Digital Identifiers and Metadata for the IVS
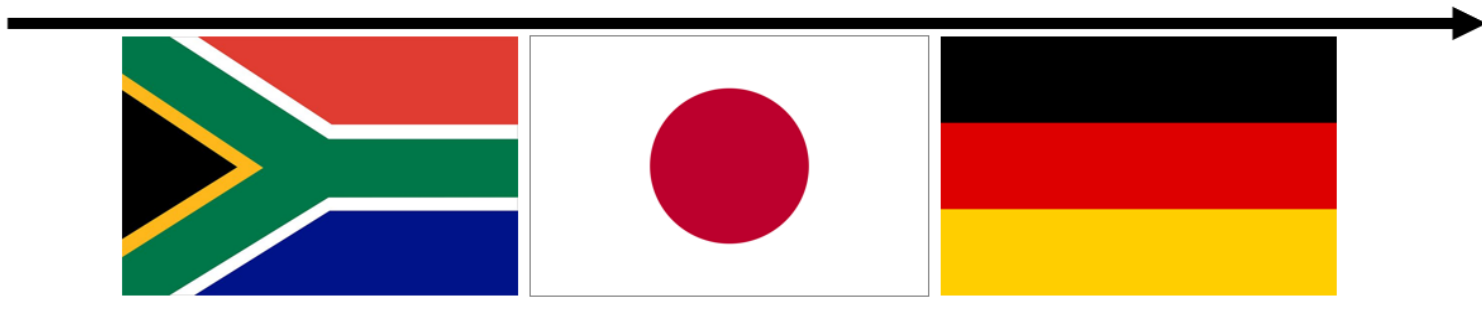
Glenda Coetzer[1], Marisa Nickola[1], Yu Takagi[2]

[1]*South African Radio Astronomy Observatory*
[2]*Geospatial Information Authority of Japan*

