

ISDC Metadata Management

V. Mende, B. Ritschel, H. Palm, L. Gericke, S. Freiberg ⁽¹⁾

GeoForschungsZentrum Potsdam

Data Centre

14473 Potsdam, Germany

The Information System and Data Center (ISDC) is managing more than 10TB of geoscience data and information. By now, these data are coming from 11 missions² with nearly 300 product types, approximately 16 Mio products and more than 1500 users. One main part of the “WP173: Metadata, Information and Product Retrieval” consists of the metadata management for the retrieval of product independent information and data. This paper is giving a short overview about the developed and used metadata concept.

A product type of a geoscience mission or project consists of a set of products (data file(s) + metadata). This relation is shown exemplarily in figure 1 in respect to orbit products of the new satellite mission TerraSAR-X.

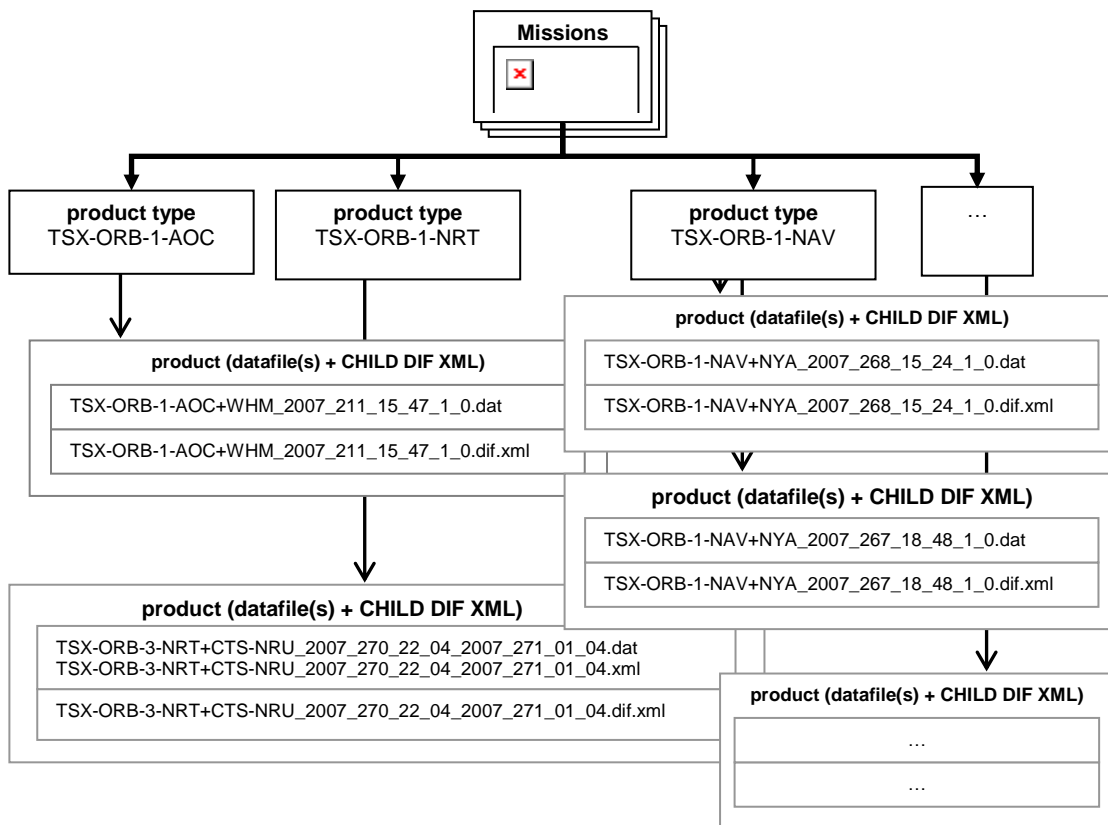


Figure 1: ISDC Relations

¹ Email: vmende@gfz-potsdam.de, rit@gfz-potsdam.de, palm@gfz-potsdam.de, lg@gfz-potsdam.de, sebast@gfz-potsdam.de

