ADVANCES IN DATA INTEGRATION AND QUALITY CONTROL IN SUPPORT OF GROUND-BASED NUCLEAR DETONATION DETECTION

Richard J. Stead, Michael L. Begnaud, and Julio C. Aguilar-Chang
Los Alamos National Laboratory

Automated QC

A first-generation QC tool, recently completed, was developed (LASS). The LASS schema is based on a structured, object-oriented database system, which allows for the incorporation of complex data types and object hierarchies. The tool is designed to perform automated quality control checks on incoming data, with the goal of identifying and correcting errors before data is stored in the database. This tool has been used successfully in both research and development environments, demonstrating its effectiveness in improving data quality and reducing the time required for manual QC. The next generation of the tool is using the features in the LASS database, which is expected to result in improved accuracy and efficiency in the QC process.

Infrastructure for QC

The need for Quality Control (QC)

It is simple in that it has only four objects in the schema: External resources of data are obtained or received by LANL certified Generalized Data Acquisition and Integration Process.

Table Schema Document, Sandia National Laboratories report SAND2002-27th Annual Seismic Research Review: Ground-Based Nuclear Explosion Monitoring Technologies

- Limitations are that it only does QC/QA (inspection, not repair), and can only do this on database objects.
- Automatically handles all documented descriptions of the schemas, no need to track all possible QC/QA checks.
- Identifies table type if the table is documented in the schema.
- Recognizes that the table has at least one column and column.
- Recognizes that the column has at least one column.

Conclusion and Recommendations

We have developed a suite of tools to facilitate data processing, which has evolved over time to meet emerging needs and capabilities. The toolset includes a variety of components, from data intake and management to data analysis and visualization. We have found that the integration of these tools into a single, cohesive package is necessary to achieve the desired results. In the future, we plan to continue to refine and improve the toolset, with a focus on enhancing its capabilities to handle larger datasets and more complex analyses. We also plan to collaborate with other institutions and groups to share our findings and experiences, and to continue to contribute to the developing field of data sciences.

The views expressed here do not necessarily reflect the opinion of the United States Government, the United States Department of Energy, or the speaker's Laboratory.