Insight ON: AI

Chi gestisce davvero la tua AI?

Il vero significato di un approccio ‘End-to-End’

InsightAI

Molti leader tecnologici hanno già vissuto questo scenario: il progetto si conclude, il nuovo sistema viene rilasciato e il partner si sposta su una nuova iniziativa. L'organizzazione resta così responsabile di una soluzione che non è stata pensata per una gestione a lungo termine. La governance è stata rinviata a un momento successivo, i team chiamati a utilizzarla quotidianamente non hanno contribuito alla sua progettazione e la documentazione descrive come la soluzione è stata sviluppata, senza però fornire indicazioni concrete su come gestirla nel tempo.

Le partnership focalizzate esclusivamente sulla fase di delivery sono ancora molto diffuse nei progetti tecnologici. Sebbene questo approccio possa rendere particolarmente complessa la gestione delle sfide che emergono dopo il go-live, per alcune tipologie di software le criticità possono comunque rimanere gestibili. Con l'AI, però, raramente è così. Un sistema di AI in produzione presenta infatti una serie di esigenze operative continue che vanno oltre quelle del software tradizionale, e questi requisiti non sempre si manifestano in modo evidente quando iniziano a sorgere problemi. Molte organizzazioni che sviluppano soluzioni di AI con una collaborazione limitata alla sola fase di implementazione non comprendono appieno le implicazioni di questa scelta fino a quando non si trovano ad affrontarle direttamente: il partner ha già concluso il proprio incarico e il sistema che rimane in gestione si rivela più complesso da mantenere, governare e rendere affidabile nel tempo.

Questo è il delivery gap: il divario che si crea tra il lancio di un sistema e una soluzione di AI che operi in modo affidabile, governato e conforme alle regole. Le organizzazioni che riescono a colmare questo divario tendono a condividere un approccio comune: collaborano con partner che, fin dal primo giorno, progettano e pianificano pensando a ciò che accadrà dopo dodici mesi, non solo al momento del go-live.

Nel 2026, colmare il delivery gap è più urgente che mai. A partire dal 2 agosto, gli obblighi previsti dall'EU AI Act per i sistemi ad alto rischio diventeranno pienamente applicabili alle organizzazioni che operano nell'Unione Europea. La normativa richiede non solo di sviluppare l'AI in modo responsabile, ma anche di gestirla responsabilmente nel tempo. Ciò significa effettuare valutazioni di conformità, disporre di processi documentati, garantire un monitoraggio continuo e definire chiaramente le responsabilità legate alle decisioni automatizzate. In questo contesto, il go-live non rappresenta più il traguardo finale; per le organizzazioni che hanno impostato le proprie collaborazioni sull'AI con un perimetro limitato alla sola fase di delivery, questo divario operativo e gestionale rappresenta una sfida concreta da affrontare e colmare.

L'accordo provvisorio introduce inoltre una tempistica definita per l'applicazione differita delle norme relative ai sistemi di AI ad alto rischio: le nuove date previste sono il 2 dicembre 2027 per i sistemi di AI ad alto rischio autonomi (stand-alone) e il 2 agosto 2028 per i sistemi di AI ad alto rischio integrati all'interno di prodotti. Inoltre, l'accordo provvisorio ripristina l'obbligo per i fornitori di registrare i sistemi di AI nel database europeo dei sistemi ad alto rischio, anche nei casi in cui ritengano che tali sistemi possano beneficiare di un'esenzione dalla classificazione come AI ad alto rischio.L'accordo reintroduce inoltre il principio della stretta necessità per il trattamento delle categorie particolari di dati personali (dati sensibili) quando tale trattamento è finalizzato a garantire l'individuazione e la correzione di potenziali bias nei sistemi di AI.

