WP.7
ENGLISH ONLY

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

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

Topic 2: Functions and attributes of statistical metadata

METAPLUS

Statistics Sweden[1]

A new metadata system for documenting micro data has been developed at statistics Sweden. The model adheres to international standards such as ISO 11179 and its basic attributes such as register, variable, object class/population and value domain are introduced. The system can be used in standardisation and harmonisation work, but also in several other situations like survey design, leading to reduction of respondent burden. Some of the future development plans i.e. a connection to the production process are mentioned.

1. MetaPlus is a part of Statistics Sweden's metadata system. This paper describes the role of MetaPlus in the system, its theoretical foundation, the development process and its intended use.

I. STATISTICS SWEDEN'S METADATA SYSTEM

2. Somewhat simplified metadata can be described as data on data, or information on data. Metadata describes the contents and the technical structure of a dataset. It is an important tool when interpreting the meaning, possibilities and limitations of data. Metadata may be subdivided into different types: technical or contents (quality) metadata; formalised or free text metadata; general or specific metadata; public or private metadata. All these types of metadata can be found at Statistics Sweden.

3. The original metadata type is the formalised, technical metadata found in record layouts, which are necessary to be able to use various software products to access data. Statistics Sweden's SCBDOK system is an example of contents metadata in a free text format.

4. Metadata is of great importance to internal users within Statistics Sweden, but perhaps even more important to external users in order to make correct interpretations and analyses of the statistics. Metadata facilitates communication between client and producer when ordering and executing a commission. Current and future researchers are supported by having an ample access to metadata. Producers of statistics need support from metadata to, e.g., maintain and develop production systems.

[1] Prepared by Klas Blomqvist, klas.blomqvist@scb.se, Lars-Göran Lundell, Catrin Karling, and Eva Svensson eva.svensson@scb.se

5. Statistics Sweden's metadata system contains several parts that are interrelated to make sure that information that has once been entered into the system is re-used to the fullest possible extent. Documenting must not be unnecessarily time consuming or cumbersome.

A. Description of the statistics

6. 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.

7. The documentation is in free text format following a standard template containing the main chapter's *General information* and *Quality declaration.*

B. SCBDOK

8. 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.

9. 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

C. MetaPlus

10. 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.

D. Metadata for aggregate statistics

11. The metadata system comprises a separate part containing metadata for tables in the Statistical database (SSD).

E. Metadata responsibility

12. Each subject matter unit is responsible for creating its documentations and their contents. Completed documentations are checked and approved of by a *documentation responsible* staff member on the department level. The system, templates, manuals, and external publishing are handled centrally.

F. Publishing metadata

13. Finalised documentations are published on Statistics Sweden's web site, www.scb.se, currently only in Swedish: Statistikdatabasen/ Dokumentationer.

G. Process metadata

14. The relation between MetaPlus and process data has not yet been fully investigated, for more information see section 6.1, future plans.

II. BACKGROUND

15. During 2001 to 2003 Statistics Sweden carried out several analyses of the metadata domain, including current and future demands. It became clear that Metadok 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.

16. 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.

17. 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.

18. 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.

19. 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.

20. 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.

21. 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.

22. 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.

III. METAPLUS USAGE AREAS

23. The MetaPlus system is constructed to have a central role in the production system at Statistics Sweden. It can be seen as a hub that different activities are connected to, a multi-purpose tool from which new functions can be folded out when needed. The original reason for developing MetaPlus was to replace the old Metadok system for documenting final observation registers, but from a development perspective the task has been wider than that: to also find new areas where the system can be of use.

A. Range of use

24. There are several fields in the production of statistics where MetaPlus can be applied. These all focus on making the production more efficient. MetaPlus provides new possibilities to get an overview of the Statistics Sweden data storage. That gives the possibility to consider already in the design phase if previously collected data can be used. By doing so new, potentially unnecessary data collections can be avoided. This will reduce the respondent burden, also leading to a more effective production process, making it possible to use resources for other tasks. The quality aspects comparability and access are improved by the overview that the system provides. The accessibility is improved in two aspects: directly, since it enables users to search the micro data metadata, and indirectly, when the staff gets easy access to what their colleagues are doing.

