

Generic Statistical Information Model (GSIM): Specification

(Version 1.1, December 2013)

About this document

This is aimed at metadata specialists, information architects and solutions architects. This document includes descriptions of information in a statistical organization. There are also a number of annexes, which include information about Exchange Channels, a glossary, UML class diagrams and the GSIM extension methodology.



This work is licensed under the Creative Commons Attribution 3.0 Unported License. To view a copy of this license, visit <http://creativecommons.org/licenses/by/3.0/>. If you re-use all or part of this work, please attribute it to the United Nations Economic Commission for Europe (UNECE), on behalf of the international statistical community.

I. Introduction	3
II. Information in the Statistical Business Process	5
<i>A. Identifying and Evaluating Statistical Needs</i>	<i>5</i>
<i>B. Designing and Managing Statistical Programs.....</i>	<i>6</i>
<i>C. Designing Processes.....</i>	<i>7</i>
<i>D. Running Processes</i>	<i>12</i>
<i>E. Exchanging Information</i>	<i>13</i>
<i>F. Collecting Information.....</i>	<i>15</i>
<i>G. Processing and Analyzing Information</i>	<i>15</i>
<i>H. Disseminating Information.....</i>	<i>17</i>
III. Foundational Information	19
<i>A. Concepts</i>	<i>19</i>
<i>B. Population.....</i>	<i>20</i>
<i>C. Node and Node Set.....</i>	<i>21</i>
<i>D. Statistical Classification</i>	<i>23</i>
<i>E. Variable</i>	<i>25</i>
<i>F. Represented Variable.....</i>	<i>25</i>
<i>G. Instance Variable.....</i>	<i>27</i>
<i>H. Information Resources.....</i>	<i>27</i>
<i>I. Data Sets</i>	<i>28</i>
<i>J. Dimensional and Unit Data Structures.....</i>	<i>30</i>
<i>K. Referential Metadata Sets.....</i>	<i>31</i>
IV. Technical information	33
<i>A. Identity and Administrative Details</i>	<i>33</i>
<i>B. Information Providers, Information Consumers, Organizations, and Individuals</i>	<i>34</i>
Annex A: Exchange Channels	35
<i>A. Administrative Registers</i>	<i>35</i>
<i>B. Web Scraping for Data Collection.....</i>	<i>36</i>
<i>C. Survey Data Collection.....</i>	<i>38</i>
Annex B: Glossary.....	41
Annex C: UML Diagrams.....	70
Annex D: Extending the model	240
<i>A. GSIM Extension Methodology</i>	<i>240</i>
<i>B. Administrative Attributes</i>	<i>243</i>

Figure 1. Identify and Evaluate Statistical Needs 5

Figure 2. Design and Manage Statistical Program 6

Figure 3. Process Steps can be as large or small as needed 8

Figure 4. Design Processes..... 9

Figure 5. Use of re-usable Business Services..... 11

Figure 6. Run Process..... 12

Figure 7. Exchange Channels 13

Figure 8. Exchange Channels for collecting information..... 15

Figure 9. Exchange Channel for disseminating information..... 17

Figure 10. Concepts..... 19

Figure 11. Populations and Units 20

Figure 12. Node and Node Set inheritance..... 21

Figure 13. Statistical Classifications 23

Figure 14. Variable..... 25

Figure 15. Represented Variable 26

Figure 16. Instance Variable 27

Figure 17. Information Resources 27

Figure 18. Data Sets 28

Figure 19. Data Structures 30

Figure 20. Referential Metadata Sets 31

Figure 21. Identifiable Artefacts 33

Figure 22. Agents 34

Figure 23. Administrative Register 35

Figure 24. Web Scraping Channel 37

Figure 24. Questionnaire 38

I. Introduction

1. The GSIM Specification is the most detailed level of the Generic Statistical Information Model (GSIM). It provides a set of standardized, consistently described information objects, which are the inputs and outputs in the design and production of statistics. Each information object is defined and its attributes and relationships are specified. GSIM is the result of a collaboration involving statistical organizations across the world in order to develop and maintain a generic reference model suitable for all organizations and meet the strategic goals (in particular the modernization effort) of the official statistics community. For contextual information, an introduction to GSIM and information on using GSIM, please refer to the GSIM Brochures, Communication and Implementing GSIM documents.

2. There is a widespread interest across statistical organizations in being able to trace how statistical information (for example, data and metadata) "flow" through statistical business processes (into and out of processes). Interested parties include broad statistical systems (like the European Statistical System), National Statistical Systems (both centralized and decentralized) and smaller task teams working inside National Statistical Offices.

3. GSIM covers the whole statistical process and is designed to support both current and new ways of producing statistics. Section II describes tasks (for example identifying statistical needs, managing statistical programs, dissemination) which statistical organizations undertake and how the model describes the information flows in those tasks. This section also contains descriptions of designing and running processes to show how GSIM has models the explicit separation between the design and execution of statistical processes.

4. There is an increasing business need to record reliable, structured information about the processes used to produce specific statistical outputs. In order to maximize transparency and reproducibility of results, it is important for a statistical organization to understand the processes it undertakes and their inputs and outputs. Section III describes the foundational information objects (that is, the conceptual and structural metadata objects) that are used as inputs and outputs in a statistical business process.

5. There are a number of technical information objects in GSIM. These objects are the fundamental building blocks that support many of the other objects and relationships in the model. They provide features which are reusable by other objects to support functionality such as identity, versioning etc. These objects are described in Section IV of this document.

6. This document provides a description of GSIM in the context of a statistical organization. It has a number of annexes which provide further details for the reader. These annexes are:

- Annex A: Exchange Channels - This annex provides further information about the three subtypes of Exchange Channel focused on data collection.
- Annex B: Glossary - The annex gives readers definitions and explanatory descriptions for the GSIM information objects.
- Annex C: UML diagrams - This annex includes all detailed UML models of GSIM.

- Annex D: Extending the model - This annex provides information for implementers on how to extend GSIM for organization specific purposes. It also contains the set of recommended attributes for the administration of GSIM objects.

7. Note: GSIM information objects have been given in italics in the descriptions that follow. The diagrams included in this section are stylized representations of the model. **The colours of the boxes in diagrams represent which group the information object belongs to (Blue for Business Group, Red for Exchange Group, Green for Concepts Group, Yellow for Structures Group and Orange for the Base Group).** In many cases there is more detail to be found in the UML. Detailed information on each information object in the model, including a glossary and UML class diagrams can be found in Annexes B and C of this document.

II. Information in the Statistical Business Process

8. This section looks at different ways that information objects are used within the statistical business process. It considers eight different scenarios, identifying the information objects used and the relationships between those objects.

A. Identifying and Evaluating Statistical Needs

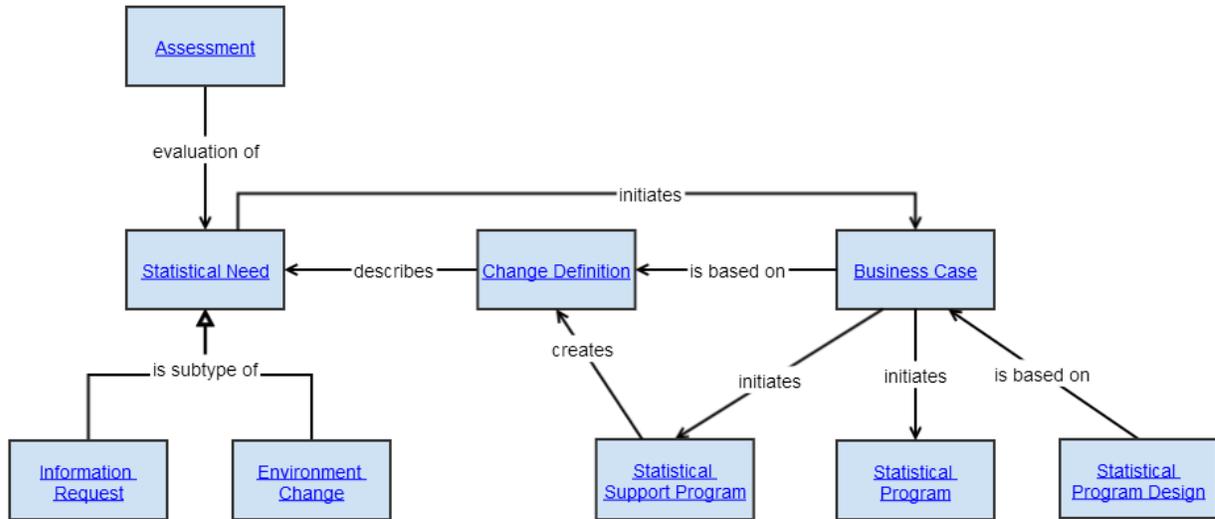


Figure 1. Identify and Evaluate Statistical Needs

9. An organization will react and change due to a variety of needs. A *Statistical Need* presents itself to the statistical organization in the form of an *Environment Change* or an *Information Request*.

10. *Environment Change* indicates that there needs to be an externally motivated change. This may be specific to the organization in the form of reduced budget or new demands from stakeholders, or may be a broader change such as the availability of new methodology or technology.

11. When an organization receives an *Information Request* this will identify the information that a person or organization in the user community requires for a particular purpose. This community may include users within the organization as well as external to it. For example, a the team responsible for compiling National Accounts may need a new *Business Process* to be initiated to produce new inputs to their compilation process. This request will commonly be defined in terms of a *Subject Field* that defines what the user wants to measure. When an *Information Request* is received it will be discussed and clarified with the user. Once clarified, a search will be done to check if the data already exist. Discovering these *Data Sets* may be enabled by searching for *Concepts* and *Classifications*. Each of these activities are described by a *Process Step*.

12. The *Statistical Need* - whether an *Information Request* or *Environmental Change* - will be formalized into a *Change Definition*, typically created by a *Statistical Support Program* (a "statistical

change program"). The *Change Definition* identifies the specific nature of the change in terms of its impacts on the organization or specific *Statistical Programs* or *Statistical Support Programs*. This *Change Definition* is used as an input into a *Business Case*. A successful outcome will either initiate a new *Statistical Program* or a new *Statistical Support Program* that will create a new *Statistical Program Design* that redefines the way an existing *Statistical Program* is carried out.

13. A *Statistical Need* can also be internally driven. At any point in the statistical business process, an organization may undertake an evaluation to determine utility or effectiveness of the business process or its inputs and outputs. An *Assessment* will be undertaken to evaluate any resources, processes or outputs and may refer to any object described in the model. *Assessments* include gap analyses undertaken in the context of *Business Cases* and evaluations undertaken to determine whether a statistical output meets the need for which it was first created.

B: Designing and Managing Statistical Programs

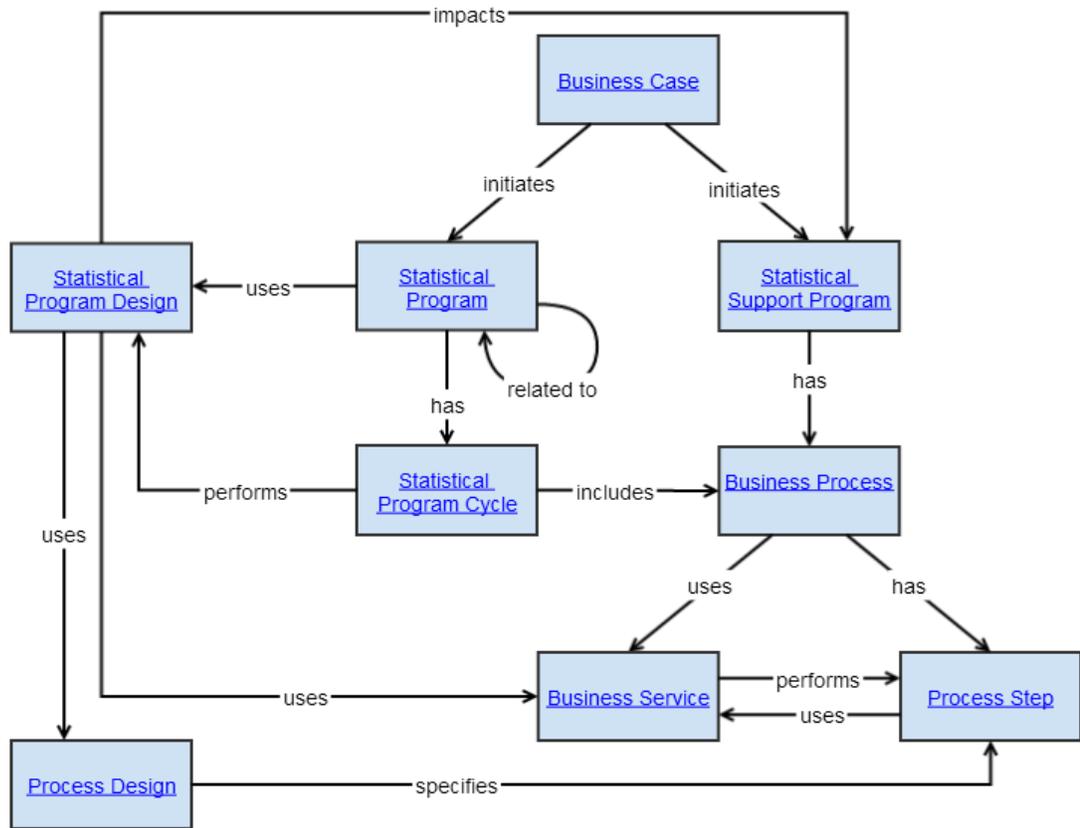


Figure 2. Design and Manage Statistical Program

14. A statistical organization will respond to a perceived *Statistical Need* by creating a *Business Case*. Responding to the *Business Case* will involve one of three things: the creation of a new *Statistical Support Program*, the creation of a new *Statistical Program*, or the evolution of an existing *Statistical Program Design* to be implemented by an existing *Statistical Program*.

15. *Statistical Support Programs* undertake the activities of the statistical organization such as statistical change programs, data management programs, metadata management programs, methodological research programs, etc. A good example is a program which manages classifications.

16. *Statistical Programs* are those programs that an organization undertakes to produce statistics (for example, a retail trade survey). *Statistical Programs* are cyclical - they perform cycles of collection, production and dissemination of products. Each such cycle is represented by a *Statistical Program Cycle* object. The *Statistical Program Cycle* is a repeating activity to produce statistics at a particular point in time (for example, the retail trade survey for March 2012).

17. *Statistical Programs* require *Statistical Program Designs* to achieve their objectives. These designs cover the design of all activities to be undertaken, notably at the level of *Business Processes*. Within a *Statistical Program Cycle*, several *Business Processes* would typically be performed. These can be understood to correspond to the processes and sub-processes found in the Generic Statistical Business Process Model (GSBPM). These *Business Processes* may be repeated within a cycle. Each iteration can be made up of multiple activities of the same or different types. As an example of this, within a single cycle, the *Statistical Program* might perform three iterations of data collection and processing, then analyze the data and disseminate the resulting statistical *Products*. Each of these activities could be understood to be a separate *Business Process*.

18. The *Statistical Program Design* specifies the way in which *Business Processes* will be conducted. This includes the use of re-usable *Business Services* (possibly sourced from outside the statistical organization), or through the design and use of more traditional processes. In the latter case, *Process Design* objects would be used to specify *Process Steps*. (Although re-usable *Business Services* are also specified by *Process Designs* and *Process Steps*, these will already exist, and not need new design work as part of the *Statistical Program Design*.)

19. It should be noted that *Statistical Program Designs* specify what *Process Steps* will need *Process Designs*, and also which *Business Services* would be used, but do not do the low-level specification of how such *Process Steps* and *Business Services* are executed. These specifications are found in the *Process Design* object.

C: Designing Process Steps

20. Before explaining the objects which GSIM uses to represent the design of *Process Steps*, it is important to discuss the nature of processes more generally. The types of objects provided by GSIM perform specific functions. In GSIM, *Business Processes* have *Process Steps*. Each *Process Step* can be as "large scale" or "small scale" as the designer of a particular *Business Process* chooses (see Figure below).

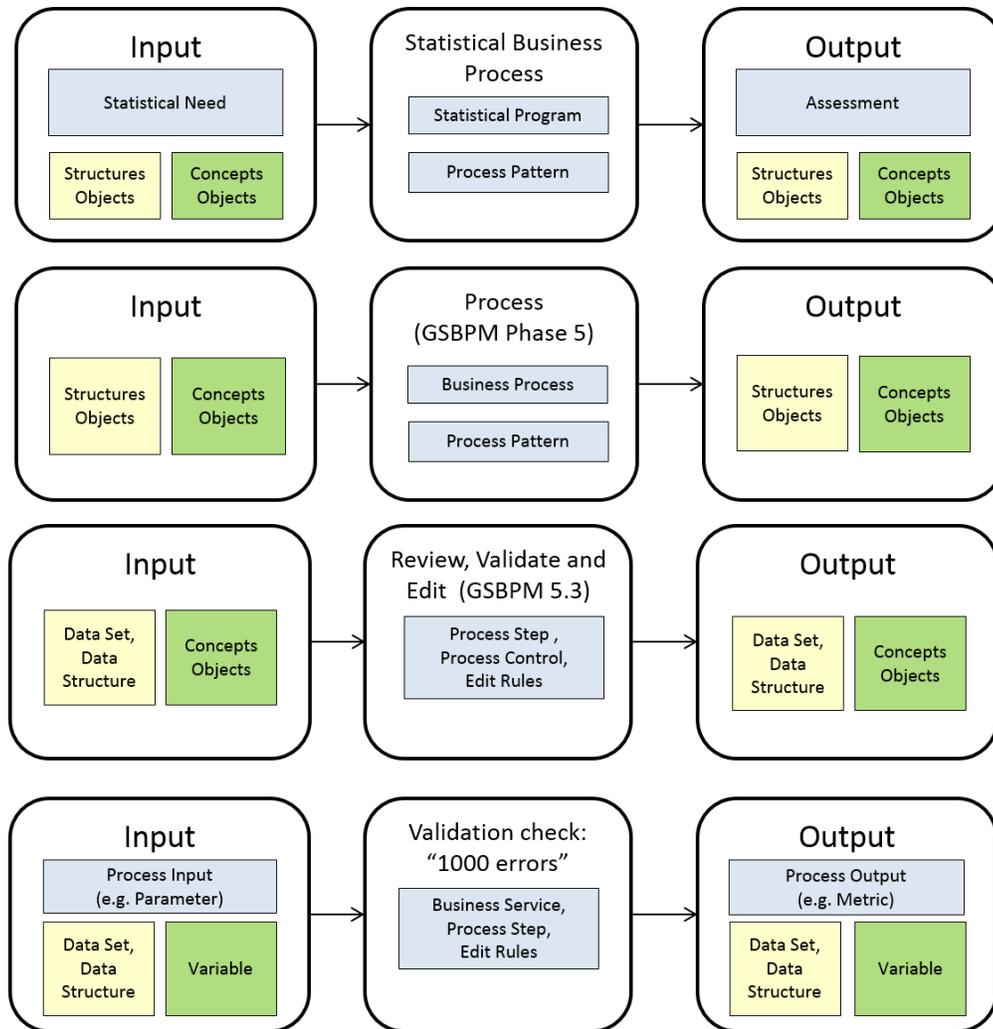


Figure 3. Process Steps can be as large or small as needed

21. *Process Steps* can contain "sub-steps", those "sub-steps" can contain further "sub-steps" within them and so on indefinitely. Typically, the outputs of one *Process Step* become inputs to the next *Process Step*. There can also be conditional flow logic applied to the sequence of *Process Steps*, based on parameters which have been passed in, or conditions met by the outputs of a previous *Process Step*.

22. The design of a *Process Step* thus can be understood to use other *Process Steps* and even other *Business Services* which have already been designed and made available for re-use. In a more traditional scenario, the *Process Step* is designed and then executed. In future, it is foreseen that re-usable *Business Services* will be increasingly common, having been designed and implemented by another external organization. The next sections describe these two scenarios.

i. Designing *Process Steps*

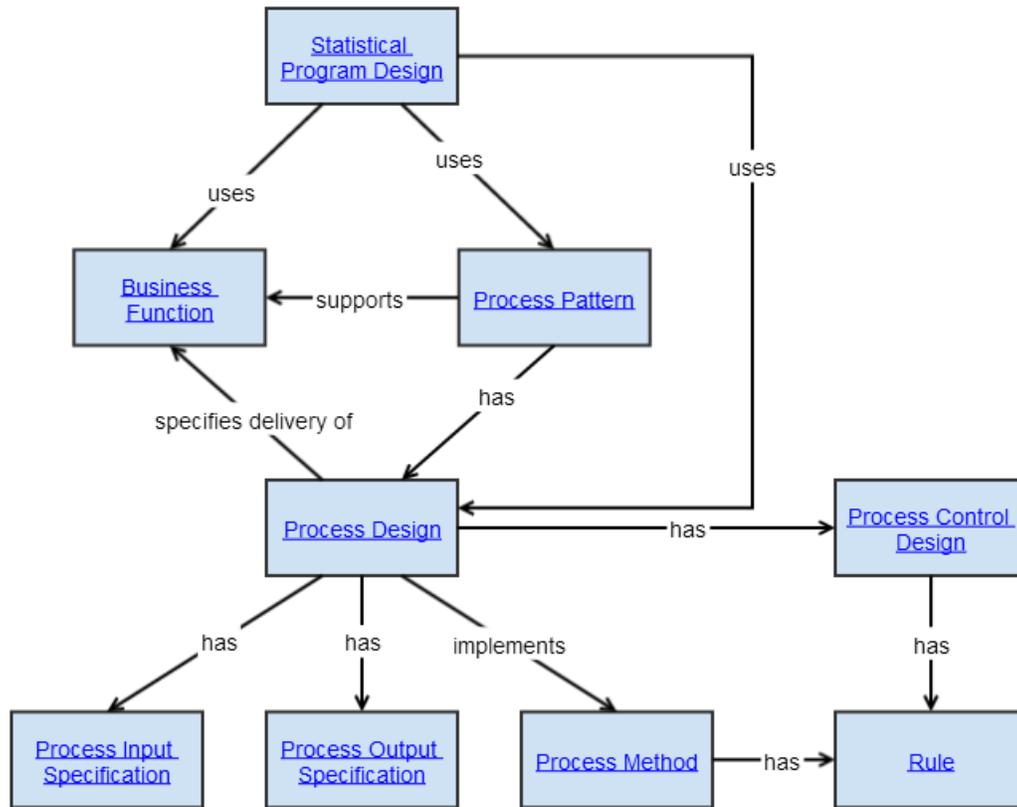


Figure 4. Design Process Steps

23. A *Statistical Program Design* is associated with a top level *Process Step* whose *Process Design* contains all the sub-steps and process flows required to put that *Statistical Program* into effect. Each *Process Step* in a statistical *Business Process* has been included to serve some purpose. This is captured as the *Business Function* associated with the *Process Step*. An example of a *Business Function* could be "impute missing values in the data". In order to support this *Business Function*, an imputation process is needed, which will require a *Process Design*.

24. In line with the GSIM design principle of separating design and production, GSIM assumes that *Process Steps* will be designed during a design phase. Having divided a planned statistical *Business Process* into *Process Steps*, the next requirement is to specify a *Process Design* for each step. The *Process Design* identifies how each *Process Step* will be performed. A *Process Design* may use a *Process Pattern* which is a nominated set of *Process Designs* and associated flows (*Process Control Designs*) which have been highlighted for reuse.

25. *Process Designs* specify several things: they identify the different types of inputs and outputs represented by the *Process Input Specification* and *Process Output Specification*. Examples of *Process Inputs* include data, metadata such as *Statistical Classifications*, imputation and editing *Rules*, parameters, etc. *Process Outputs* can be reports of various types (processing metrics, reports about

data validation and quality, etc.), edited *Data Sets*, new *Data Sets*, new or revised instances of metadata, etc.

26. To continue the example, the process designer would specify the inputs in the *Process Input Specification* as imputation *Rules* and the *Data Set* for which imputation is desired. The *Process Output Specification* would include an edited *Data Set* containing the imputed values, plus a report detailing which values had been imputed.

27. The *Process Design* specifies the control logic, that is the sequencing and conditional flow logic among different sub-processes (*Process Steps*). This flow is described in the *Process Control Design*. When creating a *Process Design*, a *Process Control Design* that provides information on "what should happen next" is specified. Sometimes one *Process Step* will be followed by the same step under all circumstances. In such cases the *Process Control Design* simply records what *Process Step* comes next. However, sometimes there will be a choice of which *Process Step* will be executed next. In this case, the *Process Control Design* will detail the set of possible "next steps" and the criteria to be applied in order to identify which *Process Step(s)* should be performed next.

28. The *Process Design* associated with that *Process Step* will identify the *Process Method* that will be used to perform the *Business Function* associated with the *Process Step*. For example, if the *Business Function* is 'impute missing values in the data', the *Process Method* might be 'nearest neighbour imputation'.

29. A *Process Method* specifies the method to be used, and is associated with a set of *Rules* to be applied. For example, any use of the *Process Method* 'nearest neighbour imputation' will be associated with a (parameterized) *Rule* for determining the 'nearest neighbour'. In that example the *Rule* will be mathematical (for example, based on a formula). *Rules* can also be logical (for example, if Condition 1 is 'false' and Condition 2 is 'false' then set the 'requires imputation' flag to 'true', else set the 'requires imputation flag' to 'false').

30. The resulting *Process Design* and *Process Control Design* objects (along with related *Process Input Specifications* and *Process Output Specifications*) would be used in the implementation of the *Process Step*.

ii. Using Re-Usable *Business Services*

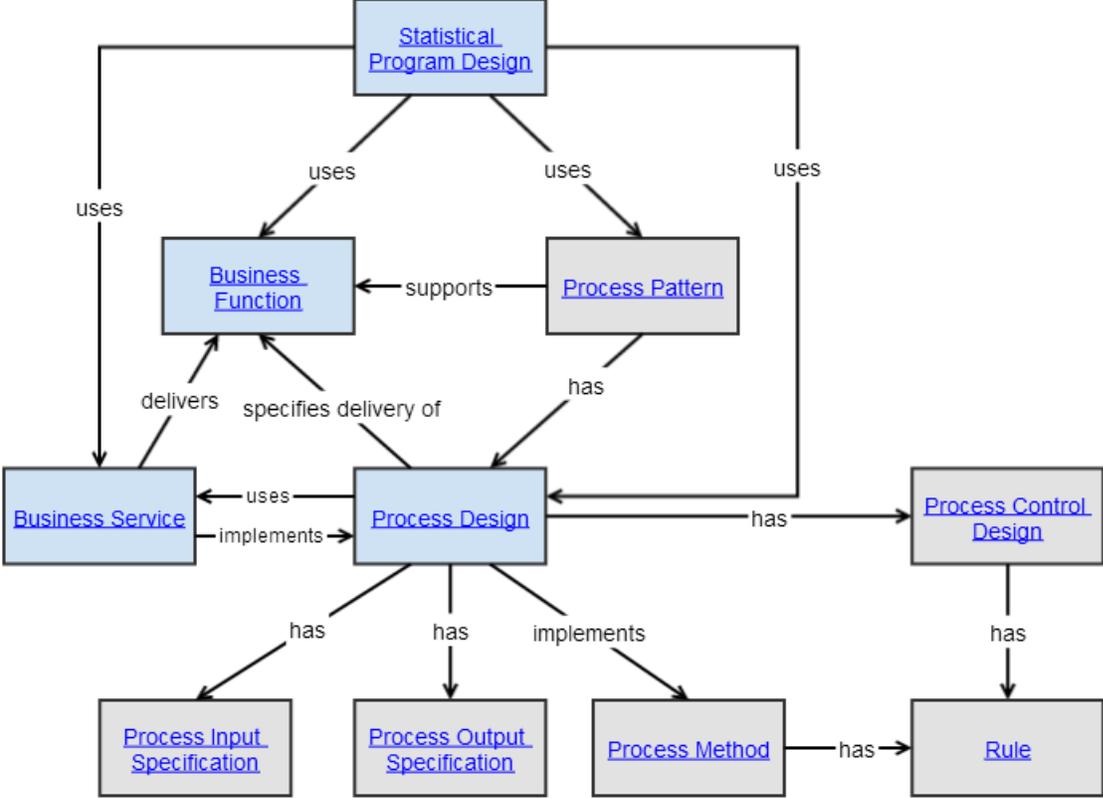


Figure 5. Use of re-usable *Business Services*

31. It is not always necessary for the *Statistical Program* to design its own *Process Steps* from the beginning. The Common Statistical Production Architecture (CSPA) describes how statistical organizations can create statistical services that are easily reused in other statistical organizations. In GSIM terms, a statistical service is a *Business Service*. A *Business Service* is a means of performing a *Business Function* (an ability that an organization possesses, typically expressed in general and high level terms and requiring a combination of organization, people, processes and technology to achieve).

32. The increased sharing and reuse of *Business Services* means that the resources needed to meet new demands for statistical production could be considerably reduced, and the time needed to produce new statistical products could be lessened. To facilitate this, CSPA introduced the concept of a statistical services catalogue, where different statistical organizations could list the statistical services they have developed, with the intent of sharing them with other statistical organizations.

33. *Business Services* have already been designed, with all of the normal input types, output types, process control design, and other properties already specified. Thus, a *Business Service* can act in a fashion similar to a *Process Step* designed within the organization, but without the effort required in the traditional scenario.

D: Running Processes

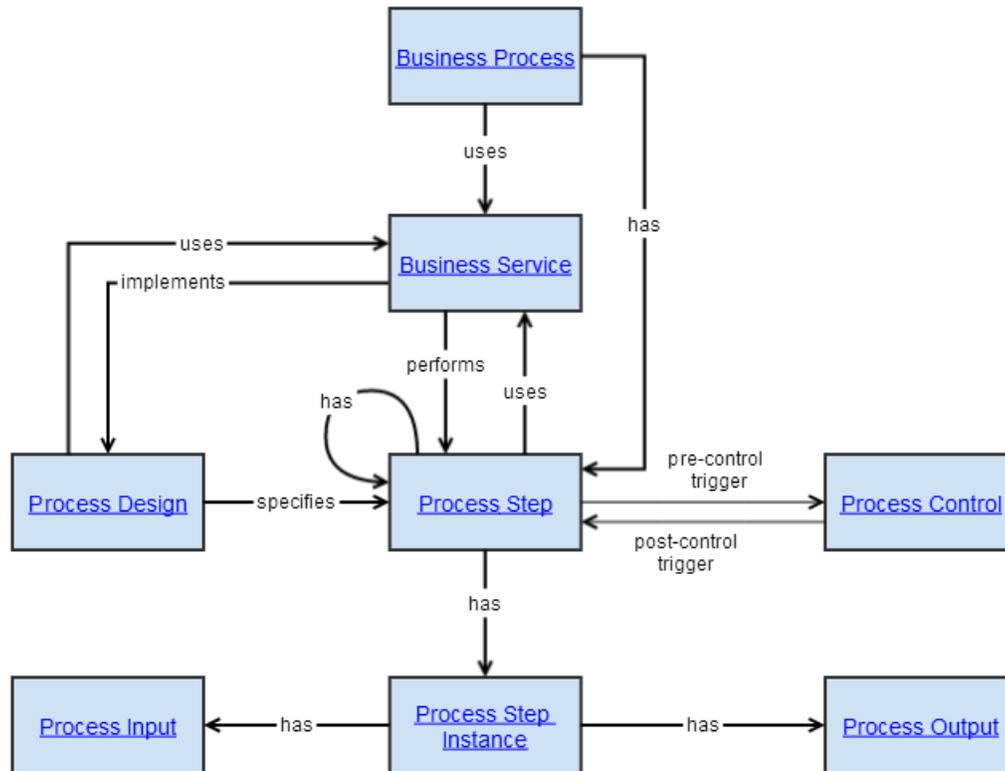


Figure 6. Run Process

34. A *Statistical Program* needs to execute processes to realize some *Business Functions*. This can be done in two ways: a *Process Step* can be directly executed by a *Business Process*, or a re-usable *Business Service* can be used by the *Business Process*, as an intermediate trigger for the execution of the *Process Step*.

35. In order to understand how this works, we characterize the nature of *Process Steps* in more detail. *Process Steps* are the resources which have been specified in a *Process Design*, and which can be executed multiple times. *Process Steps* can exist at many levels of granularity, and can involve the use of other *Process Steps* as sub-processes. The navigation among the sub-processes is performed during execution as indicated by a *Process Control*, which is itself an implementation of a *Process Control Design*.

36. Individual executions of a *Process Step* are represented by the *Process Step Instance*. It is at this level that specific instances of the inputs and outputs are used. In the *Process Design*, the types of inputs and outputs are specified (*Process Input Specification* and *Process Output Specification*) - the actual instances of inputs and outputs are associated with the *Process Step Instance*, and are represented by the *Process Input* and *Process Output* objects. Inputs can be of any type of information - rules, parameters, data sets, metadata of many kinds, etc. Outputs are similarly of many different types, and often include process metrics and various types of reports, as well as data and metadata.

37. At the time the *Process Design* is executed someone or something needs to apply the designated method and rules. The *Process Design* can designate the *Business Service* that will implement the *Process Method* at the time of execution. A *Business Service* represents a service delivered by a piece of software (as described in the section above) or a person. Putting a publication on the statistical institute's website or putting collected response forms in a shared data source for further processing are both examples of *Business Services*.

38. It should be noted that this model supports both automated and manual processes, and processes which might involve sub-processes of either type.

E: Exchanging Information

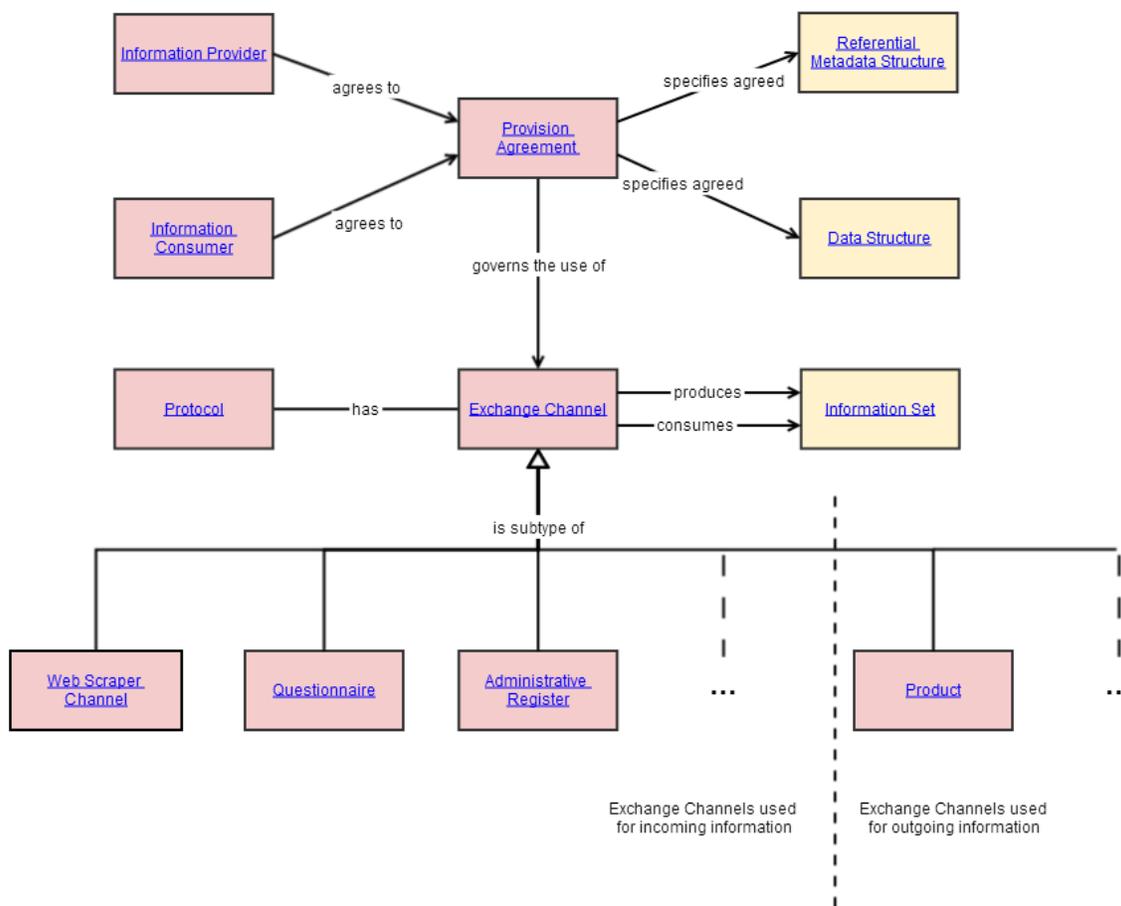


Figure 7. Exchange Channels

39. Statistics organizations collect data and referential metadata from *Information Providers*, such as survey respondents and providers of *Administrative Registers*, and disseminate data to *Information Consumers*, such as government agencies, businesses and members of the public. Each of these exchanges of data and referential metadata uses an *Exchange Channel*, which describes the means to receive (data collection) or send (dissemination) information. *Information Providers* and *Information*

Consumers can be *Organizations* or *Individuals* who are either within or external to the statistical organization.

40. Different *Exchange Channels* are used for collection and dissemination. Examples of collection *Exchange Channels* include *Questionnaire*, *Web Scraper Channel* and *Administrative Register*. The only example of a dissemination *Exchange Channel* currently contained in GSIM is *Product*. Additional *Exchange Channels* can be added by organizations depending on their needs.

41. The use of an *Exchange Channel* is governed by a *Provision Agreement* between the statistics office and the *Information Provider* (collection) or the *Information Consumer* (dissemination). The *Provision Agreement*, which may be explicitly or implicitly agreed, provides the legal or other basis by which the two parties agree to exchange data. The parties also use the *Provision Agreement* to agree the *Data Structure* and *Referential Metadata Structure* of the information to be exchanged.

42. The mechanism for exchanging information through an *Exchange Channel* is specified by a *Protocol* (e.g. SDMX web service, data file exchange, face to face interview).

43. To collect data, a statistical organization receives data and referential metadata from the *Information Provider* in a manner consistent with the *Protocol* and the *Provision Agreement*, and the *Exchange Channel* produces an *Information Set*. To disseminate data, the *Exchange Channel* consumes an *Information Set*, which is then provided to the *Information Consumer* in a manner consistent with the *Protocol* and the *Provision Agreement*. More information about collection and dissemination can be found in the following sections.

F: Collecting Information

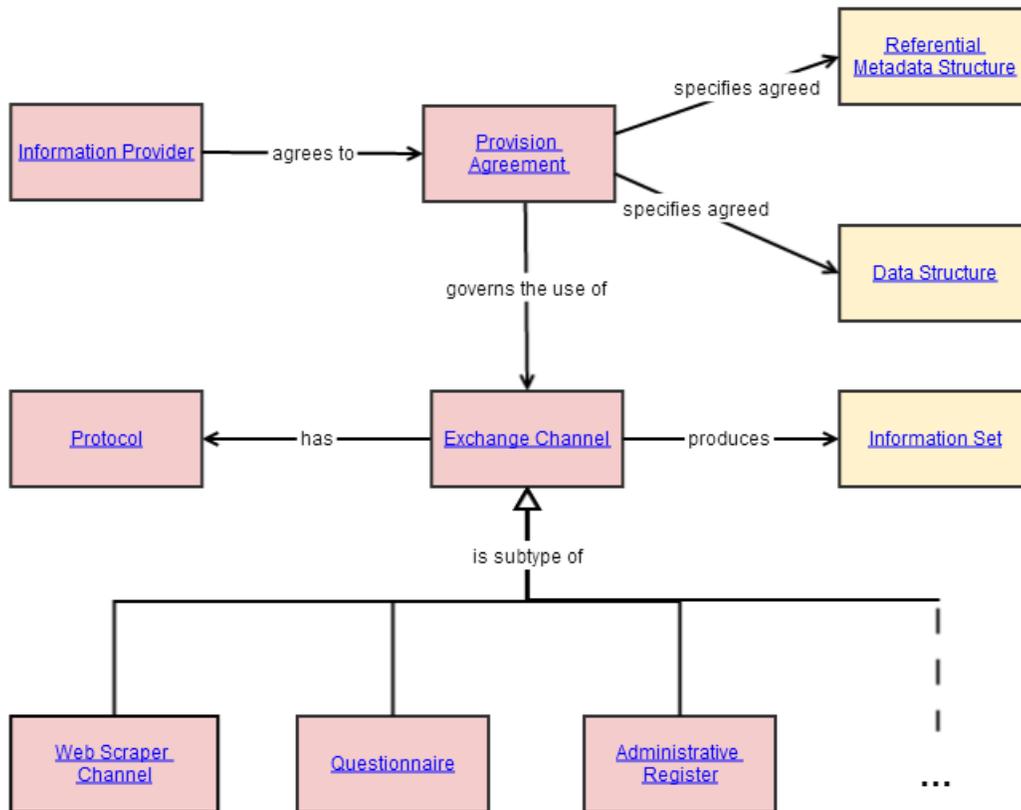


Figure 8. Exchange Channels for collecting information

44. GSIM models three collection *Exchange Channel* examples: *Questionnaire*, *Web Scraper Channel* and *Administrative Register*. Each of these is detailed in Annex A. Statistics organizations may collect data and referential metadata from *Information Providers* using additional *Exchange Channels*, such as file transfer, web services and data scanning. Statistical organizations can extend GSIM to add channels relevant to their context.

45. The use of an *Exchange Channel* for collection is governed by a *Provision Agreement* between the statistical organization and the *Information Provider*. The two parties use the *Provision Agreement* to agree the *Data Structure* and *Referential Metadata Structure* of the data to be exchanged. The mechanism for collecting information through the *Exchange Channel* is specified by a *Protocol* (e.g. face to face interview, data file exchange, web robot). The collecting organization uses the collected information to produce an *Information Set*, which may contain data or referential metadata.

G: Processing and Analyzing Information

46. GSIM is very flexible in describing the processing and analysis of information.

47. One can understand the statistical production process from a data-centric perspective¹. Statistical organizations strive to produce high-quality accurate data that is supported by the metadata needed to make the data optimally useful. For this reason, it is appropriate to think of the evolution of data as it passes through the production process. The focus of many activities is driven by the metadata, but at the end of the production process the metadata is a supporting resource from the perspective of the data and ultimately a statistical product. The relationship of the data and metadata is one which is important to understand.

48. Collected data comes into a statistical organization through an *Exchange Channel*. Regardless of how the data is collected and where it comes from, it is a resource which will begin a process of evolution through many different stages. The initial data is described as a *Data Set* with relevant *Data Structures*. *Data Sets* are stored in an organised way in a *Data Resource*. The *Data Sets* are the primary inputs and outputs of a set of *Process Steps*, as conducted by a *Statistical Program*.

49. As the statistical organization moves from raw input data to an increasingly refined set of data, it can be understood that each phase of this processing adds additional *Datasets* to the *Data Resource*. There are many different *Process Methods* which may inform these activities. These are implemented through the different *Process Steps* that the statistical organization undertakes.

50. At a certain point (and this can take place at different places within the production process, depending on the type of edits being performed) the data will be analysed for the production of statistical *Products*. The analysis of the data can be understood as using *Data Sets* from the *Data Resource* as inputs to processes such as confidentiality routines or to produce explanations of the data. The operations performed during analysis will vary based on what the ultimate *Products* are - confidentialised unit-record data may be a *Product*, or we may be publishing aggregated indicators and tables to address specific policy issues, and these involve different types of analysis - but the process is still one of further evolving the information held in the *Data Resource*.

51. In the past, there was an assumption that a data collection will be followed by processing, analysis, and dissemination of the statistical *Products*. This is a time-consuming and resource-intensive process. One way to make the functions of a statistical organization more efficient is to re-use data to produce new *Products* as they are asked for, lowering the cost and shortening the time needed for production. In this sense, the *Data Resource* can be understood as an organizational asset, to be managed and exploited to the greatest extent possible.

¹ Although these paragraphs focus of data, the same descriptions can be applied to referential metadata.

H: Disseminating Information

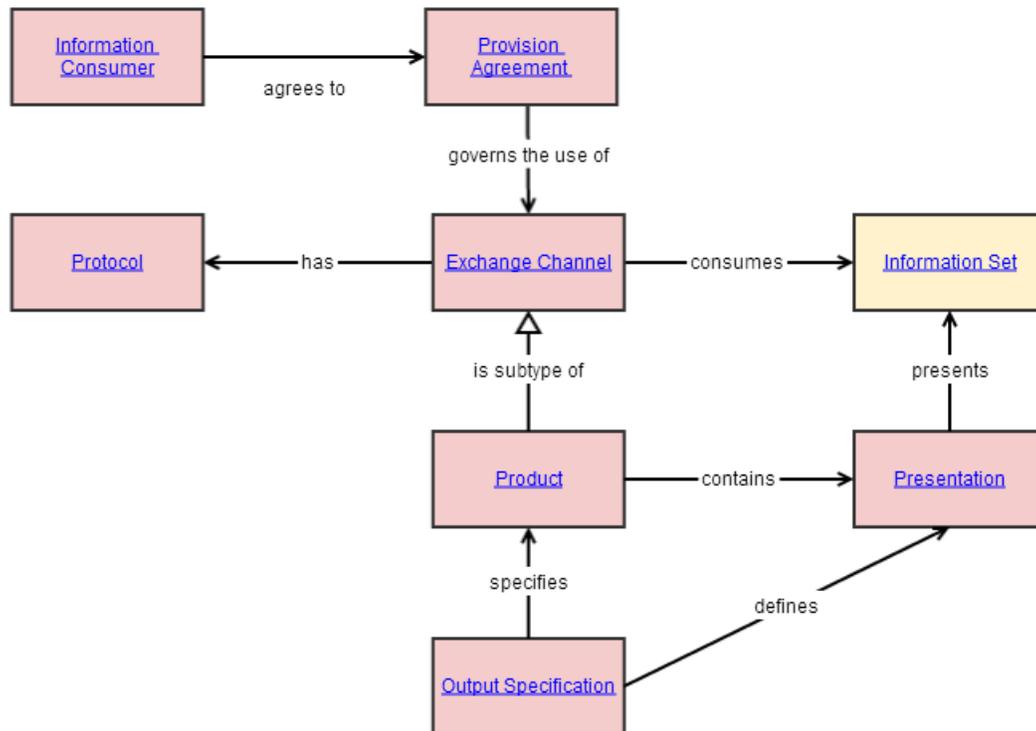


Figure 9. Exchange Channel for disseminating information

52. A statistical organization disseminates statistical information to an *Information Consumer*.

53. The *Information Consumer* accesses a set of information via a *Product* (or potentially via another *Exchange Channel*), which contains one or more *Presentations*. Each *Presentation* will typically provide a view of data and associated metadata to define and describe the structure of the presented data, and perhaps referential metadata in the form of textual media, such as quality reports.

54. A *Presentation* can take different forms - for example, it could be a screen visualization of a table of data in graphical form displayed in an HTML page, a downloadable PDF, or an SDMX file in XML format.

55. An *Output Specification* defines what is contained in the *Presentation*. A *Product*, which packages *Presentations*, may be a statistical organization's standard specific output as one might see in:

- a regular statistical bulletin (e.g. a monthly publication of the Retail Prices Index),
- a dynamically generated package of statistical content which is generated following the receipt of a query from an *Information Consumer* who wishes to access the organization's data via a published API (Application Programming Interface) or
- some data exploration facility which might be built into the statistical organization's website.

56. The *Output Specification* also defines the information required from the *Information Set* for the *Presentation*. The specifications are frequently determined by an internal (to the organization) process which would have specifically standard, static outputs to produce (such as the aforementioned statistical bulletins). For dynamically delivered products, aspects of the specification could be determined by the *Information Consumer* at run time, via machine to machine dynamic, as exemplified in the API scenario above. In either case, the requests would result in the *Output Specification* specifying *Information Set* data and/or referential metadata that will be included in each *Presentation*.

57. The mechanism for providing a *Product* is specified by a *Protocol* (e.g. SDMX-ML, DDI XML, PDF etc.). This formatting information forms part of the *Output Specification* to generate the *Product* and its *Presentations* in the appropriate format.

58. The *Information Consumer* can be one of many forms depending upon the scenario of the request. The *Information Consumer* could be a person accessing the statistical organization's website and visually inspecting the contents of a web page, or it could be a computer program requesting the information via an API using an SDMX query. The *Information Consumer's* access to the information would be subject to a *Provision Agreement*, which would set out the conditions of access and use. This might be in the form of passive acceptance of the terms and conditions of use of the data from a website the *Information Consumer* is accessing, or in the case of access to a greater level of detail via an API, it might be a more involved registration process.

III. Foundational Information

59. The GSIM Concepts and Structures groups include information objects which are foundational to the statistical *Business Process*. That is, these objects are the conceptual and structural objects which are used as the *Process Inputs* and *Process Outputs* to the process. The Concepts area of GSIM includes the sets of information objects that describe and define the terms used when talking about the real-world phenomena that the statistics measure in their practical implementation. The Structures area includes the set of information objects used in relation to data and referential metadata and their structures. The objects described in this section of the document are used to provide information that helps users of data and metadata understand the results of *Business Processes* and *Statistical Programs*.

A: Concepts

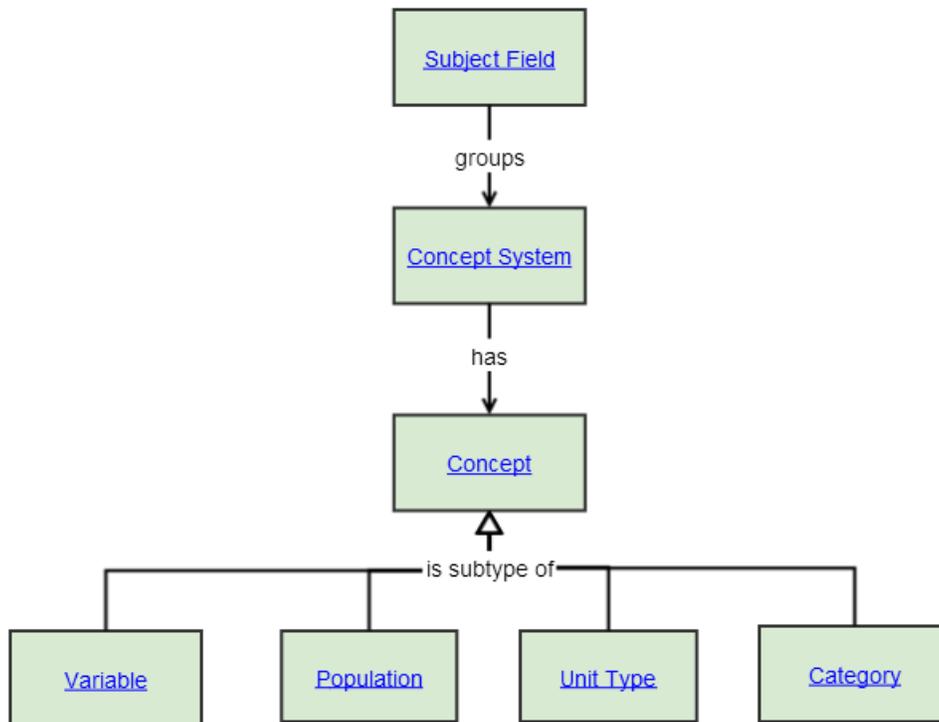


Figure 10. Concepts

60. At an abstract level, a *Concept* is defined in GSIM as a 'unit of thought differentiated by characteristics'. *Concepts* are used in different ways throughout the statistical lifecycle, and each different role of a *Concept* is described using a different information object (which are subtypes of *Concept*). A *Concept* can be used in these situations:

- (a) As a characteristic. The *Concept* is used by a *Variable* to describe the particular characteristic that is to be measured about a *Population*. For example, to measure the *Concept* of gender in a population of adults in the Netherlands, the *Variable* combines this *Concept* with the *Unit Type* person.
- (b) As a *Unit Type* or a *Population*. To describe the set of objects that information is to be obtained about in a statistical survey. For example, the *Population* of adults in Netherlands, based on the *Unit Type* of persons.
- (c) As a *Category* to further define details about a *Concept*. For example, Male and Female for the *Concept* of Gender. *Codes* can be linked to a *Category* via a *Node* (i.e., a *Code Item* or *Classification Item*), for use within a *Code List* or *Statistical Classification*.

61. *Concept Systems* are sets of *Concepts* which are structured by the relations between those *Concepts*. A *Subject Field* groups *Concept Systems* on the basis of their field of special knowledge (for example, labour market, tourism).

B. Population

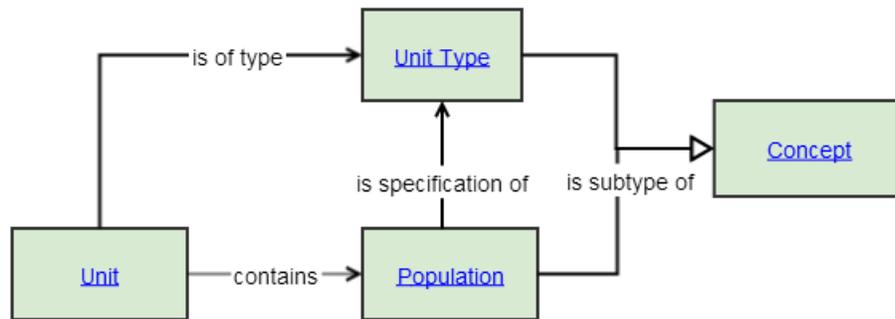


Figure 11. Populations and Units

62. There are several kinds of *Populations* depending on what *Process Step* it is used in. For example a statistical organization may refer to a target, survey, frame, or analysis population. The objects of interest in a statistical process are *Units* (for example, a particular person or a business). Data are collected about *Units*. There are two ways in which a unit is specified in the model. A *Unit* is an individual entity associated with a *Population* about which information may be obtained. A *Unit Type* (for example persons or businesses) is a way of identifying an abstract type of *Unit* that a *Variable* is measuring.

