Proposed metadata standards for FAIR access to GNSS data

Anna Miglio[1], Andras Fabian[1], Carine Bruyninx[1], Stefanie De Bodt[2], Juliette Legrand[1], Paula Oset Garcia[2], and Inge Van Nieuwerburgh[2]

[1]Royal Observatory of Belgium, Brussels, Belgium

[2]Ghent University, Ghent, Belgium

Accurate positioning for activities such as navigation, mapping, and surveying rely on permanent stations located all over the world and continuously tracking Global Navigation Satellite Systems (GNSS, such as Galileo, GPS, GLONASS).

The Royal Observatory of Belgium maintains repositories containing decades of observation data from hundreds of GNSS stations belonging to Belgian and European networks (e.g., the EUREF public repository).

However, current procedures for accessing GNSS data do not adequately serve user needs. For example, in the case of the EUREF repository, despite the fact that its GNSS data originate from a significant number of data providers and could be handled in different ways, provenance information is lacking and data licenses are not always available.

In order to respond to user demands, GNSS data and the associated metadata need to be standardised, discoverable and interoperable i.e., made FAIR (Findable, Accessible, Interoperable, and Re-usable). Indeed, FAIR data principles serve as guidelines for making scientific data suitable for reuse, by both people and machines, under clearly defined conditions.

We propose to identify existing metadata standards that cover the needs of the GNSS community to the maximum extent and to extend them and/or to develop an application profile, considering also best practices at other GNSS data repositories.

Here we present two proposals for metadata to be provided to the users when querying and/or downloading GNSS data from GNSS data repositories.

We first consider metadata containing station-specific information (e.g., station owner, GNSS equipment) and propose an extension of GeodesyML, an XML implementation of the eGeodesy model aligned with international standards such as ISO19115-1:2014 and OGC's GML. The proposed extension contains additional classes and properties from domain specific vocabularies when necessary, and includes extra metadata such as data license, file provenance information, etc. to comply with FAIR data principles. All proposed changes to GeodesyML are optional and therefore guarantee full backwards compatibility.

Secondly, we consider metadata related to GNSS observation data i.e. RINEX data files. We propose an application profile based on the specifications of the Data Catalog Vocabulary (DCAT), a RDF vocabulary that, by design, facilitates the interoperability between data portals (supporting DCAT-based RDF documents) and enables publishing metadata directly on the web by using

different formats.

In particular, our proposal (GNSS-DCAT-AP) includes new recommended metadata classes to describe the specific characteristics of GNSS observation data: the type of RINEX file (e.g., compression format, frequency); the RINEX file header and information regarding the GNSS station including the GNSS antenna and receiver; the software used to generate the RINEX file. Additional optional classes allow the inclusion of information regarding the GNSS antenna, receiver and monument associated with the GNSS station and extracted from the IGS site log or GeodesyML files