Mission on *Activity E6.2*

Metadata Introduction to Subject Matter Units

Metadata overall and focus on Single Integrated Metadata Standard (SIMS)

ICBS, 29 September to 2 October 2014

Mogens Grosen Nielsen
mgn@dst.dk
Kim Duncan Bendix
kdb@dst.dk
Statistics Denmark

STATISTICS DENMARK

Agenda

1. Metadata overall
2. Focus on Single Integrated Metadata Structure
3. Questions and discussion

STATISTICS DENMARK

Vision

- Statistical information must help users in the turbulent sea of information
  - The place to find quality information
  - Evidence based decisions
- Metadata about content and quality must
  - Allow users to locate relevant data
  - help users in their knowledge processes
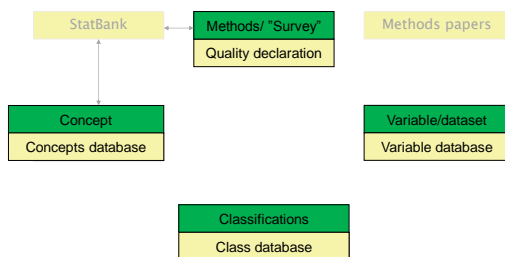 - give users precise information about products

STATISTICS DENMARK

Aren't our metadata good enough?

- Dedicated work since 1995
- Followed best international practice
- Main elements in place
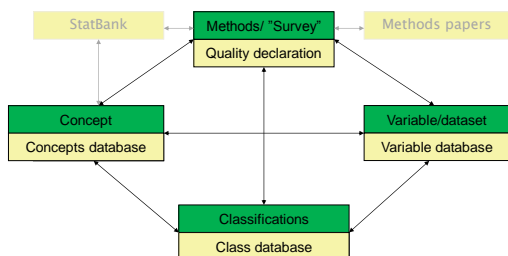

Metadata elements

StatBank	Methods/ "Survey"	Methods papers
	Quality declaration	

Concept	Variable/dataset
Concepts database	Variable database

Classifications
Class database

However...

- Consistent claims of users:

Metadata / documentation isn't good enough

STATISTICS
DENMARK

Integration of metadata


```
 StatBank        Methods/ "Survey"        Methods papers
                 Quality declaration

 Concept                                  Variable/dataset
 Concepts database                        Variable database

                 Classifications
                 Class database
```
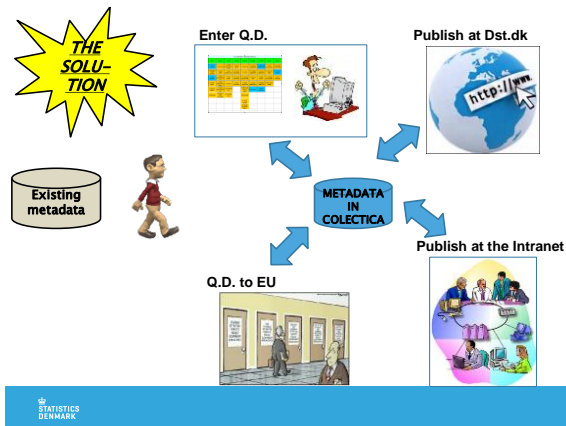
Challenges of fulfilling user needs - Denmark

- Difficult to understand content of quality declarations
- Comparability across domains is needed
- Expert knowledge too sectorial
- Presentation of metadata fragmented and incomplete
- Concepts database incomplete & inconsistent
- Classifications and code-lists in many places
- Explanation and documentation of breaks & revisions

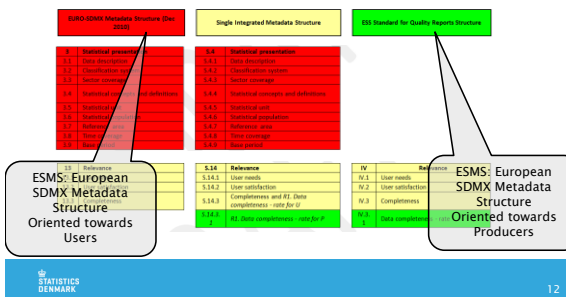## Challenges of fulfilling user needs - Denmark

- Need for better integration between variables, concepts, classifications and quality declarations including links to relevant statistical information

STATISTICS DENMARK

Single Integrated Metadata Structure (SIMS) and reporting formats: ESMS and ESQRS

Single Integrated Metadata Structure (SIMS) and reporting formats: ESMS and ESQRS

ESMS: European SDMX Metadata Structure Oriented towards Users

ESMS: European SDMX Metadata Structure Oriented towards Producers

STATISTICS DENMARK

Dissemination – front page

|	
1. Introduction (Introduction text)	
2. Data Description (Data description text)	+
3. Relevance (Introduction text)	+
4. Accuracy and reliability (Accuracy and reliability text)	+
5. Timeliness (Timeliness text)	+
6. Comparability and Coherence Comparability (text)	+
7. Accessibility and Clarity (Accessibility text)	+
8. Contact (Contact text)	+

