SCHEDA DEL MODULO

L’IA come specchio: le risposte dipendono da come chiedi

Modulo adattivo | ~20 min

Davanti all’intelligenza artificiale (IA) si oscilla tra fascino e timore – due atteggiamenti che le attribuiscono un’intelligenza che non ha. Comprenderla diventa decisivo per usarla. Con l’idea di «specchio» di Shannon Vallor, questo modulo mostra che l’IA non pensa: ci restituisce i nostri dati, i nostri bias e perfino il nostro stato interiore. Le sue falle – allucinazioni, compiacenza, confusione – parlano di noi. La sua lezione: la qualità di ciò che si ottiene dipende da ciò che si porta.

OBIETTIVI DIDATTICI

01 Comprendere che cos’è l’IA – un algoritmo che automatizza, non un cervello dotato di intenzione o di coscienza.

02 Vedere l’IA come uno specchio che restituisce i nostri dati, bias, confusione e stato interiore del momento.

03 Mantenere lo spirito critico davanti a tre riflessi: l’allucinazione, la compiacenza e la risposta generica predefinita.

04 Migliorare il dialogo: contesto duraturo, permesso di dissentire, ragionamento esplicitato, rilettura delle proprie richieste.

DESTINATARI

Chiunque usi l’IA ogni giorno e voglia affinare il modo in cui la interroga.

- **Prerequisito:** nessuno

SVOLGIMENTO DEL MODULO

In autonomia, con un allenamento adattato a chi apprende.

1. **Autovalutazione**: su 9 dimensioni, per fare il punto sulle proprie pratiche e orientare l’allenamento.
2. **Video dell’esperto**: di 10 minuti, con 3 pause di riflessione.
3. **Casi pratici**: 2 primi casi scelti in base all’autovalutazione, poi altri 16 facoltativi. Ogni risposta riceve un riscontro didattico: i suoi effetti e una pista per migliorare.
4. **Impegno con una data**: un’azione da provare, scritta dal partecipante o scelta tra le proposte.
5. **Scheda promemoria da scaricare**: punti chiave, domande, impegno, note personali e checklist.

UN ESEMPIO TRA I CASI PRATICI PROPOSTI

Un cliente scontento ha scritto a Kwame, addetto alla clientela in un’agenzia di autonoleggio, e alla sua collega Farah. Farah pensa che tocchi a Kwame rispondere; Kwame pensa che tocchi a Farah. Kwame racconta la sua versione all’assistente IA, che gli dà ragione. Che cosa deve fare Kwame?

VALUTAZIONE E COMPLETAMENTO

Soglia: 80%

Il punteggio tiene conto delle risposte ai casi pratici, dell’impegno formulato e, se previsto, delle attività di apprendimento complementari. Per completare il modulo servono un impegno e un punteggio di almeno 80%.

METODI & MODELLI

- L’IA come specchio – Shannon Vallor, The AI Mirror (2024)
- Compiacenza algoritmica (sycophancy) – concetto tecnico riconosciuto
- Dare all’IA il permesso di esprimere disaccordo – metodo Afthonios

FONTI & RIFERIMENTI

- Stanford – studio sulla compiacenza (sycophancy) delle IA (Science, 2026): i modelli danno ragione all’utente ~49% di volte in più rispetto a un essere umano
- Shannon Vallor – The AI Mirror (2024), professoressa di etica dell’IA all’Università di Edimburgo
- Vectara, «Hallucination Leaderboard» (aggiornamento maggio 2026): tasso di allucinazione nel riassunto di documenti dall’1,8% del modello migliore al 24,2% del meno affidabile
- Dhuliawala et al., Findings of ACL (2024) – verifica strutturata e affidabilità delle risposte dell’IA

Modulo creato da Afthonios.com.

Contenuto verificato il 15 ottobre 2026