

Sei pronto per un LLM **privato** in azienda?

15 domande strutturate per capire se l'AI on-premise è la scelta giusta per la tua azienda — e da dove partire. Compilabile in 10 minuti, senza chiacchiere di marketing.

Perché questa checklist

Nel 2026 l'AI generativa è ovunque. Ma la stragrande maggioranza dei progetti aziendali AI arriva a fine anno con budget bruciati e zero valore in produzione. Il motivo è quasi sempre lo stesso: si è partiti dalla tecnologia, non dai processi.

Questa checklist ribalta l'approccio. In 15 domande strutturate, capisci se hai davvero le condizioni per portare un LLM on-premise in azienda — dai dati che tratti, ai casi d'uso realistici, alla infrastruttura minima, fino al team che dovrà gestirlo.

Alla fine ottieni un punteggio grezzo e un percorso consigliato. Non è consulenza, è auto-diagnostica.

Come usarla. Rispondi onestamente Sì o No a ogni checkpoint. Ogni Sì vale 1 punto. Alla fine sommi i punti e leggi il verdetto nella pagina finale.

Le 6 aree coperte

1. Dati e sensibilità

Che tipo di informazioni tratti e quanto sono strategiche

2. Casi d'uso realistici

Dove l'AI porta valore concreto nei tuoi processi

3. Infrastruttura minima

Cosa serve davvero in termini di hardware e rete

4. Sicurezza e compliance

GDPR, AI Act, NIS2, tracciabilità

5. Team e competenze

Chi si occupa di cosa, quando serve un partner

6. Budget e timeline

Range realistici per pilota e deploy completo

1. Dati e sensibilità

Se i tuoi dati sono critici, l'AI on-premise passa da "opzione" a "requisito". Se sono comuni, il cloud SaaS resta la scelta più economica.

- ☐ **Trattate dati coperti da NDA cliente** (progetti riservati, informazioni industriali, contratti sensibili).
- ☐ **Il vostro settore è regolamentato** (sanità, finanza, difesa, PA, dati particolari GDPR).
- ☐ **Avete brevetti, know-how tecnico o formule proprietarie** nei documenti aziendali.
- ☐ **Volete evitare che i dati passino per server USA** (Schrems II, sovranità del dato).
- ☐ **Avete già dovuto rispondere a domande interne o audit** sul "dove finiscono i nostri dati" quando qualcuno ha proposto ChatGPT/Copilot.

2. Casi d'uso realistici

L'AI on-premise non serve per "fare qualcosa con l'AI". Serve quando c'è almeno un processo ricorrente che oggi consuma tempo. Se non trovi 2-3 casi concreti, il pilota non parte.

- ☐ **Il vostro team dedica ore ogni settimana** a leggere/estrarre informazioni da PDF (contratti, fatture, DDT, capitolati).
- ☐ **Ricevete molti ticket/richieste ripetitive** (help desk interno, assistenza clienti) che potrebbero essere pre-filtrati.
- ☐ **Producete offerte, preventivi o proposte** partendo da template + personalizzazioni ricorrenti.
- ☐ **Avete una knowledge base interna** (wiki, SharePoint, documenti di processo) che i colleghi consultano spesso.
- ☐ **Fate riunioni settimanali** con verbali che qualcuno deve poi trascrivere e distribuire.

3. Infrastruttura minima

Un LLM aziendale non richiede un datacenter. Ma serve un livello base di infrastruttura IT. Se manca il minimo, si parte da lì.

- ☐ **Avete una rete aziendale strutturata** con firewall, VPN, segmentazione dei server dai client.
- ☐ **Fate backup automatici periodici** con verifica di restore (non solo backup "che gira da solo").
- ☐ **Avete uno spazio server** dove installare hardware nuovo (server room, armadio rack, cloud privato dedicato).
- ☐ **Avete un partner IT di riferimento** che gestisce l'infrastruttura (interno o esterno).

Range hardware realistici

CONFIGURAZIONE	COSTO TIPICO	MODELLI SUPPORTATI
Workstation GPU consumer (RTX 4090)	4-8k €	LLM 7-13B (Llama 3, Mistral, Qwen)
Server GPU professionale (RTX 6000 Ada)	18-25k €	LLM 30-70B, uso team ristretto
Server A100/H100 dedicato	60-120k €	LLM 70B+ con più utenti concorrenti
Cloud privato europeo (Hetzner, OVH)	200-800 €/mese	Alternativa scalabile senza CAPEX iniziale

4. Sicurezza e compliance

L'AI on-premise elimina i problemi di sovranità del dato, ma la responsabilità della sicurezza rimane in azienda. Servono presidi minimi.

