Abstract description of the model represented in taxonomies following the DPM approach
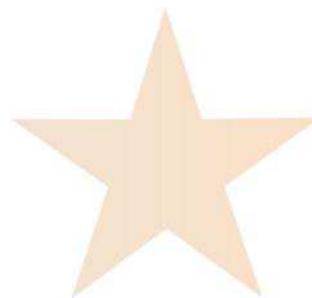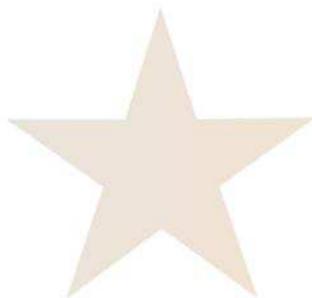
Changes

1	26/06/2012	-	Initial version
public release v1	20/12/2012	-	Updated version number

Table of Contents

1 Introduction

The Data Point Modeling is a methodology for the development of financial data models that describe characteristics of the information exchanged in the context of supervisory reporting processes. These data models are often referred to as Data Point Models (DPM[1]) or as meta-data (data that describes data). Data Point Models are formally represented by XBRL taxonomy files that are built following certain conventions to warrant an accurate and univocal representation of the data model.

The purpose of this document is to define the set of characteristics that DP models describe and thus, the information represent by XBRL taxonomies following this approach. In other words, this document defines the meta-model (a model that describes models) of the information exchanged. This document will make use of UML diagrams to formally represent the different parts of the meta-model.

The following picture represents the relationship between meta-model, meta-data and data, and the means used to describe them:

Meta-model (DP formal model)	Describes →	Data model Metadata DPM	Describes →	Financial Data
Documented in ↓		Published in ↓		Exchanged in ↓
UML diagrams (this document)		XBRL Taxonomy files		XBRL Instance files (or equivalent mechanism)
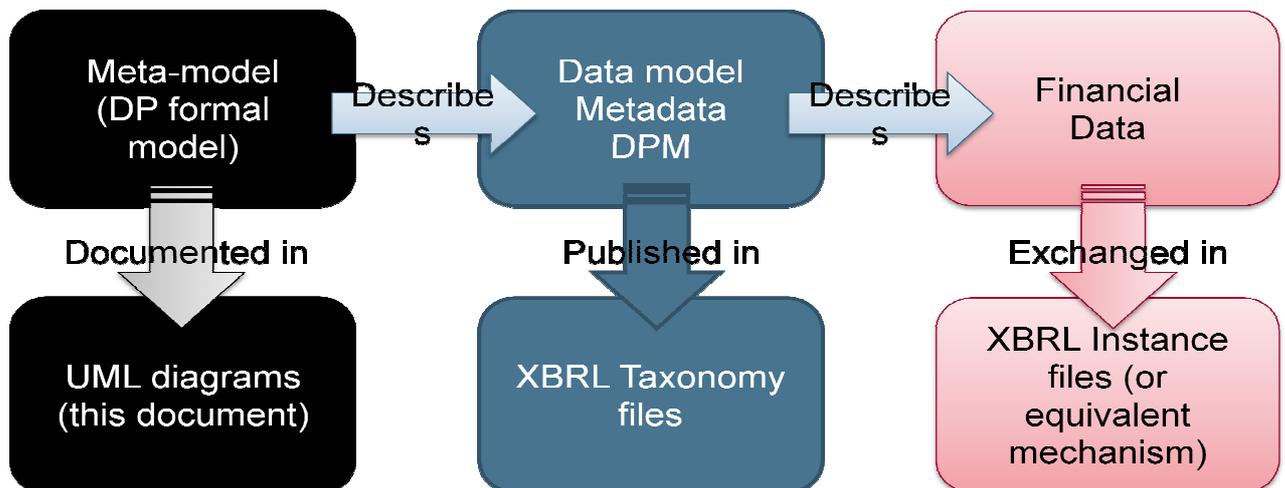
At the right side, the representation of financial data is done using XBRL instance documents as part of the reporting exchange process. The characteristics of each financial concepts subject of this exchange process are described in the data model or meta-data; its formal representation is done using XBRL taxonomy files. The definition of those characteristics is the meta-model; its description if the subject of this document with the support of UML diagrams.