L'accordo provvisorio rinvia inoltre al 2 agosto 2027 il termine entro il quale le autorità competenti nazionali dovranno istituire i sandbox normativi per l'AI. Allo stesso tempo, riduce da sei a tre mesi il periodo concesso ai fornitori per implementare soluzioni di trasparenza relative ai contenuti generati artificialmente, fissando la nuova scadenza al 2 dicembre 2026. L'intesa raggiunta dai co-legislatori chiarisce inoltre le competenze dell'AI Office nella supervisione dei sistemi di AI basati su modelli di uso generale (general-purpose AI), nei casi in cui sia il modello sia il relativo sistema siano sviluppati dallo stesso fornitore. L'accordo definisce anche una serie di eccezioni in cui la competenza continuerà a rimanere in capo alle autorità nazionali, tra cui i settori dell'applicazione della legge, della gestione delle frontiere, delle autorità giudiziarie e delle istituzioni finanziarie.

Gli analisti legali e gli esperti del settore sottolineano che questo rinvio non rappresenta una "pausa" o un rinvio della conformità normativa. I requisiti fondamentali – come la governance dei dati, la documentazione tecnica e la supervisione umana – rimangono invariati. La raccomandazione principale dell'EU AI Office e degli studi legali specializzati è di considerare questa estensione di 16 mesi come un periodo da utilizzare per costruire e rafforzare le capacità necessarie alla conformità, non come un periodo di attesa. La complessità del percorso verso la conformità normativa, soprattutto per i sistemi di AI ad alto rischio, richiede infatti spesso importanti cambiamenti operativi, che possono richiedere oltre un anno per essere pianificati e implementati correttamente.

Strategia, engineering e operations: perché i partner per l'AI devono coprire l'intero percorso

Il delivery gap si presenta con tale frequenza perché la maggior parte delle collaborazioni nell'ambito dell'AI è limitata a una sola fase di un percorso composto da tre elementi: strategia, implementazione e gestione operativa. Quando l'ambito di un progetto si ferma a una di queste fasi, possono emergere tre tipologie di criticità che compromettono il successo dell'iniziativa nel lungo periodo.

La prima si verifica quando la consulenza si ferma alla strategia. Vengono prodotti una roadmap coerente, un business case ben strutturato e una visione convincente del potenziale della soluzione. Tuttavia, la concreta realizzazione tecnica diventa responsabilità di qualcun altro. Le domande che determinano se un sistema di AI funzionerà davvero in produzione – ad esempio come il modello si comporterà su dati reali invece che su dataset di test, quali saranno i costi inferenziali su larga scala o come si integrerà con i sistemi già presenti in azienda – richiedono il coinvolgimento, fin dall'inizio, di persone con competenze ingegneristiche. Le discrepanze tra la strategia e la realtà tecnica possono creare problemi in qualsiasi progetto tecnologico, ma nell'AI tendono a emergere più tardi e a essere più costose da correggere.

La seconda criticità riguarda i partner che si fermano al go-live. Il sistema viene rilasciato, funziona correttamente il primo giorno e il progetto si conclude. Tuttavia, un sistema di AI in produzione non è un prodotto finito. A differenza del software tradizionale, può deteriorarsi nel tempo in modi non immediatamente evidenti: il modello può andare incontro a fenomeni di drift, le misure di controllo e sicurezza richiedono manutenzione continua man mano che cambiano le modalità di utilizzo e gli output possono apparire plausibili pur non essendo più affidabili. Garantire prestazioni adeguate richiede monitoraggio costante, governance attiva e una chiara responsabilità sul fatto che le decisioni generate dal sistema continuino a essere corrette e appropriate. Se questi aspetti non rientrano nell'ambito della collaborazione, l'organizzazione si ritrova con una soluzione funzionante ma senza il supporto necessario per mantenerla tale nel tempo.

