



Giacomo Mantovani

Senior AI Engineer & sviluppatore full-stack

Italia · Da remoto · Italiano (madrelingua), inglese (fluente)

Progetto e realizzo **sistemi AI che girano in produzione**: agenti, RAG, AI per documenti e voce, con il backend e il cloud intorno. Quattro anni di progetti per **banche, assicurazioni, cliniche e aziende di prodotto**, da tech lead e referente tecnico del cliente. Entro nel tuo team, lavoro nel tuo codice e ne rispondo fino alla messa in produzione.

1M+

richieste l'anno smistate da un agente LLM, 81% di precisione contro 75% umano

-84%

di lavoro manuale su oltre 2.500 bilanci al giorno

2M+

prodotti ricercabili a parole, ogni articolo verificato sulla richiesta

-45%

di costo AI per articolo su un motore di contenuti automatico

PROGETTI SELEZIONATI

Agente di smistamento help desk

Grande banca italiana · tech lead, team di 4

Agente LLM su Azure OpenAI, RAG e LangChain che smista **oltre 1M di richieste l'anno in 400+ categorie**. Tempi di risoluzione da 5-6 giorni a 2-3.

Verifica documenti di successione

Banca · progettazione e backend, team di 3

Documenti legali confrontati con i dati della banca; i controlli non superati vanno a un operatore con il motivo. **Tempi di lavorazione -80%**.

Receptionist telefonica AI

IGEA, SaaS sanitario · CTO e co-founder

Agente vocale multi-tenant che prenota, sposta e smista le chiamate per le cliniche, conforme al GDPR, con bassa latenza e audio in cache per ridurre i costi.

Estrazione dati dai bilanci

Banca · architetto della pipeline, team di 5

Pipeline OCR + LLM con circa 30 controlli di validazione per documento, FastAPI e Kafka su OpenShift. **Oltre 2.500 documenti al giorno, -84% di lavoro manuale**.

Ricerca prodotti conversazionale

Estyl, AI per la moda · tech lead

Ricerca vettoriale ibrida, filtri strutturati e re-ranking su **oltre 2M di prodotti** su Google Cloud, con una pipeline di valutazione rigorosa, riconoscimento dell'intento e preferenze per utente.

Piattaforme AI per contenuti e finanza

Mokka Studios · senior AI engineer

Pipeline LLM autonoma dalla ricerca all'articolo pubblicato (**-45% di costo**); backend AWS e report finanziari AI per una piattaforma di investimento.

ESPERIENZA

2026 - oggi

Senior AI Engineer · Mokka Studios — AI e backend su prodotti in produzione

2026 - oggi

CTO e co-founder · IGEA — agente vocale AI per cliniche

Gen - Lug 2026

Tech Lead AI Engineer · Estyl — assistente moda conversazionale

2022 - 2025

AI Engineer · Sprint Reply — AI generativa per banche, assicurazioni e finanza

2015 - 2022

LM AI & Data Engineering, L Ingegneria Informatica · Università di Pisa

STACK

Python

FastAPI

TypeScript

LangChain / LangGraph

Azure OpenAI

Anthropic

OpenAI

RAG e DB vettoriali

Azure Document Intelligence

Vapi / LiveKit

Google Cloud

AWS

Azure

Cloudflare

Kafka

Docker

CI/CD

Claude Code / Codex

COME LAVORO

- Valutazione prima del lancio, su casi reali
- Tetti di spesa e costo per richiesta monitorato
- Dati in UE, GDPR, audit trail, log senza dati sensibili
- Consegna chiara: repo, cloud e brand restano tuoi