

**UNITED NATIONS STATISTICAL COMMISSION  
ECONOMIC COMMISSION FOR EUROPE  
CONFERENCE OF EUROPEAN STATISTICIANS**

**UNECE Workshop on the Common Metadata Framework  
(Vienna, Austria, 4-6 July 2007)**

**CASE STUDY - SWEDEN<sup>1</sup>**

**1. INTRODUCTION**

- Organization details:

**Statistics Sweden**

Mail: SCB, Box 24300, SE-104 51 STOCKHOLM

Phone: + 46 8 506 940 00

Fax: + 46 8 661 52 61

E-mail: [scb@scb.se](mailto:scb@scb.se)

Visitors: Karlavägen 100, Stockholm

***Results from operations in 2006:***

146 publications

356 press releases

10 800 commissions

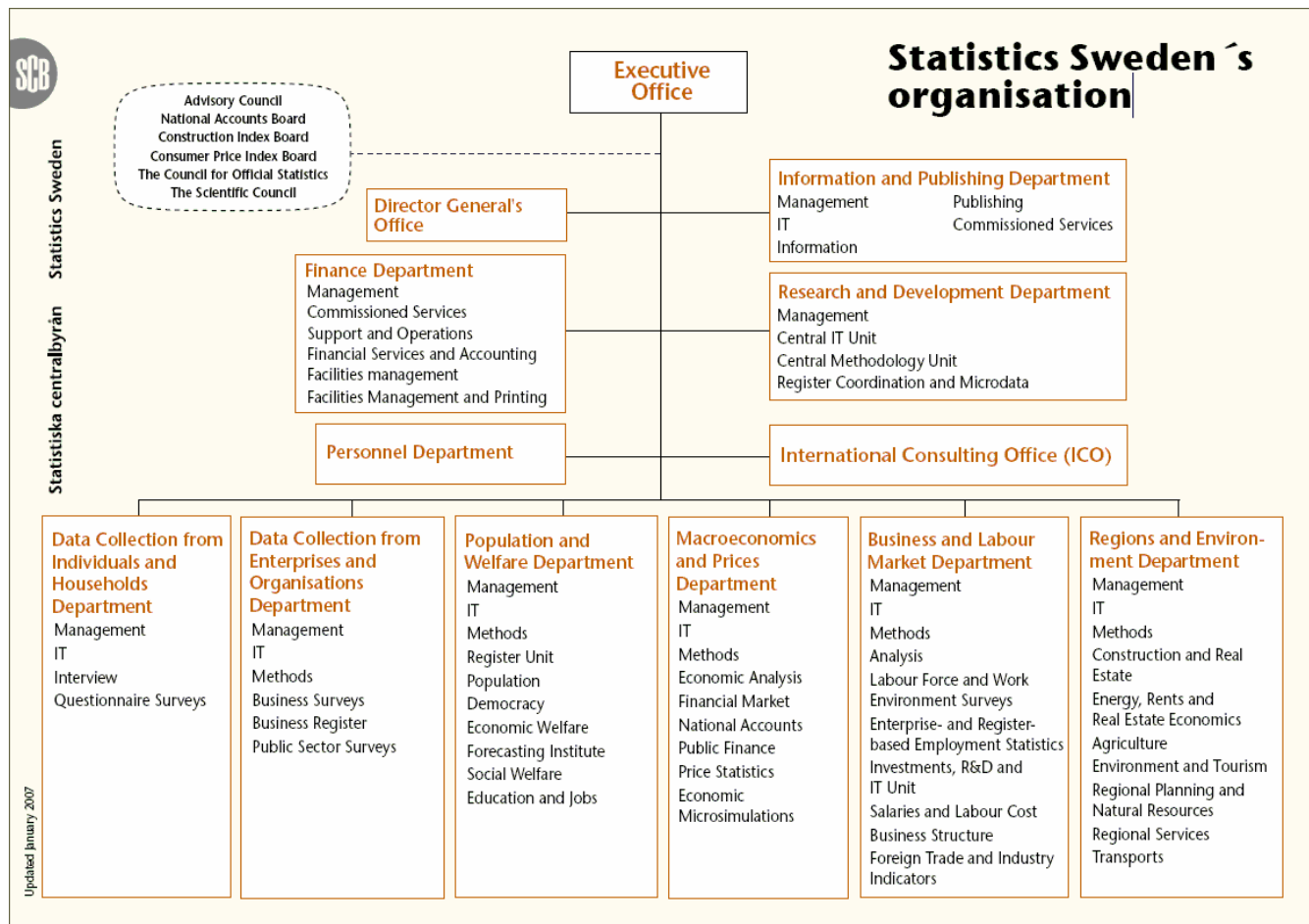
3,7 million visitors at <http://www.scb.se/>