General

25. MetaPlus supplies information that makes it possible for the user to match data from various registers and surveys. When a new survey is created or when a request comes to Statistics Sweden MetaPlus can be used to search the metadata on micro data for relevant statistical units, populations and variables to se if the issue can be solved by using already collected data (completely or partly).

26. Information on steps and/or occurrences that affect (may affect) comparability over time can be documented and reported in MetaPlus. For handling time series and information concerning history and breaks in time series events (e.g. changes in definitions, etc.) can be stored in the system.

Survey design

27. The search function in MetaPlus can be used when constructing sample frames to find suitable frame data. Information on estimation, correction, editing, coding and imputation can be stored in the system and be of great help for users who are interested in using a data material and want to know of its quality and how the survey has been carried through.

System development

28. MetaPlus contains metadata, i.e. information concerning content. No data are stored in the system. The system contains information on where data are physically stored. This information and information on the data formats used to store the variables can be used when modelling new databases. Value sets in MetaPlus can be used as code tables in the production systems, e.g., when editing and coding.

Documentation

29. MetaPlus is designed to provide a good work situation for those who are carrying out the documentation work. The main idea is that information already documented should be easily re-useable. The metadata from previous survey rounds can be reused when documenting the latest one. Metadata needs only be created and stored once. This means that not only standards and classifications, but also metadata created and stored by others in the system may be used. Metadata reused in this way may be linked to several documentations. This means that the majority of the information in MetaPlus only needs to be documented once.

30. In order to improve metadata quality MetaPlus allows documentation work to be ongoing in parallel with the production process. This means that metadata can be captured when it is to be used.

Links to other metadata systems

31. In order to be able to handle the different types of use for MetaPlus presented above the system has to be able to co-operate with other parts of the metadata and administrative systems at Statistics Sweden, such as: SCBDOK, About the Statistics, the Quality Declaration, the Product Database, the Archiving System, the Personal Data Act Administration System, the System Development Model, and in the future a Questionnaire Repository. Hence, the system is designed to allow additions of new functionality.

The search function

32. Several of the examples above require a complex search function to provide the metadata overview. The system contains a well structured search function that makes it possible to combine different metadata objects in the search providing this functionality.

Harmonization

33. The MetaPlus system can be used as a harmonization and standardisation tool. This gives possibilities to promote shared usage of data in different statistical activities. The system therefore contains standards for object classes, variables and value domains with clear and harmonized definitions. This means that when documenting a survey in the system a certain degree of harmonization and standardisation is reached since metadata already stored is used or a connection to existing metadata is made. MetaPlus is the main tool for the harmonization work since it provides a good overview of all the data at Statistics Sweden. It thereby gives possibilities to identify areas where co-ordination is needed, and a harmonization project for that field can be initiated.

Improving production efficiency

34. As mentioned several times above the MetaPlus system provides an overview of the data used in the production at Statistics Sweden. This means that data can be used more efficiently, which in its turn will reduce the respondent burden by using previously collected data to a larger extent. The harmonisation of the production also leads to a more efficient production process.

Statistical Data warehouse

35. In order to build a structured data storage (a statistical data warehouse) formalized metadata is a prerequisite. Metadata is the principal that provides and withholds the structure in the data storage. Metadata makes is possible to find data and provides support for usage of the data.

VI. THE METAPLUS DATA MODEL

36. The MetaPlus data model is based on international standards, experiences from other statistical organisations and previous experiences in the field at Statistics Sweden.

37. The ISO/IEC standard 11179, *Information technology – Metadata registries* has been the foundation for building the conceptual model. The ongoing work in the Neuchâtel group is the other main influence for the conceptual model (The Neuchâtel group is a co-operation between a number of national statistical organisations, among those Statistics Sweden).

38. The major differences between the MetaPlus model and the ISO 11179 and Neuchatel models are found in the terminology. The concepts used in MetaPlus apply to concepts in the Statistics Sweden environment.

The MetaPlus concepts *register*, *register variant* and *register version* do not have any direct equivalents in other models.

39. Compared to the models developed by other statistical organisations there is a clear difference in how MetaPlus distinguishes between the conceptual model and the physical adaptation. This is most clearly seen when looking at the distinction between the concepts *variable* (concept) and *column* (physical adaptation).

40. One of the bases for the model and a difference compared to earlier systems at used at Statistics Sweden is to promote reuse. A variable created and defined once can be reused when documenting a later version of the same survey, but also by other surveys using the same variable.