La terza criticità è spesso meno visibile nelle fasi iniziali, ma rappresenta uno degli ostacoli più rilevanti nella pratica: l'integrazione tra le moderne tecnologie di AI e i sistemi legacy su cui si basano molte aziende. La ricerca The Digital Sovereignty Trilemma ha rilevato che il 41% dei responsabili IT afferma che l'impossibilità di modernizzare o migrare le applicazioni legacy sta limitando concretamente la propria organizzazione. Si tratta di una sfida diffusa e concreta. Anche le aziende che non sono direttamente frenate dai sistemi esistenti raramente operano in ambienti completamente nuovi: la maggior parte dispone di decenni di applicazioni, processi e architetture dati che non sono stati progettati pensando all'AI.

L'integrazione con questi sistemi è anche l'ambito in cui le collaborazioni focalizzate esclusivamente sulla delivery tendono più spesso a lasciare attività incompiute. Le connessioni più complesse – quelle tra le nuove capacità offerte dall'AI e le piattaforme che supportano le altre funzioni aziendali – vengono spesso implementate quanto basta per consentire il lancio della soluzione, rinviando a una fase successiva gli interventi più profondi. Una volta terminato il progetto, quel lavoro rimandato diventa responsabilità dell'organizzazione.

Per questo motivo, una vera capacità end-to-end richiede che lo stesso partner accompagni l'azienda lungo tutte le fasi del percorso. L'intelligenza operativa nasce da una strategia fondata sulla realtà tecnica, da un'implementazione progettata per funzionare fin dal primo giorno e da un modello operativo pensato sin dall'inizio per mantenere i sistemi efficienti, conformi alle normative e in costante miglioramento. E, aspetto altrettanto importante, significa collaborare con un partner capace di individuare e progettare casi d'uso che generino un reale valore di business, non semplicemente applicazioni tecnologicamente possibili.

Costruire competenze AI durature: l'esperienza di un leader mondiale dell'istruzione

Ciò che significa davvero colmare il delivery gap nella pratica va ben oltre il semplice corretto utilizzo della tecnologia. Siamo stati contattati da un importante fornitore globale di servizi educativi, attivo in quasi 200 Paesi e specializzato nella fornitura di contenuti digitali, sistemi di valutazione e soluzioni di apprendimento supportate dalla tecnologia su scala internazionale. La divisione engineering di questa organizzazione si è rivolta a Insight AI con una domanda che molte aziende tendono a porsi troppo tardi: non solo se l'AI potesse funzionare, ma se i propri team fossero realmente in grado di gestirla in modo sicuro, governarla correttamente e migliorarla nel tempo in autonomia.

Quando Insight AI ha avviato la collaborazione con i team tecnici dell'organizzazione, un sondaggio iniziale tra gli ingegneri coinvolti ha evidenziato un aspetto che i dati sull'adozione dell'AI tendono spesso a nascondere. La maggior parte del personale utilizzava già regolarmente strumenti di intelligenza artificiale, quindi, sulla carta, i livelli di adozione apparivano positivi. Tuttavia, si tratta di una situazione comune a molte organizzazioni: usare frequentemente strumenti AI non significa necessariamente utilizzarli in modo efficace o sicuro. Analizzando più a fondo le modalità di utilizzo, è emerso che la valutazione strutturata dei risultati prodotti dall'AI avveniva in modo discontinuo e che le tecniche più avanzate – quelle che spesso generano i benefici più significativi – erano sfruttate solo in misura limitata.

Ancora più significativo è stato comprendere cosa stesse realmente frenando l'adozione. Le preoccupazioni legate alla sicurezza e alla conformità normativa rappresentavano l'ostacolo più citato, trasversalmente a ruoli e funzioni. Gli ingegneri che non avevano chiarezza su quali dati potessero essere condivisi con gli strumenti di AI tendevano a evitarne completamente alcuni oppure a prendere decisioni autonome senza disporre di linee guida comuni. Il problema non era la mancanza di attenzione o responsabilità, ma l'assenza di standard condivisi, indicazioni operative e della fiducia necessaria per utilizzare l'AI in modo corretto e conforme. In questo contesto, introdurre ulteriore tecnologia non sarebbe stato sufficiente a risolvere la situazione.

