

Paper for Consideration by S-S-100WG

PRIMAR view on producer codes

Submitted by:	PRIMAR
Executive Summary:	PRIMAR views on S-100WG2_8.5_S-100_ProducerCodes
Related Documents:	S-100WG2_8.5_S-100_ProducerCodes_Rev1-1.pdf
Related Projects:	

Analysis/Discussion

When considering file based data exchange distribution, product identification and administration is crucial. When exchanging large amounts of data we need solutions for administration of the products and their versioning.

As of today, we have a functional model for naming ENC's in a way that makes product identification and administration manageable for system and ECDIS developers, producers, distributors and end users. This model has been evolved and tested well during many years of ENC production and distribution. One important aspect is that apart from immediately giving producer/scale band information, the file names are also used as product IDs in data distribution services. The current naming conventions support effective file management by humans and computers without the need to read and interpret the content of catalogue files. The current file naming convention has its origins back to when MS-DOS was the dominating operating systems on computers and it has a limitation on filenames to 8+3 characters.

It is of great importance that the products have unique filenames when handling data from different producers. The proposal being discussed would open for the possibility that different datasets from different producers could be given the same name. This must be avoided.

Immediate producer recognition when looking at a filename is an advantage when handling large amounts of data from different producers on a day to day basis. So would immediate product recognition be.

As S-100 evolves and new products are being identified and produced, this opens for the possibility that the product IDs (file names) in the distribution model could end up being equal (e.g. NO123456.000, NO123456.102, NO123456.104....). The alternative solution proposed (KR+101+1+00000) would solve this issue.

Regarding the issue with only 9 variants available for different producers on the same national code the following solutions could be used:

1. Extend to use ISO ALPHA-2 Code + one additional character (US1, US2, US3...) or
2. Extend to use ISO ALPHA-2 Code + two additional characters (US00, US01..., US11...)

Various combinations using ISO ALPHA-3 Code could also be considered.

While proposing changes to the naming convention, we can also reserve additional characters for Dataset ID. It can enable a data producer more flexibility in defining his product naming scheme. We propose extending the number of optional characters to 10.

Conclusions

Producer Codes should be kept because:

- Immediate recognition is an advantage.
- They help avoid duplicated file names.
- They are used as part of product IDs.

The alternative solution proposed (KR+101+1+00000) could be changed to (KR~~X(X)~~ +101+1+0000000000). The advantages of choosing one of the amended alternative solutions are that:

- File names could continue to be used as product IDs.
- No duplication of file names is possible.
- Equal product IDs on different products would be avoided.
- Immediate producer/product/scale band information without searching metadata is an advantage when handling large amount of similar products from different producers.
- Added flexibility for data producer defining product naming scheme.

Conclusions

PRIMAR recommend choosing one of the amended alternative solutions proposed and extend the number of optional characters to 10. Alternatively further investigation could be done also with regards to considering using ISO ALPHA-3 Code.

Action Required of S-100WG

Consider the arguments and proposal in this document.