Statement of Work for Dr Peter Fox

National Center for Atmospheric Research

MaDS:

To be submitted to the NASA ROSES-2008 AIST Program

PI: Dr. Gregory Leptoukh (Gregory.Leptoukh@nasa.gov)

Aerosol measurements from satellites play an important role in improving understanding of aerosol impact on climate change and providing monitoring capabilities for air pollution assessment. Multiple sensors measure aerosols using different technologies and spatio-temporal resolution. To compare these measurements with each other and the relevant models and ground-based observations, analysis systems have been constructed to support a consistent joint aerosol data processing and analysis. Some of these systems provide easy interfaces and access to the results.

However, some concern has been expressed by researchers and application users regarding the provenance of online analysis results. For example, how are science data quality fields used in the analysis? How does a user know what combination of parameters, quality information, thresholds, are scientifically "safe" to use? One such analysis system is Giovanni, the Goddard online visualization and analysis tool. Giovanni's ease of use, normally a positive characteristic, also raises the question of whether it is easy to misuse.

One aspect is knowledge about science quality usage recommendations for a single parameter analysis. Another, much more complicated, is science knowledge required for multi-parameter analysis. Supporting data intercomparison from different instruments or models produces cases when a certain combination of selected parameters and Giovanni services is either should not be allowed or should be qualified in a certain way. For example, for the data merging service in Giovanni, only compatible parameters from different instruments should used together – one should not merge aerosol optical depth with precipitation.

These rules are currently devised by scientists and then explicitly incorporated into the application. However, even today, a sheer number of combinations of parameters, sensors, and services, makes impossible the straightforward "hardcoding" of the known rules into an interface. This approach is not very scalable as we support more intercomparison amongst more parameters from new Decadal Survey instruments, specifically, aerosol measurements by ACE, together with data from VIIRS onboard NPP and NPOESS, with the data from the currently sensor data from MODIS, OMI, POLDER, etc.

Instead, we seek a flexible approach whereby the characteristics of the data set variables and their related quality variables can be encoded in a way that these intercomparison rules can be derived.

Tasks and Deliverables

<u>Year 1</u>

Co-I Dr. Fox and HAO team will work with PI team to document the use cases and identify the required ontology and provenance representations. HAO will also assist in capturing scientist knowledge about the essential science and quality characteristics of parameter measurements and work to encode this knowledge in such a way that a computer can retrieve it based on user input or machine inference HAO will either install or provide required semantic web technology infrastructure to demonstrate initial phase of use case. HAO team will participate in presentations and publications of work.

Year 2

Co-I Dr. Fox and HAO team will continue and complete the analysis of the first use case and commence work on the second use case and work with the PI team to present only the safe (valid) analysis comparisons, or in some cases the caveats regarding speculative comparisons, representing them semantically. HAO team will participate in presentations and publications of work.

Year 3

Co-I Dr. Fox and HAO team will work with the PI team in demonstrating use cases and work with PI team to provide tunable options for quality screening/weighting (if any) to the user. HAO will assist in generating the Giovanni workflow to perform the selected operation and record the provenance associated with it. HAO team will participate in presentations and publications of work.

Total Cost: $