---

[1] The acronym DPM is used in the sense of a methodology for the design of models (Data Point Modelling) and sometimes is used to refer to the models designed using this methodology (Data Point Model).

EBA European Banking Authority

2 Basic definitions

2.1 Facts

Flows of information in reporting processes involve the exchange of data. The most basic data particle exchanged in these flows is referred to as a *fact*. Facts are not strictly part of the data model, but they are necessary to understand the meaning of its most basic component: the *data point*. A *data point* is the financial concept that a *fact* describes.

A fact measures a financial concept of a certain entity (e.g.: a credit institution) in a moment of time (or in the case of flow nature concepts, an interval of time). We will refer to this as the context[2] of the fact.

The result of the measure of a data point in a certain context is the *value* of the fact. The value is usually an amount, but could be of a different type: a description, a Boolean value, a date… In the case of numeric information, the fact also includes a unit (for instance, the currency of a monetary amount) and its precision (a measure of the accuracy of the value reported).

In addition to the value, a fact may include some textual descriptions in one or several languages: the footnotes.

---
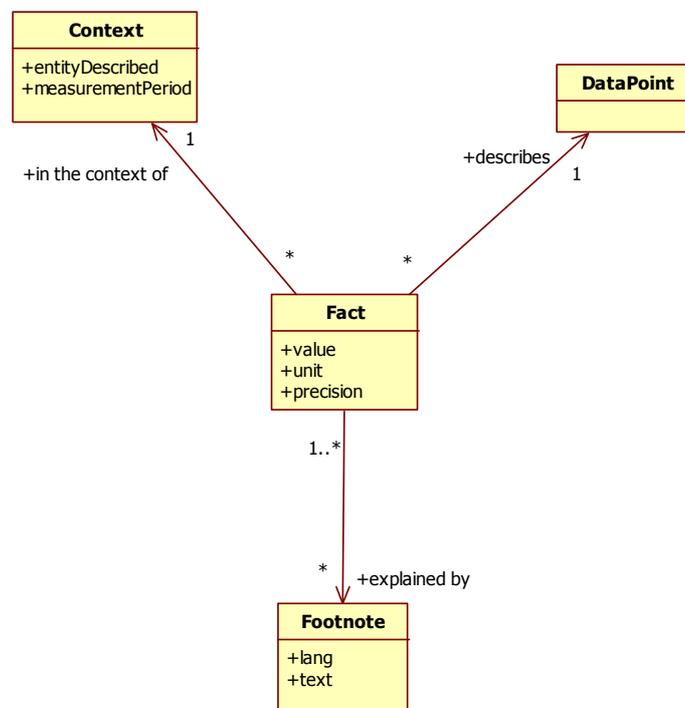
[2] Note that this definition of context is not directly related to the context XML element used in XBRL instances.

2.2 Data points

In the context of supervisory exchange processes, a data point represents an individual data requirement. Data points are expressed as a composition of characteristics that univocally identify the financial concept to be measured. These characteristics are:

- A metric: the characteristic that defines the nature of the measure to be performed. Each metric determines a data type and a period type. The data type (monetary, percentage, Boolean…) corresponds to nature of the values of the facts to be reported. The period type establishes whether the measure is performed at a specific instant of time (stocks) or during an interval of time (flows).

- A number of dimensional characteristics that qualify and complement the metric and provide the proper context to understand the financial phenomenon represented. Each dimensional characteristic is composed of a dimension and a member. The dimension represents an identifying property and the member is the value given to that property in the context of the data point.