***Number of employed:*** Approx. 1400 employees (approx. 550 in Stockholm, approx. 700 in Örebro and approx. 150 field interviewers around the country)

***Turnover:*** 934 million SEK, of which 451 million SEK in appropriation

---

<sup>1</sup> Prepared by Klas Blomqvist, [klas.blomqvist@scb.se](mailto:klas.blomqvist@scb.se)



- Explanation of the overall strategy, programme providing frame for projects, basic metadata management principles used

## About Statistics Sweden

1. Statistics Sweden is a central government authority for official statistics and other government statistics and in this capacity also responsible for coordinating and supporting the Swedish system for official statistics.

## Statistics Sweden's mission

2. In order to supply customers with good quality statistics to be used as the basis for decision-making, debate and research, Statistics Sweden will:
  - develop, produce and communicate statistics,
  - actively contribute to statistical cooperation on an international level and
  - coordinate and support the Swedish system for official statistics.

## Statistics Sweden's mission statement

3. Statistics Sweden refines data to statistical information for customers through advanced methodological competence, broad knowledge of subject matter and modern techniques in a cost-effective way. The statistics should be impartial, relevant and of good quality, at the same time as being based on scientific principles. We should facilitate the provision of data and protect the collected microdata. In cooperation with others, Statistics Sweden develops common statistical systems, both nationally and internationally. Statistics Sweden is an attractive workplace where the staff are happy with and enjoy their work.

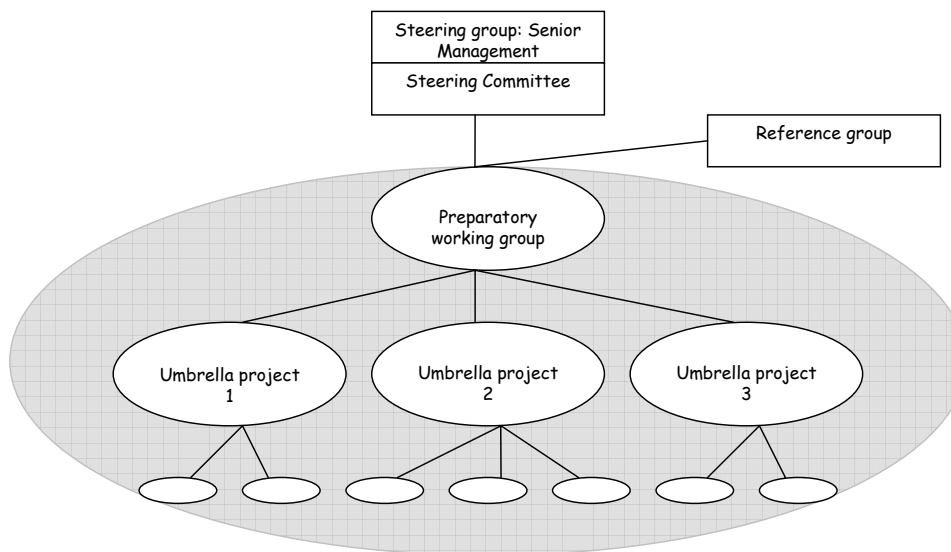
## Statistics Sweden's vision

4. Statistics Sweden is of world class standard in refining data to statistical information, adapted to the customers' needs.
5. Statistics Sweden is already among the leaders in many areas, among others quality work, dissemination, international consultancy, register data and sample surveys, such as the Labour Force Survey (LFS), Consumer Price Index (CPI) and the Survey of Living Conditions (SLC). At the same time, there are areas that need to be strengthened, for example in the household and IT areas. The economic statistics also needs to be improved in several respects.

