**Position Roles and Responsibilities – MACSS G-014 Data Science**

The Contractor shall provide data science and services as described below. The contractor shall be responsible for providing all support services and products required to complete the detailed description of work and responsibilities as described herein. The contractor shall perform the scope of work within the period of performance authorized at contract award.

The Contractor shall provide data science services within NGA in support of GEOINT analysis. Knowledge of data science, methodologies, and processing techniques shall be employed on behalf of GEOINT analysis and community customers while providing operational support. Individual data scientists shall have applied knowledge in five distinct areas depending on mission requirements:

* Applied Mathematics – Statistical and mathematical support to provide temporal and pattern analysis, correlation of events, probability analysis, assessments of sampling, machine learning analysis of variance and error, and regression testing and analysis.
* Programming Skills – Write scripts in Visual Basic, Python, or other software for modeling processes, repeatability, efficiency, knowledge capture, hypothesis testing and process automation.
* Data Stories – Document and visualize data both temporally and spatially to assist in data integrity checks, ask the next question, and display analytical assessments.
* Scientific Discipline – Implement scientific rigor of quantitative and qualitative techniques to enhance the analysis of complex national security problems.
* Capabilities Advancement - Apply emerging and innovative technologies to answer key intelligence questions. Provide technical consultation into the development, evaluation, use, and deployment of solutions to optimize GEOINT analysis and production.

**Required:**

Education and Experience levels:

* Master of Science degree or higher in a quantitative discipline such as data science, mathematics, statistics, earth science, quantitative social sciences, geographic information science, computer science, physics, or a related field. In conjunction with proficiency and at least seven years’ experience with applied data processing and scientific analysis of large datasets and machine learning

OR

* Bachelor of Science degree in a quantitative discipline such as data science, mathematics, statistics, earth science, quantitative social sciences, geographic information science, computer science, physics, or a related field. In conjunction with proficiency and at least ten years’ experience with applied data processing and scientific analysis of large datasets and machine learning.
* Extensive knowledge of programming languages (e.g., Python, Pig, Java, JavaScript, SQL, or R), spatial analysis tools and concepts, data mining methods, database structures and analytic information extraction and visualization.
* Training in applied mathematics including statistics and mathematical modeling to support temporal and pattern analysis, correlation of events, probability analysis, assessments of sampling, analysis of variance and error, and regression testing and analysis. Demonstrated competency in one or more programming languages such as JavaScript, Python, or R for processes, repeatability, efficiency, knowledge capture, hypothesis testing, visualization, and process automation.
* Familiarization with modeling software (eg. FME, SPSS, SAS) for processes, repeatability, efficiency, knowledge capture, hypothesis testing, visualization, and process automation.
* Demonstrated experience relevant to managing data science projects and workflows
* Demonstrated experience programmatically connecting to database and web based data sources.
* Demonstrated knowledge and experience in data mining, cleansing, and exploring spatial, temporal, and non-spatial data in both structured and unstructured formats.
* Demonstrated experience with two or more cloud based technologies.
* e.g. Docker containers, Jupyter Hub, Zeppelin, Apache-Spark, Centos, Jenkins, GIT
* Proficiency in MS Office suite.

**Desired:**

* PhD in a quantitative discipline such as data science, mathematics, statistics, earth science, quantitative social sciences, geographic information science, computer science, physics, or a related field.
* Computational analytics modeling
* Experience engineering and tuning deep learning algorithms for data science. Experience with two or more programing languages:
* R, Python, Pig, Java, JavaScript, and/or SQL
* Experience running command line operations in one or more operating systems:
* Windows and/or Linux (command line)
* Experience with two or more visualization tools \ packages: e.g. (ggplot, Plotly, matplotlib, D3, Tableau, bokeh)
* Experience connecting to two or more database and web data sources:
* Database: e.g. Postgres, Oracle, SQLite, and/or Arc SDE
* Web: e.g API’s, GeoJson, REST
* Demonstrated developer experience with two or more cloud based technologies. e.g. Docker containers, Jupyter Hub, Zeppelin, Apache-Spark, Centos, Jenkins, GIT
* Domain knowledge / IC: Background in intelligence, defense, international relations, or public administration multi-disciplines.
* Demonstrated familiarity with the US Intelligence Community specifically GEOINT collection and exploitation.
* Experience collaborating with all national and service intelligence agencies/centers
* Demonstrated familiarity with the GEOINT Community and associated TPED products