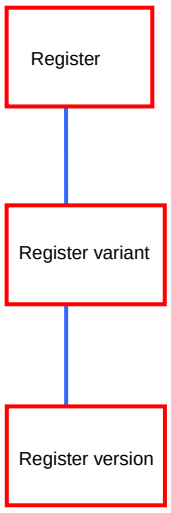
41. In contrast to earlier systems MetaPlus makes it possible to start documenting registers (surveys) and variables without having the physical storage in place. The metadata can be documented in stages, which means that one does not have to document everything when all the steps of the production process are done. This makes documentation the last step of the production process. Using MetaPlus the documentation can therefore be finalized at the same time as a phase of the production process is finished.

A. Basic structure and concepts

42. This section describes the MetaPlus structure on a basic level, the elementary concepts in MetaPlus are defined and their uses in the model are described.

The register structure

43. The metadata content in MetaPlus is connected to a hierarchical register structure: *register – register variant – register version*. In this context the term register is used to describe the data matrix that is used to produce statistics (it may be a final observation register).



44. The register structure is hierarchical. Thus the characteristics are inherited to the level below. One register has one or several register variants and the register variant can have one or several register versions.

45. The MetaPlus application is designed according to this basic structure and the other content is applied to it.

Register

46. A register represents a general container of all data necessary to produce the statistics. Every documentation in MetaPlus must be connected to a register. A register is in most cases stable over time. Source registers, from which one or a few variables are retrieved to the documented register, are not included here. They are identified as sources when documenting a certain variable.

Registers used to support the production process are documented as supporting registers. The register concept has no direct equivalent in ISO 11179.

Register variant

47. The system requires every register to have at least one register variant. When needed, a register can be partitioned into several register variants that differ in definition of register population and/or variable content. Input data registers and final observation registers are normally different register variants of a register. At different stages during the production phases various other variants may be created, e.g., output registers or variants for commissions or other purposes like commissioned work, etc. A register variant normally lasts as long as the register it belongs to. The register variant concept has no direct equivalent in ISO 11179.

Register version

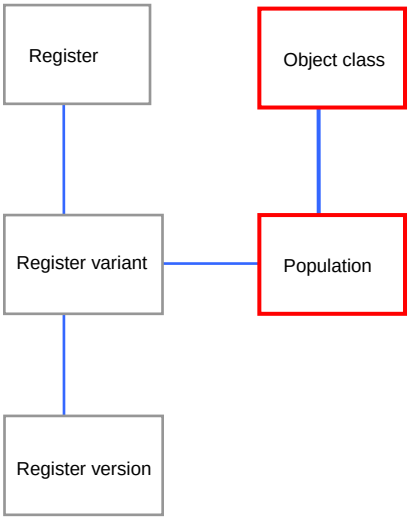
48. A register version is a physical register that can be considered as a realisation of a register variant at a certain point of time, i.e. the final observation register from a specific survey round tied to a certain population (the object set described in the register variant limited in time and space to fit a certain statistical purpose or activity). The register version is a list of objects belonging to a register population. The list can contain all the objects or some of them (a sample). The objects in the register version can be identifiable or anonymised. A register version can not be reused altogether, but may be copied to form the foundation for a new register version relating to another point of time or period of time. The register version concept has no direct equivalent in ISO 11179.

Example

Register	The population register	The business register	Energy use in manufacturing industry	The labour force survey
Register variant	Families	The SAMU-version November	Energy use in manufacturing industry	LFS withdrawal
Register version	2000-12-31	November 2002	2003	2005

Population and object class

49. The population is defined in relation to a register variant. A register variant may contain several populations. The population is a delimitation of an object class. Linking the register variant to the population is initiated by identifying its object class.



Object class

50. An object class is an abstraction of an object. An object is an independently existing phenomenon, with no relation to time and space. Whether a given object class is also a statistical object is determined by

a relation to a statistical activity. Several object classes may be used in different roles in the production process of a statistical activity. It can be used as a frame element, source, register object or target object.

51. An object class is an abstraction of a statistical object. It can be a super type of another object class, for example organisation and local unit. An object class can be a part of an other object class, for example one or several persons make up a household, or they can make up a new object class in combination with an other object class. Employment is, e.g., made up of the combination of a person and a local unit. In MetaPlus these are called complex object classes, as opposed to basic object classes.

