Data in 3 Levels

By Jan Hjelmager and Lars Erik Storgaard
Who are we?

• Name: Jan Hjelmager
  • Employer: Danish Agency for Data Supply and Infrastructure (SDFI)
  • Task: Standardization (internal, national and international)
  • Contact info:
    • Email: jnh@sdfi.dk
    • Phone: +45 72 54 51 15

• Name: Lars Erik Storgaard
  • Employer: Danish Agency for Data Supply and Infrastructure (SDFI)
  • Task: Standardization (internal, national and INSPIRE)
  • Contact info:
    • Email: laers@sdfi.dk
    • Phone: +45 72 54 52 79
Setting the scene

• A data quality project

• Prioritizing use of resources

• New metadata system

• Based on international standards

• Only on data product level

• All SDFI product **shall** be in this concept and classified in three levels
Introduction – the three product levels

All SDFI data products must be specified and distributed at a solid base level.

For some community/business critical data products, there are additional requirements.

Exceptional

Special

Basic
Introduction – the three product levels

Exceptional
- Basic
- Special
- Exceptional

• Description of data product
• Description of lineage, quality and update frequency
• Operation and support information available

Basic
- Special
- Exceptional

• Basic
• Dataset and data product spec. available
• Datamodel and use case available
• The data product fulfil special requirement relative to distribution and quality

Special
- Exceptional
• Basic, Special
• Conceptual model and guidance document available
• The data product fulfill exceptional requirement relative to distribution and quality
How to obtain the data product classification?

The classification matrix

**Criticality**
- Political (height political focus, e.g. climate agenda)
- Legal (public administration)
- Strategic (overall objectives, contractually bound)

**Usage**
- Quantitative (amount of users)
- Importance (societal or business-critical use of data products)
- Value creation (type of value creation)
How to obtain the data product classification?

The decision tree

Is there a socially critical use of the product?
- Yes → Are there exceptional requirements for operation, support or data quality?
- No → Is there a business-critical use of the product?
  - Yes → Are there exceptional requirements for operation, support or data quality?
  - No → Is there an obligation in relation to Danish/European legislation?
    - Yes → Is there something political or strategic that dictates a higher classification than Special?
    - No → Special
      - Data that must meet special requirements
    - No → Basic
      - Basic data that meets SDFI's minimum requirements
  - No → Is there something political or strategic that dictates a higher classification than Basic?
    - Yes → Exceptional
      - Data that must be able to meet extraordinary requirements
    - No → Basic
      - Basic data that meets SDFI's minimum requirements

Classification result

- Basic 28
- Special 37
- Exceptional 3

Data volume:
- 68 different data products
  - 3 GNSS
  - 2 Elevation
  - 6 Orthophoto
  - 10 Raster
  - 47 Vector
Continuous adjustment

1. Evaluate the starting situation and describe the desired data classification
2. Rate as-is data quality
3. Gap-analysis
4. Plan data quality adjustment
5. Execute quality adjustment
Standards, Technical Specifications and legal requirement

- Fulfil INSPIRE Directive
- EN/ISO 19115-1 Geographic information – Metadata – Part 1: Fundamentals
- EN/ISO 19115-2 Geographic information – Metadata – Part 2: Extensions for acquisition and processing
- EN/ISO 19157-1 Geographic information — Data quality — Part 1: General requirements
The framework

Requirements to be meet:
- Metadata requirement from INSPIRE
- Data distribution
- Not develop/invent our own vocabulary

The framework:
- A simple
- A detailed

Data volume:
- 68 different data products
  - 3 GNSS
  - 2 Elevation
  - 6 Orthophoto
  - 10 Raster
  - 47 Vector
The content of the simple framework

- Name of data product
- Abstract
- Data collection frequency
- Distribution format
- Distribution method and where from
- Data collection method and quality
- Last update and update frequency
- Operation information
- Support
- Title of data set specification
- Title of data product specification (DPS)

- Does the product fulfil the DPS?
- Information about the data model
- Use cases related to the product
- Any special requirement regarding data quality
- Any special requirement regarding distribution
- Any expeptionel requirement regarding data quality
- Any expeptionel requirement regarding distribution
- Does a conceptual model exist
- User guide
- Classification level
The content of the detailed framework (vector and raster)

**Vector**
- Completeness
  - Omision /omission
- Thematic accuracy
  - Classification
- Logic consistency
  - Conceptual consistency
  - Topological consistency
- Temporal quality
  - Temporal consistency
  - Temporal validity
- Geometric accuracy
  - Absolute accuracy

**Raster**
- Conceptual consistency
- Geometric accuracy
- Unit of distribution
- Size of unit of distribution
- Feature catalogue
- Name of data set included in the product (including version)
- Where can the product be obtained online
- Responsible for each data set in the product
- Process description and date for the process
- Responsible for the process
In summary and status

• All 68 product have been assigned to a responsible department
• All responsible of one or more product have filled a spreadsheet(s) with the required information
• The data quality content of the spreadsheet(s) (cf ISO 19157-1) are vague
• More guidance is required on filling out the spreadsheet(s)
• An organization to continue (and maintain) the framework in operation
• An governance structure have been developed to maintain the content of the framework
• Defined as-is and to-be product levels
• Pending: Transfer of product between levels

Data in 3 Levels by Jan Hjelmager and Lars Erik Storgaard

11 – 12 October 2023

14