C. Node and Node Set

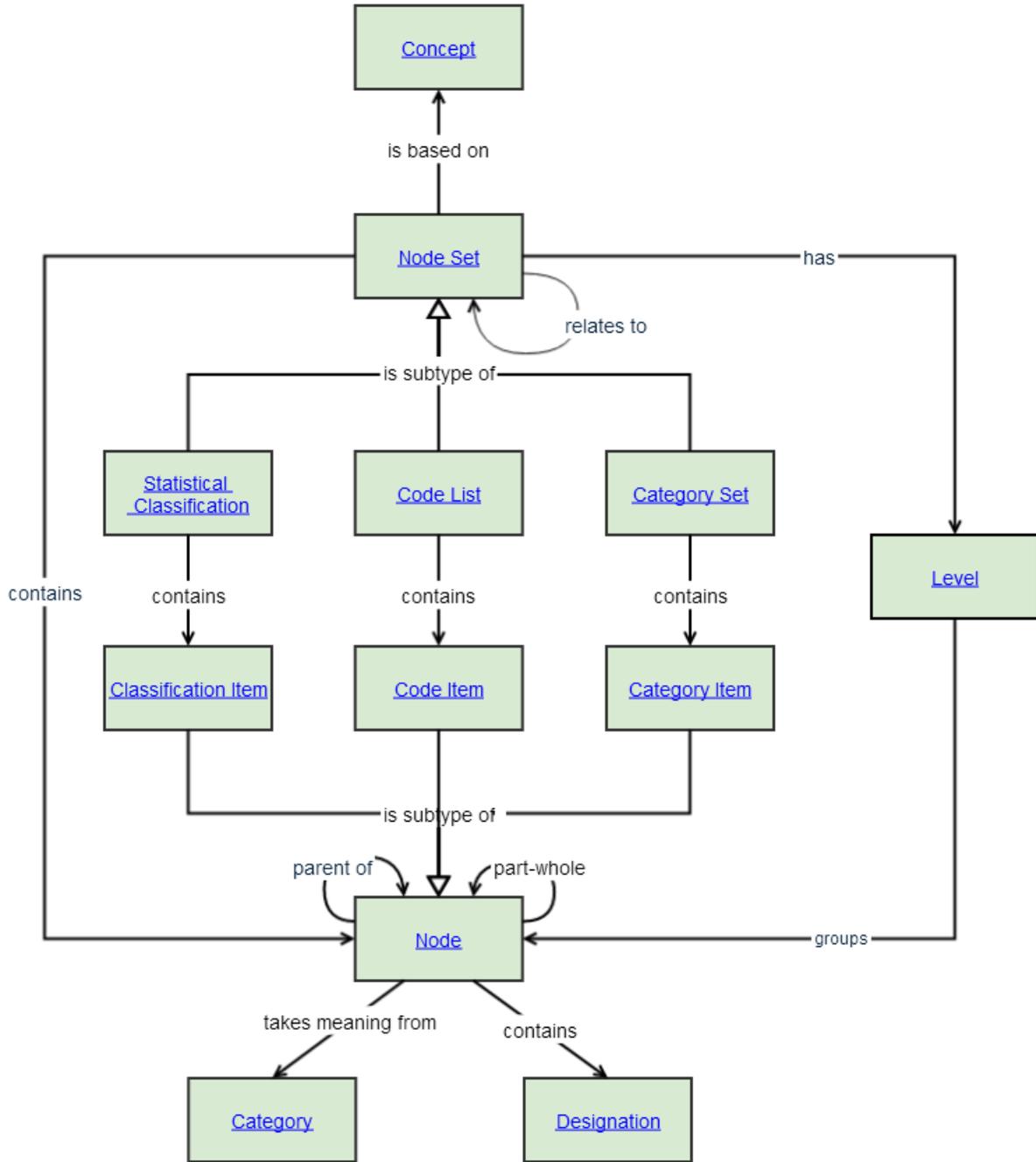


Figure 12. Node and Node Set inheritance

63. A *Category* is a particular type of *Concept* whose role is to define a characteristic. There are three ways in which a *Category* can be used. In GSIM, these are described as the three subtypes of *Node* - *Category Item*, *Code Item* and *Classification Item*. *Categories* are grouped into *Node Sets*

based on the way in which it can be used. There are three subtypes of these groups (*Node Sets*) - *Category Sets*, *Code Lists* and *Statistical Classifications*.

64. A *Category Set* is a set of *Category Items*, which contain the meaning of a *Category* without any associated representations. An example of a *Category Set* is: Male, Female.

65. In a *Code List*, the *Code Items* contain the meaning of the *Categories* combined with a *Code* representation. An example of a *Code List* is: 1. Male, 2. Female.

66. A *Statistical Classification* is similar to a *Code List*. It combines the meaning of the *Category* with a *Code* representation. However the content of a *Statistical Classification* must fulfil certain criteria and have a certain status. The *Classification Items* must be mutually exclusive and jointly exhaustive for the *Level* at which they exist at in the *Statistical Classification*. An example of a *Statistical Classification* is: 1. Male, 2. Female, 3. Intersex.

67. A *Code List* does not have to satisfy the same criteria as the *Statistical Classification*. The *Code List* can also contain additional *Code Items* to support a particular use of the *Code List*, such as the inclusion of missing values.

68. The similarities between *Statistical Classifications*, *Code Lists* and *Category Sets* are inherent through their link (as subtypes) to *Node Set*. Similarly, the three types of item which make up each group (*Classification Item*, *Code Item* and *Category Item* respectively) are subtypes of *Node*.

D. Statistical Classification

69. This section describes a *Statistical Classification* and its related management objects, as a particular view of the *Node Set* portion of GSIM. Further detail about *Statistical Classifications* in particular can be found in the GSIM Statistical Classification Model.

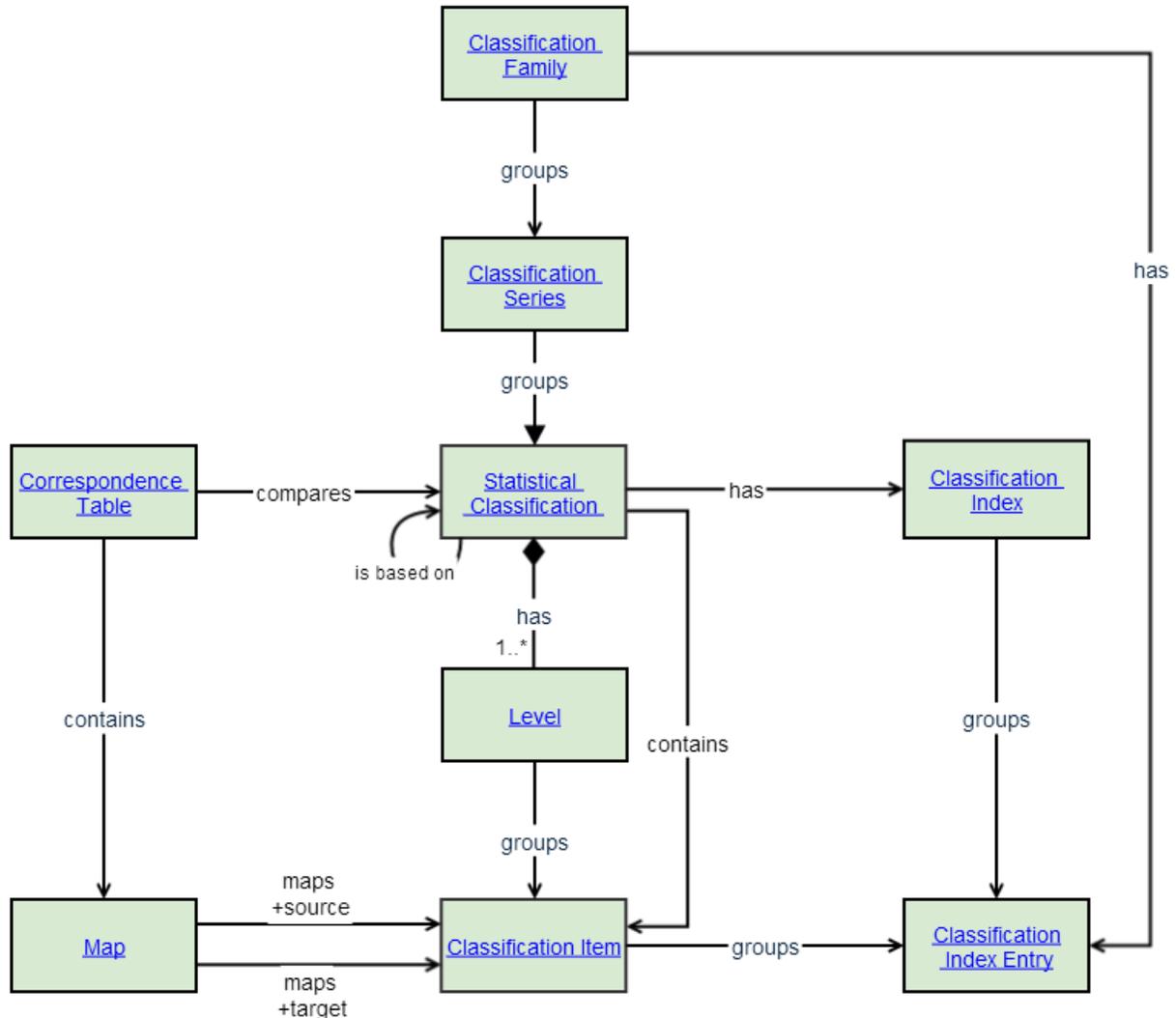


Figure 13. Statistical Classifications

70. The figure above provides an overview of the objects relating to *Statistical Classifications*.

71. A *Classification Family* is a group of *Classification Series* related based on a common *Concept* (e.g. economic activity). A *Classification Series* is an ensemble of one or more *Statistical Classifications* that are based on the same *Concept*. The *Statistical Classifications* in a *Classification Series* are related to each other as versions or updates. Typically, these *Statistical Classifications* have

the same name, for example International Standard Industrial Classification of All Economic Activities (ISIC), or International Standard Industrial Classification of Occupations (ISCO).

72. A *Statistical Classification* is a set of *Categories* which may be assigned to one or more *Represented Variables* 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. One example of a *Statistical Classification* is ISIC rev 4.

73. The *Categories* are defined to reference one or more characteristics of a particular population of interest. 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 a *Category* at the next *Level* up.

74. A *Statistical Classification* has *Categories* that are represented by *Classification Items*. These *Classification Items* are organised into *Levels* determined by the hierarchy. A *Level* is a set of *Concepts* that are mutually exclusive and jointly exhaustive; for example: section, division, group and class in ISIC rev 4.

75. A *Classification Item* combines the meaning from a *Category*, its representation (i.e., *Code*) and additional information in order to meet the *Statistical Classification* criteria, for example "Agriculture, forestry and fishing" and accompanying explanatory text such as information about what is included and excluded.

76. *Statistical Classifications* can be versions or variants. A variant type of *Statistical Classification* is based on a version type of *Statistical Classification*. In a variant the *Categories* of the version may be split, aggregated or regrouped to provide additions or alternatives to the standard order and structure of the original *Statistical Classification*.

77. A *Correspondence Table* is a set of *Maps*. These *Maps* link a *Classification Item* in a *Statistical Classification* with a corresponding *Classification Item* in another *Statistical Classification* via the *Concept* which is common to both *Classification Items*. For example, in a *Correspondence Table* displaying the relationship between ISIC rev 4 and the North American Industry Classification System (NAICS 2007 (US)), "0112 - Growing of Rice" in ISIC Rev 4 is related to "111160 - Rice Farming" in NAICS through the common concept of "growing rice".

78. 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 Programs*.

79. A *Classification Index Entry* is a word or short text (e.g. the name of a locality, an economic activity or an occupational title) describing a type of *Concept* 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 a *Statistical Classification*, they are normally associated with *Classification Items* at the lowest *Level*.

E. Variable

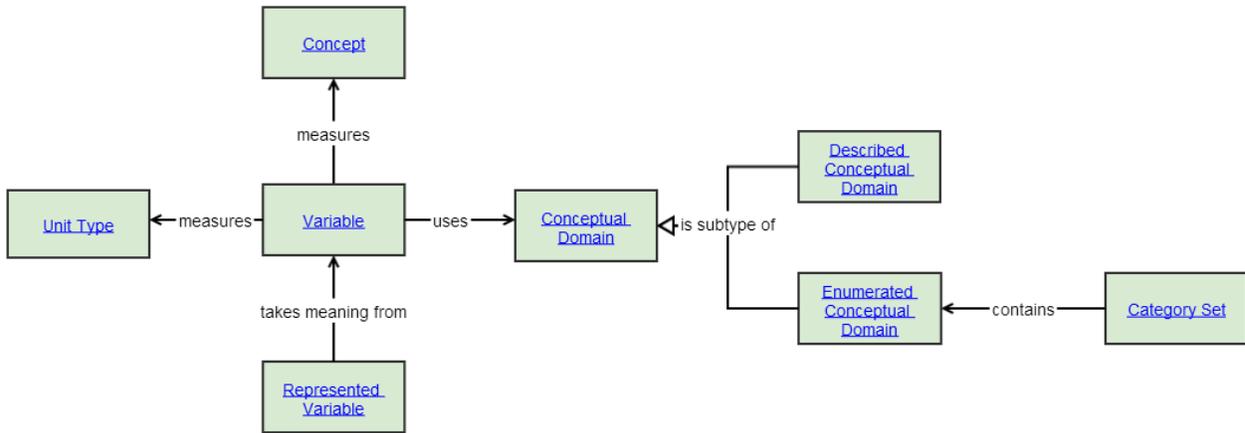


Figure 14. Variable

80. When used as part of a *Business Process*, a *Unit Type* defining a *Population* is associated with a characteristic. The association of *Unit Type* and a *Concept* playing the role of a characteristic is called a *Variable* (see Figure 14). For example, if the *Population* is adults in Netherlands, then a relevant *Variable* might be the *Concept* educational attainment combined with the *Unit Type* person.

81. The *Variable* (person's educational attainment) does not include any information on how the resulting value may be represented. This information (the *Value Domain*) is associated with the *Represented Variable*. This distinction promotes the reuse of a *Variable* definition when what is being measured is conceptually the same but it is represented in a different manner.

82. A derived variable is created by a *Process Step* that applies a *Process Method* to one or more *Process Inputs (Variables)*. The *Process Output* of the *Process Step* is the derived variable.

83. A *Conceptual Domain* is associated with a *Variable*. It has two subtypes: *Described Conceptual Domain* and *Enumerated Conceptual Domain*. An *Enumerated Conceptual Domain*, in combination with a *Category Set*, contains information on the semantics of the *Categories* used by the *Variable*.

F. Represented Variable

84. GSIM assists users in understanding both the meaning and the concrete data-representation of the object. Accordingly, GSIM distinguishes between conceptual and representation levels in the model, to differentiate between the objects used to conceptually describe information, and those that are representational.

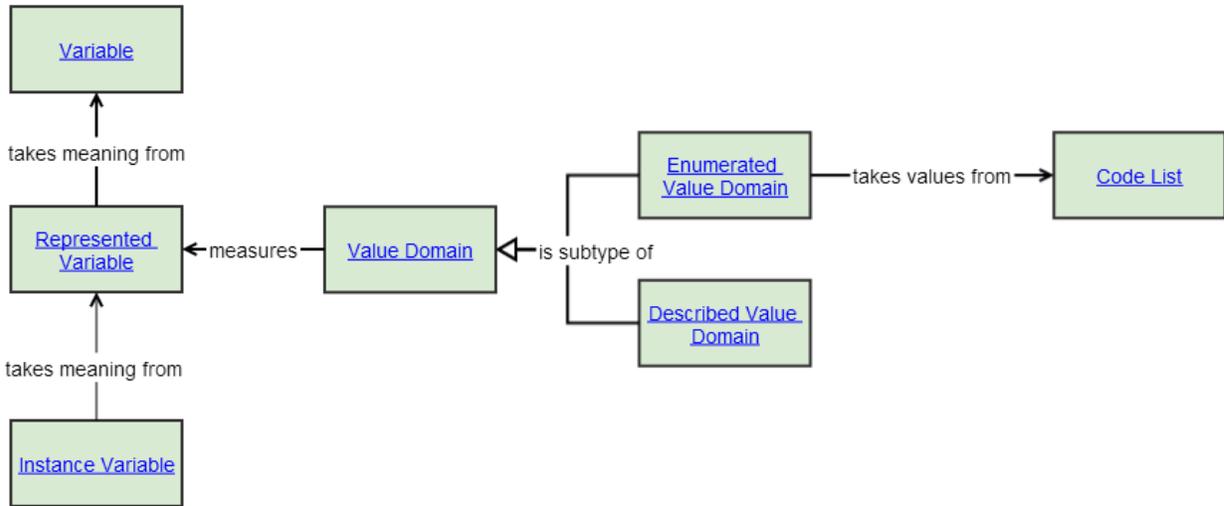


Figure 15. Represented Variable

85. The *Represented Variable* (see Figure 15) adds information that describes how the resulting values may be represented through association with a *Value Domain*. While *Conceptual Domains* are associated with a *Variable*, *Value Domains* are associated with a *Represented Variable*. These two domains are distinguished because GSIM separates the semantic aspect (*Conceptual Domain*) and the representational aspect (*Value Domain*).

86. Both the *Enumerated Value Domain* and the *Described Value Domain* (the two subtypes of *Value Domain*) give information on how the *Represented Variable* is represented. The *Enumerated Value Domain* does this in combination with a *Code List*, while the *Described Value Domain* provides a definition of how to form the values, rather than explicitly listing them.

87. The *Value Domain* includes data type and unit of measure information. The data type contains information on the allowed computations one may perform on the *Datum* (nominal-, ordinal-, interval-data, etc.), while the unit of measure (Tonnes, Count of __, Dollars, etc.) refines the measure of the *Value Domain*. For example gender codes lead to nominal statistical data, whereas age values in years lead to interval data.

G. Instance Variable

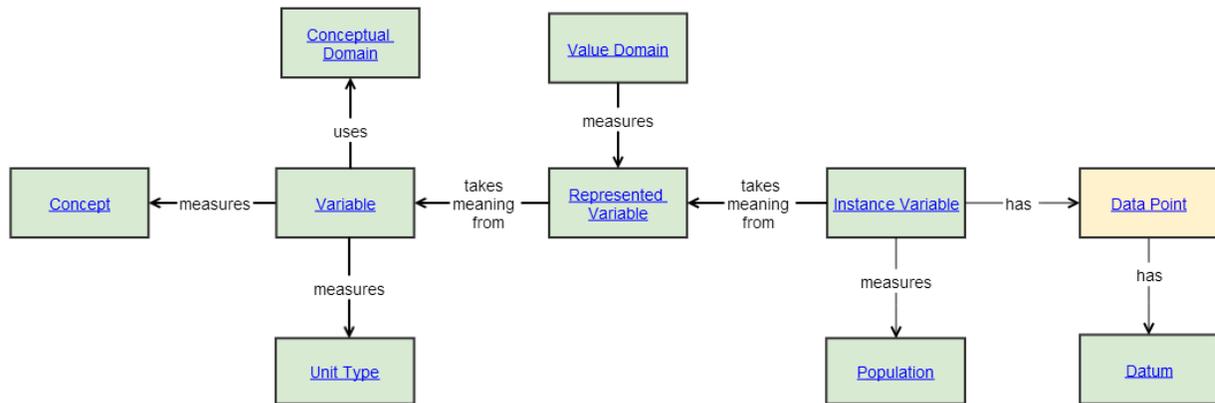


Figure 16. Instance Variable

88. An *Instance Variable* (see Figure 16) is a *Represented Variable* that has been associated with a *Data Set*. This can correspond to a column of data in a database. For example, the “age of all the US presidents either now (if they are alive) or the age at their deaths” is a column of data described by an *Instance Variable*, which is a combination of the *Represented Variable* describing "Person’s Age" and the *Value Domain* of "decimal natural numbers (in years)".

89. A *Datum* is contained within a *Data Point* in a *Data Set*. It may be defined by the measure of a *Value Domain* associated with a describing *Instance Variable*, combined with the link to a *Unit* (for unit data), or a *Population* (for dimensional data).

H. Information Resources

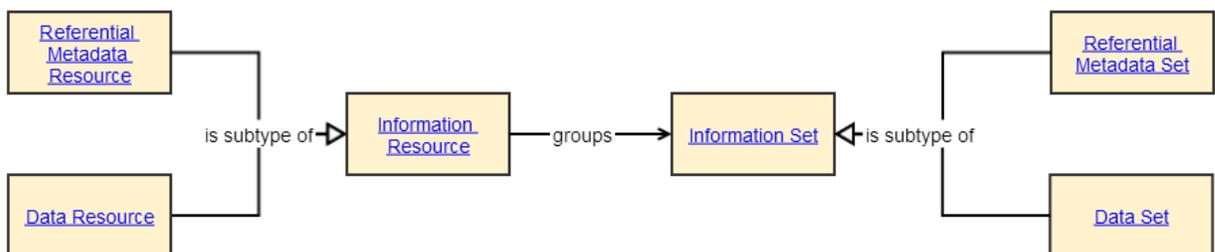


Figure 17. Information Resources

90. Statistical organizations collect, process, analyse and disseminate *Information Sets*, which are either data (*Data Sets*) or referential metadata (*Referential Metadata Sets*).

91. Each *Data Set* must be structured according to a *Data Structure* (for example, a structure for Balance of Payments, Demography, Tourism, Education etc.). In the same way, a *Referential Metadata Set* must be structured according to a *Referential Metadata Structure* (e.g. an organization’s quality framework).

92. *Information Resources* contain *Information Sets*. The main purpose of the *Information Resource* is to aid discovery and management of *Information Sets*, by providing location and other information relevant to these tasks. There are two types of *Information Resource*. *Data Resources* contain *Data Sets*, and *Referential Metadata Resources* contain *Referential Metadata Sets*.

I. Data Sets

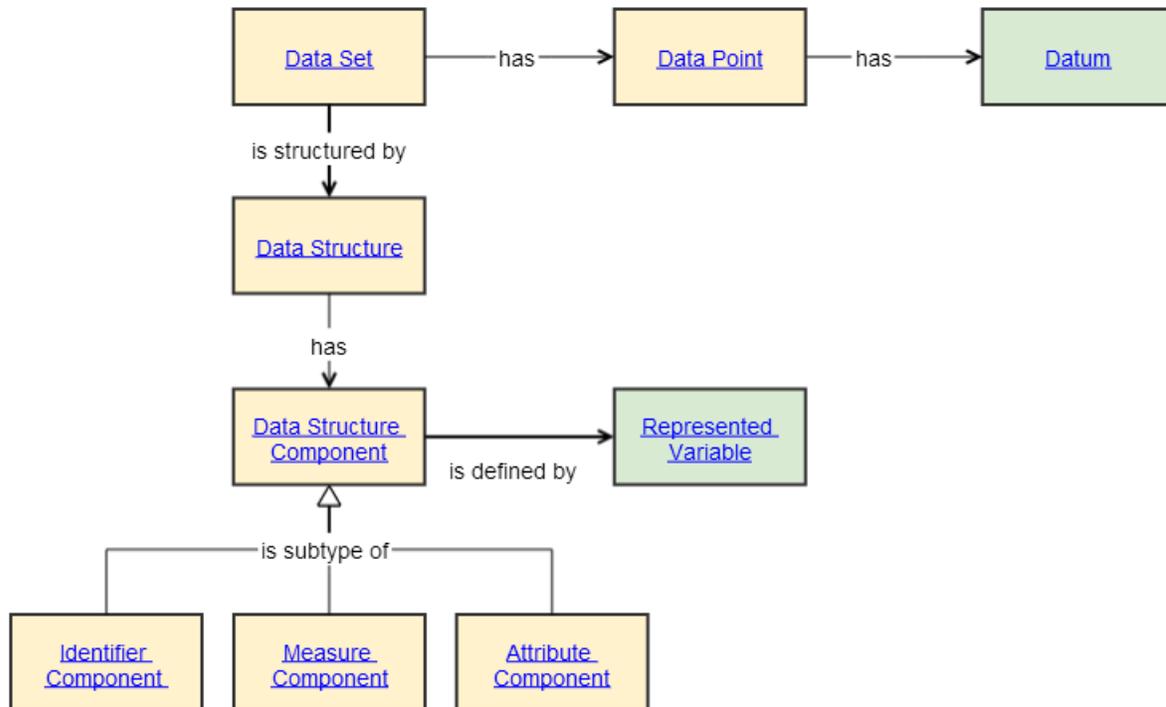


Figure 18. Data Sets

93. A *Data Set* has *Data Points*. A *Data Point* is placeholder (for example, an empty cell in a table) in a *Data Set* for a *Datum*. The *Datum* is the value that populates that placeholder (for example, an item of factual information obtained by measurement or created by a production process). A *Data Structure* describes the structure of a *Data Set* by means of *Data Structure Components* (*Identifier Components*, *Measure Components* and *Attribute Components*). These are all *Represented Variables* with specific roles.

94. *Data Sets* come in different forms, for example as Administrative Registers, Time Series, Panel Data, or Survival Data, just to name a few. The type of a *Data Set* determines the set of specific attributes to be defined, the type of *Data Structure* required (*Unit Data Structure* or *Dimensional Data Structure*), and the methods applicable to the data.

95. For instance, an administrative register is characterized by a *Unit Data Structure*, with attributes such as its original purpose or the last update date of each record. It contains a record identifying variable, and can be used to define a *Population* that is used as a frame, to replace or

complement existing surveys, or as an auxiliary input to imputation. Record matching is an example of a method specifically relevant for registers.

96. An example for a type of *Data Set* defined by a *Dimensional Data Structure* is a time series. It has specific attributes such as frequency and type of temporal aggregation and specific methods, for example, seasonal adjustment, and must contain a temporal variable.

97. Unit data and dimensional data are perspectives on data. Although not typically the case, the same set of data could be described both ways. Sometimes what is considered dimensional data by one organization (for example, a national statistical office) might be considered unit data by another (for example, Eurostat where the unit is the member state). A particular collection of data need not be considered to be intrinsically one or the other. This matter of perspective is conceptual. In GSIM, the distinction is that a *Unit Data Set* contains data about *Units* and a *Dimensional Data Set* contains data about either *Units* or *Populations*.

98. GSIM states that all *Data Sets* must have a structure associated with them. There are, however, cases where a *Data Set* has no structure – because it was not stored or lost, or it is not known. This type of data may become more prevalent for statistical organizations in the future. In order for a statistical organization to use this data, the data will need to go through a process of being structured. For example, in a case of investigation of new potential data sources for a new or changed *Statistical Need*, there will need to be a process where these new data are analyzed to determine their content and structure. It is only after this process that these new *Data Sets* can be described using the *Data Structure* objects. This unstructured data is currently described by GSIM as a *Process Input*. Organizations could extend GSIM to capture this use case by creating a new subtype of the *Information Set* object.

J. Dimensional and Unit Data Structures

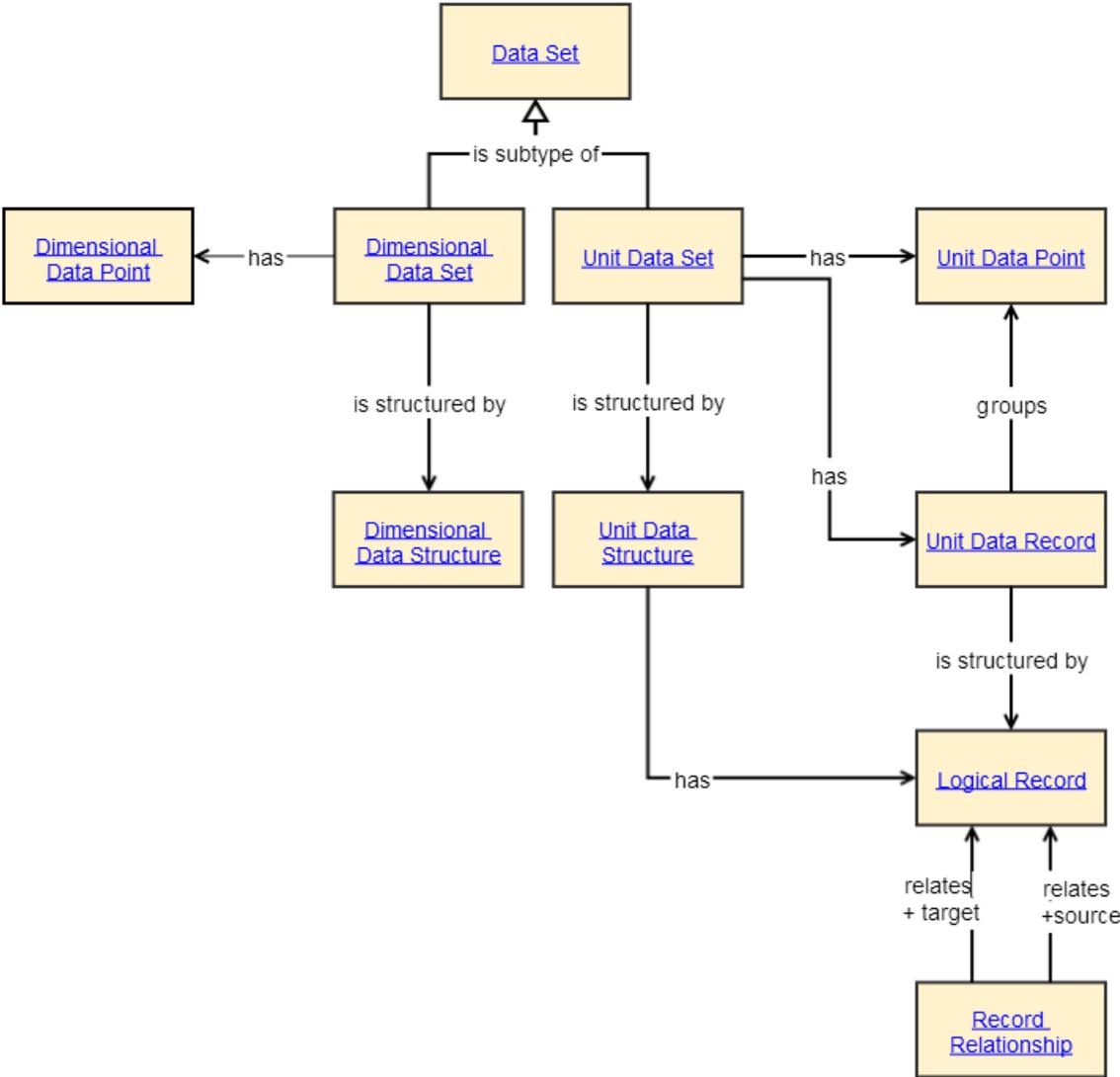


Figure 19. Data Structures

99. A *Dimensional Data Structure* describes the structure of a *Dimensional Data Set* by means of *Represented Variables* with specific roles.

100. The combination of dimensions contained in a *Dimensional Data Structure* creates a key or identifier of the measured values. For instance, country, indicator, measurement unit, frequency, and time dimensions together identify the cells in a cross-country time series with multiple indicators (for example, gross domestic product, gross domestic debt) measured in different units (for example, various currencies, percent changes) and at different frequencies (for example, annual, quarterly). The cells in such a multi-dimensional table contain the observation values.

101. A measure is the variable that provides a container for these observation values. It takes its semantics from a subset of the dimensions of the *Dimensional Data Structure*. In the previous example, indicator and measurement unit can be considered as those semantics-providing dimensions, whereas frequency and time are the temporal dimensions and country the geographic dimension. An example for a measure in addition to the plain 'observation value' could be 'pre-break observation value' in the case of a time series. Dimensions typically refer to *Represented Variables* with coded *Value Domains (Enumerated Value Domains)*, measures to *Represented Variables* with uncoded *Value Domains (Described Value Domains)*.

102. A *Unit Data Structure* describes the structure of a *Unit Data Set* by means of *Represented Variables* with specific roles. It distinguishes between the logical and physical structure of a *Data Set*. A *Unit Data Set* may contain data on more than one type of *Unit*, each represented by its own record type.

103. *Logical Records* describe the structure of such record types, independent of physical features by referring to *Represented Variables* that may include a unit identification (for example, household number). A *Record Relationship* defines source-target relations between *Logical Records*.

K. Referential Metadata Sets

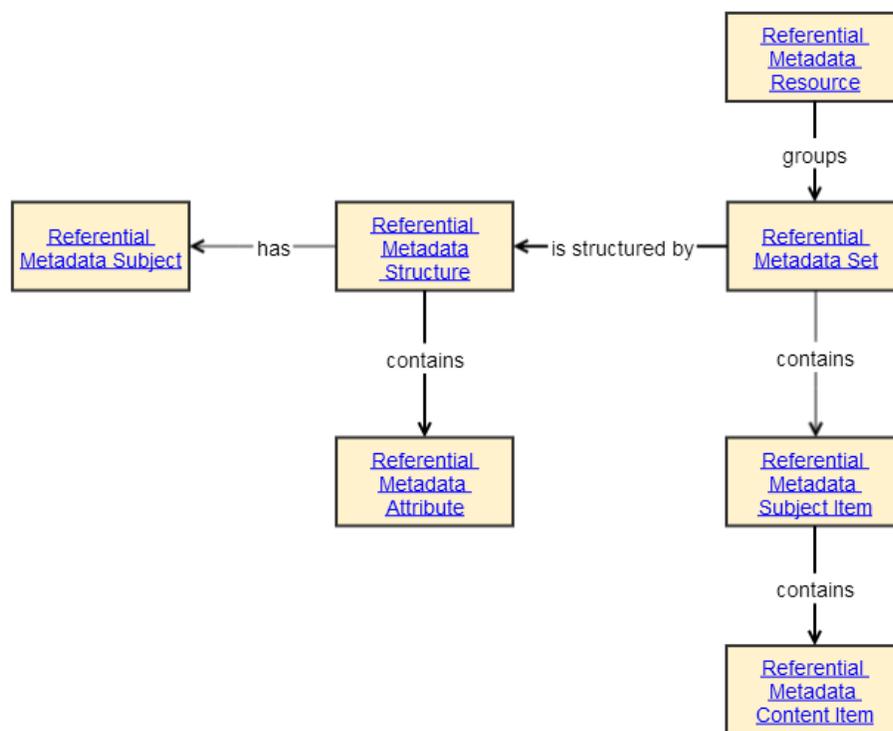


Figure 20. Referential Metadata Sets

104. Information that describes the characteristics of statistics is “referential metadata”. These metadata can be broad, such as about an entire *Statistical Program*, or narrow, such as about an individual *Data Point*. *Referential Metadata Resources*, a special type of *Information Resource*, provide top-level containers for referential metadata.

105. A *Referential Metadata Set* organizes referential metadata, whose structure is defined in a *Referential Metadata Structure*. A *Referential Metadata Structure* specifies both the *Referential Metadata Subject* for which referential metadata may be included, and a structured list of *Referential Metadata Attributes* that can be reported or authored for the given *Referential Metadata Subject*.

106. These subjects may be any GSIM object type, or any *Data Point* or set of *Data Points* created from a specific *Data Structure*.

- Example of a GSIM object type as a *Referential Metadata Subject: Product* for which there is a list specified in a *Value Domain*. The *Value Domain* specifies the list of actual *Products* for which reference metadata can be reported or authored using this *Referential Metadata Structure*.
- Examples of *Referential Metadata Attributes* include status, coverage, methodology description, and quality indicator.

107. A *Referential Metadata Set* contains the actual referential metadata reported or authored. The *Referential Metadata Subject Item* identifies the actual object e.g. actual *Product* such as Balance of Payments and International Investment Position, Australia, June 2013, or actual *Data Points* such as the *Data Points* for a single region within a *Data Set* covering all regions for a country.

108. The *Referential Metadata Content Item* is the actual metadata for the identified *Referential Metadata Subject Item*. Each *Referential Metadata Content Item* contains the reported referential metadata for one *Referential Metadata Attribute* specified in the *Referential Metadata Structure*.

Table 1. Example of Use of GSIM Referential Metadata Objects

GSIM Object	ONS Statistical bulletin: Public Sector Finances, October 2013: Table 1
<i>Referential Metadata Structure</i>	Implicit
<i>Referential Metadata Subject</i>	Data Structure Component
<i>Referential Metadata Attribute</i>	Table footnote
<i>Referential Metadata Set</i>	Footnotes
<i>Referential Metadata Subject Item</i>	<i>Data Structure Component</i> : £ billion; PS Current Budget; PS Current Budget ex APF;...
<i>Referential Metadata Content Item</i>	Footnoted text

IV. Technical information

109. These objects can be seen as the fundamental building blocks that support many of the other objects and relationships in the model. They form the nucleus for the application of GSIM objects. They provide features which are reusable by other objects to support functionality such as identity, versioning etc.

110. The GSIM Base Group consists of two sets of objects

1. Those that give identity and administrative details that are re-usable by other information objects.
2. Those that model the organizations and individual that may provide or consume data and referential metadata.

A: Identity and Administrative Details

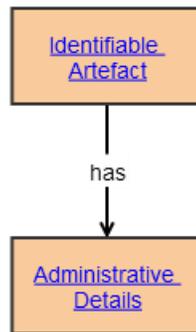


Figure 21. Identifiable Artefacts

111. The only base artefact in GSIM that gives underlying identity is the *Identifiable Artefact*. It can be inherited by any class in GSIM for which identity is required.

112. There is no attempt in GSIM to model the administration of items in repositories such as the maintenance agency, versioning, repository functions. However, the *Identifiable Artefact* does have a link to *Administrative Details* where such details can be added using the GSIM extension methodology.

B. Information Providers, Information Consumers, Organizations, and Individuals

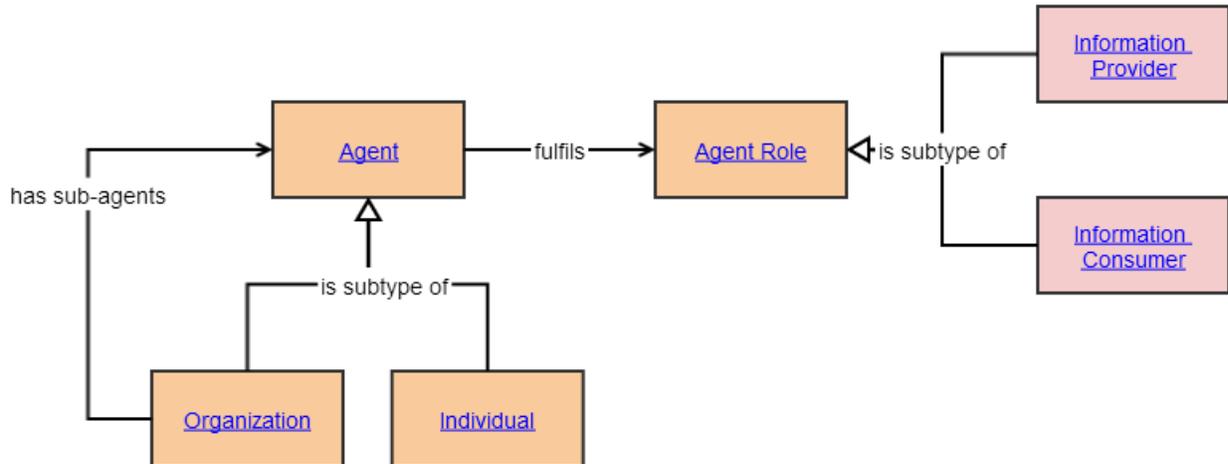


Figure 22. Agents

113. *Information Providers* and *Information Consumers* are the respective sources for and the targets of data and referential metadata collection and dissemination. Each *Agent* can play the role of *Information Provider* or *Information Consumer* in a particular context of collection or dissemination. The same *Agent* may play the role of *Information Provider* in one context and the role of *Information Consumer* in another context. For any one *Agent Role* there must be a single *Agent* that plays the role: this is actual *Organization* or *Individual* that is the *Information Provider* or *Information Consumer*.

114. If the *Agent* is an *Organization* then it is possible to specify the structure of the *Organization* in terms of sub *Organizations* or *Individuals*.

Annex A: Exchange Channels for Data Collection

115. All data collection is modelled in GSIM using the *Exchange Channel* object, which represents the mechanism by which data comes into the statistical organization. This object is always extended into sub-classes, to describe specific sources of data collection. There is a growing emphasis on the use of non-survey data sources, as these often represent sources of data which can be realized more quickly and at lower cost. The model can be extended by adding further sub-classes to represent other sources, as required.

116. Two common forms of data collection are the use of data from administrative registers, and the collection of data by programmatically "scraping" web sites for their content. To reflect this, GSIM models two non-survey data sources - *Administrative Registers* and *Web Scraper Channels*. It also models one survey data source - *Questionnaire*. The following sections describe how each of these is modelled in GSIM.

A: Administrative Registers

117. In the illustration below, we show how GSIM can model administrative registers as data sources. The sub-type of *Exchange Channel* which represents administrative register as a source is the *Administrative Register* object.

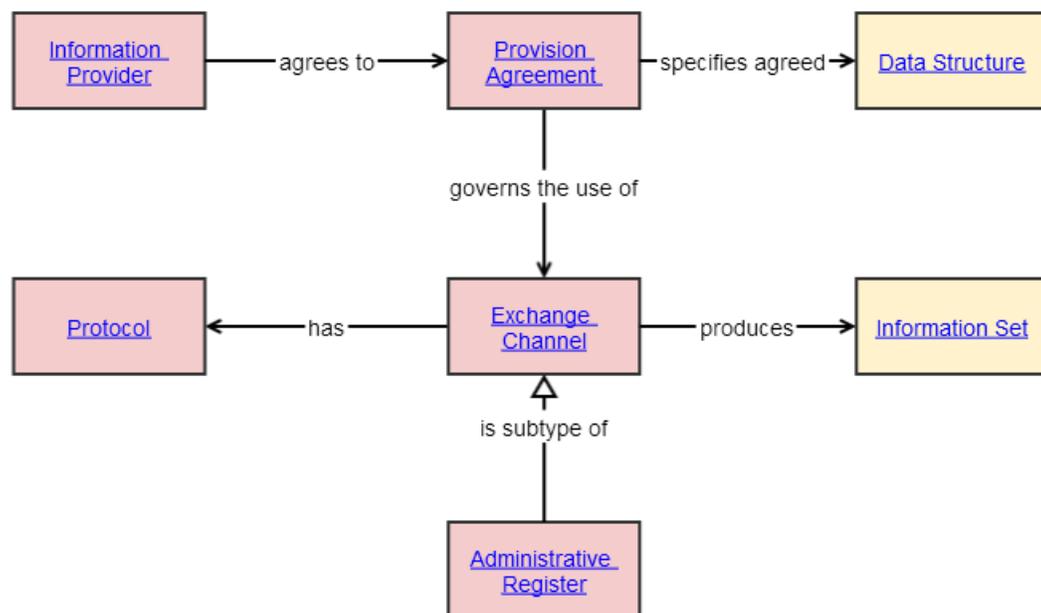


Figure 23. Administrative Register

118. The important information about *Administrative Registers* includes:

- the agreement between the statistical organization and the provider of the register data,

- the protocol for accessing the data, and
- the structure of the data to be received.

119. Each of these can be described by information objects which are inherited from the *Exchange Channel* by the subtype *Administrative Register*.

120. The agreement between the statistical organization and the *Information Provider* is represented using a *Provision Agreement* object. This shows the relationship between our *Administrative Register* and the *Organization* with which the agreement exists (the *Information Provider*). There is typically an agreed structure for the data - described in the data collection agreement - but this can sometimes be different from the structure of the data actually received. The *Information Provider* object has a relationship to the *Data Structure* object. This represents the agreed structure of the information to be collected from the administrative register.

121. The *Administrative Register* object also inherits a relationship to a *Protocol* object from its parent *Exchange Channel*. The *Protocol* object captures the details of the technical process by which the register data is to be collected. This might be through the use of a standard mechanism such as an SDMX data exchange, a technical mechanism such as a query to a database, or even a manual process.

122. The *Exchange Channel* object also allows for its *Administrative Register* sub-type to link to the *Data Set* actually collected, which references its own *Data Structure* object. By comparing the collected data and its structure against the "agreed" structure, the received data can be validated. Note that if the information being collected were referential metadata, the *Referential Metadata Structure* object would replace the *Data Structure* in the diagram, and similarly the *Referential Metadata Set* would replace the *Data Set*.

123. Once a *Data Set* is collected, we have all of the usual objects such as the *Data Point*, the *Instance Variable*, and so on. As more *Data Sets* are collected, these can in turn be stored in a *Data Resource*, which would hold all of the data coming from the *Administrative Register* over time.

B. Web Scraping for Data Collection

124. The second non-questionnaire data source to be modelled is a web scraper, as seen in the diagram below.

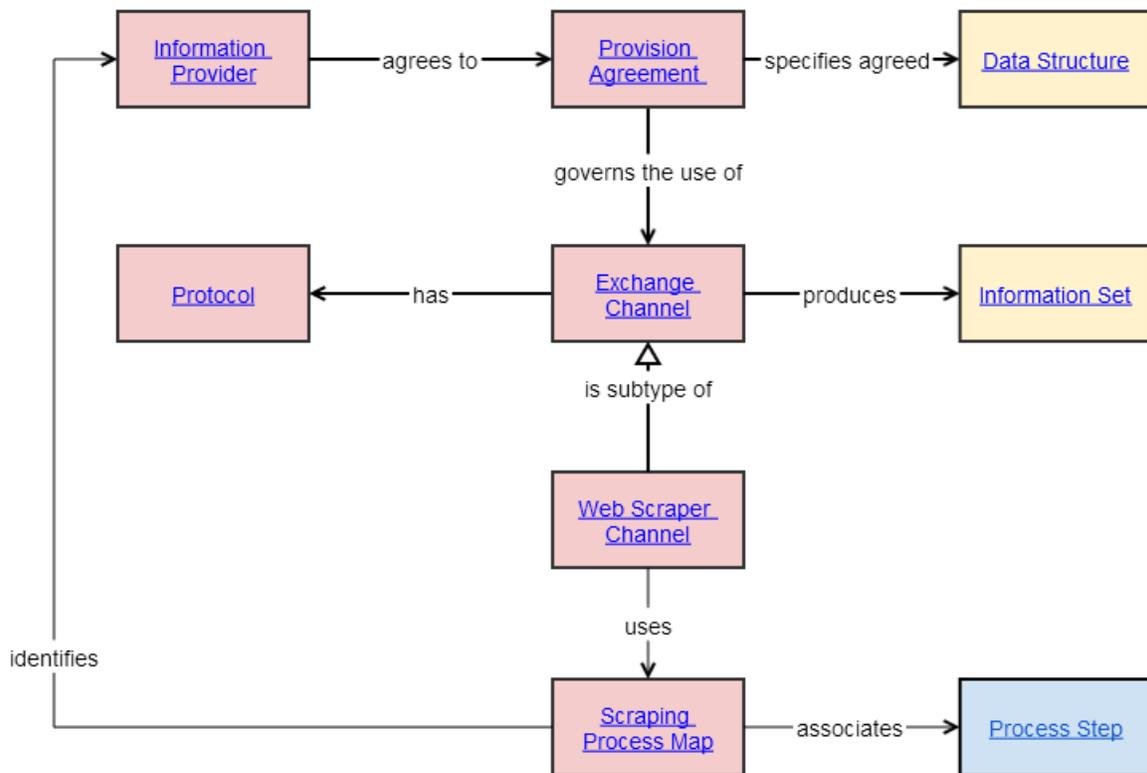


Figure 24. Web Scraping Channel

125. There will be at least a notional *Provision Agreement* between the statistical organization collecting the data through the web scraper and each of the organizations whose sites are being scraped (the *Information Provider*), even if this is only the terms and conditions of accessing the data provider's website. In many cases, Internet robots used to do web scraping are blocked from websites, and there is typically contact between the scraping organization and the data providing one, to make sure that access is not blocked, and to know when the website's structure might change.

126. Although perhaps trivial, the *Protocol* being used will need to be recorded, being either HTTP or HTTPS (by definition, the scraping tool is operating on the web).

127. Each website is scraped using a software application. Due to the varying structure of different websites, often a different software tool will be needed for each website. Further, every time the website being scraped is structurally modified, adjustments may need to be made to the software tool. The software tools themselves are represented as *Process Steps* in GSIM, these being the result of a design process administered through a *Statistical Program*, which are capable of being executed to programmatically collect the data.

128. The management of the mappings between each website and the software tool used to scrape it is important information to capture. It is necessary to be able to describe the software tools used to

scrape websites, and their link to the websites for which they are designed. This is done using the *Scraping Process Map* object. This object links a *Process Step* and one or more *Information Providers* (the organizations whose sites that software tool can scrape). A set of these gives the links needed to manage the mappings between the web scraping tools and the sites from which the data is collected.

129. As for the *Administrative Register* above, the structure of the data to be collected and the information regarding the actual data collected are captured in the *Data Set* and *Data Structure* objects.

C: Survey Data Collection

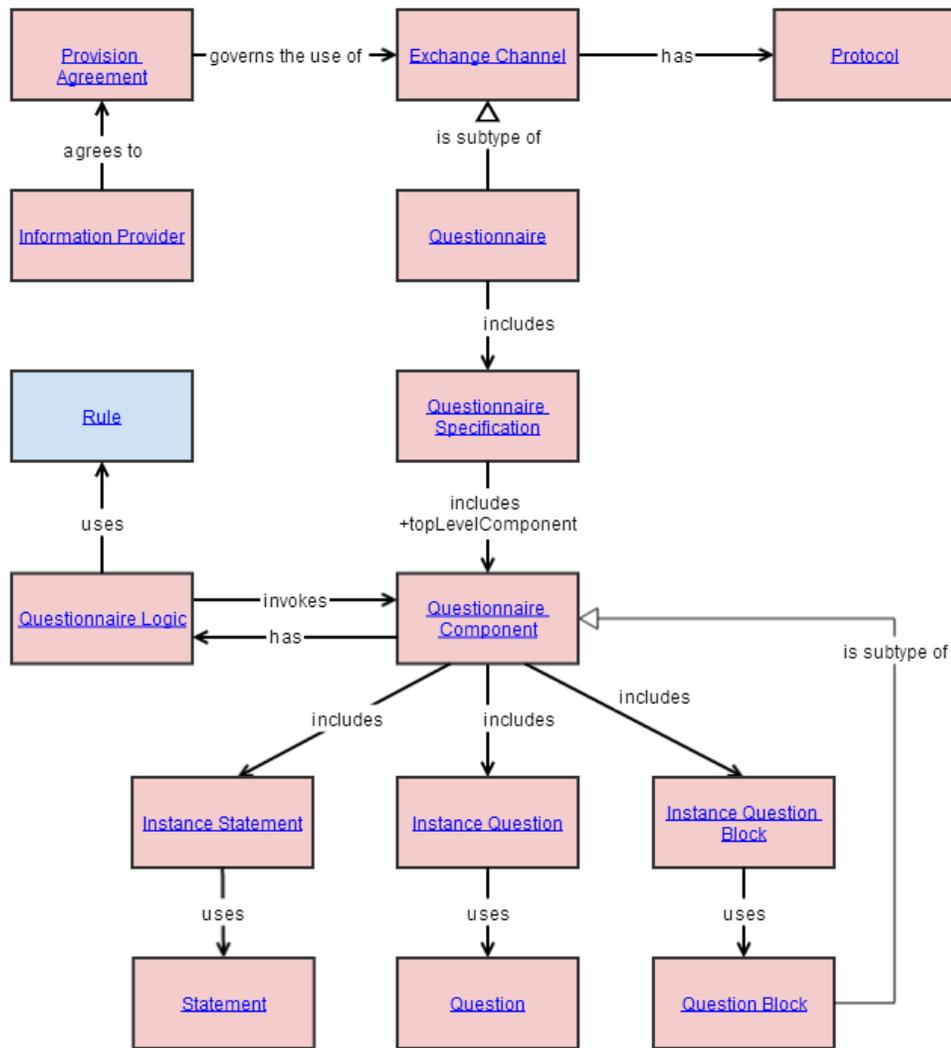


Figure 24. Questionnaire

130. Although more and more alternative data collection methods (such as *Administrative Register* sources) will be utilized by statistical organizations, it is envisaged that for the foreseeable future,

surveys will continue to make extensive use of questionnaires for the purpose of data collection. As such, *Questionnaire* is included in GSIM as a subtype of *Exchange Channel*.

131. The *Provision Agreement* establishes the relationship between the *Questionnaire* and the *Information Provider*, in the form of some agreement to provide data to the collecting organization. This is sometimes (especially in the case of collections for official business statistics) a mandatory requirement specified by law.

132. Depending upon the survey it will be used for, the *Questionnaire* could be developed as one or more generic types. Each instance of a *Questionnaire* will be constructed by reference to the *Questionnaire Specification*. A *Questionnaire* could take the form of a standard *Questionnaire Specification* (i.e. the layout would be the same or have a relatively small number of variations) for a particular survey, or at the other extreme, the *Questionnaire Specification* could be tailored to each *Information Provider* (or *Unit*) selected for the survey.

