

[Non-DoD Source] S-102 Dataset file naming

Svein Skjæveland <svein.skjaeveland@ecc.no>

Fri 2/19/2021 9:43 AM

To: Haselmaier, Lawrence H Jr CIV USN NAVOCEANO SSC MS (USA) <lawrence.haselmaier@navy.mil>

📎 1 attachments (3 MB)

S-102PT7_2021_3.1_S-102_Ed2.1.0 - draftRedline_7Cs_BSH.doc;

Hello Lawrence,

I have a suggestion related to the chapter 11.2.3 Dataset file naming that could resolve the confusion pointed out by Germany:

11.2.3 Dataset file naming

Dataset naming must follow a standard pattern to give implementers greater predictability of incoming datasets. S-102 dataset naming conventions must follow these rules.

102PPPP00000000000.H5

- 102 - the first 3 characters identify the dataset as an S-102 dataset (mandatory).
- PPPP - the fourth to seventh characters is the producer code according to the IHO Geospatial Information Registry, Producer Code Register (mandatory for S-102).
- 000000000000 - the eighth to the maximum nineteenth characters are optional and may be used in any way by the producer to provide the unique file name. The following characters are allowed in the dataset name: A to Z, 0 to 9 and the special character _ (underscore).
- H5 - denotes and HDF5 file.

DR Daniel Rohde
 What is the filling character?
 Because the producer code has only two characters.

The reason for using a 4 character production code relates to decisions made in previous S100WG meetings where the future expansion of S-62 (List of IHO Data Producer Codes) has been taken into consideration. This is also supported by the Dataset naming rules laid out in S-97 (Guidance for Product Specification Developers):



S-101 Product Specification has a bullet I think we could use also in S-102:

11.3.2 Dataset file naming

101CCCC0000000000.EEE

The file name forms a unique S-101 identifier where:

- 101 - the first 3 characters identify the dataset as an S-101 dataset (mandatory).
- CCCC - the fourth to seventh characters identify the producer code of the issuing agency (mandatory for S-101). Where the producer code is derived from a 2 or 3 character format (for instance when converting S-57 ENCs), the missing characters of the producer code must be populated with zeros ("00" or "0" respectively) for the sixth and seventh characters of the dataset file name, as required.
- 0000000000 - the eighth to the maximum seventeenth characters are optional and may be used in any way by the producer to provide the unique file name. The following characters are allowed in the dataset name: A to Z, 0 to 9 and the special character _ (underscore).
- .EEE – new datasets and new editions use 000, updates start at 001 and increment until a limit of 999 (mandatory). Re-issues use the same number as the last Update applied to the dataset. Cancellations use the next sequential number from the previous Update applied to the dataset.

The suggestion would be to replace the second bullet in the 102 spec with the second bullet from the 101 spec.

I believe this would make it more clear how to encode the sixth and seventh characters and resolve the confusion.

Attached the redline with my suggestion added to Germany's comment.

Have a nice weekend.

--

Best regards,

Svein S. Skjaeveland
Manager International Standardization
ECC AS - Primar Service Provider

Tel: +47 51 93 95 00
Tel: +47 92 66 26 97

www.ecc.no
www.primar.org