DQWG DQWG17/4/1a

Meeting 17 8 Feb 2022

Agenda Item 4.1A

**Report on the Cross check of DQ chapters of S-1xx PSs**

Submitted by review of S-1xx PSs subWG

**SUMMARY**

Executive Summary: The subWG has carried out the Cross check of DQ chapters of published S-1xx PSs including S-101, S-102, S-111, S-121, S-122, S-123, S-127 and S-129.The results show that the implementation of S-97 in these PSs is necessary to be harmonized and improved.

Action to be taken: See paragraph 5&6

Related documents: Action DQWG16/04, DQWG16-04.1A, DQWG13-05A and S-97

**1. Introduction/Overview**

According to the action DQWG 16/04, the subWG comprised of Vice Chair, NL, SE and UNH has implemented the Cross check of DQ chapters of 8 published S-1xx PSs including S-101, S-102, S-111, S-121, S-122, S-123, S-127 and S-129.

**2. Principles of the Cross check of DQ chapters**

The Cross check of DQ chapters follow principles as below:

a) The Cross check shall be carried out in accordance with IHO publication S-97 Ed 1.1.0 - “IHO Guidelines for Creating S-100 Product Specifications”。

b) Only the DQ chapters of S-1xx PSs shall be checked and the results will be presented in the form of cross check matrix.

c) DQ elements included in the other parts of each S-1xx PS (such as S-111) shall be described in the form of Notes under its cross check matrix.

**3. Implementation process of the Cross check of DQ chapters**

**3.1 Identify the DQ requirements for S-1xx PS in S-97**

The recommended data quality measures of each S-1xx PS have been identified by the subWG according to S-97 part C Table C-7-1 – Recommended data quality measures.

