

**E4**

COMPUTER  
ENGINEERING

WHEN PERFORMANCE MATTERS

# E4 COMPUTER ENGINEERING: LA RIVOLUZIONE DELLA GEN AI IN AZIENDA

Emanuele Riva, *Enterprise AI Senior Sales*

**E4 Computer Engineering** progetta e realizza soluzioni ad altissimo contenuto tecnologico per Cluster HPC, Cloud, Data Analytics, Intelligenza Artificiale e HyperConverged Infrastructure per i mercati Academic ed Enterprise. Da anni collaboriamo con i principali centri di ricerca a livello nazionale e internazionale (Cineca, CERN, ECMWF, LEONARDO) e siamo coinvolti in progetti di livello nazionale ed europeo in ambito HPC e AI (EuroHPC JU EPI, EUPEX, Horizon Europe, KDT).

Ogni soluzione E4 è **UNICA**, come ognuno dei nostri clienti; **COLLAUDATA** in ogni singolo componente; **VALIDATA** per verificare le prestazioni effettive di ciascun sistema e **SERVITA** da tecnici che forniscono assistenza nelle più estese e complesse infrastrutture di calcolo italiane ed europee.

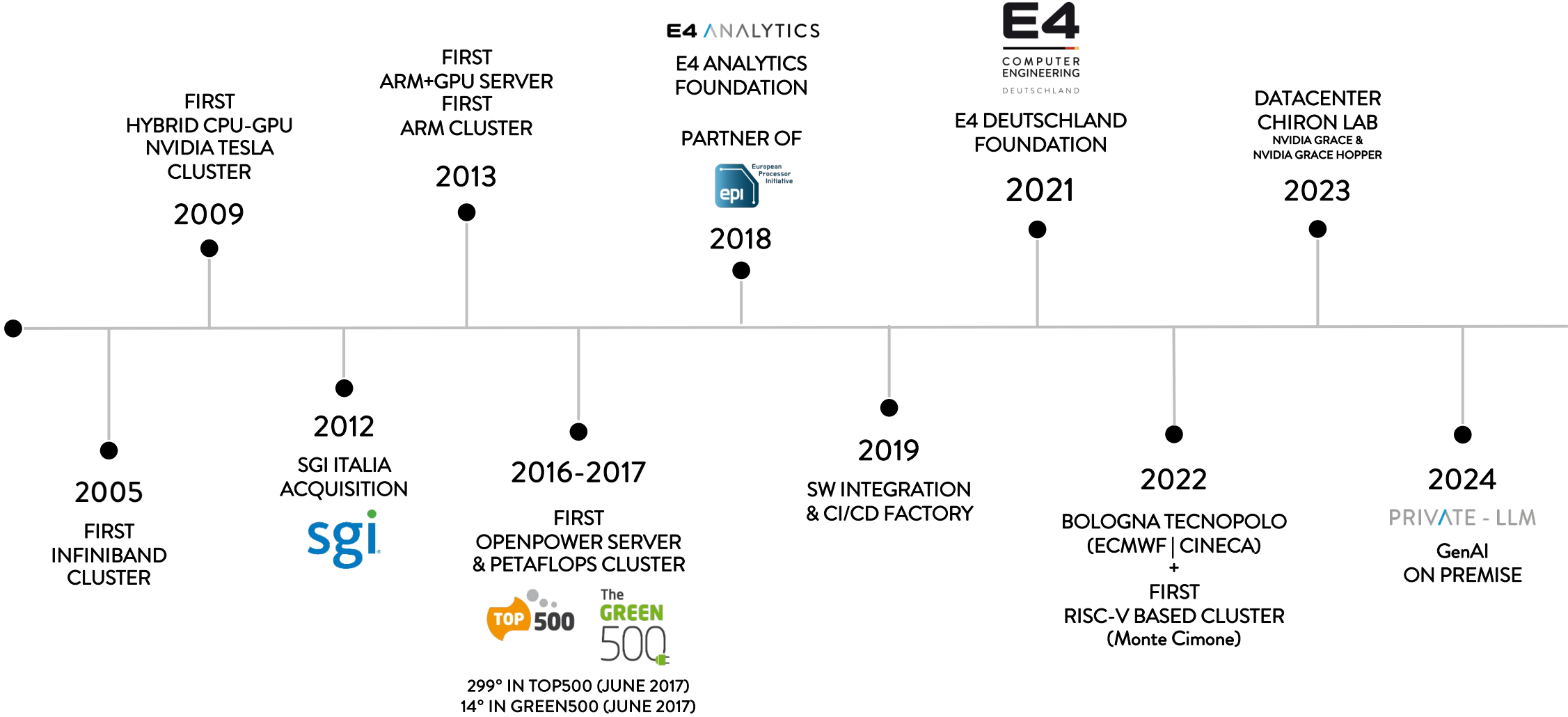
## **E4** ANALYTICS

LET YOUR DATA PAY YOUR GROWTH

Tramite la sister company **E4 Analytics**, E4 lavora per integrare la Data Science nelle organizzazioni che intraprendono la Trasformazione Digitale del loro business per migliorare prodotti/processi e ottimizzare risorse. Operiamo all'intersezione tra business e tecnologia, supportando il cliente nell'adozione di soluzioni AI: con E4 Analytics, i dati aziendali diventano una risorsa per produrre valore.

# E4 INNOVATION HISTORY

2002



## PROBABILMENTE... CI CONOSCETE GIA'...



## INFN &amp; CERN AWARD (2012-2014)

Premio di eccellenza nella collaborazione industriale per gli esperimenti ATLAS e CMS presso il Large Hadron Collider al CERN nell'ambito del Progetto della scoperta del bosone di Higgs

