Appendix 1_ Worked example of the GSIM Statistical Classification Model

58.In this appendix, a worked example for all object types and most attributes, based mainly on the Standard Industrial Classification (SIC 2007), is provided to facilitate understanding.

59. Attributes or terms used in the descriptions which are underlined, refer to an object type listed and described elsewhere in the model.

Classification Family

60. A Classification Family is a group of Classification Series related from a particular point of view. The Classification Family is related by being based on a common Concept (e.g. economic activity).

61. Different classification databases may use different types of Classification Families and have different names for the families, as no standard has been agreed upon.

Identifier: A Classification Family is identified by a unique identifier.
E.g.: IA
Name: A Classification Family has a name.
E.g.: Industrial activities
Classification Series: A Classification Family may refer to a number of Classification Series.
E.g.: Standard Industrial Classification, Classification of CPA codes
See also: Classification Series

62. For a practical example, see: http://stabas.ssb.no/MainFrames.asp?Language=en

Classification Series

63. A Classification Series is an ensemble of one or more Statistical Classifications, based on the same concept, and related to each other as versions or updates. Typically, these Statistical Classifications have the same name (for example ISIC or ISCO).

Identifier: A Classification Series is identified by a unique identifier, which may typically be an abbreviation of its name.
E.g.: SIC
Name: A Classification Series has a name as provided by the owner.
E.g.: Standard Industrial Classification
Description: Short general description of the Classification Series, including its purpose, its main subject areas etc.
E.g.: SIC is primarily a statistical standard. The standard will be the basis for coding units according to principal activity in e.g. a business register. The SIC is one of the most important standards of economic statistics, and it will make it possible to compare and analyze statistical data both at the national/international level and over time. SIC is also used for administrative purposes.
Context: A Classification Series can be designed in a specific context.
E.g.: Not relevant
Objects/units classified: A Classification Series is designed to classify a specific type of object/unit according to a specific attribute.
E.g.: Economic activities
Subject areas: Areas of statistics in which the Classification Series is implemented.
E.g.: National accounts, Energy and manufacturing
Owners: The statistical office or other authority, which created and maintains the Statistical Classification(s) related to the Classification Series. A Classification Series may have several owners.
E.g.: Statistics Norway (The national version of SIC)
Keywords: A Classification Series can be associated with one or a number of keywords.
E.g.: Industry, business, legal units
Classification Family: Classification Series may be grouped into Classification Families. Shows which Classification Family the Classification Series belongs.
E.g.: Industrial activities
Statistical Classification: A Classification Series has at least one Statistical Classification.

E.g.: SIC94, SIC 2002, SIC 2007
Current Statistical Classification: If there are several Statistical Classifications related to a Classification Series, one Statistical Classification *may be* assigned as the currently valid Statistical Classification.
E.g.: SIC 2007

64. For a practical example, see: http://www.ssb.no/metadata/classification/stabas/342101/en

Statistical Classification

65. A Statistical Classification is a set of categories which may be assigned to one or more variables registered in statistical surveys or administrative files, and used in the production and dissemination of statistics. The categories at each level of the classification structure must be mutually exclusive and jointly exhaustive of all objects/units in the population of interest.

66. The categories are defined with reference to one or more characteristics of a particular population of units of observation. A Statistical Classification may have a flat, linear structure or may be hierarchically structured, such that all categories at lower Levels are sub-categories of categories at the next Level up. Categories in Statistical Classifications are represented in the information model as Classification Items.