133. The *Questionnaire Specification* will consist of a top level *Questionnaire Component*, which will itself be made up of lower level *Questionnaire Components*, built up in a hierarchical manner. Each *Questionnaire Component* will in turn be made up of a number of *Instance Question Blocks*, *Instance Questions*, and *Instance Statements*.

134. In its simplest form, a *Questionnaire Specification* would have a single *Questionnaire Component* made up of a number of simple *Instance Question Blocks*, *Instance Questions* and *Instance Statements*, but will also have associated *Questionnaire Logic*, which will govern the navigation and validation of *Questions* and responses within the *Questionnaire Specification*. The *Questionnaire Logic* will implement a number of *Rules*, which will carry out such work as the evaluation of the response data in terms of the range of acceptable values. In most cases, the *Questionnaire Specification* will be built up using of several *Questionnaire Component* levels, each with their associated *Questionnaire Logic*.

135. *Question Block*, *Question* and *Statement* are reusable artifacts, which will be implemented in the *Questionnaire Specification* by means of the *Instance Question Blocks*, *Instance Questions* and *Instance Statements* respectively. It might be that the actual *Question Blocks*, *Questions* and *Statements* would be stored in some searchable library for use during the *Questionnaire Specification* development process.

136. *Questions* can take the form of a multiple question item, and can be hierarchical. *Question* will have a connection to one or more *Variables*, and will also be associated with a *Value Domain*, specifying the constraints of the values which can be assigned to the *Variables* in the response to the *Question*.

137. Different *Protocols* (modes of collection) would require different implementations of *Questionnaire*. For example, if a *Questionnaire Specification* is designed for collection via a web page, a similar *Questionnaire Specification* design containing all the same *Question Blocks*, and *Questions* but is intended for collection via a printed paper form, it would be implemented in a different instance of *Questionnaire*. Thus, where a multi-mode data collection strategy is adopted for

a survey, separate *Questionnaire Specifications* would be needed to be developed for each *Protocol* (mode) employed, and they would be implemented in different *Questionnaire* instances.

138. The navigation and validation aspects within the *Questionnaire* will need to be designed with the *Protocol* (mode of capture) in mind. For example, if the *Questionnaire* is to be rendered as a paper form, then the navigation will be implemented using an *Instance Statement* in the form of a text instruction to the *Information Provider* such as "If the response to gender question is 'MALE' then go to question X". If a similar *Questionnaire* were to be rendered as a web form, then the navigation could be automated and the *Information Provider* would be automatically routed to 'question X'.

Annex B: Glossary

Object	Group	Definition	Explanatory Text	Synonyms
Administrative Details	Base	A placeholder for extensions to the model based on an organization's administrative needs.	The <i>Administrative Details</i> object is designed to act as a 'placeholder' to allow for future extensions to the existing model. It allows for further information to be added about the administrative details required to maintain the other objects outlined by GSIM.	
Administrative Register	Exchange	A source of administrative information which is obtained from an external organization (or sometimes from another department of the same organization)	The <i>Administrative Register</i> is a source of administrative information obtained from external organizations. The <i>Administrative Register</i> would be provided under a <i>Provision Agreement</i> with the supplying organization. This administrative information is usually collected for an organization's operational purposes, rather than for statistical purposes.	
Agent	Base	An actor that performs a role in relation to the statistical <i>Business Process</i> .	An <i>Agent</i> may be either an <i>Organization</i> or an <i>Individual</i> . An <i>Organization</i> may be an entire organization or entities within a larger organization, such as departments or divisions. An <i>Organization</i> may have sub <i>Agents</i> , which may be either other <i>Organizations</i> within the parent <i>Organization</i> or <i>Individuals</i> that belong to that <i>Organization</i> .	
Agent Role	Base	The function or activities of an <i>Agent</i> , in regard to their involvement in the statistical <i>Business Process</i> .	An <i>Agent Role</i> may apply to either type of <i>Agent</i> - an <i>Organization</i> or <i>Individual</i> . A common example would be to identify which individuals or departments within an organization provide administrative data.	
Assessment	Business	The result of the analysis of the quality and effectiveness of any activity undertaken by a statistical organization and recommendations on how these can be improved.	An <i>Assessment</i> can be of a variety of types. One example may include a gap analysis, where a current state is determined along with what is needed to reach its target state. Alternately, an <i>Assessment</i> may compare current processes against a set of	

			<p>requirements, for example a new <i>Statistical Need</i> or change in the operating environment.</p> <p>An <i>Assessment</i> can use various information objects as inputs, whether they are the main objects that the <i>Assessment</i> is about or auxiliary information objects that help accomplish the <i>Assessment</i>.</p>	
Attribute Component	Structures	The role given to a <i>Represented Variable</i> in the context of a <i>Data Structure</i> , which supplies information other than identification or measures.	For example the publication status of an observation (e.g. provisional, final, revised)	
Business Case	Business	A proposal for a body of work that will deliver outputs designed to achieve outcomes. A <i>Business Case</i> will provide the reasoning for undertaking a <i>Statistical Support Program</i> to initiate a new <i>Statistical Program Design</i> for an existing <i>Statistical Program</i> , or an entirely new <i>Statistical Program</i> , as well as the details of the change proposed.	A <i>Business Case</i> is produced as a result of a detailed consideration of a <i>Change Definition</i> . It sets out a plan for how the change described by the <i>Change Definition</i> can be achieved. A <i>Business Case</i> usually comprises various evaluations. The <i>Business Case</i> will specify the stakeholders that are impacted by the <i>Statistical Need</i> or by the different solutions that are required to implement it.	
Business Function	Business	Something an enterprise does, or needs to do, in order to achieve its objectives.	<p>A <i>Business Function</i> delivers added value from a business point of view. It is delivered by bringing together people, processes and technology (resources), for a specific business purpose.</p> <p><i>Business Functions</i> answer in a generic sense "What business purpose does this <i>Business Service</i> or <i>Process Step</i> serve?" Through identifying the <i>Business Function</i> associated with each <i>Business Service</i> or <i>Process Step</i> it increases the documentation of the use of the associated <i>Business Services</i> and <i>Process Steps</i>, to enable future reuse.</p> <p>A <i>Business Function</i> may be defined directly with descriptive text and/or through reference to an existing catalogue of <i>Business Functions</i>. The phases and sub</p>	

			<p>processes defined within GSBPM can be used as an internationally agreed basis for cataloguing high level <i>Business Functions</i>. A catalogue might also include <i>Business Functions</i> defined at a lower level than "sub process". For example, "Identify and address outliers" might be catalogued as a lower level <i>Business Function</i> with the "Review, validate and edit" function (5.3) defined within GSBPM.</p>	
Business Process	Business	The set of <i>Process Steps</i> to perform one of more <i>Business Functions</i> to deliver a <i>Statistical Program Cycle</i> or <i>Statistical Support Program</i> .	<p>For example, a particular <i>Statistical Program Cycle</i> might include several data collection activities, the corresponding editing activities for each collection and the production and dissemination of final outputs. Each of these may be considered separate <i>Business Processes</i> for the <i>Statistical Program Cycle</i>.</p>	
Business Service	Business	A means of performing a <i>Business Function</i> (an ability that an organization possesses, typically expressed in general and high level terms and requiring a combination of organization, people, processes and technology to achieve).	<p>A <i>Business Service</i> may provide one means of accessing a particular <i>Business Function</i>. The operation of a <i>Business Service</i> will perform one or more <i>Business Processes</i>.</p> <p>The explicitly defined interface of a <i>Business Service</i> can be seen as representing a "service contract". If particular inputs are provided then the service will deliver particular outputs in compliance within specific parameters (for example, within a particular period of time).</p> <p>Note: The interface of a <i>Business Service</i> is not necessarily IT based. For example, a typical postal service will have a number of service interfaces:</p> <ul style="list-style-type: none"> - Public letter box for posting letters - Counter at post office for interacting with postal workers 	
Category	Concepts	A <i>Concept</i> whose role is to extensionally define and measure a characteristic.	<p><i>Categories</i> for the <i>Concept</i> of sex include: Male, Female</p>	class

			<p>Note: An extensional definition is a description of a <i>Concept</i> by enumerating all of its sub ordinate <i>Concepts</i> under one criterion or sub division.</p> <p>For example - the Noble Gases (in the periodic table) is extensionally defined by the set of elements including Helium, Neon, Argon, Krypton, Xenon, Radon. (ISO 1087-1)</p>	
Category Item	Concepts	An element of a <i>Category Set</i> .	A type of <i>Node</i> particular to a <i>Category Set</i> type of <i>Node Set</i> . A <i>Category Item</i> contains the meaning of a <i>Category</i> without any associated representation.	
Category Set	Concepts	A list of <i>Categories</i>	<p>A <i>Category Set</i> is a type of <i>Node Set</i> which groups <i>Categories</i> through the use of <i>Category Items</i>. The <i>Categories</i> in a <i>Category Set</i> typically have no assigned <i>Designations (Codes)</i>.</p> <p>For example: Male, Female</p>	
Change Definition	Business	A structured, well-defined specification for a proposed change.	<p>A related object - the <i>Statistical Need</i> - is a change expression as it has been received by an organization. A <i>Statistical Need</i> is a raw expression of a proposed change, and is not necessarily well-defined. A <i>Change Definition</i> is created when a <i>Statistical Need</i> is analyzed by an organization, and expresses the raw need in well-defined, structured terms.</p> <p>A <i>Change Definition</i> does not assess the feasibility of the change or propose solutions to deliver the change - this role is satisfied by the <i>Business Case</i> object. The precise structure or organization of a <i>Change Definition</i> can be further specified by rules or standards local to a given organization. It also includes the specific <i>Concepts</i> to be measured and the <i>Population</i> that is under consideration.</p>	

			<p>Once a <i>Statistical Need</i> has been received, the first step is to do the conceptual work to establish what it is we are trying to measure. The final output of this conceptual work is the <i>Change Definition</i>.</p> <p>The next step is to assess how we are going to make the measurements - to design a solution and put forward a proposal for a body of work that will deliver on the requirements of the original <i>Statistical Need</i></p>	
Classification Family	Concepts	A <i>Classification Family</i> is a group of <i>Classification Series</i> related from a particular point of view. The <i>Classification Family</i> is related by being based on a common <i>Concept</i> (e.g. economic activity).	Different classification databases may use different types of <i>Classification Families</i> and have different names for the families, as no standard has been agreed upon.	
Classification Index	Concepts	A <i>Classification Index</i> is an ordered list (alphabetical, in code order etc.) of <i>Classification Index Entries</i> . A <i>Classification Index</i> can relate to one particular or to several <i>Statistical Classifications</i> .	<p>A <i>Classification Index</i> shows the relationship between text found in statistical data sources (responses to survey questionnaires, administrative records) and one or more <i>Statistical Classifications</i>. A <i>Classification Index</i> may be used to assign the codes for Classification Items to observations in statistical collections.</p> <p>A <i>Statistical Classification</i> is a subtype of <i>Node Set</i>. The relationship between <i>Statistical Classification</i> and <i>Classification Index</i> can also be extended to include the other <i>Node Set</i> types - <i>Code List</i> and <i>Category Set</i>.</p>	
Classification Index Entry	Concepts	A <i>Classification Index Entry</i> 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 <i>Classification Item</i> applies, together with the code of the corresponding <i>Classification Item</i> . Each	A <i>Classification Item</i> is a subtype of <i>Node</i> . The relationship between <i>Classification Item</i> and <i>Classification Index Entry</i> can also be extended to include the other <i>Node</i> types - <i>Code Item</i> and <i>Category Item</i> .	

		<i>Classification Index Entry</i> typically refers to one item of the <i>Statistical Classification</i> . Although a <i>Classification Index Entry</i> may be associated with a <i>Classification Item</i> at any <i>Level</i> of a <i>Statistical Classification</i> , <i>Classification Index Entries</i> are normally associated with items at the lowest <i>Level</i> .		
Classification Item	Concepts	A <i>Classification Item</i> represents a <i>Category</i> at a certain <i>Level</i> within a <i>Statistical Classification</i> . It defines the content and the borders of the <i>Category</i> . A <i>Unit</i> can be classified to one and only one item at each <i>Level</i> of a <i>Statistical Classification</i> .		
Classification Series	Concepts	A <i>Classification Series</i> is an ensemble of one or more <i>Statistical Classifications</i> , based on the same concept, and related to each other as versions or updates. Typically, these <i>Statistical Classifications</i> have the same name (for example, ISIC or ISCO).		
Code	Concepts	A <i>Designation</i> for a <i>Category</i> .	<i>Codes</i> are unique within their <i>Code List</i> . Example: M (Male) F (Female).	
Code Item	Concepts	An element of a <i>Code List</i> .	A type of <i>Node</i> particular to a <i>Code List</i> type of <i>Node Set</i> . A <i>Code Item</i> combines the meaning of the included <i>Category</i> with a <i>Code</i> representation.	
Code List	Concepts	A list of <i>Categories</i> where each <i>Category</i> has a predefined <i>Code</i> assigned to it.	A kind of <i>Node Set</i> for which the <i>Category</i> contained in each <i>Node</i> has a <i>Code</i> assigned as a <i>Designation</i> . For example: 1 - Male 2 - Female	

Code Value	Concepts	An alpha-numeric string used to represent a <i>Code</i> .	A <i>Code Value</i> is a subtype of <i>Sign</i> - a way of denoting the value of a <i>Code</i> . This is a kind of <i>Sign</i> used for <i>Codes</i> .	
Concept	Concepts	Unit of thought differentiated by characteristics.		
Concept System	Concepts	Set of <i>Concepts</i> structured by the relations among them.	Here are 2 examples 1) Concept of Sex: Male, Female, Other 2) ISIC (the list is too long to write down)	
Conceptual Domain	Concepts	Set of valid <i>Concepts</i> .	The <i>Concepts</i> can be described by either enumeration or by an expression.	
Correspondence Table	Concepts	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; or, 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.	A Statistical Classification is a subtype of Node Set. The relationship between Statistical Classification and Correspondence Table can also be extended to include the other Node Sets - Code List and Category Set.	
Data Point	Structures	A placeholder (or cell) for the value of an <i>Instance Variable</i>	Field in a <i>Data Structure</i> which corresponds to a cell in a table. The <i>Data Point</i> is structural and distinct from the value (the <i>Datum</i>) that it holds.	
Data Resource	Structures	An organized collection of stored information made of one or more <i>Data Sets</i> .	<i>Data Resources</i> are collections of data that are used by a statistical activity to produce information. <i>Data Resource</i> is a specialization of an <i>Information Resource</i> .	data source
Data Set	Structures	An organized collection of data.	Examples of <i>Data Sets</i> could be observation registers, time series, longitudinal data, survey data, rectangular data sets, event-history data, tables, data tables, cubes, registers, hypercubes, and matrixes. A broader term	database, data file, file, table

			for <i>Data Set</i> could be data. A narrower term for <i>Data Set</i> could be data element, data record, cell, field.	
Data Structure	Structures	Defines the structure of an organized collection of data (<i>Data Set</i>).	The structure is described using <i>Data Structure Components</i> that can be either <i>Attribute Components</i> , <i>Identifier Components</i> or <i>Measure Components</i> . Examples for unit data include social security number, country of residence, age, citizenship, country of birth, where the social security number and the country of residence are both identifying components and the others are measured variables obtained directly or indirectly from the person (<i>Unit</i>).	
Data Structure Component	Structures	The role of the <i>Represented Variable</i> in the context of a <i>Data Structure</i> .	A <i>Data Structure Component</i> can be an <i>Attribute Component</i> , <i>Measure Component</i> or an <i>Identifier Component</i> . Example of <i>Attribute Component</i> : The publication status of an observation such as provisional, revised. Example of <i>Measure Component</i> : age and height of a person in a <i>Unit Data Set</i> or number of citizens and number of households in a country in a <i>Data Set</i> for multiple countries (<i>Dimensional Data Set</i>). Example of <i>Identifier Component</i> : The personal identification number of a Swedish citizen for unit data or the name of a country in the European Union for dimensional data.	
Datum	Concepts	A value.	A <i>Datum</i> is the actual instance of data that was collected or derived. It is the value which populates a <i>Data Point</i> . A <i>Datum</i> is the value found in a cell of a table.	value
Described Conceptual Domain	Concepts	A <i>Conceptual Domain</i> defined by an expression.	For example: All real numbers between 0 and 1.	Non-enumerated

				conceptual domain
Described Value Domain	Concepts	A <i>Value Domain</i> defined by an expression.	For example: All real decimal numbers between 0 and 1.	Non-enumerated value domain
Designation	Concepts	The name given to an object for identification.	The association of a <i>Concept</i> with a <i>Sign</i> that denotes it.	
Dimensional Data Point	Structures	A placeholder (or cell) for the value of an <i>Instance Variable</i> with respect to either a <i>Unit</i> or <i>Population</i> .	A <i>Dimensional Data Point</i> is uniquely identified by the combination of exactly one value for each of the dimensions (<i>Identifier Component</i>) and one measure (<i>Measure Component</i>). There may be multiple values for the same <i>Dimensional Data Point</i> that is for the same combination of dimension values and the same measure. The different values represent different versions of the data in the <i>Data Point</i> . Values are only distinguished on the basis of quality, date/time of measurement or calculation, status, etc. This is handled through the mechanisms provided by the <i>Datum</i> information object.	cell
Dimensional Data Set	Structures	A collection of dimensional data that conforms to a known structure.		
Dimensional Data Structure	Structures	Describes the structure of a <i>Dimensional Data Set</i> .	For example (city, average income, total population) where the city is the <i>Identifier Component</i> and the others are measured variables.	
Enumerated Conceptual Domain	Concepts	A <i>Conceptual Domain</i> expressed as a list of <i>Categories</i>	For instance, the Sex <i>Categories</i> : 'Male' and 'Female'	
Enumerated Value Domain	Concepts	A <i>Value Domain</i> expressed as a list of <i>Categories</i> and associated <i>Codes</i> .	Example - Sex Codes <m, male>; <f, female>; <o, other>.	
Environment Change	Business	A requirement for change that originates from a change in the operating environment of the statistical organization.	An <i>Environment Change</i> reflects change in the context in which a statistical organization operates. <i>Environment Changes</i> can be of different origins and also take different forms. They can result from a	

			precise event (budget cut, new legislation enforced) or from a progressive process (technical or methodological progress, application or tool obsolescence). Other examples of <i>Environment Changes</i> include the availability of a new <i>Information Resource</i> , the opportunity for new collaboration between organizations, etc.	
Exchange Channel	Exchange	A means of exchanging data.	An abstract object that describes the means to receive (data collection) or send (dissemination) information. Different <i>Exchange Channels</i> are used for collection and dissemination. Examples of collection <i>Exchange Channel</i> include <i>Questionnaire</i> , <i>Web Scraper Channel</i> and <i>Administrative Register</i> . The only example of a dissemination <i>Exchange Channel</i> currently contained in GSIM is <i>Product</i> . Additional <i>Exchange Channels</i> can be added to the model as needed by individual organizations.	
Identifiable Artefact	Base	An abstract class that comprises the basic attributes and associations needed for identification, naming and other documentation.	An instance of any GSIM information object is an <i>Identifiable Artifact</i> .	
Identifier Component	Structures	The role given to a <i>Represented Variable</i> in the context of a <i>Data Structure</i> to identify the unit in an organized collection of data.	An <i>Identifier Component</i> is a sub-type of <i>Data Structure Component</i> . The personal identification number of a Swedish citizen for unit data or the name of a country in the European Union for dimensional data.	
Individual	Base	A person who acts, or is designated to act towards a specific purpose.		
Information Consumer	Exchange	A person or organization that consumes disseminated data.	The <i>Information Consumer</i> accesses a set of information via a <i>Product</i> (or potentially via another Exchange Channel), which contains one or more <i>Presentations</i> . The <i>Information Consumer's</i> access to	

			the information is subject to a <i>Provision Agreement</i> , which sets out conditions of access.	
Information Provider	Exchange	An <i>Individual</i> or <i>Organization</i> that provides collected information.	An <i>Information Provider</i> possesses sets of information (that it has generated, collected, produced, bought or otherwise acquired) and is willing to supply that information (data or referential metadata) to the statistical office. The two parties use a <i>Provision Agreement</i> to agree the <i>Data Structure</i> and <i>Referential Metadata Structure</i> of the data to be exchanged via an <i>Exchange Channel</i> .	information supplier, data supplier
Information Request	Business	An outline of a need for new information required for a particular purpose.	An <i>Information Request</i> is a special case of <i>Statistical Need</i> that may come in an organized form, for example by specifying on which <i>Subject Field</i> the information is required. It may also be a more general request and require refinement by the statistical agency and formalized in a <i>Change Definition</i> .	
Information Resource	Structures	An abstract notion that is any organized collection of information.	There currently are only two concrete sub classes: <i>Data Resource</i> and <i>Referential Metadata Resource</i> . The <i>Information Resource</i> allows the model to be extended to other types of resource.	
Information Set	Structures	Organized collections of statistical content.	Statistical organizations collect, process, analyze and disseminate <i>Information Sets</i> , which contain data (<i>Data Sets</i>), referential metadata (<i>Referential Metadata Sets</i>), or potentially other types of statistical content, which could be included in addition types of <i>Information Set</i> .	
Instance Question	Exchange	The use of a <i>Question</i> in a particular <i>Questionnaire</i> .	The Instance Question is the use of a <i>Question</i> in a particular <i>Questionnaire Component</i> . This also includes the use of the <i>Question</i> in a <i>Question Block</i> , which is a particular type of <i>Questionnaire Component</i> .	
Instance Question Block	Exchange	The use of a <i>Question Block</i> in a particular <i>Questionnaire</i> .	The <i>Instance Question Block</i> is the use of a <i>Question Block</i> in a particular <i>Questionnaire Component</i> . This	

			also includes the use of a <i>Question Block</i> in another <i>Question Block</i> , as it is a particular type of <i>Questionnaire Component</i> .	
Instance Statement	Exchange	The use of a <i>Statement</i> in a particular <i>Questionnaire</i> .	The <i>Instance Statement</i> is the use of a <i>Statement</i> in a particular <i>Questionnaire Component</i> . This also includes the use of the <i>Statement</i> in a <i>Question Block</i> , which is a particular type of <i>Questionnaire Component</i> .	
Instance Variable	Concepts	The use of a <i>Represented Variable</i> within a <i>Data Set</i> . It may include information about the source of the data.	The <i>Instance Variable</i> is used to describe actual instances of data that have been collected. Here are 3 examples: 1) Gender: Dan Gillman has gender <m, male>, Arofan Gregory has gender<m, male>, etc. 2) Number of employees: Microsoft has 90,000 employees; IBM has 433,000 employees, etc. 3) Endowment: Johns Hopkins has endowment of <3, \$1,000,000 and above>, Yale has endowment of <3, \$1,000,000 and above>, etc.	
Level	Concepts	A <i>Statistical Classification</i> has a structure which is composed of one or several <i>Levels</i> . A <i>Level</i> often is associated with a <i>Concept</i> , which defines it. In a hierarchical classification the <i>Classification Items</i> of each <i>Level</i> but the highest are aggregated to the nearest higher <i>Level</i> . A linear classification has only one <i>Level</i> .	A <i>Statistical Classification</i> is a subtype of <i>Node Set</i> . The relationship between <i>Statistical Classification</i> and <i>Level</i> can also be extended to include the other <i>Node Set</i> types - <i>Code List</i> and <i>Category Set</i> .	
Logical Record	Structures	Describes a type of <i>Unit Data Record</i> for one <i>Unit Type</i> within a <i>Unit Data Set</i> .	Examples: household, person or dwelling record.	
Map	Concepts	A <i>Map</i> is an expression of the relation between a <i>Classification Item</i> in a source <i>Statistical Classification</i> and a	The use of <i>Correspondence Tables</i> and <i>Maps</i> can be extended to include all types of <i>Node</i> and <i>Node Set</i> . This means that a <i>Correspondence Table</i> could map	

		<p>corresponding <i>Classification Item</i> in the target <i>Statistical Classification</i>. The <i>Map</i> should specify whether the relationship between the two <i>Classification Items</i> is partial or complete. Depending on the relationship type of the <i>Correspondence Table</i>, there may be several <i>Maps</i> for a single source or target item.</p>	<p>between the items of <i>Statistical Classifications</i>, <i>Code Lists</i> or <i>Category Sets</i>.</p>	
Measure Component	Structures	<p>The role given to a <i>Represented Variable</i> in the context of a <i>Data Structure</i> to hold the observed/derived values for a particular <i>Unit</i> in an organized collection of data.</p>	<p>A <i>Measure Component</i> is a sub-type of <i>Data Structure Component</i>. For example age and height of a person in a <i>Unit Data Set</i> or number of citizens and number of households in a country in a <i>Data Set</i> for multiple countries (<i>Dimensional Data Set</i>).</p>	
Node	Concepts	<p>A combination of a <i>Category</i> and related attributes.</p>	<p>A <i>Node</i> is created as a <i>Category</i>, <i>Code</i> or <i>Classification Item</i> for the purpose of defining the situation in which the <i>Category</i> is being used.</p>	
Node Set	Concepts	<p>A set of <i>Nodes</i>.</p>	<p><i>Node Set</i> is a kind of <i>Concept System</i>. Here are 2 examples:</p> <p>1) <i>Sex Categories</i></p> <ul style="list-style-type: none"> • Male • Female • Other <p>2) <i>Sex Codes</i></p> <ul style="list-style-type: none"> • <m, male> • <f, female> • <o, other> 	

Organization	Base	A unique framework of authority within which a person or persons act, or are designated to act, towards some purpose.		
Output Specification	Exchange	Defines how <i>Information Sets</i> consumed by a <i>Product</i> are presented to <i>Information Consumers</i> .	The <i>Output Specification</i> specifies <i>Products</i> and defines the <i>Presentations</i> they contain. The <i>Output Specification</i> may be fully defined during the design process (such as in a paper publication or a predefined web report), or may be a combination of designed specification supplemented by user selections (such as in an online data query tool).	
Population	Concepts	The total membership of a defined class of people, objects or events.	A population is used to describe the total membership of a group of people, objects or events based on characteristics, e.g. time and geographic boundaries. Here are 3 examples – 1. US adult persons on 13 November 1956 2. US computer companies at the end of 2012 3. Universities in the US 1 January 2011.	
Presentation	Exchange	The way data and referential metadata are presented in a <i>Product</i> .	A <i>Product</i> has one or more <i>Presentations</i> , which present data and referential metadata from <i>Information Sets</i> . A <i>Presentation</i> is defined by an <i>Output Specification</i> . <i>Presentation</i> can be in different forms; e.g. tables, graphs, structured data files. Examples: <ul style="list-style-type: none"> • A table of data. Based on a <i>Data Set</i>, the related <i>Data Structure</i> is used to label the column and row headings for the table. The <i>Data Set</i> is used to populate the cells in the table. Reference metadata is used to populate footnotes and cell notes on the table. Confidentiality rules are applied to the <i>Data Set</i> to suppress any disclosive cells. • A data file based on a standard (e.g. SDMX). 	

			<ul style="list-style-type: none"> • A PDF document describing a <i>Classification</i>. • Any structural metadata object expressed in a standard format (e.g. DDI 3.1 XML). • A list of <i>Products</i> or services (e.g. a product catalogue or a web services description language (WSDL) file). • A web page containing <i>Classifications</i>, descriptions of <i>Variables</i>, etc. 	
Process Control	Business	A set of decision points which determine the flow between the <i>Process Steps</i> used to perform a <i>Business Process</i> .	The typical use of <i>Process Control</i> is to determine what happens next after a <i>Process Step</i> is executed. The possible paths, and the decision criteria, associated with a <i>Process Control</i> are specified as part of designing a production process, captured in a <i>Process Control Design</i> . There is typically a very close relationship between the design of a process and the design of a <i>Process Control</i> .	
Process Control Design	Business	The specification of the decision points required during the execution of a <i>Business Process</i> .	<p>The design of a <i>Process Control</i> typically takes place as part of the design of the process itself. This involves determining the conditional routing between the various sub-processes and services used by the executing process associated with the <i>Process Control</i> and specified by the <i>Process Control Design</i>.</p> <p>It is possible to define a <i>Process Control</i> where the next step in the <i>Process Step</i> that will be executed is a fixed value rather than a "choice" between two or more possibilities. Where such a design would be appropriate, this feature allows, for example, initiation of a step in the <i>Process Step</i> representing the GSBPM Process Phase (5) to always lead to initiation of GSBPM sub-process Integrate Data (5.1) as the next</p>	

			<p>step.</p> <p>This allows a process designer to divide a <i>Business Process</i> into logical steps (for example, where each step performs a specific <i>Business Function</i> through re-use of a <i>Business Service</i>) even if these process steps will always follow each other in the same order. In all cases, the <i>Process Control Design</i> defines and the <i>Process Control</i> manages the flow between <i>Process Steps</i>, even where the flow is "trivial". <i>Process Design</i> is left to focus entirely on the design of the process itself, not sequencing between steps.</p>	
Process Design	Business	The specification of how a <i>Process Step</i> will be performed. This includes specifying the types of <i>Process Inputs</i> required and the type of <i>Process Outputs</i> that will be produced.	<p>A <i>Process Design</i> is the design time specification of a <i>Process Step</i> that is performed as part of a run-time <i>Business Service</i>. A <i>Process Step</i> can be as big or small as the designer of a particular <i>Business Service</i> chooses. From a design perspective, one <i>Process Step</i> can contain "sub-steps", each of which is conceptualized as a (smaller) <i>Process Step</i> in its own right. Each of those "sub-steps" may contain "sub-steps" within them and so on indefinitely. It is a decision for the process designer to what extent to subdivide steps. At some level it will be appropriate to consider a <i>Process Step</i> to be a discrete task (unit of work) without warranting further subdivision. At that level the <i>Process Step</i> is designed to process particular <i>Process Inputs</i>, according to a particular <i>Process Method</i>, to produce particular <i>Process Outputs</i>. The flow between a <i>Process Step</i> and any sub steps is managed via <i>Process Control</i>.</p>	
Process Input	Business	Any instance of an information object which is supplied to a <i>Process Step Instance</i> at the time its execution is initiated.	<p><i>Process Input</i> might include information that is used as an input that will be transformed (e.g. a <i>Data Set</i>), information that is used to control specific parameters of the process (e.g. a <i>Rule</i>), and information that is</p>	

			used as reference to guide the process (e.g. a <i>Code List</i>).	
Process Input Specification	Business	A record of the types of inputs required for a <i>Process Design</i> .	<p>The <i>Process Input Specification</i> enumerates the <i>Process Inputs</i> required at the time a <i>Process Design</i> is executed. For example, if five different <i>Process Inputs</i> are required, the <i>Process Input Specification</i> will describe each of the five inputs. For each required <i>Process Input</i> the <i>Process Input Specification</i> will record the type of information object (based on GSIM) which will be used as the <i>Process Input</i> (example types might be a <i>Dimensional Data Set</i> or a <i>Classification</i>).</p> <p>The <i>Process Input</i> to be provided at the time of <i>Process Step</i> execution will then be a specific instance of the type of information object specified by the <i>Process Input Specification</i>. For example, if a <i>Process Input Specification</i> requires a <i>Dimensional Data Set</i> then the corresponding <i>Process Input</i> provided at the time of <i>Process Step</i> execution will be a particular <i>Dimensional Data Set</i>.</p>	
Process Method	Business	A specification of the technique which will be used to perform the unit of work.	<p>The technique specified by a <i>Process Method</i> is independent from any choice of technologies and/or other tools which will be used to apply that technique in a particular instance. The definition of the technique may, however, intrinsically require the application of specific <i>Rules</i> (for example, mathematical or logical formulas).</p> <p>A <i>Process Method</i> describes a particular method for performing a <i>Process Step</i>.</p>	
Process Output	Business	Any instance of an information object which is produced by a <i>Process Step</i> as a result of its execution.	<i>Process Outputs</i> have an attribute of Process Output Type, which has two possible values:	

			<ul style="list-style-type: none"> • <i>Transformed Output</i> is the result which provides the "reason for existence" of the <i>Process Step</i>. If that output were no longer required then there would be no need for the <i>Process Step</i> in its current form. Typically a <i>Transformed Output</i> is either a <i>Process Input</i> to a subsequent <i>Process Step</i> or it represents the final product from a statistical business process. • A <i>Process Metric</i> records information about the execution of a <i>Process Step</i>. For example, how long it took to complete execution of the <i>Process Step</i> and what percentage of records in the <i>Process Input</i> was updated by the <i>Process Step</i> to produce the <i>Transformed Output</i>. 	
Process Output Specification	Business	A record of the types of outputs required for a Process Design.	<p>The <i>Process Output Specification</i> enumerates the <i>Process Outputs</i> that are expected to be produced at the time a <i>Process Design</i> is executed. For example, if five different <i>Process Outputs</i> expected, the <i>Process Output Specification</i> will describe each of the five outputs. For each expected <i>Process Output</i> the <i>Process Output Specification</i> will record the type of information object (based on GSIM) which will be used as the <i>Process Output</i> (Example types might be a <i>Dimensional Data Set</i> or a <i>Classification</i>).</p> <p>The <i>Process Output</i> to be provided at the time of <i>Process Step</i> execution will then be a specific instance of the type of information object specified by the <i>Process Output Specification</i>. For example, if a <i>Process Output Specification</i> expects a <i>Dimensional Data Set</i> then the corresponding <i>Process Output</i></p>	

			provided at the time of <i>Process Step</i> execution will be a particular <i>Dimensional Data Set</i> .	
Process Pattern	Business	A nominated set of <i>Process Designs</i> , and associated <i>Process Control Designs</i> (flow), which have been highlighted for possible reuse.	In a particular <i>Business Process</i> , some <i>Process Steps</i> may be unique to that <i>Business Process</i> while others may be applicable to other <i>Business Processes</i> . A <i>Process Pattern</i> can be seen as a reusable template. It is a means to accelerate design processes and to achieve sharing and reuse of design patterns which have proved effective. Reuse of <i>Process Patterns</i> can indicate the possibility to reuse related <i>Business Services</i> . By deciding to reuse a <i>Process Pattern</i> , a designer is actually reusing the pattern of <i>Process Designs</i> and <i>Process Control Designs</i> associated with that <i>Process Pattern</i> . They will receive a new instance of the <i>Process Designs</i> and <i>Process Control Designs</i> . If they then tailor their "instance" of the <i>Process Designs</i> and <i>Process Control Designs</i> to better meet their needs they will not change the definition of the reusable <i>Process Pattern</i> .	
Process Step	Business	A <i>Process Step</i> is a work package that performs a <i>Business Process</i> . A <i>Process Step</i> implements the <i>Process Step Design</i> specified in order to produce the outputs for which the <i>Process Step</i> was designed.	Each <i>Process Step</i> is the use of a <i>Process Step Design</i> in a particular context (e.g. within a specific <i>Business Process</i>). At the time of execution a <i>Process Step Instance</i> specifies the actual instances of input objects (for example, specific <i>Data Sets</i> , specific <i>Variables</i>) to be supplied.	
Process Step Instance	Business	An executed step in a <i>Business Process</i> . A <i>Process Step Instance</i> specifies the actual inputs to and outputs from for an occurrence of a <i>Process Step</i> .	Each <i>Process Step</i> is the use of a <i>Process Step Design</i> in a particular context (e.g. within a specific <i>Business Process</i>). At the time of execution a <i>Process Step Instance</i> specifies the actual instances of input objects (for example, specific <i>Data Sets</i> , specific <i>Variables</i>) to be supplied. Each <i>Process Step Instance</i> may produce unique	

			<p>results even though the <i>Process Step</i> remains constant.</p> <p>Even when the inputs remain the same, metrics such as the elapsed time to complete execution of process step may vary from execution to execution. For this reason, each <i>Process Step Instance</i> details of inputs and outputs for that instance of implementing the <i>Process Step</i>.</p> <p>In this way it is possible to trace the flow of execution of a <i>Business Process</i> through all the <i>Process Steps</i> which were involved.</p>	
Product	Exchange	A package of content that can be disseminated as a whole.	<p>A <i>Product</i> is the only defined type of <i>Exchange Channel</i> for outgoing information. A <i>Product</i> packages <i>Presentations</i> of <i>Information Sets</i> for an <i>Information Consumer</i>. The <i>Product</i> and its <i>Presentations</i> are generated according to <i>Output Specifications</i>, which define how the information from the <i>Information Sets</i> it consumes are presented to the <i>Information Consumer</i>. The <i>Protocol</i> for a <i>Product</i> determines the mechanism by which the <i>Product</i> is disseminated (e.g. website, SDMX web service, paper publication).</p> <p>A <i>Provision Agreement</i> between the statistics office and the <i>Information Consumer</i> governs the use of a <i>Product</i> by the <i>Information Consumer</i>. The <i>Provision Agreement</i>, which may be explicitly or implicitly agreed, provides the legal or other basis by which the two parties agree to exchange data. In many cases, dissemination <i>Provision Agreements</i> are implicit in the terms of use published by the statistics office. For static <i>Products</i> (e.g. paper publications), specifications are predetermined. For dynamic products, aspects of specification could be determined</p>	

			by the <i>Information Consumer</i> at run time. Both cases result in <i>Output Specifications</i> specifying <i>Information Set</i> data or referential metadata that will be included in each <i>Presentation</i> within the <i>Product</i> .	
Protocol	Exchange	The mechanism for exchanging information through an <i>Exchange Channel</i> .	A Protocol specifies the mechanism (e.g. SDMX web service, data file exchange, web robot, face to face interview, mailed paper form) of exchanging information through an <i>Exchange Channel</i> .	
Provision Agreement	Exchange	The legal or other basis by which two parties agree to exchange data.	A <i>Provision Agreement</i> between the statistical organization and the <i>Information Provider</i> (collection) or the <i>Information Consumer</i> (dissemination) governs the use of <i>Exchange Channels</i> . The <i>Provision Agreement</i> , which may be explicitly or implicitly agreed, provides the legal or other basis by which the two parties agree to exchange data. The parties also use the <i>Provision Agreement</i> to agree the <i>Data Structure</i> and <i>Referential Metadata Structure</i> of the information to be exchanged.	
Question	Exchange	Describes the text used to elicit a response for the <i>Concept</i> to be measured.	A <i>Question</i> may be a single question used to obtain a response, or may be a multiple question, a construct which links multiple sub-questions, each with their own response. A <i>Question</i> also includes a relationship to the <i>Value Domain</i> to document the associated response criteria for the question. A single response question will have one <i>Value Domain</i> associated with it, while a 'multiple question' may have more than one <i>Value Domain</i> . A <i>Question</i> should be designed with re-use in mind, as it can be used in multiple <i>Questionnaires</i> .	Multiple Question
Question Block	Exchange	A set of <i>Questions</i> , <i>Statements</i> or instructions which are used together.	A <i>Question Block</i> should be designed for reuse, as it can be used in multiple <i>Questionnaires</i> . The <i>Question Block</i> is a type of <i>Questionnaire Component</i> . A statistical organization will often have a number of	Question Module

			<p><i>Question Blocks</i> which they reuse in a number of <i>Questionnaires</i>. Examples of <i>Question Blocks</i> include:</p> <ul style="list-style-type: none"> • Household <i>Question Block</i> • Income <i>Question Block</i> • Employment <i>Question Block</i> 	
Questionnaire	Exchange	A concrete and usable tool to elicit information from observation <i>units</i> .	<p>This is an example of a way statistical organizations collect information (an exchange channel). Each mode should be interpreted as a new <i>Questionnaire</i> derived from the <i>Questionnaire Specification</i>.</p> <p>The <i>Questionnaire</i> is a subtype of <i>Exchange Channel</i>, as it is a way in which data is obtained.</p>	
Questionnaire Component	Exchange	A record of the flow of a <i>Questionnaire Specification</i> and its use of <i>Questions</i> , <i>Question Blocks</i> and <i>Statements</i>	<p>Defines the structure of the <i>Questionnaire Specification</i>, as a combination of <i>Questions</i>, <i>Question Blocks</i> and <i>Statements</i>. It is the object which groups together all the components of a <i>Questionnaire</i>.</p> <p>A <i>Questionnaire Component</i> is recursive, in that it can refer to other <i>Questionnaire Components</i> and accompanying <i>Questionnaire Logic</i> objects at a lower level. It is only at the top level where the <i>Questionnaire Component</i> links to the <i>Questionnaire Specification</i>,</p>	Question Block
Questionnaire Logic	Exchange	Governs the sequence of <i>Questions</i> , <i>Question Blocks</i> and <i>Statements</i> based on factors such as the current location, the response to the previous questions etc., invoking navigation and validation rules to apply.		Routing
Questionnaire Specification	Exchange	The tool designed to elicit information from observation <i>Units</i> .	This represents the complete questionnaire design, with a relationship to the top level <i>Questionnaire Component</i> .	

			There may be many different <i>Questionnaire Specifications</i> , for the same surveys, or tailored to individual observation <i>Units</i> (respondents) so that there would be a different <i>Questionnaire Specification</i> for each respondent. The design would also differ depending upon the specific mode of collection the <i>Questionnaire</i> is designed for.	
Record Relationship	Structures	Describes relationships between <i>Logical Records</i> within a <i>Unit Data Structure</i> . It must have both a source <i>Logical Record</i> and a target <i>Logical Record</i> in order to define the relationship.	Example: Relationship between person and household <i>Logical Records</i> within a <i>Unit Data Set</i> .	
Referential Metadata Attribute	Structures	The role given to a <i>Represented Variable</i> to supply information in the context of a <i>Referential Metadata Structure</i> .		
Referential Metadata Content Item	Structures	The content describing a particular characteristic of a <i>Referential Metadata Subject</i> .	A <i>Referential Metadata Content Item</i> contains the actual content describing a particular characteristic of a <i>Referential Metadata Subject</i> .	
Referential Metadata Resource	Structures	An organized collection of stored information consisting of one or more <i>Referential Metadata Sets</i> .	<i>Referential Metadata Resources</i> are collections of structured information that may be used by a statistical activity to produce information. This information object is a specialization of an <i>Information Resource</i> .	
Referential Metadata Set	Structures	An organized collection of referential metadata for a given <i>Referential Metadata Subject</i> .	<i>Referential Metadata Sets</i> organize referential metadata. Each <i>Referential Metadata Set</i> uses a <i>Referential Metadata Structure</i> to define a structured list of <i>Referential Metadata Attributes</i> for a given <i>Referential Metadata Subject</i> .	
Referential Metadata Structure	Structures	Defines the structure of an organized collection of referential metadata (<i>Referential Metadata Set</i>).	A <i>Referential Metadata Structure</i> defines a structured list of <i>Referential Metadata Attributes</i> for a given <i>Referential Metadata Subject</i> . Examples of <i>Referential Metadata Attributes</i> are those that describe quality information and methodologies. Examples of subject are: objects like a <i>Questionnaire</i>	Metadata Structure Definition

			or a <i>Classification</i> , or collections of data like a <i>Data Set</i> , or any <i>Data Point</i> or set of <i>Data Points</i> created from a specific <i>Data Structure</i> .	
Referential Metadata Subject	Structures	Identifies the subject of an organized collection of referential metadata.	The <i>Referential Metadata Subject</i> identifies the subject of the metadata that can be reported using this Referential Metadata Structure. These subjects may be any GSIM object type, or any <i>Data Point</i> or set of <i>Data Points</i> created from a specific <i>Data Structure</i> . Examples: The GSIM object type may be Product for which there is a list specified in a <i>Value Domain</i> . The <i>Value Domain</i> specifies the list of actual Products for which reference metadata can be reported or authored using this <i>Referential Metadata Structure</i> .	
Referential Metadata Subject Item	Structures	Identifies the actual subject for which referential metadata is reported.	Examples are an actual <i>Product</i> such as Balance of Payments and International Investment Position, Australia, June 2013, or a collection of <i>Data Points</i> such as the <i>Data Points</i> for a single region within a <i>Data Set</i> covering all regions for a country.	
Represented Variable	Concepts	A combination of a characteristic of a population to be measured and how that measure will be represented.	Example: The pair (Number of Employees, Integer), where "Number of Employees" is the characteristic of the population (<i>Variable</i>) and "Integer" is how that measure will be represented (<i>Value Domain</i>).	
Rule	Business	A specific mathematical or logical expression which can be evaluated to determine specific behavior.	<i>Rules</i> are of several types: they may be derived from methods to determine the control flow of a process when it is being designed and executed; they may be used as the input parameters of processes (e.g., imputation rules, edit rules); and they may be used to drive the logical flow of a questionnaire. There are many forms of <i>Rules</i> and their purpose, character and expression can vary greatly.	

Scraping Process Map	Exchange	Maps a web scraping process to a specific website.	<i>Scraping Process Map</i> is an essential element of the <i>Web Scraper Channel</i> . The process being mapped can be a <i>Business Service</i> or a <i>Process Step</i> .	
Sign	Concepts	Something that suggests the presence or existence of a fact, condition, or quality.	It is a perceivable object. This object is used to denote a <i>Concept</i> as a <i>Designation</i> .	
Statement	Exchange	A report of facts in a <i>Questionnaire</i>	<i>Statements</i> are often included to provide further explanation to respondents. Example: "The following questions are about your health". The object is also used to represent completion instructions for the interviewer or respondent. <i>Statement</i> should be designed with re-use in mind as it can be used in numerous <i>Questionnaires</i> .	Interviewer Instruction Instruction
Statistical Classification	Concepts	A <i>Statistical Classification</i> is a set of <i>Categories</i> 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 <i>Categories</i> at each <i>Level</i> of the classification structure must be mutually exclusive and jointly exhaustive of all objects/units in the population of interest.	The <i>Categories</i> are defined with reference to one or more characteristics of a particular population of units of observation. A <i>Statistical Classification</i> may have a flat, linear structure or may be hierarchically structured, such that all <i>Categories</i> at lower <i>Levels</i> are sub- <i>Categories</i> of <i>Categories</i> at the next <i>Level</i> up. <i>Categories</i> in <i>Statistical Classifications</i> are represented in the information model as <i>Classification Items</i> .	
Statistical Need	Business	A requirement, request or other notification that will be considered by an organization. A <i>Statistical Need</i> does not necessarily have structure or format - it is a 'raw' need as received by the organization. A <i>Statistical Need</i> may be of a variety of types including <i>Environmental Change</i> or <i>Information Request</i> .	The <i>Statistical Need</i> is a proposed or imposed requirement, request or other notification as it has been received by an organization. A <i>Statistical Need</i> is a raw expression of a requirement, and is not necessarily well-defined. A related object - <i>Change Definition</i> - is created when a <i>Statistical Need</i> is analyzed by an organization. <i>Change Definition</i> expresses the raw need in well-defined, structured terms. Once a <i>Statistical Need</i> has been received, the first step is to do the conceptual work to establish what it is	

			<p>we are trying to measure. The final output of this conceptual work is the <i>Change Definition</i>.</p> <p>In some cases, the <i>Statistical Need</i> can result from the <i>Assessment</i> of the quality, efficiency, etc. of an existing process.</p>	
Statistical Program	Business	<p>A set of activities, which may be repeated, that describes the purpose and context of a set of <i>Business Process</i> within the context of the relevant <i>Statistical Program Cycles</i>.</p>	<p>The <i>Statistical Program</i> is one of a family of objects that provide the environmental context in which activities to produce statistics within a statistical organization are conducted. <i>Statistical Program</i> is the top level object that describes the purpose and objectives of a set of activities. <i>Statistical Program</i> will usually correspond to an ongoing activity such as a survey or output series. Some examples of <i>Statistical Program</i> are:</p> <ul style="list-style-type: none"> • Labour Force Survey - Multipurpose Household Survey - National Accounts - Demography - Overseas Arrivals and Departures <p>Related to the <i>Statistical Program</i> object there are <i>Statistical Program Design</i> and <i>Statistical Program Cycle</i> objects that hold the detailed information about the design and conduct of the <i>Business Process</i>.</p> <p>In the case of the traditional approach, an organization has received a <i>Statistical Need</i> and produced a <i>Change Definition</i> and an approved <i>Business Case</i>. The <i>Business Case</i> will specify either a change to the design or methodology of an existing <i>Statistical Program</i>, which will result in a new <i>Statistical Program Design</i>; or a change to one or more existing <i>Statistical Programs</i> (for example, to add an additional objective to the <i>Statistical Program</i>); or result in a new <i>Statistical Program</i> being created.</p>	

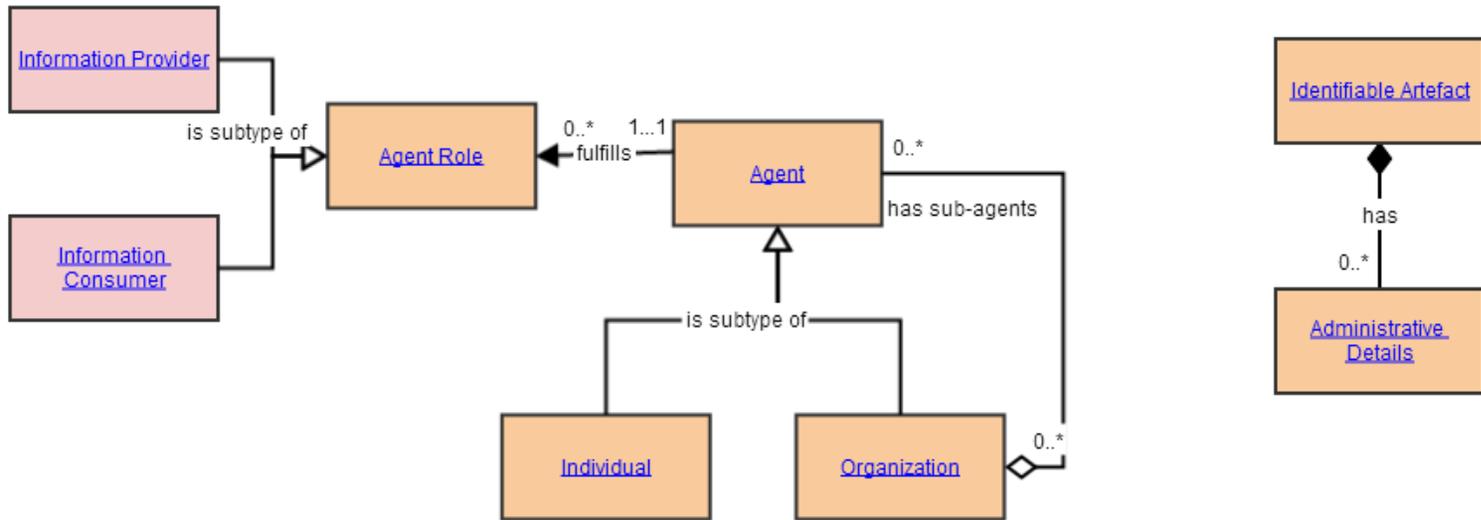
			This does not include statistical support functions such as metadata management, data management (and other overarching GSBPM processes) and design functions. These activities are conducted as part of <i>Statistical Support Programs</i> .	
Statistical Program Cycle	Business	A set of activities to investigate characteristics of a given <i>Population</i> for a particular reference period.	A <i>Statistical Program Cycle</i> documents the execution of an iteration of a <i>Statistical Program</i> according to the associated <i>Statistical Program Design</i> for a certain reference period. It identifies the activities that are undertaken as a part of the cycle and the specific resources required and processes used and description of relevant methodological information used in this cycle defined by the <i>Statistical Program Design</i> .	
Statistical Program Design	Business	The specification of the resources required, processes used and description of relevant methodological information about the set of activities undertaken to investigate characteristics of a given <i>Population</i> .	The <i>Statistical Program Design</i> is an objects that provide the operational context in which a set of <i>Business Processes</i> is conducted. A simple example is where a <i>Statistical Program</i> relates to a single survey, for example, the Labour Force Survey. The <i>Statistical Program</i> will have a series of <i>Statistical Program Design</i> objects that describe the methodology and design used throughout the life of the survey. When a methodological change is made to the survey, a new <i>Statistical Program Design</i> is created to record the details of the new design.	
Statistical Support Program	Business	A program which is not related to the post-design cyclic production of statistical products, but is necessary to support cyclical production.	This type of program will include such functions as metadata management, data management, methodological research, and design functions. These programs correspond to the horizontal functions shown in the GSBPM, as well as programs to create new or change existing <i>Statistical Programs</i> .	