## The Lotta project

6. There is intensive work at Statistics Sweden going to standardise processes and tools. The background is a decentralised organisation with some recent changes towards standardisation. The speed has increased considerably together with a more clear direction towards standardisation and process orientation. There are both short-term and long-term goals. The overall goals are to reduce costs and raise quality; some such possibilities are already in view, and more will come. There are many challenges in the current phase ending in 2007, e.g. to cover the production process broadly while also evaluating and building methods and tools for sub-processes.
7. The Lotta project has led to a complete reorganisation of development work at statistics Sweden. All resources have been tied to a central project organisation (see image below).

*The Lotta organisation:*



## The MetaPlus project

8. During 2001 to 2003 Statistics Sweden carried out several analyses of the metadata domain, including current and future demands. It became clear that Metadok, the documentation system used at the time did not fulfil a number of the requirements. Most prominent were the lack of possibility to re-use previously completed documentations and the support for longitudinal aspects.
9. Thus a project was launched in 2004 with a primary mission to develop a replacement system for Metadok that would provide support for the requirements that had been identified. Further conditions for the project included contributing to an improved overview of Statistics Sweden's data store and providing tools for harmonisation and coordination. The Classification database, which was until then a separate system, was to be integrated in the new system.

10. The project was staffed with methodologists, IT specialists and persons having practical experience from documentation and classifications. A few project members participated throughout the entire period 2004-2006, but most participants only took part for a shorter time. Important complements to the project group were active user groups that throughout the project acted as discussion partners, specified requirements, came up with ideas and tested prototypes. Furthermore, the results were regularly firmly established in the organisation through presentations to various groups and through public seminars.
11. Initially the project compiled already existing materials from previous projects and analyses, followed by a comprehensive analysis, where requirements from current and assumed future users were assembled. The complete list of demands and expectations turned out to be quite substantial.
12. The demands were structured and prioritised and formed a basis of the use cases, i.e. descriptions of what the system was to produce. This work lead to a high level view of the entire system: a conceptual model.
13. Once the conceptual model had been decided on, the development work started to create the physical database and the basic application whose primary task was to create new documentations.
14. In parallel with the development of the new system a migration from Metadok to MetaPlus was carried out. Documentations previously created in the old system were copied to the new one. Since the structures of the two systems are very different the migration was a complicated and time consuming task. It was to a large extent a manual operation primarily carried out by those who had previously created the documentations.
15. On January 1, 2007 MetaPlus was publicly available to Statistics Sweden personnel. At an early stage it was clear that it would not be possible for the project to cover all demands and expectations of the original list of requirements, given the resources and time at its disposal. Hence a decision was taken to make version 1.0 a “basic version” having the fundamental functionality, but being built to be easily expanded, adding new functions as demands arise. The project concluded its work and the system was delivered to the organisation for administration and maintenance, but at the same time instigating a new project to plan for further development.

## **Metadata management**

16. The Research and Development department is responsible for maintaining and developing the different metadata systems and templates. This also includes providing training and instructions. The department has an overall responsibility for content. In the Research and Development department there is a group consisting of members from the central methodology unit, the register coordination and microdata unit and the central IT unit that is responsible for the area classifications, metadata and content harmonisation (KMI). This means to administer and develop the MetaPlus system, provide training in metadata systems and templates, methodological work and harmonisation on classifications, content harmonisation in the MetaPlus system and other development work in the field.
17. Due to the decentralised organisation much of this work is taking place in groups or networks. There is a metadata group, a documentation network and a classification network. These are led by members from the KMI-group and consist of members from all subject matter departments (see section 5 for more information).
18. The publishing department is responsible for presenting metadata and is also responsible for the statistical databases (SSD) and the metadata belonging to it.

