintendere, omettere, o inventare. La supervisione umana resta indispensabile.

Un esempio illustrativo delle potenzialità, tratto dalla pratica clinica: la funzione di sintesi intelligente di una piattaforma gestionale ha permesso a un veterinario di scoprire quindici telefonate simili distribuite su diversi anni per lo stesso paziente — un cane che vomitava — ciascuna singolarmente insignificante, ma nel complesso indicativa di un problema cronico. Il pattern era nelle parole, non nei numeri. Nessun clinico avrebbe riletto anni di appunti telefonici per un singolo paziente; il modello lo fa in secondi.

4.6 Pattern nel territorio — Epidemiologia e sorveglianza

Infine, l'AI si applica alla scala territoriale: modelli predittivi per la diffusione di malattie trasmesse da insetti vettori, per la previsione di focolai nelle filiere animali, per il rilevamento precoce di malattie negli allevamenti. Il meccanismo è quello dell'integrazione di dati eterogenei: spostamenti di animali, dati meteorologici, esiti di laboratorio, notifiche sanitarie — fonti diverse che, combinate da un modello, possono rivelare trend invisibili ai singoli operatori.

La review di Ghavi Hossein-Zadeh (2025) documenta diverse applicazioni sperimentali in quest'area, ma sottolinea che la maggior parte è ancora in fase di sviluppo e che la traslazione dalla ricerca alla pratica di sorveglianza resta una sfida aperta. È un campo in cui l'AI potrebbe avere un impatto enorme, ma dove la qualità e l'integrazione dei dati restano il collo di bottiglia principale.

5. I rischi: tre trappole da conoscere

5.1 L'automation bias

Il rischio più insidioso nell'uso dell'AI in ambito clinico non è che il software sbaglia. È che il professionista smetta di accorgersi quando sbaglia.

L'automation bias è un fenomeno psicologico ampiamente documentato nella letteratura scientifica (Goddard et al., 2012), dall'aviazione alla medicina umana, dalla finanza al controllo industriale. Si manifesta come la tendenza a fidarsi dell'output di un sistema automatizzato anche quando la propria esperienza, il proprio giudizio o le evidenze disponibili suggeriscono il contrario. Non è pigrizia: è un meccanismo cognitivo. Il nostro cervello tende naturalmente ad allinearsi con una risposta che appare autorevole, strutturata, presentata con sicurezza. Se un software dice “normale”, il livello di attenzione cala.

In medicina umana, gli studi sull'automation bias hanno prodotto risultati preoccupanti. Radiologi e dermatologi affiancati da sistemi AI possono migliorare la propria performance complessiva, ma tendono a essere influenzati dagli errori del sistema: quando l'AI fornisce un suggerimento sbagliato, una parte dei clinici lo segue, riducendo

l'accuratezza della decisione (Tschantl et al., 2020). Non perché siano meno competenti, ma perché il suggerimento automatico abbassa la soglia di attenzione critica.

Per il veterinario che utilizza un software di refertazione AI o un sistema di supporto decisionale, questo significa una cosa concreta: il valore aggiunto dell'AI è reale solo se si mantiene attivo il proprio giudizio critico. Nel momento in cui il software diventa un oracolo — qualcosa di cui ci si fida senza verificare — il rischio di errore non diminuisce: aumenta.

5.2 Le allucinazioni in contesto clinico

Il rischio delle allucinazioni dell'AI generativa, già descritto nel capitolo precedente, assume una valenza particolare quando il contesto è clinico. Un testo generato da un modello linguistico su un argomento medico presenta due caratteristiche che lo rendono pericoloso: è formulato con il linguaggio e la struttura che un professionista si aspetta, e non porta con sé nessun segnale visibile che distingua le parti corrette da quelle inventate.

Un modello che genera una diagnosi differenziale con cinque ipotesi potrebbe avere le prime tre perfettamente corrette e le ultime due completamente inventate — e tutte e cinque saranno scritte con la stessa sicurezza e la stessa coerenza interna. Non c'è un asterisco sulle allucinazioni. Non c'è un avviso. L'unico filtro è la competenza di chi legge.

Questo ha un'implicazione paradossale: l'AI generativa è più pericolosa per chi ne sa meno. Un esperto può riconoscere l'errore; un neofita rischia di prenderlo per oro colato. E questo ci riporta al paradosso dell'adozione con cui abbiamo aperto: se quasi la metà degli utilizzatori dichiara scarsa familiarità, il rischio che le allucinazioni non vengano riconosciute non è teorico.