Subject Field	Concepts	One or more <i>Concept Systems</i> used for the grouping of <i>Concepts</i> and <i>Categories</i> for the production of statistics.	A <i>Subject Field</i> is a field of special knowledge under which a set of <i>Concepts</i> and their <i>Designations</i> is used. For example, labour market, environmental expenditure, tourism, etc.	subject area, theme
Unit	Concepts	The object of interest in a <i>Business Process</i>	Here are 3 examples - 1. Individual US person (i.e., Arofan Gregory, Dan Gillman, Barack Obama, etc.) 2. Individual US computer companies (i.e., Microsoft, Apple, IBM, etc.) 3. Individual US universities (i.e., Johns Hopkins, University of Maryland, Yale, etc.)	
Unit Data Point	Structures	A placeholder (or cell) for the value of an <i>Instance Variable</i> with respect to a <i>Unit</i> .	This placeholder may point to multiple values representing different versions of the data. Values are only distinguished on the basis of quality, date/time of measurement or calculation, status, etc. This is handled through the mechanisms provided by the <i>Datum</i> information object.	cell
Unit Data Record	Structures	Contains the specific values (as a collection of <i>Unit Data Points</i>) related to a given <i>Unit</i> as defined in a <i>Logical Record</i> .	For example (1212123, 48, American, United Kingdom) specifies the age (48) in years on the 1st of January 2012 in years, the current citizenship (American), and the country of birth (United Kingdom) for a person with social security number 1212123.	
Unit Data Set	Structures	A collection of data that conforms to a known structure and describes aspects of one or more <i>Units</i> .	Example: A synthetic unit record file is a collection of artificially constructed <i>Unit Data Records</i> , combined in a file to create a <i>Unit Data Set</i> .	micro data, unit data, synthetic unit record file
Unit Data Structure	Structures	Describes the structure of a <i>Unit Data Set</i> .	For example (social security number, country of residence, age, citizenship, country of birth) where the social security number and the country of residence are the identifying components (<i>Identifier Component</i>) and the others are measured variables obtained directly or indirectly from the person (<i>Unit</i>) and are <i>Measure Components</i> of the <i>Logical Record</i> .	file description, dataset description

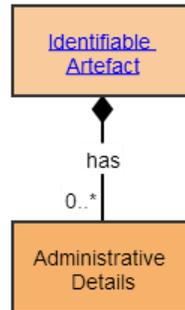
Unit Type	Concepts	A <i>Unit Type</i> is a class of objects of interest	A <i>Unit Type</i> is used to describe a class or group of <i>Units</i> based on a single characteristic, but with no specification of time and geography. For example, the <i>Unit Type</i> of “Person” groups together a set of <i>Units</i> based on the characteristic that they are ‘Persons’. It concerns not only <i>Unit Types</i> used in dissemination, but anywhere in the statistical process. E.g. using administrative data might involve the use of a fiscal unit.	Object class (ISO 11179)
Value Domain	Concepts	The permitted range of values for a characteristic of a variable	The values can be described by enumeration or by an expression	
Variable	Concepts	The use of a <i>Concept</i> as a characteristic of a <i>Population</i> intended to be measured	The Variable combines the meaning of a Concept with a Unit Type, to define the characteristic that is to be measured. Here are 3 examples - 1. Sex of person 2. Number of employees 3. Value of endowment	
Web Scraper Channel	Exchange	A concrete and usable tool to gather information from the Internet.	This is an example of a way statistical organizations collect information (an <i>Exchange Channel</i>). The <i>Web Scraper Channel</i> contains <i>Scraping Process Maps</i> , which map the channel to each website targeted for scraping.	

Annex C: UML Diagrams

Base Group



Administrative Details

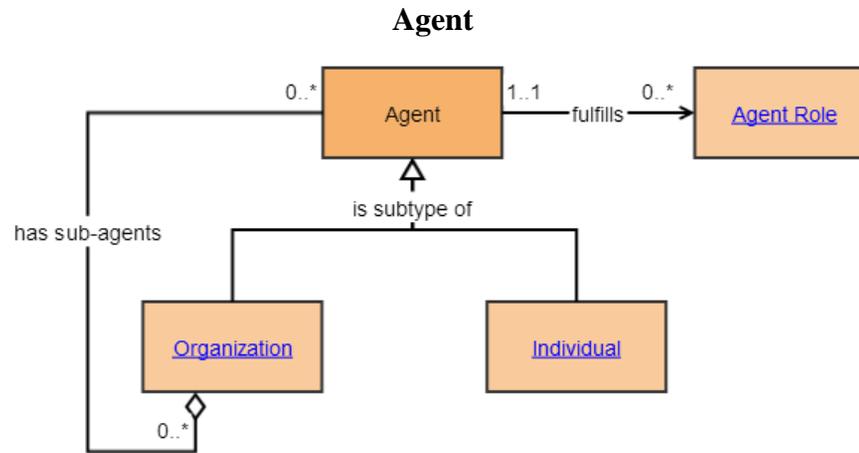


Definition

Object	Group	Definition	Explanatory Text	Synonyms
Administrative Details	Base	A placeholder for extensions to the model based on an organization's administrative needs.	The <i>Administrative Details</i> object is designed to act as a 'placeholder' to allow for future extensions to the existing model. It allows for further information to be added about the administrative details required to maintain the other objects outlined by GSIM.	

Attributes

To be defined on an 'as needs' basis.



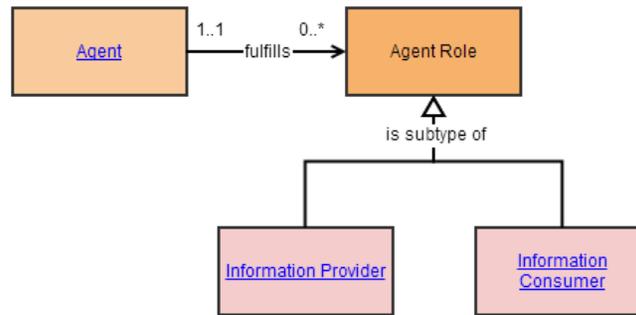
Definition

Object	Group	Definition	Explanatory Text	Synonyms
Agent	Base	An actor that performs a role in relation to the statistical <i>Business Process</i> .	An <i>Agent</i> may be either an <i>Organization</i> or an <i>Individual</i> . An <i>Organization</i> may be an entire organization or entities within a larger organization, such as departments or divisions. An <i>Organization</i> may have sub <i>Agents</i> , which may be either other <i>Organizations</i> within the parent <i>Organization</i> or <i>Individuals</i> that belong to that <i>Organization</i> .	

Attributes

Name	Description	Cardinality	Value Type
Name		0..1	Text
Description		0..1	Text

Agent Role



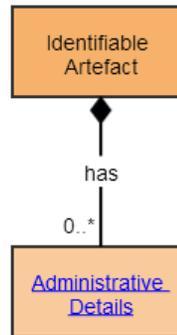
Definition

Object	Group	Definition	Explanatory Text	Synonyms
Agent Role	Base	The function or activities of an <i>Agent</i> , in regard to their involvement in the statistical <i>Business Process</i> .	An <i>Agent Role</i> may apply to either type of <i>Agent</i> - an <i>Organization</i> or <i>Individual</i> . A common example would be to identify which individuals or departments within an organization provide administrative data.	

Attributes

Name	Description	Cardinality	Value Type
Name		0..1	Text
Description		0..1	Text

Identifiable Artefact



Note: there are many relationships between *Identifiable Artefact* and other information objects; these are not shown in the above diagram.

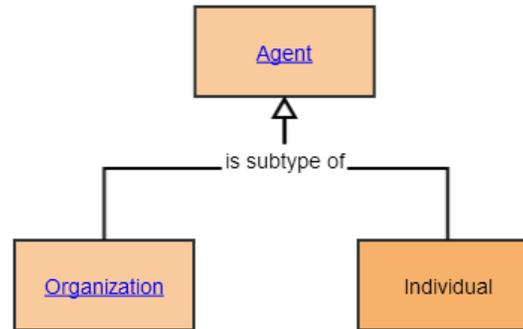
Definition

Object	Group	Definition	Explanatory Text	Synonyms
Identifiable Artefact	Base	An abstract class that comprises the basic attributes and associations needed for identification, naming and other documentation.	An instance of any GSIM information object is an <i>Identifiable Artefact</i> .	

Attribute

Name	Description	Cardinality	Value Type
id		1..1	string

Individual



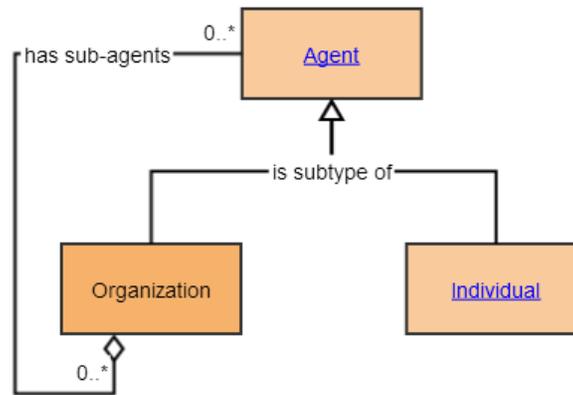
Definition

Object	Group	Definition	Explanatory Text	Synonyms
Individual	Base	A person who acts, or is designated to act towards a specific purpose.		

Attributes

Name	Description	Cardinality	Value Type
Name		0..1	Text
Description		0..1	Text

Organization



Definition

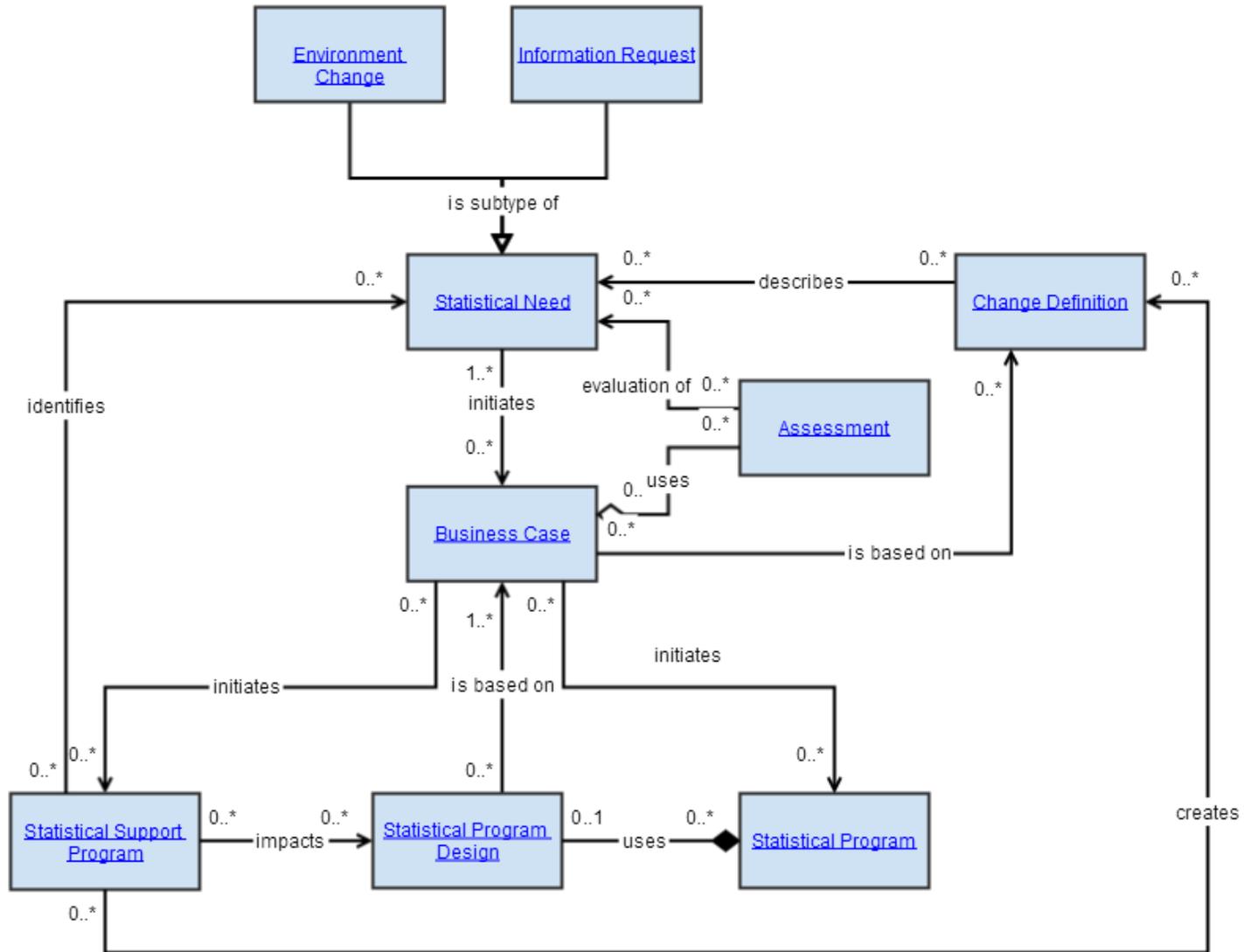
Object	Group	Definition	Explanatory Text	Synonyms
Organization	Base	A unique framework of authority within which a person or persons act, or are designated to act, towards some purpose.		

Attributes

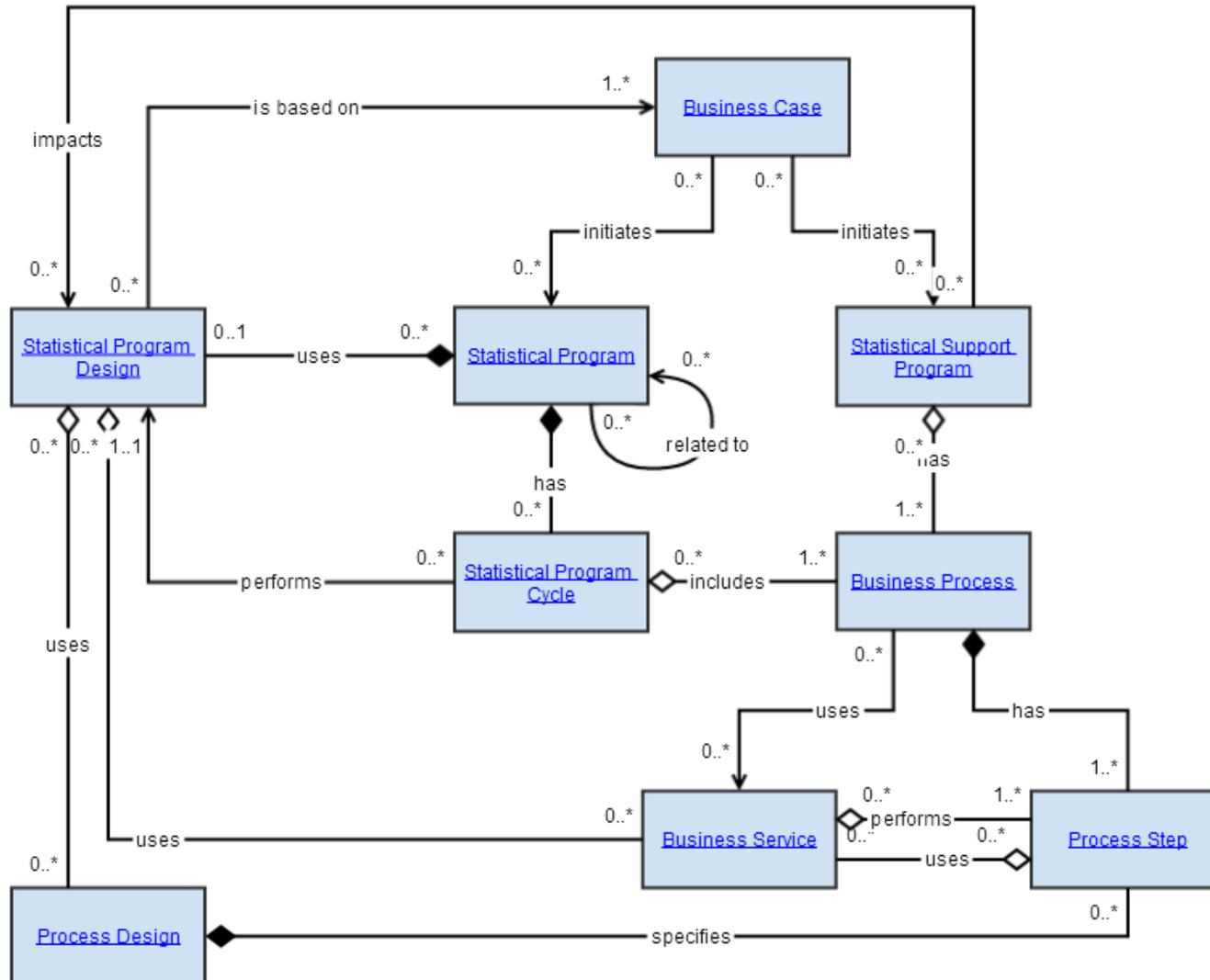
Name	Description	Cardinality	Value Type
Name		0..1	Text
Description		0..1	Text

Business Group

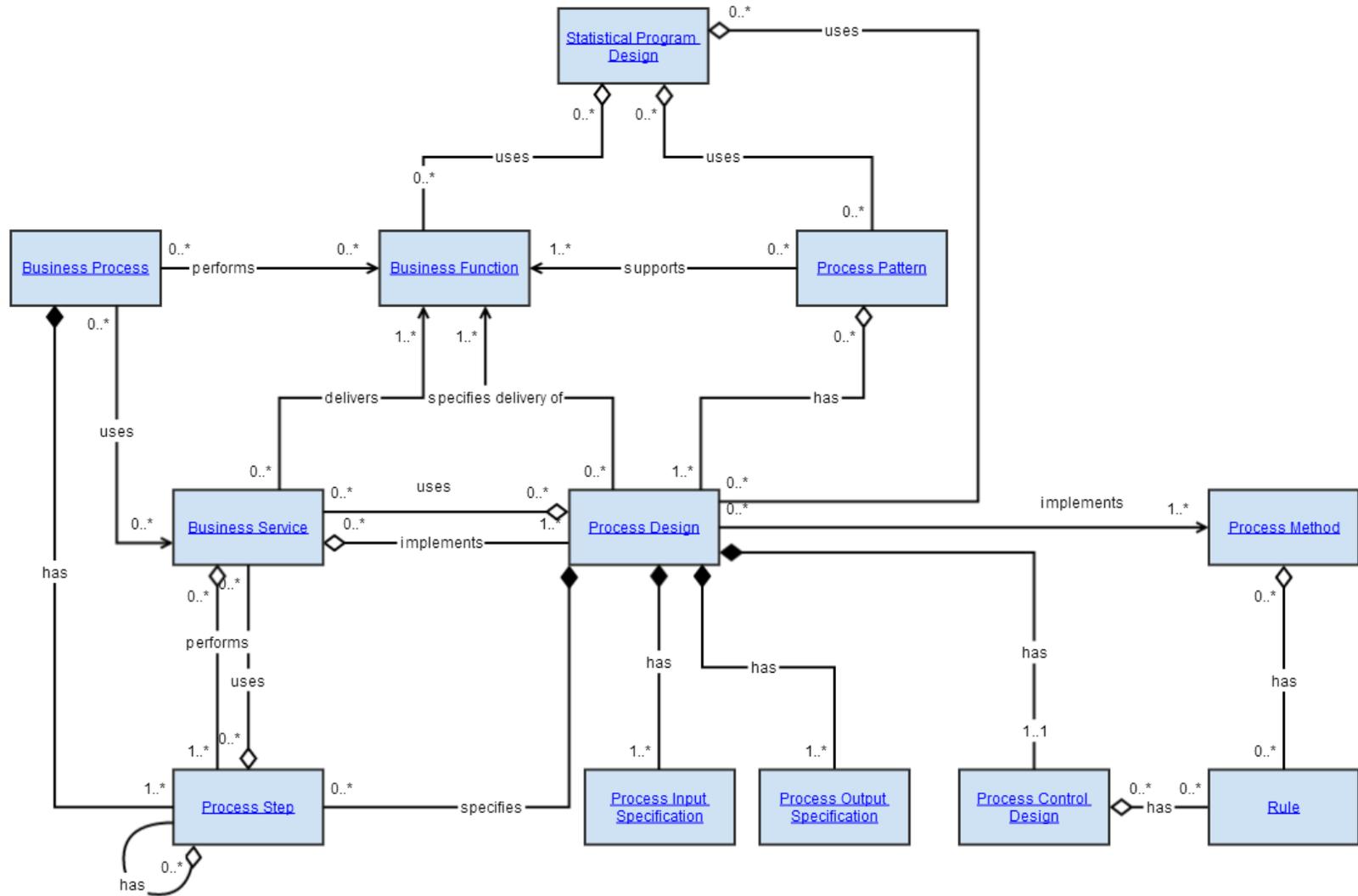
Identifying and Evaluating Statistical Needs



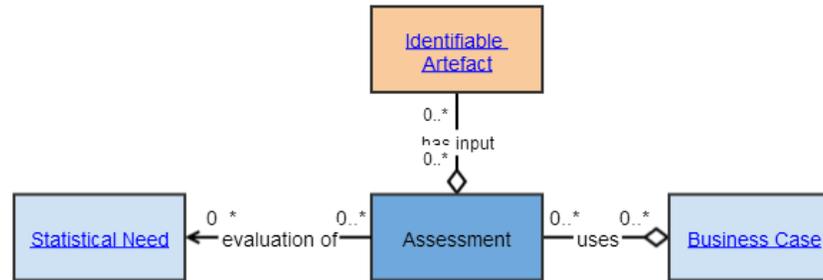
Designing and Managing Statistical Programs



Designing Processes



Assessment



Definition

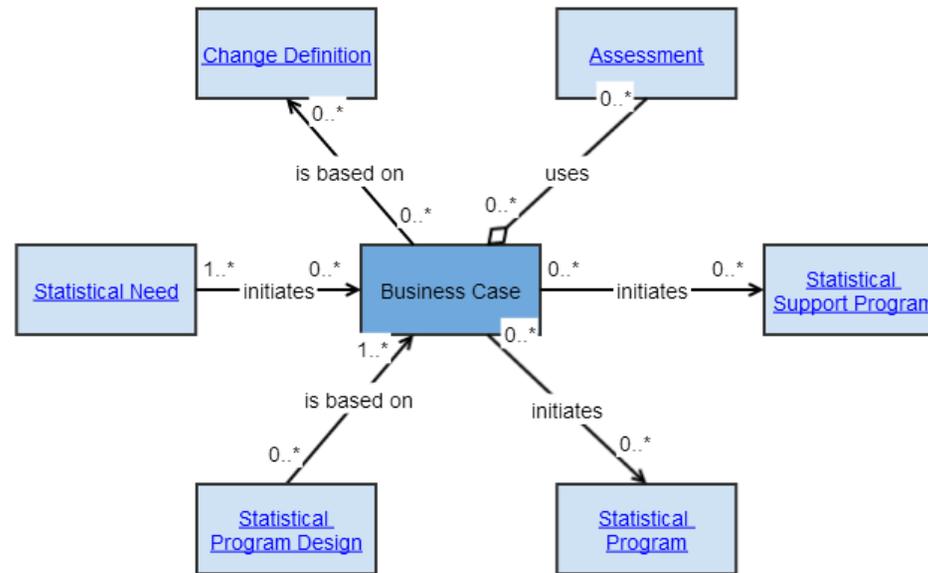
Object	Group	Definition	Explanatory Text	Synonyms
Assessment	Business	The result of the analysis of the quality and effectiveness of any activity undertaken by a statistical organization and recommendations on how these can be improved.	An <i>Assessment</i> can be of a variety of types. One example may include a gap analysis, where a current state is determined along with what is needed to reach its target state. Alternately, an <i>Assessment</i> may compare current processes against a set of requirements, for example a new <i>Statistical Need</i> or change in the operating environment. An <i>Assessment</i> can use various information objects as inputs, whether they are the main objects that the <i>Assessment</i> is about or auxiliary information objects that help accomplish the <i>Assessment</i> .	

Attributes

Name	Description	Cardinality	Value Type
Name		1..1	Text
Description		1..1	Text
Date assessed		1..*	Date
Subject Matter Domain		0..*	Text
Issues		0..*	Text

Results		0..*	Text
Recommendations		0..*	Text

Business Case



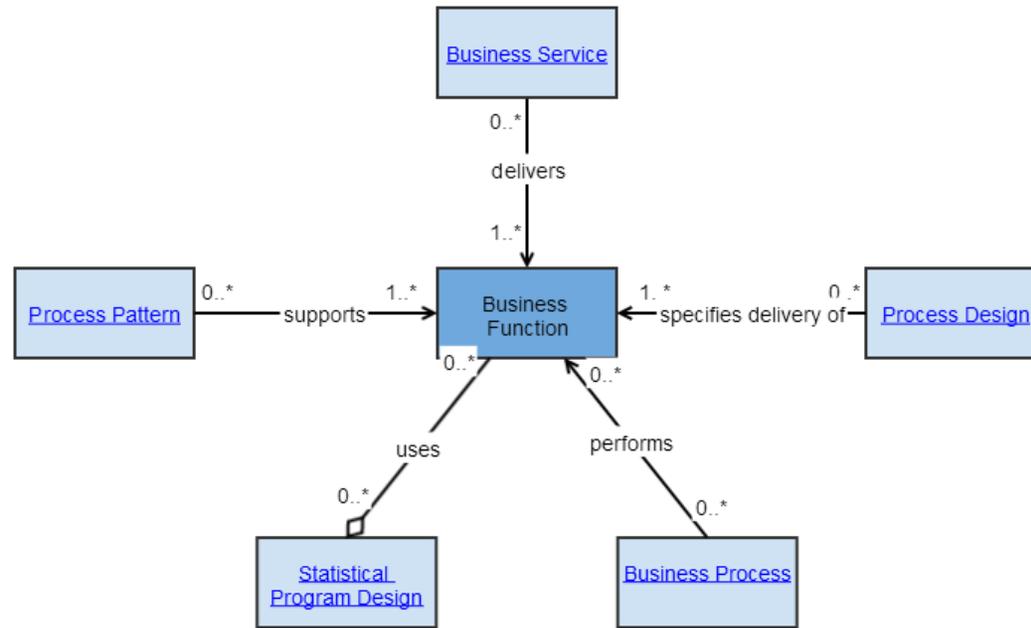
Definition

Object	Group	Definition	Explanatory Text	Synonyms
Business Case	Business	A proposal for a body of work that will deliver outputs designed to achieve outcomes. A <i>Business Case</i> will provide the reasoning for undertaking a <i>Statistical Support Program</i> to initiate a new <i>Statistical Program Design</i> for an existing <i>Statistical Program</i> , or an entirely new <i>Statistical Program</i> , as well as the details of the change proposed.	A <i>Business Case</i> is produced as a result of a detailed consideration of a <i>Change Definition</i> . It sets out a plan for how the change described by the <i>Change Definition</i> can be achieved. A <i>Business Case</i> usually comprises various evaluations. The <i>Business Case</i> will specify the stakeholders that are impacted by the <i>Statistical Need</i> or by the different solutions that are required to implement it.	

Attributes

Name	Description	Cardinality	Value Type
Name		1..1	Text
Description		1..1	Text
Date initiated		0..1	Date
Date approved		0..1	Date
Date implementation commenced		0..1	Date
Type		1..*	e.g. new program, permanent (indefinite) change to existing program, temporary change to existing program, cease program
Outcomes (objectives)		1..*	Text
Outputs (deliverables)		1..*	Text

Business Function



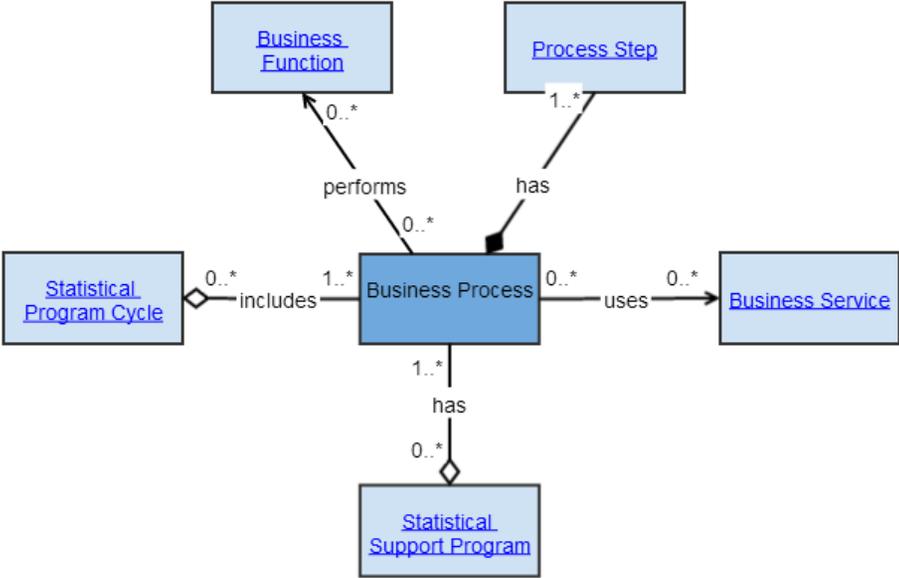
Definition

Object	Group	Definition	Explanatory Text	Synonyms
Business Function	Business	Something an enterprise does, or needs to do, in order to achieve its objectives.	<p>A <i>Business Function</i> delivers added value from a business point of view. It is delivered by bringing together people, processes and technology (resources), for a specific business purpose.</p> <p><i>Business Functions</i> answer in a generic sense "What business purpose does this <i>Business Service</i> or <i>Process Step</i> serve?" Through identifying the <i>Business Function</i> associated with each <i>Business Service</i> or <i>Process Step</i> it increases the documentation of the use of the associated <i>Business Services</i> and <i>Process Steps</i>, to enable future reuse.</p> <p>A <i>Business Function</i> may be defined directly with descriptive text and/or through reference to an existing catalogue of <i>Business Functions</i>. The phases and sub processes defined within GSBPM can be used as an internationally agreed basis for cataloguing high level <i>Business Functions</i>. A catalogue might also include <i>Business Functions</i> defined at a lower level than "sub process". For example, "Identify and address outliers" might be catalogued as a lower level <i>Business Function</i> with the "Review, validate and edit" function (5.3) defined within GSBPM.</p>	

Attributes

Name	Description	Cardinality	Value Type
Name		1..1	Text
Description		1..1	Text

Business Process



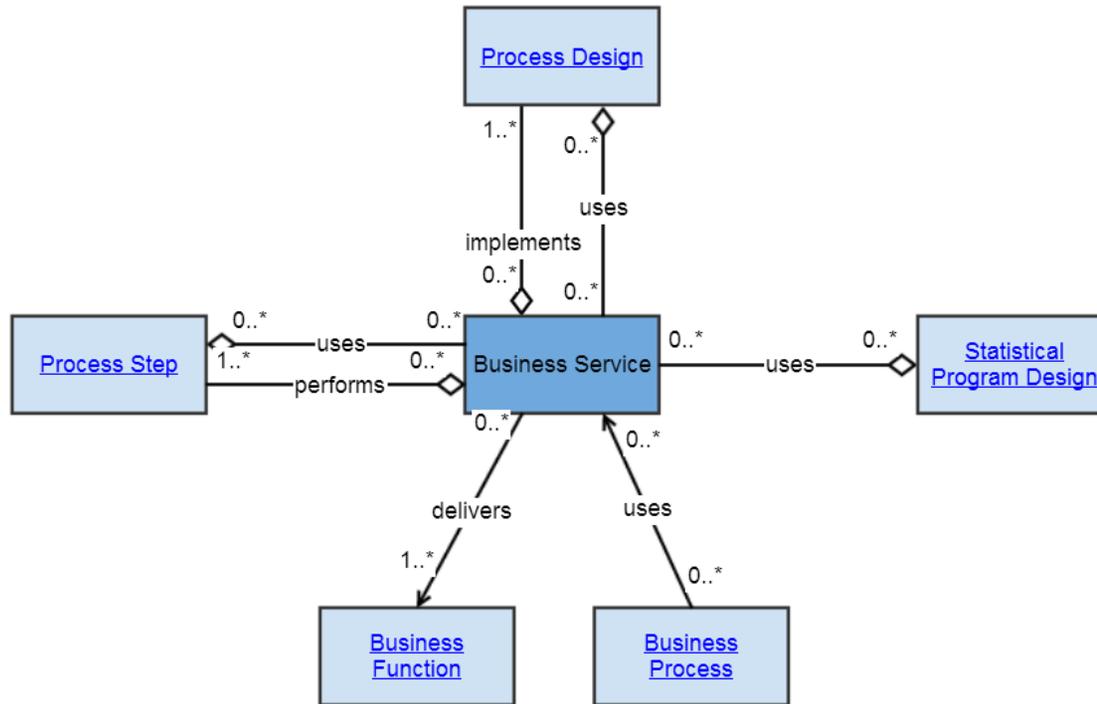
Definition

Object	Group	Definition	Explanatory Text	Synonyms
Business Process	Business	The set of <i>Process Steps</i> to perform one of more <i>Business Functions</i> to deliver a <i>Statistical Program Cycle</i> or <i>Statistical Support Program</i> .	For example, a particular <i>Statistical Program Cycle</i> might include several data collection activities, the corresponding editing activities for each collection and the production and dissemination of final outputs. Each of these may be considered separate <i>Business Processes</i> for the <i>Statistical Program Cycle</i> .	

Attributes

Name	Description	Cardinality	Value Type
Name		1..1	Text
Description		1..1	Text
Date initiated	First date of validity	0..1	Date
Date ended	Last date of validity	0..1	Date
Status		1..1	Extensible redefined list (e.g. New Proposal, New-Under Development, Current, Completed, Cancelled, Transferred to another Organization)

Business Service



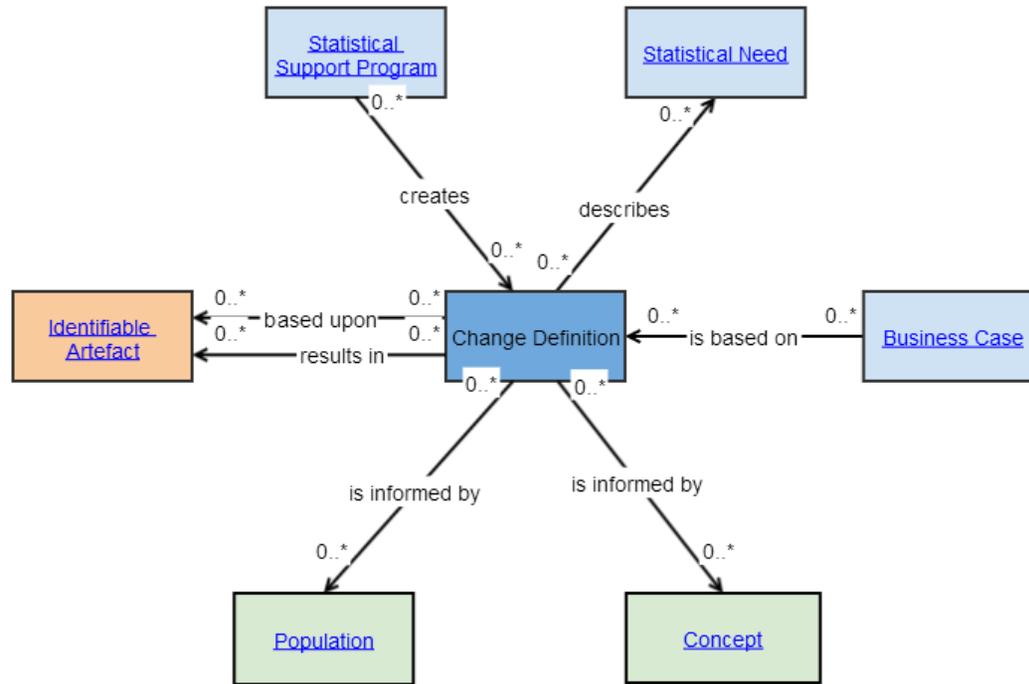
Definition

Object	Group	Definition	Explanatory Text	Synonyms
Business Service	Business	A means of performing a <i>Business Function</i> (an ability that an organization possesses, typically expressed in general and high level terms and requiring a combination of organization, people, processes and technology to achieve).	<p>A <i>Business Service</i> may provide one means of accessing a particular <i>Business Function</i>. The operation of a <i>Business Service</i> will perform one or more <i>Business Processes</i>.</p> <p>The explicitly defined interface of a <i>Business Service</i> can be seen as representing a "service contract". If particular inputs are provided then the service will deliver particular outputs in compliance within specific parameters (for example, within a particular period of time).</p> <p>Note: The interface of a <i>Business Service</i> is not necessarily IT based. For example, a typical postal service will have a number of service interfaces:</p> <ul style="list-style-type: none"> - Public letter box for posting letters - Counter at post office for interacting with postal workers 	

Attributes

Name	Description	Cardinality	Value Type
Name		1..1	Text
Description		1..1	Text
Service Interface	Specifies how to communicate with the service.	0..*	Text
Location	Specifies where the service can be accessed.	0..1	Text

Change Definition



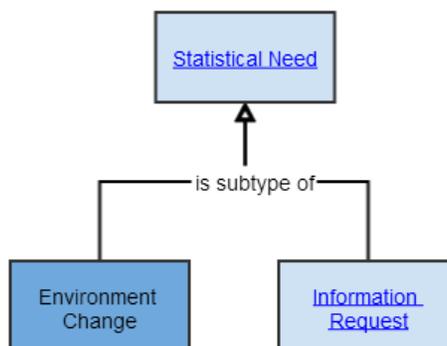
Definition

Object	Group	Definition	Explanatory Text	Synonyms
Change Definition	Business	A structured, well-defined specification for a proposed change.	<p>A related object - the <i>Statistical Need</i> - is a change expression as it has been received by an organization. A <i>Statistical Need</i> is a raw expression of a proposed change, and is not necessarily well-defined. A <i>Change Definition</i> is created when a <i>Statistical Need</i> is analyzed by an organization, and expresses the raw need in well-defined, structured terms.</p> <p>A <i>Change Definition</i> does not assess the feasibility of the change or propose solutions to deliver the change - this role is satisfied by the <i>Business Case</i> object. The precise structure or organization of a <i>Change Definition</i> can be further specified by rules or standards local to a given organization. It also includes the specific <i>Concepts</i> to be measured and the <i>Population</i> that is under consideration.</p> <p>Once a <i>Statistical Need</i> has been received, the first step is to do the conceptual work to establish what it is we are trying to measure. The final output of this conceptual work is the <i>Change Definition</i>.</p> <p>The next step is to assess how we are going to make the measurements - to design a solution and put forward a proposal for a body of work that will deliver on the requirements of the original <i>Statistical Need</i></p>	

Attributes

Name	Description	Cardinality	Value Type
Name	A human-readable identifier for the object	0..1	Text
Description	A human-readable description of the object	0..1	Text

Environment Change

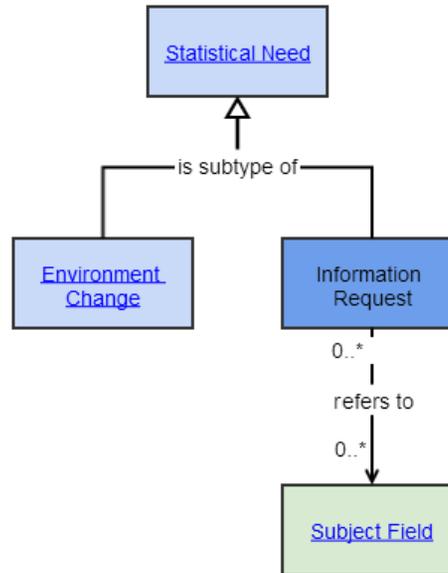


Definition

Object	Group	Definition	Explanatory Text	Synonyms
Environment Change	Business	A requirement for change that originates from a change in the operating environment of the statistical organization.	An <i>Environment Change</i> reflects change in the context in which a statistical organization operates. <i>Environment Changes</i> can be of different origins and also take different forms. They can result from a precise event (budget cut, new legislation enforced) or from a progressive process (technical or methodological progress, application or tool obsolescence). Other examples of <i>Environment Changes</i> include the availability of a new <i>Information Resource</i> , the opportunity for new collaboration between organizations, etc.	

Attributes Name	Description	Cardinality	Value Type
Change origin		1..1	Text
Legal changes		0..*	Text
Method changes		0..*	Text
Software changes		0..*	Text
Other changes		0..*	Text

Information Request



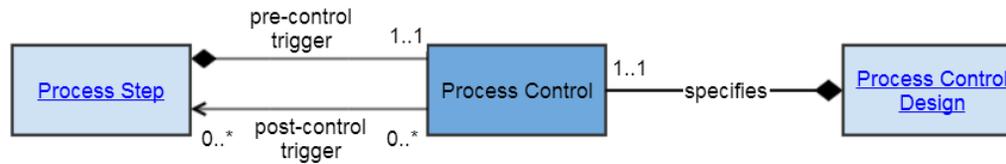
Definition

Object	Group	Definition	Explanatory Text	Synonyms
Information Request	Business	An outline of a need for new information required for a particular purpose.	An <i>Information Request</i> is a special case of <i>Statistical Need</i> that may come in an organized form, for example by specifying on which <i>Subject Field</i> the information is required. It may also be a more general request and require refinement by the statistical agency and formalized in a <i>Change Definition</i> .	

Attributes

Name	Description	Cardinality	Value Type
Information requestor		0..1	Text
Contact details of requestor		0..1	Text
Coverage of information required		1..1	Text
Date information required		0..1	Date

Process Control



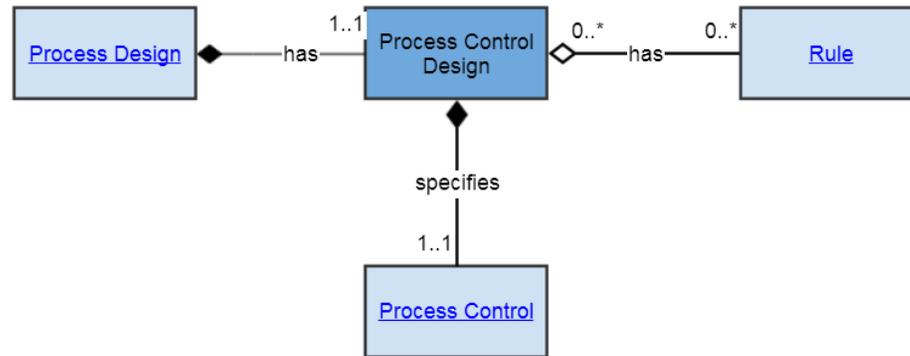
Definition

Object	Group	Definition	Explanatory Text	Synonyms
Process Control	Business	A set of decision points which determine the flow between the <i>Process Steps</i> used to perform a <i>Business Process</i> .	The typical use of <i>Process Control</i> is to determine what happens next after a <i>Process Step</i> is executed. The possible paths, and the decision criteria, associated with a <i>Process Control</i> are specified as part of designing a production process, captured in a <i>Process Control Design</i> . There is typically a very close relationship between the design of a process and the design of a <i>Process Control</i> .	

Attributes

Name	Description	Cardinality	Value Type
Name		1..1	Text
Description		1..1	Text
Start Event	The event which triggered the control.	0..1	Text
Status	Success or error, typically using a coded value.	0..1	Text

Process Control Design



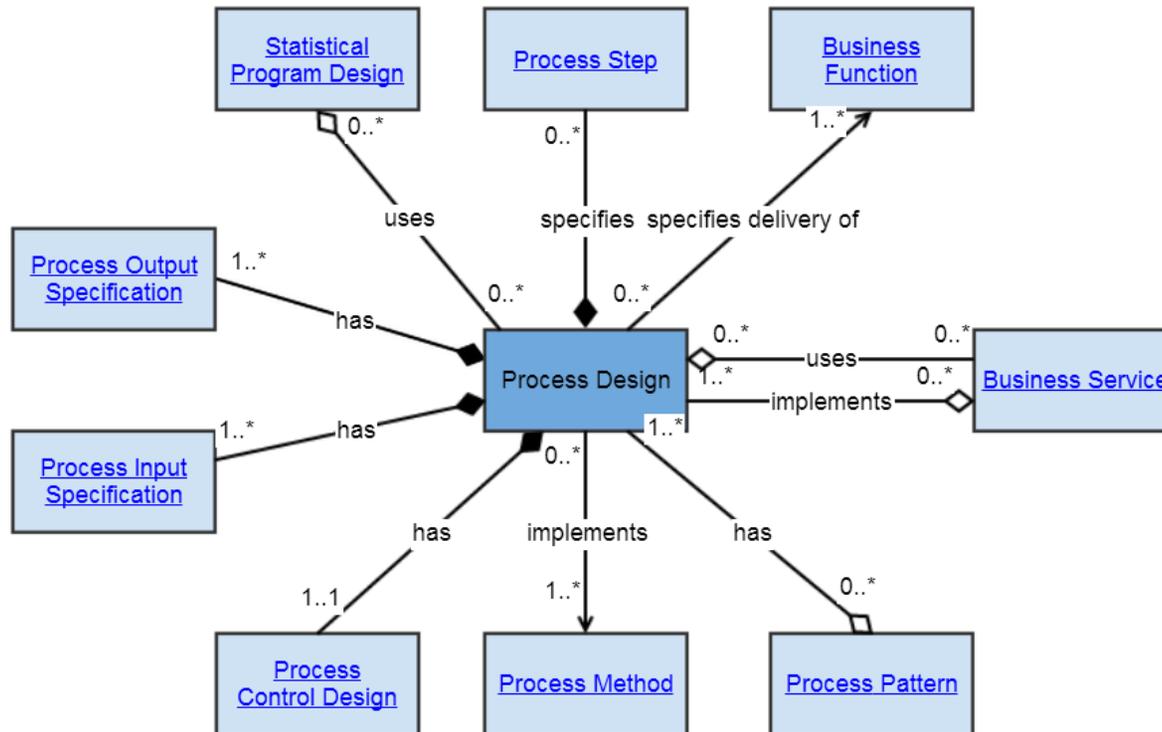
Definition

Object	Group	Definition	Explanatory Text	Synonyms
Process Control Design	Business	The specification of the decision points required during the execution of a <i>Business Process</i> .	<p>The design of a <i>Process Control</i> typically takes place as part of the design of the process itself. This involves determining the conditional routing between the various sub-processes and services used by the executing process associated with the <i>Process Control</i> and specified by the <i>Process Control Design</i>.</p> <p>It is possible to define a <i>Process Control</i> where the next step in the <i>Process Step</i> that will be executed is a fixed value rather than a "choice" between two or more possibilities. Where such a design would be appropriate, this feature allows, for example, initiation of a step in the <i>Process Step</i> representing the GSBPM Process Phase (5) to always lead to initiation of GSBPM sub-process Integrate Data (5.1) as the next step.</p> <p>This allows a process designer to divide a <i>Business Process</i> into logical steps (for example, where each step performs a specific <i>Business Function</i> through re-use of a <i>Business Service</i>) even if these process steps will always follow each other in the same order. In all cases, the <i>Process Control Design</i> defines and the <i>Process Control</i> manages the flow between <i>Process Steps</i>, even where the flow is "trivial". <i>Process Design</i> is left to focus entirely on the design of the process itself, not sequencing between steps.</p>	

Attributes

Name	Description	Cardinality	Value Type
Name		1..1	Text
Description		1..1	Text

Process Design



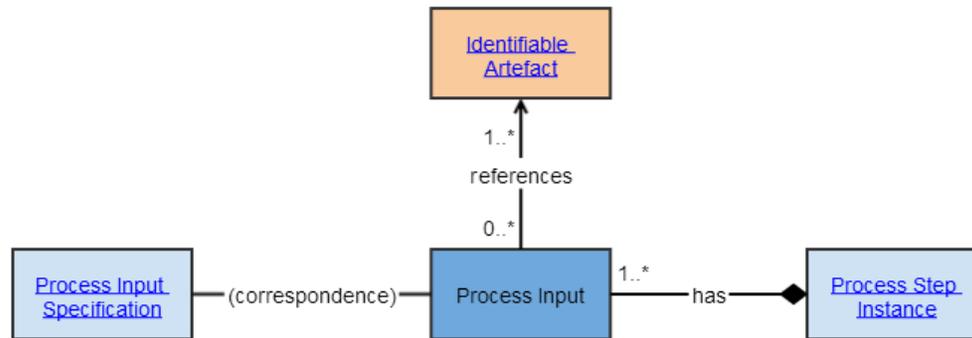
Definition

Object	Group	Definition	Explanatory Text	Synonyms
Process Design	Business	The specification of how a <i>Process Step</i> will be performed. This includes specifying the types of <i>Process Inputs</i> required and the type of <i>Process Outputs</i> that will be produced.	A <i>Process Design</i> is the design time specification of a <i>Process Step</i> that is performed as part of a run-time <i>Business Service</i> . A <i>Process Step</i> can be as big or small as the designer of a particular <i>Business Service</i> chooses. From a design perspective, one <i>Process Step</i> can contain "sub-steps", each of which is conceptualized as a (smaller) <i>Process Step</i> in its own right. Each of those "sub-steps" may contain "sub-steps" within them and so on indefinitely. It is a decision for the process designer to what extent to subdivide steps. At some level it will be appropriate to consider a <i>Process Step</i> to be a discrete task (unit of work) without warranting further subdivision. At that level the <i>Process Step</i> is designed to process particular <i>Process Inputs</i> , according to a particular <i>Process Method</i> , to produce particular <i>Process Outputs</i> . The flow between a <i>Process Step</i> and any sub steps is managed via <i>Process Control</i> .	

Attributes

Name	Description	Cardinality	Value Type
Name		1..1	Text
Description		1..1	Text

Process Input



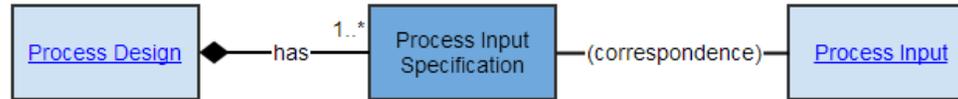
Definition

Object	Group	Definition	Explanatory Text	Synonyms
Process Input	Business	Any instance of an information object which is supplied to a <i>Process Step Instance</i> at the time its execution is initiated.	<i>Process Input</i> might include information that is used as an input that will be transformed (e.g. a <i>Data Set</i>), information that is used to control specific parameters of the process (e.g. a <i>Rule</i>), and information that is used as reference to guide the process (e.g. a <i>Code List</i>).	

Attributes

Name	Description	Cardinality	Value Type
Name		1..1	Text
Description		1..1	Text
Process Input Type	e.g. Parameter Input, Process Support Input, Transformable Input	0..1	Text

Process Input Specification



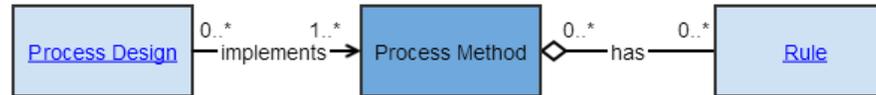
Definition

Object	Group	Definition	Explanatory Text	Synonyms
Process Input Specification	Business	A record of the types of inputs required for a <i>Process Design</i> .	<p>The <i>Process Input Specification</i> enumerates the <i>Process Inputs</i> required at the time a <i>Process Design</i> is executed. For example, if five different <i>Process Inputs</i> are required, the <i>Process Input Specification</i> will describe each of the five inputs. For each required <i>Process Input</i> the <i>Process Input Specification</i> will record the type of information object (based on GSIM) which will be used as the <i>Process Input</i> (example types might be a <i>Dimensional Data Set</i> or a <i>Classification</i>).</p> <p>The <i>Process Input</i> to be provided at the time of <i>Process Step</i> execution will then be a specific instance of the type of information object specified by the <i>Process Input Specification</i>. For example, if a <i>Process Input Specification</i> requires a <i>Dimensional Data Set</i> then the corresponding <i>Process Input</i> provided at the time of <i>Process Step</i> execution will be a particular <i>Dimensional Data Set</i>.</p>	

Attributes

Name	Description	Cardinality	Value Type
Name		1..1	Text
Description		1..1	Text
Type	This denotes the type of object which can be used as an input.	1..1	Text

Process Method



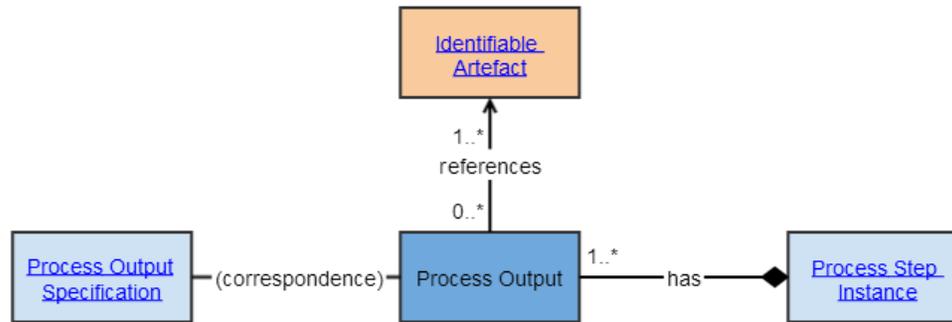
Definition

Object	Group	Definition	Explanatory Text	Synonyms
Process Method	Business	A specification of the technique which will be used to perform the unit of work.	<p>The technique specified by a <i>Process Method</i> is independent from any choice of technologies and/or other tools which will be used to apply that technique in a particular instance. The definition of the technique may, however, intrinsically require the application of specific <i>Rules</i> (for example, mathematical or logical formulas).</p> <p><i>A Process Method</i> describes a particular method for performing a <i>Process Step</i>.</p>	

Attributes

Name	Description	Cardinality	Value Type
Name		1..1	Text
Description		1..1	Text

Process Output



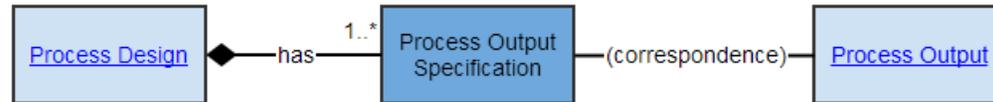
Definition

Object	Group	Definition	Explanatory Text	Synonyms
Process Output	Business	Any instance of an information object which is produced by a <i>Process Step</i> as a result of its execution.	<p><i>Process Outputs</i> have an attribute of Process Output Type, which has two possible values:</p> <ul style="list-style-type: none"> • <i>Transformed Output</i> is the result which provides the "reason for existence" of the <i>Process Step</i>. If that output were no longer required then there would be no need for the <i>Process Step</i> in its current form. Typically a <i>Transformed Output</i> is either a <i>Process Input</i> to a subsequent <i>Process Step</i> or it represents the final product from a statistical business process. • A <i>Process Metric</i> records information about the execution of a <i>Process Step</i>. For example, how long it took to complete execution of the <i>Process Step</i> and what percentage of records in the <i>Process Input</i> was updated by the <i>Process Step</i> to produce the <i>Transformed Output</i>. 	