- A temporal reference that helps to determine the specific instant or interval of time in the context a given reporting period. The temporal reference enables the distinction, for instance, of information that refers to end of the reporting period, the beginning of the fiscal year, a quarter, a month, …
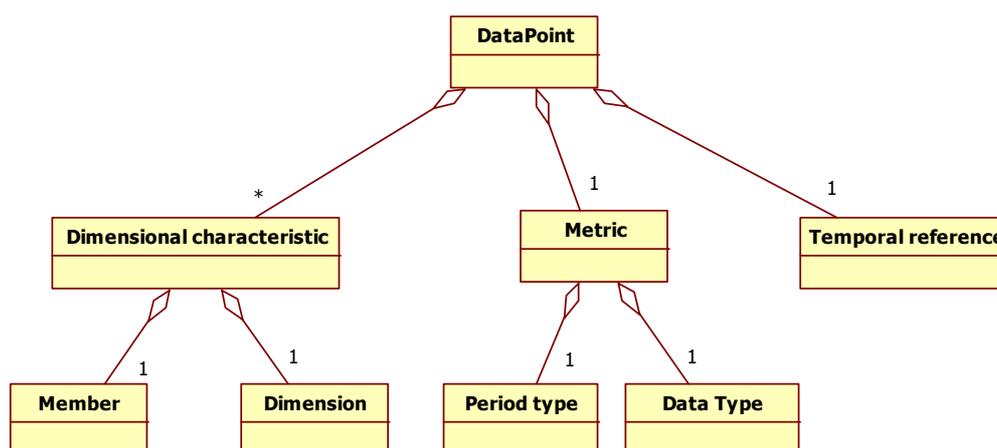
Individual data points do not have an explicit representation in the model. As it will be described later, they are defined as part of sets of data requirements (sets of data points) with the help of the dictionary of concepts (dimensions, members, metrics…), tables and dimensional restrictions. However, it should be noted that the representation of a data point is independent of the table it is part of; in fact, a single data point can be part of one or more tables. In all those cases, its identification is unique an independent of the table. Accordingly, a fact is reported only once.

2.3 Public elements

Public elements are a generalization of certain concepts of the model that are identified by a code in a certain scope and include additional information such as readable names, definitions and legal references. Examples of public elements are metrics, dimensions, members or tables. The name, definition and legal references of a public element can be expressed in different languages:
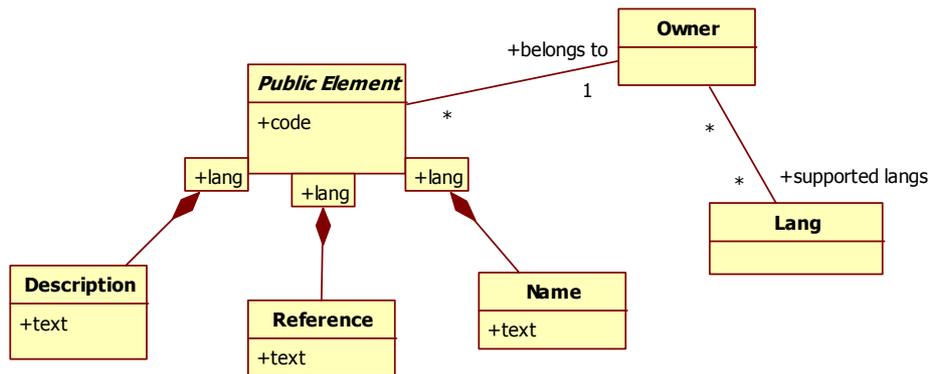
Each public element is defined by an institution or organization, represented in the model as the owner. The owner includes a list of supported languages. The owner is expected to maintain the labels for the concepts it defines and uses in the supported languages.

3 Dictionary concepts

The core concepts of the dictionary are metrics, dimensions, domains and domain members. Secondary concepts are families and perspectives (auxiliary concepts meant to group dimensions for presentation purposes).