Cosa è cambiato

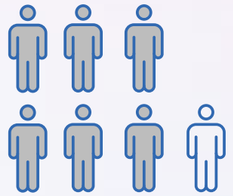
L'obiettivo del programma era costruire ciò che fino a quel momento era mancato ai team ingegneristici: le basi necessarie per un utilizzo dell'intelligenza artificiale efficace, responsabile e sostenibile nel tempo. Per raggiungerlo, sono stati organizzati workshop basati su attività reali di sviluppo software e quality assurance, affiancati da sessioni di coaching direttamente integrate nei progetti, nei repository di codice e nei processi di testing. A supporto di questo percorso è stato inoltre creato un insieme di strumenti e linee guida di governance – guide operative, framework decisionali, librerie di prompt e playbook di best practice – progettati fin dall'inizio per continuare a generare valore anche dopo la conclusione del programma.

Per evitare che standard e competenze si disperdessero una volta terminata la collaborazione, è stata inoltre creata una rete interna di ambassador, composta dai professionisti più esperti, con il compito di presidiare l'evoluzione delle best practice, promuoverne l'adozione e affiancare i colleghi durante le attività di coaching.

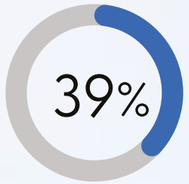
Il risultato più significativo, soprattutto in vista della scadenza del 2 agosto, riguarda però la governance. Al termine del programma, i professionisti senior erano in grado di identificare e applicare controlli specifici, come la corretta configurazione delle licenze aziendali, l'utilizzo di frammenti di codice anonimizzati per i contenuti sensibili e la divulgazione sistematica dell'uso dell'AI nelle pull request. Un team è arrivato persino a individuare un problema in uno strumento dotato di capacità di scrittura, che stava sovrascrivendo il contenuto dei ticket prima che la situazione si trasformasse in un incidente reale. Si tratta di un esempio concreto di consapevolezza operativa dei rischi: una competenza che non si acquisisce leggendo una policy, ma che richiede tempo, pratica ed esperienza sul campo.

"Leggi ogni riga con attenzione. L'AI può produrre codice che sembra corretto, ma che in realtà testa qualcosa di completamente diverso da ciò che dovrebbe. Se un test non fallisce quando gli fornisci un input non valido, significa che non sta realmente testando nulla." QA Engineer

I risultati ottenuti hanno superato ampiamente gli obiettivi iniziali:



6 su 7 sviluppatori affermano che gli strumenti di AI li aiutano a risparmiare tempo reale nelle attività quotidiane



39% riduzione del tempo necessario per arrivare al primo commit di codice, rispetto a un obiettivo iniziale del 15-20%

2-8 ore risparmiate per ogni test automatizzato creato

10-30 minuti risparmiati per ciascun bug report

Dall’adozione all’autonomia

La dimostrazione più concreta che queste competenze sono state effettivamente trasferite è arrivata dopo la conclusione del programma. Il team QA, il gruppo che all’inizio del percorso aveva espresso la valutazione più prudente in merito alle proprie competenze sull’AI, ha successivamente sviluppato due agenti progettati per affrontare specifiche criticità dei propri processi operativi. Il primo ha ridotto il tempo necessario per le attività manuali di verifica della conformità in materia di accessibilità da un massimo di 38 ore per pagina a circa 30 minuti. Il secondo ha abbattuto i tempi di revisione del design per singolo componente, passando da fino a tre ore a meno di dieci minuti. L’aspetto più significativo è che nessuna di queste soluzioni è stata sviluppata da Insight AI. Sono state progettate e realizzate direttamente dagli ingegneri dell’organizzazione, grazie a competenze e capacità che oggi appartengono stabilmente al patrimonio interno dell’azienda.

