Paper for Consideration by S-101PT12

"File-less" resources management (including cancellations) for S-101

Submitted by:	PRIMAR
Executive Summary:	Proposal to define more clearly in the S-101 Product Specification documentation options for "file-less" published data management, including "file-less" cancellation.
Related Documents:	S-101 v.1.2.0 PS, S-100 v.5.1.0
Related Projects:	

Introduction / Background

To our understanding, and what we can find described in the S-101 Product Specification, it is agreed that at this time and those versions of the S-101 PS - v. 1.2.0 and v. 2.0.0, there is only one considered and well described in the S-101 documentation dataset cancellation implementation allowed, and it is file based cancellation.

Below are extracts from the S-101 v.1.2.0 Product Specification which supports this understanding:

11.2 Exchange Set

S-101 datasets are grouped into Exchange Sets. Each Exchange Set consists of one or more ENC datasets with a single Exchange Catalogue XML file containing metadata. Optionally, an associated XML Metadata file (S100_ResourcePurpose (ISOMetadata) – see S-100 Part 17, clause 17-4.5) may be included (see clause 11.5). It may also include one or more support files. The S-101 Exchange Set structure is the same as that described in S-100 Part 17, clause 17-4.1.

Units of Delivery:	Exchange Set
Transfer Size:	Unlimited
Medium Name:	Digital data delivery

Other Delivery Information:

Each Exchange Set has a single Exchange Catalogue which contains the discovery metadata for each dataset and references to any support files. See S-100 Part 17, clauses 17-4.4 and 17-4.5.

11.3 Dataset

11.3.1 Datasets

Four types of dataset files may be produced and contained within an Exchange Set:

- New dataset and New Edition of a dataset (Base dataset): Including new information which has not been previously distributed by updates. Each New Edition of a dataset must have the same name as the dataset that it replaces. A New Edition can also be ENC data that has previously been produced for this area and at the same optimum display scale. The encoding structure is located in Annex B – clause B-5.

- Update: Changing some information in an existing dataset. Each Update dataset file must have the same name as the original base cell file, with an extension number greater than or equal to

001 (see EEE in clause 11.3.2 below). They must not extend the geographical area covered by the base cell file to which they apply (see clause 4.5.2). The encoding structure for an Update is located in Annex B – clause B-6.

- Re-issue of a dataset: Including all the updates applied to the original dataset up to the date of the reissue. A Re-issue is intended to avoid unnecessary loading of the Base cell and all applicable updates individually for new users of the dataset, therefore does not contain any new information additional to that previously issued by updates, and can be issued at any time. The encoding structure is located in Annex B – clause B-5.

- Cancellation: The dataset is cancelled and is deleted from the system. The encoding structure for a Cancellation file is located in Annex B – clause B-7.

Annex B - B-7 describes only a file based cancellation and the cancellation file's required structure:

B-7 Dataset Cancellation Structure

```
Dataset cancelation file
    |
    |--<1>- Dataset General Information record
        |
        |--<1>-DSID (13\\*1): Dataset Identification field
```

B-7.1 Field content

B-7.1.1 Dataset Identification field - DSID

| Subfield name	Label	Value	Format	Comment
Record name	RCNM	{10}	b11	{10} – Data Set Identification
Record identification number	RCID	{1}	b14	Only one record
Encoding specification	ENSP	"S-100 Part 10a"	A()	Encoding specification that defines the encoding
Encoding specification edition	ENED	"5.1"	A()	Edition of the encoding specification
Product identifier	PRSP	"INT.IHO.S-101.1.2"	A()	Unique identifier for the data product as specified in the product specification
Product edition	PRED	"1.2"	A()	Edition of the product specification
Application profile	PROF	"2"	A()	"2" – Update dataset profile
Dataset file identifier	DSNM		A()	The file identifier including the extension but excluding any path information
Dataset title	DSTL		A()	The title of the dataset
Dataset reference date	DSRD		A(8)	The reference date of the dataset Format: YYYYMMDD according to ISO 8601
Dataset language	DSLG	"EN"	A()	The (primary) language used in this dataset
Dataset abstract	DSAB	omitted	A()	The abstract of the dataset
Dataset edition	DSED	"0"	A()	0 – Indicates the cancellation
Dataset topic category	*DSTC	{14}{18}	b11	A set of topic categories