All the concepts in the dictionary are public elements. In addition to the properties and language specific information of public elements, dictionary elements may define a currency period, so that obsolete concepts can be filtered by applications. The currency period does not have an impact on the reporting process; it is meant to make easier the management of the concepts of the dictionary.

3.1 Metrics

Metrics define the nature of the measure to be performed. In other words, metrics define the property or amount to be obtained. Consequently, metrics carry some business meaning and determine the data type and the period type:

- The data type establishes the set of possible values of the facts reported according to that metric: monetary information, percentages, dates, texts…

- The period type defines whether the property / amount to be measured corresponds to a specific moment in time (instant type) or whether its nature requires it to be obtained by taking measures during an interval of time (duration type).
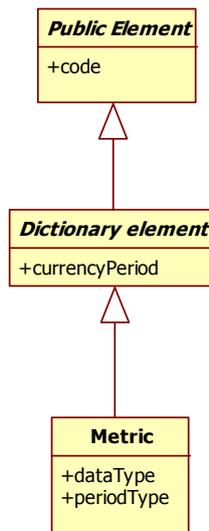
Each data point in the model must define one and only one metric.

3.2 Dimensions, domains and domain members

Dimensions are properties used to qualify metrics and thus, narrow down the definition of the business concept represented by a data point. In other words, dimensions give meaning to the context in which metrics are obtained.

The values given to dimensions are called members. Members are grouped in domains. Thus, domains are sets of members that share a certain semantic identity. For instance, UK, Spain, German and Spain are members of the geographical area domain.

The members of a domain can be defined by enumeration (explicit domain) or by defining a data type plus some additional constraints (typed domain). An example of explicit domain could be the geographical area or the currency. An example of typed domain could be ISIN codes.

A certain dimension can only be given as values the members of a single domain. Thus, it is said that a dimension is associated to a domain. If the domain associated to a dimension is explicit, the dimension is said to be explicit; otherwise, the dimension is typed. However, a domain can be associated to multiple dimensions. For instance, the dimensions "Origin of the transaction" and "Destination of the transactions" can both be associated to the geographical area domain.

A data point may be associated to any[3] number of dimensions with their corresponding values. But dimensions cannot be defined twice for a single data point, or be given multiple values.

In the case of explicit domains, a member can be chosen to be the default one[4]. The consequence is that for each data point, any dimension associated to such domains is considered to be given the default value unless a different value is explicitly assigned.

---

[3] Zero dimensions is also a possible choice.
[4] In the XBRL Specification, two dimensions sharing the same domain can have different default members. However, this possibility has been constrained in the model as it has been considered unnecessary and confusing in the modelling of data.
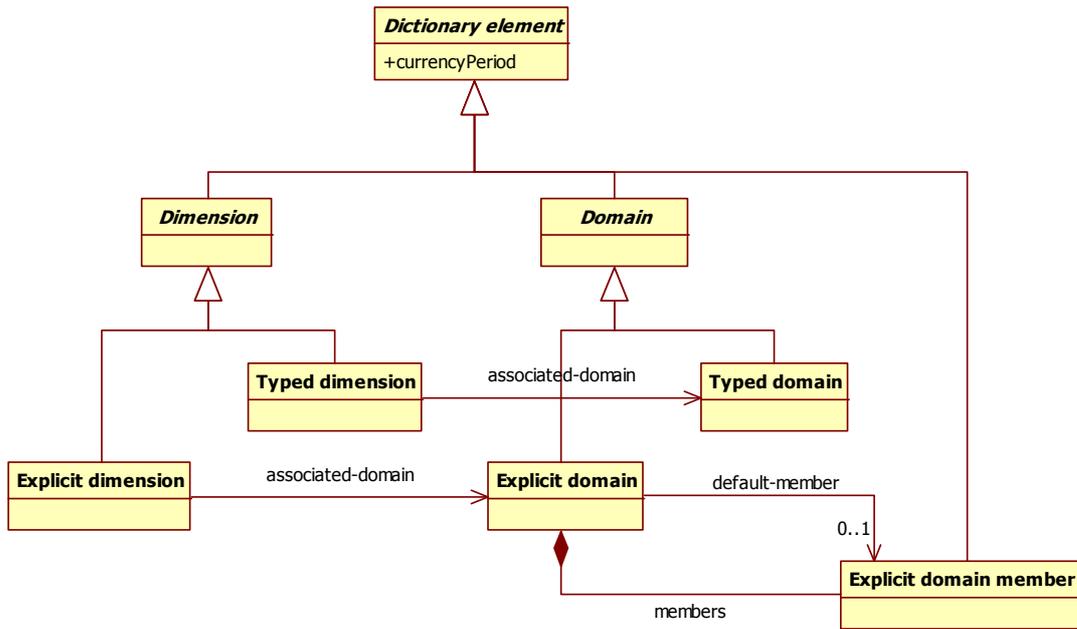
3.3 Hierarchies

