

+ BE YOURSELF – BECAUSE THAT’S HOW WE WANT YOU TO +

At Anexia, we take responsibility for all the **challenges of the digital world** – each and every day. That’s because we see ourselves as the **“Digital Transformation Engine”**. And how do we succeed in this mission? It’s simple. We are a family of facilitators with innovative ideas. We’re capable of, want to, and allowed to make things happen. That’s what makes us so unique! Do you want to be a part of our digital revolution and make history with us? **#joinourrevolution**

AI/ML ENGINEER – LLM SYSTEMS (M/W/D)

Vienna, Graz, Klagenfurt, Remote | full time

You are passionate about artificial intelligence and enjoy working with cutting-edge LLM technologies? You work in a structured and independent manner and love building intelligent systems that create real value? You are a communicative team player who thrives on implementing innovative AI solutions? Then this position is perfect for you!

#PROFILE



- You will drive forward the development and deployment of AI/ML systems, particularly focusing on Large Language Models (LLMs) and intelligent agent frameworks
- You implement and operate **LLM inference solutions** using runtimes like Ollama, vLLM, and TensorRT-LLM
- You develop **agentic applications** with frameworks such as Dify, LangFlow, Haystack, LlamaIndex, LangChain, LangGraph or Pydantic AI
- You **integrate AI capabilities** from various providers (like OpenAI, Anthropic, Groq, and Cerebras) into production systems
- You implement **LLMOps practices** using tools like LangFuse and LangSmith for monitoring, tracing, and observability
- You **build and optimize ML pipelines** using Python, TensorFlow, PyTorch, and scikit-learn for various AI applications
- You **containerize and deploy AI applications** using Docker and Podman

- You work with SQL databases and implement evaluation frameworks like DeepEval to ensure model quality
 - You work closely with cross-functional teams and enjoy presenting and continuously improving your AI solutions
-



#WANTED

- Experience with RAG (Retrieval-Augmented Generation), prompt engineering or multi-agent systems is highly valued
 - Strong understanding of Large Language Models and experience with HuggingFace ecosystem
 - Hands-on experience with at least one agentic framework (LangChain, LangGraph, Haystack, LlamaIndex, Dify, LangFlow, or Pydantic AI)
 - Experience with Python and modern ML frameworks (TensorFlow, PyTorch, scikit-learn)
 - Proficiency with LLM inference runtimes such as Ollama, vLLM, or TensorRT-LLM
 - Experience with LLM APIs (OpenAI, Anthropic, or similar providers)
 - Strong knowledge of SQL and database operations
 - Understanding of LLMops practices and model evaluation methodologies
 - Experience with containerization (Docker/Podman) and orchestration tools
 - Experience with version control systems and CI/CD tools is an advantage
 - Knowledge of DevOps practices (Ansible) is a plus
 - Very good written and spoken German and English skills
-

#REWARD

Anexia grew up as a family, and with our growth strategy, we intend to keep it that way. We like to take responsibility, respect one another, and we know that we can achieve whatever we put our minds to. We look after Anexia, and Anexia

looks after us. Enthusiasm, experience, and competence count, so you can expect a fair salary and numerous benefits.

salary range



Our benefits



We are more than just a workplace. Therefore, we offer our employees many benefits (which partly depend on your position and location): good transport connections and flexible working hours. Period. Not to mention further training, a mentoring program, free parking, a mobile phone and notebook, snacks and drinks, company cars, employee events, a fun corner, ...

#JOINOURREVOLUTION

APPLY NOW

#ABOUT

Anexia is offering high-quality and individual cloud- and managed hosting solutions as well as individual software and app development. Anexia is also running its own, independent data centers. Founded in 2006, Anexia is headquartered in Klagenfurt and has offices in Vienna, Graz, Karlsruhe, Nuremberg and New York City. The company serves numerous international customers.



Over 350 employees at more than 5 office locations in Austria, Germany and the USA



We are ISO-certified and have had a "very good credit rating" for years



Customers such as Lufthansa, McDonalds, TeamViewer, and many more trust in our quality

anexia

netcup®

TELEMATICA

DATASIX
DATACENTER

You have questions?

Daniela KOZÓ | Telephon: +43-50-556 | E-mail: career@anexia-it.com