

# Quattro fattori da considerare per scegliere il giusto modello di IA

Di fronte all'evoluzione rapida e continua delle tecnologie, le aziende sono chiamate oggi a massimizzare l'utilizzo di intelligenza artificiale (IA) e intelligenza artificiale generativa (IA gen) per generare valore. La scelta del modello di IA è un passaggio cruciale perché solo con il giusto modello sarà possibile ottimizzare le iniziative di IA e raggiungere gli obiettivi prefissati. Questa checklist prende in esame quattro aspetti chiave da valutare per la scelta del modello di IA e illustra come [Red Hat® Enterprise Linux® AI](#) offre modelli per lo sviluppo di soluzioni adatte agli scenari di utilizzo aziendali.

## 1 La dimensione del modello

Scegliere un modello con una dimensione congrua alle esigenze dell'organizzazione e ai suoi specifici scenari di utilizzo è un punto di partenza fondamentale poiché ha un impatto diretto sui costi di gestione del modello, sulle competenze necessarie per gestirlo e sull'accuratezza dei risultati.

I modelli più ampi dispongono di più parametri e garantiscono quindi risultati ottimali anche in scenari complessi. Questi modelli (e i dati che li alimentano) richiedono però maggiori risorse di elaborazione e infrastrutture complesse, il che può incidere significativamente sui costi operativi e sulle competenze necessarie per gestirli.

Inoltre, quando si tratta di ottimizzare questi modelli con i dati pertinenti per l'azienda, per scenari di utilizzo specifici o per un determinato settore, tali costi e competenze sono destinati ad aumentare esponenzialmente.

Per un approccio più efficiente conviene utilizzare modelli più piccoli e specializzati, come i [modelli linguistici di grandi dimensioni \(LLM\) della famiglia Granite](#) inclusi in Red Hat Enterprise Linux AI. L'addestramento di questi modelli sui dati aziendali richiede meno tempo, minori competenze specialistiche ed è più economico.

## 2 La trasparenza del modello

Un altro fattore chiave che influisce sull'accuratezza e sull'efficienza dei modelli di IA gen è la trasparenza sui dati utilizzati per l'addestramento, in particolare su aspetti quali origine, proprietà, qualità, accessibilità, affidabilità, pertinenza, diversità e volume.

Utilizzare un modello poco trasparente mette l'organizzazione a rischio. Un rischio che va dalla generazione di risultati errati dovuti alla presenza di dati inesatti in fase di addestramento a possibili conseguenze legali derivanti dall'utilizzo di modelli addestrati su materiale protetto da copyright.

Red Hat consiglia di adottare modelli con sorgenti di dati e dataset trasparenti, come i modelli Granite inclusi in Red Hat Enterprise Linux AI che offrono dataset per l'addestramento completi, trasparenza sui contenuti del modello, oltre a supporto e tutela di Red Hat.

In questo modo le organizzazioni possono portare avanti le proprie iniziative di IA gen in tutta sicurezza con modelli che godono del supporto di un'azienda affidabile, sono stati sviluppati e addestrati eticamente e nel rispetto di tutte le normative e tutelano i dati sensibili.

### 3 Modello open source o proprietario?

Esistono diversi tipi di licenza per gli LLM, tra cui licenze open source (ad esempio Apache 2.0, GPL, ecc.) e licenze commerciali che solitamente sono proprietarie o basate su sottoscrizione.

La scelta di un modello open source o proprietario può avere un impatto significativo su diversi aspetti:

- ▶ Costi ed efficienza operativa.
- ▶ Livelli di personalizzazione possibili.
- ▶ Proprietà delle applicazioni e dei servizi sviluppati utilizzando tale modello e dei dati forniti al modello.

I modelli proprietari sono più costosi, difficili da addestrare/ottimizzare per i propri scenari di utilizzo o per il proprio settore, pongono limiti alla personalizzazione e sono in generale meno efficienti dei modelli open source più piccoli, che beneficiano dell'adozione e del contributo della community. Inoltre, per vendere i servizi e le applicazioni sviluppate utilizzando un modello con licenza commerciale di norma occorre pagare un'ulteriore costo al proprietario del modello, che spesso acquisisce anche la proprietà di tutti i dati forniti al modello.

Gli LLM Granite inclusi in Red Hat Enterprise Linux AI assicurano l'efficienza operativa e finanziaria, gli utenti possono utilizzare, modificare e distribuire i modelli liberamente e mantengono la proprietà completa dei propri dati.

### 4 L'addestramento del modello

La personalizzazione è essenziale perché i modelli soddisfino le esigenze degli scenari di utilizzo aziendali con efficienza, precisione e a costi ridotti. Le diverse tecniche di ottimizzazione dei modelli garantiscono risultati più o meno efficienti ed efficaci.

Red Hat Enterprise Linux AI propone strumenti di allineamento dei modelli accessibili tramite la distribuzione supportata e basata sul ciclo di vita di [InstructLab](#).

A differenza delle tecniche di addestramento degli LLM tradizionali, il metodo InstructLab è in grado di semplificare lo sviluppo dei modelli e garantirne la scalabilità. Il metodo prevede un processo di generazione di dati sintetici guidato dalla tassonomia, un framework di ottimizzazione multifase per migliorare le prestazioni del modello e la presenza di una community open source in cui gli sviluppatori possono contribuire liberamente ai modelli con nuove competenze e conoscenze.

Gli strumenti di allineamento dei modelli InstructLab:

- ▶ Offrono una soluzione a costi contenuti per migliorare l'allineamento degli LLM.
- ▶ Promuovono il miglioramento dei modelli open source con un approccio basato sulla community.
- ▶ Ampliano la base di utenti che può accedere alla personalizzazione dei modelli fornendo strumenti, addestramento e conoscenze accessibili.
- ▶ Permettono agli utenti di mantenere la completa proprietà dei dataset per l'addestramento e la proprietà intellettuale sui modelli.

#### Scopri di più

[Leggi l'ebook](#) per scoprire in che modo Red Hat aiuta le aziende a introdurre l'IA.

#### Inizia subito

[Contatta Red Hat](#) per scoprire in che modo può aiutare la tua organizzazione a intraprendere un percorso di adozione dell'IA.



#### Informazioni su Red Hat

Red Hat consente la standardizzazione in diversi ambienti e lo sviluppo di applicazioni cloud native, oltre a favorire l'automazione, la protezione e la gestione di ambienti complessi grazie a [pluripremiati](#) servizi di consulenza, formazione e supporto.

**f** [facebook.com/RedHatItaly](#)  
**X** [twitter.com/RedHatItaly](#)  
**in** [linkedin.com/company/red-hat](#)

**ITALIA**  
[it.redhat.com](#)  
[italy@redhat.com](mailto:italy@redhat.com)

**EUROPA, MEDIO ORIENTE,  
E AFRICA (EMEA)**  
00800 7334 2835  
[it.redhat.com](#)  
[europa@redhat.com](mailto:europa@redhat.com)