Attributes

Name	Description	Cardinality	Value Type
Name		1..1	Text
Description		1..1	Text
Process Output Type	Transformed Output or Process Metric	0..1	Text

Process Output Specification



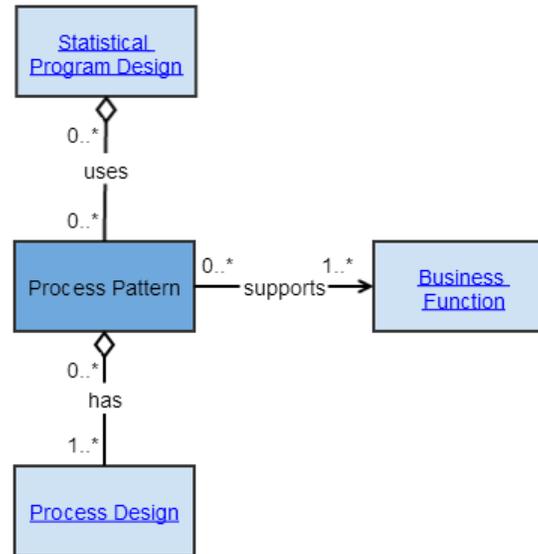
Definition

Object	Group	Definition	Explanatory Text	Synonyms
Process Output Specification	Business	A record of the types of outputs required for a Process Design.	<p>The <i>Process Output Specification</i> enumerates the <i>Process Outputs</i> that are expected to be produced at the time a <i>Process Design</i> is executed. For example, if five different <i>Process Outputs</i> expected, the <i>Process Output Specification</i> will describe each of the five outputs. For each expected <i>Process Output</i> the <i>Process Output Specification</i> will record the type of information object (based on GSIM) which will be used as the <i>Process Output</i> (Example types might be a <i>Dimensional Data Set</i> or a <i>Classification</i>).</p> <p>The <i>Process Output</i> to be provided at the time of <i>Process Step</i> execution will then be a specific instance of the type of information object specified by the <i>Process Output Specification</i>. For example, if a <i>Process Output Specification</i> expects a <i>Dimensional Data Set</i> then the corresponding <i>Process Output</i> provided at the time of <i>Process Step</i> execution will be a particular <i>Dimensional Data Set</i>.</p>	

Attributes

Name	Description	Cardinality	Value Type
Name		1..1	Text
Description		1..1	Text
Type	This denotes the type of object which can be used as an input.	1..1	Text

Process Pattern



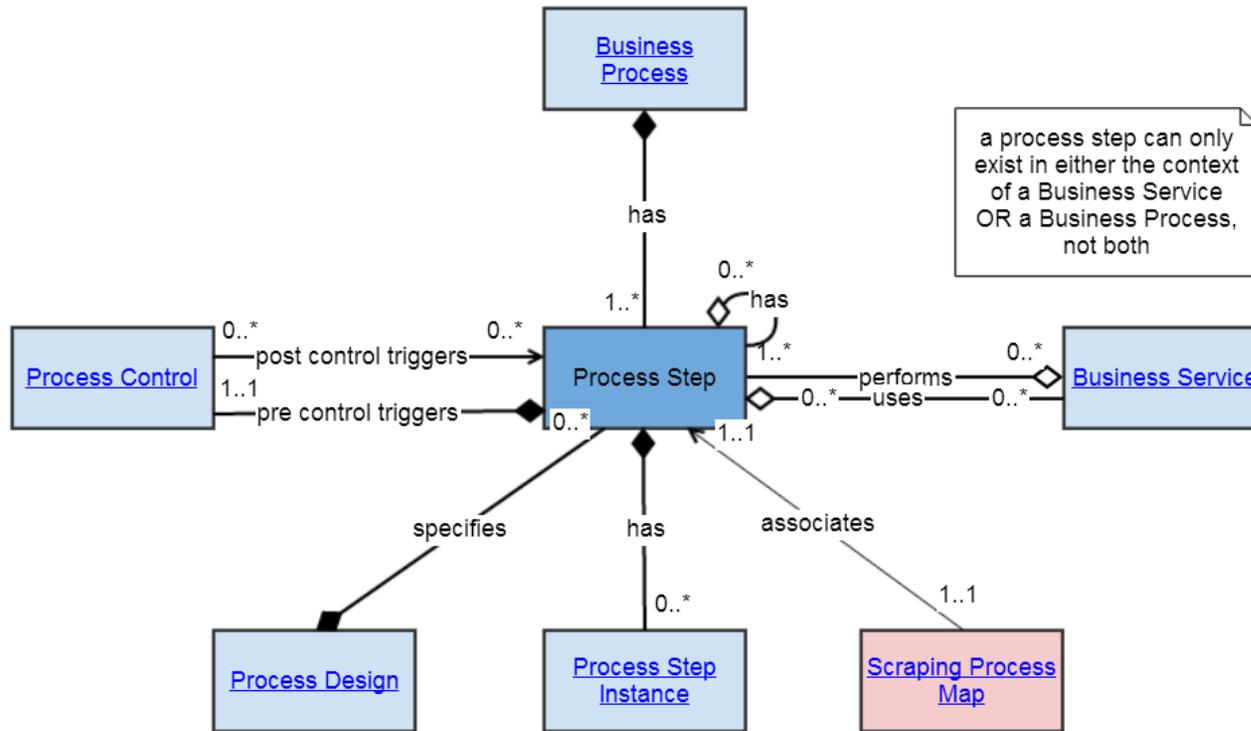
Definition

Object	Group	Definition	Explanatory Text	Synonyms
Process Pattern	Business	A nominated set of <i>Process Designs</i> , and associated <i>Process Control Designs</i> (flow), which have been highlighted for possible reuse.	<p>In a particular <i>Business Process</i>, some <i>Process Steps</i> may be unique to that <i>Business Process</i> while others may be applicable to other <i>Business Processes</i>. A <i>Process Pattern</i> can be seen as a reusable template. It is a means to accelerate design processes and to achieve sharing and reuse of design patterns which have proved effective. Reuse of <i>Process Patterns</i> can indicate the possibility to reuse related <i>Business Services</i>.</p> <p>By deciding to reuse a <i>Process Pattern</i>, a designer is actually reusing the pattern of <i>Process Designs</i> and <i>Process Control Designs</i> associated with that <i>Process Pattern</i>. They will receive a new instance of the <i>Process Designs</i> and <i>Process Control Designs</i>. If they then tailor their "instance" of the <i>Process Designs</i> and <i>Process Control Designs</i> to better meet their needs they will not change the definition of the reusable <i>Process Pattern</i>.</p>	

Attributes

Name	Description	Cardinality	Value Type
Name		1..1	Text
Description		1..1	Text

Process Step



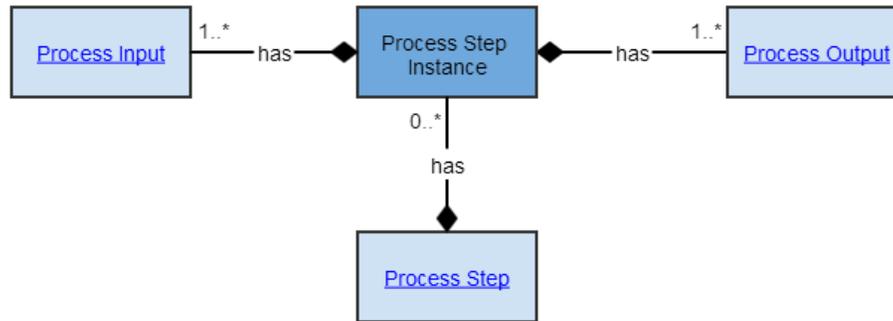
Definition

Object	Group	Definition	Explanatory Text	Synonyms
Process Step	Business	A <i>Process Step</i> is a work package that performs a <i>Business Process</i> . A <i>Process Step</i> implements the <i>Process Step Design</i> specified in order to produce the outputs for which the <i>Process Step</i> was designed.	Each <i>Process Step</i> is the use of a <i>Process Step Design</i> in a particular context (e.g. within a specific <i>Business Process</i>). At the time of execution a <i>Process Step Instance</i> specifies the actual instances of input objects (for example, specific <i>Data Sets</i> , specific <i>Variables</i>) to be supplied.	

Attributes

Name	Description	Cardinality	Value Type
Name		1..1	Text
Description		1..1	Text

Process Step Instance

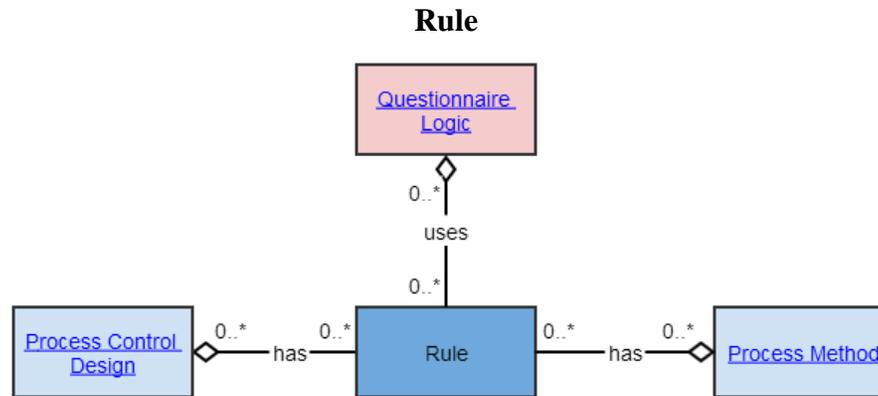


Definition

Object	Group	Definition	Explanatory Text	Synonyms
Process Step Instance	Business	An executed step in a <i>Business Process</i> . A <i>Process Step Instance</i> specifies the actual inputs to and outputs from for an occurrence of a <i>Process Step</i> .	<p>Each <i>Process Step</i> is the use of a <i>Process Step Design</i> in a particular context (e.g. within a specific Business Process). At the time of execution a <i>Process Step Instance</i> specifies the actual instances of input objects (for example, specific Data Sets, specific Variables) to be supplied.</p> <p>Each <i>Process Step Instance</i> may produce unique results even though the <i>Process Step</i> remains constant.</p> <p>Even when the inputs remain the same, metrics such as the elapsed time to complete execution of process step may vary from execution to execution. For this reason, each <i>Process Step Instance</i> details of inputs and outputs for that instance of implementing the <i>Process Step</i>.</p> <p>In this way it is possible to trace the flow of execution of a <i>Business Process</i> through all the <i>Process Steps</i> which were involved.</p>	

Attributes

Name	Description	Cardinality	Value Type
Name		1..1	Text
Description		1..1	Text
Start Time	The time a process instance starts.	0..1	Date
End Time	The time a process instance ends.	0..1	Date
Trigger Event	The event which gives the signal to start the process execution.	0..1	Number



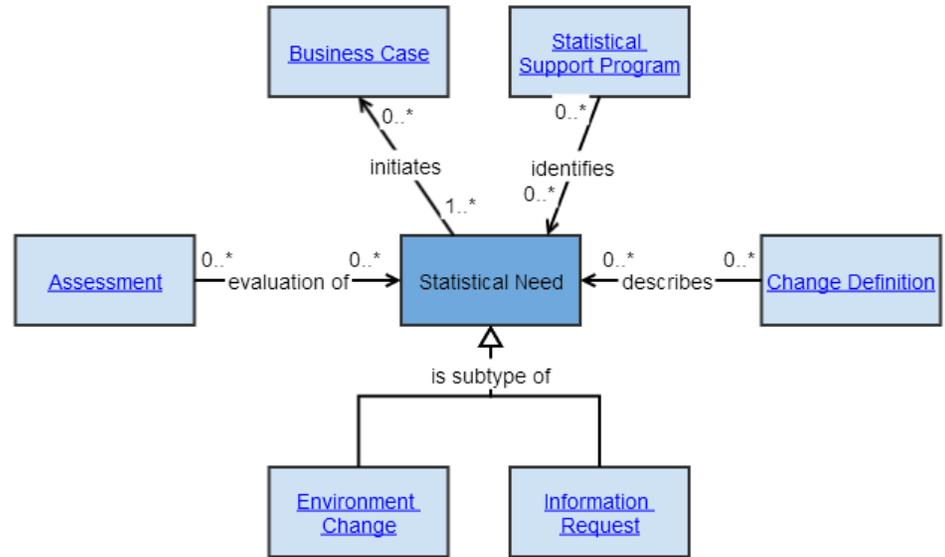
Definition

Object	Group	Definition	Explanatory Text	Synonyms
Rule	Business	A specific mathematical or logical expression which can be evaluated to determine specific behavior.	<i>Rules</i> are of several types: they may be derived from methods to determine the control flow of a process when it is being designed and executed; they may be used as the input parameters of processes (e.g., imputation rules, edit rules); and they may be used to drive the logical flow of a questionnaire. There are many forms of <i>Rules</i> and their purpose, character and expression can vary greatly.	

Attributes

Name	Description	Cardinality	Value Type
Name		1..1	Text
Description	A description of the rule, including information about its use and relationship to methods, etc.	1..1	Text
Algorithm	The rule expressed as an algorithm.	0..1	Text
Rule Type	A type taken from a controlled vocabulary.	0..1	Text
System Executable Indicator	Whether the rule is formatted to be executed by a system, or is only documentary.	0..1	Binary

Statistical Need



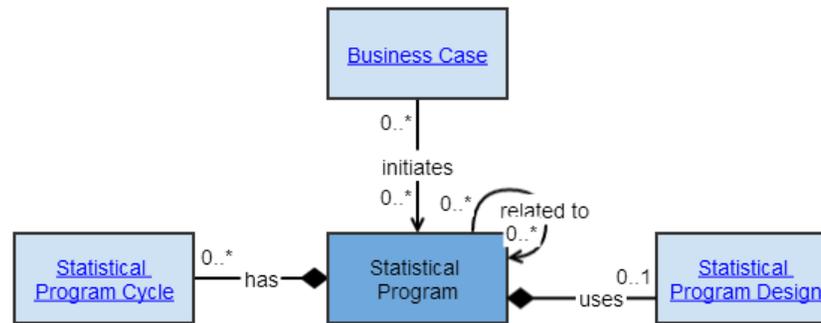
Definition

Object	Group	Definition	Explanatory Text	Synonyms
Statistical Need	Business	A requirement, request or other notification that will be considered by an organization. A <i>Statistical Need</i> does not necessarily have structure or format - it is a 'raw' need as received by the organization. A <i>Statistical Need</i> may be of a variety of types including <i>Environmental Change</i> or <i>Information Request</i> .	<p>The <i>Statistical Need</i> is a proposed or imposed requirement, request or other notification as it has been received by an organization. A <i>Statistical Need</i> is a raw expression of a requirement, and is not necessarily well-defined. A related object - <i>Change Definition</i> - is created when a <i>Statistical Need</i> is analyzed by an organization. <i>Change Definition</i> expresses the raw need in well-defined, structured terms.</p> <p>Once a <i>Statistical Need</i> has been received, the first step is to do the conceptual work to establish what it is we are trying to measure. The final output of this conceptual work is the <i>Change Definition</i>.</p> <p>In some cases, the <i>Statistical Need</i> can result from the <i>Assessment</i> of the quality, efficiency, etc. of an existing process.</p>	

Attributes

Name	Description	Cardinality	Value Type
Name		1..1	Text
Description		0..1	Text
Date created		0..1	Date
Type		0..1	e.g. Information request; External environment change; Internal environment change; Other
Status		0..1	Met/Unmet

Statistical Program



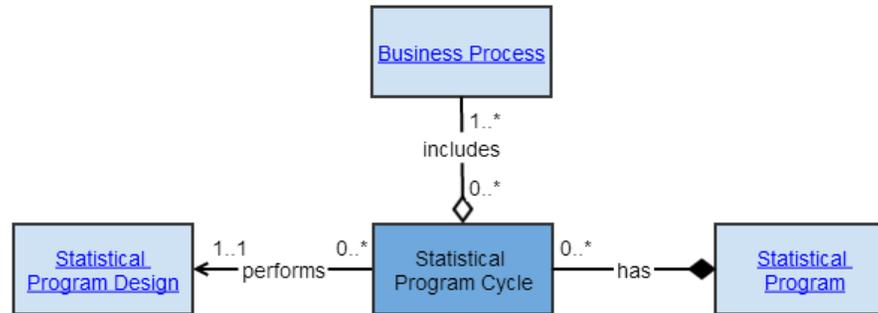
Definition

Object	Group	Definition	Explanatory Text	Synonyms
Statistical Program	Business	A set of activities, which may be repeated, that describes the purpose and context of a set of <i>Business Process</i> within the context of the relevant <i>Statistical Program Cycles</i> .	<p>The <i>Statistical Program</i> is one of a family of objects that provide the environmental context in which activities to produce statistics within a statistical organization are conducted. <i>Statistical Program</i> is the top level object that describes the purpose and objectives of a set of activities. <i>Statistical Program</i> will usually correspond to an ongoing activity such as a survey or output series. Some examples of <i>Statistical Program</i> are:</p> <ul style="list-style-type: none"> • Labour Force Survey - Multipurpose Household Survey - National Accounts - Demography - Overseas Arrivals and Departures <p>Related to the <i>Statistical Program</i> object there are <i>Statistical Program Design</i> and <i>Statistical Program Cycle</i> objects that hold the detailed information about the design and conduct of the <i>Business Process</i>.</p> <p>In the case of the traditional approach, an organization has received a <i>Statistical Need</i> and produced a <i>Change Definition</i> and an approved <i>Business Case</i>. The <i>Business Case</i> will specify either a change to the design or methodology of an existing <i>Statistical Program</i>, which will result in a new <i>Statistical Program Design</i>; or a change to one or more existing <i>Statistical Programs</i> (for example, to add an additional objective to the <i>Statistical Program</i>); or result in a new <i>Statistical Program</i> being created.</p> <p>This does not include statistical support functions such as metadata management, data management (and other overarching GSBPM processes) and design functions. These activities are conducted as part of <i>Statistical Support Programs</i>.</p>	

Attributes

Name	Description	Cardinality	Value Type
Name	The official name of the Statistical Program	1..1	Text
Description	The general description of the scope, function, and goals of the program	1..1	Text
Date initiated		0..1	Date
Date ended		0..1	Date
Status	The current condition of the program	1..1	e.g. New Proposal, New-Under Development, Current, Completed, Cancelled, Transferred to Another Organization
History	A description of the precursors of the program within the organization	0..1	Text
Subject Matter Domain		1..*	Text
Source of funding		0..1	Text
Budget		0..1	Numeric
Legislative reference	Any legislative materials, e.g. parliamentary tabling documents	0..*	Text
Legal framework		0..*	Text

Statistical Program Cycle



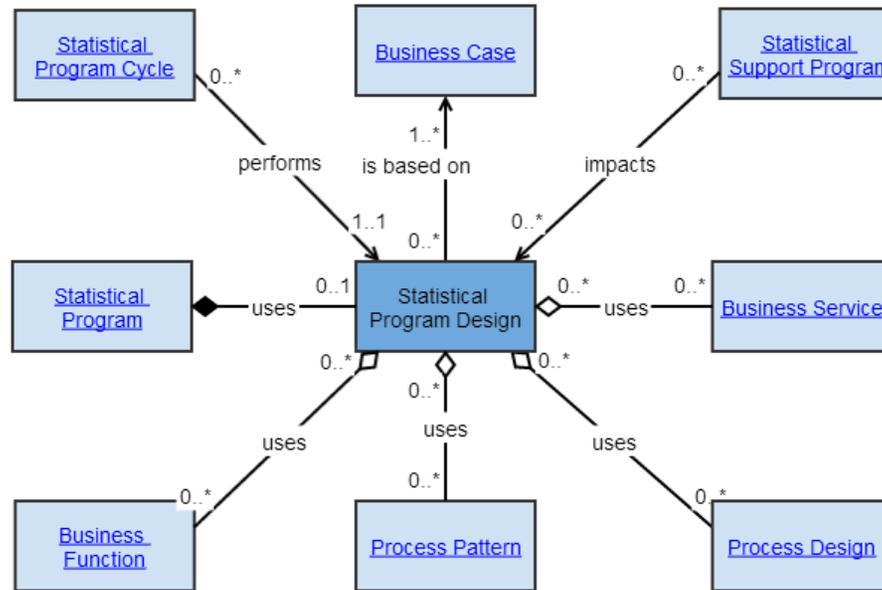
Definition

Object	Group	Definition	Explanatory Text	Synonyms
Statistical Program Cycle	Business	A set of activities to investigate characteristics of a given <i>Population</i> for a particular reference period.	A <i>Statistical Program Cycle</i> documents the execution of an iteration of a <i>Statistical Program</i> according to the associated <i>Statistical Program Design</i> for a certain reference period. It identifies the activities that are undertaken as a part of the cycle and the specific resources required and processes used and description of relevant methodological information used in this cycle defined by the <i>Statistical Program Design</i> .	

Attributes

Name	Description	Cardinality	Value Type
Name		1..1	Text
Description		1..1	Text
Reference period start		1..1	Date
Reference period end		1..1	Date

Statistical Program Design



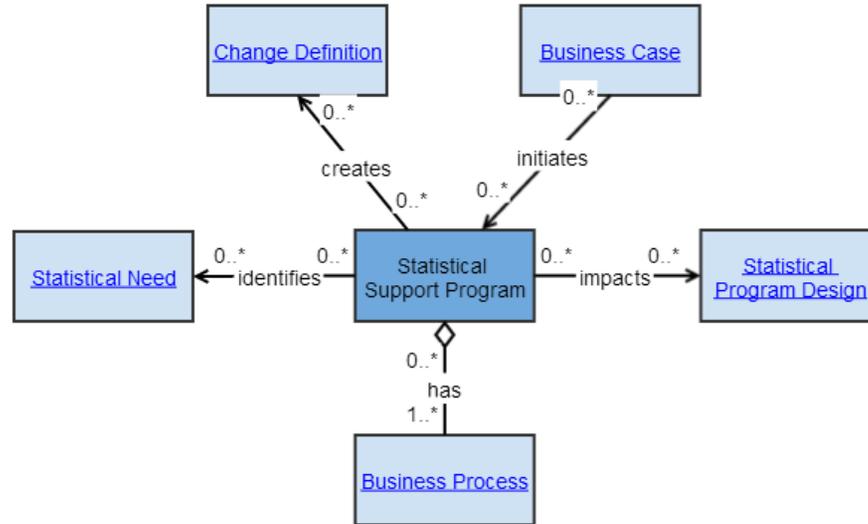
Definition

Object	Group	Definition	Explanatory Text	Synonyms
Statistical Program Design	Business	The specification of the resources required, processes used and description of relevant methodological information about the set of activities undertaken to investigate characteristics of a given <i>Population</i> .	<p>The <i>Statistical Program Design</i> is an objects that provide the operational context in which a set of <i>Business Processes</i> is conducted.</p> <p>A simple example is where a <i>Statistical Program</i> relates to a single survey, for example, the Labour Force Survey. The <i>Statistical Program</i> will have a series of <i>Statistical Program Design</i> objects that describe the methodology and design used throughout the life of the survey. When a methodological change is made to the survey, a new <i>Statistical Program Design</i> is created to record the details of the new design.</p>	

Attributes

Name	Description	Cardinality	Value Type
Name		1..1	Text
Description		1..1	Text
Date initiated	First date of validity	0..1	Date
Date ended	Last date of validity	0..1	Date
Status		1..1	Extensible redefined list e.g. New Proposal, New-Under Development, Current, Completed, Cancelled, Transferred to Another Organization
Conceptual framework	Describe the conceptual framework for the Statistical Program (e.g. SNA)	0..*	Text

Statistical Support Program



Definition

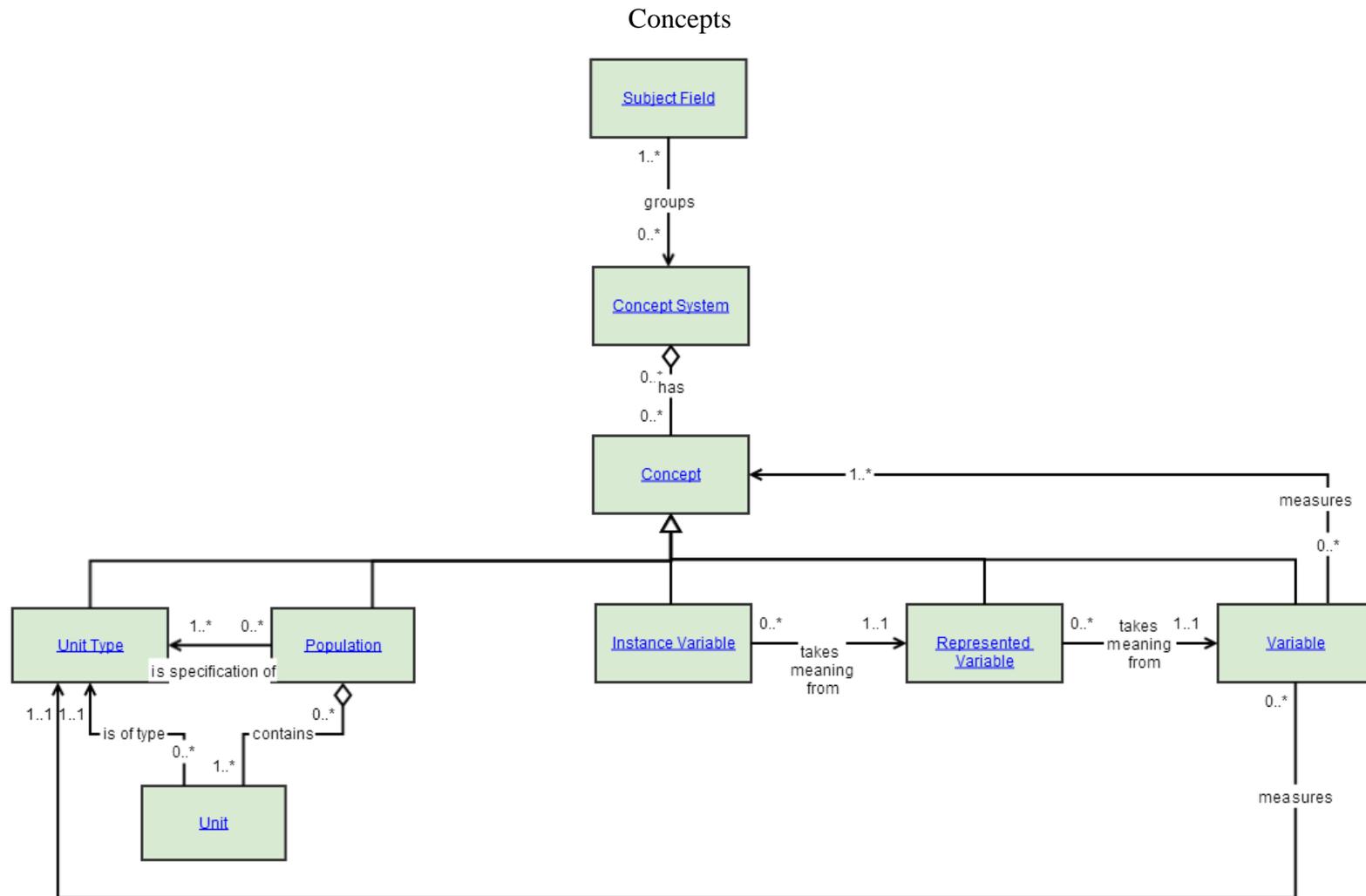
Object	Group	Definition	Explanatory Text	Synonyms
Statistical Support Program	Business	A program which is not related to the post-design cyclic production of statistical products, but is necessary to support cyclical production.	This type of program will include such functions as metadata management, data management, methodological research, and design functions. These programs correspond to the horizontal functions shown in the GSBPM, as well as programs to create new or change existing <i>Statistical Programs</i> .	

Attributes

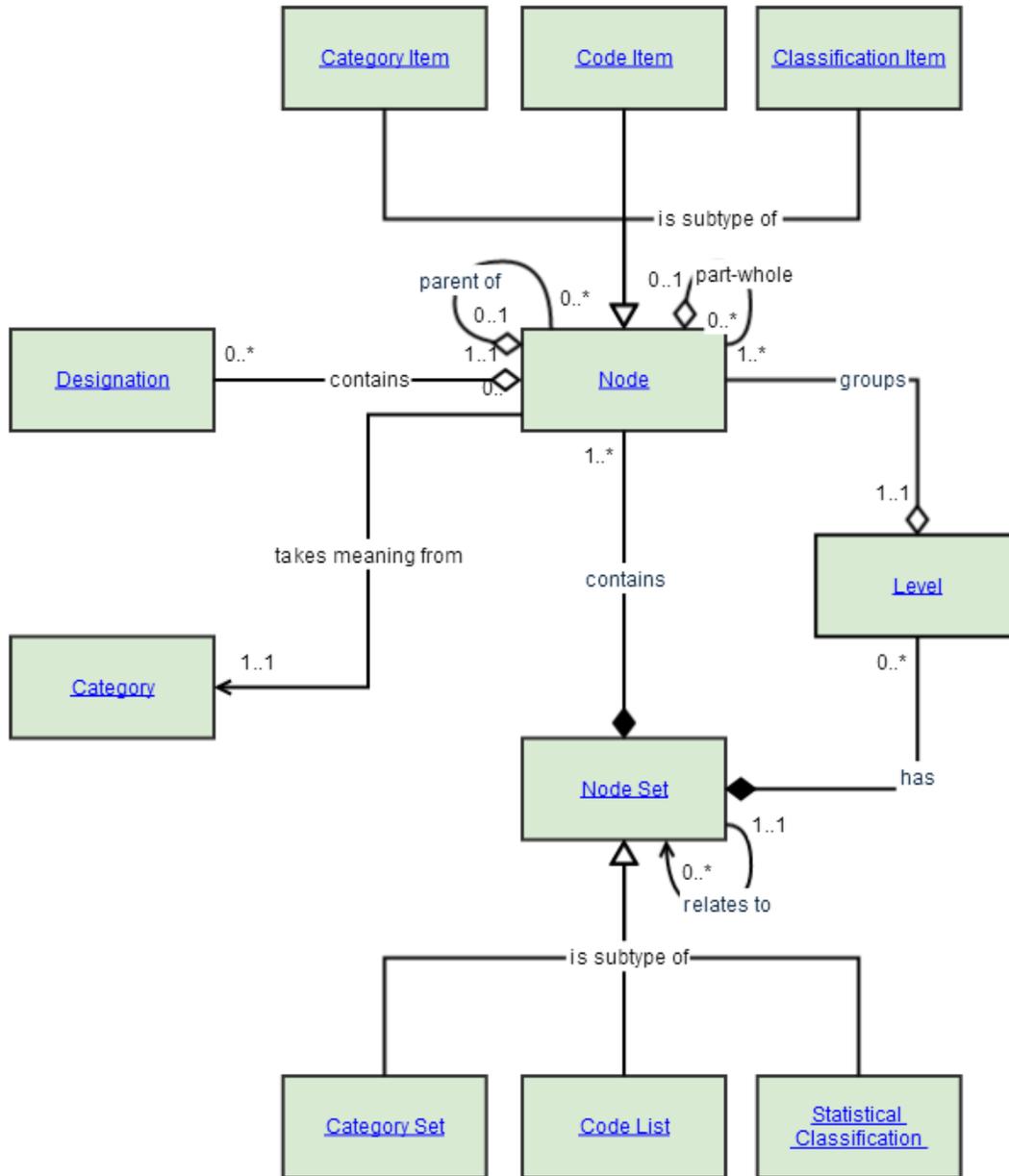
Name	Description	Cardinality	Value Type
Name	The official name of the support program	1..1	Text
Description	The general description of the scope, function, and goals of the program	1..1	Text

Date initiated		0..1	Date
Date ended		0..1	Date
Status	The current condition of the program	1..1	e.g. New Proposal, New-Under Development, Current, Completed, Cancelled, Transferred to Another Organization
History	A description of the precursors of the program within the organization	0..1	Text
Subject Matter Domain	An indication of how the program fits into the organization's scheme for describing these, if applicable.	0..1	Text
Significant events	A description of the real-world events which lead to the creation of the program	0..1	Text

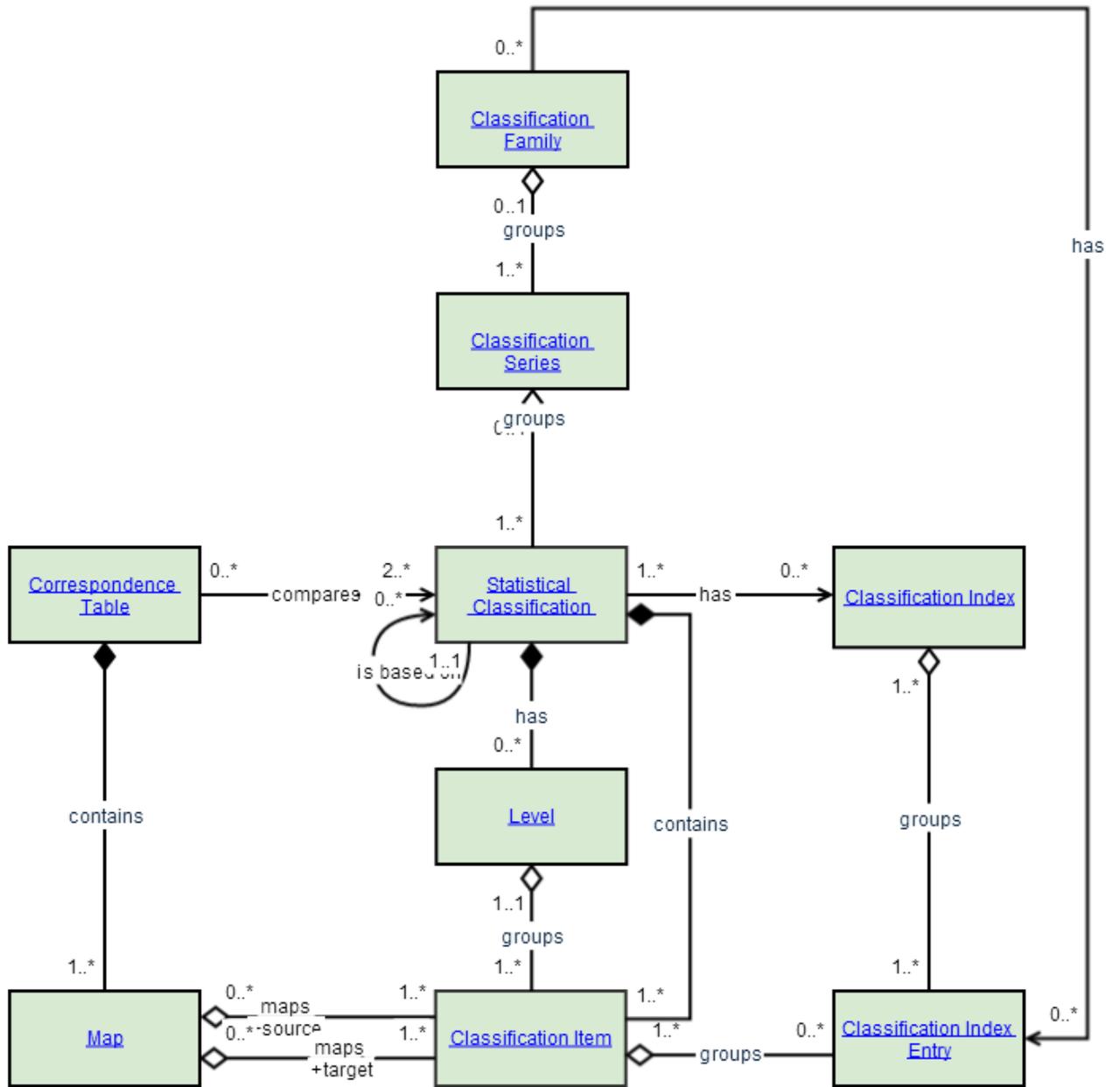
Concepts Group



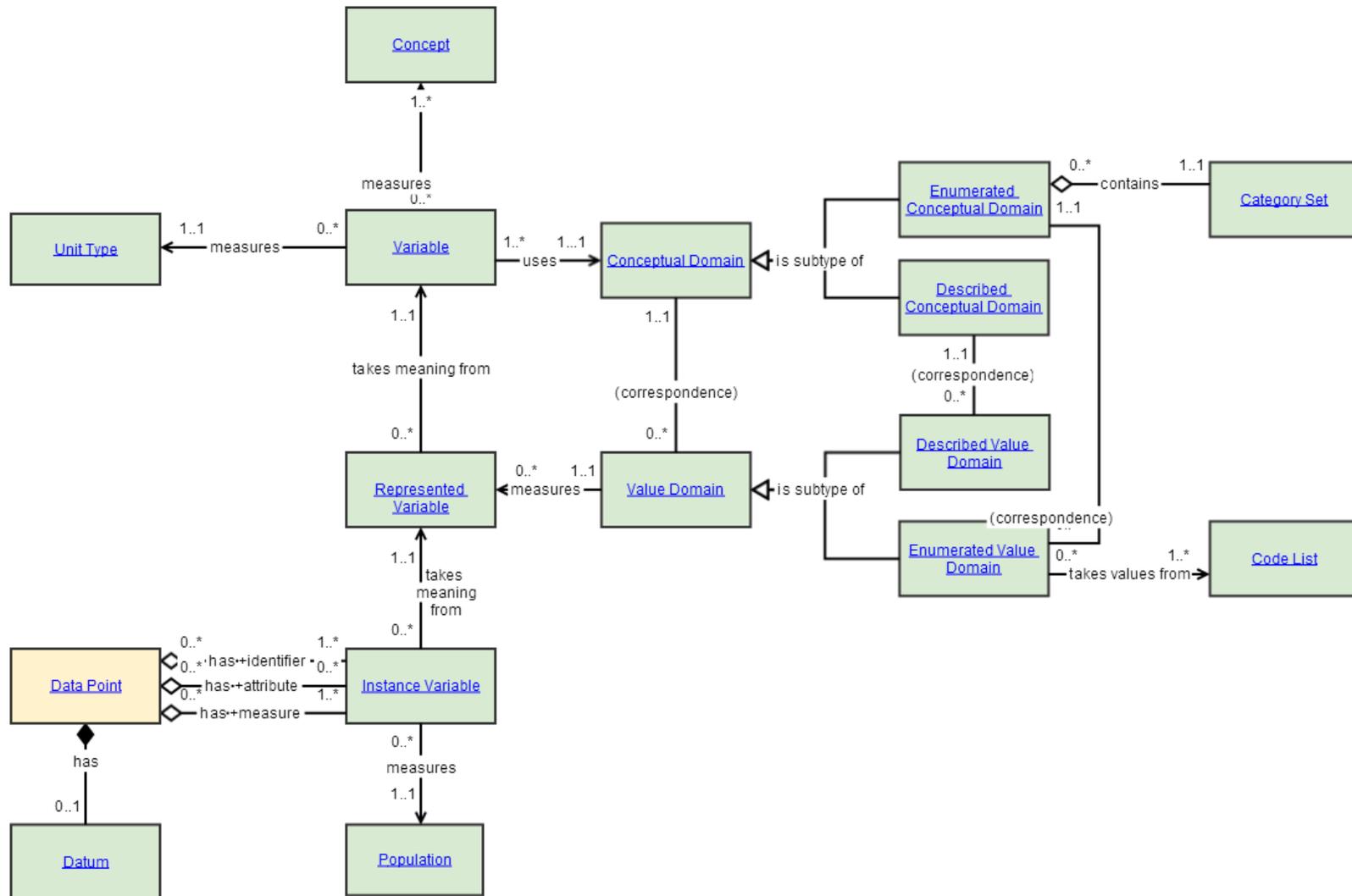
Nodes and Node Sets



Statistical Classifications



Variables



Category

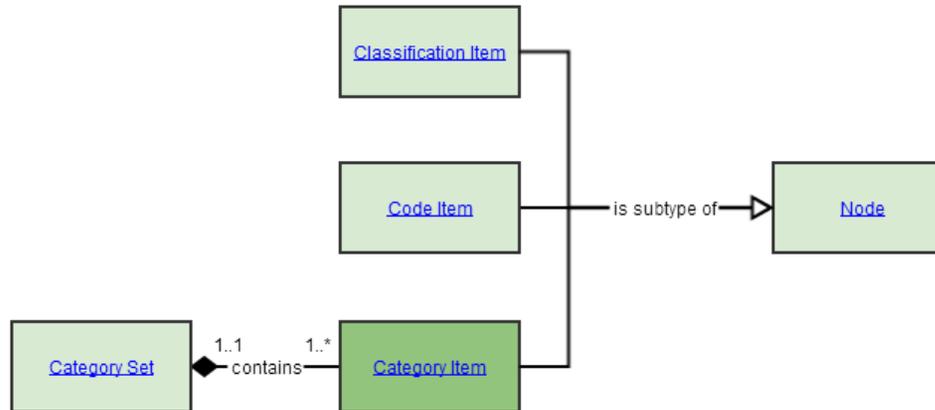


Definition

Object	Group	Definition	Explanatory Text	Synonyms
Category	Concepts	A <i>Concept</i> whose role is to extensionally define and measure a characteristic.	<p><i>Categories</i> for the <i>Concept</i> of sex include: Male, Female</p> <p>Note: An extensional definition is a description of a <i>Concept</i> by enumerating all of its sub ordinate <i>Concepts</i> under one criterion or sub division.</p> <p>For example - the Noble Gases (in the periodic table) is extensionally defined by the set of elements including Helium, Neon, Argon, Krypton, Xenon, Radon. (ISO 1087-1)</p>	class

Attributes Name	Description	Cardinality	Value Type
Name		0..1	Text
Description		0..1	Text

Category Item



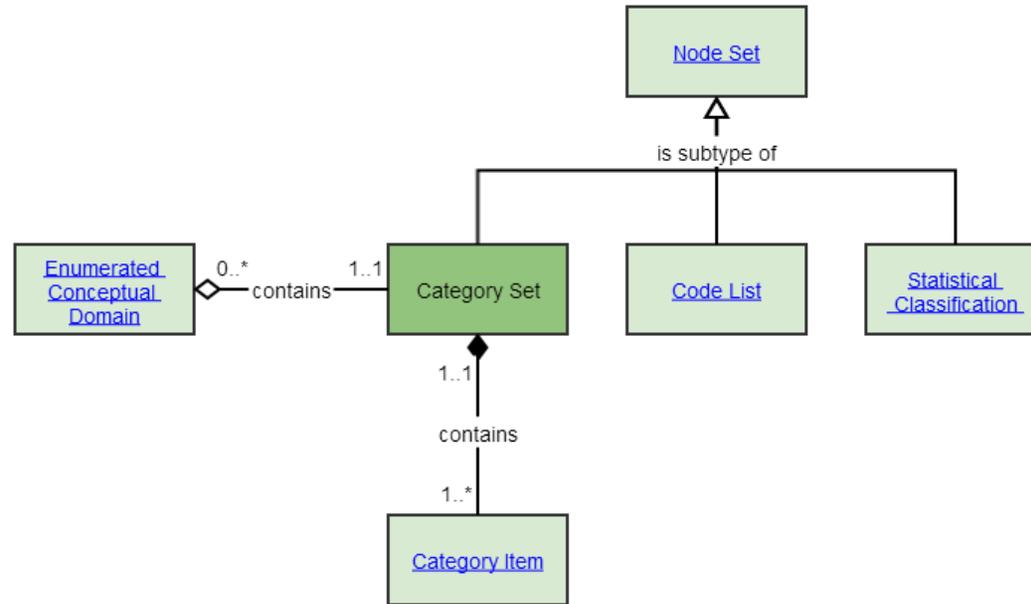
Definition

Object	Group	Definition	Explanatory Text	Synonyms
Category Item	Concepts	An element of a <i>Category Set</i> .	A type of <i>Node</i> particular to a <i>Category Set</i> type of <i>Node Set</i> . A <i>Category Item</i> contains the meaning of a <i>Category</i> without any associated representation.	

Attributes

Name	Description	Cardinality	Value Type
Name		0..1	Text
Description		0..1	Text

Category Set



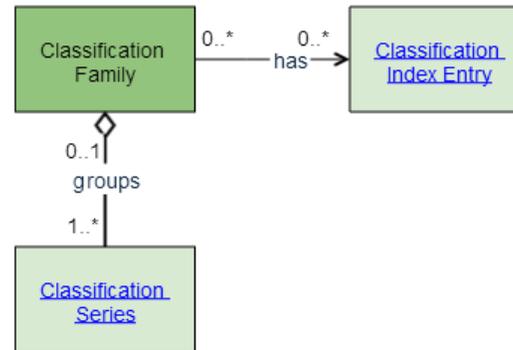
Definition

Object	Group	Definition	Explanatory Text	Synonyms
Category Set	Concepts	A list of <i>Categories</i>	A <i>Category Set</i> is a type of <i>Node Set</i> which groups <i>Categories</i> through the use of <i>Category Items</i> . The <i>Categories</i> in a <i>Category Set</i> typically have no assigned <i>Designations (Codes)</i> . For example: Male, Female	

Attributes

Name	Description	Cardinality	Value Type
Name		0..1	Text
Description		0..1	Text

Classification Family



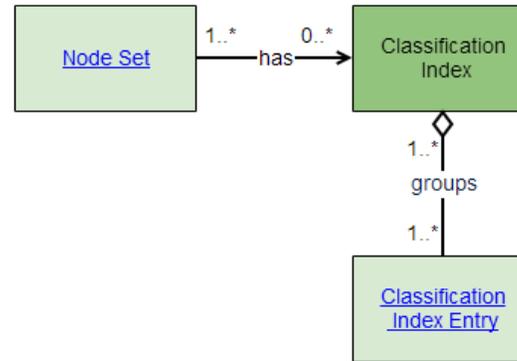
Definition

Object	Group	Definition	Explanatory Text	Synonyms
Classification Family	Concepts	A <i>Classification Family</i> is a group of <i>Classification Series</i> related from a particular point of view. The <i>Classification Family</i> is related by being based on a common <i>Concept</i> (e.g. economic activity).	Different classification databases may use different types of <i>Classification Families</i> and have different names for the families, as no standard has been agreed upon.	

Attributes

Name	Description	Cardinality	Value Domain
Name	A Classification Family has a name.	1..1	Textual
Description		0..1	Text

Classification Index



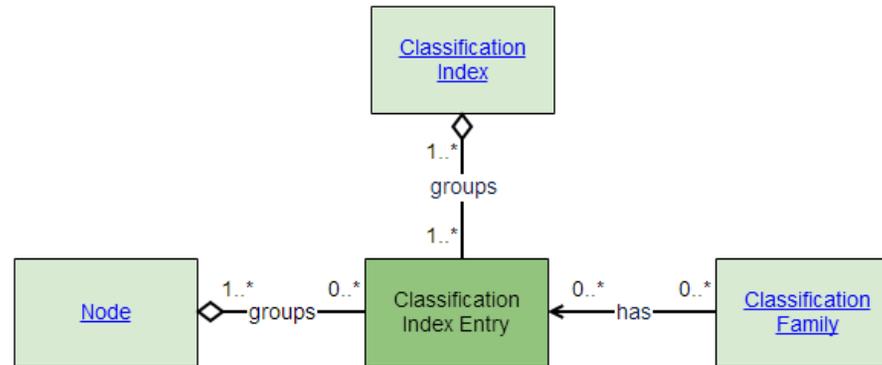
Definition

Object	Group	Definition	Explanatory Text	Synonyms
Classification Index	Concepts	A <i>Classification Index</i> is an ordered list (alphabetical, in code order etc.) of <i>Classification Index Entries</i> . A <i>Classification Index</i> can relate to one particular or to several <i>Statistical Classifications</i> .	<p>A <i>Classification Index</i> shows the relationship between text found in statistical data sources (responses to survey questionnaires, administrative records) and one or more <i>Statistical Classifications</i>. A <i>Classification Index</i> may be used to assign the codes for Classification Items to observations in statistical collections.</p> <p>A <i>Statistical Classification</i> is a subtype of <i>Node Set</i>. The relationship between <i>Statistical Classification</i> and <i>Classification Index</i> can also be extended to include the other <i>Node Set</i> types - <i>Code List</i> and <i>Category Set</i>.</p>	

Attributes

Name	Description	Cardinality	Value Domain
Name		0..1	Text
Description		0..1	Text
Release date	Date when the current version of the Classification Index was released.	0..1	Date
Maintenance unit	The unit or group of persons within the organization responsible for the Classification Index, i.e. for adding, changing or deleting Classification Index Entries.	0..1	Textual
Contact persons	Person(s) who may be contacted for additional information about the Classification Index.	0..1	Textual
Publications	A list of the publications in which the Classification Index has been published.	0..n	Textual
Languages	A Classification Index can exist in several languages. Indicates the languages available. If a 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.	0..n	Textual
Corrections	Verbal summary description of corrections, which have occurred within the Classification Index. Corrections include changing the item code associated with a Classification Index Entry.	0..n	Textual
Coding Instructions	Additional information which drives the coding process for all entries in a Classification Index.	0..n	Textual

Classification Index Entry



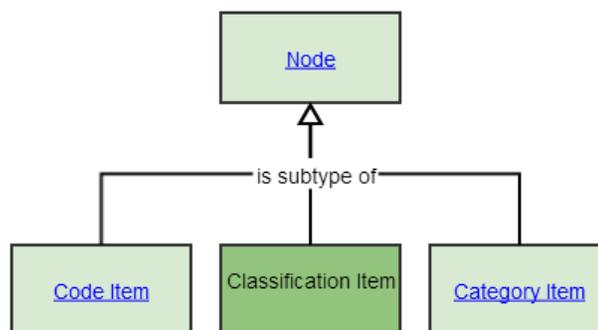
Definition

Object	Group	Definition	Explanatory Text	Synonyms
Classification Index Entry	Concepts	A <i>Classification Index Entry</i> 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 <i>Classification Item</i> applies, together with the code of the corresponding <i>Classification Item</i> . Each <i>Classification Index Entry</i> typically refers to one item of the <i>Statistical Classification</i> . Although a <i>Classification Index Entry</i> may be associated with a <i>Classification Item</i> at any <i>Level</i> of a <i>Statistical Classification</i> , <i>Classification Index Entries</i> are normally associated with items at the lowest <i>Level</i> .	A <i>Classification Item</i> is a subtype of <i>Node</i> . The relationship between <i>Classification Item</i> and <i>Classification Index Entry</i> can also be extended to include the other <i>Node</i> types - <i>Code Item</i> and <i>Category Item</i> .	

Attributes

Name	Description	Cardinality	Value Domain
Name		0..1	Text
Description		0..1	Text
Text	Text describing the type of object/unit or object property.	1..*	Textual
Valid from	Date from which the Classification Index Entry became valid. <i>The date must be defined if the Classification Index Entry belongs to a floating Classification Index.</i>	0..1	Date
Valid to	Date at which the Classification Index Entry became invalid. <i>The date must be defined if the Classification Index Entry belongs to a floating Classification Index and is no longer valid.</i>	0..1	Date
Coding Instructions	Additional information which drives the coding process. Required when coding is dependent upon one or many other factors.	0..n	Textual

Classification Item



Definition

Object	Group	Definition	Explanatory Text	Synonyms
Classification Item	Concepts	A <i>Classification Item</i> represents a <i>Category</i> at a certain <i>Level</i> within a <i>Statistical Classification</i> . It defines the content and the borders of the <i>Category</i> . A <i>Unit</i> can be classified to one and only one item at each <i>Level</i> of a <i>Statistical Classification</i> .		

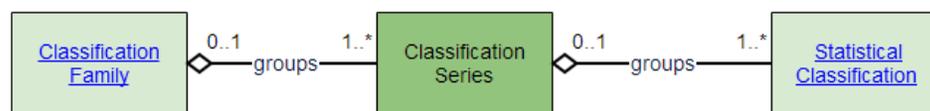
Attributes

Name	Description	Cardinality	Value Domain
Name		0..1	Text
Description		0..1	Text
Code	A <i>Classification Item</i> is identified by an alphabetical, numerical or alphanumerical code, which is in line with the code structure of the classification <i>Level</i> . The code is unique within the <i>Statistical Classification</i> to which the item belongs.	1..1	Textual
Official name	A <i>Classification Item</i> 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 <i>Statistical Classification</i> to which the item belongs, except for	1..1	Textual

Name	Description	Cardinality	Value Domain
	categories that are identical at more than one level in a hierarchical classification.		
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.	0..n	Textual
Explanatory notes	<p>A Classification Item may be associated with explanatory notes, which further describe and clarify the contents of the Category. Explanatory notes consist of:</p> <p>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.</p> <p>Includes: Specifies the contents of the Category.</p> <p>Includes also: A list of borderline cases, which belong to the described Category.</p> <p>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.</p>	0..1	Textual
Level number	The number of the <u>Level</u> to which the item belongs.	1..1	Numeric
Generated	Indicates whether or not the item has been generated to make the level to which it belongs complete.	0..1	Binary
Currently valid	If updates are allowed in the Statistical Classification, an 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 item is currently valid.	0..1	Binary
Valid from	Date from which the item became valid. The date must be defined if the item belongs to a floating Statistical classification.	0..1	Date
Valid to	Date at which the item became invalid. The date must be defined if the item belongs to a floating Statistical classification and is no longer valid.	0..1	Date
Future events	The future events describe a change (or a number of changes) related to an invalid item. These changes may e.g. have turned the now invalid item into	0..n	Textual

Name	Description	Cardinality	Value Domain
	one or several successor items. This allows the possibility to follow successors of the item in the future.		
Changes from previous version of the Statistical Classification	Describes the changes, which the item has been subject to from the previous version to the actual Statistical Classification.	0..1	Textual
Updates	Describes the changes, which the item has been subject to during the life time of the actual Statistical Classification.	0..1	Textual
Parent item	The item at the next higher level of the classification Statistical Classification of which the actual item is a sub item.	0..1	Identifier
Sub items	Each item, which is not at the lowest level of the Statistical Classification, might contain one or a number of sub items, i.e. items at the next lower level of the Statistical Classification.	0..n	Identifier
Linked items	Items of other classification versions or variants with which the item is linked, either as source or target, through Correspondence Tables.	0..n	Identifier
Case laws	Refers to identifiers of one or more case law rulings related to the Classification Item.	0..n	Textual
Case law descriptions	Refers to descriptions of the above case laws	0..n	Textual
Case law dates	Refers to date of above case laws	0..n	Date

Classification Series



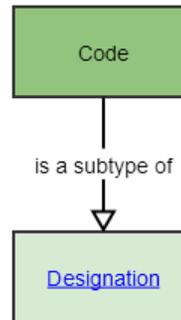
Definition

Object	Group	Definition	Explanatory Text	Synonyms
Classification Series	Concepts	A <i>Classification Series</i> is an ensemble of one or more <i>Statistical Classifications</i> , based on the same concept, and related to each other as versions or updates. Typically, these <i>Statistical Classifications</i> have the same name (for example, ISIC or ISCO).		

