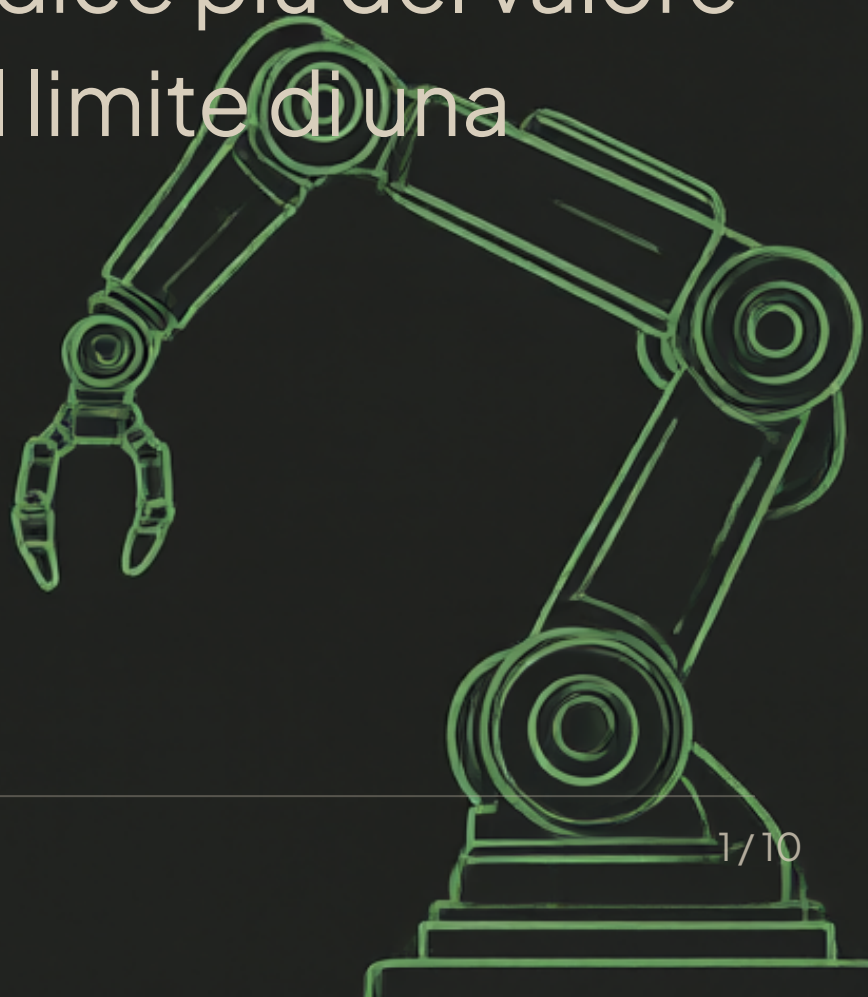
TECNICAI

Il miglior modello di IA al mondo ha passato l'10,8 per cento delle prove da professionista.

Un benchmark indipendente ha misurato l'IA su pratiche vere. Il risultato dice più del valore di un professionista che del limite di una macchina.



27 ottobre 2026. Esce DAYJOB: Tax.

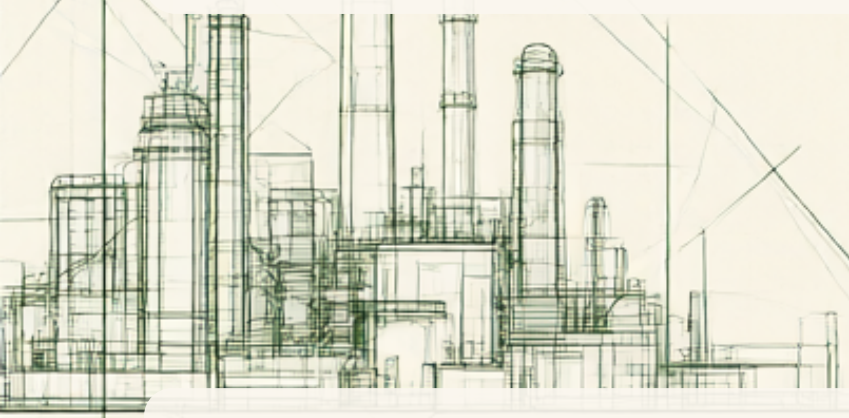
Lo pubblica Surge AI, fornitore indipendente di dati per l'addestramento.