Il rischio non riguarda solo le allucinazioni del modello, ma anche le affermazioni di chi lo commercializza. Nel 2025, la startup OpenVet ha dichiarato che il proprio sistema AI aveva ottenuto un punteggio del 100% al NAVLE, l'esame abilitante per la professione veterinaria in Nord America. L'International Council for Veterinary Assessment (ICVA), che amministra il NAVLE, ha contestato formalmente l'affermazione, definendola falsa e chiarendo che l'esame è somministrato esclusivamente a candidati umani in condizioni controllate e con contenuti non accessibili pubblicamente. OpenVet aveva in realtà testato il proprio modello su domande di pratica o materiali disponibili, non sull'esame reale. Il comunicato stampa è stato successivamente rimosso. È un esempio emblematico di come l'hype tecnologico possa superare la realtà, e di perché le validazioni indipendenti siano indispensabili.

5.3 Il vuoto regolamentare

Un terzo elemento di rischio riguarda il quadro normativo — o, più precisamente, la sua assenza nel contesto veterinario.

L'Unione Europea ha adottato nel 2024 il Regolamento sull'Intelligenza Artificiale (EU AI Act, Reg. 2024/1689), entrato in vigore il 1° agosto 2024 ma con un'applicazione scaglionata: alcune disposizioni (come il divieto dei sistemi a rischio inaccettabile) sono già operative, il grosso delle regole si applica dal 2 agosto 2026, mentre gli obblighi per i sistemi ad alto rischio integrati in prodotti regolati (tra cui i dispositivi medici) scatteranno dal 2 agosto 2027. Il Regolamento classifica i sistemi AI in base al rischio e impone obblighi stringenti ai sistemi ad alto rischio, tra cui quelli integrati in dispositivi medici o destinati al supporto decisionale clinico. In ambito sanitario umano, un software AI diagnostico deve rispettare requisiti di trasparenza, qualità dei dati, robustezza tecnica e supervisione umana. L'Italia, con la Legge 23 settembre 2025, n. 132, è stato il primo Paese UE a dotarsi di un quadro normativo nazionale sull'AI, allineato all'AI Act europeo.

Ma ecco il punto critico per la vostra professione: i dispositivi veterinari non rientrano automaticamente nel Regolamento europeo sui dispositivi medici (EU MDR/IVDR), che si applica a prodotti destinati all'uso su esseri umani. Di conseguenza, il software AI per la diagnostica veterinaria non segue il percorso di certificazione previsto per i dispositivi medici umani e si colloca in un quadro normativo meno armonizzato e meno presidiato. Questo non significa che sia del tutto escluso da ogni forma di regolazione — possono applicarsi norme generali sulla sicurezza dei prodotti e, in alcuni casi, disposizioni dell'AI Act — ma nella pratica il risultato è che, mentre un software AI per uso umano deve rispettare requisiti stringenti di validazione clinica, trasparenza e sorveglianza post-commercializzazione, per applicazioni veterinarie non esiste oggi un livello di controllo comparabile e uniformemente applicato.

Le principali società specialistiche stanno prendendo posizione: ACVR e ECVDI hanno pubblicato una position statement congiunta sull'AI (ACVR & ECVDI, 2025) in cui chiedono trasparenza nello sviluppo dei modelli, buone pratiche di machine learning, reporting sistematico degli errori e supervisione clinica obbligatoria. È un segnale che la governance del settore è ancora in evoluzione e che la comunità scientifica veterinaria sta cercando di colmare il divario normativo dall'interno della professione.

La domanda di fondo, ancora senza risposta univoca, è: chi risponde se l'AI sbaglia e il veterinario si è fidato? Il produttore del software? Il veterinario? Entrambi? In assenza di un quadro normativo dedicato, la responsabilità resta di fatto interamente sul professionista, che utilizza strumenti privi di certificazione formale sulla base della propria valutazione.

La conseguenza pratica è anche assicurativa. L'utilizzo di software AI non certificato come dispositivo medico in ambito diagnostico introduce un profilo di rischio che molte polizze di responsabilità professionale non contemplano ancora in modo esplicito. Questo non implica automaticamente l'assenza di copertura, ma può generare aree di incertezza: se un errore diagnostico è influenzato da un suggerimento errato di un sistema AI, la valutazione della responsabilità e dell'eventuale copertura assicurativa dipenderà dalle

condizioni specifiche della polizza e dal comportamento del professionista. In particolare, il rischio aumenta quando l'AI viene utilizzata in modo acritico o senza adeguata verifica clinica. È un punto che merita attenzione da parte degli ordini professionali e delle compagnie assicurative, e che il singolo veterinario dovrebbe verificare nella propria polizza prima di integrare strumenti AI nel flusso diagnostico.