Hierarchies are sets of members of an explicit domain arranged in a hierarchical disposition. A node of a hierarchy can define basic arithmetical relationships (=, <= or >=) in relation to its child nodes. Child nodes can determine a sign (+ or -) that represents whether the member referenced should account positively or negatively to the arithmetical relationship defined in the parent node. Hierarchies are a compact way of defining three ideas:

-   Subsets of the members of a domain (sub-domains)

-   Schematic arrangements of members of a domain for presentation purposes

-   Basic arithmetic relationships

Approach for market risk			
=		1	Total/NA
=	+	2	Standardised approaches
=	+	3	Normal treatment
	+	4	Maturity-based approach
	+	4	Duration-based approach
=	+	4	Other
	+	5	Debt securities under the first category
	+	5	Debt securities under the second category
	+	5	Debt securities under the third category
	+	5	Debt securities under the fourth category
	+	5	Securitisation exposures subject to 1250% risk weighting or deduction and unrated liquidity facilities
	+	3	Particular approach for CIUs
	+	3	Additional requirements for options

COREP sample hierarchy

The UML diagram for hierarchies follows. Hierarchies are public elements, and thus, can include names, references to legal documentation…
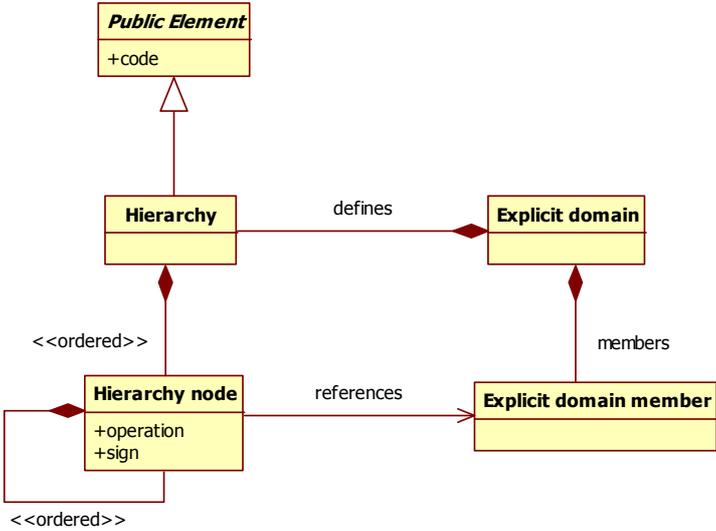
A special case of hierarchies are those where referenced members are metrics:
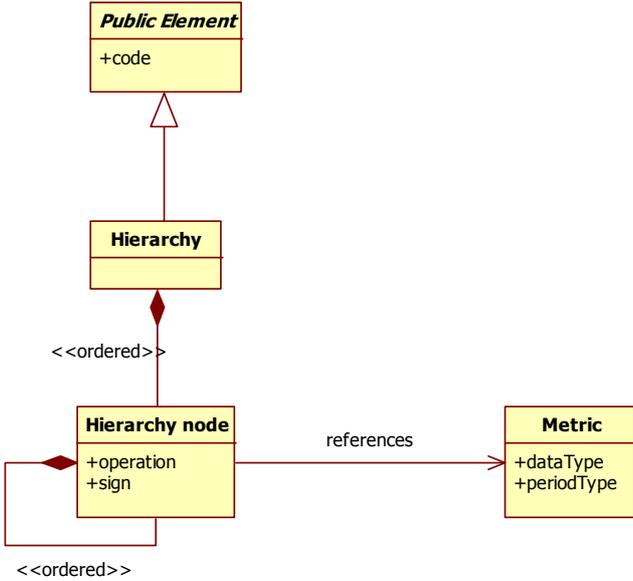
3.4 Perspectives and families of dimensions