Vuoi sviluppare questo livello di competenza AI anche all’interno della tua organizzazione?

I nostri specialisti possono illustrarti l’approccio adottato in questo programma, mostrando come il trasferimento delle competenze, la governance e l’applicazione pratica siano stati progettati per generare valore ben oltre la fase di implementazione iniziale.

[Parla con uno specialista AI]

Progettare con l’obiettivo finale in mente

Questo programma dimostra concretamente cosa significa sviluppare competenze durature a livello organizzativo: mettere le persone che gestiscono un sistema di AI nelle condizioni di utilizzarlo in modo efficace, governarlo correttamente e farlo evolvere nel tempo. Lo stesso principio si applica anche all’ingegnerizzazione delle soluzioni: la vera domanda non è solo se un sistema sia stato costruito per essere lanciato, ma se sia stato progettato per essere gestito, adattato e governato nel lungo periodo.

Abbiamo messo in pratica questi principi nello sviluppo di AURA, una piattaforma basata sull’intelligenza artificiale nata per risolvere una sfida concreta all’interno della nostra organizzazione. La trasformazione del lavoro svolto nei progetti in case study destinati ai clienti era infatti un processo lento e fortemente manuale, che spesso non consentiva di disporre dei contenuti necessari quando emergeva una nuova opportunità commerciale. Grazie ad AURA, i team di vendita possono accedere rapidamente a casi di successo pertinenti durante le trattative e creare contenuti e narrazioni in più lingue capaci di valorizzare le competenze dell’azienda, contribuendo direttamente ai risultati commerciali. Lo sviluppo della piattaforma ha rappresentato anche un’opportunità per applicare ai nostri sistemi gli stessi principi ingegneristici che adottiamo nei progetti realizzati per i clienti.

Il principio alla base di queste scelte è semplice: gli elementi che determinano l’affidabilità di un sistema nel tempo raramente emergono durante una dimostrazione. Diventano invece evidenti mesi dopo, quando i dati trattati diventano più sensibili, l’utilizzo cresce e il contesto normativo si fa più rigoroso. Per questo motivo, governance e conformità non sono state considerate aspetti da aggiungere successivamente, ma componenti integrate fin dalle prime fasi del progetto. Abbiamo progettato la piattaforma affinché i dati siano anonimizzati per impostazione predefinita e gli accessi regolati in base ai ruoli. Le modifiche all’infrastruttura sono completamente tracciabili e replicabili, anziché gestite manualmente e difficili da verificare. Inoltre, il monitoraggio operativo è stato integrato nativamente nella soluzione, invece di essere introdotto in un secondo momento.

Nessuna di queste scelte modifica il funzionamento della piattaforma nel giorno del rilascio, ma sono proprio questi elementi a determinare se AURA possa essere gestita in modo responsabile su larga scala, adattata all’evoluzione delle esigenze aziendali e governata efficacemente in un contesto normativo sempre più stringente, comprese le disposizioni che l’EU AI Act considera oggi obblighi continuativi e non interventi una tantum.

AURA è tuttora operativa, continua a evolversi ed è governata dagli stessi principi progettuali definiti fin dall’inizio. Nel tempo, la piattaforma è stata ampliata per creare narrazioni tematiche o verticali di settore a partire da gruppi di case study correlati: un’evoluzione che non sarebbe stata possibile senza un’architettura progettata fin dall’origine per crescere e adattarsi nel tempo.

Cosa significa essere conformi da agosto 2026

Gli obblighi previsti dall'EU AI Act per i sistemi ad alto rischio non possono essere soddisfatti con una semplice approvazione iniziale. Richiedono un livello di conformità continuo, basato su documentazione costantemente aggiornata, monitoraggio permanente, supervisione umana delle decisioni automatizzate e chiari meccanismi di responsabilità che possano essere dimostrati concretamente e non solo dichiarati. Per le organizzazioni le cui collaborazioni in ambito AI si concludono con la fase di implementazione, soddisfare questi requisiti rappresenta una sfida operativa significativa, per la quale spesso non sono adeguatamente preparate.