- ☐ **Avete un registro dei trattamenti** aggiornato ex art. 30 GDPR.
- ☐ **Avete definito ruoli e responsabilità privacy** (Titolare, eventuale DPO o Referente Privacy interno).
- ☐ **Avete una procedura scritta** per la gestione di data breach ex art. 33 GDPR (notifica al Garante entro 72h).
- ☐ **Avete letto l'AI Act europeo** e sapete se i vostri casi d'uso rientrano in categoria "rischio limitato" o "alto rischio".
- ☐ **Se in settore regolamentato:** avete verificato con il compliance officer che l'AI on-premise sia compatibile con le vostre policy specifiche.

5. Team e competenze

Un LLM aziendale ha bisogno di due profili minimi: chi lo installa/mantiene e chi decide i casi d'uso. Se mancano entrambi, serve un partner.

- ☐ **Avete un IT manager (interno o consulente stabile)** con esperienza su server Linux e reti aziendali.
- ☐ **Avete uno sponsor a livello direzione** che spinga il progetto e sblocchi decisioni.
- ☐ **Avete identificato un "product owner" del progetto AI** (chi decide priorità dei casi d'uso, chi valida i risultati).
- ☐ **I dipendenti coinvolti nei casi d'uso sono disponibili a testare** e dare feedback nelle prime settimane di uso.

6. Budget e timeline

L'AI on-premise ha un CAPEX iniziale ma costi ricorrenti bassi. Il ROI arriva nell'arco di 12-24 mesi se ci sono casi d'uso reali.

Range investimento tipici (2026)

SCENARIO	INVESTIMENTO INIZIALE	MANUTENZIONE ANNUA
Pilota single use case (6-10 sett.)	15-30k €	3-6k €
Deploy 3-4 use case integrati	40-80k €	8-15k €
Enterprise fine-tuning profondo	80-200k €	15-30k €

Timeline realistica

FASE	DURATA
Analisi + scelta modello + setup infrastruttura	3-4 settimane
Pilota single use case in produzione	6-10 settimane totali
Estensione a 3-4 use case integrati	3-5 mesi totali
Fine-tuning e ottimizzazione continua	Iterazioni 6-12 mesi

Il tuo punteggio

Somma i checkpoint a cui hai risposto "Sì". Massimo 22 punti.

Non ancora pronto.

0 - 8

Prima serve consolidare processi, dati e infrastruttura. Un pilota AI oggi finirebbe in un cassetto. Parti da un audit interno.

Parti da un pilota mirato.

9 - 14

Hai il minimo per iniziare. Scegli UN caso d'uso ad alto valore, misura risultati per 8-12 settimane, poi decidi estensione.

Pronto per un deploy full.

15 - 22

Hai maturità tecnica, casi d'uso multipli, team e budget. Puoi partire con un progetto integrato su 3-4 use case in parallelo.

Cosa fare adesso

Il punteggio è un punto di partenza, non una sentenza. Il modo migliore per validarlo è confrontarlo con qualcuno che abbia visto molti casi simili al tuo.

Discovery gratuita di 30 minuti

Un senior tecnico di Hexagon ti aiuta a interpretare il punteggio nel tuo contesto specifico.
Nessuna vendita, nessun impegno.

[Prenota su hexagonsistemi.it](https://hexagonsistemi.it)

Chi siamo

Hexagon snc è una software house italiana con sede a Solbiate Olona (VA). Dal 1995 progettiamo gestionali su misura, modernizziamo software esistenti, integriamo AI on-premise nei processi aziendali. Serviamo aziende in tutta Italia con presenza on-site nell'Alto Milanese e Varesotto.

CONTATTI

Telefono: 0331 375357

WhatsApp: +39 348 392 5267

Email: info@hexagonsistemi.it

Web: hexagonsistemi.it

SEDI

Legale:

Via Genova 6, 21058 Solbiate Olona (VA)

Operativa:

Piazza G. Marconi 7, 21058 Solbiate Olona (VA)