52. The model describes *register objects*, i.e. the object class that data is related to in the physical register and the *target objects*, i.e. the objects that make up the survey population.

53. In the practical work we try to stimulate reuse of already created object classes and keep total amount down to a minimum. It is therefore only administrators who can create new object classes. At present the system contains the following object classes:

Basic object classes, hierarchically ordered

Actor
- Person
- Organisation
  - Enterprice
    - Kind of activity unit
    - Local unit
      - Local kind of activity unit

Utility
- Firm objects
  - Real estate
    - Building
      - Flat
    - Assessed unit
      - Valuation unit
  - Land and water
- Moveable objects
  - Dwelling
  - Financial object
  - Viecle
  - Product
    - Goods
    - Services

Complex object classes

Employment	(person, local unit)
Household	(person, person)
Product offering	(product, local unit)
RTBFamily	(person, person)
Transport	(viecle, product, firm object)
Traffic accident	(viecle, viecle, firm object)

The *object class* concepts in MetaPlus and ISO 11179 are equivalents.

Population

54. A population describes the set of objects included in a statistical activity, regardless of a survey round. The population description may relate to the population that is the origin of the register objects, *the register population*, or to the population that is the origin of the target objects, the *survey population*. In many cases the two populations are identical. In practice a population may be completely reused for several surveys and survey rounds. There is, however, no restriction to prevent a user from creating a new population when needed. The population concept has no direct equivalent in ISO 11179.

Context population

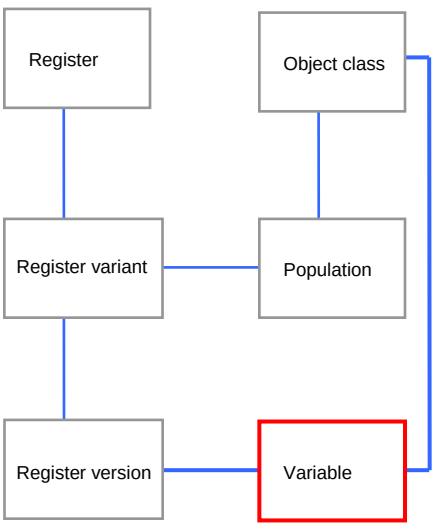
55. The model uses one further definition of population: the context population that describes the set of objects that a survey collects data on during a certain survey round. The context population is defined in relation to a register version. A register version may contain several context populations. A context population denotes both a context register population and a context survey population. The realisation of the register population and the survey population gives the actual populations for a specific survey round. A specific context population is normally not possible to reuse, since it is coupled to one statistical activity and a specific survey round (register version). The context population concept has no direct equivalent in ISO 11179.

    Example:

Register	The population register	The business register	Energy use in manufacturing industry	The labour force survey
Object class	Person	Local unit	Local unit	Person
Population	Persons registered as living in Sweden	All local units in Sweden	Enterprises and local units in NACE 10-37	Persons registered as living in Sweden aged 16-64 years
Context population	Persons registered as living in Sweden 2000-12-31	All local units in Sweden 2002-11-27	Enterprises and local units in NACE 10-37 2006-01-01	Persons registered as living in Sweden aged 16-64 years 2005-12-01

The variable structure

56. Variables are added at the register version level. It can be done in different ways: using an already defined variable, copy and modify a variable or define a new variable. The variable is specified by the link to the object class created when defining the population. This also links the variable to its object class.



57. There are three concepts related to variables in the conceptual model: variable, object variable and context variable.

Variable

58. A variable is an expression of a characteristic (attribute) that can be linked to one or several statistical object types. It may vary on a conceptual level independent of object or use in a specific statistical activity. A variable can be used in several surveys and survey rounds, but new variables can be created when needed. At present the system contains about 3 000 variables. The MetaPlus *variable* concept corresponds closely to the *property* concept of the *data element* used in ISO 11179.

Object variable

59. An object variable is a conceptual definition of a characteristic (variable) connected to a specific object class with no given level of detail. Object variables are created when needed by the system and are made up by combining object class and variable. The *object variable* in MetaPlus is about the same as the *data element* concept in ISO11179.