Table 1 - DQ requirements for S-1xx PSs in S-97

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Data Quality Measure**  | **DQ measure / description**  | **Applicable to spatial representation types**  | **101** | **102** | **111** | **121** | **122** | **123** | **127** | **129** |
| Completeness / Commission  | numberOfExcessItems / This data quality measure indicates the number of items in the dataset, that should not have been present in the dataset.  | All S-100 based PS  | Y | Y | Y | Y | Y | Y | Y | Y |
| Completeness / Commission  | numberOfDuplicateFeatureInstances / This data quality measure indicates the total number of exact duplications of feature instances within the data.  | All S-100 based PS  | Y | Y | Y | Y | Y | Y | Y | Y |
| Completeness / Omission  | numberOfMissingItems / This data quality measure is an indicator that shows that a specific item is missing in the data.  | All S-100 based PS  | Y | Y | Y | Y | Y | Y | Y | Y |
| Logical Consistency / Conceptual Consistency  | numberOfInvalidSurfaceOverlaps / This data quality measure is a count of the total number of erroneous overlaps within the data. Which surfaces may overlap and which must not is application dependent. Not all overlapping surfaces are necessarily erroneous.  | PS with geometric surfaces.  | Y | Y | Y | Y | Y | Y | Y | Y |
| Logical Consistency / Domain Consistency  | numberOfNonconformantItems / This data quality measure is a count of all items in the dataset that are not in conformance with their value domain.  | All S-100 based PS  | Y | Y | Y | Y | Y | Y | Y | Y |
| Logical Consistency / Format Consistency  | physicalStructureConflictsNumber / This data quality measure is a count of all items in the dataset that are stored in conflict with the physical structure of the dataset.  | All S-100 based PS  | Y | Y | Y | Y | Y | Y | Y | Y |
| Logical Consistency / Topological Consistency  | rateOfFaultyPointCurveConnections / This data quality measure indicates the number of faulty link-node connections in relation to the number of supposed link-node connections. This data quality measure gives the erroneous point-curve connections in relation to the total number of point-curve connections.  | PS with curves.  | Y | N | N | Y | Y | Y | Y | Y |
| Logical Consistency / Topological Consistency  | numberOfMissingConnectionsUndershoots / This data quality measure is a count of items in the dataset within the parameter tolerance that are mismatched due to undershoots.  | PS with curves  | Y | N | N | Y | Y | Y | Y | Y |
| Logical Consistency / Topological Consistency  | numberOfMissingConnectionsOvershoots / This data quality measure is a count of items in the dataset within the parameter tolerance that are mismatched due to overshoots.  | PS with curves.  | Y | N | N | Y | Y | Y | Y | Y |
| Logical Consistency / Topological Consistency  | numberOfInvalidSlivers / This data quality measure is a count of all items in the dataset that are invalid sliver surfaces. A sliver is an unintended area that occurs when adjacent surfaces are not digitized properly. The borders of the adjacent surfaces may unintentionally gap or overlap to cause a topological error.  | PS with geometric surfaces.  | Y | N | N | Y | Y | Y | Y | Y |
| Logical Consistency / Topological Consistency  | numberOfInvalidSelfIntersects / This data quality measure is a count of all items in the dataset that illegally intersect with themselves.  | PS with curves / geometric surfaces.  | Y | N | N | Y | Y | Y | Y | Y |
| Logical Consistency / Topological Consistency  | numberOfInvalidSelfOverlap / This data quality measure is a count of all items in the dataset that illegally self-overlap.  | PS with curves / geometric surfaces.  | Y | N | N | Y | Y | Y | Y | Y |
| Positional Accuracy / Absolute or External Accuracy  | Root Mean Square Error / Standard deviation, where the true value is not estimated from the observations but known a priori.  | PS with objects that have coordinative values associated.  | Y | Y | Y | Y | Y | Y | Y | Y |
| Positional Accuracy / Vertical Position Accuracy  | linearMapAccuracy2Sigma / Half length of the interval defined by an upper and lower limit in which the true value lies with probability 95%.  | PS with objects that have a vertical coordinative values associated.  | Y | Y | Y | Y | Y | Y | Y | Y |
| Positional Accuracy / Horizontal Position Accuracy  | linearMapAccuracy2Sigma / Half length of the interval defined by an upper and lower limit in which the true value lies with probability 95%.  | PS with objects that have a horizontal coordinative values associated.  | Y | Y | Y | Y | Y | Y | Y | Y |
| Positional Accuracy / Gridded Data Position Accuracy  | Root mean square error of planimetry / Radius of a circle around the given point, in which the true value lies with probability P.  | PS with objects that have a gridded coordinative values associated.  | N | Y | Y | N | N | N | N | N |
| Temporal Quality / Temporal Consistency  | Correctness of ordered events or sequences, if reported.  | PS with objects that have a time value associated.  | Y | Y | Y | Y | Y | Y | Y | Y |
| Thematic Accuracy / Thematic Classification Correctness  | miscalculationRate / This data quality measure indicates the number of incorrectly classified features in relation to the number of features that are supposed to be there. [Adapted from ISO 19157] This is a RATE which is a ratio, and is expressed as a REAL number representing the rational fraction corresponding to the numerator and denominator of the ratio. For example, if there are 1 items that are classified incorrectly and there are 100 of the items in the dataset then the ratio is 1/100 and the reported rate = 0.01.  | All S-100 based PS.  | Y | Y | Y | Y | Y | Y | Y | Y |
| Aggregation Measures / Aggregation Measures  | DataProductSpecificationPassed / This data quality measure is a boolean indicating that all requirements in the referred data Product Specification are fulfilled.  | PS that a require a complete pass of all elements of a dataset/dataset series/spatial object types  | Y | Y | Y | Y | Y | Y | Y | Y |
| Aggregation Measures / Aggregation Measures  | DataProductSpecificationFailRate / This data quality measure is a number indicating the number of data Product Specification requirements that are not fulfilled by the current product/dataset in relation to the total number of data Product Specification requirements.  | PS that a require a complete pass of all elements of a dataset/dataset series/spatial object types  | Y | Y | Y | Y | Y | Y | Y | Y |

Y= required, N= not required.

**3.2 Implement the Cross check of DQ chapter of S-1xx PSs one by one**

The Cross check of DQ chapters of S-1xx PSs including S-101 Ed1.0.0, S-102 Ed2.0.0, S-111 Ed1.0.0, S-121 Ed1.0.0, S-122 Ed1.0.0, S-123 Ed1.0.0, S-127 Ed1.0.0 and S-129 Ed1.0.0 have been implemented one by one to verify whether these recommended DQ measures are included.

**3.2.1 S-101**

**3.2.1.1 DQ chapters of S-101**

The DQ chapter of S-101 is as follow:





**3.2.1.2 Result of cross check**

|  |  |
| --- | --- |
| **Recommendations** | **S-101** |
| 1.Completeness | N |
| 2.Conceptual consistency | N |
| 3. Domain consistency | N |
| 4. Format consistency | N |
| 5. Topological consistency | N |
| 6. Positional Accuracy | N |
| 7. Thematic Accuracy | N |
| 8. Temporal Quality | N |
| 9. Aggregation | N |
| 10. Introduction to DQ Paragraph | Y |

Y=YES, N=NO, N/A=not application

**NOTES:**

a）In terms of “10. Introduction to DQ Paragraph”，S-101 follows the template provided by S-97 but makes a distinction of the positional accuracy to depth and positional (horizontal). We don’t agree that this is necessary.