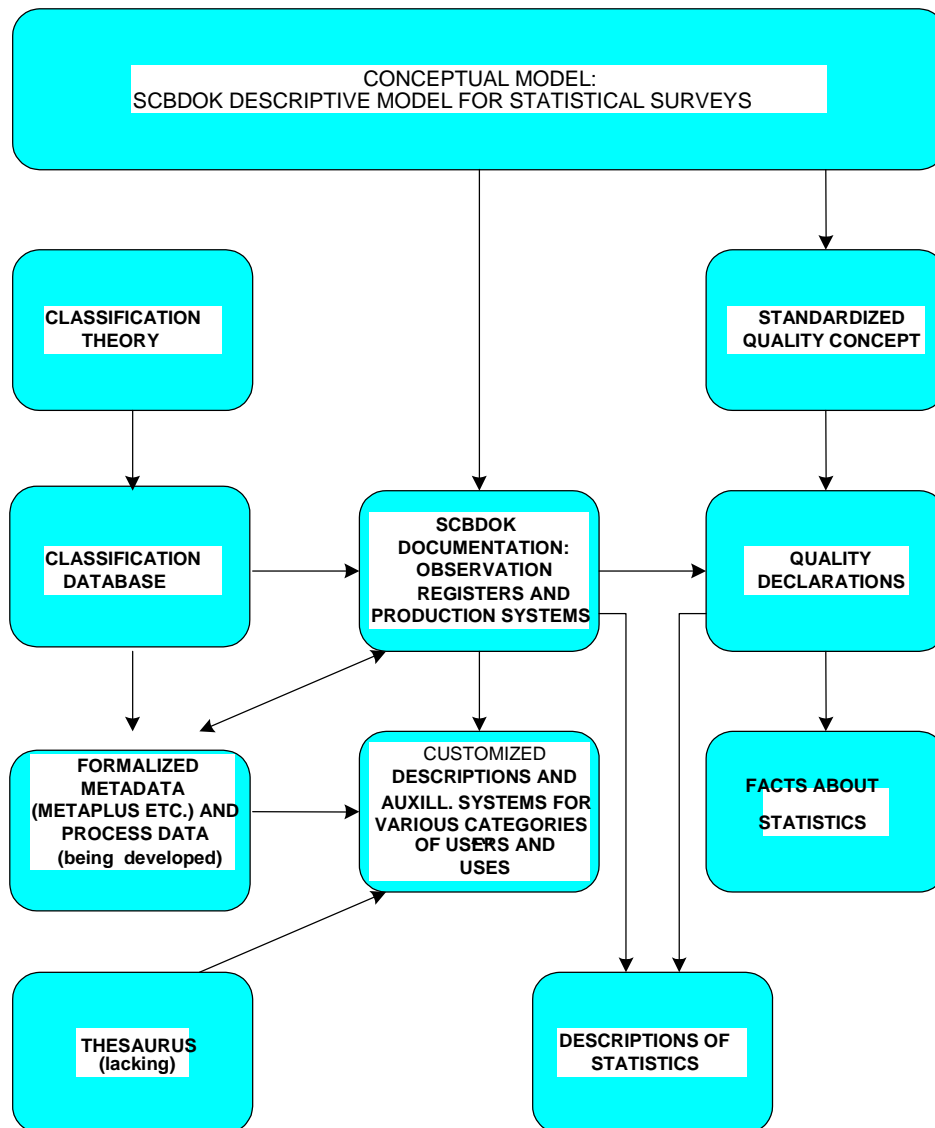
- Overall objective (main goals) so far
  - Document metadata once, metadata should be collected when “made”, reuse
  - Standardised metadata should be used when it is possible

- One metadata repository for micro and macro data
- Lessons learned (future challenges)
  - Involve users at an early stage in development. Makes them feeling a part of the development
  - Inform a lot
  - Make some sort of prototype at an early stage, it is very difficult and abstract for users to describe use cases for a system with this complexity
  - How and what process metadata are we going to collect?

## 2. THE STATISTICAL METADATA SYSTEMS AND THE STATISTICAL CYCLE

- Description of the metainformation systems that are the subject of this case study and how they fit within other organizational systems.