Context variable

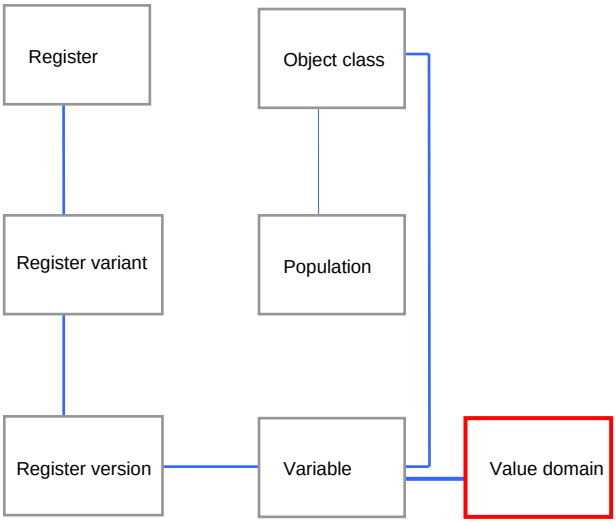
60. A context variable is the combined representation of an object class, a variable and a value domain for a specific register version (or survey round). The context variable defines the metadata for a variable that are specific for that register version. A context variable can exist for all or some of the objects in a survey population. A context variable is not reusable. Each documentation creates a new context population and a new set of context variables. The *context variable* in MetaPlus corresponds to the ISO11179 *data element*.

*Example**:***

Register	The population register	The business register	Energy use in manufacturing industry	The labour force survey
Variable	Number of children	Activity	Energy use	Local unit number
Object variable	Person number of children	The local unit's activity	The local unit's energy use	The persons local unit number
Context variable	Person number of children 2000-12-31	The local unit's activity according to SNI 92 5-digit level 2002-11-27	The local unit's energy use 2006-01-01	The persons local unit number 2005-12-01

Value domains and classifications

61. In MetaPlus a value domain is connected to a variable in a register version. There are different ways to create this link. One way is to use a value domain that has been connected to the variable by someone else. Another way is to connect it to a value domain that is stored in MetaPlus. From this value domain one can create a new value domain by modifying it, based on the original value domain. It is also possible to create a new value domain, for example by importing a text file.



62. In the complete conceptual model there are three terms used to describe value domains: conceptual value domain, value domain and value. The classification and value domain part of MetaPlus is based on the Neuchâtel terminology for classifications.

Conceptual value domain

63. A conceptual value domain is made up by a definition and a level of detail for a value domain and the meaning of its categories. The conceptual value domain includes a classification or a type of measure unit. The defined level of detail is connected to a code in the value domain. A conceptual value domain can be reused in several value domains. The *conceptual value domain* in MetaPlus is about the same as *conceptual domain* in ISO11179.

Value domain

64. A value domain shows the level of detail for a variable and its representation, i.e. a measure unit or alphanumeric specification and the definition of the different categories given a certain level of detail and how it is alphanumerically represented (by a code). Given the knowledge of what value domain a variable has the value of a single object are interpretable. One value domain can be reused by several context variables. MetaPlus allows for three ways to define a value domain:
    1. Enumerated
    2. Continuous
    3. Descriptive

    The *value domain* in MetaPlus is equivalent to *value domain* in ISO11179.