Attributes

Name	Description	Cardinality	Value Domain
Name	A Classification Series has a name as provided by the owner.	0..1	Text
Description	Short general description of the Classification Series, including its purpose, its main subject areas etc.	0..1	Text
Context	Classification Series can be designed in a specific context.	0..1	Text
Objects/units classified	A Classification Series is designed to classify a specific type of object/unit according to a specific attribute.	1..1	Text
Subject areas	Areas of statistics in which the Classification Series is implemented.	1..1	Text
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.	0..1	Text
Keywords	A Classification Series can be associated with one or a number of keywords.	0..n	Text

Code



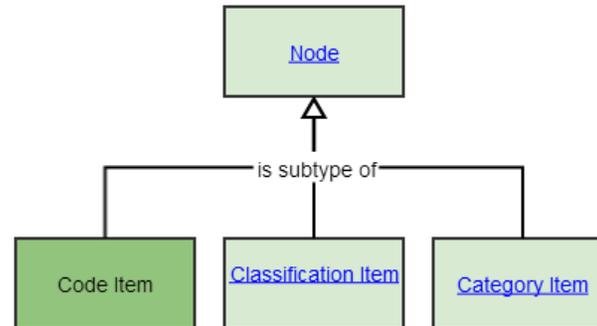
Definition

Object	Group	Definition	Explanatory Text	Synonyms
Code	Concepts	A <i>Designation</i> for a <i>Category</i> .	<i>Codes</i> are unique within their <i>Code List</i> . Example: M (Male) F (Female).	

Attributes

Name	Description	Cardinality	Value Type
Name		0..1	Text
Description		0..1	Text

Code Item



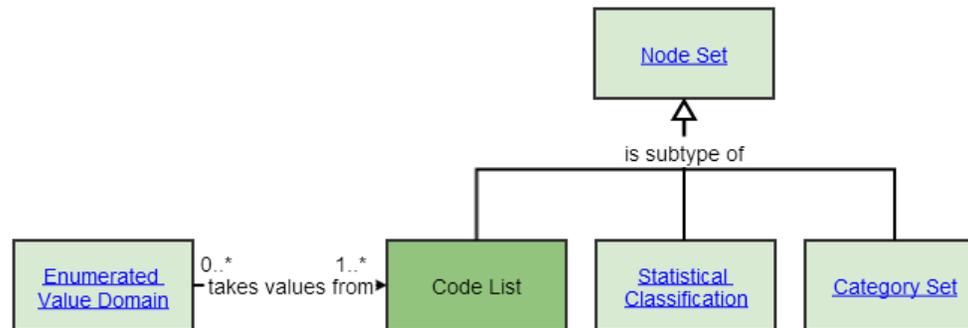
Definition

Object	Group	Definition	Explanatory Text	Synonyms
Code Item	Concepts	An element of a <i>Code List</i> .	A type of <i>Node</i> particular to a <i>Code List</i> type of <i>Node Set</i> . A <i>Code Item</i> combines the meaning of the included <i>Category</i> with a <i>Code</i> representation.	

Attributes

Name	Description	Cardinality	Value Type
Name	A human-readable identifier for the object	0..1	Text
Description	A human-readable description of the object	0..1	Text

Code List



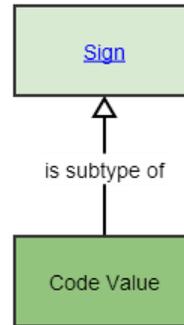
Definition

Object	Group	Definition	Explanatory Text	Synonyms
Code List	Concepts	A list of <i>Categories</i> where each <i>Category</i> has a predefined <i>Code</i> assigned to it.	<p>A kind of <i>Node Set</i> for which the <i>Category</i> contained in each <i>Node</i> has a <i>Code</i> assigned as a <i>Designation</i>.</p> <p>For example: 1 - Male 2 - Female</p>	

Attributes

Name	Description	Cardinality	Value Type
Name		0..1	Text
Description		0..1	Text

Code Value



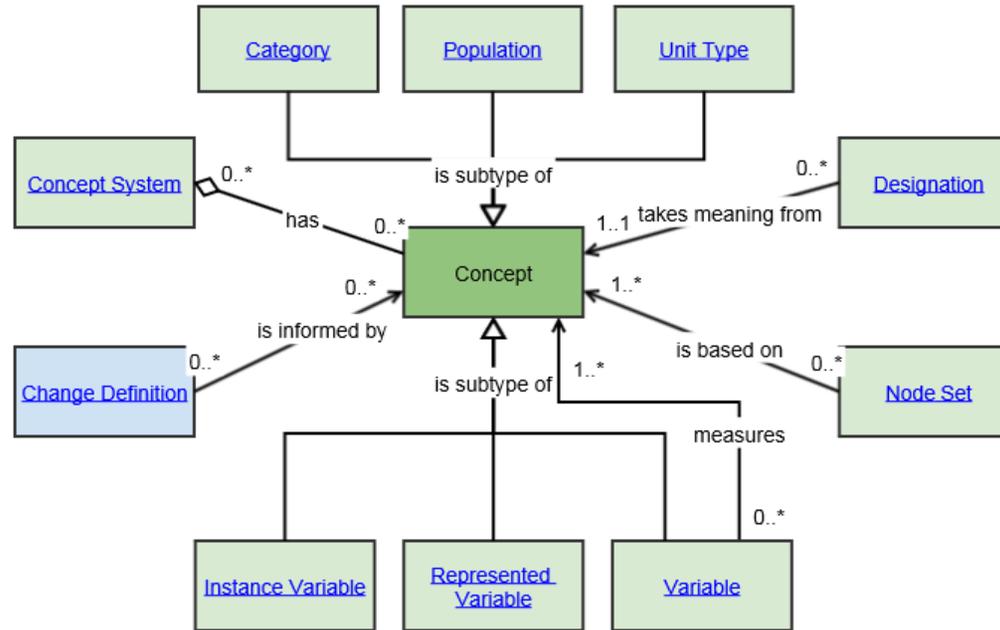
Definition

Object	Group	Definition	Explanatory Text	Synonyms
Code Value	Concepts	An alpha-numeric string used to represent a <i>Code</i> .	A <i>Code Value</i> is a subtype of <i>Sign</i> - a way of denoting the value of a <i>Code</i> . This is a kind of <i>Sign</i> used for <i>Codes</i> .	

Attributes

Name	Description	Cardinality	Value Type
Value	The value which is used to denote the code.	1..1	numeric

Concept



Definition

Object	Group	Definition	Explanatory Text	Synonyms
Concept	Concepts	Unit of thought differentiated by characteristics.		

Attributes

Name	Description	Cardinality	Value Type
Name		0..1	Text
Description		0..1	Text
Definition	Representation of a concept by a descriptive statement which serves to differentiate it from related concepts.	1..*	Text (Definitions may be rendered in different languages)

Concept System



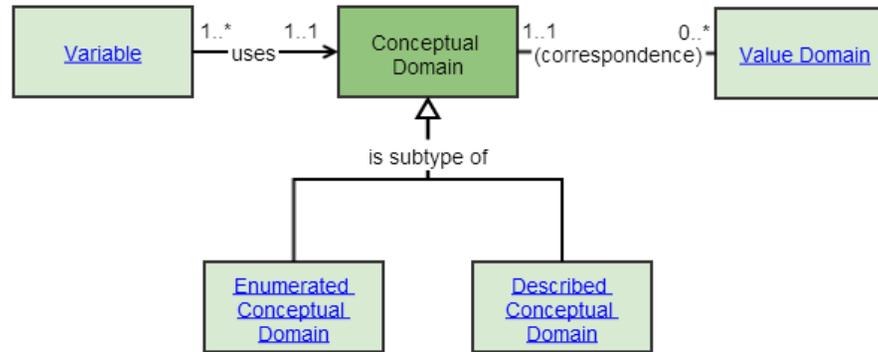
Definition

Object	Group	Definition	Explanatory Text	Synonyms
Concept System	Concepts	Set of <i>Concepts</i> structured by the relations among them.	Here are 2 examples 1) Concept of Sex: Male, Female, Other 2) ISIC (the list is too long to write down)	

Attributes

Name	Description	Cardinality	Value Type
Name	A human-readable identifier for the object	0..1	Text
Description	A human-readable description of the object	0..1	Text

Conceptual Domain



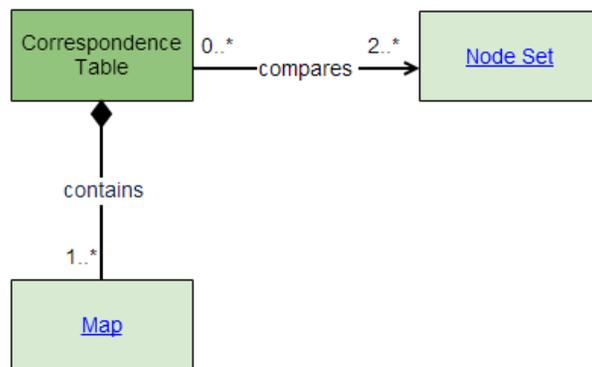
Definition

Object	Group	Definition	Explanatory Text	Synonyms
Conceptual Domain	Concepts	Set of valid <i>Concepts</i> .	The <i>Concepts</i> can be described by either enumeration or by an expression.	

Attributes

Name	Description	Cardinality	Value Type
Name		0..1	Text
Description		0..1	Text

Correspondence Table



Definition

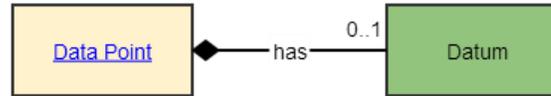
Object	Group	Definition	Explanatory Text	Synonyms
Correspondence Table	Concepts	A Correspondence Table expresses the relationship between two Statistical Classifications. These are typically: two versions from the same Classification Series; Statistical Classifications from different Classification Series; a variant and the version on which it is based; or, 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.	A Statistical Classification is a subtype of Node Set. The relationship between Statistical Classification and Correspondence Table can also be extended to include the other Node Sets - Code List and Category Set.	

Attributes

Name	Description	Cardinality	Value Domain
Identifier	A Correspondence Table is identified by a unique identifier, which may typically include the identifiers of the versions or variants involved.	1..1	Textual
Name	A Correspondence Table has a name as provided by the owner.	1..1	Textual
Description	The description contains information about the scope and aim of the correspondence and the principles on which it is based.	0..1	Textual
Owners	The statistical office, other authority or section that created and maintains the Correspondence Table. A Correspondence Table may have several owners.	0..1	Textual
Maintenance unit	The unit or group of persons who are responsible for the Correspondence Table, i.e. for maintaining and updating it.	0..1	Textual
Contact persons	The person(s) who may be contacted for additional information about the Correspondence Table.	0..n	Textual
Publications	A list of the publications in which the Correspondence Table has been published.	0..n	Textual
Source	The <u>Statistical Classification</u> from which the correspondence is made.	0..n	ID
Target	The <u>Statistical Classification</u> (s) to which the correspondence is directed. There may be multiple Target Statistical Classifications associated with the Correspondence Table.	0..n	ID
Source level	The correspondence is normally restricted to a certain <u>Level</u> in the source Statistical Classification. In this case, target items are assigned only to source items on the given level. If no level is indicated, target items can be assigned to any level of the source Statistical Classification.	0..1	ID
Target level	The correspondence is normally restricted to a certain <u>Level</u> in the target Statistical Classification. In this case, source items are assigned only to target items on the given level. If no level is indicated, source items can be assigned to any level of the target Statistical Classification.	0..n	ID
Relationship type	A correspondence can define a 1:1, 1:N, N:1 or M:N relationship between source and target items.	0..1	Textual
Floating	If the source and/or target Statistical Classifications of a correspondence table are floating classifications, the date of the correspondence table must be noted.	0..1	Textual

Name	Description	Cardinality	Value Domain
	The correspondence table expresses the relationships between the two Statistical Classifications as they existed on the date specified in the table.		

Datum



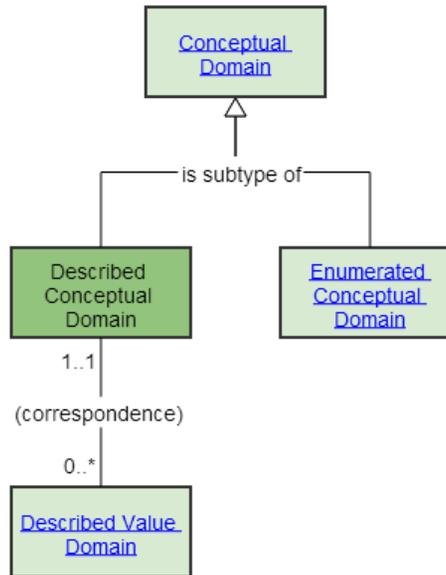
Definition

Object	Group	Definition	Explanatory Text	Synonyms
Datum	Concepts	A value.	A <i>Datum</i> is the actual instance of data that was collected or derived. It is the value which populates a <i>Data Point</i> . A <i>Datum</i> is the value found in a cell of a table.	value

Attributes

Name	Description	Cardinality	Value Type
Name		0..1	Text
Description		0..1	Text

Described Conceptual Domain



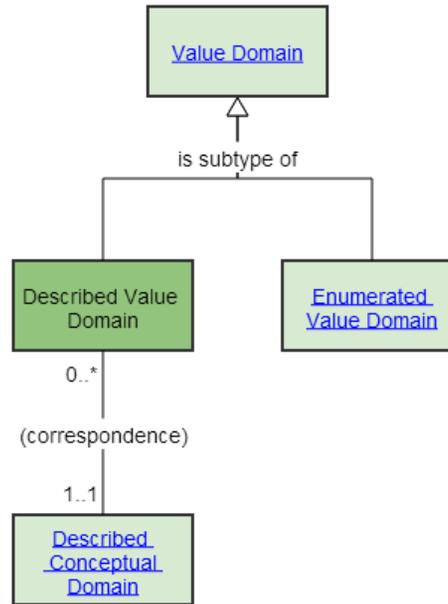
Definition

Object	Group	Definition	Explanatory Text	Synonyms
Described Conceptual Domain	Concepts	A <i>Conceptual Domain</i> defined by an expression.	For example: All real numbers between 0 and 1.	Non-enumerated conceptual domain

Attributes

Name	Description	Cardinality	Value Type
Name		0..1	Text
Description		0..1	Text

Described Value Domain



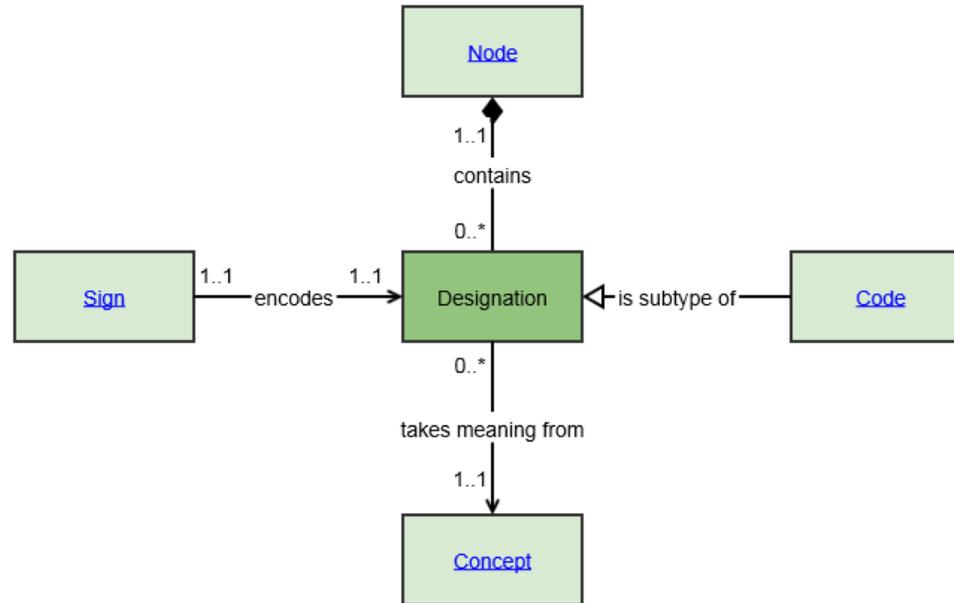
Definition

Object	Group	Definition	Explanatory Text	Synonyms
Described Value Domain	Concepts	A <i>Value Domain</i> defined by an expression.	For example: All real decimal numbers between 0 and 1.	Non-enumerated value domain

Attributes

Name	Description	Cardinality	Value Type
Name		0..1	Text
Description		0..1	Text
Unit of measure	Units by which some quantity is measured	1..1	Text
Data type		1..1	Text

Designation



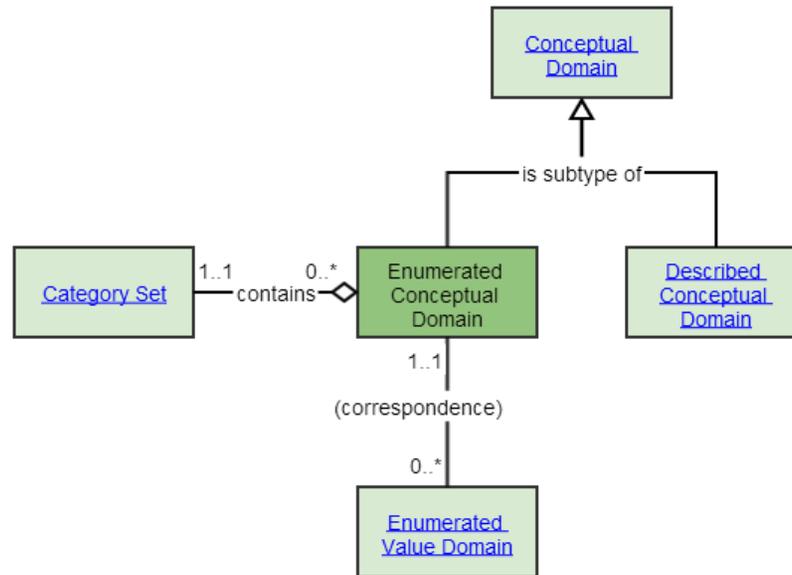
Definition

Object	Group	Definition	Explanatory Text	Synonyms
Designation	Concepts	The name given to an object for identification.	The association of a <i>Concept</i> with a <i>Sign</i> that denotes it.	

Attributes

Name	Description	Cardinality	Value Type
Name		0..1	Text
Description		0..1	Text

Enumerated Conceptual Domain



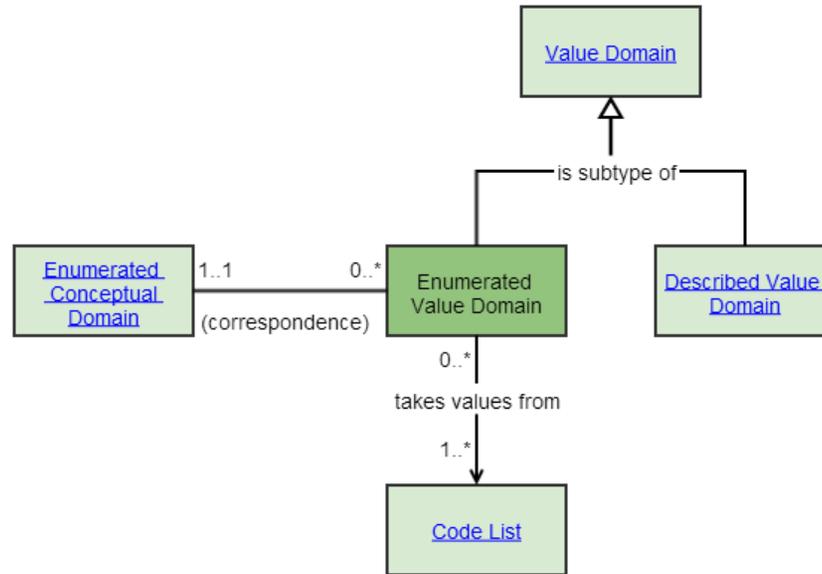
Definition

Object	Group	Definition	Explanatory Text	Synonyms
Enumerated Conceptual Domain	Concepts	A <i>Conceptual Domain</i> expressed as a list of <i>Categories</i>	For instance, the Sex <i>Categories</i> : 'Male' and 'Female'	

Attributes

Name	Description	Cardinality	Value Type
Name		0..1	Text
Description		0..1	Text

Enumerated Value Domain



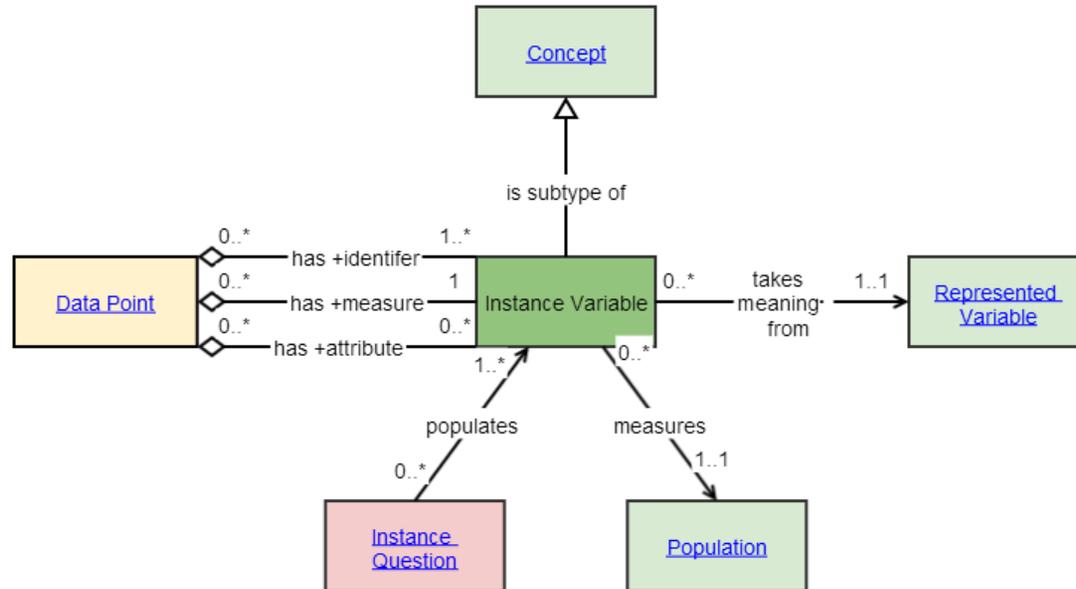
Definition

Object	Group	Definition	Explanatory Text	Synonyms
Enumerated Value Domain	Concepts	A <i>Value Domain</i> expressed as a list of <i>Categories</i> and associated <i>Codes</i> .	Example - Sex Codes <m, male>; <f, female>; <o, other>.	

Attributes

Name	Description	Cardinality	Value Type
Name		0..1	Text
Description		0..1	Text

Instance Variable



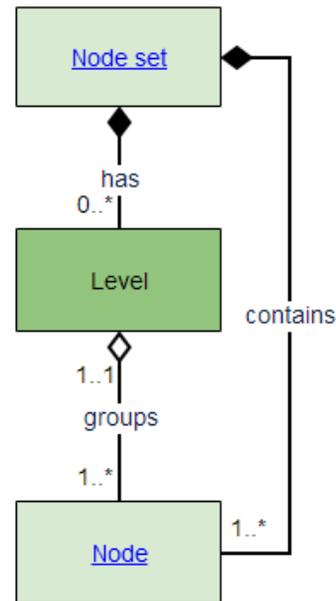
Definition

Object	Group	Definition	Explanatory Text	Synonyms
Instance Variable	Concepts	The use of a <i>Represented Variable</i> within a <i>Data Set</i> . It may include information about the source of the data.	The <i>Instance Variable</i> is used to describe actual instances of data that have been collected. Here are 3 examples: 1) Gender: Dan Gillman has gender <m, male>, Arofan Gregory has gender<m, male>, etc. 2) Number of employees: Microsoft has 90,000 employees; IBM has 433,000 employees, etc. 3) Endowment: Johns Hopkins has endowment of <3, \$1,000,000 and above>, Yale has endowment of <3, \$1,000,000 and above>, etc.	

Attributes

Name	Description	Cardinality	Value Type
Name		0..1	Text
Description		0..1	Text

Level



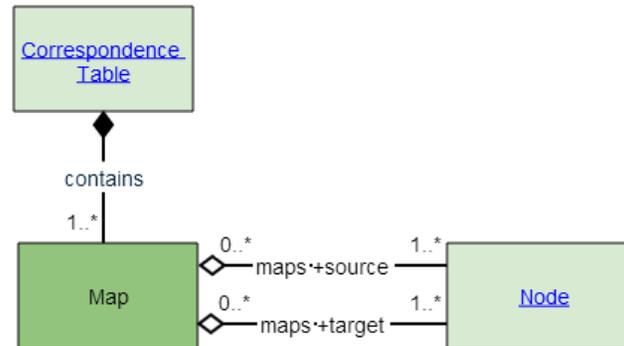
Definition

Object	Group	Definition	Explanatory Text	Synonyms
Level	Concepts	A <i>Statistical Classification</i> has a structure which is composed of one or several <i>Levels</i> . A <i>Level</i> often is associated with a <i>Concept</i> , which defines it. In a hierarchical classification the <i>Classification Items</i> of each <i>Level</i> but the highest are aggregated to the nearest higher <i>Level</i> . A linear classification has only one <i>Level</i> .	A <i>Statistical Classification</i> is a subtype of <i>Node Set</i> . The relationship between <i>Statistical Classification</i> and <i>Level</i> can also be extended to include the other <i>Node Set</i> types - <i>Code List</i> and <i>Category Set</i> .	

Attributes

Name	Description	Cardinality	Value Domain
Level number	The number associated with the Level. Levels are numbered consecutively starting with level 1 at the highest (most aggregated) Level.	0..1	Textual
Level name	The name given to the Level.	0..1	Textual
Description	Text describing the content and particular purpose of the Level.	0..1	Textual
Number of items	The number of items (Categories) at the Level.	0..1	Textual
Code type	Indicates whether the item code at the Level is alphabetical, numerical or alphanumerical.	0..1	Textual
Code structure	Indicates how the code is constructed of numbers, letters and separators.	0..1	Textual
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).	0..1	Textual
Items	An ordered list of the Categories (Classification Items) that constitute the Level.	1..n	Textual

Map



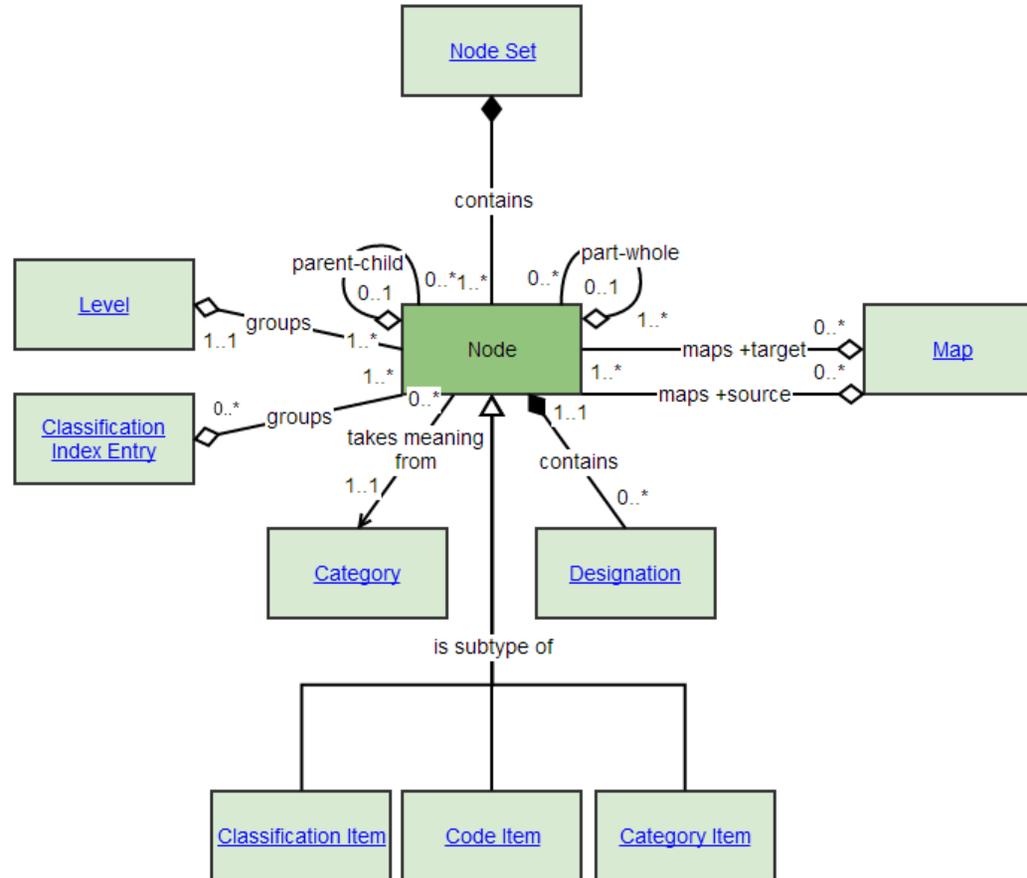
Definition

Object	Group	Definition	Explanatory Text	Synonyms
Map	Concepts	A <i>Map</i> is an expression of the relation between a <i>Classification Item</i> in a source <i>Statistical Classification</i> and a corresponding <i>Classification Item</i> in the target <i>Statistical Classification</i> . The <i>Map</i> should specify whether the relationship between the two <i>Classification Items</i> is partial or complete. Depending on the relationship type of the <i>Correspondence Table</i> , there may be several <i>Maps</i> for a single source or target item.	The use of <i>Correspondence Tables</i> and <i>Maps</i> can be extended to include all types of <i>Node</i> and <i>Node Set</i> . This means that a <i>Correspondence Table</i> could map between the items of <i>Statistical Classifications</i> , <i>Code Lists</i> or <i>Category Sets</i> .	

Attributes

Name	Description	Cardinality	Value Domain
Valid from	Date from which the Map became valid. The date must be defined if the Map belongs to a floating correspondence table.	0..1	Date
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.	0..1	Date

Node



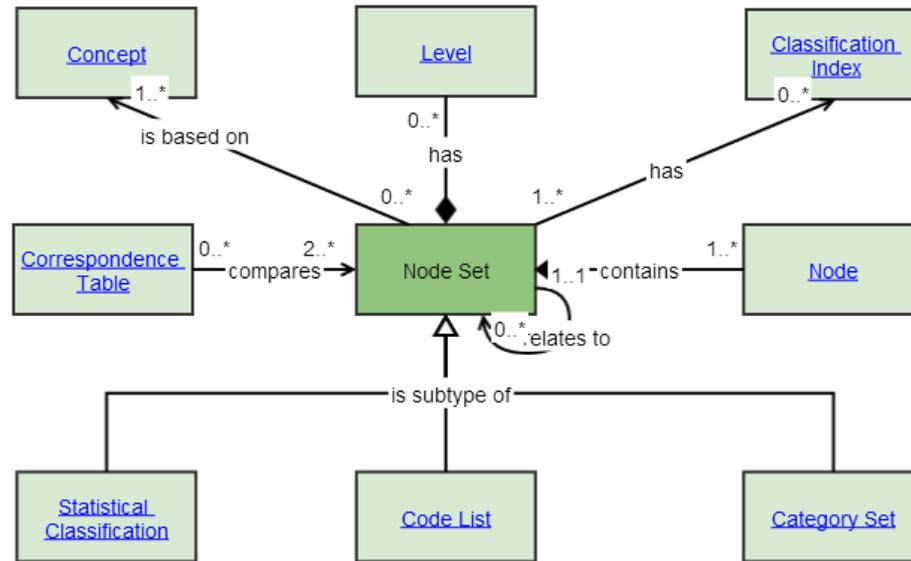
Definition

Object	Group	Definition	Explanatory Text	Synonyms
Node	Concepts	A combination of a <i>Category</i> and related attributes.	A <i>Node</i> is created as a <i>Category</i> , <i>Code</i> or <i>Classification Item</i> for the purpose of defining the situation in which the <i>Category</i> is being used.	

Attributes

Name	Description	Cardinality	Value Type
Aggregation Type	To define the parent/child relationship between Nodes, it tells us whether we are applying the part whole relationship, or the super/sub type relationships.	0..1	Text
Alias	A type of explanatory note that can be used to define alternative labels for the category contained within the Node.	0..*	Text
Name		0..1	Text
Description		0..1	Text
Annotation	A human-readable internal note intended for the developers/maintainers of GSIM.	0..*	Text

Node Set



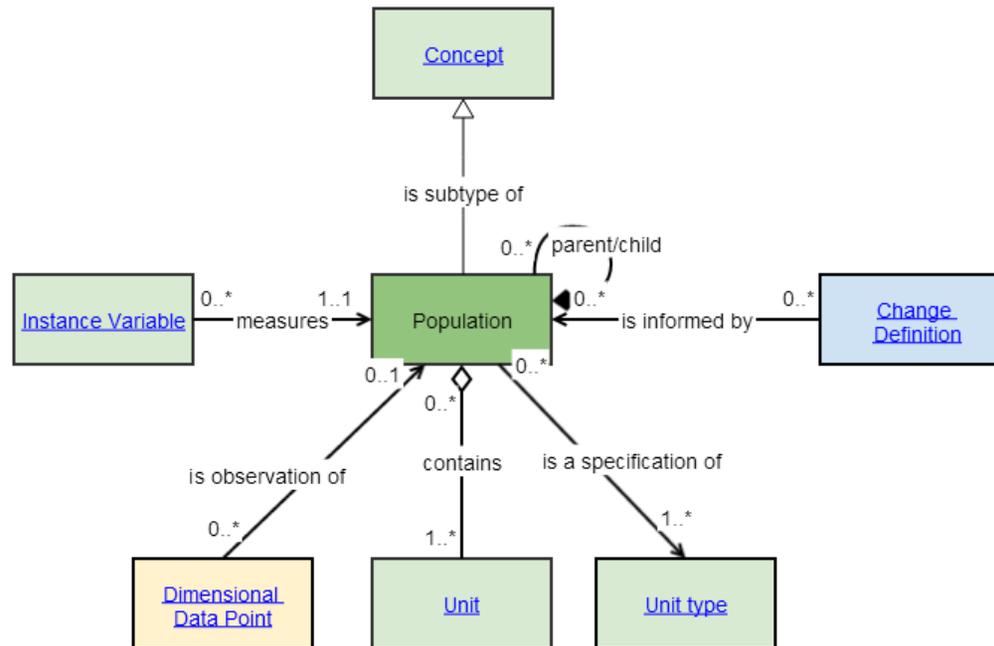
Definition

Object	Group	Definition	Explanatory Text	Synonyms
Node Set	Concepts	A set of <i>Nodes</i> .	<p><i>Node Set</i> is a kind of <i>Concept System</i>. Here are 2 examples:</p> <p>1) <i>Sex Categories</i></p> <ul style="list-style-type: none"> • Male • Female • Other <p>2) <i>Sex Codes</i></p> <ul style="list-style-type: none"> • <m, male> • <f, female> • <o, other> 	

Attributes

Name	Description	Cardinality	Value Type
Name		0..1	Text
Version	The version of the object assigned by the owning agency.	1..1	Version designator (defaults to "1.0")
Agency	The organization or legal entity which owns and maintains the object.	1..1	Entity designator
Description		0..1	Text
Annotation	A human-readable internal note intended for the developers/maintainers of GSIM.	0..n	Text
Valid From	The effective date on which the object is published.	0..1	Date
Valid To	The effective date on which the object is withdrawn from publication.	0..1	Date

Population



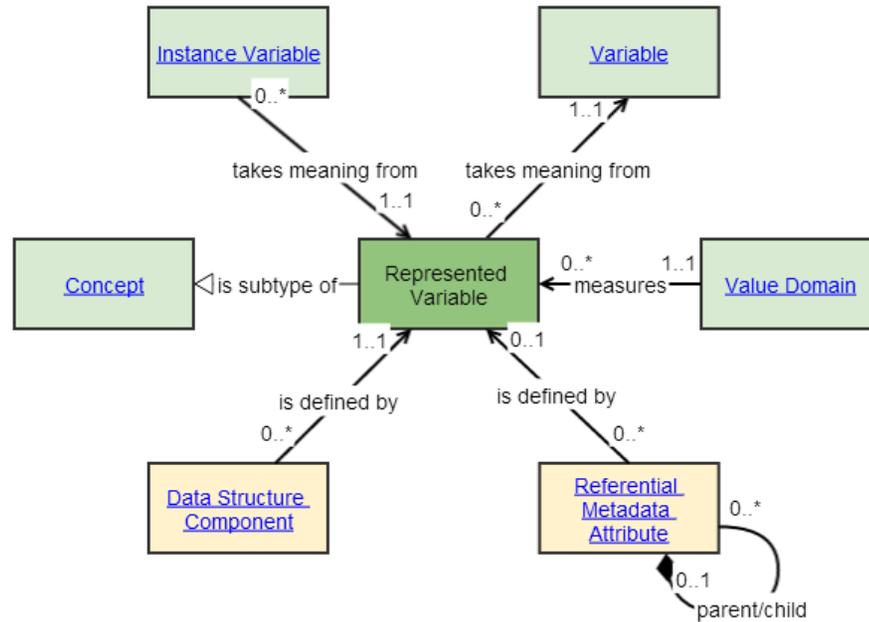
Definition

Object	Group	Definition	Explanatory Text	Synonyms
Population	Concepts	The total membership of a defined class of people, objects or events.	<p>A <i>Population</i> is used to describe the total membership of a group of people, objects or events based on characteristics, e.g. time and geographic boundaries.</p> <p>Here are 3 examples –</p> <ol style="list-style-type: none"> 1. US adult persons on 13 November 1956 2. US computer companies at the end of 2012 3. Universities in the US 1 January 2011. 	

Attributes

Name	Description	Cardinality	Value Type
Name		0..1	Text
Description		0..1	Text
Reference Period	The time period to which the population is associated	0..1	DateRange
Geography	The geographical area to which the population is associated	0..1	Text

Represented Variable



Definition

Object	Group	Definition	Explanatory Text	Synonyms
Represented Variable	Concepts	A combination of a characteristic of a population to be measured and how that measure will be represented.	<p>Example:</p> <p>The pair (Number of Employees, Integer), where "Number of Employees" is the characteristic of the population (<i>Variable</i>) and "Integer" is how that measure will be represented (<i>Value Domain</i>).</p>	

Attributes

Name	Description	Cardinality	Value Type
Name		0..1	Text
Description		0..1	Text

Sign



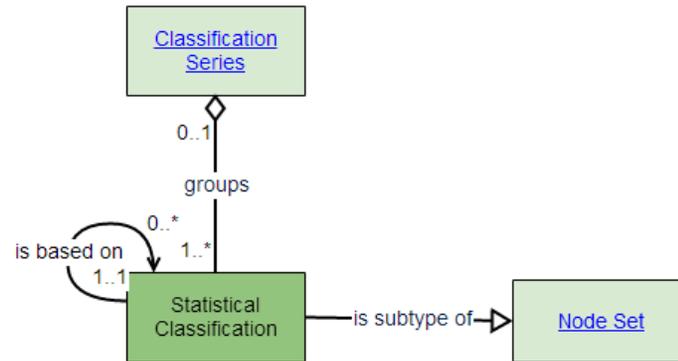
Definition

Object	Group	Definition	Explanatory Text	Synonyms
Sign	Concepts	Something that suggests the presence or existence of a fact, condition, or quality.	It is a perceivable object. This object is used to denote a <i>Concept</i> as a <i>Designation</i> .	

Attributes

Name	Description	Cardinality	Value Type
Name		0..1	Text
Description		0..1	Text
Value	A human-readable value for the object.	1..1	Text

Statistical Classification



Definition

Object	Group	Definition	Explanatory Text	Synonyms
Statistical Classification	Concepts	A <i>Statistical Classification</i> is a set of <i>Categories</i> 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 <i>Categories</i> at each <i>Level</i> of the classification structure must be mutually exclusive and jointly exhaustive of all objects/units in the population of interest.	The <i>Categories</i> are defined with reference to one or more characteristics of a particular population of units of observation. A <i>Statistical Classification</i> may have a flat, linear structure or may be hierarchically structured, such that all <i>Categories</i> at lower <i>Levels</i> are sub- <i>Categories</i> of <i>Categories</i> at the next <i>Level</i> up. <i>Categories</i> in <i>Statistical Classifications</i> are represented in the information model as <i>Classification Items</i> .	

Attributes

Name	Description	Cardinality	Value Domain
Name	A <i>Statistical Classification</i> has a name as provided by the owner or maintenance unit.	1..1	Textual
Introduction	The introduction provides a detailed description of the <i>Statistical Classification</i> , the background for its creation, the classification variable and objects/units classified, classification rules etc.	0..1	Textual
Description		0..1	Text
Release date	Date on which the <i>Statistical Classification</i> was released.	0..1	Date
Termination date	Date on which the <i>Statistical Classification</i> was superseded by a successor version or otherwise ceased to be valid.	0..1	Date
Current	Indicates whether or not the <i>Statistical Classification</i> is currently valid.	0..1	Textual
Maintenance unit	The unit or group of persons within the organization who are responsible for the <i>Statistical Classification</i> (i.e., for maintaining, updating and changing it).	0..1	Textual
Contact persons	Person(s) who may be contacted for additional information about the <i>Statistical Classification</i> .	0..n	Textual
Legal base	Indicates that the <i>Statistical Classification</i> is covered by a legal act or by some other formal agreement.	0..n	Textual
Publications	A list of the publications, including print, PDF, HTML and other electronic formats, in which the <i>Statistical Classification</i> has been published.	0..n	Textual
Name types	A list of the defined types of alternative item names available for the <i>Statistical Classification</i> . Each name type refers to a list of alternative item names.	0..n	Textual
Languages available	A <i>Statistical Classification</i> 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.	0..n	Textual
Copyright	<i>Statistical Classifications</i> may have restricted copyrights. Such <i>Statistical Classifications</i> might be excluded from downloading. Notes the copyright statement that should be displayed in official publications to indicate the copyright owner.	0..n	Textual

Dissemination allowed	Indicates whether or not the <i>Statistical Classification</i> may be published or otherwise disseminated (e.g. electronic dissemination).	0..1	Textual
Version	Indicates if the <i>Statistical Classification</i> is a version.	0..1	Textual
Update	Indicates if the <i>Statistical Classification</i> is an update.	0..1	Textual
Floating	Indicates if the <i>Statistical Classification</i> is a floating classification. In a floating statistical classification, a validity period should be defined for all <i>Classification Items</i> which will allow the display of the item structure and content at different points of time.	0..1	Textual
Predecessor	For those <i>Statistical Classifications</i> that are versions or updates, notes the preceding <i>Statistical Classification</i> of which the actual <i>Statistical Classification</i> is the successor.	0..1	Textual
Successor	Notes the <i>Statistical Classification</i> that superceded the actual <i>Statistical Classification</i> .	0..1	Textual
Changes from previous version or update	A summary description of the nature and content of changes from the preceding version or update. Specific changes are recorded in the <i>Classification Item</i> object under the “Changes from previous version and updates” attribute.	0..n	Textual
Derived from	A <i>Statistical Classification</i> can be derived from one of the classification versions of another <i>Classification Series</i> . The derived <i>Statistical Classification</i> 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 <i>Classification Items</i> , rearranging them in a different structure. Indicates the classification version from which the actual <i>Statistical Classification</i> is derived.	0..1	Textual
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 a version. Note whether <i>Classification Items</i> 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 <i>Classification Item</i> names and explanatory notes that do not involve structural changes are permissible within a version.	0..1	Textual

Updates	Summary description of changes which have occurred since the most recent classification version or classification update came into force.	0..n	Textual
Variants available	Identifies any variants associated with this version.	0..1	Textual
Variant	For those <i>Statistical Classifications</i> that are variants, notes the <i>Statistical Classification</i> on which it is based and any subsequent versions of that <i>Statistical Classification</i> to which it is also applicable.	0..1	Textual
Changes from base Statistical Classification	Describes the relationship between the variant and its base <i>Statistical Classification</i> , including regroupings, aggregations added and extensions.	0..n	Textual
Purpose of variant	If the <i>Statistical Classification</i> is a variant, notes the specific purpose for which it was developed.	0..1	Textual

Subject Field



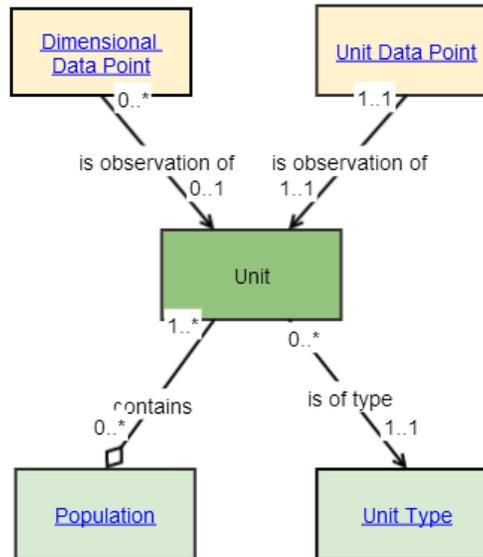
Definition

Object	Group	Definition	Explanatory Text	Synonyms
Subject Field	Concepts	One or more <i>Concept Systems</i> used for the grouping of <i>Concepts</i> and <i>Categories</i> for the production of statistics.	A <i>Subject Field</i> is a field of special knowledge under which a set of <i>Concepts</i> and their <i>Designations</i> is used. For example, labour market, environmental expenditure, tourism, etc.	subject area, theme

Attributes

Name	Description	Cardinality	Value Type
Name		0..1	Text
Description		0..1	Text

Unit

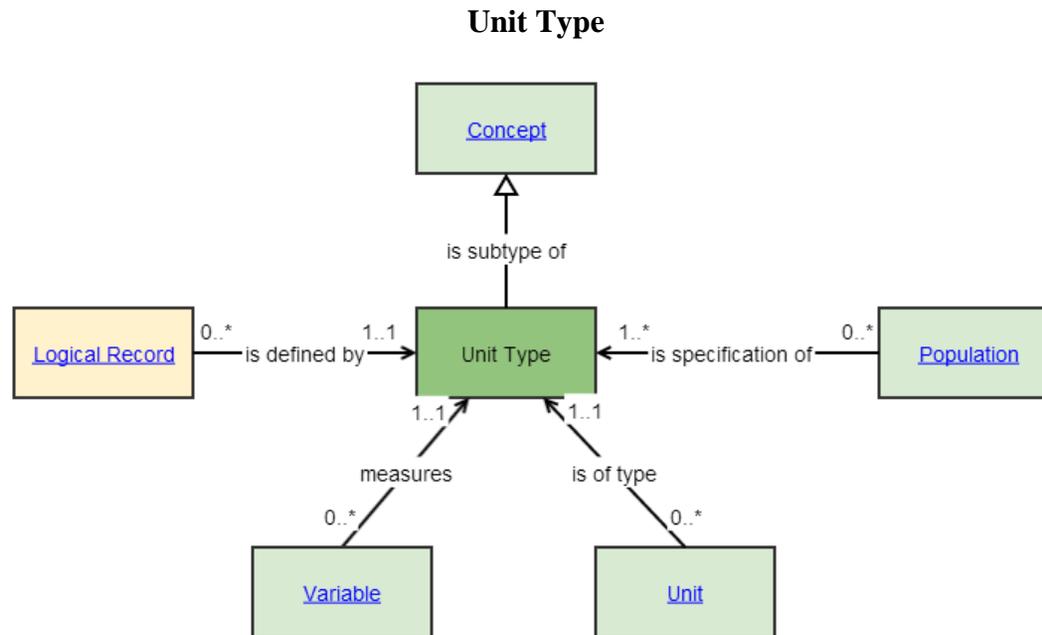


Definition

Object	Group	Definition	Explanatory Text	Synonyms
Unit	Concepts	The object of interest in a <i>Business Process</i>	Here are 3 examples - 1. Individual US person (i.e., Arofan Gregory, Dan Gillman, Barack Obama, etc.) 2. Individual US computer companies (i.e., Microsoft, Apple, IBM, etc.) 3. Individual US universities (i.e., Johns Hopkins, University of Maryland, Yale, etc.)	