Abstract

The International VLBI Service for Geodesy and Astrometry (IVS) is committed to providing quality data and products in support of geodetic, geophysical and astrometric research and operational activities through good Research Data Management (RDM) practices. The use of Digital Object Identifiers (DOIs) can contribute significantly to making data and products generated by the IVS community Findable, Accessible, Interoperable and Reusable (FAIR) (Fig. 1). A DOI is permanent and can therefore always be used to locate the data object to which it refers. An exploratory study to determine best practices for attributing DOIs to datasets and products of the IVS community has been initiated. We report on progress with the investigation and provide recommendations.
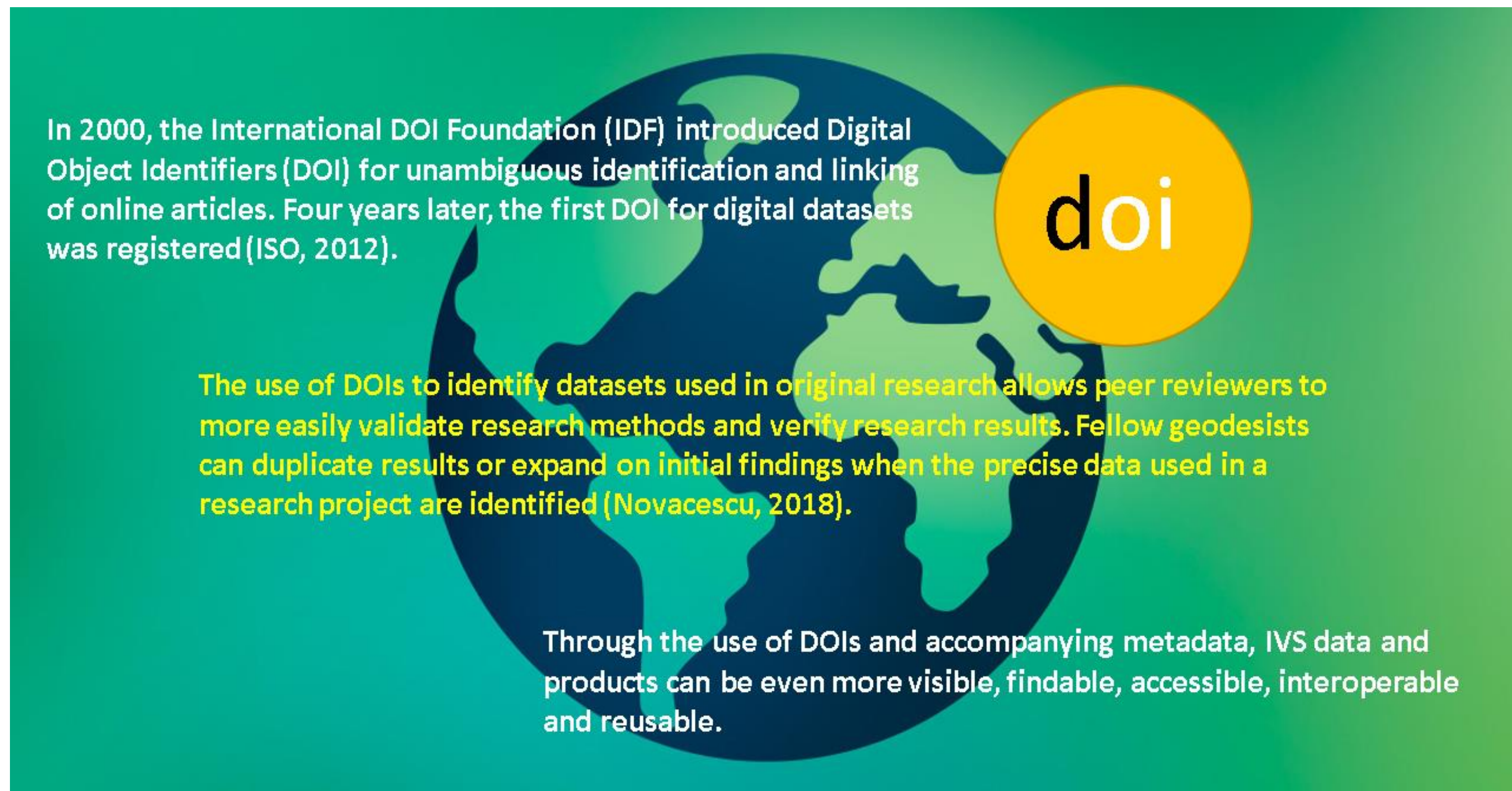
Introduction

In 2000, the International DOI Foundation (IDF) introduced Digital Object Identifiers (DOI) for unambiguous identification and linking of online articles. Four years later, the first DOI for digital datasets was registered (ISO, 2012).

The use of DOIs to identify datasets used in original research allows peer reviewers to more easily validate research methods and verify research results. Fellow geodesists can duplicate results or expand on initial findings when the precise data used in a research project are identified (Novacescu, 2018).

Through the use of DOIs and accompanying metadata, IVS data and products can be even more visible, findable, accessible, interoperable and reusable.

Fig. 1 Digital Object Identifiers (DOI) (Paskin, 2010; Wanchoo, 2017)

Anatomy of a DOI

A DOI looks like this without a resolver (URL):

DOI: 10.70024/Pqz1104tcZXP

A DOI looks like this with a resolver (URL):

https://doi.org/10.70024/Pqz1104tcZXP

Resolver | Prefix | Suffix

The resolver and prefix are provided by a Registration Agency (RA) and the suffix (internal unique opaque name for the archived data source) is provided by the client (Wanchoo, 2017).