Value

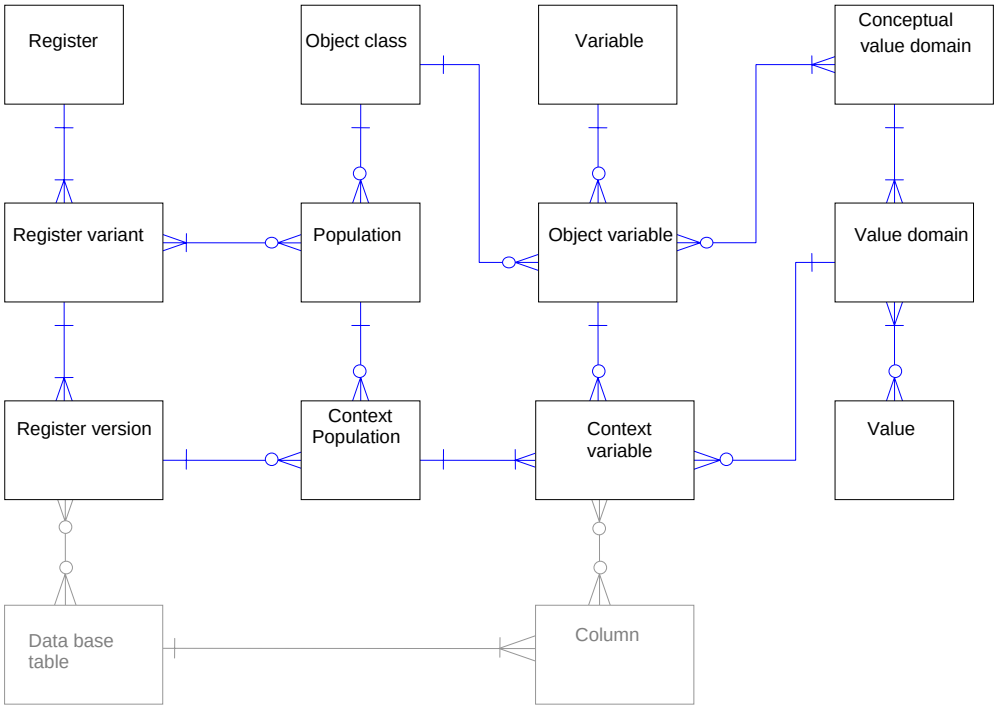
65. Values are the individual representations of an enumerated value domain, without a measure unit. A value can be reused in several value domains. The *value* in MetaPlus is equivalent to *value meaning* in ISO11179.

Example:

Register	The population register	The business register	Energy use in manufacturing industry	The labour force survey
Conceptual value domain	Continuous	SNI	Continuous	Text
Value domain	Number	SNI 1992 5-digit	Quantity	Text number format

B. The conceptual model

66. From the basic concepts described in the section above the MetaPlus conceptual model can be built. See below:



67. The figure is drawn to make each column represent ”families” of concepts: register, object/population, variable and value domain/classification. The rows represent stability over time, or the possibilities for reuse: the higher up the better possibility for reuse.

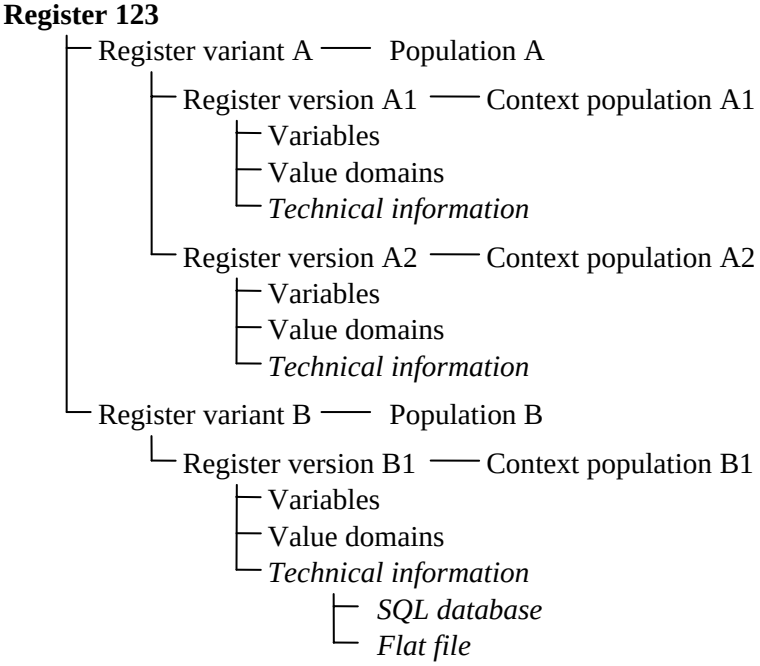
68. In addition to the above defined concepts, two new ones are introduced in the bottom row, meaning that they are the ones least possible to reuse. They are drawn in grey in the figure since they are not necessary for the rest of the model. It represents the part of the model that makes up the physical storage, the technical metadata.

69. One advantage of keeping technical metadata separated from content metadata is that one register version may be stored physically in several copies for different systems: e.g., one copy in a relational database, and another copy in a flat file. The whole documentation can be kept intact when the new copy is created, just adding a reference to a new database table. This also means that variable definitions can be standardised and harmonised without making changes in the production environment.

70. Similarly a column denotes the physical storage of a context variable. Bear in mind that a context variable can be stored in one or several columns and also that a column can contain one or several context variables.