*Summary of different types of documentation and documentation tools and their inter-relationship at Statistics Sweden. (The figure is old but the main ideas are the same):*



## Description of the statistics

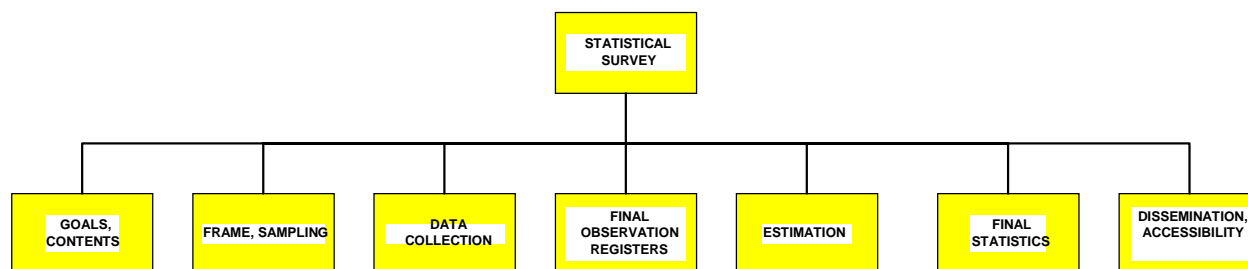
19. Descriptions of the statistics are to be available for all official Swedish statistics, and according to the Swedish Statistics Act they must be produced for all statistics products. The purpose is to provide brief information on the quality of the statistics and other basic facts. The descriptions serve an important function as a quality declaration for the published statistics.
20. The documentation is in free text format following a standard template containing the main chapter's *General information* and *Quality declaration*.

## SCBDOK

21. In 1994 Statistics Sweden decided that all observation registers and production systems under its responsibility must be documented in the SCBDOK system. The purpose is to provide a detailed account of the process of creating a statistical register, from data collection to presentation.
22. The SCBDOK documentation is created in free text format following a standard template containing the following six chapters:
0. General information
  1. Contents outline
  2. Data collection
  3. Final observation registers
  4. Statistical processing and presentation
  5. Data processing system
  6. Log file

SCBDOK 3.0			
<b>0</b>	<b>General information</b>	<b>1</b>	<b>Contents outline</b>
0.1	Policy area	1.1	Observation parameters
0.2	Domain of interest	1.2	Statistical target parameters
0.3	Part of the Official Statistics of Sweden	1.3	Output, statistics and microdata
0.4	Responsible person	1.4	Documentation och metadata
0.5	Producer	<b>2</b>	<b>Data collection</b>
0.6	Mandatory duty to submit data to the survey		
0.7	Confidentiality and processing rules for personal data		
0.8	Appraisal and disposal rules		
0.9	EU regulations		
0.10	Objectives and background		
0.11	Use of the statistics		
0.12	Design and implementation	2.1	Frame and frame procedures
0.13	Planned changes in future surveys	2.2	Sampling procedures
		2.3	Measurement instruments
		2.4	Data collection procedures
		2.5	Data preparation
<b>3</b>	<b>Final observation registers</b>	<b>4</b>	<b>Statistical processing and presentation</b>
3.1	Production versions	4.1	Estimation: assumptions and calculation formulas
3.2	Long-term (archive, terminal) storage versions	4.2	Presentation procedures
3.3	Experiences from the latest survey round		
<b>5</b>	<b>Data processing system</b>	<b>6</b>	<b>Log files</b>

*The statistical production process as in SCBDOK, please note that this description differ to the description made in the Lotta project, see below:*



23. The SCBDOK template covers the whole production process.

## MetaPlus

24. Previously Statistics Sweden used the in-house developed Metadok system to document final observation registers in a formalised way, and the Classification Database (KDB) to document national and international classifications. From 2007 both those systems were replaced by the new MetaPlus system. Chapter 3 of the SCBDOK documentation is derived from MetaPlus.