Identifier: A Statistical Classification is identified by a unique identifier. The identifier of a Statistical Classification principally considered to be a version or update is typically an abbreviation of its name. It is often distinguished from other versions/updates of the same Classification Series by reference to a revision number or to the year when it came into force. The identifier of a Statistical Classification that is considered to be variant typically refers to (contains) the identifier of its base Statistical Classification.
E.g.: SIC2007
Name: A Statistical Classification has a name as provided by the owner or maintenance unit.
E.g.: Standard Industrial Classification (SIC 2007)
Introduction: The introduction provides a detailed description of the Statistical Classification, the background for its creation, the classification variable and objects/units classified, classification rules etc. See Appendix 2 for a checklist of possible topics to be included in the introduction.
E.g.: Standard Industrial Classification is primarily a statistical standard. In practice, this means the standard will be the basis for coding units according to the most important activities in Statistics Norway's Business register and in the Central Coordinating Register for Legal Entities. SIC2007 is one of the most important standards of economic statistics, and it makes it possible to compare and analyze statistical data at the national/international level and over time.
Release date: Date on which the Statistical Classification was released.
E.g.: 01.01.2009
Termination date: Date on which the Statistical Classification was superseded by a successor version or otherwise ceased to be valid.
Current: Indicates whether or not the Statistical Classification is currently valid.
E.g.: Yes
Maintenance unit: The unit or group of persons within the organisation who are responsible for the Statistical Classification (i.e. for maintaining, updating and changing it).
E.g.: 810 - Division for statistical populations
Contact persons: Person(s) who may be contacted for additional information about the Statistical Classification.
E.g.: Ida Skogvoll, isk@ssb.no
Legal base: Indicates that the Statistical Classification version is covered by a legal act or by some other formal agreement.
E.g.: Council Regulation (EEC) No. 1893/2006
Publications: A list of the publications, including print, PDF, HTML and other electronic formats, in which the Statistical Classification has been published.
E.g.: http://www.ssb.no/a/publikasjoner/pdf/nos_d383/nos_d383.pdf
Name types: A list of the defined types of alternative item names available for the Statistical Classification. Each name type refers to a list of alternative item names.
E.g.: Short titles (for use in our dissemination tables)
Languages available: A Statistical Classification can exist in one or several languages. Indicates the languages available, whether the version is completely or partially translated, and which part is available in which language.
E.g.: Norwegian (bokmål), Norwegian (nynorsk), English
Copyright: Statistical Classifications may have restricted copyrights. Such Statistical Classifications might be excluded from downloading. Notes the copyright statement that should be displayed in official publications to indicate the copyright owner.
E.g.: Not relevant
Dissemination allowed: Indicates whether or not the Statistical Classification may be published or otherwise disseminated (e.g. electronic dissemination).
E.g.: Yes
Classification Series: A Statistical Classification is a version or update of one specific Classification Series
E.g.: Standard Industrial Classification

Levels: The structure of a Statistical Classification is defined by its Levels (classification levels). Include here links to the relevant Levels.
E.g.: Section, Division, Group, Class, Subclass
Classification Items: A Statistical Classification is composed of categories structured in one or more Levels. Each category is represented by a Classification Item, which defines the content and the borders of the category.
E.g.: **01.24** Growing of pome fruits and stone fruits
Correspondence Tables: A Statistical Classification may be linked to other classification versions or classification variants through Correspondence Tables. Include here links to any relevant Correspondence Tables.
E.g.: Correspondence table SIC 2007/SIC 2002
Classification Indexes: A Statistical Classification can be associated with one or a number of Classification Indexes in which Index Entries are linked to the appropriate Classification Item. Include here links to any relevant Classification Indexes.
E.g.: SN07XTD_en
Version: Indicates if the Statistical Classification is a version.
E.g.: Yes
Update: Indicates if the Statistical Classification is an update.

Floating: Indicates if the Statistical Classification is a floating Statistical Classification. In a floating Statistical Classification, a validity period should be defined for all Classification Items which will allow the display of the item structure and content at different points of time.
E.g.: No
Predecessor: For those Statistical Classifications that are versions or updates, notes the preceding Statistical Classification of which the actual Statistical Classification is the successor. .
E.g.: SIC 2002
Successor: Notes the Statistical Classification that superseded the actual Statistical Classification.
Derived from: A Statistical Classification can be derived from one of the classification versions of another Classification Series. The derived Statistical Classification can either inherit the structure of the classification version from which it is derived, usually adding more detail, or use a large part of its Classification Items, rearranging them in a different structure. Indicates the classification version from which the actual Statistical Classification is derived.
E.g.: NACE Rev.2
Changes from previous version or update: A summary description of the nature and content of changes from the previous version or update. Specific changes are recorded in the Classification Item object under the "changes from previous version and update" attribute.
E.g.: Information presented in **Publications** and in Correspondence Table
Updates possible: Indicates whether or not updates are allowed within the classification version, i.e. without leading to a new version. Indicate here what structural changes, if any, are permissible within the version. Note whether Classification Items can be added to the structure and whether they can be revalidated or invalidated. Such changes are more likely to be permissible in floating classifications. Also indicate whether changes to such things as Classification Item names and explanatory notes that do not involve structural changes are permissible within the version.
E.g.: No
Updates: Summary description of changes which have occurred since the most recent classification version or classification update came into force.
E.g.: Not relevant:
Variants available: Identifies any variants associated with this version.
E.g.: Variant of SIC - Environmental accounts (SIC2007)
Variant: For those Statistical Classifications that are variants, notes the Statistical Classification on which it is based and any subsequent versions of that Statistical Classification to which it is also applicable.
Changes from base Statistical Classification: Describes the relationship between the variant and its base Statistical Classification, including regroupings, aggregations added and extensions.
Purpose of variant: If the Statistical Classification is a variant, notes the specific purpose for which it was developed.
See also: Classification Series, Level, Classification Item, Correspondence Tables, Classification Index.

