

AI - EVAL

Evaluation, Observability e Guardrails

Senza Golden Dataset e release gate ogni release AI è un esperimento sui Suoi utenti.

DURATA

2 giorni

GRUPPO

fino a 12 persone

FORMATO

In-house · Remoto · Nel lab · Tedesco o inglese

QUESTO ESISTE DOPO

- ▶ Un Golden Dataset per il proprio sistema: casi curati con risultati attesi, casi limite e richieste volutamente difficili
- ▶ Metriche di qualità implementate per groundedness, rilevanza e completezza, incluso un LLM-as-Judge calibrato con controprova a campione
- ▶ Tracing sull'intera catena dalla richiesta attraverso retrieval e chiamata al modello fino alla risposta, con costi e latenza per ogni passo
- ▶ Un budget di costi e latenza per tipo di richiesta con soglie di allarme
- ▶ Release gate definiti: quali valori misurati un deployment deve superare e quando scatta un rollback

AGENDA

Giorno 1

La diagnosi del volo alla cieca · Costruire il Golden Dataset · Implementare le metriche: groundedness, rilevanza, completezza · Il primo run di valutazione · Risultato della giornata: una baseline solida

Giorno 2

Tracing: rendere visibile la catena · Budget di costi e latenza · Guardrail prima e dopo il modello · Release gate e riconoscimento delle regressioni · Risultato della giornata: il gate tiene



IL SUO TRAINER

Dino Bordonaro

Microsoft MVP per otto anni consecutivi (Microsoft Azure · Cloud and Datacenter Management), Microsoft Certified Trainer, quasi 30 anni di IT. Costruisce sistemi AI sovrani su hardware proprio (Azure Local, Foundry Local, GPU NVIDIA) e forma esattamente su questo. Oltre 800 professionisti formati, ogni concetto collaudato nel proprio enterprise lab in un datacenter certificato TÜV.

15.800 € forfettario, IVA esclusa

Il track mette il Suo team in condizione di costruire e gestire il proprio sistema di misura. Non è un audit del Suo sistema produttivo né un servizio di managed monitoring, per questo parliamo di un progetto.

AI - EVAL