25. MetaPlus is primarily developed for documenting final observation registers, describing the micro data, meaning that at the moment it is mainly used in the statistical computation phase. It can also be used to describe raw data, in the data collection phase and aggregated data used in tables which would be in the statistical computation phase. One of the ideas with the system is to use it in the design phase for example to get variable definitions and construct the sample. How MetaPlus is going to be used in the production is not clear at this point, several projects are looking into that at the present and more work will be done in this field during autumn.

- An overview of the process model and how/when it was developed. How it is mapped to CMF Lifecycle model (i.e. 1 - survey planning and design; 2 - survey preparation; 3 - Data collection; 4 - Input processing; 5 - Derivation, Estimation, Aggregation; 6 - Analysis; 7 - Dissemination; 8 - Post survey evaluation).

METIS	Statistics Sweden*
1 Survey planning and design	1 Target, customer demand
2 Survey preparation	2 Frame and sample
3 Data collection	3 Data collection
4 Input processing	4 Data preparation
5 Derivation, estimation, aggregation	5 Statistical computation
6 Analysis	
7 Dissemination	6 Dissemination and communication
8 Post survey evaluation	7 Evaluation/customer satisfaction

\*All the levels have sub levels, see figure at paragraph 28 below.

- Description of different metadata groups, like process metadata, quality metadata etc.

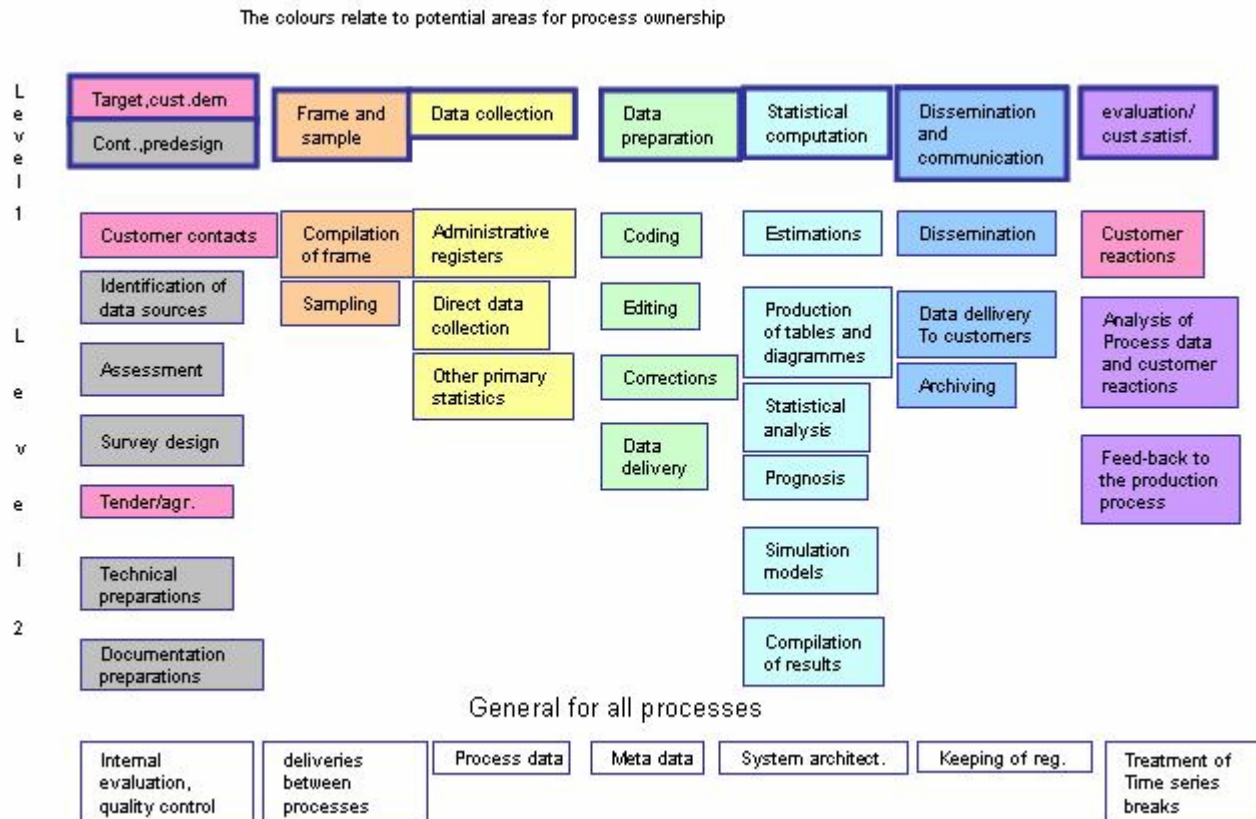
26. Currently no central systems collect process metadata. Some production systems do collect local process metadata, but there is no overview of its contents or usage.