Families are groups of dimensions only relevant for presentation purposes. Considering the number of dimensions in the model can grow over the years, these groups can help to establish certain classification and facilitate the management of the dimensions in the dictionary.
Perspectives represent different criteria of grouping: for financial purposes, for prudential purposes, for statistical purpose… Perspectives, therefore, establish an association between dimensions and families. Given a perspective, one dimension can be associated with a single family.
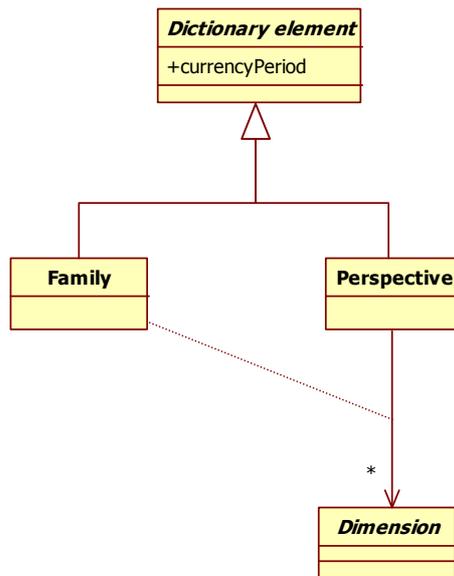
4 Reporting requirements layer

Frameworks, taxonomies, tables, modules and other concepts constitute the layer of the model where actual reporting requirements are specified in terms of the concepts defined in the dictionary.