Cinquanta richieste reali di contribuenti.

1.789 file di archivio del cliente.

Un solo compito: lavorare la pratica come farebbe un consulente.

Fonte: surgehq.ai, benchmark DAYJOB: Tax, 27/10/2026



Il modello riceve la richiesta e tutti i documenti del cliente.

Deve cercare, calcolare, consegnare.

La valutazione non si ferma al calcolo.

Controlla anche se l'IA ha notato i problemi che il cliente non ha mai sollevato.

È la voce del test che si chiama: «identifying issues the client never raised».



Primo in classifica: 10,8 per cento.

Secondo: 9,6 per cento. Terzo: 9,2 per cento.

Soglia di superamento: 95 per cento dei criteri.

Tradotto: su una pratica vera, oggi, l'intervento umano serve ancora. Non è un'abitudine di mercato: è il numero.

Fonte: surgehq.ai, pass rate medio dei tentativi valutati, 27/10/2026

UNA PROVA TRA LE CINQUANTA

Anna vende casa. Due date, entrambe valide.

Il caso è americano: un'agevolazione sulle plusvalenze di chi vende l'abitazione.

L'IA controlla le due date: corrette entrambe.

Poi c'è un requisito di residenza che scade ad agosto 2027. Se la vendita ritarda, l'agevolazione si perde.

Lo nota un modello su due. Astra non lo nomina mai, in cinque tentativi. Claude Opus lo indica sempre.

> Agents haven't yet passed the audit.

Frase testuale del comunicato Surge AI, 27/10/2026 (traduzione: gli agenti non hanno ancora superato la revisione). I nomi dei modelli sono quelli del comunicato.

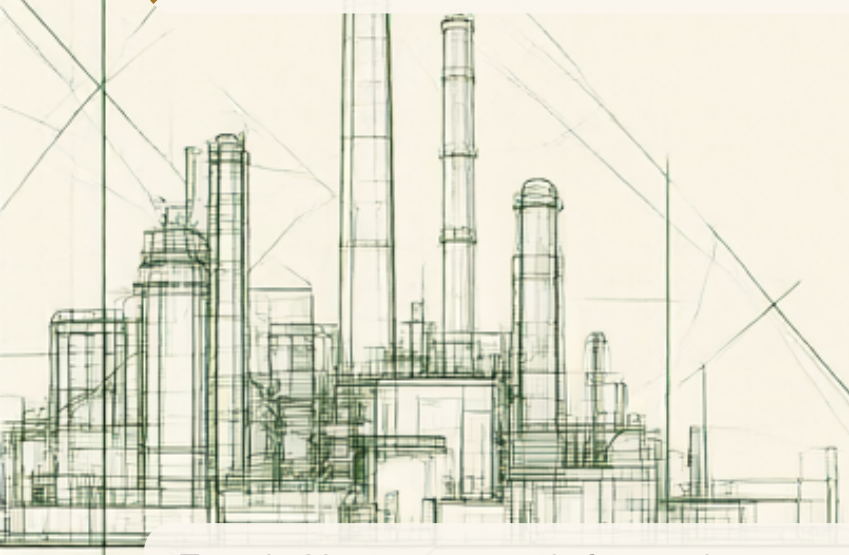
Fonte: surgehq.ai, benchmark DAYJOB: Tax, 27/10/2026

L'IA non sbaglia il calcolo. Sbaglia l'accorgersi.

Esegue quello che le chiedi.

Non ti segnala il rischio che non ti è venuto in mente di nominare.

> In pratica,
> il cliente paga il professionista per le domande che non sa di dover fare. Il benchmark non ha misurato la velocità della macchina: ha misurato esattamente quella distanza.