67. For a practical example, see: http://www.ssb.no/metadata/classversion/stabas/8118001/en

Level

68. A Statistical Classification has a structure which is is composed of one or several Levels. A Level often is associated with a Concept, which defines it. In a hierarchical Statistical Classification the Classification Items of each Level but the highest are aggregated to the nearest higher Level. A linear Statistical Classification has only one Level.

Identifier: A Level is identified by a unique identifier.
E.g.: SIC2007L5

Level number: The number associated with the Level. Levels are numbered consecutively starting with Level 1 at the highest (most aggregated) Level.
E.g.: 5
Level name: The name given to the Level.
E.g.: Subclass
Description: Text describing the content and particular purpose of the Level.
E.g.: Subclass is the most detailed level and describes the national level in SIC.
Number of Classification Items: The number of Classification Items (Categories) at the Level.
E.g.: 817
Code type: Indicates whether the code at the Level is alphabetical, numerical or alphanumerical.
E.g.: Numerical
Code structure: Indicates how the code at the Level is constructed of numbers, letters and separators.
E.g.: nn.nnn
Dummy code: Rule for the construction of dummy codes from the codes of the next higher Level (used when one or several Categories are the same in two consecutive Levels).
E.g.: Not relevant
Items: An ordered list of the Categories (Classification Items) that constitute the Level.
E.g.: http://www.ssb.no/metadata/codelist/stabas/8118013/en
See also: Statistical Classification, Classification Item

Correspondence table

69. A Correspondence Table expresses the relationship between two Statistical Classifications. These are typically: two versions from the same Classification Series; Statistical Classifications from different Classification Series; a variant and the version on which it is based; different versions of a variant. In the first and last examples, the Correspondence Table facilitates comparability over time. Correspondence relationships are shown in both directions.

Identifier: A Correspondence Table is identified by a unique identifier, which may typically include the identifiers of the versions or variants involved.
E.g.: C20072002
Name: A Correspondence Table has a name as provided by the owner.
E.g.: SN2007, SN2002
Description: The description contains information about the scope and aim of the Correspondence Table and the principles on which it is based.
E.g.: The Correspondence Table shows the changes between SIC versions 2002 and 2007 and makes comparability over time possible.
Owners: The statistical office, other authority or section that created and maintains the Correspondence Table. A Correspondence Table may have several owners.
E.g.: Statistics Norway
Maintenance unit: The unit or group of persons who are responsible for the Correspondence Table, i.e. for maintaining and updating it.
E.g.: 810 - Division for statistical populations
Contact persons: The person(s) who may be contacted for additional information about the Correspondence Table.
E.g.: Ida Skogvoll, isk@ssb.no
Publications: A list of the publications in which the Correspondence Table has been published.
E.g.: The Correspondence Table is only published in the classification database
Source: The Statistical Classification from which the correspondence is made.
E.g.: SIC 2007
Target: The Statistical classification(s) to which the correspondence is directed. There may be multiple target Statistical Classifications associated with the Correspondence Table.
E.g.: SIC 2002
Source level: The correspondence is normally restricted to a certain Level in the source Statistical Classification. In this case target Classification Items are assigned only to source Classification Items on the given Level. If no Level is indicated target Classification Items can be assigned to any Level of the source Statistical Classification.
E.g.: Level 5
Target level: The correspondence is normally restricted to a certain Level in the target Statistical Classification. In this case source Classification Items are assigned only to target Classification Items on the given Level. If no Level is indicated, source Classification Items can be assigned to any Level of the target Statistical Classification.
E.g.: Level 5
Relationship type: A correspondence can define a 1:1, 1:N, N:1 or M:N relationship between source and target Classification Items.
E.g.: M:N

Floating: If the source and/or target Statistical Classifications of a Correspondence Table are floating Statistical Classifications, the date of the Correspondence must be noted. The Correspondence Table expresses the relationships between the two Statistical Classifications as they existed on the date specified in the Correspondence Table.
E.g.: No
See also Statistical Classification, Classification Item, Level, Map

70. For a practical example, see: http://stabas.ssb.no/CorrTabFrames.asp?ID=8364101&Language=en

Classification index

71. A Classification Index is an ordered list (alphabetical, in code order etc.) of Classification Index Entries. A Classification Index can relate to one particular or to several Statistical Classifications. A Classification Index shows the relationship between text found in statistical data sources (responses to survey questionnaires, administrative records) and one or more statistical classifications. A Classification Index may be used to assign the codes for Classification Items to observations in statistical collections.