Attributes

Name	Description	Cardinality	Value Type
Name		0..1	Text
Description		0..1	Text



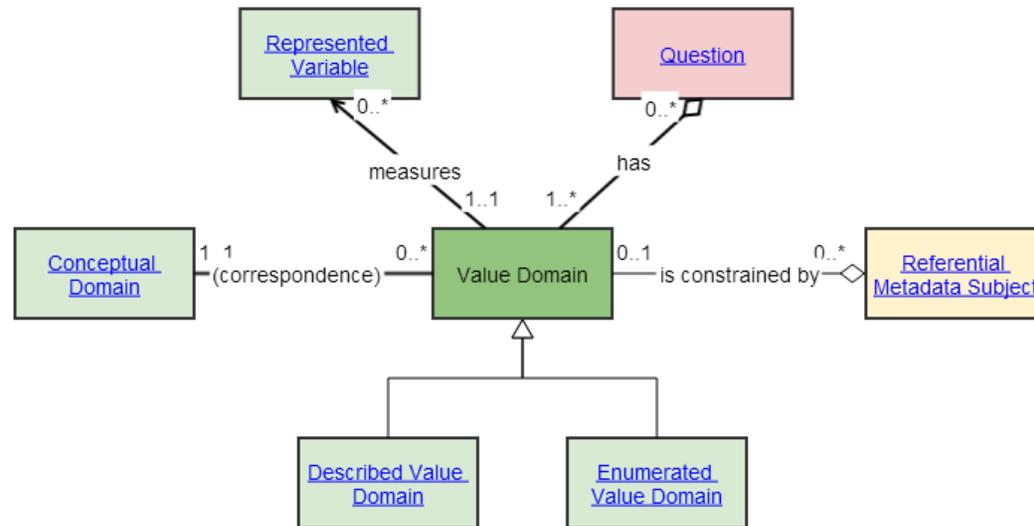
Definition

Object	Group	Definition	Explanatory Text	Synonyms
Unit Type	Concepts	A <i>Unit Type</i> is a class of objects of interest	A <i>Unit Type</i> is used to describe a class or group of <i>Units</i> based on a single characteristic, but with no specification of time and geography. For example, the <i>Unit Type</i> of “Person” groups together a set of <i>Units</i> based on the characteristic that they are ‘Persons’. It concerns not only <i>Unit Types</i> used in dissemination, but anywhere in the statistical process. E.g. using administrative data might involve the use of a fiscal unit.	Object class (ISO 11179)

Attributes

Name	Description	Cardinality	Value Type
Name		0..1	Text
Description		0..1	Text

Value Domain



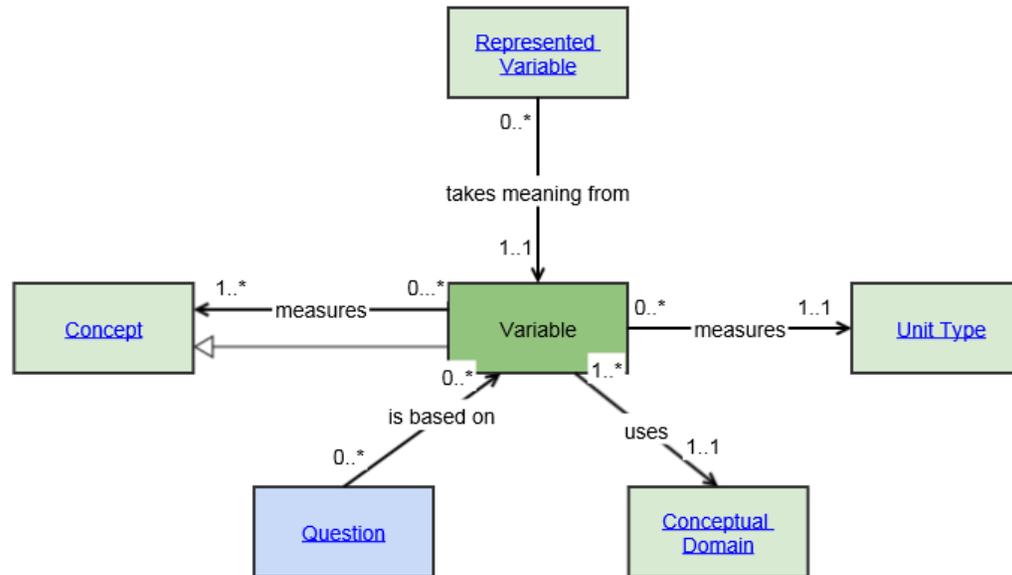
Definition

Object	Group	Definition	Explanatory Text	Synonyms
Value Domain	Concepts	The permitted range of values for a characteristic of a variable	The values can be described by enumeration or by an expression	

Attributes

Name	Description	Cardinality	Value Type
Name		0..1	Text
Description		0..1	Text

Variable



Definition

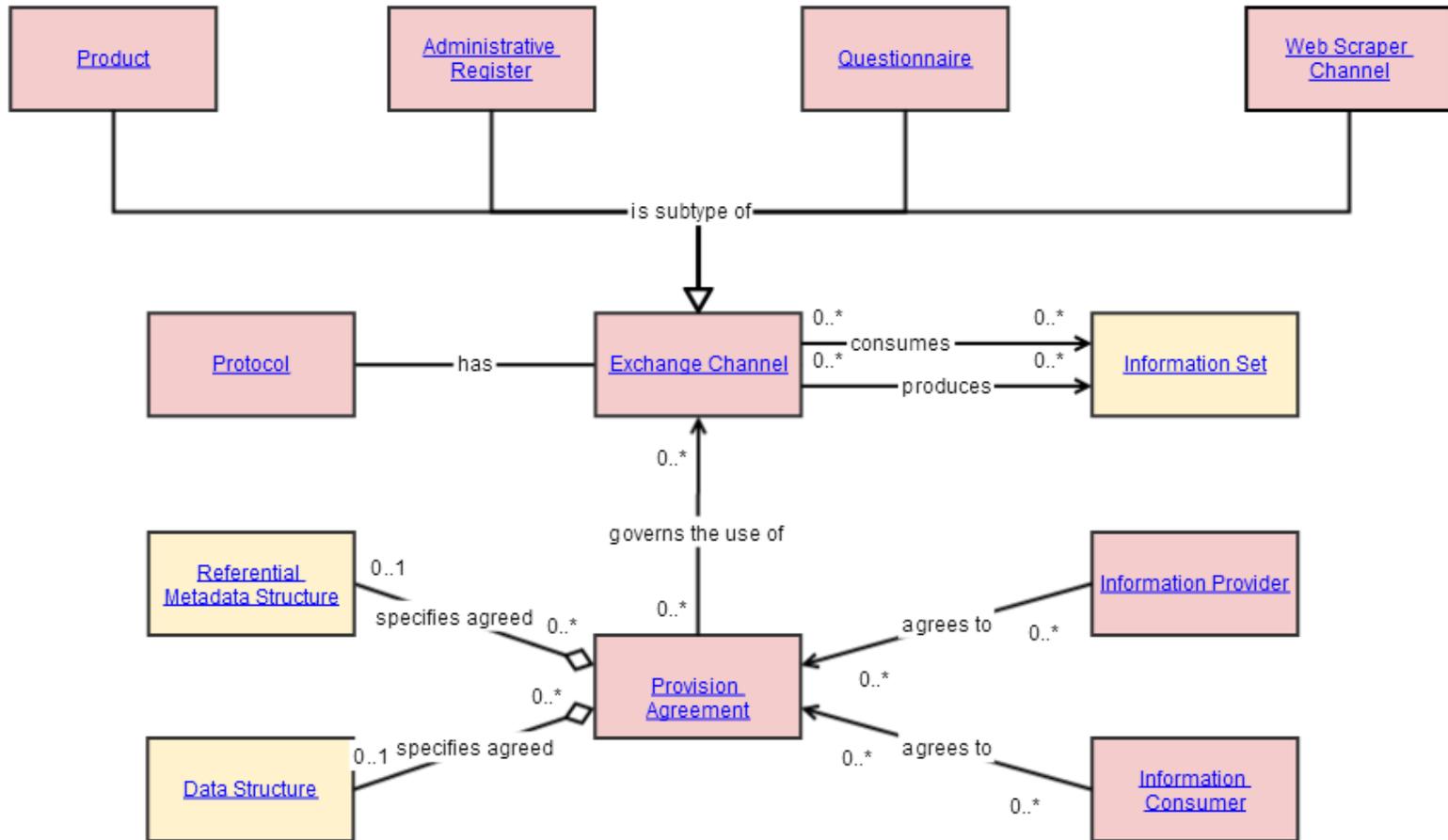
Object	Group	Definition	Explanatory Text	Synonyms
Variable	Concepts	The use of a <i>Concept</i> as a characteristic of a <i>Population</i> intended to be measured	The Variable combines the meaning of a Concept with a Unit Type, to define the characteristic that is to be measured. Here are 3 examples - 1. Sex of person 2. Number of employees 3. Value of endowment	

Attributes

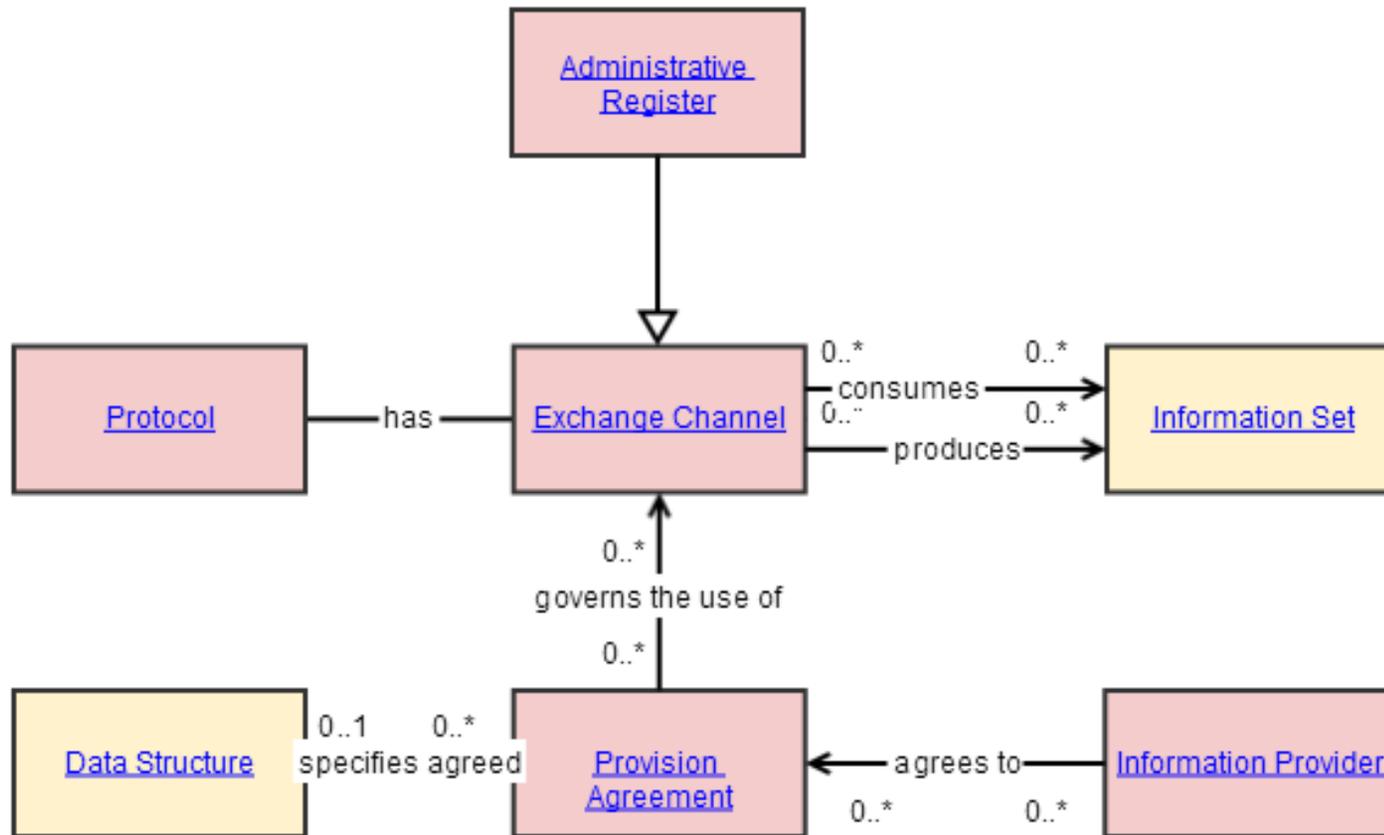
Name	Description	Cardinality	Value Type
Name		0..1	Text
Description		0..1	Text

Exchange Group

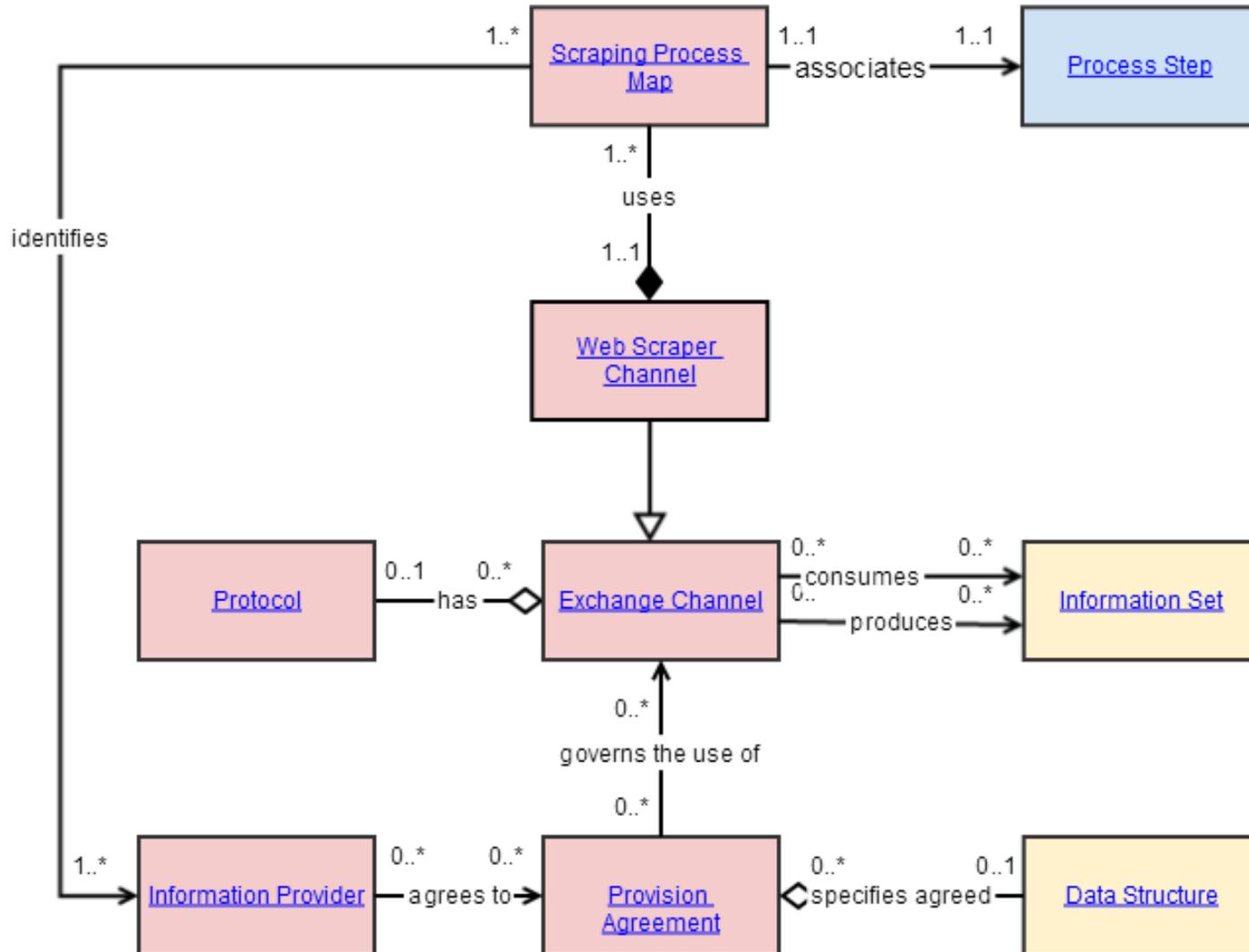
Exchange Channel



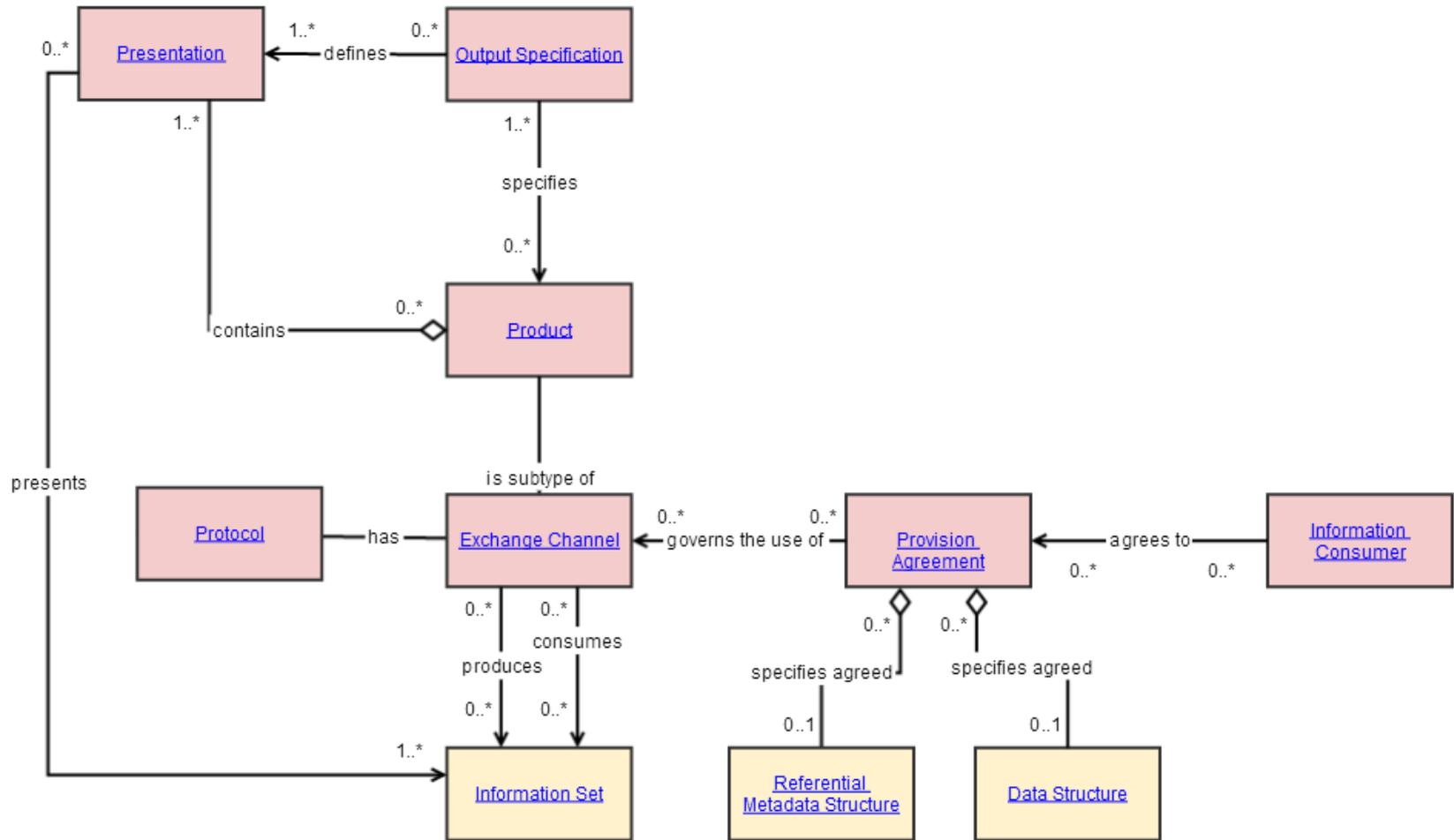
Administrative Register



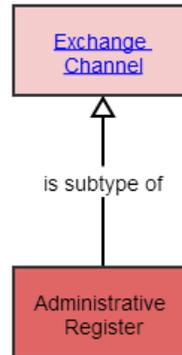
Web Scraping Channel



Disseminations



Administrative Register



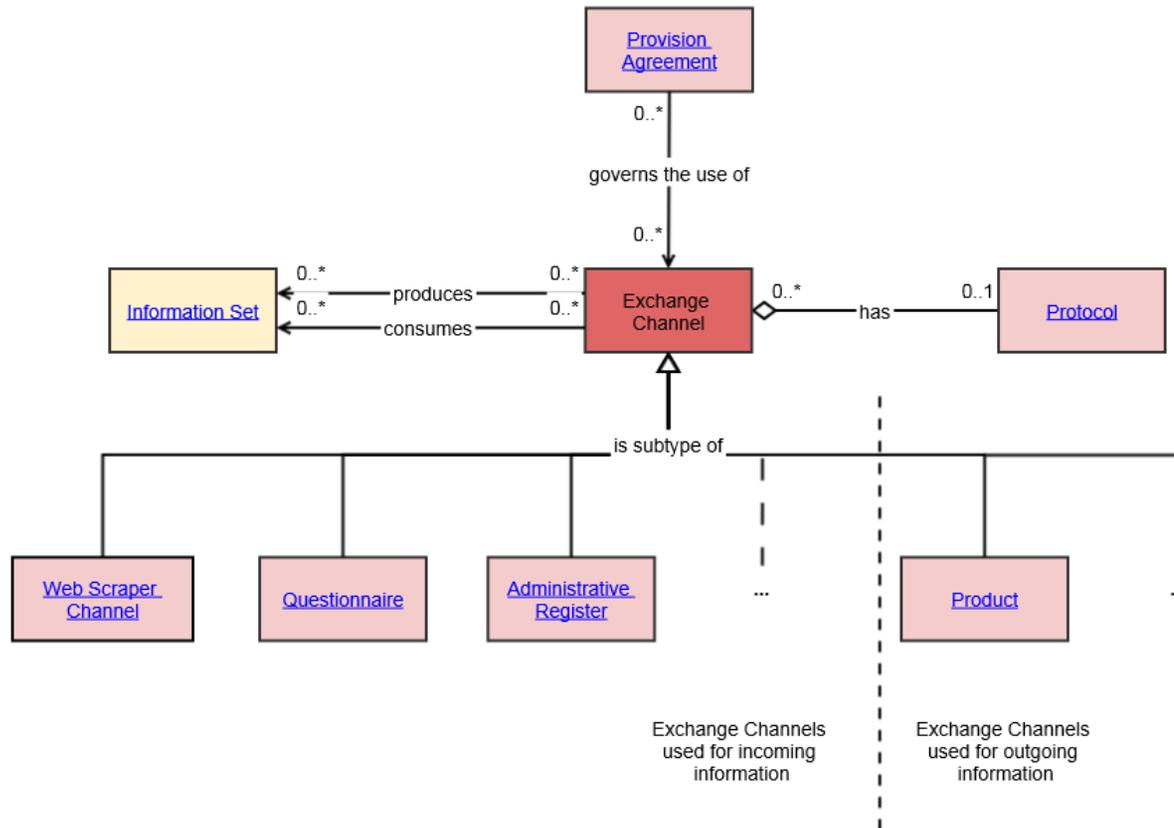
Definition

Object	Group	Definition	Explanatory Text	Synonyms
Administrative Register	Exchange	A source of administrative information which is obtained from an external organization (or sometimes from another department of the same organization)	The <i>Administrative Register</i> is a source of administrative information obtained from external organizations. The <i>Administrative Register</i> would be provided under a <i>Provision Agreement</i> with the supplying organization. This administrative information is usually collected for an organization's operational purposes, rather than for statistical purposes.	

Attributes

Name	Description	Cardinality	Value Type
Name		1..1	Text
Description		0..1	Text
SupplierIdentifier	An identifier for the supplier of the Administrative Register	0..1	Text

Exchange Channel



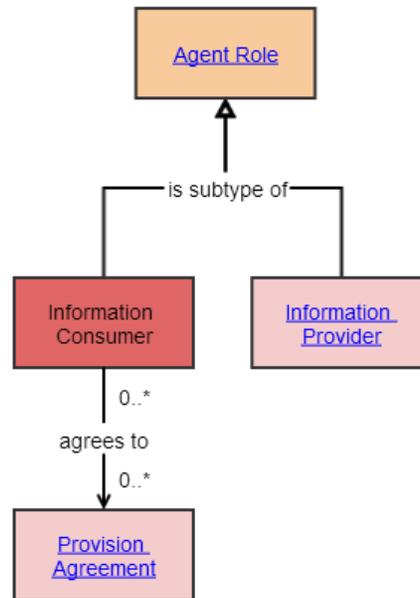
Definition

Object	Group	Definition	Explanatory Text	Synonyms
Exchange Channel	Exchange	A means of exchanging data.	An abstract object that describes the means to receive (data collection) or send (dissemination) information. Different <i>Exchange Channels</i> are used for collection and dissemination. Examples of collection <i>Exchange Channel</i> include <i>Questionnaire</i> , <i>Web Scraper Channel</i> and <i>Administrative Register</i> . The only example of a dissemination <i>Exchange Channel</i> currently contained in GSIM is <i>Product</i> . Additional <i>Exchange Channels</i> can be added to the model as needed by individual organizations.	

Attributes

Name	Description	Cardinality	Value Type
Name		0..1	text
Description		0..1	text
Direction	Direction of the exchange channel: collect or disseminate	1..1	text

Information Consumer



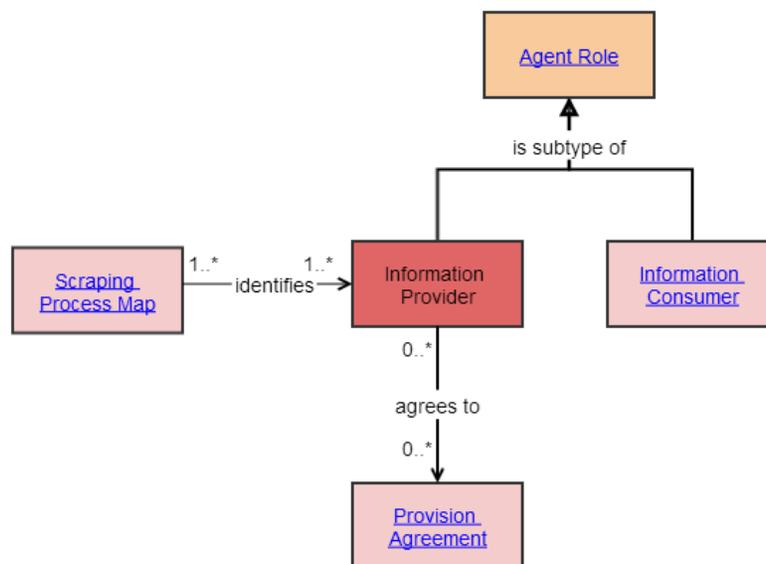
Definition

Object	Group	Definition	Explanatory Text	Synonyms
Information Consumer	Exchange	A person or organization that consumes disseminated data.	The <i>Information Consumer</i> accesses a set of information via a <i>Product</i> (or potentially via another Exchange Channel), which contains one or more <i>Presentations</i> . The <i>Information Consumer's</i> access to the information is subject to a <i>Provision Agreement</i> , which sets out conditions of access.	

Attributes

Name	Description	Cardinality	Value Type
Name		0..1	Text
Description		0..1	Text

Information Provider



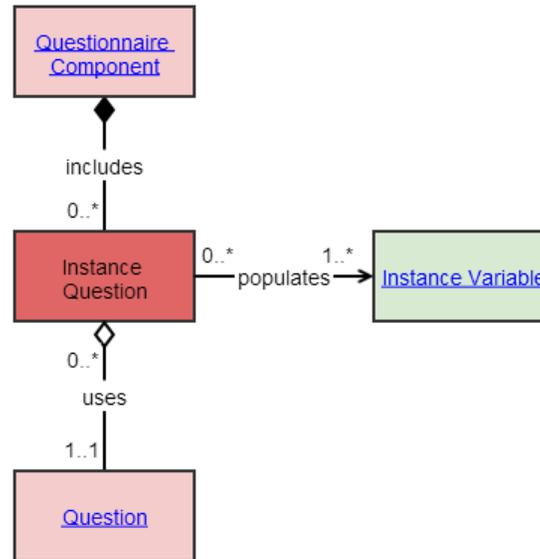
Definition

Object	Group	Definition	Explanatory Text	Synonyms
Information Provider	Exchange	An <i>Individual</i> or <i>Organization</i> that provides collected information.	An <i>Information Provider</i> possesses sets of information (that it has generated, collected, produced, bought or otherwise acquired) and is willing to supply that information (data or referential metadata) to the statistical office. The two parties use a <i>Provision Agreement</i> to agree the <i>Data Structure</i> and <i>Referential Metadata Structure</i> of the data to be exchanged via an <i>Exchange Channel</i> .	information supplier, data supplier

Attributes

Name	Description	Cardinality	Value Type
Name		0..1	Text
Description		0..1	Text

Instance Question



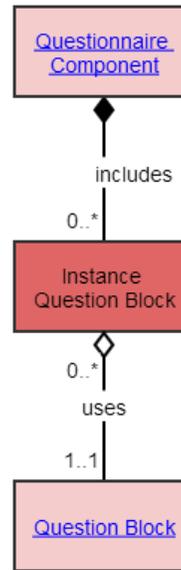
Definition

Object	Group	Definition	Explanatory Text	Synonyms
Instance Question	Exchange	The use of a <i>Question</i> in a particular <i>Questionnaire</i> .	The Instance Question is the use of a <i>Question</i> in a particular <i>Questionnaire Component</i> . This also includes the use of the <i>Question</i> in a <i>Question Block</i> , which is a particular type of <i>Questionnaire Component</i> .	

Attributes

Instance Question inherits all attributes from the Question it is associated with.

Instance Question Block



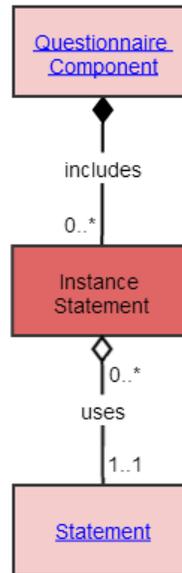
Definition

Object	Group	Definition	Explanatory Text	Synonyms
Instance Question Block	Exchange	The use of a <i>Question Block</i> in a particular <i>Questionnaire</i> .	The <i>Instance Question Block</i> is the use of a <i>Question Block</i> in a particular <i>Questionnaire Component</i> . This also includes the use of a <i>Question Block</i> in another <i>Question Block</i> , as it is a particular type of <i>Questionnaire Component</i> .	

Attributes

Instance Question Block inherits all attributes from the Question Block it is associated with.

Instance Statement



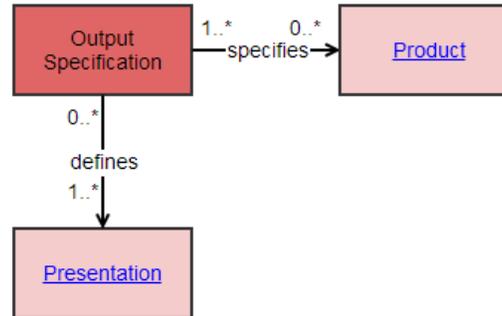
Definition

Object	Group	Definition	Explanatory Text	Synonyms
Instance Statement	Exchange	The use of a <i>Statement</i> in a particular <i>Questionnaire</i> .	The <i>Instance Statement</i> is the use of a <i>Statement</i> in a particular <i>Questionnaire Component</i> . This also includes the use of the <i>Statement</i> in a <i>Question Block</i> , which is a particular type of <i>Questionnaire Component</i> .	

Attributes

Instance Statement inherits all attributes from the Statement it is associated with.

Output Specification



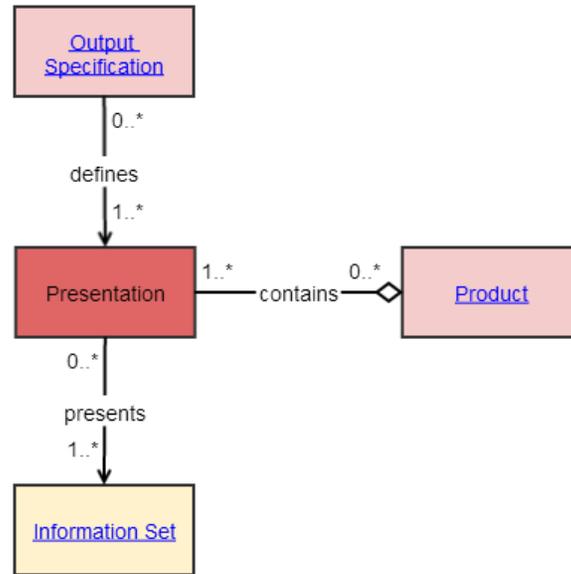
Definition

Object	Group	Definition	Explanatory Text	Synonyms
Output Specification	Exchange	Defines how <i>Information Sets</i> consumed by a <i>Product</i> are presented to <i>Information Consumers</i> .	The <i>Output Specification</i> specifies <i>Products</i> and defines the <i>Presentations</i> they contain. The <i>Output Specification</i> may be fully defined during the design process (such as in a paper publication or a predefined web report), or may be a combination of designed specification supplemented by user selections (such as in an online data query tool).	

Attributes

Name	Description	Cardinality	Value Type
Name		0..1	Text
Description		0..1	Text

Presentation



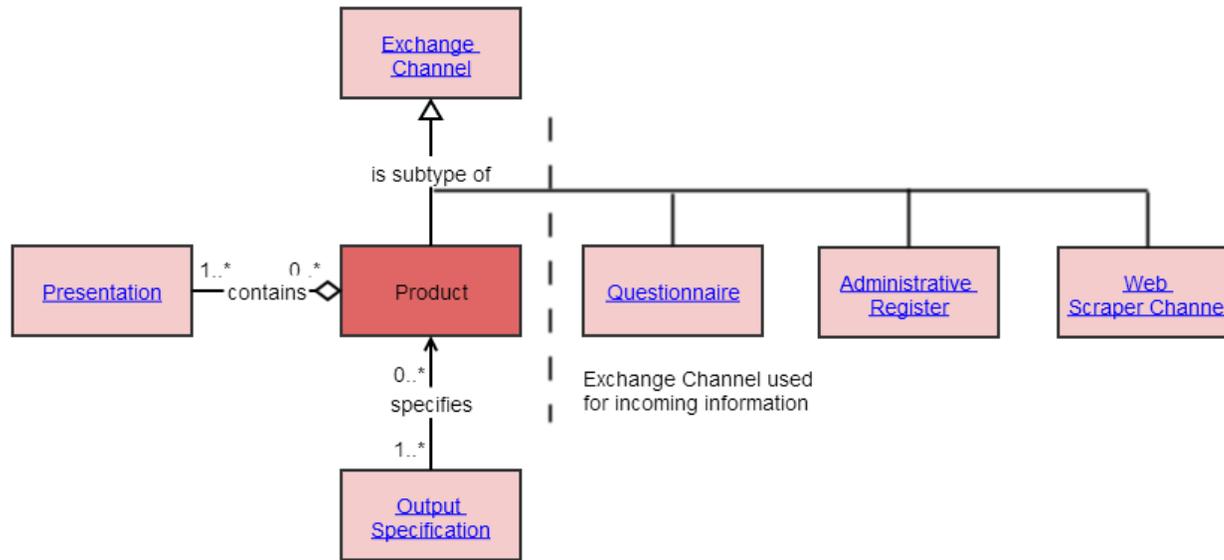
Definition

Object	Group	Definition	Explanatory Text	Synonyms
Presentation	Exchange	The way data and referential metadata are presented in a <i>Product</i> .	<p>A <i>Product</i> has one or more <i>Presentations</i>, which present data and referential metadata from <i>Information Sets</i>. A <i>Presentation</i> is defined by an <i>Output Specification</i>.</p> <p><i>Presentation</i> can be in different forms; e.g. tables, graphs, structured data files.</p> <p>Examples:</p> <ul style="list-style-type: none"> • A table of data. Based on a <i>Data Set</i>, the related <i>Data Structure</i> is used to label the column and row headings for the table. The <i>Data Set</i> is used to populate the cells in the table. Reference metadata is used to populate footnotes and cell notes on the table. Confidentiality rules are applied to the <i>Data Set</i> to suppress any disclosive cells. • A data file based on a standard (e.g. SDMX). • A PDF document describing a <i>Classification</i>. • Any structural metadata object expressed in a standard format (e.g. DDI 3.1 XML). • A list of <i>Products</i> or services (e.g. a product catalogue or a web services description language (WSDL) file). • A web page containing <i>Classifications</i>, descriptions of <i>Variables</i>, etc. 	

Attributes

Name	Description	Cardinality	Value Type
Name		0..1	Text
Description		0..1	Text

Product



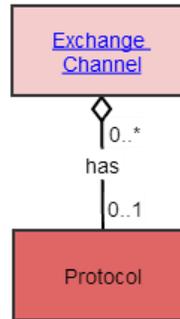
Definition

Object	Group	Definition	Explanatory Text	Synonyms
Product	Exchange	A package of content that can be disseminated as a whole.	<p>A <i>Product</i> is the only defined type of <i>Exchange Channel</i> for outgoing information. A <i>Product</i> packages <i>Presentations of Information Sets</i> for an <i>Information Consumer</i>. The <i>Product</i> and its <i>Presentations</i> are generated according to <i>Output Specifications</i>, which define how the information from the <i>Information Sets</i> it consumes are presented to the <i>Information Consumer</i>. The <i>Protocol</i> for a <i>Product</i> determines the mechanism by which the <i>Product</i> is disseminated (e.g. website, SDMX web service, paper publication).</p> <p>A <i>Provision Agreement</i> between the statistics office and the <i>Information Consumer</i> governs the use of a <i>Product</i> by the <i>Information Consumer</i>. The <i>Provision Agreement</i>, which may be explicitly or implicitly agreed, provides the legal or other basis by which the two parties agree to exchange data. In many cases, dissemination <i>Provision Agreements</i> are implicit in the terms of use published by the statistics office.</p> <p>For static <i>Products</i> (e.g. paper publications), specifications are predetermined. For dynamic products, aspects of specification could be determined by the <i>Information Consumer</i> at run time. Both cases result in <i>Output Specifications</i> specifying <i>Information Set</i> data or referential metadata that will be included in each <i>Presentation</i> within the <i>Product</i>.</p>	

Attributes

Name	Description	Cardinality	Value Type
Name		0..1	Text
Description		0..1	Text

Protocol



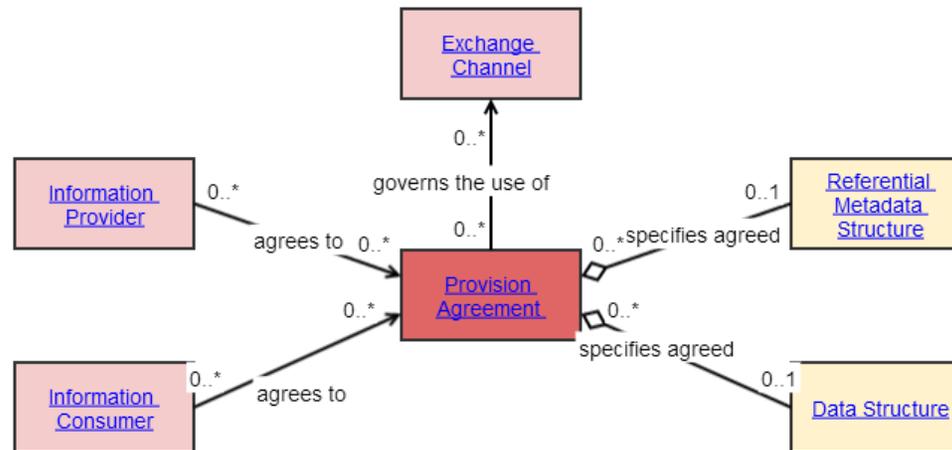
Definition

Object	Group	Definition	Explanatory Text	Synonyms
Protocol	Exchange	The mechanism for exchanging information through an <i>Exchange Channel</i> .	A Protocol specifies the mechanism (e.g. SDMX web service, data file exchange, web robot, face to face interview, mailed paper form) of exchanging information through an <i>Exchange Channel</i> .	

Attributes

Name	Description	Cardinality	Value Type
Name		0..1	Text
Description		0..1	Text

Provision Agreement



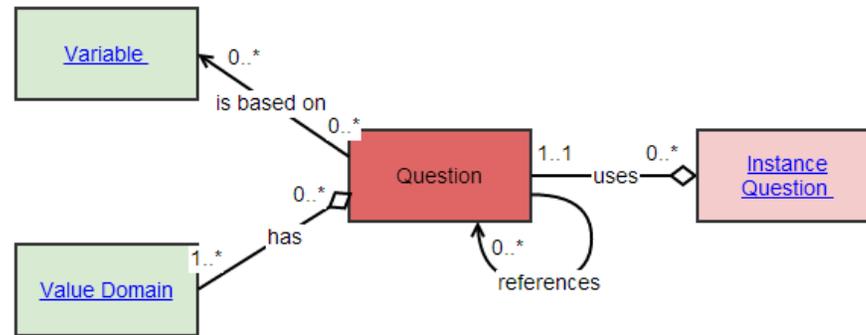
Definition

Object	Group	Definition	Explanatory Text	Synonyms
Provision Agreement	Exchange	The legal or other basis by which two parties agree to exchange data.	A <i>Provision Agreement</i> between the statistical organization and the <i>Information Provider</i> (collection) or the <i>Information Consumer</i> (dissemination) governs the use of <i>Exchange Channels</i> . The <i>Provision Agreement</i> , which may be explicitly or implicitly agreed, provides the legal or other basis by which the two parties agree to exchange data. The parties also use the <i>Provision Agreement</i> to agree the <i>Data Structure</i> and <i>Referential Metadata Structure</i> of the information to be exchanged.	

Attributes

Name	Description	Cardinality	Value Type
Name		0..1	text
Description		0..1	text

Question



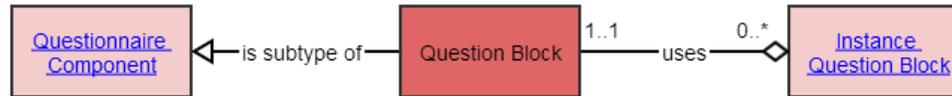
Definition

Object	Group	Definition	Explanatory Text	Synonyms
Question	Exchange	Describes the text used to elicit a response for the <i>Concept</i> to be measured.	<p>A <i>Question</i> may be a single question used to obtain a response, or may be a multiple question, a construct which links multiple sub-questions, each with their own response.</p> <p>A <i>Question</i> also includes a relationship to the <i>Value Domain</i> to document the associated response criteria for the question. A single response question will have one <i>Value Domain</i> associated with it, while a 'multiple question' may have more than one <i>Value Domain</i>.</p> <p>A <i>Question</i> should be designed with re-use in mind, as it can be used in multiple <i>Questionnaires</i>.</p>	Multiple Question

Attributes

Name	Description	Value Type	Cardinality
Name		Text	0..1
Description		Text	0..1
QuestionText	The text which describes the information which is to be obtained	Text	1..1
Question Purpose	A description of the purpose of the question, whether the question has a specific expected function	Text	0..1

Question Block



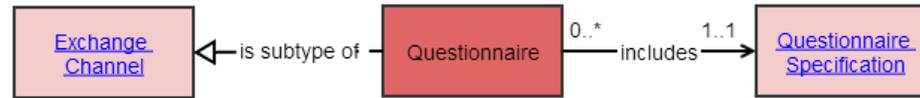
Definition

Object	Group	Definition	Explanatory Text	Synonyms
Question Block	Exchange	A set of <i>Questions, Statements</i> or instructions which are used together.	<p>A <i>Question Block</i> should be designed for reuse, as it can be used in multiple <i>Questionnaires</i>. The <i>Question Block</i> is a type of <i>Questionnaire Component</i>. A statistical organization will often have a number of <i>Question Blocks</i> which they reuse in a number of <i>Questionnaires</i>. Examples of <i>Question Blocks</i> include:</p> <ul style="list-style-type: none"> • Household <i>Question Block</i> • Income <i>Question Block</i> • Employment <i>Question Block</i> 	Question Module

Attributes

Name	Description	Cardinality	Value type
Name		0..1	text
Description		0..1	text

Questionnaire



Definition

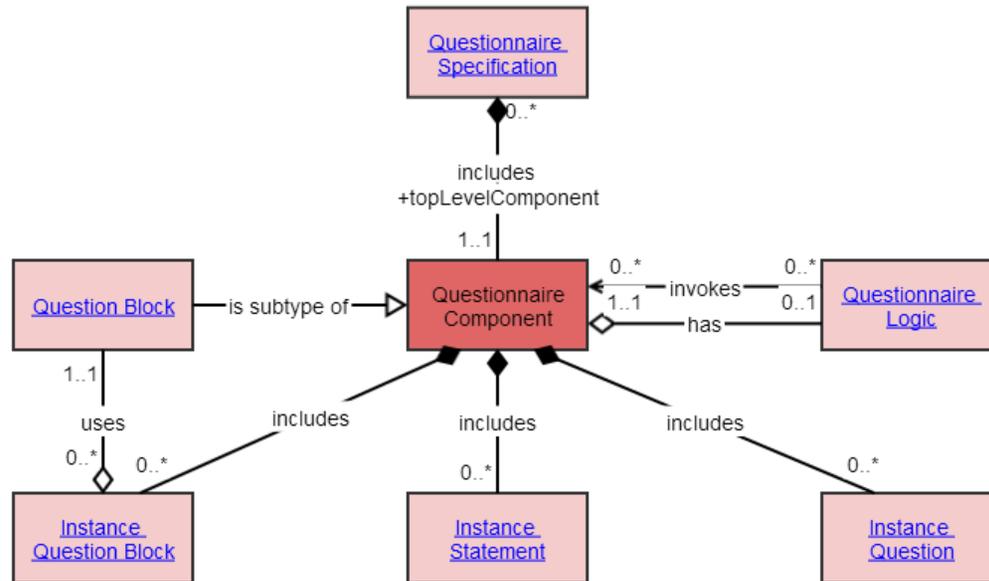
Object	Group	Definition	Explanatory Text	Synonyms
Questionnaire	Exchange	A concrete and usable tool to elicit information from observation <i>units</i> .	This is an example of a way statistical organizations collect information (an exchange channel). Each mode should be interpreted as a new <i>Questionnaire</i> derived from the <i>Questionnaire Specification</i> . The <i>Questionnaire</i> is a subtype of <i>Exchange Channel</i> , as it is a way in which data is obtained.	

Attributes

Name	Description	Cardinality	Value Type
Name		1..1	text
Description	Background information about the questionnaire	0..1	text
DateIssued	Date the questionnaire was created.	1..1	date
DateReplacement	Planned date for the replacement of the questionnaire for other version. It can contain the date in which the new version was issued.	1..1	date
Media	Description of the kind of media conceived for the use of the Questionnaire (printed, electronic, etc.)	1..1	Text
Survey	Information on the survey which the questionnaire will be used by.	0..*	Text
SupportArtifacts	A list of devices, software programs, storage media, gadgets or other tools needed to support the use of the Questionnaire.	0..*	Text

Name	Description	Cardinality	Value Type
DetailDocument	Reference to a document containing details of the implementation of the main elements of the Questionnaire	1..1	Link

Questionnaire Component



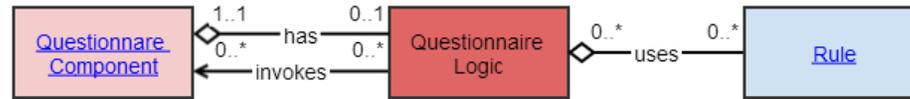
Definition

Object	Group	Definition	Explanatory Text	Synonyms
Questionnaire Component	Exchange	A record of the flow of a <i>Questionnaire Specification</i> and its use of <i>Questions</i> , <i>Question Blocks</i> and <i>Statements</i>	Defines the structure of the <i>Questionnaire Specification</i> , as a combination of <i>Questions</i> , <i>Question Blocks</i> and <i>Statements</i> . It is the object which groups together all the components of a <i>Questionnaire</i> . <i>A Questionnaire Component</i> is recursive, in that it can refer to other <i>Questionnaire Components</i> and accompanying <i>Questionnaire Logic</i> objects at a lower level. It is only at the top level where the <i>Questionnaire Component</i> links to the <i>Questionnaire Specification</i> ,	Question Block

Attributes

Name	Description	Cardinality	Value Type
Name		1..1	Text
Description		0..1	Text
Component Sequence	The order in which instance question, and instance statement appear in the Questionnaire Component	0..*	number

Questionnaire Logic



Definition

Object	Group	Definition	Explanatory Text	Synonyms
Questionnaire Logic	Exchange	Governs the sequence of <i>Questions</i> , <i>Question Blocks</i> and <i>Statements</i> based on factors such as the current location, the response to the previous questions etc., invoking navigation and validation rules to apply.		Routing

Attributes

Name	Description	Cardinality	Value Type
Name		0..1	text
Description		0..1	text
RoutingInformation	Routing information, which will also use responses from Rule.	1..n	text

Questionnaire Specification



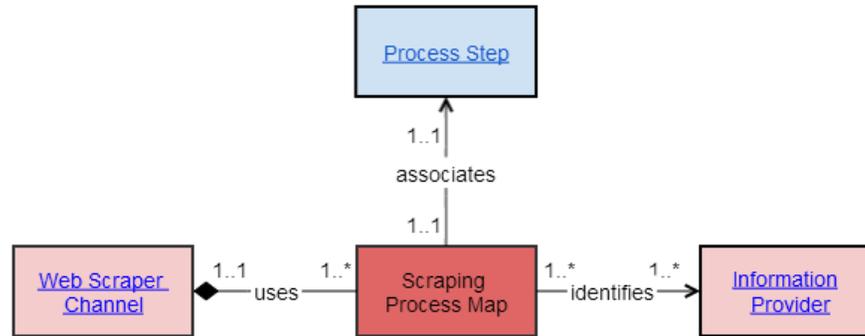
Definition

Object	Group	Definition	Explanatory Text	Synonyms
Questionnaire Specification	Exchange	The tool designed to elicit information from observation <i>Units</i> .	This represents the complete questionnaire design, with a relationship to the top level <i>Questionnaire Component</i> . There may be many different <i>Questionnaire Specifications</i> , for the same surveys, or tailored to individual observation <i>Units</i> (respondents) so that there would be a different <i>Questionnaire Specification</i> for each respondent. The design would also differ depending upon the specific mode of collection the <i>Questionnaire</i> is designed for.	

Attributes

Name	Description	Cardinality	Value Type
Name		1..1	Text
Description		0..1	Text
Valid From	The date from which the object can be used	0..1	Date
Valid To	The date after which the object cannot be used	0..1	Date

Scraping Process Map



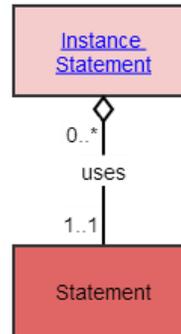
Definition

Object	Group	Definition	Explanatory Text	Synonyms
Scraping Process Map	Exchange	Maps a web scraping process to a specific website.	<i>Scraping Process Map</i> is an essential element of the <i>Web Scraper Channel</i> . The process being mapped can be a <i>Business Service</i> or a <i>Process Step</i> .	

Attributes

Name	Description	Cardinality	Value Type
Name		0..1	text
Description		0..1	text
Date_Issued	Date the mapping was issued.	1..1	date
Date_Replacement	Date the mapping was replaced.	0..1	date

Statement



Definition

Object	Group	Definition	Explanatory Text	Synonyms
Statement	Exchange	A report of facts in a <i>Questionnaire</i>	<p><i>Statements</i> are often included to provide further explanation to respondents. Example:</p> <p>"The following questions are about your health".</p> <p>The object is also used to represent completion instructions for the interviewer or respondent.</p> <p><i>Statement</i> should be designed with re-use in mind as it can be used in numerous <i>Questionnaires</i>.</p>	Interviewer Instruction Instruction

Attributes

Name	Description	Cardinality	Value type
Name		Text	0..1
Description		Text	0..1
Statement Text	The information, note, fact or instruction text making up the statement	0..1	Text

Web Scraper Channel



Definition

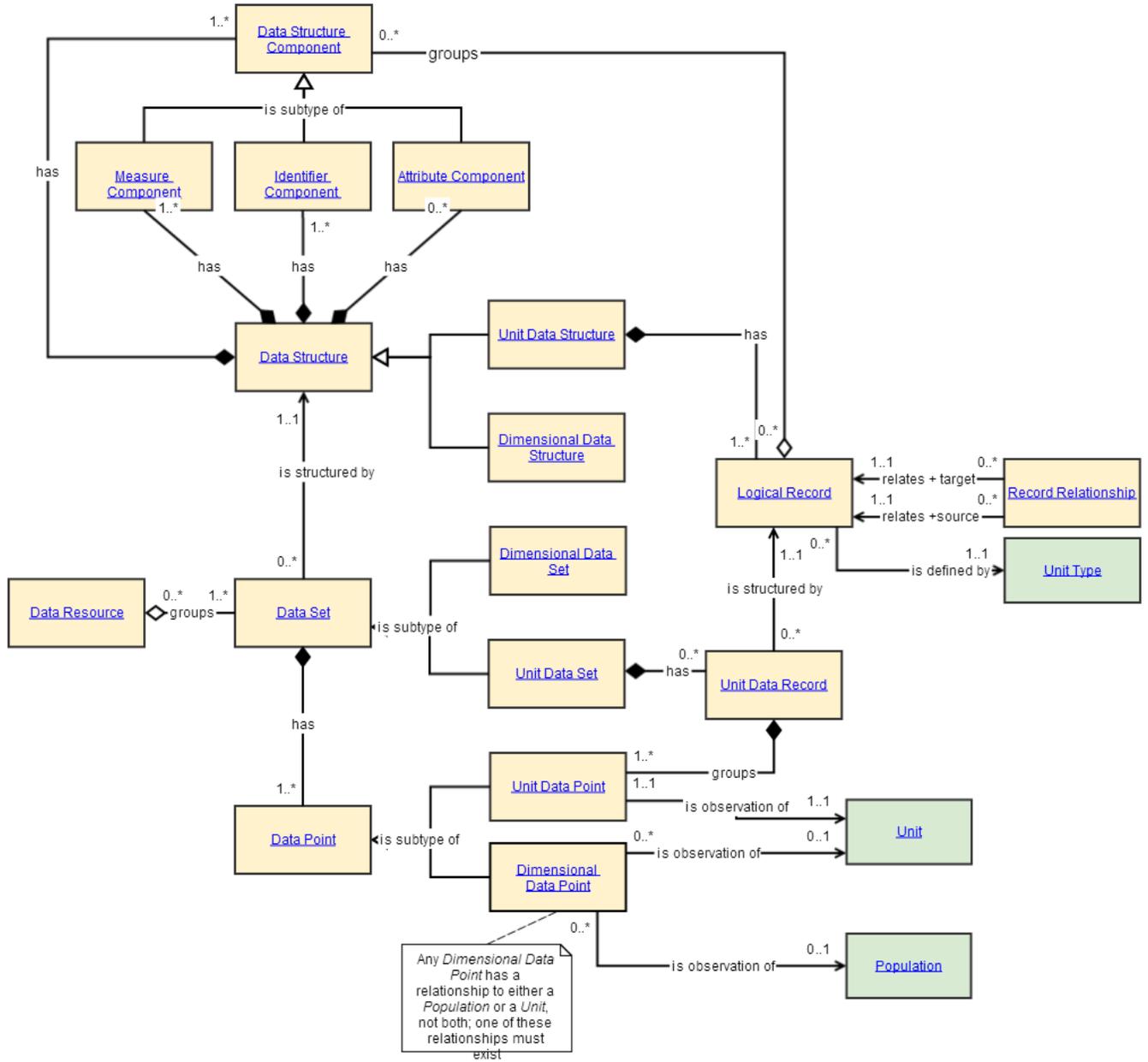
Object	Group	Definition	Explanatory Text	Synonyms
Web Scraper Channel	Exchange	A concrete and usable tool to gather information from the Internet.	This is an example of a way statistical organizations collect information (an <i>Exchange Channel</i>). The <i>Web Scraper Channel</i> contains <i>Scraping Process Maps</i> , which map the channel to each website targeted for scraping.	