Fig. 2 Anatomy of a DOI (Paskin, 2010; Wanchoo, 2017)

Growth of DOI Minting

Minting of DOIs have grown considerable since the concept was introduced in 2000. Fig. 3 shows the number of DOI resolutions per month over time; the yellow line shows the moving average (DOI Foundation, 2023).
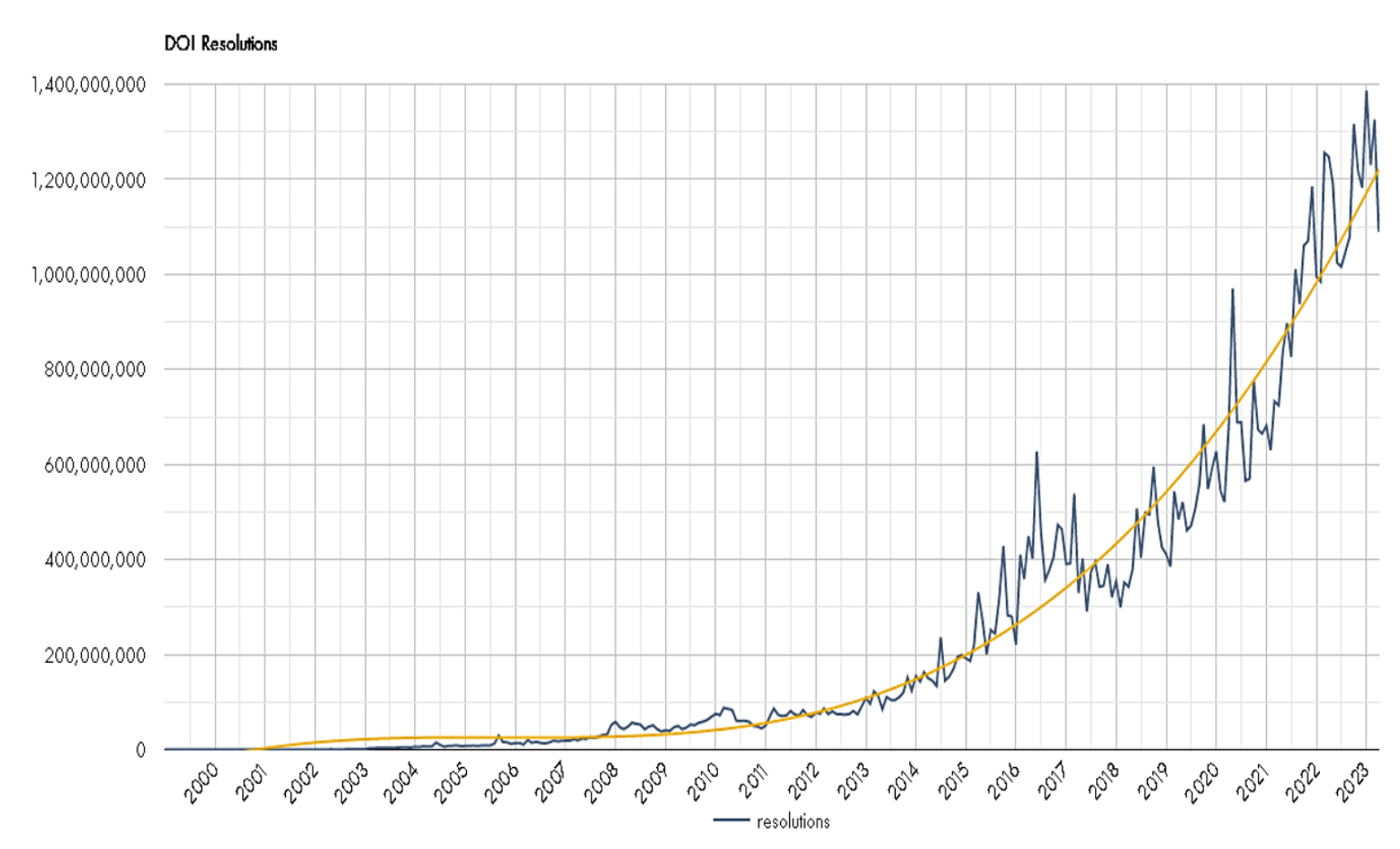