Selected fields and standards

- External
  - User oriented reporting
  - SIMS level 1 and 2 (ESMS)
- Internal
  - Producer oriented reporting
  - SIMS leve 1 + 2 + 3 + 4 (ESQRS)
- Recommendation for ICBS
  - Start with SIMS level 1 and 2 (short run) and extend with SIMS level 3 and 4 (long run)
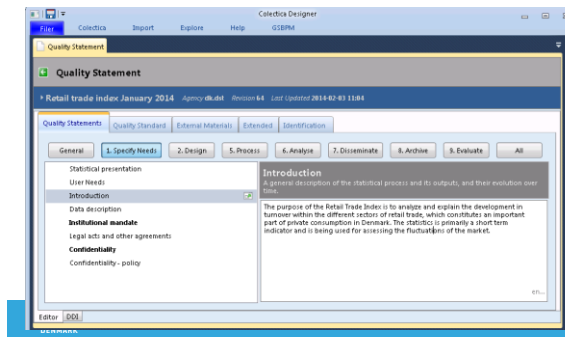
Work processes and quality declarations

STATISTICS DENMARK

5
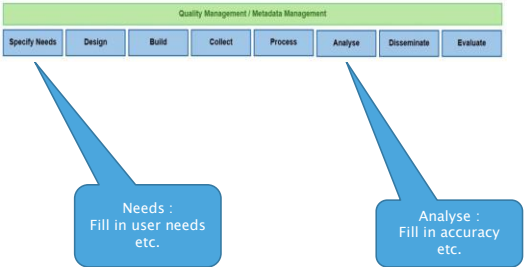
Work processes and quality declarations in Colectica – Retail Trade index
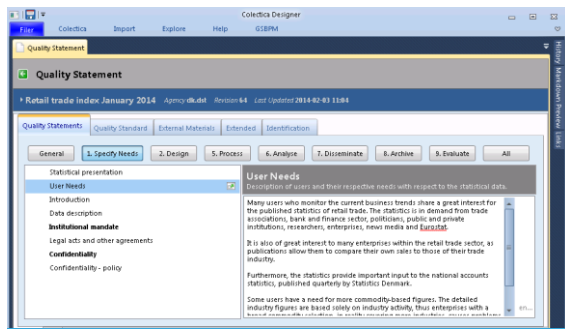
Work processes and quality declarations in Colectica – Retail Trade index

Transmission to Eurostat and customized presentation in reports and at dst.dk

- Many views into quality declarations to support various users
- Example: Extract from report with summary information for Retail Trade Index

Introduction
The purpose of the Retail Trade Index is to analyze and explain the development in turnover within the different sectors of retail trade, which constitutes an important part of private consumption in Denmark. The statistics is primarily a short term indicator and is being used for assessing the fluctuations of the market.

Data description
The retail trade index shows the development of turnover from sales of commodities to private persons in Danish retail enterprises. The retail trade index is published as value and volume indices. The value indices show the development of the turnover in current prices. Value indices are published on 43 industries within the retail trade sector and for three main commodity groups: …

Relevance

Software elements

Architecture
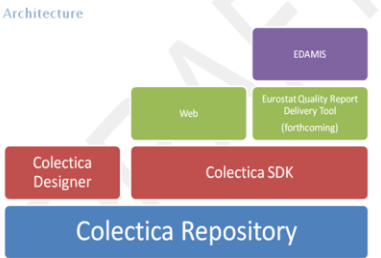
ESMS elements

*

1	Contact
1.1	Contact organisation
1.2	Contact organisation unit
1.3	Contact name
1.4	Contact person function
1.5	Contact mail address
1.6	Contact email address
1.7	Contact phone number
1.8	Contact fax number

ESMS elements

*

3	Statistical presentation
3.1	Data description
3.2	Classification system
3.3	Sector coverage
3.4	Statistical concepts and definitions
3.5	Statistical unit
3.6	Statistical population
3.7	Reference area
3.8	Time coverage
3.9	Base period

4	Unit of measure

5	Reference period

6	Institutional mandate
6.1	Legal acts and other agreements
6.2	Data sharing

STATISTICS
DENMARK

ESMS elements

7	Confidentiality
7.1	Confidentiality - policy
7.2	Confidentiality - data treatment

8	Release policy
8.1	Release calendar
8.2	Release calendar access
8.3	User access

9	Frequency of dissemination

ESMS elements

•

10	Dissemination format
10.1	News release
10.2	Publications
10.3	On-line database

10.4	Micro-data access
10.5	Other

11	Accessibility of documentation
11.1	Documentation on methodology

11.2	Quality documentation

12	Quality management
12.1	Quality assurance
12.2	Quality assessment

ESMS elements

13	Relevance
13.1	User needs
13.2	User satisfaction
13.3	Completeness

14	Accuracy and reliability
14.1	Overall accuracy
14.2	Sampling error

14.3	Non-sampling error

STATISTICS DENMARK

ESMS elements

15	Timeliness and punctuality
15.1	Timeliness

| 15.2 | Punctuality |

16	Comparability
16.1	Comparability - geographical

| 16.2 | Comparability - over time |

17	Coherence
17.1	Coherence - cross domain

| 17.2 | Coherence - internal |

ESMS elements

18	Cost and burden

19	Data revision
19.1	Data revision - policy
19.2	Data revision - practice

20	Statistical processing
20.1	Source data
20.2	Frequency of data collection
20.3	Data collection
20.4	Data validation
20.5	Data compilation

| 20.6 | Adjustment |

21	Comment

The End!

STATISTICS DENMARK