Attributes

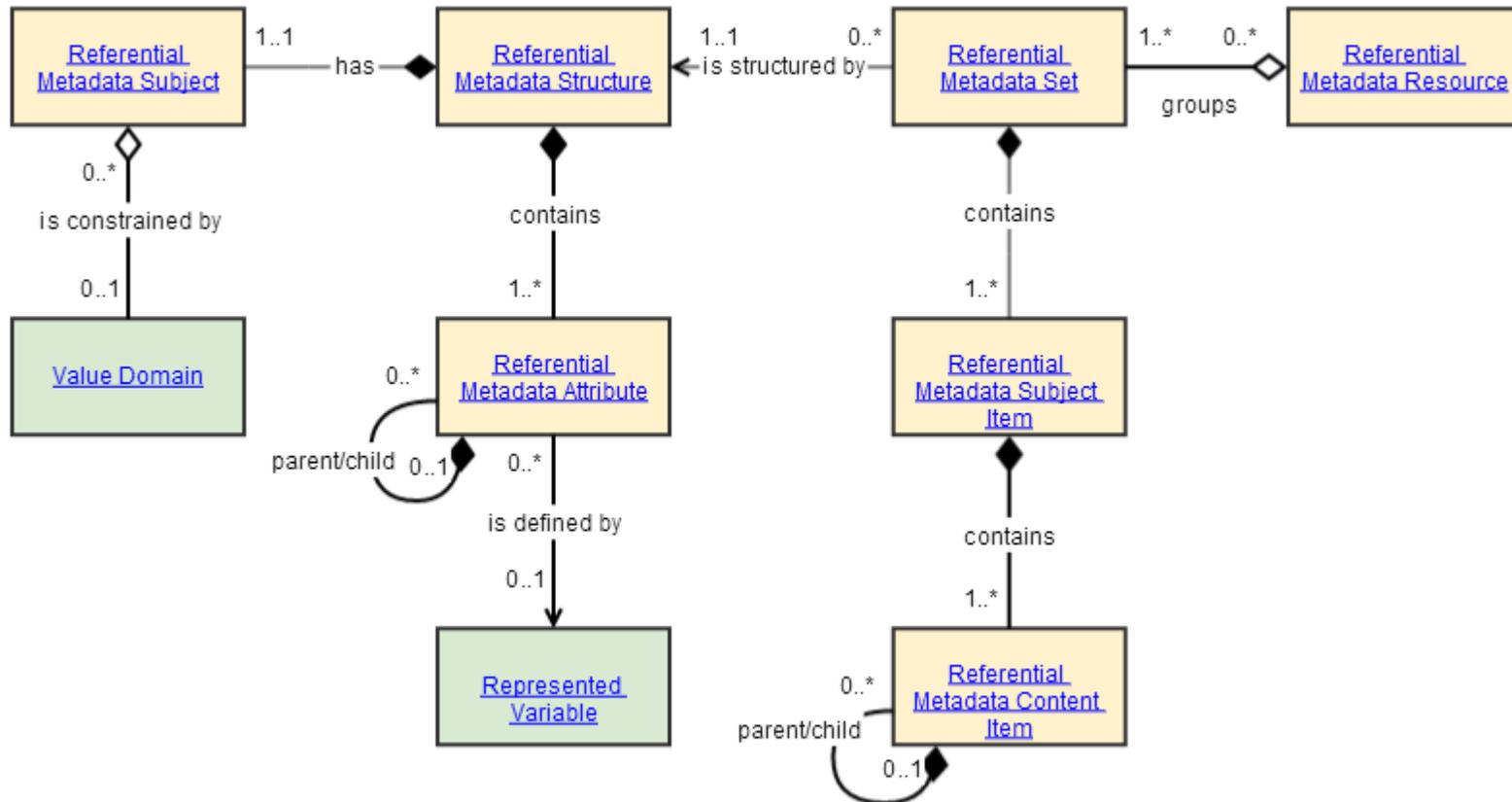
Name	Description	Cardinality	Value Type
Name		0..1	text
Description		0..1	text

Structures Groups

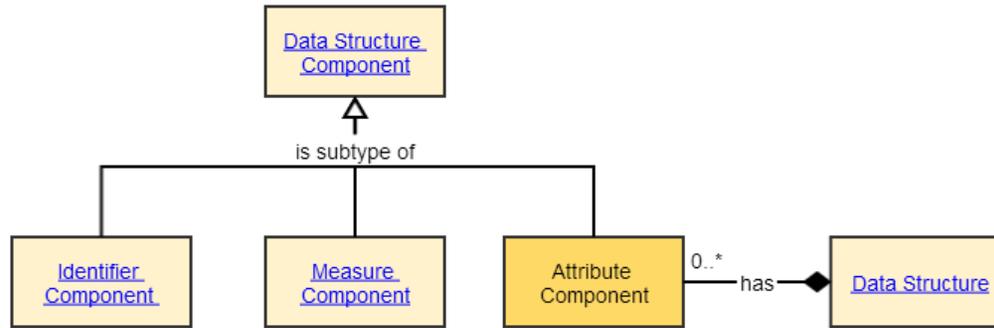
Data Sets



Referential Metadata Sets



Attribute Component



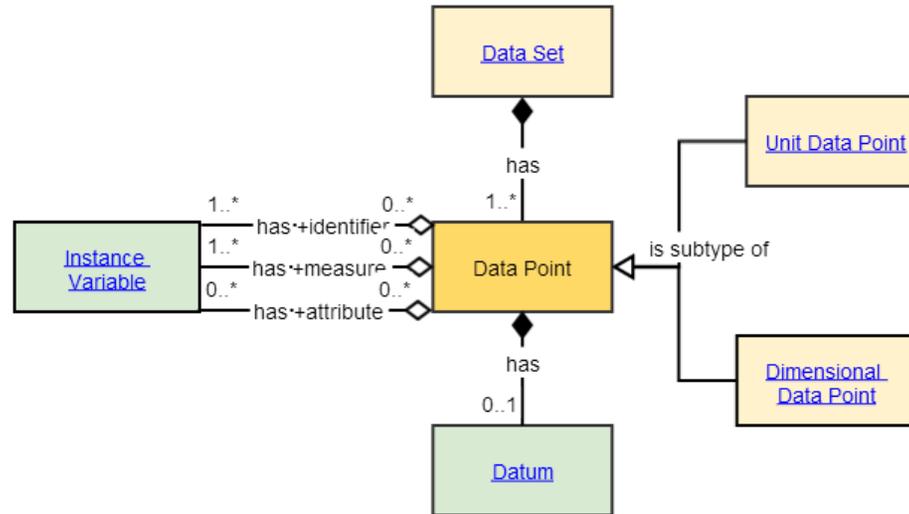
Definition

Object	Group	Definition	Explanatory Text	Synonyms
Attribute Component	Structures	The role given to a <i>Represented Variable</i> in the context of a <i>Data Structure</i> , which supplies information other than identification or measures.	For example the publication status of an observation (e.g. provisional, final, revised)	

Attributes

Name	Description	Cardinality	Value Type
Name		0..1	Text
Description		0..1	Text

Data Point



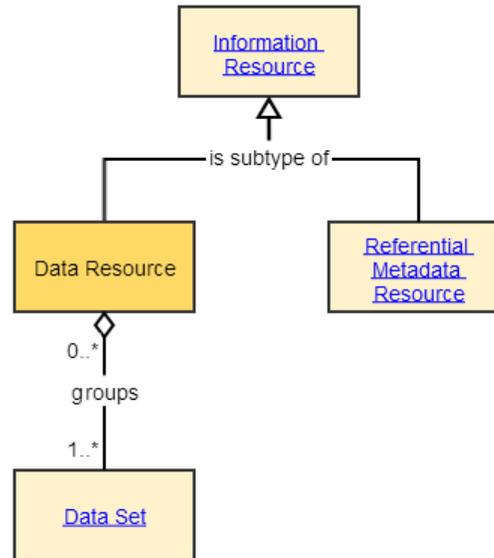
Definition

Object	Group	Definition	Explanatory Text	Synonyms
Data Point	Structures	A placeholder (or cell) for the value of an <i>Instance Variable</i>	Field in a <i>Data Structure</i> which corresponds to a cell in a table. The <i>Data Point</i> is structural and distinct from the value (the <i>Datum</i>) that it holds.	

Attributes

Name	Description	Cardinality	Value Type
Name		0..1	Text
Description		0..1	Text

Data Resource



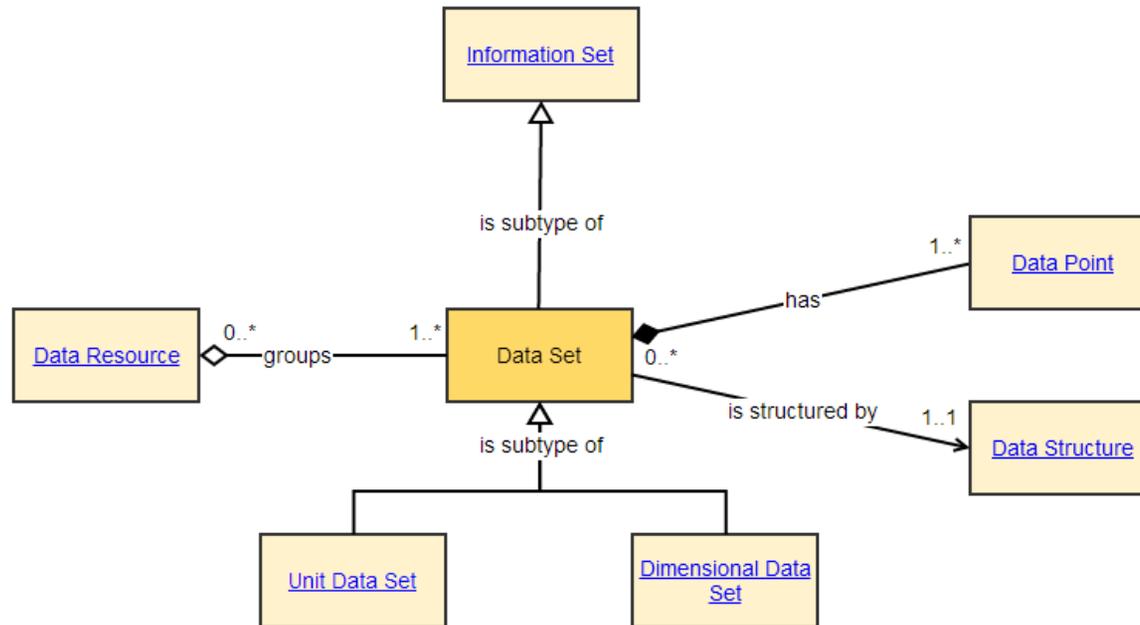
Definition

Object	Group	Definition	Explanatory Text	Synonyms
Data Resource	Structures	An organized collection of stored information made of one or more <i>Data Sets</i> .	<i>Data Resources</i> are collections of data that are used by a statistical activity to produce information. <i>Data Resource</i> is a specialization of an <i>Information Resource</i> .	data source

Attributes

Name	Description	Cardinality	Value Type
Name		1..1	Text
Description		1..1	Text

Data Set



Definition

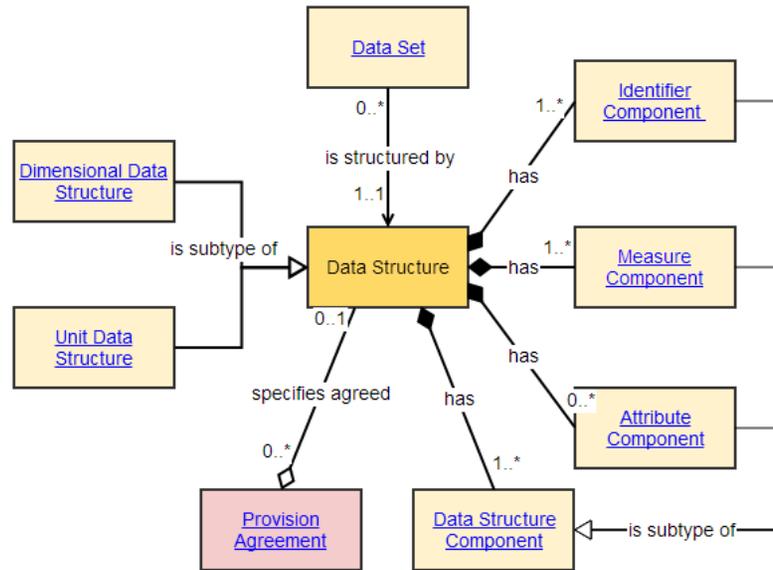
Object	Group	Definition	Explanatory Text	Synonyms
Data Set	Structures	An organized collection of data.	Examples of <i>Data Sets</i> could be observation registers, time series, longitudinal data, survey data, rectangular data sets, event-history data, tables, data tables, cubes, registers, hypercubes, and matrixes. A broader term for <i>Data Set</i> could be data. A narrower term for <i>Data Set</i> could be data element, data record, cell, field.	database, data file, file, table

Attributes

Name	Description	Cardinality	Value Type
Name		0..1	Text
Description		0..1	Text

Name	Description	Cardinality	Value Type
Name	A human-readable identifier for the object	0..1	Text
Description	A human-readable description of the object	0..1	Text

Data Structure



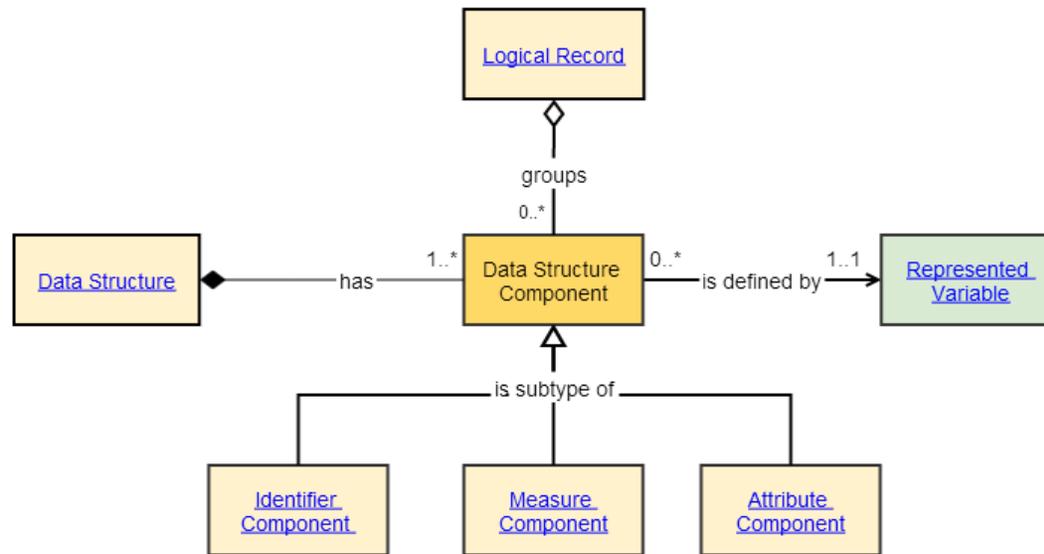
Definition

Object	Group	Definition	Explanatory Text	Synonyms
Data Structure	Structures	Defines the structure of an organized collection of data (<i>Data Set</i>).	The structure is described using <i>Data Structure Components</i> that can be either <i>Attribute Components</i> , <i>Identifier Components</i> or <i>Measure Components</i> . Examples for unit data include social security number, country of residence, age, citizenship, country of birth, where the social security number and the country of residence are both identifying components and the others are measured variables obtained directly or indirectly from the person (<i>Unit</i>).	

Attributes

Name	Description	Cardinality	Value Type
Name		0..1	Text
Description		0..1	Text

Data Structure Component



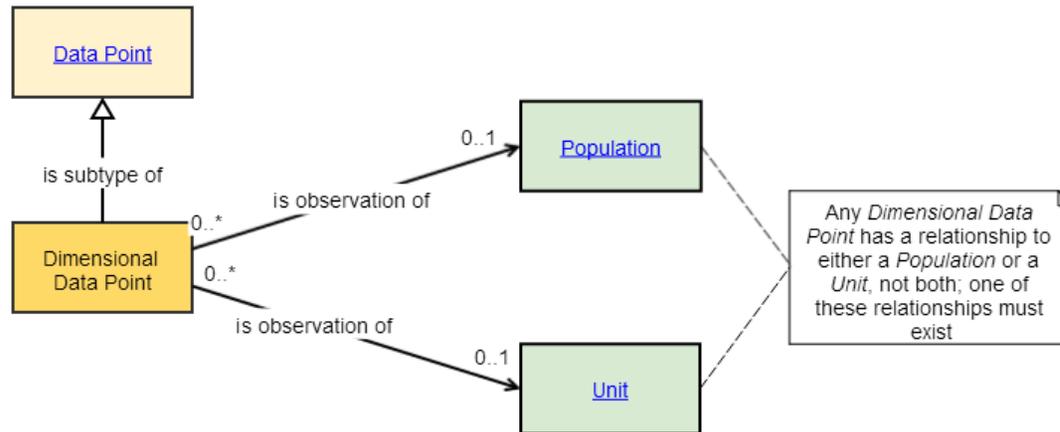
Definition

Object	Group	Definition	Explanatory Text	Synonyms
Data Structure Component	Structures	The role of the <i>Represented Variable</i> in the context of a <i>Data Structure</i> .	<p>A <i>Data Structure Component</i> can be an <i>Attribute Component</i>, <i>Measure Component</i> or an <i>Identifier Component</i>.</p> <p>Example of <i>Attribute Component</i>: The publication status of an observation such as provisional, revised.</p> <p>Example of <i>Measure Component</i>: age and height of a person in a <i>Unit Data Set</i> or number of citizens and number of households in a country in a <i>Data Set</i> for multiple countries (<i>Dimensional Data Set</i>).</p> <p>Example of <i>Identifier Component</i>: The personal identification number of a Swedish citizen for unit data or the name of a country in the European Union for dimensional data.</p>	

Attributes

Name	Description	Cardinality	Value Type
Name		0..1	Text
Description		0..1	Text

Dimensional Data Point



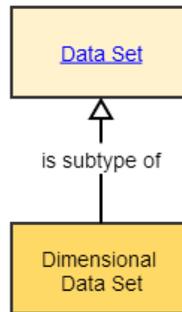
Definition

Object	Group	Definition	Explanatory Text	Synonyms
Dimensional Data Point	Structures	A placeholder (or cell) for the value of an <i>Instance Variable</i> with respect to either a <i>Unit</i> or <i>Population</i> .	A <i>Dimensional Data Point</i> is uniquely identified by the combination of exactly one value for each of the dimensions (<i>Identifier Component</i>) and one measure (<i>Measure Component</i>). There may be multiple values for the same <i>Dimensional Data Point</i> that is for the same combination of dimension values and the same measure. The different values represent different versions of the data in the <i>Data Point</i> . Values are only distinguished on the basis of quality, date/time of measurement or calculation, status, etc. This is handled through the mechanisms provided by the <i>Datum</i> information object.	cell

Attributes

Name	Description	Cardinality	Value Type
Name		0..1	Text
Description		0..1	Text

Dimensional Data Set



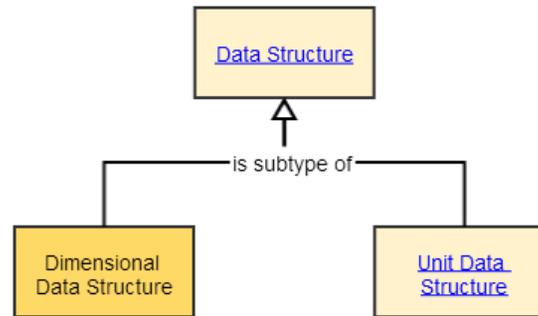
Definition

Object	Group	Definition	Explanatory Text	Synonyms
Dimensional Data Set	Structures	A collection of dimensional data that conforms to a known structure.		

Attributes

Name	Description	Cardinality	Value Type
Name		0..1	Text
Description		0..1	Text

Dimensional Data Structure



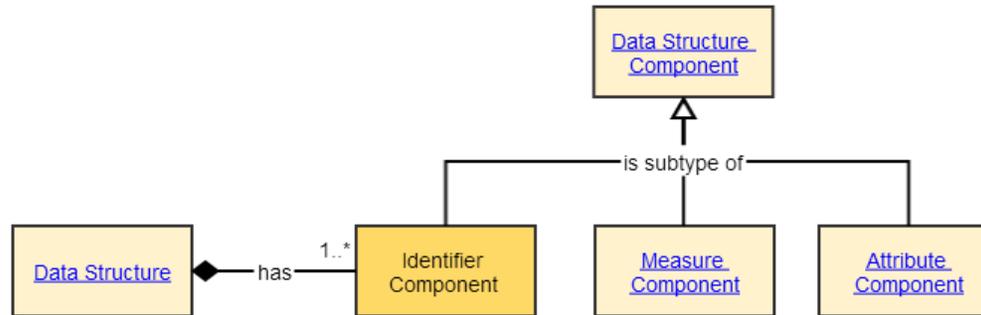
Definition

Object	Group	Definition	Explanatory Text	Synonyms
Dimensional Data Structure	Structures	Describes the structure of a <i>Dimensional Data Set</i> .	For example (city, average income, total population) where the city is the <i>Identifier Component</i> and the others are measured variables.	

Attributes

Name	Description	Cardinality	Value Type
Name		0..1	Text
Description		0..1	Text

Identifier Component



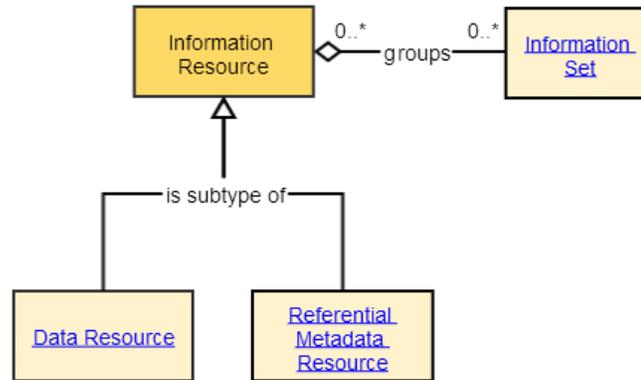
Definition

Object	Group	Definition	Explanatory Text	Synonyms
Identifier Component	Structures	The role given to a <i>Represented Variable</i> in the context of a <i>Data Structure</i> to identify the unit in an organized collection of data.	An <i>Identifier Component</i> is a sub-type of <i>Data Structure Component</i> . The personal identification number of a Swedish citizen for unit data or the name of a country in the European Union for dimensional data.	

Attributes

Name	Description	Cardinality	Value Type
Name		0..1	Text
Description		0..1	Text

Information Resource



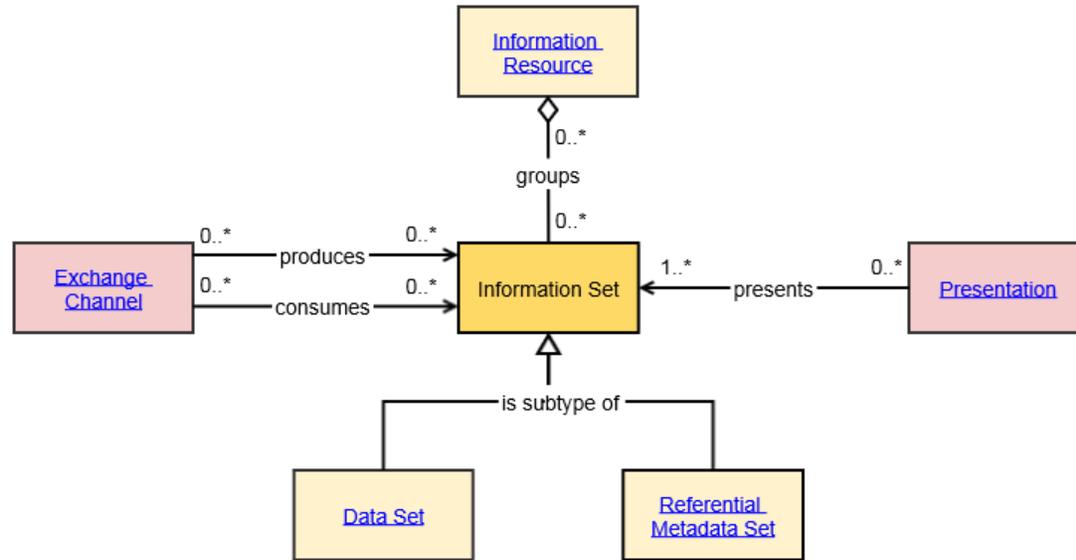
Definition

Object	Group	Definition	Explanatory Text	Synonyms
Information Resource	Structures	An abstract notion that is any organized collection of information.	There currently are only two concrete sub classes: <i>Data Resource</i> and <i>Referential Metadata Resource</i> . The <i>Information Resource</i> allows the model to be extended to other types of resource.	

Attributes

Name	Description	Cardinality	Value Type
Name		0..1	Text
Description		0..1	Text
owner	Identification of the person, institution or group which owns the information resource	0..1	Text
location	A description of the location where the data resource can be found, it could be a physical address or a logical address (like an URI)	1..*	Text

Information Set



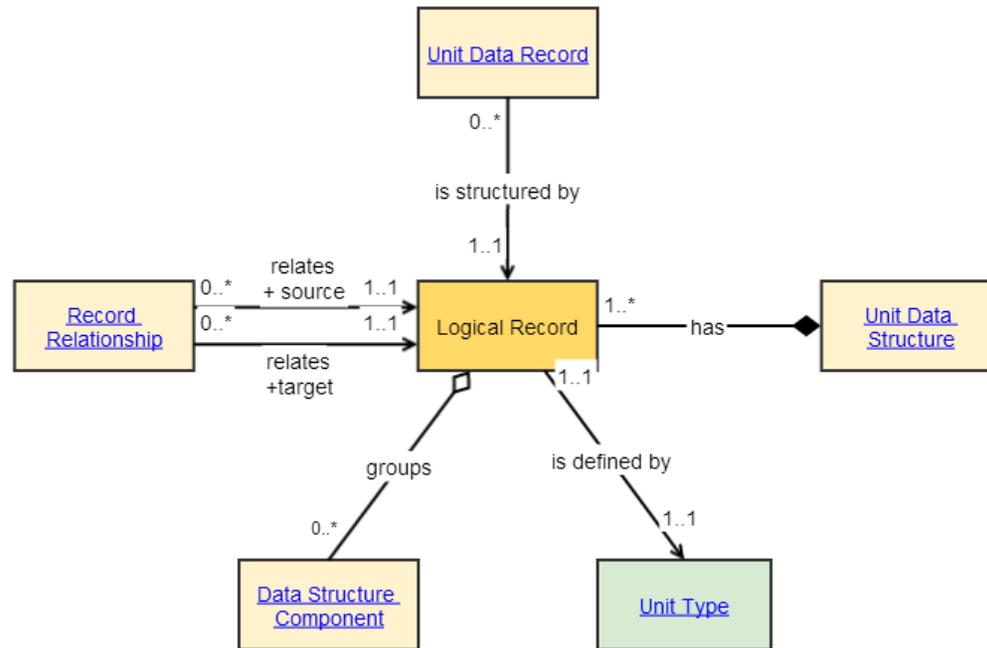
Definition

Object	Group	Definition	Explanatory Text	Synonyms
Information Set	Structures	Organized collections of statistical content.	Statistical organizations collect, process, analyze and disseminate <i>Information Sets</i> , which contain data (<i>Data Sets</i>), referential metadata (<i>Referential Metadata Sets</i>), or potentially other types of statistical content, which could be included in addition types of <i>Information Set</i> .	

Attributes

Name	Description	Cardinality	Value Type
Name		0..1	Text
Description		0..1	Text

Logical Record



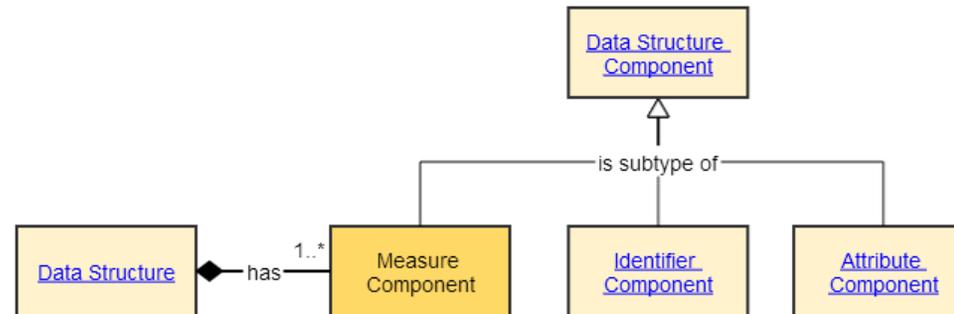
Definition

Object	Group	Definition	Explanatory Text	Synonyms
Logical Record	Structures	Describes a type of <i>Unit Data Record</i> for one <i>Unit Type</i> within a <i>Unit Data Set</i> .	Examples: household, person or dwelling record.	

Attributes

Name	Description	Cardinality	Value Type
Name		0..1	Text
Description		0..1	Text

Measure Component



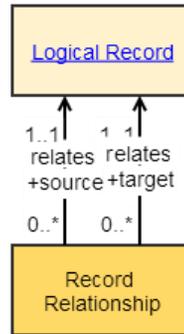
Definition

Object	Group	Definition	Explanatory Text	Synonyms
Measure Component	Structures	The role given to a <i>Represented Variable</i> in the context of a <i>Data Structure</i> to hold the observed/derived values for a particular <i>Unit</i> in an organized collection of data.	A <i>Measure Component</i> is a sub-type of <i>Data Structure Component</i> . For example age and height of a person in a <i>Unit Data Set</i> or number of citizens and number of households in a country in a <i>Data Set</i> for multiple countries (<i>Dimensional Data Set</i>).	

Attributes

Name	Description	Cardinality	Value Type
Name		0..1	Text
Description		0..1	Text

Record Relationship



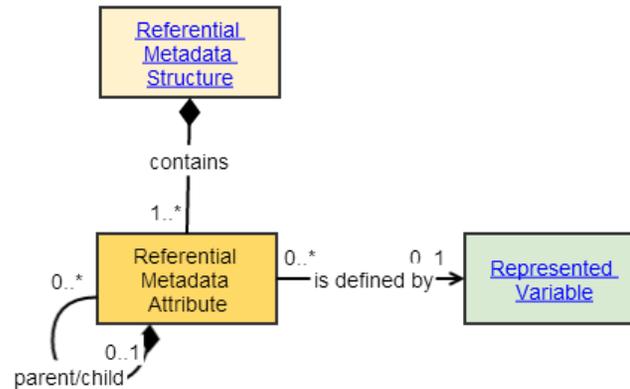
Definition

Object	Group	Definition	Explanatory Text	Synonyms
Record Relationship	Structures	Describes relationships between <i>Logical Records</i> within a <i>Unit Data Structure</i> . It must have both a source <i>Logical Record</i> and a target <i>Logical Record</i> in order to define the relationship.	Example: Relationship between person and household <i>Logical Records</i> within a <i>Unit Data Set</i> .	

Attributes

Name	Description	Cardinality	Value Type
Name		0..1	Text
Description		0..1	Text

Referential Metadata Attribute



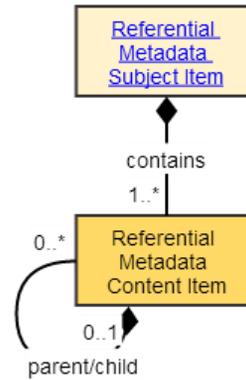
Definition

Object	Group	Definition	Explanatory Text	Synonyms
Referential Metadata Attribute	Structures	The role given to a <i>Represented Variable</i> to supply information in the context of a <i>Referential Metadata Structure</i> .		

Attributes

Name	Description	Cardinality	Value Domain
Name		0..1	Text
Description		0..1	Text
isContainer	Boolean indicating whether or not this attribute actually will contain a value when reported in a Metadata Set	0..1	Binary

Referential Metadata Content Item



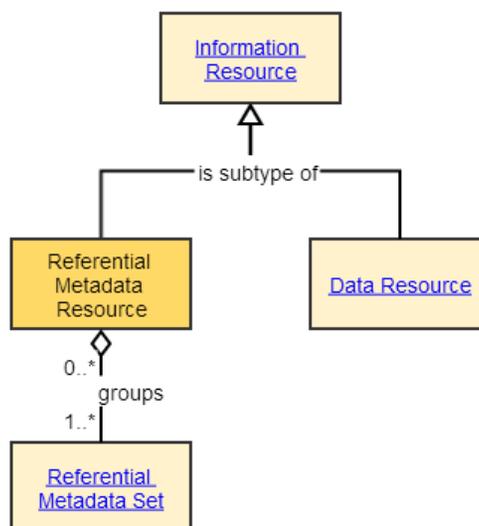
Definition

Object	Group	Definition	Explanatory Text	Synonyms
Referential Metadata Content Item	Structures	The content describing a particular characteristic of a <i>Referential Metadata Subject</i> .	A <i>Referential Metadata Content Item</i> contains the actual content describing a particular characteristic of a <i>Referential Metadata Subject</i> .	

Attributes

Name	Description	Cardinality	Value Type
Name		0..1	Text
Description		0..1	Text

Referential Metadata Resource



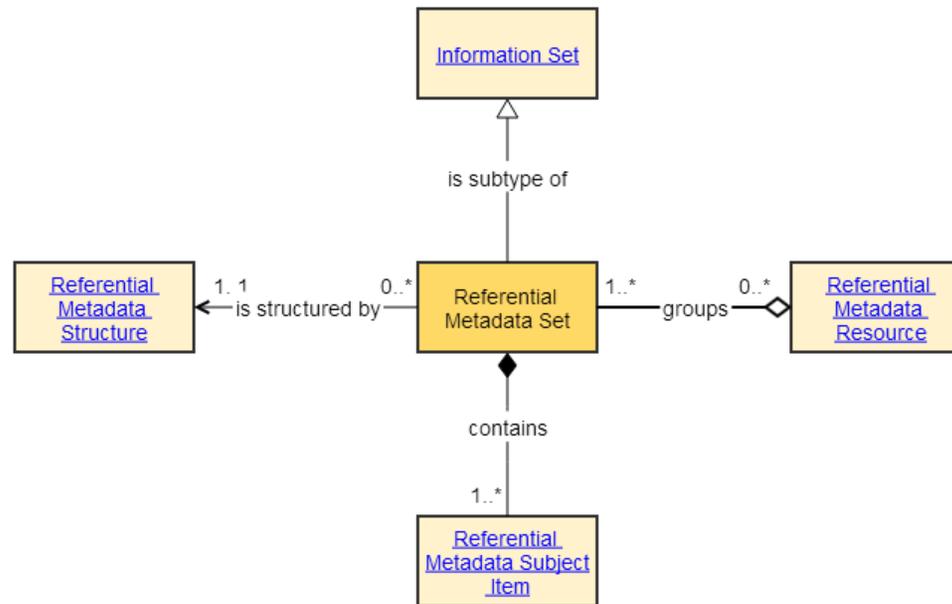
Definition

Object	Group	Definition	Explanatory Text	Synonyms
Referential Metadata Resource	Structures	An organized collection of stored information consisting of one or more <i>Referential Metadata Sets</i> .	<i>Referential Metadata Resources</i> are collections of structured information that may be used by a statistical activity to produce information. This information object is a specialization of an <i>Information Resource</i> .	

Attributes

Name	Description	Cardinality	Value Type
Name		0..1	Text
Description		0..1	Text

Referential Metadata Set



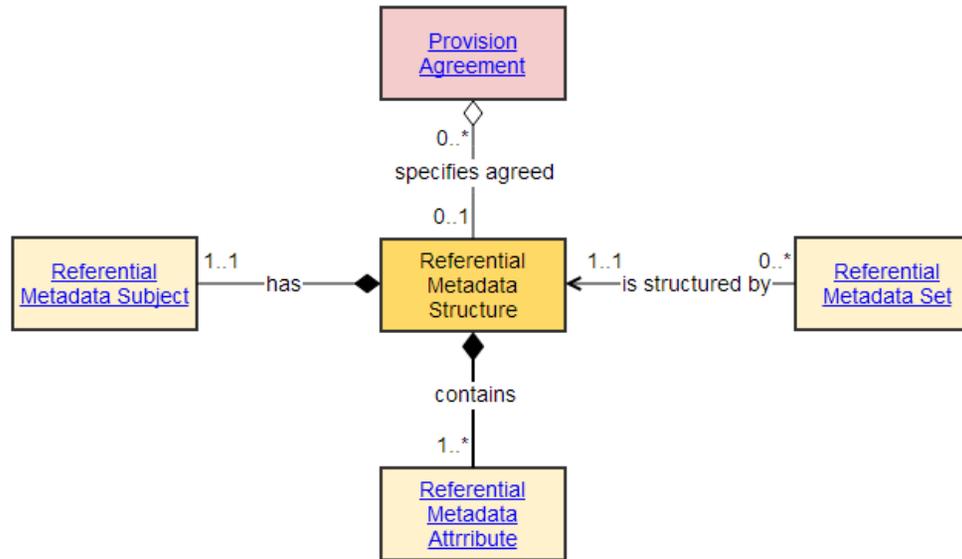
Definition

Object	Group	Definition	Explanatory Text	Synonyms
Referential Metadata Set	Structures	An organized collection of referential metadata for a given <i>Referential Metadata Subject</i> .	<i>Referential Metadata Sets</i> organize referential metadata. Each <i>Referential Metadata Set</i> uses a <i>Referential Metadata Structure</i> to define a structured list of <i>Referential Metadata Attributes</i> for a given <i>Referential Metadata Subject</i> .	

Attributes

Name	Description	Cardinality	Value Type
Name		0..1	Text
Description		0..1	Text

Referential Metadata Structure



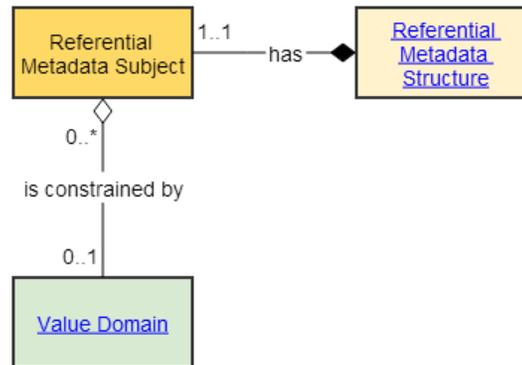
Definition

Name	Group	Definition	Explanatory Text	Synonyms
Referential Metadata Structure	Structures	Defines the structure of an organized collection of referential metadata (<i>Referential Metadata Set</i>).	<p><i>A Referential Metadata Structure defines a structured list of Referential Metadata Attributes for a given Referential Metadata Subject.</i></p> <p>Examples of <i>Referential Metadata Attributes</i> are those that describe quality information and methodologies. Examples of subject are: objects like a <i>Questionnaire</i> or a <i>Classification</i>, or collections of data like a <i>Data Set</i>, or any <i>Data Point</i> or set of <i>Data Points</i> created from a specific <i>Data Structure</i>.</p>	Metadata Structure Definition

Attributes

Name	Description	Cardinality	Value Domain
Name		0..1	Text
Description		0..1	Text

Referential Metadata Subject



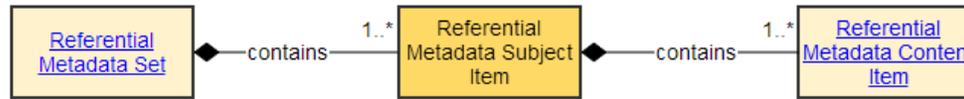
Definition

Name	Group	Definition	Explanatory Text	Synonyms
Referential Metadata Subject	Structures	Identifies the subject of an organized collection of referential metadata.	<p>The <i>Referential Metadata Subject</i> identifies the subject of the metadata that can be reported using this <i>Referential Metadata Structure</i>. These subjects may be any GSIM object type, or any <i>Data Point</i> or set of <i>Data Points</i> created from a specific <i>Data Structure</i>.</p> <p>Examples: The GSIM object type may be Product for which there is a list specified in a <i>Value Domain</i>. The <i>Value Domain</i> specifies the list of actual Products for which reference metadata can be reported or authored using this <i>Referential Metadata Structure</i>.</p>	

Attributes

Name	Description	Cardinality	Value Domain
Name	A human-readable identifier for the object	0..1	Text
Description	A human-readable description of the object	0..1	Text

Referential Metadata Subject Item



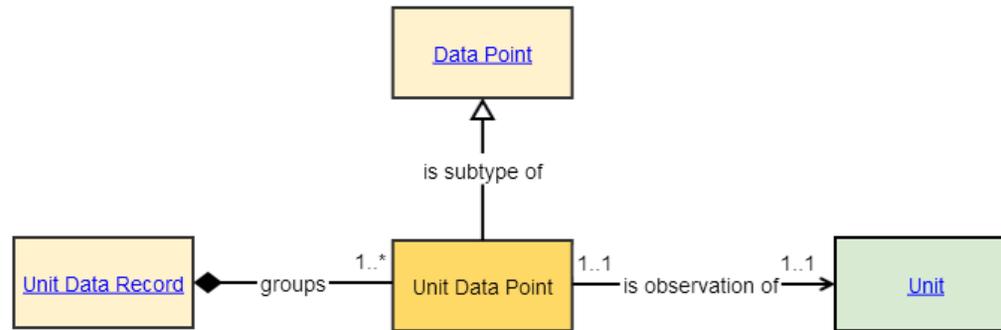
Definition

Object	Group	Definition	Explanatory Text	Synonyms
Referential Metadata Subject Item	Structures	Identifies the actual subject for which referential metadata is reported.	Examples are an actual <i>Product</i> such as Balance of Payments and International Investment Position, Australia, June 2013, or a collection of <i>Data Points</i> such as the <i>Data Points</i> for a single region within a <i>Data Set</i> covering all regions for a country.	

Attributes

Name	Description	Cardinality	Value Type
Name		0..1	Text
Description		0..1	Text

Unit Data Point



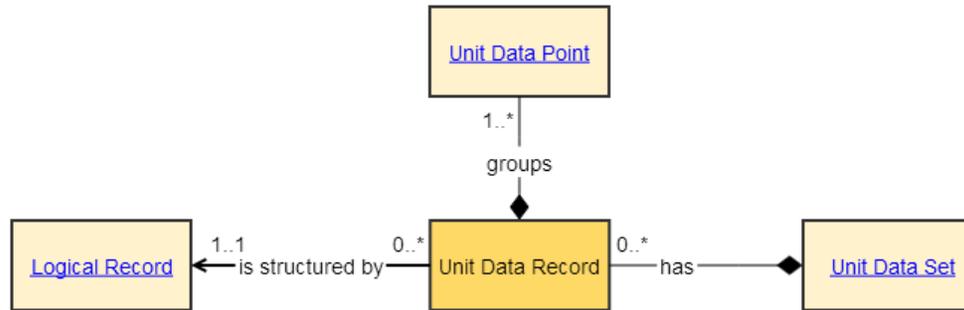
Definition

Object	Group	Definition	Explanatory Text	Synonyms
Unit Data Point	Structures	A placeholder (or cell) for the value of an <i>Instance Variable</i> with respect to a <i>Unit</i> .	This placeholder may point to multiple values representing different versions of the data. Values are only distinguished on the basis of quality, date/time of measurement or calculation, status, etc. This is handled through the mechanisms provided by the <i>Datum</i> information object.	cell

Attributes

Name	Description	Cardinality	Value Type
Name		0..1	Text
Description		0..1	Text

Unit Data Record



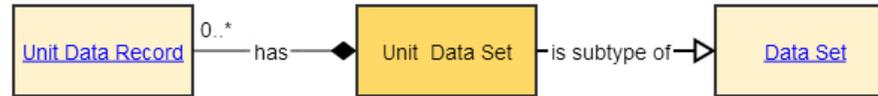
Definition

Object	Group	Definition	Explanatory Text	Synonyms
Unit Data Record	Structures	Contains the specific values (as a collection of <i>Unit Data Points</i>) related to a given <i>Unit</i> as defined in a <i>Logical Record</i> .	For example (1212123, 48, American, United Kingdom) specifies the age (48) in years on the 1st of January 2012 in years, the current citizenship (American), and the country of birth (United Kingdom) for a person with social security number 1212123.	

Attributes

Name	Description	Cardinality	Value Type
Name		0..1	Text
Description		0..1	Text

Unit Data Set



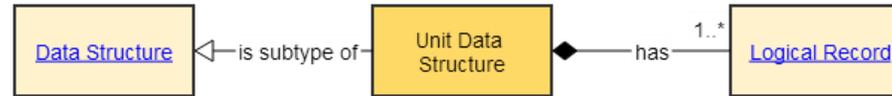
Definition

Object	Group	Definition	Explanatory Text	Synonyms
Unit Data Set	Structures	A collection of data that conforms to a known structure and describes aspects of one or more <i>Units</i> .	Example: A synthetic unit record file is a collection of artificially constructed <i>Unit Data Records</i> , combined in a file to create a <i>Unit Data Set</i> .	micro data, unit data, synthetic unit record file

Attributes

Name	Description	Cardinality	Value Type
Name		0..1	Text
Description		0..1	Text

Unit Data Structure



Definition

Object	Group	Definition	Explanatory Text	Synonyms
Unit Data Structure	Structures	Describes the structure of a <i>Unit Data Set</i> .	For example (social security number, country of residence, age, citizenship, country of birth) where the social security number and the country of residence are the identifying components (<i>Identifier Component</i>) and the others are measured variables obtained directly or indirectly from the person (<i>Unit</i>) and are <i>Measure Components</i> of the <i>Logical Record</i> .	file description, dataset description

Attributes

Name	Description	Cardinality	Value Type
Name		0..1	Text
Description		0..1	Text

Annex D: Extending the model

140. One of the GSIM design principles is that GSIM can easily be adapted and extended to meet users' needs. It is expected that some implementers may wish to extend GSIM, by adding detail and indicating which information objects are used, and exactly how.

141. Examples of when this could be needed are:

- (a) A statistical organization wants to specify types of *Rules* (for example, Methodological Rules and Process Control Rules)
- (b) A statistical organization wants to add another specialization of *Instrument*

142. Note that there are many points in GSIM where additional detail is expected to be added. These extensions can be done using the modelling techniques which GSIM itself uses. The following guidelines are intended to help modellers employ a common technique when extending and implementing the conceptual model, so that the use of GSIM itself within specific organizations is done in a common and understandable fashion.

143. For people who have experience in modelling with the standard UML tools, the recommended technique should be straightforward. However, not all staff have this experience. For those with less familiarity, a 'metamodel template' is also provided which allows non-modellers to capture the same information in a form that relies on plain text.

A. GSIM Extension Methodology

Namespaces

144. As part of the GSIM v1.1 release, the Enterprise Architect file which contains the UML models will be released. In this file there are five 'namespaces' (or 'packages') – one for each of the GSIM Groups.

145. Any organization extending GSIM should establish one or more namespaces which are specific to and owned/maintained by that organization. This provides a clean separation between GSIM itself, and the extensions that have been made to it.

146. In many cases, the extensions might provide useful input to future development of GSIM itself, so should be made available to the maintenance agency (UNECE Standards Modernisation Committee). In other cases, they may be too organization-specific for this purpose.

147. The classes native to GSIM would be imported into the organization-specific namespace(s), and extensions made from them. Any new information objects would also be modelled in this namespace. In the same way that GSIM itself is organized into namespaces, it is recommended that if more than one organization-specific namespace is created by the extender, these should be organized along similar lines.

New Classes

148. New classes may be created using the same style of modelling as is found in GSIM itself. GSIM uses a fairly standard but restricted set of the features of UML. The best guide to this style is to study the GSIM UML models. Such things as multiple inheritances have been avoided, and there is a distinct style in terms of how relationship roles are named.

Extensions/restrictions to existing classes

149. Any class within GSIM can be imported and then extended/restricted. Classes can be extended with new properties and relationships, and the existing properties and relationships can be over-ridden.

150. The extended classes inherit all properties and relationships from their parents, so these do not need to be explicitly modelled unless:

- (a) they are required for clearer understanding (they will appear preceded by a slash ["/"]); or
- (b) they have been changed - that is, over-ridden.

151. Extension and restriction in the UML models are shown with an open-headed arrow pointing from the extending/restricting class to the class that it inherits from, and of which it is a sub-type. The details of what is allowed are provided below:

Extension of existing classes:

152. Create a new sub-type, with its own name, a definition, explanatory text, and examples, and then specify any additional type-specific additions to the set of properties or relationships which that information object possesses.

153. Note: There are some common attributes, which exist for all GSIM information objects, and these will be present by inheritance. The same is true for administrative attributes added to the GSIM Base *Administrative Details* information object.

Restriction of classes:

154. The information object to be restricted is imported into the organization-specific namespace and then sub-classed. Any existing relationships or properties may be over-ridden, unless they are required by the inherited cardinalities. This is done by simply re-stating the property or relationship, and changing its details. Even within required cardinalities, so long as a restriction still produces a valid instance of its parent, the change is allowed. For example, a property with a cardinality of 1..* may be restricted to having a cardinality of 1, but not less than that, since at least one instance of that property is required.

155. Note: If a class in GSIM is to be both extended and restricted, the same sub-type is used, with over-rides and additions made as desired.

156. It is possible, using this mechanism, to express exactly what information objects within an organization are used and not used. If there is no relationship to an information object, or if its cardinality has been reduced to 0 for all properties and relationships, it is simply not used.

Documentation

157. GSIM itself should be used as an example of how to document extensions and restrictions. This means providing the information in the metamodel template (see below) and providing the definitions and descriptions/examples in tabular form, as well as providing an overall narrative of each UML diagram produced.

Box 1. Metamodel Template

Information Object Name

Version:

Package:

Definition:

Explanatory Text:

Synonyms:

Constraints:

Attributes

Name	Description	Cardinality	Value Type

Relationships (repeat as needed)

Name:

Target Object:

Relationship Type:

Description:

Source Role:

Source cardinality:

Target Role:

Target Cardinality:

Constraints:

B. Administrative Attributes

158. GSIM does not model the information used by statistical organizations to administer and maintain their metadata - there are too many potential differences. Such administrative attributes are also very dependent on implementation, and GSIM is a conceptual model.

159. To support the use of administrative attributes, GSIM provides an information object - *Administrative Details* - which can be extended to include whatever set of administrative attributes are needed by an implementer of the GSIM.

160. In order, to encourage commonality of practice, GSIM recommends a set of administrative attributes based on the ISO/IEC 11179 standard. The following table shows the set of recommended attributes for the administration of GSIM information objects.

Table 3. Recommended Attributes

Name	Description	Mandatory	Value Domain
Identification attributes			
Name	A term which designates a concept, in this case an information object. The identifying name will be the preferred designation. There will be many terms to designate the same information object, such as synonyms and terms in other languages.	Yes	Text
ID	The unique identifier of the information object; assigned by the owner agency.	Yes	Number
Governance attributes			
Version	The version designator of the information object assigned by the owner agency.	Yes	Number
Owner Agency	The organization or legal entity that owns and maintains the information object.	Yes	Text
Organization Unit	The organization unit, within an agency, which owns (has rights to create, update, delete) the information object.	No	Controlled vocabulary
Valid From	The date on which the information object is effective or valid.	Yes	Date
Valid Until	The date on which the information object is no longer effective or valid.	Yes	Date

Created Date	The date on which the information object was created	Yes	Date
Created User Id	The person who created the information object	Yes	Controlled vocabulary
Last Update Date	The date on which the information object was last changed.	No	Date
Last Update User Id	The person who last changed the information object.	No	Controlled vocabulary
Administrative status ²	indicator for access to an item: under review, open for use, or removed	No	Controlled vocabulary
Life cycle status	indicator for the quality of an item: incomplete, valid, superseded, or retired	No	Controlled vocabulary
Content attributes			
Description	A statement which describes an information object. It also delineates the information object's scope.	Yes	Text
Annotation	A comment or instruction which provides additional explanations about the information object and how to use it.	No	Text
Topic	The subject or theme the information object is related to. This is included to support search.	No	Controlled vocabulary
Keyword	Terms related to the information object. These are included to support search.	No	Controlled vocabulary
Technical implementation attribute			
IsStructured	Identifies if the description can be executed by a machine.	No	Boolean

161. Implementers can use the GSIM extension methodology to include the recommended set of administrative attributes. The *Administrative Details* information object in GSIM has been purposefully left blank as a stub to be extended.

162. In this case, all that is needed is to create a namespace and to import the *Administrative Details* information object into it. The *Administrative Details* information object is then sub-classed, and the

² Administered Status refers to the availability of an item description, whereas the Life-Cycle Status refers to the quality and relevance of an item description

attributes listed above are added. Figure 23 shows what would appear in a UML diagram if this is done.

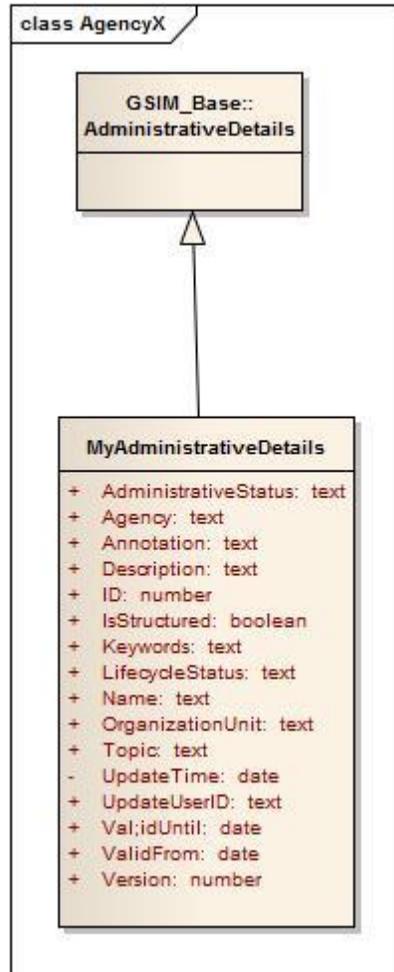


Figure 1. Extension of Administrative Details.

Note: The fields containing controlled vocabularies are shown in the diagram as text. These text strings would agree with a maintained list appropriate to the field which uses them.