² CHAMP; GRACE, TerraSAR-X, GASP, GNSS, GGP, GPS-PDR and others

In order to describe and manage the data products we are using an evolution of the NASA's Directory Interchange Format (DIF) standard (version 9.x). The Global Change Master Directory (GCMD) defines metadata as: *"Descriptive information that characterizes a set of quantitative and/or qualitative measurements and distinguishes that set from other similar measurement sets."* (L. Olsen)

For the management of product types this format is excellent. The ISDC DIF conform base schema of the parent DIF XML documents is stipulated in the "base-dif.xsd" file. The ISDC XML schema has been defined on the basis of the GCMD XML schema definition: http://gcmd.nasa.gov/Aboutus/xml/dif/dif_v9.7.1.xsd. In order to describe single products, it was necessary to extend the DIF standard and to modify the GCMD XML schema. Even the structure of the ISDC DIF XML schema is different from the GCMD schema, the ISDC parent DIF XML documents are valid in relation to the GCMD schema. Additionally, the ISDC is using a parent – child DIF combination. The metadata of product types are described in associated parent DIF XML files according to the base-dif.xsd schema. The parent DIF XML files are validated and stored in an Oracle XML database. The product (data file) specific metadata are documented in child DIF XML files. Each product type has its own schema for the child DIF XML files. So, child DIF documents are used in order to describe the data file specific properties. The complex XML type <Data_Parameters> in the child DIF XML document provides the specific extension of the parent DIF XML structures. This includes specific metadata of the product, like data filename, data file size, revision, satellite ID and others. In order to realize this data model, we are using the redefine XML technique for the definition of complex XML types for the <Data_Parameters>. By redefining the ISDC "base-dif.xsd" schema all child DIF XML documents are derived. Using the GCMD XML schema, this approach would not be possible because of the definition of XML reference structures.

The extended metadata of the child DIF XML documents are stored in product type related tables in a relational DB. The connection between the child DIF XML files and the parent DIF XML document is given by the equality of parts of the <Entry_ID> element in both, the product type and the related product metadata documents. Additionally the content of the <Parent_DIF> element in the child DIF XML document refers to the appropriate parent DIF document. The relation between the schemata, the XML metadata files and the storage structures is shown in figure 2.

Using the parent DIF XML structures it is possible to realize a thematic search concerning the content of the different product type documents as well as to provide interoperability to other catalogue systems. Now It is possible to transform the XML DIF files into ISO 19115 standard documents in order to use OGC compliant Web Services like the "deegree CWS 2.0". Furthermore an efficient harmonisation with other catalogue systems can be realized by international standards. The structure of XML easily allows extending the DIF standard in future. Using the parent-child DIF concept, only a small amount of mandatory metadata must be put both in the parent and child DIF XML documents.