**3.2.1.3 DQ elements included in the other parts of S-101**

No DQ element is included in the other parts of S-101.

**3.2.2 S-102**

**3.2.2.1 DQ chapters of S-102**

The DQ chapter of S-102 is as follow:



**3.2.2.2 Result of Cross check**

|  |  |
| --- | --- |
| **Recommendations** | **S-102** |
| 1.Completeness | Y |
| 2.Conceptual consistency | Y |
| 3. Domain consistency | Y |
| 4. Format consistency | Y |
| 5. Topological consistency | N/A |
| 6. Positional Accuracy | Y |
| 7. Thematic Accuracy | Y |
| 8. Temporal Quality | Y |
| 9. Aggregation | N |
| 10. Introduction to DQ Paragraph | N |

Y=YES, N=NO, N/A=not application

**NOTES:**

a）In terms of “10. Introduction to DQ Paragraph”，S-102 doesn’t follow the template provided by S-97 but only includes the first paragraph of the template.

**3.2.2.3 DQ elements included in the other parts of S-102**

No DQ element is included in the other parts of S-102.

**3.2.3 S-111**

**3.2.3.1 DQ chapters of S-111**

The DQ chapter of S-111 is as follow:



**3.2.3.2 Result of cross check**

|  |  |
| --- | --- |
| **Recommendations** | **S-111** |
| 1.Completeness | Y |
| 2.Conceptual consistency | N |
| 3. Domain consistency | N |
| 4. Format consistency | N |
| 5. Topological consistency | N/A |
| 6. Positional Accuracy | Y |
| 7. Thematic Accuracy | Y |
| 8. Temporal Quality | N |
| 9. Aggregation | Y |
| 10. Introduction to DQ Paragraph | N |

Y=YES, N=NO, N/A=not application

**NOTES:**

a）In terms of “10. Introduction to DQ Paragraph”，S-111 doesn’t follow the template provided by S-97 but only includes the first paragraph of the template.

b）S-111 does include an “Logical Consistency” , however, it has nothing to do with the proposed in S-97.

**3.2.3.3 DQ elements included in the other parts of S-111**

Conceptual consistency, Domain consistency, Positional Accuracy and Temporal Quality are included in Annex D.

**3.2.4 S-121**

**3.2.4.1 DQ chapters of S-121**

The DQ chapter of S-121 is as follow:



**3.2.4.2 Result of cross check**

|  |  |
| --- | --- |
| **Recommendations** | **S-121** |
| 1.Completeness | Y |
| 2.Conceptual consistency | N |
| 3. Domain consistency | Y |
| 4. Format consistency | Y |
| 5. Topological consistency | Y |
| 6. Positional Accuracy | Y |
| 7. Thematic Accuracy | Y |
| 8. Temporal Quality | Y |
| 9. Aggregation | N |
| 10. Introduction to DQ Paragraph | Y |

Y=YES, N=NO, N/A=not application

**NOTES:**

-

**3.2.4.3 DQ elements included in the other parts of S-121**

No DQ element is included in the other parts of S-121.

**3.2.5 S-122**

**3.2.5.1 DQ chapters of S-122**

The DQ chapter of S-122 is as follow:



**3.2.5.2 Result of cross check**

|  |  |
| --- | --- |
| **Recommendations** | **S-122** |
| 1.Completeness | N |
| 2.Conceptual consistency | N |
| 3. Domain consistency | N |
| 4. Format consistency | N |
| 5. Topological consistency | N |
| 6. Positional Accuracy | N  |
| 7. Thematic Accuracy | N |
| 8. Temporal Quality | N |
| 9. Aggregation | N |
| 10. Introduction to DQ Paragraph | N |

Y=YES, N=NO, N/A=not application

**NOTES:**

a）In terms of “10. Introduction to DQ Paragraph”，S-122 does include an introduction, however, this is their own and has nothing to do with the proposed in S-97.

b) In fact, the whole DQ chapter of S-122 is only related to data validation.

**3.2.5.3 DQ elements included in the other parts of S-122**

No DQ element is included in the other parts of S-122.

