

# Automatic assessment of metadata quality in ISO 191\*\*

Susanne Feistel<sup>1</sup>, Ulrike Kleeberg<sup>2</sup>, Jörn Kohlus<sup>3</sup>, Rainer Lehfeldt<sup>4</sup>, Carsten Schirnack<sup>5</sup>, Susanne Tamm<sup>6</sup>

## Creating accessible quality information

### Document your workflow:

1. Purpose of your research
2. Subject of your research
3. Data sampling or sources
4. Geo-reference your data
5. Data processing steps
6. Quality Control
7. Availability of your data

### Put quality information in ISO metadata:

- ➔ MD\_Identification.purpose
- ➔ MD\_Identification.abstract
- ➔ LI\_Source, LI\_Lineage
- ➔ EX\_Extent, MD\_Identification.spatialResolution
- ➔ LI\_ProcessSteps
- ➔ DQ\_QualityElement, DQ\_EvaluationMethod
- ➔ MD\_Distribution, MD\_Constraints

\*\* ISO 19107, ISO 19111, ISO 19115 (2014), ISO 19139, ISO 19156, ISO 19157

## Creating a quality flag

### Create standardized string as quality flag:

- Batch xml analysis via software
- Manually via web application form
- Manually via web application xml upload

The result: **a comparable alphanumeric string combining quality scheme and flags, e.g. „SDN::1“** for „good quality“ as defined by Seadatanet.

### Put quality flag in ISO metadata:

DQ\_StandaloneQualityReportInformation (o,n)

## Good Practice

Quality information may be stored in a practical number of specified recommended fields of the metadata standard. Therefore data portals with specific filter options can increase dissemination and availability of data with documented quality information.

Search and filter options for data with specific quality requirements

Documenting data with standardized metadata for quality.

		ISO 191**										
		obligation occurrence	LI_Source	CI_Citation	MD_Identification.purpose	MD_Identification.abstract	EX_Extent	LI_Lineage	DQ_Element	DQ_EvaluationMethod	MD_Distribution	MD_Constraint.accessConstraint
ABOUT THE DATA	WHY											
	1. Purpose of your research											
	WHAT											
	2. Subject of your research											
	3. Data sources											
	WHERE & WHEN											
	4. Geo-reference your data											
HOW												
3. Data sampling												
5. Data processing steps												
6. Quality Control												
WHO												
AVAILABILITY												
7. Availability of your data												
	Good Practice											
	Documentation											
	Media											
	Products											
	Modelling											
	Experimental											
	Lab data without geo-reference											
	combining multiple source data & create visualization, maps, etc.											
	combining multiple source data & run algorithms											
	In-situ Sampling											
	in the broadest sense											

x + x = (Probably) Good Quality  
x = Questionable Quality  
  = Unknown Quality

## Discussion

### Pro:

- applicable for all data types
- minimal quality information
- motivation to document
- self assessment of quality
- objective assessment
- easy implementation

### Contra:

- very generic
- no content analysis possible

### Next:

- data type specific quality information

<sup>1</sup> susanne.feistel@io-warnemuende.de

<sup>2</sup> ulrike.kleeberg@hzg.de

<sup>3</sup> joern.kohlus@lkn.landsh.de

<sup>4</sup> rainer.lehfeldt@baw.de

<sup>5</sup> cschirnack@geomar.de

<sup>6</sup> susanne.tamm@bsh.de