Tables and axes

Data required by supervisors to credit institutions is described in legal normative by means of bi-dimensional forms usually referred to as templates or business templates. These business templates are represented in the model with the help of tables and table groups.
A table consists of the combination of one or more axes with a certain spatial disposition (x, y or z). Each axis defines a set of ordinates arranged hierarchically. Most ordinates define a place in the table for data requirements (a column in the case of X axes or a row in the case of Y axes); some ordinates, called abstract, are just used to group other ordinates.
Each ordinate defines a set of characteristics (the ones described in chapter 0) that identify the data requirements nested under such ordinate and its descendants, though a child ordinate may override any of those characteristics.
The set of potential data requirements represented by a table are the data points result of the Cartesian product of non abstract ordinates in its axes. Each data point obtains its identifying characteristics from their corresponding ordinates in each of the axis in the table.
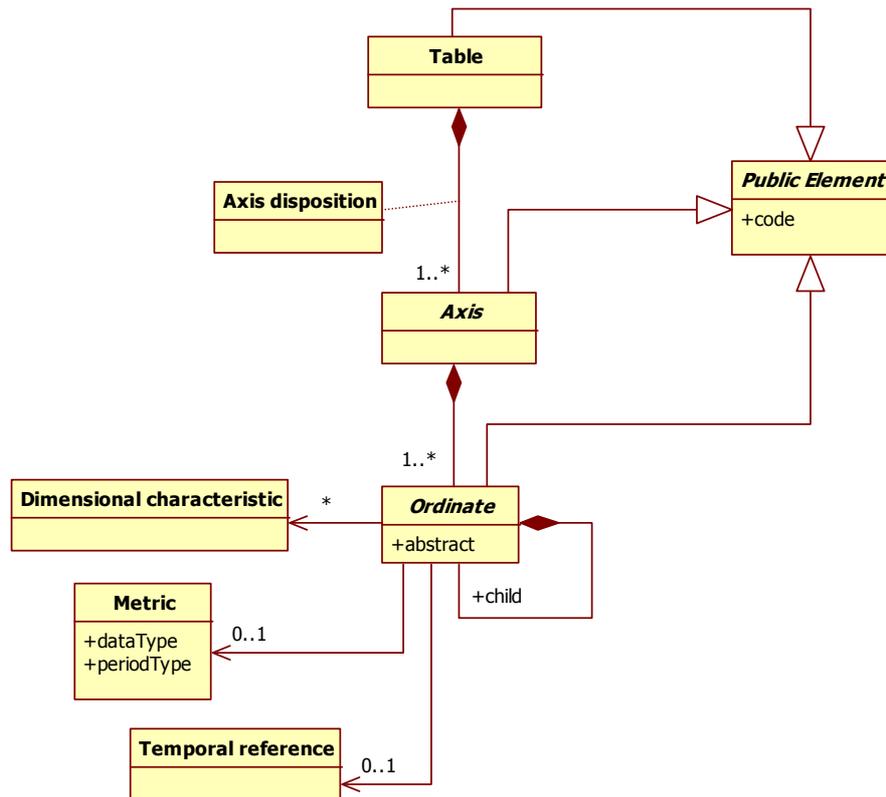
Tables, axes and ordinates are public elements. Thus, they can have readable labels or references to legal documentation. Consequently, the model has two types of labels / references: the global ones (the ones defined at dictionary level) and the ones defined in the context of a table. Global labels must be independent of the table, and so, are usually longer than table specific ones, that are meant to be showed in the context of a table.

4.2 Types of axes: open and predefined axes

The model considers two types of axes:

- Predefined axes are those whose list of ordinates is defined a priori; its carnality is fixed. Predefined axes are defined either by explicitly enumerating its ordinates (explicit ordinates) or by including a reference to a hierarchy in the model.

- Open axes, as opposed to predefined ones, are the ones that can have a variable number of entries. The number of entries depends on the data reported. For instance, the list of subsidiaries of a heading company can be expressed in a table where each row represents the information of a subsidiary. Subsidiaries are represented as an open y-axis. The dimensions represented in open axes in a table are sometimes referred to as key dimensions, as any row (or column if x-axes are used) in the table is identified by the values of the key dimensions. An open table (a table with open axes) cannot have more than one row (or column) with the same values for all its key dimensions.
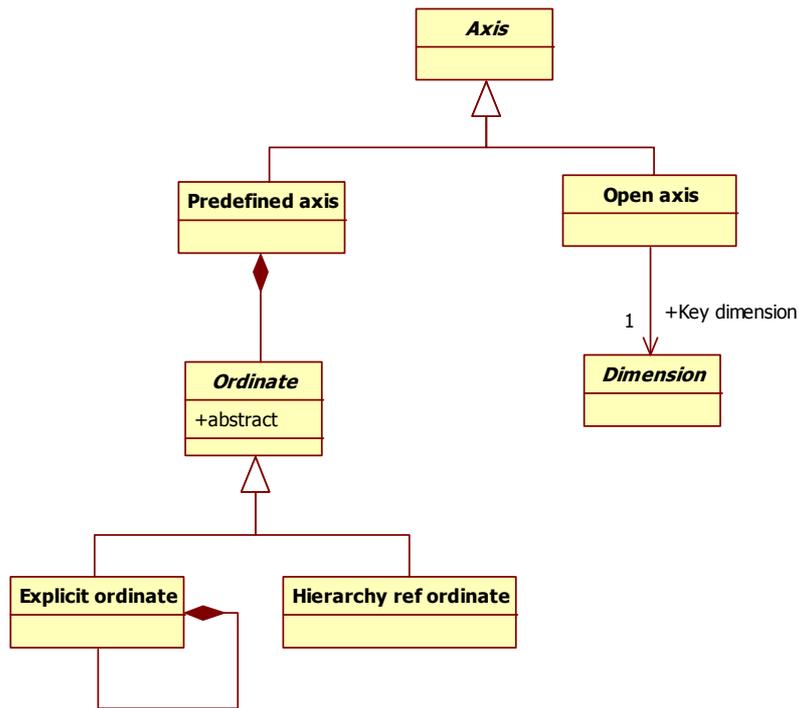
4.3 Hypercubes