Fig. 3 DOI growth rate (DOI Foundation, 2023)

DOI Processing Workflow

Different approaches for processing a DOI can be followed (Wanchoo, 2017). An example of a DOI processing workflow is provided in Fig. 4.
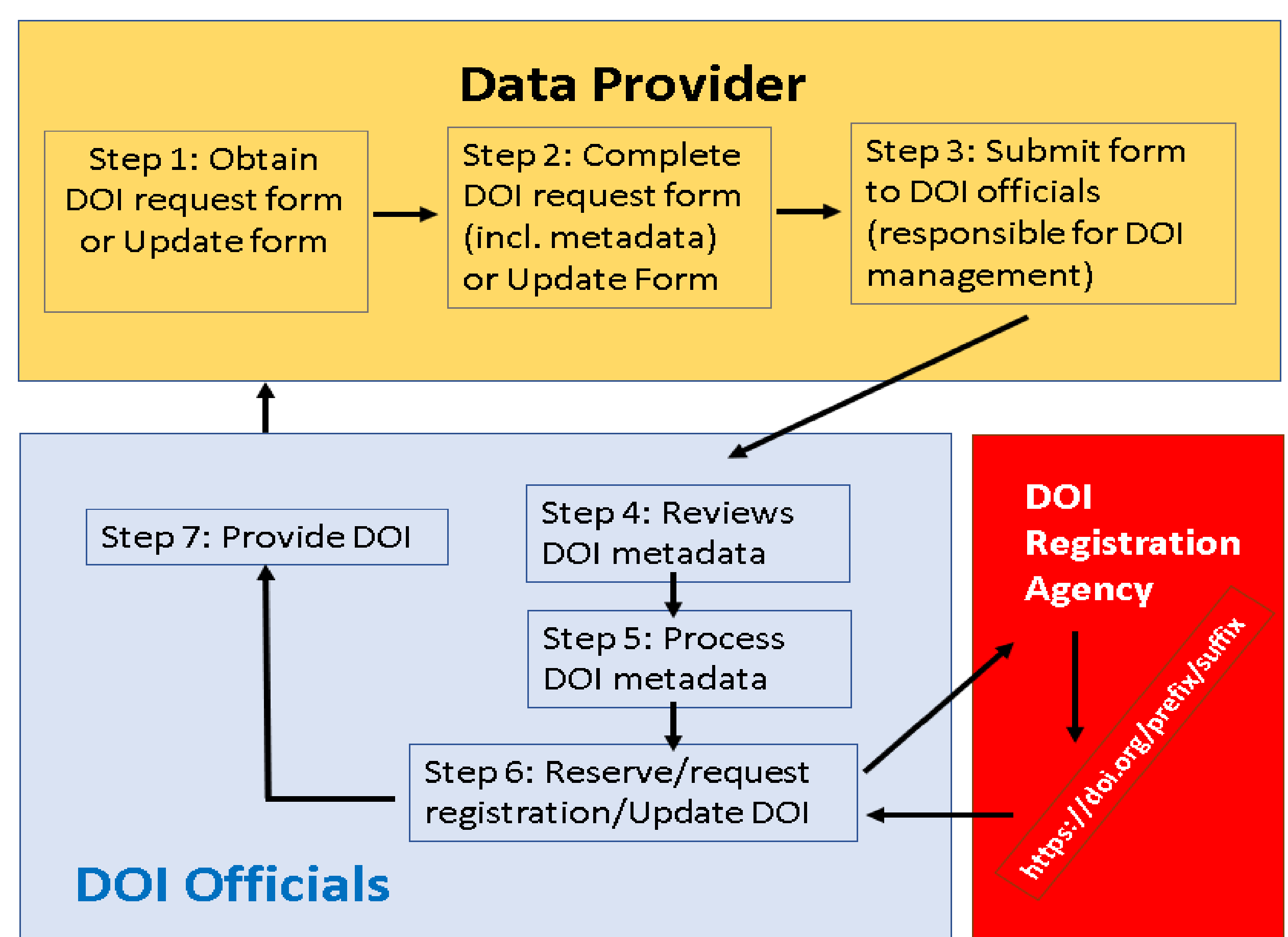

