DQWG14-10j

Paper for Consideration by Data Quality Working Group

Data integrity, marine boundaries from a MSDI perspective.

Submitted by:	MSDIWG - Chair
Executive Summary:	Informative paper describing data security, data integrity, marine boundaries from a MSDI perspective
Related Documents:	C-17 Spatial Data Infrastructures: “The Marine Dimension” - Guidance for Hydrographic Offices, Edition 2.0.0, January 2017. S-63. IHO Data Protection Scheme, Edition 1.2.0 - January 2015 IHO S-100 Part 15 (within S-100 edition 4.0.0)

Introduction / Background

At the MSDIWG9 meeting in February 2018 in Brazil, the WG discussed data security from a MSDI perspective. It was decided to establish a small breakout group in order to consider the issue of “data security” in the context of MSDI datasets.

Analysis/Discussion

After the MSDIWG9 meeting the group delivered an input paper. The findings are presented in this paper. The group’s position was that the issue is not primarily one of data “security” – security tends to be focused around unauthorised usage, for example stealing a car is unauthorised use of the car. The current model of data security within marine chart data is represented by IHO S-63. The conclusion the MSDIWG came to when looking at these issues from the MSDI perspective was that the main priority is actually data “integrity”, also dealt with comprehensively by IHO S-63. Data integrity establishes two pieces of knowledge for data users, (1) knowing who a piece of data came from and (2) the knowledge that the data has not changed in its journey to the end user.

Why is this more important for MSDI? Because the core concept of MSDI is reuse of marine geospatial data outside its traditional use case of primary SOLAS navigation and within a much broader sphere of activity. The nature of some of the datasets may well be sensitive, not because they are confidential but because there is a high impact cost of them being wrong. If an MSDI provider wrongly attributes a dataset to a particular official body or incorrectly reproduces a dataset (either by visualising it poorly or providing a copy of the incorrect data) the repercussions can be large.

By way of example consider that one of the fundamental datasets recently under consideration are UNCLOS maritime limits and boundaries (other examples exist but this is a robust, simple example which is useful for the purposes of illustrating the problem). UNCLOS official limits and boundaries are a foundation dataset and often used to further denote other official limits and boundaries such as marine protected areas, fishing zones and many others, defining rights and responsibilities as part of a harmonised marine cadastral system. These datasets are simple, by comparison with the complex geospatial data which make up ENC but because they represent the results of, often long standing, political agreements and treaties their economic and political weight can be enormous and the impact of their incorrect reproduction within an MSDI environment is of concern. The challenge, technically, is to provide the means and mechanisms, therefore, to protect the data integrity and assure the end user of the provenance of the data they are receiving.

Is there a ready-made solution?

- Ongoing the IHO and MSDI community needs to consider this issue
- Consider adapting existing mechanisms:
  - Stream based may not be suitable for “data centric” models
  - IHO S-63 (and S-101) relies on a specific end user system
  - Other standards exist but may need adaptation
  - All data integrity systems require a “trust network” to define identity.

How should this challenge be addressed? There are two components. There is clearly a need for a sound technical solution – in the ENC world IHO S-63 was developed and implemented globally within the ECDIS community for precisely this purpose but very much in an ECDIS context, ignoring the wider uses of digital hydrographic data. The S-63 defined a bespoke global network of bodies together with a system of digital assurance in the form of digital signatures which delivers a measure of data integrity to every ENC end user. Users know the origin of the data they are using and that it is complete. IHO S-100 has adopted a modernised version of the S-63 scheme within its new Part 15 (within S-100 edition 4.0.0) and this will provide a similar mechanism for digital signatures without the requirement to enforce data encryption as well. This may meet some of the needs of the MSDI community but this remains to be tested by stakeholders.

The other important element to consider is the communication and promotion of the importance of data integrity among the end user community. This should not be underestimated, particularly in the context of MSDI data. The MSDI user community is far more diverse than the ENC world and there is no standardised “end user system” (like an ECDIS). MSDI data originators often have no concrete idea who is using data or for exactly what purpose. This implies that whatever data integrity measure is used it needs to remain an integral part of the data itself and be delivered with it in a non-transformed way so that it can be verified once the data gets to the end user. The promotion of the importance of data integrity should be an integral part of the MSDI picture with users fully understanding the origin of the data they are using.

Conclusions

As with the development of ENC we are just beginning to understand the nature of this problem as we move to implementation of MSDI in many different areas. The “grim reality” of implementation is forcing the MSDI community us to consider the impact of incomplete, corrupted or wrongly attributed data.

In MSDI often integrity has a higher priority than security. Because often MSDI is built with the express purpose of promulgating data so most (not all) use is “authorised”

Security:

- Unauthorised use (e.g stealing a car, downloading a pirate movie)
- To demonstrate “authorised use” some form of “permission” is required.

Integrity:

- Who sent me this?
- Is it complete?
- Different from “is it correct?”

The conclusion was that the issue is “integrity” which relies on two things, knowing where a piece of data came from and the knowledge that it has not changed in its journey to the end user. This is also dealt with by IHO S-63 in the form of its digital signatures.

There are many ways of addressing this problem and the IHO community needs to establish a method, range of methods, tools, technologies, processes and supporting structures to address the issue as a priority.

Recommendations

The DQWG should be made aware about data integrity and using marine boundaries from a MSDI perspective is a good use case to illustrate the issues. The potential impacts and the need to provide means and mechanisms to protect the data integrity and assure the end user of the provenance of the data they are receiving should be considered further.

Action Required of Data Quality Working Group

The DQWG is invited to:

a. note this report;

b. discuss it at the meeting;

c. and to take any actions as deemed necessary.