Il programma di sviluppo delle competenze AI realizzato presso il fornitore globale di servizi educativi mostra concretamente cosa significhi essere pronti dal punto di vista normativo. Lo dimostrano il 90% dei professionisti senior in grado di identificare e applicare controlli specifici nel proprio contesto operativo, il team che ha individuato un rischio di governance prima che si trasformasse in un incidente in produzione e l'insieme di standard, strumenti di governance e figure di riferimento interne che continuano a guidare l'utilizzo dell'AI mentre la tecnologia evolve.

Questa solidità operativa sta diventando sempre più anche un vantaggio competitivo. Lo studio **Digital Sovereignty Trilemma** ha rilevato che il 55% delle organizzazioni considera la complessità normativa – tra GDPR, DORA ed EU AI Act – una delle principali sfide strategiche. Quando la maggior parte delle aziende incontra difficoltà in un determinato ambito, chi riesce a gestirlo efficacemente ottiene un vantaggio concreto: oggi il 60% dei clienti richiede evidenze di conformità durante i processi di procurement e il 43% delle organizzazioni ha già utilizzato le proprie credenziali in materia di sovranità digitale e conformità per acquisire o mantenere opportunità di business. Le organizzazioni in grado di dimostrare una conformità continua, e non semplicemente il completamento di una fase progettuale, saranno quelle meglio posizionate man mano che gli obblighi normativi diventeranno più stringenti.

Guardare oltre la fase di implementazione

Sia il programma di sviluppo delle competenze AI realizzato presso il provider educativo sia il progetto AURA mostrano, da prospettive diverse, cosa significhi costruire pensando al lungo termine. Nel primo caso, l'obiettivo era sviluppare competenze interne, definire standard, rafforzare la governance e creare la fiducia necessaria affinché l'organizzazione potesse continuare a evolvere autonomamente anche dopo la conclusione del coinvolgimento di Insight AI. Nel caso di AURA, invece, la sostenibilità nel lungo periodo è stata integrata fin dalle prime fasi di progettazione. Le decisioni che determinano la capacità di una piattaforma di operare, evolvere e rimanere conforme nel tempo sono state prese durante lo sviluppo, anziché essere rimandate a una fase successiva. In entrambi i casi, la domanda di fondo è la stessa: chi sta pensando a ciò che accadrà dopo il rilascio, ancora prima che il progetto abbia inizio?

Per la maggior parte delle organizzazioni, la risposta a questa domanda viene definita molto prima che un sistema entri in produzione. Dipende dalla capacità del partner tecnologico di considerare fin dall'inizio gli aspetti operativi: se la strategia tiene conto delle reali esigenze ingegneristiche, se l'architettura è progettata per essere gestita e fatta evolvere nel tempo e se la responsabilità sul funzionamento continuo della soluzione è parte integrante del progetto, anziché un tema da affrontare in un secondo momento.

È proprio questa continuità di responsabilità – dalla strategia all'ingegnerizzazione, fino alla gestione operativa del sistema – a rappresentare il vero significato di un approccio end-to-end.

Vuoi scoprire come i principi di governance e ingegnerizzazione dell'AI prendono forma in un sistema reale?

I nostri specialisti possono mostrarti come AURA è stata progettata per garantire operatività, scalabilità e governance nel lungo periodo, attraverso un approccio basato su compliance by design, monitoraggio continuo e scelte architetturali pensate per supportarne l'evoluzione nel tempo.
[Richiedi una demo di AURA]

Leggi **"Come scalare l'AI: dalla fase pilota alla produzione"** per approfondire l'altra faccia della sfida: trasformare l'AI da progetto pilota a soluzione pienamente operativa e comprendere quali competenze, processi e scelte tecnologiche siano necessari per riuscirci.

it.insight.com/ai