C. Content structure

71. The content structure in a fictitious register can be described like this:



72. The register in the example has two variants, A and B since they represent two different populations. Register variant A has two versions, or survey rounds (documented as A1 and A2). Version B only has one documented register version (B1). Register variant B is physically stored in two instances, one relational data base (SQL), and one flat file (e.g., in a long term archived file).

73. The documentation work can be carried out in any order, provided that the basic structure, register – register variant – register version, is in place. When an existing system is to be documented it may be natural to start with the technical part and then add the conceptual parts. In other situations it can be more natural to start documenting variables and value domains.

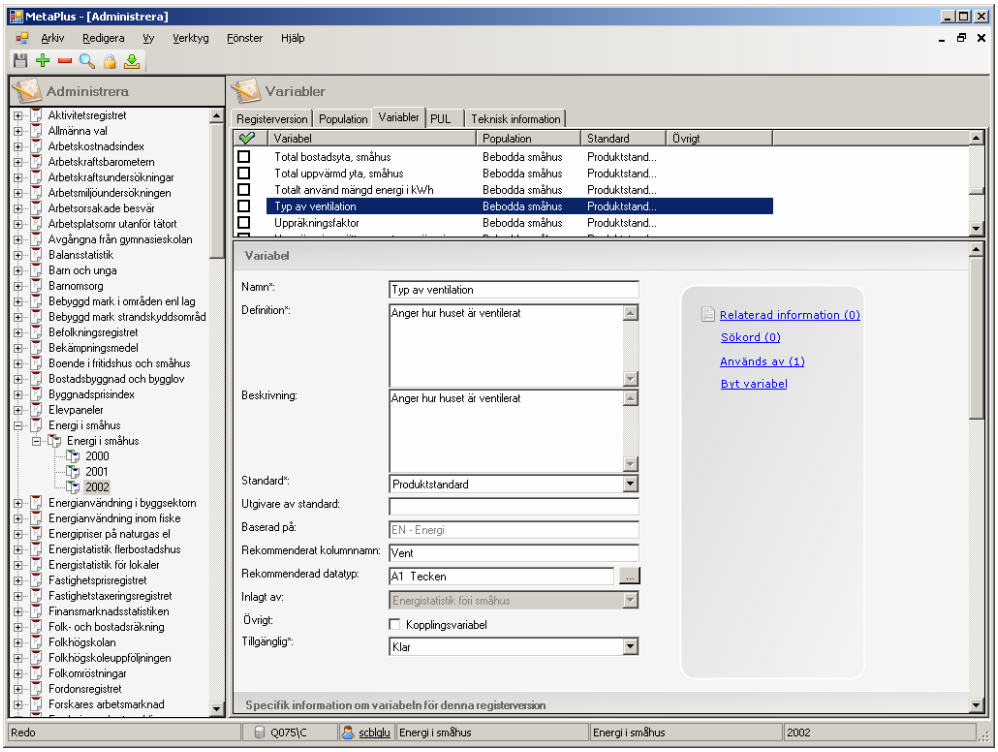
V. APPLICATIONS AND FUNCTIONS

74. The heart of MetaPlus is the central database. A number of applications and functions are using this database.

A. The main MetaPlus application

75. Entering data and maintaining data is carried out through the main application. It is operated by those who document, check and maintain documentations, work with classifications, and others.

76. The application is designed to let the operator have the register structure available on the left-hand side of the screen in most situations. The user interface makes use of tabs to let the operator choose to work with populations, variables, etc.



77. The application uses common functions built into the system to keep track of a user's permissions to view or change information.

B. Web based user interface

78. A separate application using a web browser user interface to search and present completed documentations is currently under development. It is planned to be used on Statistics Sweden's intranet as well as on its public web site. No data entry or maintenance will be possible with this application. Using this application it will be possible to download classifications, complete or partly.

C. Special functions

a. The Personal Data Act

79. According to the Personal Data Act all owners of registers containing information on individuals must on demand provide a person with information on what the register(s) contain on him/her. MetaPlus includes functions to support the retrieval and handling of necessary information.

b. Archiving

80. Most registers are required by law to be archived. Archiving must include data as well as metadata. Statistics Sweden has developed a dedicated support system for these tasks. That system uses MetaPlus to retrieve the necessary metadata and to locate the data to be archived.

c. Private applications

81. The classifications and value sets stored in MetaPlus can, and should, be used directly by production systems, instead of creating own, local code tables. Currently several projects are working on the development of a general, common systems architecture. Pending the results of these projects it is not possible to use MetaPlus directly from a private production system.