27. The main tool for Quality metadata in use at Statistics Sweden is the quality declaration. It is used for quality of the statistics in its aggregated form, macro data. An ongoing project is looking at ways to add a quality description for micro data. It will probably somehow be connected to MetaPlus

### 3. STATISTICAL METADATA IN EACH PHASE OF THE STATISTICAL CYCLE

- Detailed description of each phase of the statistical cycle and examples of the metadata used and/or created.

28. The statistical production process as described in the Lotta project.



29. Please note that this is not a final version of the description of the production process at Statistics Sweden, the work on describing the production process is ongoing at the moment. As mentioned above the SCBDOK template covers the whole production process. However, the phases of the process description in SCBDOK are not updated according to the figure above. MetaPlus describes the content in the final observation registers, but can be used to describe any data matrix at any other production stage as well i.e. the input phase or the aggregated data used in the tables.

30. The statistical databases (SSD) contain metadata for macro data.

31. The quality declaration template describes the quality of the statistics (macro data). Quality information is going to be linked to MetaPlus

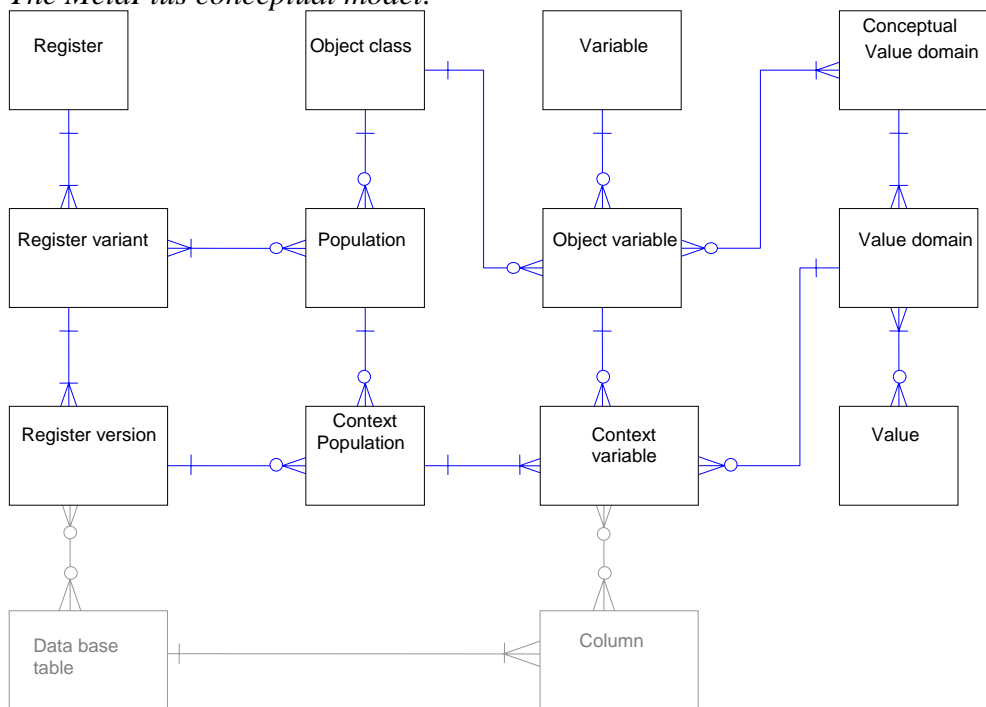
### 4. SYSTEMS AND DESIGN ISSUES

32. The various parts of the Swedish metadata system are not fully integrated. Some parts (SCBDOK, Description of the Statistics) are simply templates to be filled out in Microsoft Word. The completed documents are then presented via simple web methods. The part of the metadata system that handles documentation of the variables on the micro data level, MetaPlus, is built as a traditional client/server system, using a relational database as its core and a "fat client" application for data entry and maintenance, or web browser based "thin client" for presentation. Metadata for aggregate data are documented in a separate system: SSD. The development tools used are all from Microsoft: SQL Server and Visual Basic .NET.