Identifier: A Classification Index is identified by a name. If there are several Classification Indexes in different languages, the language should be part of the Classification Index name. If there are several Classification Indexes for different purposes, the purpose should be part of the Classification Index name.. If there are several Classification Indexes that differ by languages and by purpose, the language and the purpose should be part of the Classification Index name.
E.g.: SN07XTD_en
Release date: Date when the current version of the Classification Index was released.
E.g.: 01.01.2009
Maintenance unit: The unit or group of persons within the organisation responsible for the Classification Index, i.e. for adding, changing or deleting Classification Index entries.
E.g.: 810 - Division for statistical populations
Contact persons: Person(s) who may be contacted for additional information about the Classification Index.
E.g.: Ida Skogvoll, isk@ssb.no
Publications: A list of the publications in which the classification index has been published.
Languages: A classification index can exist in several languages. Indicates the languages available. If an Classification Index exists in several languages, the number of entries in each language may be different, as the number of terms describing the same phenomenon can change from one language to another. However the same phenomena should be described in each language.
E.g.: Norwegian (bokmål), English
Corrections: Verbal summary description of corrections, which have occurred within the Classification Index. Corrections include changing the item code associated with an index entry.
Coding instructions: Additional information which drives the coding process for all entries in a Classification Index.
Statistical Classification: A Classification Index is related to one particular Statistical Classification.
E.g.: SIC 2007

Classification item

72. A classification item represents a Category at a certain Level within a Statistical Classification. It defines the content and the borders of the Category. A statistical object/unit can be classified to one and only one Classification Item at each level of a Statistical Classification.

Code: A Classification Item is identified by an alphabetical, numerical or alphanumerical
code, which is in line with the code structure of the Level. The code is unique
within the Statistical Classification to which the Classification Item belongs.
E.g.: 01.620
Official name: A Classification Item has a name as provided by the owner or maintenance unit. The name describes the content of the Category. The name is unique within the Statistical Classification to which the Classification Item belongs, except for Categories that are identical at more than one Level in a hierarchical Statistical Classification.
E.g.: Support activities for animal production
Alternative names: A Classification Item can be expressed in terms of one or several alternative names. Each alternative name is associated with a name type.
E.g.: Short text: Support activities to agriculture
Explanatory notes: A Classification Item may be associated with explanatory notes, which further describe and clarify the contents of the category. Explanatory notes consist of:

General note: Contains either additional information about the Category, or a general description of the Category, which is not structured according to the "includes", "includes also", "excludes" pattern.
E.g.: Not relevant
Includes: Specifies the contents of the Category.
E.g.: Includes services, associated with the keeping of farm animals, in activities that increase reproduction, growth and performance in farm animals, testing of farm animals (control), maintenance of grazing areas, castration, cleaning of barns, insemination and covering, clipping of sheep, housing and care of farm animals.
E.g.: Includes activities in connection with agriculture carried out on a contract basis
Includes also: A list of borderline cases, which belong to the described Category.
E.g.: Includes also: Includes also shoeing of horses.
Excludes: A list of borderline cases, which do not belong to the described Category. Excluded cases may contain a reference to the Classification Items to which the excluded cases belong.
E.g.: Excludes: Renting out of areas exclusively for housing of farm animals is grouped under Other letting of real estate. Veterinary services is grouped under 75.000 Veterinary activities, Vaccination of farm animals is grouped under 75.000 Veterinary activities. Renting out of farm animals (e.g. cattle) is grouped under 77.390 Renting and leasing of other machinery, equipment and tangible goods n.e.c. Housing of domestic pets is grouped under 96.09 Other personal service activities n.e.c.
Level number: The number of the Level to which the Classification Item belongs.
E.g.: 5
Generated: Indicates whether or not the Classification Item has been generated to make the Level to which it belongs complete.
E.g.: No
Currently valid: If updates are allowed in the Statistical Classification, a Classification Item may be restricted in its validity, i.e. it may become valid or invalid after the Statistical Classification has been released. Indicates whether or not the Classification Item is currently valid.
E.g.: Not relevant
Valid from: Date from which the Classification Item became valid. The date must be defined if the Classification Item belongs to a floating Statistical Classification.
E.g.:
Valid to: Date at which the Classification Item became invalid. The date must be defined if the Classification Item belongs to a floating Statistical Classification and is no longer valid.
E.g.:
Future events: The future events describe a change (or a number of changes) related to an invalid Classification Item. These changes may e.g. have turned the now invalid Classification Item into one or several successor Classification Items. In describing these changes, terminology from the Typology of item changes, found in Appendix (1) should be used. This allows the possibility to follow successors of the Classification Item in the future.
E.g.: Not relevant
Changes from previous version of the Statistical Classification: Describes the changes, which the Classification Item has been subject to from the previous to the actual Statistical Classification. In describing these changes, terminology from the Typology of item changes, found in Appendix (1), should be used.
E.g.: Name and Code change**.** In SIC2002 this item was called " 01.420 Animal husbandry service activities, except veterinary activities". The history of the item is also documented in the Correspondence Table.
Updates: Describes the changes, which the Classification Item has been subject to during the life time of the actual Statistical Classification.
E.g.: Not relevant
Statistical Classification: The Statistical Classification to which the Classification Item belongs.
E.g.: SIC 2007
Parent item: The Classification Item at the next higher Level of the Statistical Classification of which the actual Classification Item is a sub item.
E.g.: 01.62 Support activities for animal production
Sub items: Each Classification Item, which is not at the lowest Level of the Statistical Classification, might contain one or a number of sub items, i.e. Classification Items at the next lower Level of the Statistical Classification.
E.g.: Not relevant
Linked items: Classification Items of other Statistical Classification with which the Classification Item is linked, either as source or target, through Correspondence Tables.
E.g.: 01.420 Animal husbandry service activities, except veterinary activities (SIC2002)
Case law: Refers to identifiers of one or more case law rulings related to the Classification Item.
E.g.: Not relevant
Case law descriptions: Case law descriptions refers to descriptions of the above case laws.
E.g.: Not relevant
Case law dates: Refers to dates of above case laws.
E.g.: Not relevant
See also Level, Classification Index Entry, Correspondence Item, Statistical Classification

Map

73. A map is an expression of the relation between a Classification Item in a source Statistical Classification and a corresponding Classification Item in the target Statistical Classification. The Map should specify whether the relationship between the two Classification Items is partial or complete. Depending on the relationship type of the Correspondence Table, there may be several maps for a single source or target Classification Item.

Source item: The source item refers to the Classification Item in the source Statistical Classification.
E.g.: 01.620 Support activities for animal production
Target item: The target item refers to the Classification Item in the target Statistical Classification.
E.g.: 01.420 Animal husbandry service activities, except veterinary activities
Partial/complete: specifies whether the relationship between the two Classification Items is partial or complete
E.g.: Complete
Valid from: Date from which the Map became valid. The date must be defined if the Map belongs to a floating Correspondence Table.
Valid to: Date at which the Map became invalid. The date must be defined if the Map belongs to a floating Correspondence Table and is no longer valid.
See also Statistical Classification, Classification Item, Correspondence Table

Classification Index Entry

74. A Classification Index Entry is a word or a short text (e.g. the name of a locality, an economic activity or an occupational title) describing a type of object/unit or object property to which a Classification Item applies, together with the code of the corresponding Classification Item. Each Classification Index Entry typically refers to one item of the Statistical Classification. Although a Classification Index Entry may be associated with a Classification Item at any level of the Statistical Classification, Classification Index Entries are normally associated with Classification Items at the lowest Level.

Text: Text describing the type of object/unit or object property.
E.g.: Animal husbandry
Statistical Classification: Identifies the Statistical Classification(s) to which the Classification Index Entry is associated.
E.g.: SIC2007
Codes: For each Statistical Classification to which the Classification Index Entry is associated, enter the code of the Classification Item in that Statistical Classification with which the Classification Index Entry is associated.
E.g.: 01.620 Support activities for animal production
Valid from: Date from which the Classification Index Entry became valid. The date must be defined if the Classification Index Entry belongs to a floating Classification Index.
Valid to: Date at which the Classification Index Entry became invalid. The date must be defined if the Classification Index Entry belongs to a floating Classification Index and is no longer valid.
Coding instructions: Additional information which drives the coding process. Required when coding is dependent upon one or many other factors.
See also Classification Item, Classification Index, Statistical Classification

---

null