VI. FUTURE PLANS

82. In addition to the development plans already mentioned there are several ideas on expanded functionality. This chapter gives a few examples.

A. Metadata and the production process

83. As mentioned above MetaPlus was initially constructed to replace the Metadok system and thereby provide chapter 3 for SCBDOK. Several other issues has been taken into account besides this main goal; one such issue was to link MetaPlus to the production process and to handle process data. The issue of what this exactly means is currently not clear since Statistics Sweden is going through intensive work standardising our processes and tools, a project called Lotta (further reading: *PROCESS REENGINEERING AT STATISTICS SWEDEN*, Sundgren, WP 22 Meeting on the Management of Statistical Information Systems MSIS 2007, Geneva, 8-10 May 2007 and *Reducing costs and raising quality through standardised processes and tools at Statistics Sweden*, Bergdahl and Elvers). A part of this process involves describing the different parts of the production process and investigating what process data is required in the different sub processes. Since this work is not yet finalised process data and the use of MetaPlus in the production is not clear at this stage.

84. The MetaPlus model was not constructed for only handling metadata for micro data and our experience so far is that MetaPlus can be used to describe data matrixes at different stages in the production process. There will probably be a close connection between a future system for handling process data and MetaPlus.

B. Integrated metadata system

85. The various parts of Statistics Sweden’s metadata system (SCBDOK, Description of the statistics, MetaPlus, Metadata for the Statistical database) are loosely coupled, sometimes overlapping systems. They should form a well integrated documentation and metadata system where users, e.g., can search for information using MetaPlus's search system and then drill down to seamlessly retrieve information from SCBDOK.

C. Distributed use

86. The first version of MetaPlus is an internal system to document data produced at Statistics Sweden. Future versions should support creation of documentations at any organisation that produces official statistics.

D. Language support

87. MetaPlus version 1.0 only supports Swedish. The model, the application and the data are all in Swedish. Future versions should support at least Swedish and English.

E. Form and query database

88. Single survey queries (interview, web or paper based) should be documented and linked to variables. Questionnaires (sets of queries) should be documented in the same way and linked to register versions.

F. MONA

89. MONA (Micro data ON-line Access) is Statistics Sweden's system to give researchers access to unidentifiable micro data via the Internet. MONA users obviously need metadata. For security reasons the MONA system resides in an environment that uses a separate, isolated network, which makes it unable to reach the MetaPlus database. The current plans include creating a replicated copy of the MetaPlus database available in the MONA environment.

G. Standardised access and export

90. All MetaPlus information is stored in a database, available through SQL commands for users with adequate access permissions. Statistics Sweden is currently analysing the possibilities of adapting a service oriented system architecture (SOA). This would probably mean that the exchange format will be based on XML standards, which would in that case also be possible to use to export entire models or single items.

H. Link to modelling tool

91. As much as possible of existing variables or entire structures should by re-used when modelling new databases. This may be considered a special case of export.

VII. OVERVIEW OF EXAMPLES

Register	The population register	The business register	Energy use in manufacturing industry	The labour force survey
Register variant	Families	The SAMU-version November	Energy use in manufacturing industry	LFS withdrawal
Register version	2000-12-31	November 2002	2003	2005
Object class	Person	Local unit	Local unit	Person
Population	Persons registered as living in Sweden	All local units in Sweden	Enterprises and local units in NACE 10- 37	Persons registered as living in Sweden aged 16-64 years
Context population	Persons registered as living in Sweden 2000-12-31	All local units in Sweden 2002-11-27	Enterprises and local units in NACE 10- 37 2006-01-01	Persons registered as living in Sweden aged 16-64 years 2005-12-01
Variable	Number of children	Activity	Energy use	Local unit number
Object variable	Person number of children	The local unit's activity	The local unit's energy use	The persons local unit number
Context variable	Person number of children 2000-12-31	The local unit's activity according to SNI 92 5digit level 2002-11-27	The local unit's energy use 2006-01-01	The persons local unit number 2005-12-01
Conceptual value domain	Continuous	SNI	Continuous	Text
Value domain	Number	SNI 1992 5-digit	Quantity	Text number format