**3.2.6 S-123**

**3.2.6.1 DQ chapters of S-123**

The DQ chapter of S-123 is as follow:



**3.2.6.2 Result of cross check**

|  |  |
| --- | --- |
| **Recommendations** | **S-123** |
| 1.Completeness | N |
| 2.Conceptual consistency | N |
| 3. Domain consistency | N |
| 4. Format consistency | N |
| 5. Topological consistency | N |
| 6. Positional Accuracy | N  |
| 7. Thematic Accuracy | N |
| 8. Temporal Quality | N |
| 9. Aggregation | N |
| 10. Introduction to DQ Paragraph | N |

Y=YES, N=NO, N/A=not application

**NOTES:**

a）In terms of “10. Introduction to DQ Paragraph”，S-123 does include an introduction, however, this is their own and has nothing to do with the proposed in S-97.

b) In fact, the whole DQ chapter of S-123 is only related to data validation and is the same as S-122.

**3.2.6.3 DQ elements included in the other parts of S-123**

No DQ element is included in the other parts of S-123.

**3.2.7 S-127**

**3.2.7.1 DQ chapters of S-127**

The DQ chapter of S-127 is as follow:



**3.2.7.2 Result of cross check**

|  |  |
| --- | --- |
| **Recommendations** | **S-127** |
| 1.Completeness | Y |
| 2.Conceptual consistency | Y |
| 3. Domain consistency | Y |
| 4. Format consistency | Y |
| 5. Topological consistency | Y |
| 6. Positional Accuracy | Y |
| 7. Thematic Accuracy | Y |
| 8. Temporal Quality | Y |
| 9. Aggregation | Y |
| 10. Introduction to DQ Paragraph | Y |

Y=YES, N=NO, N/A=not application

**NOTES:**

-

**3.2.7.3 DQ elements included in the other parts of S-127**

No DQ element is included in the other parts of S-127.

**3.2.8 S-129**

**3.2.8.1 DQ chapters of S-129**

The DQ chapter of S-129 is as follow:



**3.2.8.2 Result of cross check**

|  |  |
| --- | --- |
| **Recommendations** | **S-129** |
| 1.Completeness | N |
| 2.Conceptual consistency | N |
| 3. Domain consistency | N |
| 4. Format consistency | N |
| 5. Topological consistency | N |
| 6. Positional Accuracy | N  |
| 7. Thematic Accuracy | N |
| 8. Temporal Quality | N |
| 9. Aggregation | N |
| 10. Introduction to DQ Paragraph | N |

Y=YES, N=NO, N/A=not application

**NOTES:**

a）In terms of“10. Introduction to DQ Paragraph”，S-129 does include an introduction, but it is to mainly explain why DQ cannot be evaluated in the context of UKC (as it should be already validated based on other product specifications). However, S-97 states: “All S-100-based Product Specifications should include comprehensive ways of capturing information about the quality of the data.”

**3.2.8.3 DQ elements included in the other parts of S-129**

No DQ element is included in the other parts of S-129.

**4. Cross Check Matrix**

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Recommendations** | **S-101** | **S-102** | **S-111** | **S-121** | **S-122** | **S-123** | **S-127** | **S-129** |
| 1.Completeness | N | Y | Y | Y | N | N | Y | N |
| 2.Conceptual consistency | N | Y | N | N | N | N | Y | N |
| 3. Domain consistency | N | Y | N | Y | N | N | Y | N |
| 4. Format consistency | N | Y | N | Y | N | N | Y | N |
| 5. Topological consistency | N | N/A | N/A | Y | N | N | Y | N |
| 6. Positional Accuracy | N | Y | Y | Y | N | N | Y | N |
| 7. Thematic Accuracy | N | Y | Y | Y | N | N | Y | N |
| 8. Temporal Quality | N | Y | N | Y | N | N | Y | N |
| 9. Aggregation | N | N | Y | N | N | N | Y | N |
| 10. Introduction to DQ Paragraph | Y | N | N | Y | N | N | Y | N |

Y=YES, N=NO, N/A=not application

**NOTES:**

a) DQ chapters of the 8 published S-1xx PSs are not in a harmonized way.

b) Some S-1xx PSs like S-122 and S-123 confuse data quality and data validation.

c) DQ chapters of some S-1xx PSs do not conform to the S-97.

d) Except S-127, No S-1xx PSs implement the table C-7-1 – Recommended data quality measures of S-97.

**5. Recommendations**

It is recommended to:

a) Develop a template by combining the DQ chapters of S-102 and S-127 for use in all S-1xx PSs.

b) Prepare suggestions to S-1xx PSs respectively.

c) Continue the cross check of DQ chapters of newly released S-1xx PSs such as S-104 Ed 1.0.0 and the new Edition of the published S-1xx PSs such as S-102 Ed 2.1.0.

d) Carry out the cross check of data validation of the published S-1xx PSs.

**6. Action**

The DQWG is requested to:

1. **Note** the information provided;
2. **Approve** the recommendations above.