Non a caso, lo studio di Li & Lai (2026) riporta che il 93,8% dei veterinari intervistati chiede una supervisione regolamentare sull'AI. E la preoccupazione principale, espressa dal 54,3%, riguarda proprio l'affidabilità e l'accuratezza degli strumenti. Questa richiesta appare tanto più fondata alla luce del vuoto normativo descritto: il professionista è oggi chiamato a valutare da solo l'affidabilità di strumenti che, in medicina umana, sarebbero sottoposti a un processo di certificazione indipendente.

6. Conclusione: usare lo strumento da professionisti

La metafora che dà il titolo al convegno in cui questo documento viene presentato è quella delle mangrovie: alberi le cui radici crescono sia nell'aria che nell'acqua, in un intreccio così fitto che a un certo punto non si distingue più dove finisce la pianta e dove inizia l'elemento liquido. L'intelligenza artificiale nella medicina veterinaria sta facendo esattamente questo: si sta intrecciando con la pratica clinica, la diagnostica, la gestione dei dati, la comunicazione, in modo così pervasivo che spesso non ci si accorge nemmeno di usarla.

Le radici aeree delle mangrovie non sostituiscono il terreno. Lo raggiungono. Lo stabilizzano. Permettono all'albero di crescere dove altrimenti non potrebbe — in quell'ambiente ibrido tra terra e acqua dove nessun altro albero sopravvive.

L'intelligenza artificiale è la radice aerea della vostra professione. Estende ciò che potete fare: vedere pattern invisibili nei dati di routine, predire malattie prima che si manifestino, personalizzare terapie, sorvegliare mandrie intere, redigere documentazione in una frazione del tempo. Ma il terreno — il giudizio clinico, l'esperienza accumulata in anni di pratica, la conoscenza del singolo paziente, la relazione col proprietario — quello resta insostituibile. E resta vostro.

La sfida per il medico veterinario, oggi, non è adottare l'AI o rifiutarla. È usarla sapendo come funziona, sapendo dove può sbagliare, e mantenendo sempre acceso il proprio spirito critico. Sapere che un modello predittivo trova pattern nei dati ma dipende dalla qualità di quei dati. Sapere che un modello generativo produce testo plausibile, non verità verificate. Sapere che il proprio cervello tende a fidarsi di ciò che appare autorevole, e che questa tendenza va contrastata attivamente.

Questa consapevolezza non può essere lasciata all'improvvisazione individuale. Serve formazione strutturata, e i primi passi si stanno muovendo: Huang e Chu (2026) hanno

proposto un framework curricolare specifico per l'alfabetizzazione AI nelle facoltà di medicina veterinaria, articolato su competenze tecniche, etiche e cliniche. È un segnale che la comunità accademica sta prendendo sul serio la necessità di preparare i professionisti di domani non solo a usare questi strumenti, ma a comprenderli e a valutarli criticamente.

Usare l'intelligenza artificiale, insomma, come il professionista consapevole che ognuno di voi è. Con curiosità, con apertura, ma senza mai spegnere il motore critico che è il vero strumento insostituibile della vostra professione.

Disclaimer. *Le opinioni e le analisi contenute in questo documento sono espresse a titolo esclusivamente personale dall'autore e non rappresentano la posizione di FNOVI, ENPAV, né di alcuna istituzione o organizzazione. Le informazioni fornite hanno finalità divulgativa e non intendono sostituire la consulenza di specialisti in intelligenza artificiale o in medicina veterinaria. L'autore ha fatto ogni ragionevole sforzo per verificare l'accuratezza dei dati riportati, ma non può garantirne la completezza o l'assenza di errori.*

L'autore. *Francesco Masera è autore del volume Intelligenza artificiale: come funziona davvero (2026, prefazione di Paolo Savona).*

Email: fmasera@gmail.com

LinkedIn: [linkedin.com/in/francesco-masera](https://www.linkedin.com/in/francesco-masera)

Bibliografia

- ACVR & ECVI. (2025). American College of Veterinary Radiology and European College of Veterinary Diagnostic Imaging position statement on artificial intelligence. Journal of the American Veterinary Medical Association, 263(6). <https://doi.org/10.2460/javma.25.01.0027>
- Aubreville, M., Bertram, C. A., Donovan, T. A., Klopfleisch, R., Maier, A., & others. (2022). Computer-assisted mitotic count using a deep learning-based algorithm improves interobserver reproducibility and accuracy. Veterinary Pathology, 59(2), 211–226. <https://doi.org/10.1177/03009858211067478>
- Bertin, F.-R., Lawrence, J., Niessen, S. J. M., Pinard, C. J., Reagan, K. L., & Rentko, V. (2026). Use of, and attitudes toward, artificial intelligence in members of the American College of Veterinary Internal Medicine and the European College of Veterinary Internal Medicine — Companion Animals. Journal of Veterinary Internal Medicine, 40(1), aalafo36. <https://doi.org/10.1093/jvimsj/aalafo36>
- Bradley, R., Tagkopoulos, I., Kim, M., Kokkinos, Y., Panagiotakos, T., Kennedy, J., De Meyer, G., Watson, P., & Elliott, J. (2019). Predicting early risk of chronic kidney disease in cats using routine clinical laboratory tests and machine learning. Journal of Veterinary Internal Medicine, 33(6), 2644–2656. <https://doi.org/10.1111/jvim.15623>

