



Als Tochterunternehmen eines der größten unabhängigen Softwarehäuser Deutschlands ist unser Kunde auf fachspezifische Software für Privatanwender spezialisiert. Im Fachsegment ist unser Kunde mit den zugehörigen Web- und Mobile-Anwendungen Marktführer und wird regelmäßig mit Preisen ausgezeichnet. Für diesen Kunden suchen wir im Rahmen der Festanstellung am Standort Hannover zum nächstmöglichen Zeitpunkt einen:

DevOps-Engineer für KI Infrastruktur (m/w/d)

Job-ID: CF-00007279

Job-Ort: Hannover

Remoteanteil / Home Office: bis zu 100% Remote möglich

Hauptaufgaben:

- Hosting und Betrieb von hoch verfügbaren und skalierbaren lokalen KI-Modellen
- Einrichtung und Betreuung von Services und Datenbanken in unserem Rechenzentrum mit eigenen GPU-Servern
- Management von Kubernetes-Clustern, Deployments und KI-Komponenten mit durchdachtem Logging & Monitoring
- Bereitstellung und Automatisierung von robusten CI/CD-Prozessen mit flexiblen Stages und Environments für sichere Entwicklung
- Begleitung von Rollouts unserer KI-Anwendungen
- Administration und Performanceoptimierung von Vektordatenbanken (z. B. PGVector, Weaviate, Qdrant) mit Caching-Mechanismen (z.B. Redis)

Ihre fachlichen Voraussetzungen:

- Fundierte praktische Erfahrungen mit Kubernetes, Cluster-Management, Containerisierung, CI/CD und Self-Hosting von KI-Modellen
- Gute Kenntnisse im Bereich Networking (Ingress/Egress), Logging/Monitoring und AppSec/Datensicherheit
- Erfahrung mit Azure Devops, Git, Jenkins und Co.
- Gute Programmierkenntnisse in Python
- Praktische Erfahrung mit LLMs, RAG-Architekturen und AI Engineering (Feintuning, Prompt Engineering) sind von Vorteil
- Fließende Deutschkenntnisse in Wort und Schrift sind erforderlich

Ihre persönlichen Voraussetzungen:

- Teamfähigkeit
- Kreativität & Interesse an neuen Technologien
- Abstraktes Denkvermögen und analytische Fähigkeiten

Unser Mandant bietet:

- Mitarbeit an ambitionierten, innovativen Projekten in einem engagierten und sympathischen Team
- Spannende, abwechslungsreiche Aufgaben und zahlreiche Weiterbildungsmöglichkeiten
- Ausgezeichnete Entwicklungsperspektiven
- 30 Tage Urlaub, vermögenswirksame Leistungen, Kindergartenzuschuss
- Mitarbeiterrabatte von bekannten Marken (z.B. Freizeitparks, Kinos, VPN-Anbietern, Küchengeräten, Sportprodukten uvm.) durch die Kooperation mit Corporate Benefits
- Modernste Technik und Büroausstattung, offene und dynamische Strukturen



Ihr Ansprechpartner:

Erik Lauer

auteega GmbH
Kaiserring 14-16
68161 Mannheim

Telefon: +49 151 624 364 00
E-Mail: erik.lauer@auteega.com
Jetzt bewerben