

CONTRACTOR SUPPORT TO DATA SCIENCE & AI: DATA SCIENTIST/AI ENGINEER

Location: Norfolk, VA, USA (On-site)

Period of Performance: Base Period: 1 January 2026 – 31 December 2026, Option Period One: 1 January 2027 – 31 December 2027, Option Period Two: 1 January 2028 – 31 December 2028, Option Period Three: 1 January 2029 – 31 December 2029, Option Period Four: 1 January 2030 – 31 December 2030.

Taskings

1. Contribute to the development and implementation of an enabling data science and AI capability at HQ SACT and across the NATO Enterprise, with a specific focus on scalable data engineering and software systems to support AI initiatives.
2. Design, develop, and maintain robust data pipelines and architectures to manage the ingestion, transformation, and processing of structured and unstructured data for large Language Model (LLM)-based applications and other AI systems.
3. Lead efforts to optimize data delivery and automate data engineering processes, proposing enhancements to infrastructure to improve scalability, efficiency, and reliability in support of LLM deployments.
4. Build API-based infrastructure and frameworks that enable seamless integration of LLMs and ML models with operational systems, ensuring performance, security, and interoperability with NATO environments.
5. Support the development, testing, and validation of microservices and containerized applications to operationalize AI/ML capabilities, including deployment of LLM use cases within NATO.
6. Implement distributed data storage and processing systems (e.g., cloud-based or hybrid architectures) that align with NATO standards and enable scalable use of LLMs across the enterprise.
7. Develop tools and systems to improve data accessibility, enabling data scientists and analysts to efficiently interact with and query data for training, inference, and analytics.
8. Coordinate with data scientists, software engineers, and system architects to align data engineering workflows with broader AI/ML objectives, ensuring timely delivery of clean, high-quality data for LLM training and inference.
9. Establish mechanisms for real-time data processing and streaming, enabling LLMs to operate effectively in dynamic and responsive applications, such as operational decision support or strategic analysis.
10. Conduct preprocessing, cleansing, and transformation of raw data into formats optimized for training, fine-tuning, and inference within LLM infrastructure.
11. Implement robust monitoring, logging, and performance optimization tools for data pipelines and APIs, ensuring reliability and traceability of LLM-enabled workflows.
12. Collaborate with teams to support federated learning approaches and cross-domain data sharing, ensuring compliance with NATO data sovereignty, security, and ethical guidelines.
13. Provide subject matter expertise on data engineering and software development to (military and civilian) staff within HQ SACT or the NATO Enterprise, and develop proofs of concept for LLM-based applications as directed.
14. Research, recommend, and implement best practices for deploying LLMs in secure, cloud-based environments such as Microsoft Azure or AWS, while considering NATO-specific data policies and standards.

15. Evaluate operational requirements and objectives, recommending appropriate engineering solutions for integrating LLMs into NATO workflows and systems.
16. Stay abreast of new developments in AI engineering, including innovations in LLM technologies, data architectures, distributed computing, and API development, to bring cutting-edge capabilities into implementation within NATO.
17. Provide technical training and mentoring to NATO staff, supporting educational efforts in AI engineering, data pipeline design, API development, and digital literacy.
18. Foster a culture of innovation and data-driven decision-making across NATO by building scalable systems that enable the effective exploitation of LLMs and advanced analytics.
19. Perform additional tasks as required by the Contracting Officer's Technical Representative (COTR) related to the LABOR category.

Essential Qualifications

a. Essential qualifications:

1. A Bachelor's degree or higher at a nationally recognized/certified university in Data Science, Data Analytics, AI engineering, or a related discipline such as Mathematics, Physics, Computer Science, Software Engineering OR 4 years minimum professional experience in the area of Data Science, including providing analysis and advice in the field of data science, within the last 5 years.
2. Minimum 4 years of proven work experience as a Data Scientist, Machine Learning Engineer, Data Engineer, or Software Engineer, with a strong emphasis on distributed systems, cloud-based architectures, developing operational AI/ML solutions, and designing API-based infrastructures, microservices architectures, and containerized applications (e.g., Docker, Kubernetes).
3. Demonstrated experience working with GenAI, in particular LLMs, including preprocessing data, fine-tuning, and deployment in secure and scalable environments to include AI/ML frameworks such as TensorFlow, PyTorch, or scikit-learn.
4. Proven expertise in programming languages such as Python, Java, or Scala, with demonstrated experience in software engineering practices (e.g., version control, CI/CD pipelines, containerization).
5. Experience building and optimizing data pipelines, ETL processes, and real-time streaming solutions using tools like Apache Airflow, Kafka, Spark, or equivalent.
6. Knowledge of applied AI principles, particularly in implementing AI systems for operational decision support and analyzing unstructured data (e.g., text, imagery).
7. Ability to architect and maintain scalable data lakes, data warehouses, or distributed storage systems (e.g., Delta Lake, Snowflake, Hadoop, or NoSQL solutions).
8. Demonstrated understanding of data security, privacy, and sovereignty issues, particularly in military or international environments, ensuring compliance with NATO operational and ethical standards.
9. Experience building visually impactful reports, dashboards, and analytics using tools such as Tableau, MS Power BI, or Kibana, supporting informed decision-making for high-level stakeholders.
10. Professional experience in NATO environments or familiarity with NATO processes, organizational culture, and decision-making structures.
11. Ability to translate operational problems into practical AI/ML solutions tailored for military and civilian teams.
12. Proven ability to collaborate effectively within multidisciplinary teams, including coordinating with data scientists, software engineers, and system architects on cross-functional projects.
13. Strong oral and written communication skills, with the ability to brief non-technical audiences and mentor staff in AI engineering, data science, and software development concepts.

Desirable Qualifications

1. Experience leveraging open-source frameworks and publicly available datasets to develop innovative AI and data engineering solutions for operational or analytical use cases.
2. Proficiency in presenting data-driven insights clearly to non-technical audiences, showcasing an ability to craft compelling narratives and actionable recommendations for senior leadership.
3. Understanding of military staff workflows and processes, alongside familiarity with federated learning techniques for enabling secure collaboration across NATO nations while preserving sovereignty of sensitive datasets.
4. Exposure to agile project management methods and tools (e.g., Loop, JIRA, Trello) for coordinating and tracking progress across multi-disciplinary AI/ML projects.
5. Eligibility for NATO security clearance and prior experience working with classified or sensitive data, including understanding security protocols for processing, handling, and securing such data.
6. Exposure to cross-domain data sharing and API-driven interoperability, ensuring effective integration across systems while adhering to security and ethical guidelines within military or international environments.
7. Familiarity with principles of ethical AI development, including considerations for bias mitigation, responsible data handling, and alignment with NATO's ethical frameworks for AI deployment.