**Input paper: [[1]](#footnote-1)** ENAV30-5.1.1.4

**Input paper for the following Committee(s):** **Purpose of paper:**

(Select as appropriate)

ARM  ENG  PAP  Input

ENAV VTS  Information

**Agenda item** [[2]](#footnote-2) 5.1

**Technical domain/ Task number** 2 WG1

**Author(s)/Submitter(s)** DLR, GLA

Harmonization of S-100 Dataset Handling in Maritime Services

# Summary

In the past years, the development of many new product specifications of the IHO S-100 family had been initiated, for example, S-124 for representing navigational warning data or S-201 and S-125 for AtoN information. These data models are envisioned to be utilized in the different Maritime Services, defined by the IMO (cf. IMO MSC.1/Circ.1610). In the IALA committees, different task groups started to work on the development of technical service specifications (in accordance with IALA G1128) related to VTS and AtoN information. Additionally, a task group working on a specification of a service for the provision of navigational warnings continuously reported its activities to IALA.

A common problem encountered in these groups is the handling of S-100 data sets in combination with the dynamic communication behaviour of the technical services. Currently, there is no guidance available on the harmonization of interfaces and implementations of these services regarding S-100 data set handling.

## Purpose of the document

The purpose of this input paper is to identify open issues in the harmonization of technical service specifications of S-100 based technical services. Furthermore, the committee is requested to consider using the information provided in this document to update IALA’s technical guidance documents (specifically IALA G1157) with recommendations on how to implement S-100 data set handling (see section 5).

## Related documents

MSC.1/Circ.1610 - INITIAL DESCRIPTIONS OF MARITIME SERVICES IN THE CONTEXT OF E-NAVIGATION

IHO S-100 Universal Hydrographic Data Model V4.0.0

IALA G1157 - Web Service Based S-100 Data Exchange

IALA G1128 The Specification of e-Navigation Technical Services

EC 63173-2:2022 Maritime navigation and radiocommunication equipment and systems - Data interfaces - Part 2: Secure communication between ship and shore (SECOM)

# Background

As described in the S-100 standard, S-100 data is typically organized in data sets, that can include several instances of real-world entities, and meta-information. Existing data management systems utilize S-100 data sets to organize entities by specific properties such as their spatial geometry (e.g., all AtoN information of a specific geographical area is contained in a single data set for this area). On the other hand, technical services that make the data available to end-users provide the functionality to query for specific data (which will then be delivered as a “user-defined” data set). To this point, it is still not defined how services should deal with possible discrepancies between queried data and data in existing data sets.

SECOM is a standard dealing with the web-service based exchange of S-100 data. It includes a list of predefined query parameters (geometry, time period, UNLOCODE, …) to be used in technical services to filter the provided data. However, it is not evident, how the filtering procedure should be implemented. For example, if there is a match of geometries with only a subset of the entities in the original data set, it is not clear whether the complete original data set or only a subset of the data set should be returned. Furthermore, a match is not defined unambiguously: Is it enough if the filter geometry intersects the geometry of a data set entity, does it have to include the geometry, or is query filtering only based on the metadata of a data set? Also, if only a one-dimensional filter geometry is provided (e.g., a route), it is not defined if data entries or data sets at a certain distance (e.g., < 15 nm) should also be included in the results.

Finally, some of the S-XXX data models include the possibility to model (delta-)changes in data (e.g., removal or update of a navigational warning). Currently, there exists no guidance on how this should be implemented in services with regard to subscriptions or queries that include a time period filter.

The harmonization of these service properties is necessary to avoid confusion in the implementation and usage of the services.

## Example

Looking at an example of a service for the provision of AtoN information, it might be the case that this service gathers data from three different local authorities and manages them in three S-100 *data sets* as shown Figure 1. Each of the data sets can contain multiple *entries*, which represent the actual AtoNs . In this case, a data set only includes AtoN information for a specific geographical area (German Bight, Elbe River, Weser River).