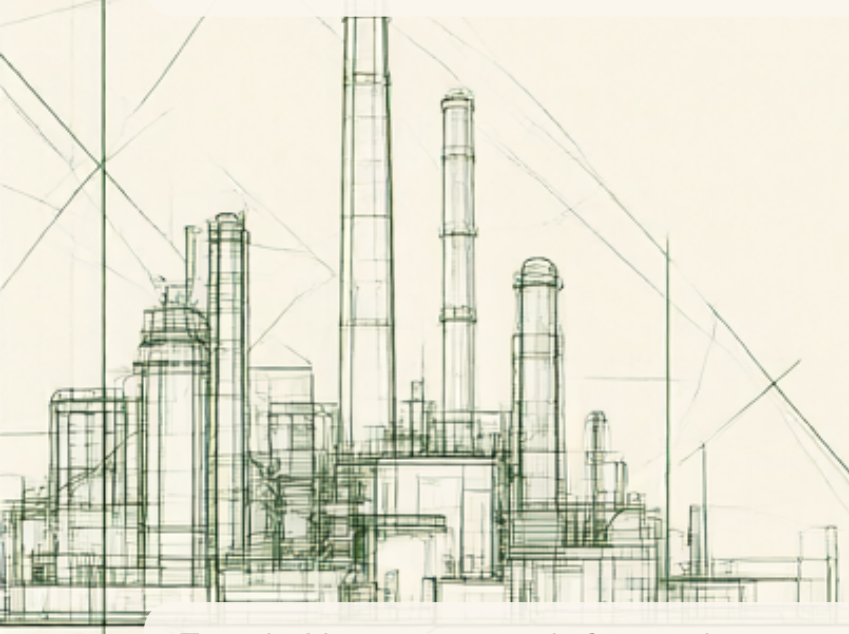


L'IA resta utile. Risponde, cerca, calcola, scrive.

È un assistente veloce e istruito.

È un assistente che non alza la mano.

La firma, la responsabilità e il giudizio restano dove stanno oggi: sulla persona che ha studiato e sul professionista che controlla.



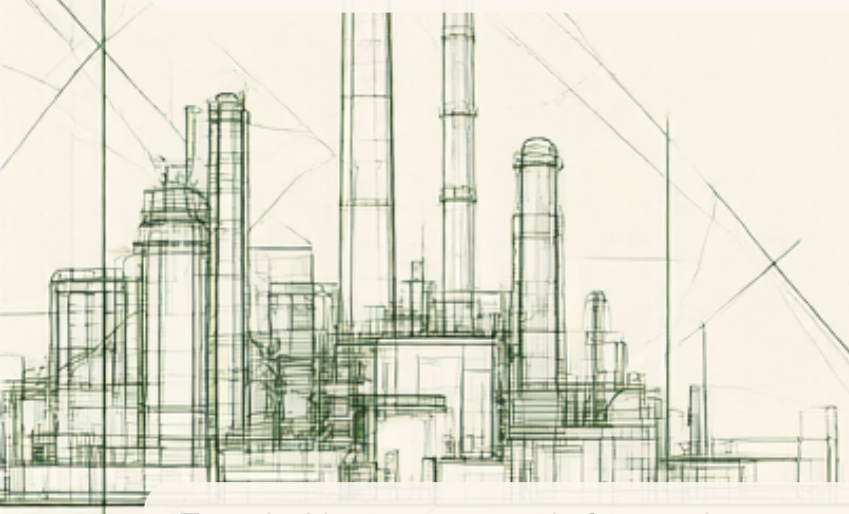
La stessa identica cosa succede con l'IA in ufficio.

Le fai scrivere una relazione, analizzare una fattura, riassumere un documento di sicurezza.

Ricevi quello che hai chiesto.

Non ricevi quello a cui non avevi pensato.

Chi usa l'IA senza un metodo non sta delegando il lavoro. Sta delegando il controllo.



Tre domande da fare alla tua IA.

1. Mi dice cosa manca, o solo quello che le chiedo?

2. Chi verifica la sua risposta prima che diventi un documento?

3. Qual è il compito che sto delegando senza accorgermi che chiede giudizio?

La terza è quella che fa più paura. Ed è la più utile.



MI PRESENTO:

La macchina ricorda e calcola.
La persona giudica.

Dubbi sull'IA che usi ogni giorno al lavoro?
Scrivici: info@tecnicai.it

Iscritto all'Ordine dei Periti Industriali di Padova · RSPP · Professionista antincendio n° PD01774P00344 · Docente dal 2012 · Fondatore TecnicAI · Co-fondatore Omeda Studios.

Contenuto informativo, non consulenza legale. Altri documenti gratuiti su tecnicai.it/news