Data Provider

Step 1: Obtain DOI request form or Update form

Step 2: Complete DOI request form (incl. metadata) or Update Form

Step 3: Submit form to DOI officials (responsible for DOI management)

Step 7: Provide DOI

Step 4: Reviews DOI metadata

Step 5: Process DOI metadata

Step 6: Reserve/request registration/Update DOI

DOI Officials

DOI Registration Agency

https://doi.org/prefix/suffix

Fig. 4 Example of DOI processing workflow (Wanchoo, 2017)

Metadata for IVS Data and Products

Research data require metadata to be FAIR. Metadata have requirements for mandatory and optional elements (Fig.5). These elements define data for accurate consistent identification, description and citation.
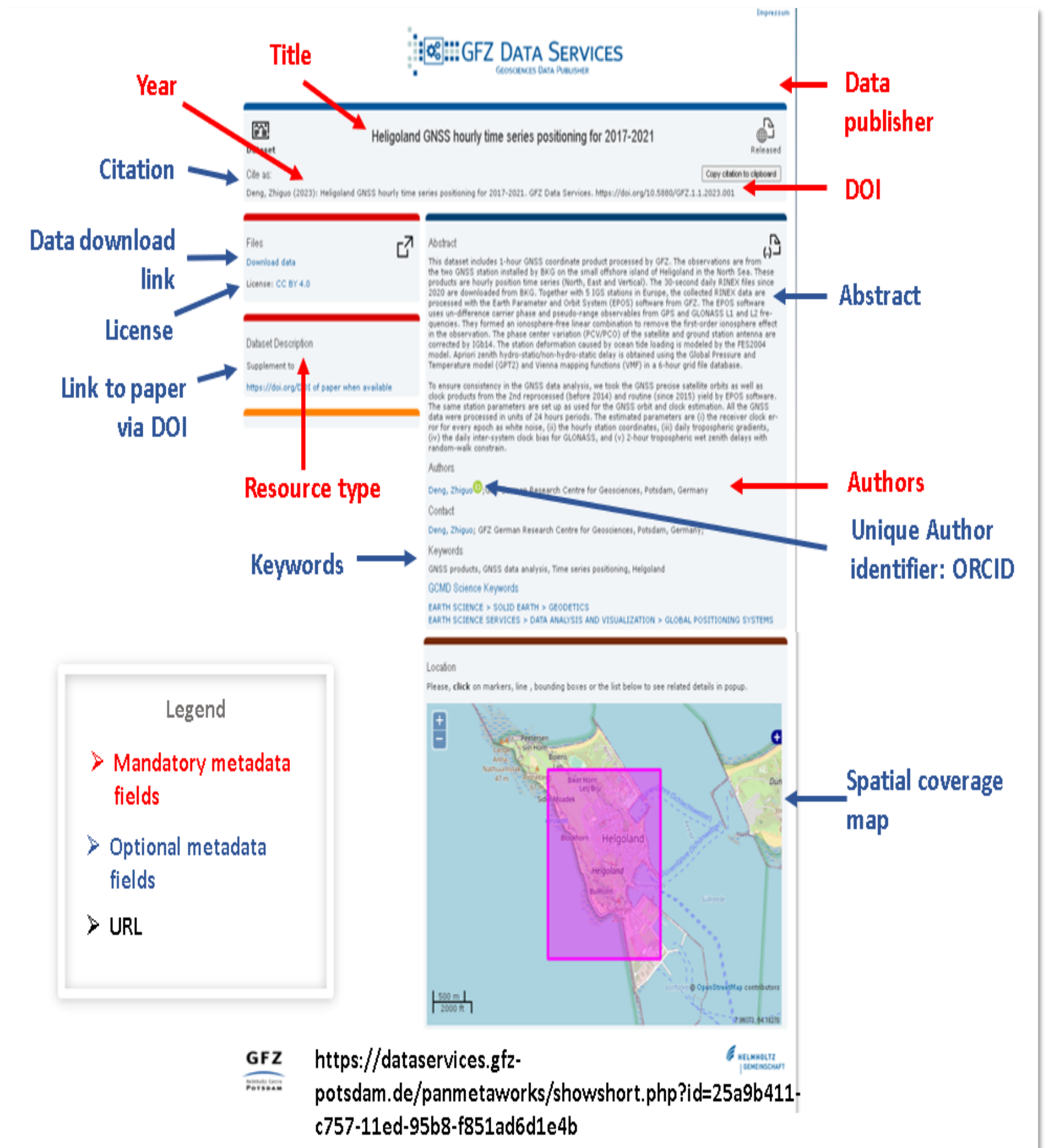

