The Challenge of Handling Big Data Sets in the Sensor Web

Christian Autermann, Christoph Stasch, and Simon Jirka

52°North GmbH, Germany (c.autermann@52north.org)

More and more Sensor Web components are deployed in different domains such as hydrology, oceanography or air quality in order to make observation data accessible via the Web. However, besides variability of data formats and protocols in environmental applications, the fast growing volume of data with high temporal and spatial resolution is imposing new challenges for Sensor Web technologies when sharing observation data and metadata about sensors.
Variability, volume and velocity are the core issues that are addressed by Big Data concepts and technologies. Most solutions in the geospatial sector focus on remote sensing and raster data, whereas big in-situ observation data sets relying on vector features require novel approaches.

Hence, in order to deal with big data sets in infrastructures for observational data, the following questions need to be answered:
1. How can big heterogeneous spatio-temporal datasets be organized, managed, and provided to Sensor Web applications?
2. How can views on big data sets and derived information products be made accessible in the Sensor Web?
3. How can big observation data sets be processed efficiently?

We illustrate these challenges with examples from the marine domain and outline how we address these challenges. We therefore show how big data approaches from mainstream IT can be re-used and applied to Sensor Web application scenarios.