33. The tools available to users for creating documentations are the Word templates, the MetaPlus application and the SSD application.
34. Development of the entire Swedish Metadata system has been carried out in-house, taking international standards (ISO11179, Neuchâtel) into consideration.
35. Statistics Sweden is currently considering a complete revision of its IT architecture, probably adopting the SOA principles. Awaiting those decisions the metadata repository is not accessible from other tools than those for data entry and presentation. Some export is possible.

*The MetaPlus conceptual model:*



## 5. ORGANIZATIONAL AND CULTURAL ISSUES

- List of roles in metadata/statistical lifecycle management: description (subject matter statistician, survey manager, metadata manager etc.), responsibilities, organisation group (Subject matter, IT, methodology & standards, central metadata group etc.)
36. The producing departments are responsible for documenting their statistics, all statistics (in principle) shall be documented in SCBDOK (including MetaPlus/Metadok) and have a quality declaration and a description of the statistics. This means that they are responsible for the content. This work is mainly carried out by subject matter statisticians, the survey manager is responsible.
- Description of the team/individuals involved in development and maintenance of metainformation systems.

### The KMI group

37. The Research and Development department is responsible for maintaining and developing the different metadata systems and templates. This also includes providing training and instructions. The department has an overall responsibility for contents. In the Research and Development unit there is a group consisting of members from the central methodology unit, the register coordination and microdata unit and the central IT unit that is responsible for classifications, metadata and content harmonisation (KMI). This means to

administer and develop MetaPlus, provide training in metadata systems and templates and development in the field

### **The documentation network**

38. All subject matter departments have at least one representative in the documentation network. It is used as a means for distributing and collecting information on the documentation system. The group plans the contents of training sessions and discusses relevant problems that arise.

### **The classification network**

39. The responsibility for classifications is decentralised at Statistics Sweden. The purpose of the network is to create a forum where classification related issues can be raised and discussed. Exchange of experiences is one of the goals, but in the long run also practical work on harmonisation of classification variants. It is used as a means for distributing and collecting information related to classifications.

### **The metadata group**

40. All subject matter divisions have at least one representative in the metadata group. IT handles broader issues concerning metadata. The purpose of the network is to create a forum where general metadata related issues can be raised and discussed

- Training approaches and knowledge management

41. The documentation network discusses training issues. Training is offered by the KMI-group and provided by experts in the field, usually experts on the tool or template in question together with methodologists.

- Partnerships and cooperation between agencies

42. Statistics Sweden is a member of the Neuchâtel group

## **6. ATTACHMENTS & LINKS**

Note: The links to attachments below will work once all files in the zip folder at <http://www.unece.org/stats/documents/ece/ces/ge.40/2007/wp.12.e.zip> are downloaded and extracted to the same location as this paper.

- a) [Bergdahl and Elvers “Reducing costs and raising quality through standardised processes and tools at Statistics Sweden”](#)
- b) [Blomqvist, Lundell, Karling and Svensson “MetaPlus”](#)
- c) [Blomqvist and Kristiansson “Developing a system for description of microdata at Statistics Sweden”](#)
- d) [Sundgren ”PROCESS REENGINEERING AT STATISTICS SWEDEN”](#)
- e) [Sundgren “Reality as a statistical construction – Helping users find statistics relevant for them”](#)
- f) [Sundgren “Manual to SCBDOK, Version 3.0”](#)
- g) [Sundgren ”DOCUMENTATION TEMPLATES AND METADATA MODELS AT STATISTICS SWEDEN”](#)

\*\*\* END \*\*\*