GNSS Station Metadata Revisited in Response to Evolving Needs

Carine BRUYNINX, Andras FABIAN, Juliette LEGRAND, Anna MIGLIO
Royal Observatory of Belgium, Av. Circulaire 3, Brussels, Belgium
C.Bruyninx@oma.be

EGU2020-18634

M³G web site:
https://gnss-metadata.eu

Royal Observatory of Belgium
GNSS Research Group

@be_gnss
www.gnss.be

Abstract

The IGS International GNSS Service) site log format is the worldwide standard for exchanging GNSS station metadata. It contains, among other things, a description of the GNSS site and its surroundings, the contact persons, and an historical overview of the GNSS equipment. This information is valuable for reliable GNSS data analysis and interpretation of the results.

This IGS site log is also used within the EUREF Permanent Network (EPN, http://epncb.oma.be) and the GNSS component of the European Plate Observing System (EPOS, https://www.epos-eu.org/).

However, due to their specific needs, these networks collect additional GNSS metadata. The EPN collects the information on the data provided by the station, while EPOS is collecting data licenses and Digital Object Identifiers (DOI).

This poster investigates the possibility to use Digital Object Identifiers to provide access to some of the additional GNSS metadata (including the data license) and explores how this can be done in practice.

Need for additional GNSS station metadata

Growing need to

- acknowledge merits of data provider
- provide statistics on data usage for funders
- provide user information on data access restrictions

Digital Object Identifiers (DOIs) for data

DOI is a character string (standardized by ISO) used to uniquely identify an object such as journal articles, research reports and **data sets,** e.g. https://doi.org/10.24414/FST8-P256

Prefix: assigned by the DOI allocation agency

Suffix: assigned by the agency minting the DOI (responsible for the content and its long-term storage)

How to get a DOI for data sets?

Allocation Agents, who are members of Registration Agencies, assign DOIs to objects. CrossRef (papers) and DataCite (data) are the most known DOI Registration Agencies.

Clients, like universities, may sign a contract with an allocation agency to become minting agencies able to "register" or "mint", DOIs. E.g. ROB mints DOIs thanks to a contract with an allocation agent (TU Delft).

DataCite DOI Metadata Scheme

Rich metadata scheme

Useful information is stored in the DOI metadata (see https://schema.datacite.org/)

Mandatory

Identifier, Creator, Publisher, Title, PublicationYear, resourceType

Data citation:
Creator (PublicationYear): Title. Publisher. (resourceTypeGeneral). Identifier

Optional

FundingReference, Format, Rights, rightsIdentifier (e.g. CC-BY-4.0, CC-BY-NC-4.0, ...), rightsURI (e.g. http://creativecommons.org/licenses/by/4.0)

Information on funding, data format, data license, ...

Optional

relatedIdentifier → DOI
relationType: IsPartOf, IsIdenticalTo, IsSourceOf, HasPart, ...

Allows to build relations between DOIs

Is it worth the effort to assign DOIs to GNSS data?

+ reliable long-term citation
+ facilitates access to, sharing and reuse of data
+ allows to include information on data license
+ allows cross-linking between
  - publications and underlying data
  - agency-specific GNSS network DOI and international GNSS network DOI (possible to set up hierarchical relations between DOIs)

Learn from other communities !

Federation of Digital Seismograph Networks (FDSN): "FDSN recommendations for seismic network DOIs and related FDSN services", http://www.fdsn.org/pdf/V1.0-21Jul2014-DOIFDSN.pdf

Community-agreement on
- content of mandatory DOI metadata fields (use for data citation)
- recommended metadata and proposal for homogenized approach on how to complete them (resource type, description, format, contributor, location, size, related identifier, ...

with practical examples.

Several DOI services offered by FDSN to lookup DOI:

e.g. Input Seismic network code (unique identifier) → output : DOI

Questions when trying to apply DOI to GNSS data

Do we need a DOI for each GNSS station?
No, but each agency owning a GNSS network should have a DOI for its GNSS network.

Who requests the DOI? The repository or the data owner?
The repository (analogy: paper → publisher ; data → repository)

If GNSS data are available from multiple repositories, does each repository have its own DOI for the same GNSS data?
No necessarily. If a new DOI is however requested by the repository for a GNSS network which already has a DOI, then use "IsIdenticalTo" or "IsVariantFormOf" metadata fields to link to the already existing DOI.

New GNSS data are collected every day. Can we associate DOIs to datasets that are still growing?
See recommendations in https://schema.datacite.org

If I use GNSS data from the IGS network, must I then cite ALL the DOIs from the underlying agency networks?
No, the IGS network data can have one DOI for all station data.
The metadata of this IGS DOI must however indicate the relation to the DOIs of the underlying networks by e.g. using "HasPart" or "References" metadata fields.

How does the user know if a DOI has been attributed to a specific GNSS data set?
GNSS-community needs to develop DOI services to lookup DOI (e.g. based on unique 9-char station identification or unique network name – as done in seismological community).

Is someone already collecting DOIs for GNSS data?
ROB's M³G system (https://gnss-metada.eu/) allows station managers to set up their agency network (unique network name) and collects the DOI for that the GNSS network.

Can the data access conditions be provided through the DOI?
Yes, use "Right" metadata fields, which allows to indicate the data license.

Conclusions

Attributing DOIs to GNSS datasets is important because it allows to
- cite GNSS data (crucial for crediting researchers and funders)
- include information on data license

However, because of large amount of GNSS stations (organized in local/regional/global networks) + GNSS data available from multiple data repositories

→ no community-agreed standards on how to assign DOIs to GNSS data
→ only few GNSS data with DOI
→ no good practice to cite GNSS data using DOI in papers

Consequently, researchers and other data providers are not encouraged to make their data freely available.

DataCite metadata scheme allows to deal with complexity of GNSS data

Similar to what was done within the seismological community, GNSS community must agree on
- common standards for completing the DOI metadata
- provision of DOI services that enable data users to inquire if a GNSS data set is associated with a DOI and harvest DOI metadata such as data access restrictions.

Discussion started in GGOS Working Group on DOI (chair: Kirsten Elger)

Long-term goal, but GNSS community needs to get started!

https://doi.org/10.5194/egusphere-egu2020-18634