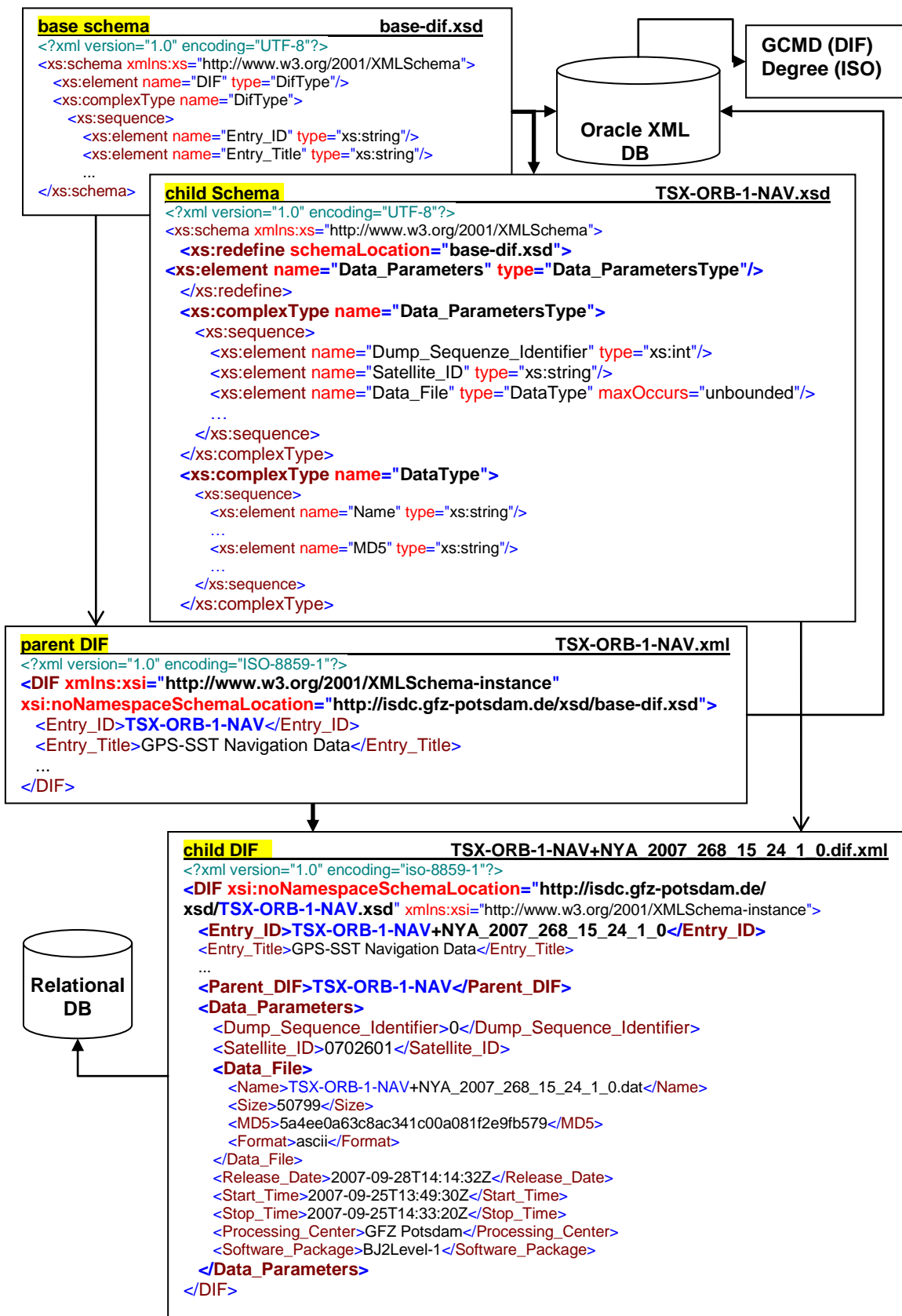


Figure 2: Schema, DIF XML and Storage

References

Deegree Project: <http://www.deegree.org>

Global Change Master Directory: <http://gcmd.nasa.gov/>
XML & XSLT: www.w3.org/TR/xslt , <http://selfhtml.org/>

[1] Braune, S., Czegka, W., Klump, J., Palm, H., Ritschel, B., Lochter, F. A. (2003): Anwendungen ISO-19115-konformer Metadaten in in Katalogsystemen aus dem Bereich umwelt- und geowissenschaftlicher Geofachdaten. - Zeitschrift für Geologische Wissenschaften, 31,1, 37-44

[2] Ritschel,B., Bruhns, C., Kopischke R., Mende V., Palm H., Freiberg S., Gericke L. The ISDC concept for long-term sustainability of geoscience data and information, PV 2007 Conference, Symposiums-Proceeding p.49, (Oberpfaffenhofen 2007)

[3] Voges, U., Senkler, K. (2005): OpenGIS® Catalogue Services Specification 2.0 - ISO19115/ISO19119 Application Profile for CSW 2.0.; OpenGIS Consortium, Wayland, Massachusetts

[4] Burgess, Ph., Palm, H., Ritschel, B., Bruhns, Ch., Freiberg, S., Gericke, L., Kase, St., Kopischke, R., Loos, St., Lowisch, St.,
Implementing modern data dissemination concepts in the ISDC Portal,
(GEOTECHNOLOGIEN Science Report ; xx), Status Seminar Erkundung des Systems Erde aus dem Weltraum““ (Oberpfaffenhofen 2006), xxx-xxx.