# CHAT GPT vs OPEN SOURCE LLM

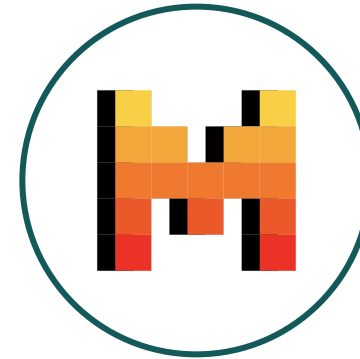


CHATGPT

VS



LLAMA3

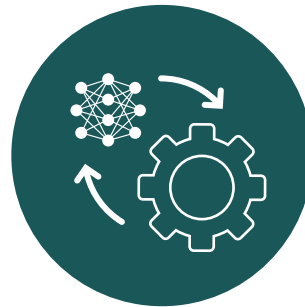


MISTRAL

# DOMAIN-SPECIFIC LLM



LLM



RAG / FINE TUNING



DEPLOY

# PRIVATE - LLM



LA TUA AI



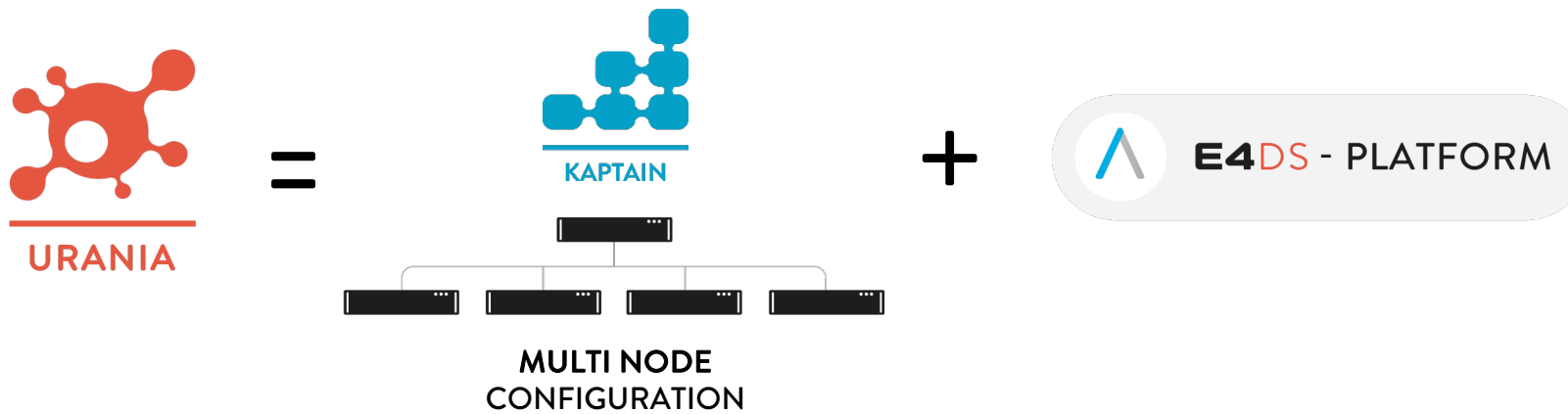
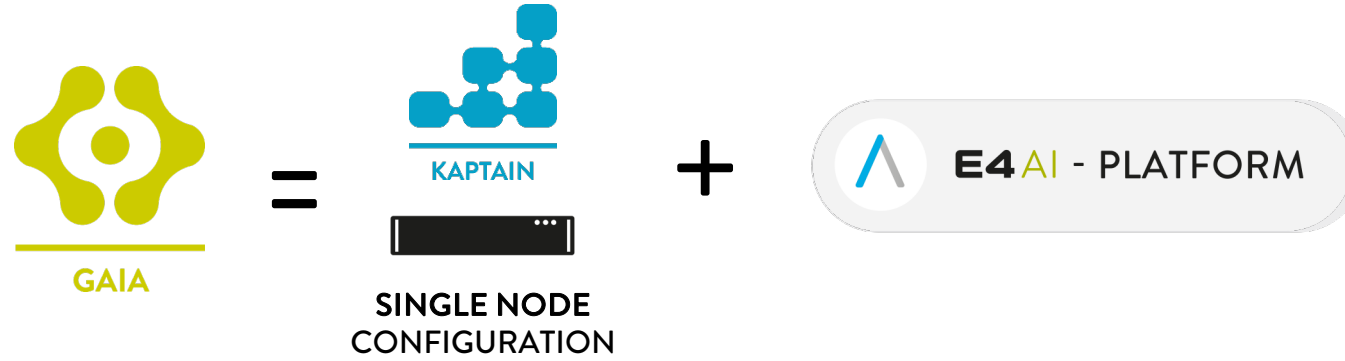
I TUOI DATI