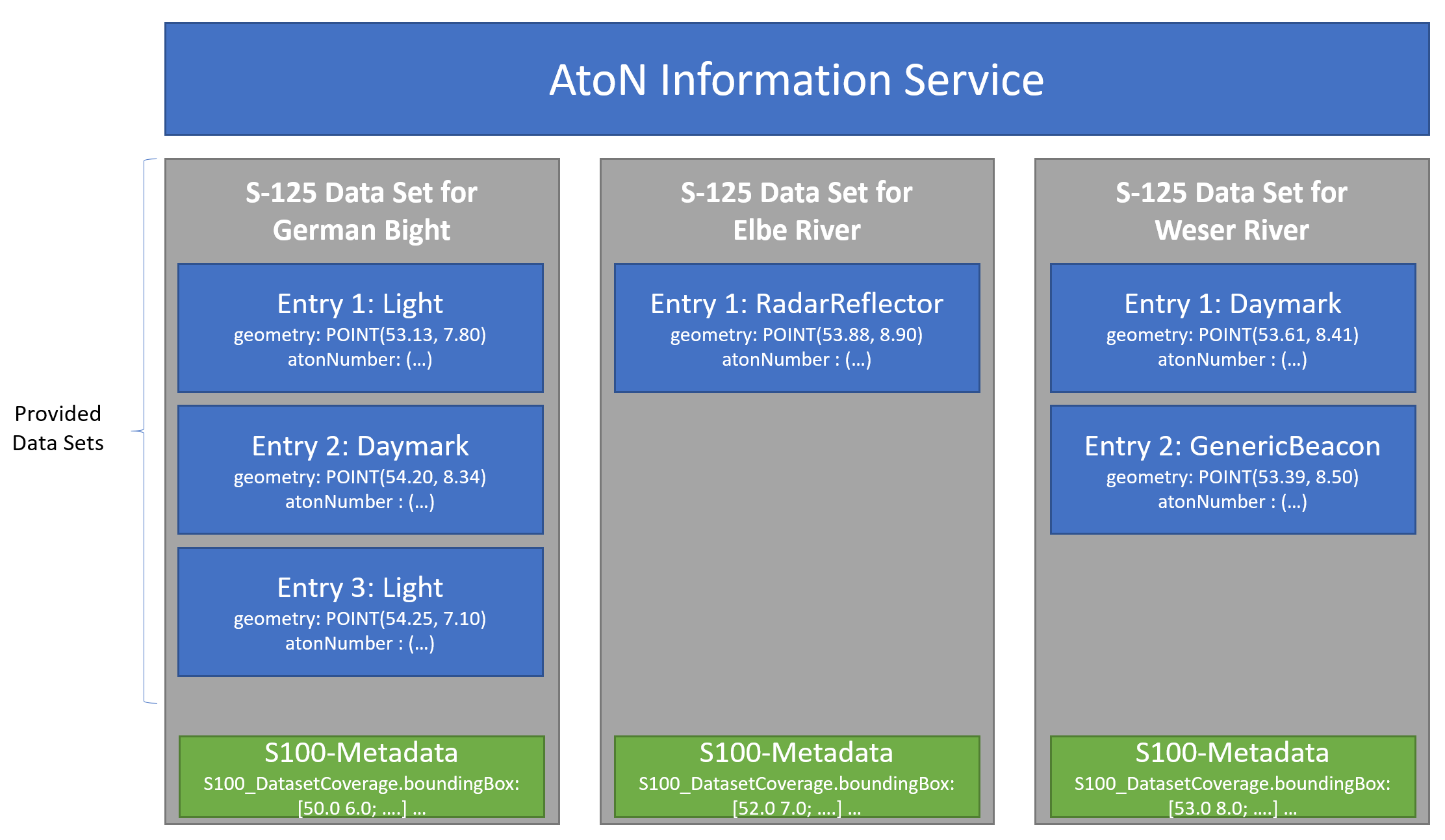


Figure 1: Example Data Sets of an AtoN information service.

A mariner on a vessel may only want to retrieve AtoN information along a planned route from the Weser River to the Elbe River. Requesting the service for AtoN information with the given route may lead to several questions for the implementation of the service:

* The route is a list of waypoints in the Elbe and the Weser. How to tell which AtoN data is “along the route” as all the AtoNs have point geometries, not directly colliding with the geometry of the route?
* Given that the mariner needs some data from the Elbe and Weser Rivers, should the service return the two complete data sets? If some of the AtoNs in the Weser are too far away from the route, should the service generate a new data set with only the relevant data? How is the generated data set distinguished from the original data set, does it contain different metadata?
* If there is an update to Entry 2 in the Weser River data set, will the mariner (if he has subscribed to the service) receive a new version of the data set, or only an update of Entry 2?

These exemplary questions may occur slightly different in use cases of other services but deal with the same problems (as described in the previous section). The easiest solution for the service architects would be to define these properties in their technical design specification. However, this would lead to many, slightly different client implementations at the service consumer side (e.g., in an ECDIS). In the worst case, an AtoN information service may return data for a different spatial area than a NW service with datasets for the same geographical areas, when given the same route. This can result in dangerous confusion and misinterpretation of data.

# Discussion

Therefore, the following guidelines are being proposed to harmonize the implementation of maritime services regarding S-100 data set handling.