Further are extracts from S-100 v.5.1.0 regarding the Resource-less published Resources management through the Exchange set:
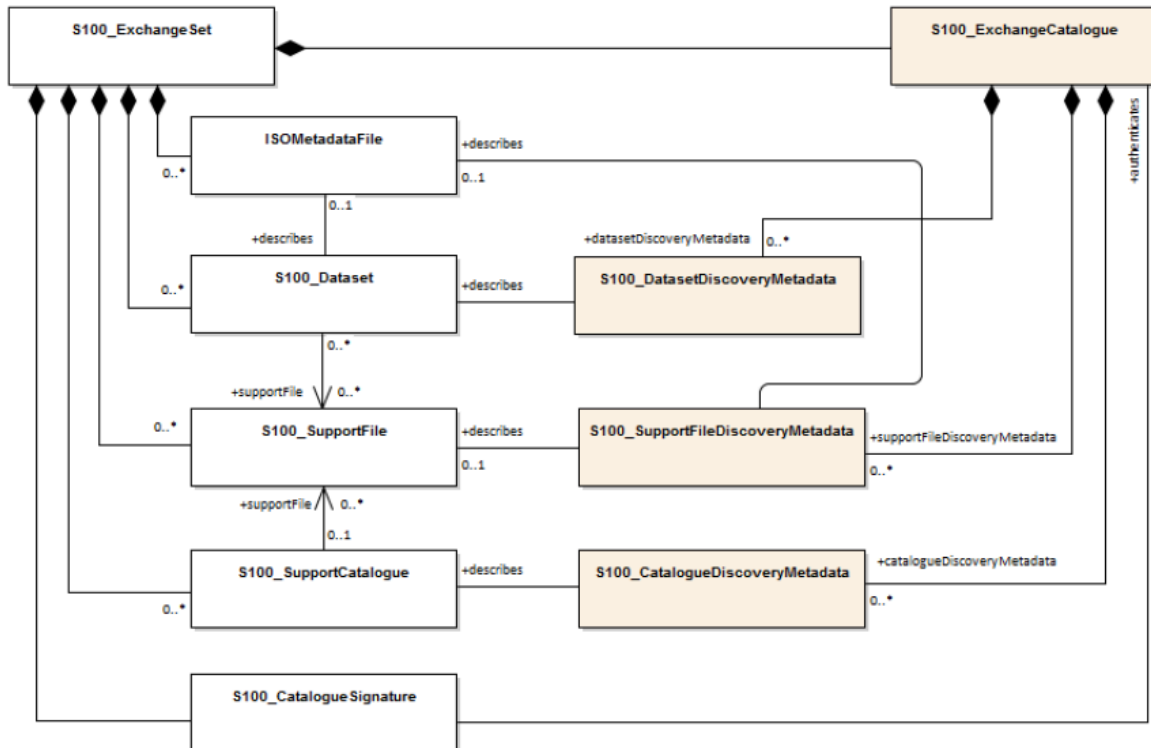
Figure 17-2 – S-100 Exchange Set

The conceptual model depicted in Figure 17-2 is very flexible and can be implemented in a variety of ways as virtually all components, except for the S-100_ExchangeCatalogue, are optional. This level of flexibility is essential to properly support the mainstream use case of exchanging geospatial data, as well as the use cases for releasing dataset and support file cancellation notices or new Catalogue releases without any data files present.

