Streamlining Data Access Services and Data Analysis Services Integration

Few are the data analysis services capable of understanding and consuming data coming from multiple data access services. This lack of interoperability between data access services and data analysis services is indeed a major roadblock for science because it prevents the reuse and repurposing of both data and analytical software to support new scientific discoveries. In this presentation, we discuss this problem in light of the Earth, Life, and Semantic Web (ELSEWEB) project funded through NASA’s ACCESS Program. The project uses the University of Kansas’ Lifemapper system as its analytical Web Service platform, which models potential future species distributions under scenarios of climate change. In an effort to broaden the range of scenarios to include land cover/land use change, ELSEWEB aims to streamline the flow of highly heterogeneous geographic, social, and geological data hosted at UNM’s Earth Data Analysis Center (EDAC) through a collection of OGC Web Coverage Services into Lifemapper. In turn, this integration will enable new modeling of complex factors associated with biotic change such as health and infectious disease, that depend not only on climate change and species distributions, but also on other human/environmental interactions.

In this presentation we discuss the integration of Lifemapper and EDAC data and model Servicesis provided by a third party semantic system, known as VisKo, to (1) translate Lifemapper data requirements to EDAC service invocations and (2) pipe the data output from EDAC into Lifemapper. VisKo is supported by a knowledge base of web service descriptions that contains information about interface requirements as well as invocation details, including service parameters. VisKo ontologies are designed to capture the knowledge required by the system to orchestrate and execute service pipelines that perform scientist’s required computation.

The ELSEWEB project also aims to expand VisKo’s original goal of building visualization pipelines into a framework that can also handle data analytics. For example, EDAC’s OGC services provide a rich set of information about spatial and temporal coverage, as well as semantic descriptions of the data, such as cloud or aerosol. Because this additional knowledge is relevant when deciding what data can be fed to Lifemapper, VisKo is set to translate OGC service metadata into semantic content that VisKo can use to reason with and match up the right EDAC OGC services with the right Lifemapper modeling services.