Data Quality Standards

Workshop 2 - International Standards Seminar I
Commission: Standards Network

FIG General Assembly in Munich
Monday October 9, 2006 14:00-15:30 Room 21a

Dr. Gerhard Joos
dotGIS – Beratender Ingenieur für Geoinformation

Gerhard.Joos@dotGIS.de
http://www.dotGIS.de

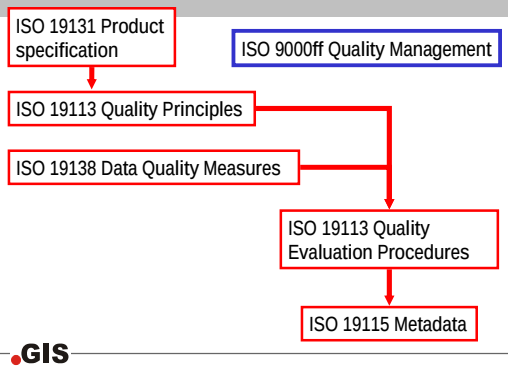
Importance of Quality

- Gain confidence in geodata
- Reduce users' complaints
- Get customer's satisfaction
- Minimize consecutive costs caused by decisions or actions based on erroneous data

Why standards for data quality?

- Data quality has to be described in an objective, application independent way
- Data quality information of different datasets has to be comparable
- Data quality information has to be exchangable
- Data quality information is part of metadata
- Data quality description for negotiation of certain quality levels
- Data quality information has to be considered for analyses on the data (e.g. uncertainty and error propagation)
- Implementations need certain parameters for the algorithms

Quality related standards in ISO/TC 211

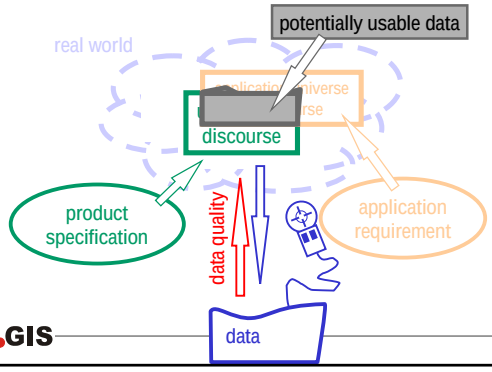


ISO 9000 Quality Management

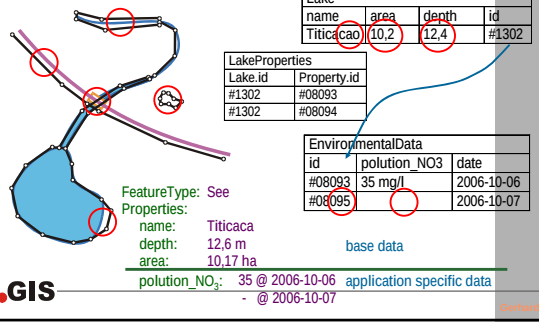
- focus on the management aspects to achieve costumer's satisfaction
- process oriented
- generic: applicable to any product or service

⇨ has to be specified for geospatial data

Concept of data quality in ISO 19113



Classification of error types ⇨ data quality elements



GIS

ISO 19113 Quality principles

data quality element

data quality subelement

- completeness
  - commission
  - omission
- logical consistency
  - conceptual consistency
  - domain consistency
  - format consistency
  - topological consistency
- positional accuracy
  - absolute or external accuracy
  - relative or internal accuracy
  - gridded data position accuracy
- temporal accuracy
  - accuracy of a time measurement
  - temporal consistency
  - temporal validity
- thematic accuracy
  - classification correctness
  - non-quantitative attribute correctness
  - quantitative attribute accuracy

GIS

ISO 19138 Data quality measures

- Defines the structure for building data quality measures
- Provides data quality base measures to be used for constructing data quality measures
  - Counting related (e.g. percentage of missing features)
  - Uncertainty related (e.g. CEP 95)
- Gives a list of 75 standardized data quality measures

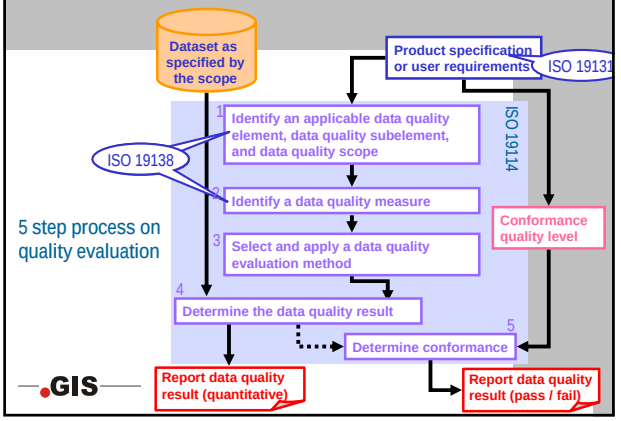
GIS

ISO 19138 Data quality measures

Line	Component	Description
1	Name	number of incorrectly classified features
2	Alias	–
3	Data quality element	thematic accuracy
4	Data quality subelement	classification correctness
5	Data quality basic measure	Error count
6	Definition	number of incorrectly classified features
7	Description	–
8	Parameter	–
9	Data quality value type	Integer
10	Data quality value structure	–
11	Source reference	–
12	Example	–
13	Identifier	62

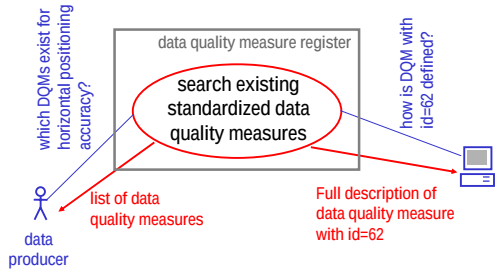
GIS

Evaluating and Reporting Quality Evaluation Results [ISO 19114]



GIS

Use case: register for data quality measures



GIS

Advantages of a register for data quality measures

- Avoid unnecessary duplication of descriptions for data quality measures
  - Reduces the amount of quality related metadata
- Helps the user to identify standardized data quality measures
- Online interface for users and systems

Summary

- Information on the quality of geospatial data is important for their usage
- Standardized quality description makes quality results comparable
- ISO/TC 211 provides a suite of standards related to quality of geospatial data
- Despite of a printed standard an online register for data quality measures is desirable

Thank you for listening!

Questions?