This approach ensures that an Exchange Set Catalogue is always included in any S-100 conformant Exchange Set, providing the essential discovery metadata about any included resources and their intended use.
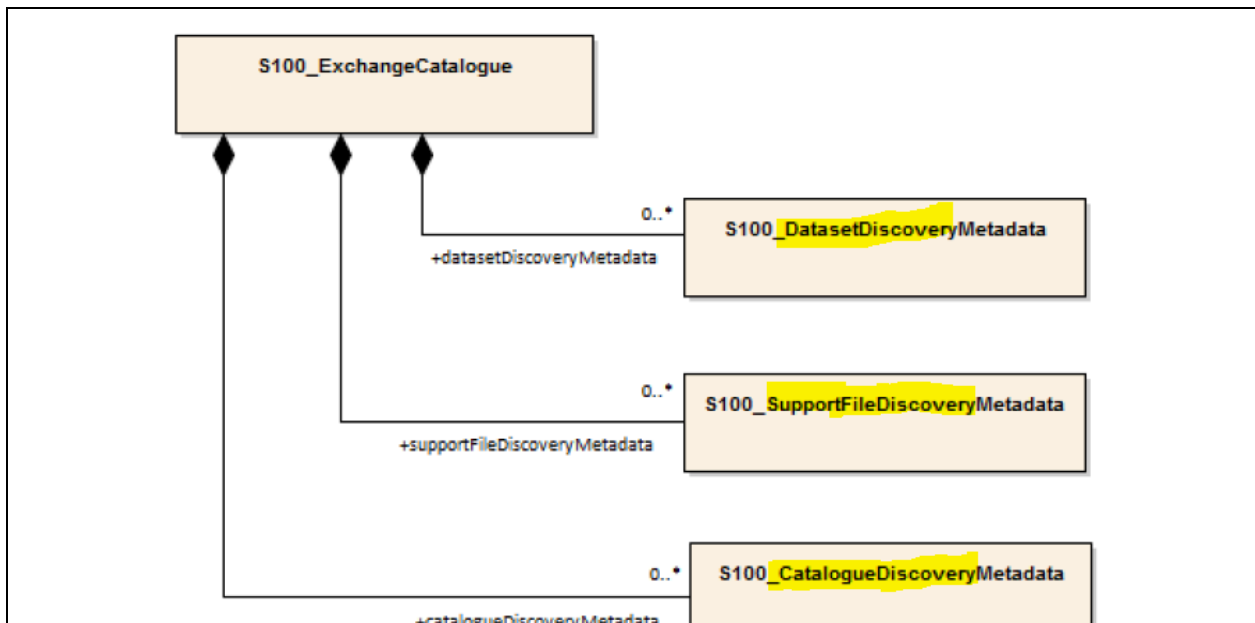
Figure 17-6 – S-100 Exchange Set Catalogue

The discovery metadata subsections have attributes which enable important information about the datasets and accompanying support files to be examined without the need to process the data, for example encryption/compression flags. Similarly, other Catalogues can be included in the S-100 Exchange Set, in support of the datasets, such as feature, portrayal, coordinate reference systems, codelists etc. In addition, the S100 Exchange Set Catalogue provides mechanisms for managing the lifecycle of records, support resources and catalogues. For example, the S100_SupportFileRevisionStatus and S100_Purpose enumerations support a revision control mechanism not only for delivering new versions and revisions, but also for cancelling such resources. This provides the ability to cancel records, support resources and catalogues using the S100 Exchange Set Catalogue records, rather than publishing incremental versions of the actual resources.

Analysis/Discussion

Reading S-101 PS document clauses 11.2, 11.3.1, 11.3.3 and Annex B, B-7 it is clear how to cancel a dataset with an update file included in the Exchange Set.

Nowhere in the S-101 PS document it is mentioned that the S-101 PS either supports "resource-less" dataset or additional resources management or not, including cancellations. Based on that and according to S-100, it can be assumed that the "file-less" resources management is also allowed in S-101 and can be implemented.

There are still many questions towards the S-100 standard itself, how the "resource-less" resources management should work in general. There is expected more development on that before it can be implemented properly regarding the ISO standards.

We have concerns that an implementation of the above in the Product Specification level at this time is very risky. Referring here to the IHO CL43/2023. If such an implementation is developed and products produced, then it needs to be supported by any other implementation throughout the production and supply chain. With almost no descriptions in Product Specification documentation at this stage, it is a high risk that the implementation is not stable or well thought through to work as intended at each stakeholders implementation for the various resources.

To clarify that the decision not to use "file-less" data management at this stage in S-101is intentional, we propose to add a text to the clause 11.3.1 clarifying that for the current edition of S-101 PS the "resource-less" or "file-less" data management is not allowed. This would clearly state that S-101PS does not allow for the current version one of the two available S-100 options of the published resources management.

As a side note, it would be good to receive confirmation from the Group, that the S-101 PS should or can restrict only the S-101 datasets and it's linked Support files (TXT and TIF). And does this then mean or is clear for any implementation that the S-101 FC, PC and any other general Support files for S-101 to work properly, would still be allowed to use both allowed resources management options described in S-100 standard.

Conclusions

Based on the above it is proposed to add the following text, along with any other agreed edits to it from the discussions during the meeting, to the clause 11.3.1 "Cancellation" bullet in the S-101 PS (yellow dashed line in the S-101PS excerpt above).

> "*Published S-101 base datasets and their textual or pictorial support files "file-less" management (including cancellations and deletions) is not allowed in S-101 Product Specification for this edition of S-101. That is - producing the S-100 Exchange Set which only includes the mandatory file "Catalog.xml" with the dataset and/or support files discovery metadata without including the appropriate dataset, update or support file in the S-100 Exchange Set, is not allowed.*"

Action Required of S-101PT10

The S-101PT10 is invited to:

Discuss proposed clarification and agree on the text to be added to the Product Specification.