## Data Set Generation

In comparison to the static organization of S-100 data sets in the classical sense, services might need to dynamically generate data sets based on query parameters that are supplied by a service consumer. The dynamic generation of data sets increases efficiency of communication as a service provider would only receive the requested data. It can also support a less cluttered portrayal of information as irrelevant features are not retrieved. However, this depends on the service and may not be applicable to all services (e.g., ENC provision).

**Considerations for services that do not utilize dynamically generated data sets:**

* If no query parameters are supplied in data request by the service consumer, the service should either return all original data sets or an empty response (if querying all data once is not an intended functionality of the service)
* If query parameters are specified, they are interpreted on a per data set basis. This can happen on the metadata level via the S100\_DataSetDiscoveryMetadata or via a custom analysis of the data entries. Either way, if there is a match with the metadata or an entry of a dataset, the whole data set must be returned in its original version.

**Considerations for services that utilize dynamically generated data sets:**

The original data at the service provider side may be organized into several basic data sets (e.g., a data set may contain AtoN information for a specific waterway). In the following, we refer to these data sets as *original data sets*. An original data set typically contains multiple entries or features. To harmonize the handling of dynamically generated and original data sets, we propose the following requirements:

* If query parameters are supplied that restrict the original data sets to only subsets, a single generated data set with all the matching entries of the original data sets should be provided.
* If query parameters and a reference to an original data set is supplied, a single generated data set with the matching entries of the referenced data set should be provided. This generated data set must have a different identifier as the original data set.
* If the service provider does not organize data in different original data sets, all matching data entries will be provided as a single generated data set.
* The derived metadata in a generated data set must be consistent with the data entries.
* The service provider must make sure that the signatures of the data are valid for the generated data sets.
* If not specified otherwise in the service specification, the service consumer is responsible for the management of generated data sets. This means that he needs no keep track of which data he already requested and which data may needs to be requested when entering a different waterway/area etc.

In addition, it should be considered whether a distinction between the original data sets, and the ones generated on runtime by user requests, is desirable. A differentiation between the two (if any), can be defined at the technical service specification level, but a harmonized way of dealing with them is strongly advised. For example, a specific keyword in the S-100 data set file identifier or title could be employed.

## Query Parameters

To avoid confusion regarding the results of data queries, the following requirements for the interpretation of query parameters are proposed:

* If not defined otherwise in the service specification, a service should return all entries/data sets that *interact* with a supplied geometry parameter. This refers to all entries with geometry that are not disjoint with the provided filter geometry (cf. OGC 09-026r2, section 7.8.3).
* Dataset entries/data sets may only be valid in a specific time range. If this is the case, and it is not specified otherwise in the service specification
  + a service should only provide those entries/data sets, which are valid after a *validFrom* parameter. In the case that data entries/data sets are valid until cancelled by another entry / data set, all of those that were valid after *validFrom* and were not cancelled should be returned. A *validFrom* parameter in the future will only be accepted if data entries/data sets have explicit values for their validity period.
  + a service should only provide those entries /data sets, which are valid till a *validTo* parameter. In the case that data entries/data sets are valid until cancelled by another entry/data set, a *validTo* parameter will not be accepted.
* For querying data by specific S-XXX data models, query parameters should use the format “SXXX” as defined in SECOM, section 5.6.7.

**Considerations for services that utilize dynamically generated data sets:**

* Dataset entries may only be valid in a specific numeric range. If this is the case, and it is not specified otherwise in the service specification
  + a service should only provide those entries in a generated data set, which have values larger than the start value of the range.
  + a service should only provide those entries in a generated data set, which have values smaller than the end value of the range.

## Changes in Data Sets

Several services need to provide data that is continuously extended, modified or removed. The following requirements should harmonize the way services deal with changes in data sets.

* If a service consumer subscribes to a service, he is responsible to pull the current version of the required data. The subscription does only provide updates if the queried data sets change.
* The service provider must keep track of the query parameters of a subscription request and only provide updates to data entries/data sets that match the subscription query.

**Considerations for services that utilize dynamically generated data sets:**

* Updates in data sets are often done per data set entry. This means that, that anytime data entries are updated, a data set (containing only the updated entries) is generated and sent to the subscribers of those entries.
* Deletions of data entries should be made available to the service consumers that have subscribed to a relevant dataset. Therefore, unless it is not specified otherwise in the service specification, the employed data products should support a way of referencing deleted data entries.

# Action requested of the Committee

The Committee is requested to:

1. Note this input paper and consider revising IALA G1157 on Web Service Based S-100 Data Exchange, adding additional guidance on the discussed matters.
2. If necessary, provide input to IEC to request clarification of the discussed matters in EC 63173-2:2022 or to the IHO S-100 WG to consider updating the standard.

1. Input document number, to be assigned by the Committee Secretary [↑](#footnote-ref-1)
2. Leave open if uncertain [↑](#footnote-ref-2)