



Tooling for interoperability and integrity between primary data sources and data infrastructures

Aleksi Aalto and Tommi Vuorinen

University of Helsinki, Institute of Seismology, Department of Geosciences and Geography, Helsinki, Finland
(aleksi.j.aalto@helsinki.fi)

Geoscience knowledge derives from a set of sparse, heterogeneous, and noisy information. Institute of Seismology governs various different types of seismological data, related observations, and results. A major part of this data is contributed to various data infrastructures.

Protocols and methods for contributing data to data infrastructures are typically agreed for the whole data infrastructure, but the methodology used by the infrastructure members for internal operations is often ad-hoc in nature due to the differing needs of infrastructure members. This may lead to interoperability issues with e.g. semantics that deteriorate the data quality.

Tooling has been developed for the Institute of Seismology to enhance internal and referential integrity between primary data sources and the data infrastructures. This includes modeling the business logic of data production in seismological networks and general research workflows. Standard data transfer formats are used as a basis and extended appropriately. The tooling allows identifying issues between separate primary data sources, rapid prototyping for research, and improves the effectiveness of continuous analysis operations.