Fig. 5 Occurrence of mandatory and optional metadata attributes on a landing page (adapted from GFZ, 2023).

DOIs for IVS: Exploratory Study Timeline & Progress

- 2013 Heinkelmann et al. present "Making it Citable: Data in IVS" (21st EVGA Analysis Workshop, Finland)
- 2014 IVS Community compiles document "IVS Data Input to ITRF2014: Metadata Identification"
- 2020 First dataset DOI for IVS contribution to ITRF2014 minted by GFZ Potsdam, Germany
  DataCite and CrossRef for data citation mentioned at WDS IPO, Koganei, Japan
  IVS DB discussion on attribution of DOIs for IVS data and products
  Yu-san Takagi (GSI) and Glenda Coetzer (SARAO) investigate best practices for attribution of DOIs for IVS
- 2021 First ideas on best practices presented and high-level inventory of IVS components, data and products presented to IVS DB
- 2022 Recommendations for establishment of IVS task team/working group presented to IVS DB

Progress:

- Defined basic structure of IVS infrastructure, data and products
- Conducted inventory of all IVS data and products
- Determined we need to mint DOIs for products first

Recommendations

1. Establish task team/working group to manage minting of DOIs for IVS
2. Determine who will be responsible for the administration of DOI minting
3. Implement GeodesyML for geodetic datasets and products
4. Use controlled vocabularies for metadata
5. Include persistent identifiers (PIDs), e.g. DOIs, ORCID, Fundref, etc. as mandatory metadata elements
6. Harmonise DOI metadata across all data centre
7. Follow hierarchical approach for PID allocation, e.g. a single PID for an entire IVS infrastructure with several children-PIDs for telescopes, etc.
8. Mint IVS products for ITRF2020 before minting IVS data
9. First DOI Level 2 data - observations for ITRF2020 computations (conduct trial run for minting of L2 data and determine best practice)
10. IVS should be the descriptor in the DOI itself (IVS tagged in DOI, e.g. suffix)
11. Establish workflow for DOI minting
12. Create landing pages for DOIs
13. Investigate DOI automation

References

DOI Foundation, 2023. *DOI System exceeds 50 million assigned identifiers"*. https://www.doi.org/
GFZ. 2023. Heligoland GNSS. https://dataservices.gfz-potsdam.de/panmetaworks/showshort.php
Paskin, N. 2010.Digital Object Identifier (DOI) System. *Encyclopaedia of Library and Information Sciences*. 3rd ed. Taylor and Francis, pp. 1586–1592.
Wanchoo, L. et al.2017. NASA EOSDIS Data identifiers: approach and system. *Data Science Journal*, **16**: 15, pp. 1-11. DOI: 10.5334/dsj-2017-015