



## Super Report: EPA Common Framework

### **Background**

The Environmental Protection Agency (EPA) holds large amounts of data across the organization in support of different areas of the agency. Some of this data is provided by third parties, and some of it is developed internally. These systems have no common framework for publishing a working report across multiple applications to derive usable information for knowledge management and decision support.

EPA sought to remove this obstacle for the purpose of developing a component by which the public can easily search for and access information about the air, water or soil quality of specific locations. To this end, Intervise designed and built an application to serve as a prototype for future development based around the creation of a complex ontology for defining appropriate classes to create usable data from these systems and to facilitate the ability to cross reference against outside forms of data.

### **Services Provided**

Intervise began by building an environmental ontology which related facilities, chemicals, sensors and readings. Additionally, Intervise captured geography for the purpose of relationship tracking. Next, data provided by the EPA for facilities and chemicals was semantically enabled and location encoded. Because EPA does not currently track sensors or readings, Intervise located and utilized web services available from the United State Geological Survey (USGA) in order to complete this step. The semantically enabled data was then loaded into a general purpose semantic web processor for fact checking. Finally, Intervise created and implemented custom visualization screens based on EPA's specific requirements and preferences. This was the only phase that required a minor programming effort. In the interest of efficiency and cost control, all other components were leveraged and left available for re-use in future development projects.

Specifically, Intervise gathered related data from relational databases and web services. Processing this data resulted in semantically enabled data or OWL

instances. The resulting system is a fully SOAP/RESTful web service implementation with accredited multi-level security. Intervise facilitated machine-to-machine interaction between systems that previously did not communicate and implemented XPDL workflow engine in order to relate data.

Natural language processing requires an ontology that categorizes and relates data. This ontology must adhere to the OWL-2 EL++ specification to allow reasonable retrieval times from systems; therefore, Intervise provided a full OWL based EL++ ontology to support Facilities, Chemicals, Sensors and Readings. This ontology is instantiated and allows a variety of access mechanism including basic inferences, keywords and attribute searches.

### **Outcome**

As a result, these EPA systems can now facilitate a search across systems which will return more specific and relevant information and can be used to create a "Super Report" which cross-references information from all applicable systems and allows data to be visualized in numerous and customizable ways. The implementation of this technology provided a prototype which EPA can leverage in order to facilitate a user-friendly search for air, water or soil quality of specific locations.