LA TUA SICUREZZA

PORTA L'AI VERSO I TUOI DATI, NEL CUORE DELLA TUA AZIENDA

# PRIVATE LLM SOLUTIONS

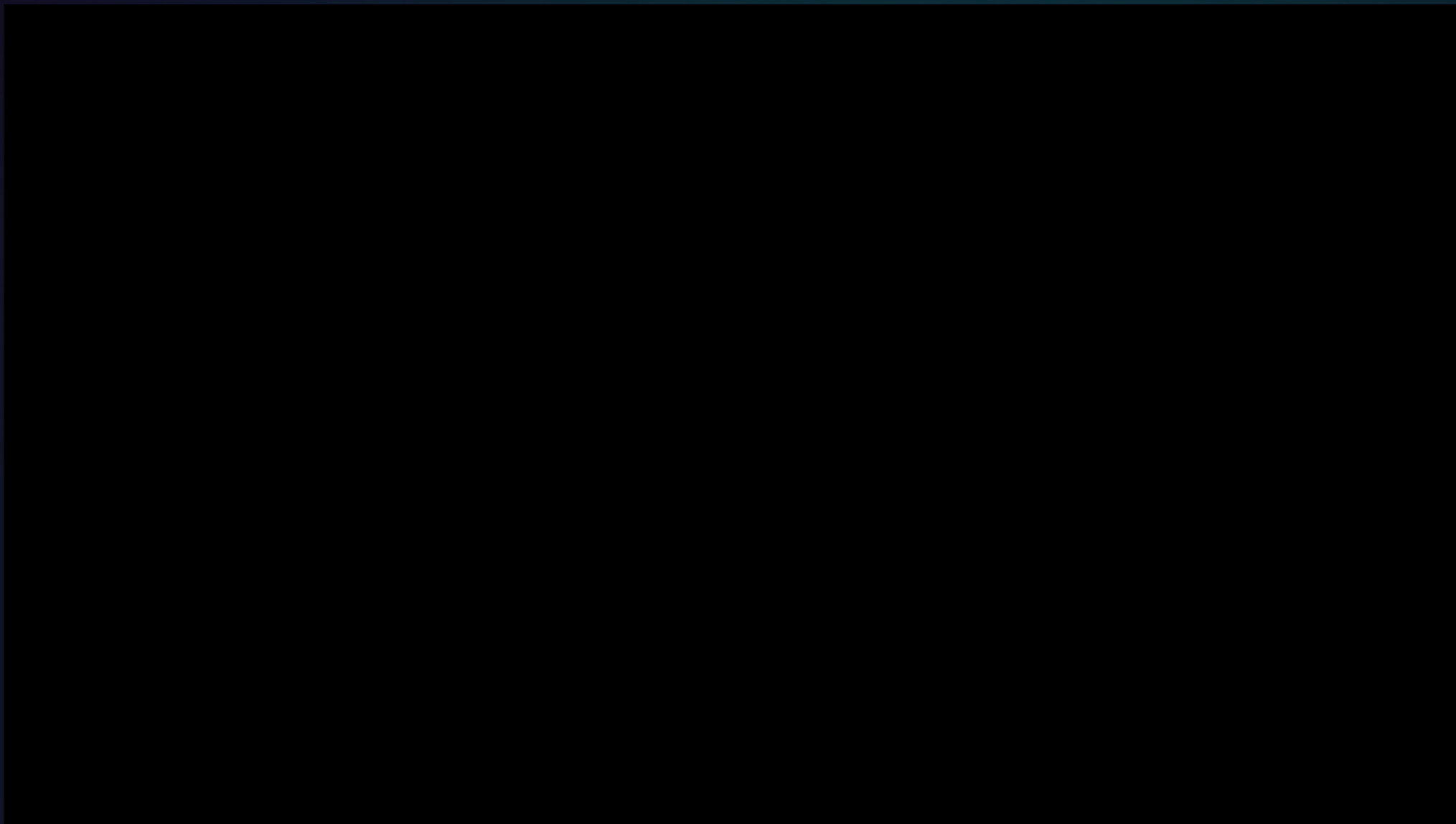


**E4**

COMPUTER  
ENGINEERING

# PRIVATE LLM BASE RAG SEMPLICE

**E4** ANALYTICS



poweredge-r760xa-spec-sheet.pdf

1 / 3 | - 135% + | [ ] [ ]

[ ] [ ] [ ]

**DELL**Technologies

Specifiche tecniche



## PowerEdge R760xa

Server scalabile e a prestazioni elevate per applicazioni GPU intensive

Dell PowerEdge R760xa è un server progettato appositamente per migliorare le prestazioni di accelerazione in tutta la più ampia gamma di esigenze del cliente.

### Innova su larga scala con i workload più impegnativi ed emergenti

Ottimizza le prestazioni dei carichi di lavoro e incrementa i risultati sfruttando la scalabilità di pari passo con la crescita dell'elaborazione, supportando in modo flessibile un'ampia gamma di GPU in un server a doppio socket/raffreddamento ad aria 2U, accelerando le applicazioni dal training e l'inferenza dell'intelligenza artificiale ai gemelli digitali fino a una scheda grafica ad alte prestazioni e agli ambienti di collaborazione ad alta densità per utenti che necessitano di prestazioni elevate.

