

Paper for Consideration by the S-100 Working Group

General Data Model in S-100

Submitted by:	SPAWAR Atlantic
Executive Summary:	S-100 needs a general data model.
Related Documents:	S-100 Version 3.0 Specification
Related Projects:	IHO S-100/S-101 Test Bed Project

Introduction / Background

In the latest iteration of the SPAWAR S-100 Viewer, the team re-examined the specification from the implementer's perspective and attempted to update the Viewer to be as strictly conformant as possible. This exercise revealed several issues that were not apparent from a simple read-through of the specification. One of the most significant issues turns out to be a complete lack of a general data model.

Analysis / Discussion

A fundamental tenant of S-100 is that it contains all the necessary parts to create a product specification. The corollary to this tenant is that a product specification is *entirely* composed of parts from S-100, with no part of the product specification extending beyond S-100. In practical terms, this means that software can be written to fully support every S-100 based product merely by implementing the S-100 specification itself.

Product specifications use two methods to constrain S-100 for its needs. The first method is to only include the parts of S-100 that are needed (all other parts are implicitly excluded). The second method is to use the Feature Catalogue to explicitly define a set of allowable feature types. While some might view the Feature Catalogue as a means for extension, it is actually a means of restriction. S-100 defines the concept of feature type, but does not restrict what feature types are allowed or what attributes and associations any feature type might contain. The product specification uses a Feature Catalogue to put explicit constraints on the allowable feature types, as well as the allowable attributes and associations for those feature types.

The issue that we encountered was that product specifications were *implicitly extending* beyond what S-100 itself allows. This occurred, not because the product specification required more functionality than S-100 provides, but because S-100 inadequately provides for a few key components required by all product specifications. To illustrate this, the remainder of the paper will use S-101 to highlight the issue. It is important to note, however, that this issue affects *all* product specifications.

For this discussion, S-101 restricts S-100 as follows:

- States that only the ISO 8211 interchange format is used (HDF5 and GML are implicitly excluded)
- Defines a Feature Catalog that restricts the specific feature types (information types, associations, etc.) that may be within the dataset

Starting with the interchange format restriction, S-101 makes use of S-100 Part 10a (ISO 8211 encoding). Part 10a defines how certain S-100 concepts (features, information, attributes, etc.) are encoded in an ISO 8211 file. Part 10a *does not* provide any concrete translation rules between the ISO 8211 and any other part of S-100. Instead, it only relies on implicit connections between similar concepts. For example, Part 10a defines a "Feature Type Record" and other parts of S-100 also define a "Feature Type". However, there is no explicit relationship defined.

Next, S-101 makes use of S-100 Part 5 (Feature Catalogue) to constrain feature instances to specific types, along with specific attributes and associations (which are themselves constrained to specific types). In order to define these restrictions, Part 5 assumes a common relationship between concepts like features, information, and attributes.

However, the ability for Part 5 to define the language for these constraints is dependent on the similarity of terms to other parts of S-100. For example, Part 5 defines “Feature Type”, but its use of “Feature Type” is only implicitly connected to the same term used in other parts of S-100.

Up to this point, we have avoided any mention of S-100 Part 3 (General Feature Model and Application Schemas). The General Feature Model is purported to provide the necessary connection between the other parts of S-100. However, the General Feature Model is more abstract than this. Instead, it defines *the means* to define those connections. In other words, the General Feature Model is a *meta-model*, one level of abstraction removed from what is needed to address the above issue. To define the actual connections (make the connections explicit), a product specification is required to create an Application Schema, based on the General Feature Model. The Application Schema defines a logical data model that makes explicit the relationships between Part 5, Part 9, Part 10a (and other parts).

However, the Application Schema violates the tenant discussed at the beginning of the paper. Because each product can define its own logical data model, it means that there is no practical way for applications to be developed for S-100 in general. The Application Schema (logical data model) must be individually implemented for each product specification. Additionally, because each product can have its own logical data model, it means that there is no ability to define a generalized interoperability mechanism. Instead, the interoperability requirements between two or more products specifications must also be individually implemented.

Recommendation

As the specification currently stands, it is impossible to implement a truly general-purpose S-100 application. To enable this capability, a fundamental change to Part 3 is recommended. Instead of using the General Feature Model and product-specific Application Schemas; S-100 should define a General Data Model.

The purpose of the General Data Model is to define concrete relationships between all of the objects defined by S-100. This model is basically the same as the model that is already *implied* by individual parts of S-100. An explicit General Data Model provides consistent, product-independent rules for:

- Mapping between interchange files and datasets
- Mapping datasets to portrayal input
- Enforcing restrictions defined by a Feature Catalogue

Additionally, a General Data Model is necessary for

- The implementation of a common interoperability specification
- The implementation of a common alarms and indications specification

The General Data Model does not, however, mandate a particular implementation in software. It is still up to the implementer to decide how to represent the General Data Model in code. All that the General Data Model ensures is that every implementation is working with the same explicitly-stated expectations.

With the General Data Model, Application Schemas are not necessary. The Feature Catalogue already places all necessary restrictions on the General Data Model for a given product specification. This also ensures that the Portrayal Catalogue rules can be safely written for a product without having to take the entire General Data Model into account.

Action Required of S-100 WG

The S-100 working group is invited to:

- a. note the paper
- b. discuss the recommendations
- c. The S-100 working group is invited to approve the formation of a team that will develop an update for Part 3 of the S-100 specification, as described in this paper.