



A Query Language for Handling Big Observation Data Sets in the Sensor Web

Christian Autermann (1), Christoph Stasch (1), Simon Jirka (1), and Roland Koppe (2)

(1) 52°North GmbH, Münster, Germany, (2) Alfred Wegener Institute, Helmholtz Centre for Polar and Marine Research, Bremerhaven, Germany

The Sensor Web provides a framework for the standardized Web-based sharing of environmental observations and sensor metadata. While the issue of varying data formats and protocols is addressed by these standards, the fast growing size of observational data is imposing new challenges for the application of these standards. Most solutions for handling big observational datasets currently focus on remote sensing applications, while big in-situ datasets relying on vector features still lack a solid approach. Conventional Sensor Web technologies may not be adequate, as the sheer size of the data transmitted and the amount of metadata accumulated may render traditional OGC Sensor Observation Services (SOS) unusable.

Besides novel approaches to store and process observation data in place, e.g. by harnessing big data technologies from mainstream IT, the access layer has to be amended to utilize and integrate these large observational data archives into applications and to enable analysis. For this, an extension to the SOS will be discussed that establishes a query language to dynamically process and filter observations at storage level, similar to the OGC Web Coverage Service (WCS) and its Web Coverage Processing Service (WCPS) extension. This will enable applications to request e.g. spatial or temporal aggregated data sets in a resolution it is able to display or it requires.

The approach will be developed and implemented in cooperation with the The Alfred Wegener Institute, Helmholtz Centre for Polar and Marine Research whose catalogue of data comprises marine observations of physical, chemical and biological phenomena from a wide variety of sensors, including mobile (like research vessels, aircrafts or underwater vehicles) and stationary (like buoys or research stations). Observations are made with a high temporal resolution and the resulting time series may span multiple decades.