- Sfrutta un'architettura potente e i due processori Intel Xeon di quarta o quinta generazione con un **numero elevato di core fino a 64** e le più recenti innovazioni on-chip per ottimizzare le operazioni di intelligenza artificiale e ML
- Fino a quattro acceleratori PCIe Gen5 double-width o fino a 12 acceleratori PCIe single-width\* per superare i limiti di densità attuali e futuri dell'elaborazione
- Supporto per adattatori GPU PCIe da NVIDIA, AMD e Intel, per risultati superiori con un'unica piattaforma

### Throughput I/O accelerato

- Implementa tecnologie di ultima generazione tra cui DDR5, NVLink Gen4, PCIe Gen 5 e SSD NVMe E3.S per superare i limiti del flusso di dati e delle possibilità di elaborazione.
- Scalabilità verso l'alto per supportare fino a 32 slot DIMM di memoria DDR5, fino a otto unità e slot di espansione PCIe Gen 5
- Il design con raffreddamento ad aria e acceleratori anteriori consente un raffreddamento migliore e supporta acceleratori con TDP più elevati (fino a 350 W)

### Architettura resiliente agli attacchi informatici per operazioni e ambiente IT Zero Trust

## Code Converter

Run Once Run Batch

programming language

input

input

Clear

▶ Execute

## AI Completion



AI will give you what you want here.



## ASSISTENTE ALLA MANUTENZIONE

Addestrare un LLM utilizzando i propri manuali di manutenzione e i codici di errore, inclusi i dati relativi a macchine o processi proprietari che non si desidera esporre pubblicamente.



## ASSISTENTE ALL'ASSISTENZA CLIENTI

Forma un LLM per aiutare il personale del servizio clienti a rispondere alle domande dei clienti, comprese le informazioni specifiche dei tuoi ticket di help desk del passato, le tue politiche e procedure attuali, mantenendo le tue conoscenze proprietarie lontane dagli LLM pubblici



## ASSISTENTE DI CODIFICA

Addestra un LLM sul tuo codice sorgente, sui tuoi standard e sulle convenzioni utilizzate dal tuo team di sviluppo in modo che l'assistente di codifica generi codice conforme agli standard che funzioni bene e si adatti al resto della tua libreria di sviluppo.



### STUDI LEGALI

Gli studi legali possono utilizzare GAIA e URANIA per creare un assistente di chat che li aiuterà ad analizzare i documenti legali in modo più efficiente. I loro dati sensibili che non possono essere condivisi pubblicamente saranno mantenuti privati e sicuri.



### INDUSTRIA DELL'ASSISTENZA SANITARIA

Le organizzazioni sanitarie gestiscono una grande quantità di dati sensibili che devono essere mantenuti riservati. GAIA e URANIA possono essere utilizzati per creare modelli linguistici specifici per analizzare le cartelle cliniche dei pazienti, identificare modelli e approfondimenti e migliorare i risultati dei pazienti.



### SERVIZI FINANZIARI

GAIA e URANIA possono aiutare le organizzazioni finanziarie a sviluppare modelli linguistici personalizzati in grado di analizzare i dati finanziari, rilevare frodi e anomalie e fornire approfondimenti per migliorare le prestazioni finanziarie e il servizio clienti.



### AGENZIE GOVERNATIVE

Le agenzie governative gestiscono i dati sensibili che devono essere mantenuti riservati e l'utilizzo di GAIA e URANIA può aiutarle a creare modelli linguistici personalizzati per analizzare i dati provenienti da varie fonti e migliorare i processi decisionali.



### MARKETING E PUBBLICITÀ

Le agenzie di marketing e pubblicità spesso analizzano grandi quantità di dati sui consumatori per sviluppare campagne di marketing mirate. Utilizzando GAIA e URANIA, le agenzie di marketing possono garantire che i dati dei loro clienti rimangano riservati, pur essendo in grado di analizzarli in modo efficace



### AEROSPAZIALE

GAIA e URANIA possono aiutare le aziende aerospaziali a migliorare la simulazione implementando LLM riservati e specifici per il dominio per creare scenari che non sono solo istantanee statiche, ma simulazioni viventi che cambiano e si evolvono.

**E4**

COMPUTER  
ENGINEERING

**GRAZIE**



# CONTACTS

Email contacts

[info@e4company.com](mailto:info@e4company.com)

[support@e4company.com](mailto:support@e4company.com)

[sales@e4company.com](mailto:sales@e4company.com)

E4 Computer Engineering SpA

Via Martiri della Libertà, 66 . 42019 Scandiano (RE) - Italy

Tel. +39 0522 991811