- Callegari, A. J., Tsang, J., Park, S., Swartzfager, D., Kapoor, S., Choy, K., & Lim, S. (2024). Functional drug susceptibility testing predicts treatment outcome in dogs with relapsed B-cell lymphoma. *Frontiers in Oncology*, 14, 1304144. <https://doi.org/10.3389/fonc.2024.1304144>
- European Parliament and Council. (2024). Regulation (EU) 2024/1689 laying down harmonised rules on artificial intelligence (Artificial Intelligence Act). *Official Journal of the European Union*, L 2024/1689.
- Ghavi Hossein-Zadeh, N. (2025). Artificial intelligence in veterinary and animal science: Applications, challenges, and future prospects. *Computers and Electronics in Agriculture*, 235, 110395. <https://doi.org/10.1016/j.compag.2025.110395>
- Goddard, K., Roudsari, A., & Wyatt, J. C. (2012). Automation bias: A systematic review of frequency, effect mediators, and mitigators. *Journal of the American Medical Informatics Association*, 19(1), 121–127. <https://doi.org/10.1136/amiainl-2011-000089>
- Huang, Y.-T., & Chu, C. P. (2026). Curriculum framework for artificial intelligence literacy in veterinary education. *Frontiers in Veterinary Science*, 13, 1801756. <https://www.frontiersin.org/journals/veterinary-science/articles/10.3389/fvets.2026.1801756>
- Joslyn, S. K., et al. (2025). Commentary: Comparison of radiological interpretation made by veterinary radiologists and state-of-the-art commercial AI software for canine and feline radiographic studies. *Frontiers in Veterinary Science*, 12, 1615947. <https://doi.org/10.3389/fvets.2025.1615947>
- Legge 23 settembre 2025, n. 132. Disposizioni e deleghe al Governo in materia di intelligenza artificiale. *Gazzetta Ufficiale della Repubblica Italiana*, Serie Generale n. 225, 25 settembre 2025.
- Li, S., & Lai, X. (2026). The adoption paradox for veterinary professionals in China: High use of artificial intelligence despite low familiarity. *Frontiers in Veterinary Science*, 13, 1727001. <https://doi.org/10.3389/fvets.2026.1727001>
- Masera, F. (2026). *Intelligenza artificiale: come funziona davvero. Guida completa per chi parte da zero*. Prefazione di P. Savona. <https://amzn.eu/d/ooq7bwfb>
- Ndiaye, Y. S., Cramton, P., Chernev, C., Ockenfels, A., & Schwarz, T. (2025). Comparison of radiological interpretation made by veterinary radiologists and state-of-the-art commercial AI software for canine and feline radiographic studies. *Frontiers in Veterinary Science*, 12, 1502790. <https://doi.org/10.3389/fvets.2025.1502790>
- Okur, D. T., Cengiz, M., Küçükaslan, İ., Peker, C., Çiplak, A. Y., Tohumcu, V., & Aydın, Ş. (2026). Performance of large language models versus clinicians and novices in veterinary theriogenology decision support. *Journal of the American Veterinary Medical Association*. Advance online publication. <https://doi.org/10.2460/javma.25.09.0615>
- Parasuraman, R., & Manzey, D. H. (2010). Complacency and bias in human use of automation: An attentional integration. *Human Factors*, 52(3), 381–410. <https://doi.org/10.1177/0018720810376055>
- Rao, K., et al. (2026). Peripheral blood mononuclear cell gene expression signatures predict long-term survivorship in canine DLBCL. *Scientific Reports*, 16. <https://doi.org/10.1038/s41598-026-44677-0>
- Tschandl, P., Rinner, C., Apalla, Z., Argenziano, G., Codella, N., Halpern, A., Janda, M., Lallas, A., Longo, C., Malvehy, J., Paoli, J., Puig, S., Rosendahl, C., Soyer, H. P., Zalaudek, I., & Kittler, H. (2025). The 2025 update of the European Society of Medical Oncology (ESMO) clinical practice guidelines for the management of cutaneous melanoma. *Annals of Oncology*, 36(1), 1–22. <https://doi.org/10.1016/j.annonc.2024.10.001>

H. (2020). Human–computer collaboration for skin cancer recognition. *Nature Medicine*, 26(8), 1229–1234. <https://doi.org/10.1038/s41591-020-0942-0>