Not all the potential combinations of the axes in a table correspond to data requirements. Business users identify non required data in templates using gray cells. In the model, the set of actual data requirements is represented with the help of valid hypercubes. Each cell in a table that corresponds to a required data point must have its representation in one the hypercubes of the table.

Hypercubes are defined as a set of metrics, a set of typed dimension and a set of explicit dimension together with the set of valid members.
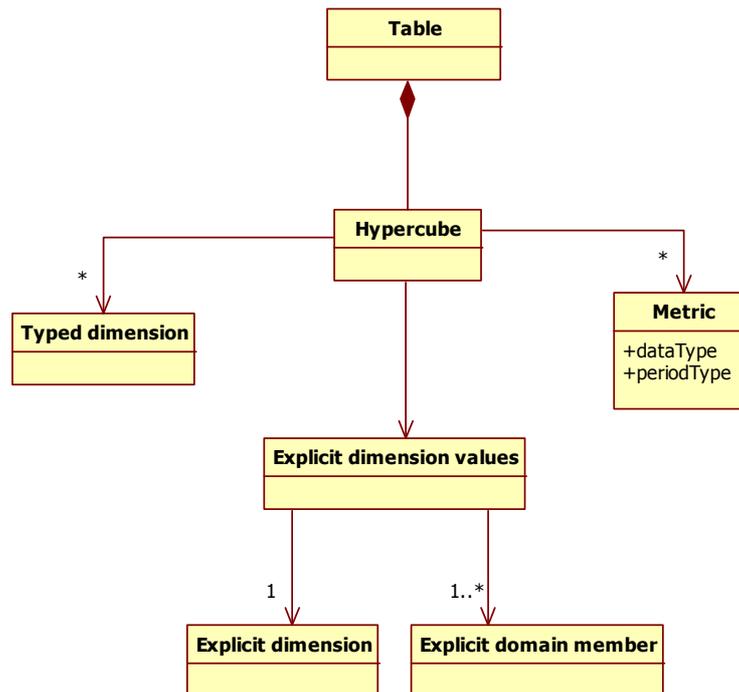
4.4 Table groups and modules

In most of the cases, business templates can be represented directly using tables. However, in some occasions, business templates are a patchwork of sets of data requirements that are mixed in a complex template in order to reduce the total number of templates. In those cases, the set of data requirements defined by one of these templates cannot be represented just using a single table, but a set of them. In order to maintain a clear link between these complex templates and the tables that represent their content, the concept of table group has been introduced in the model.

A table group is just a set of tables that usually corresponds to one business template. If required, table groups could be used also to represent more complex hierarchical representations of a set of tables. For instance, some FINREP and COREP templates have been split into smaller ones because of its complexity. In fact, a table group can be composed of other table groups in addition to tables.

Modules are predefined sets of tables and table groups that are used in a certain process. For instance, in reporting processes, a module defines sets of information that must be reported together. Instance documents include a reference to the module (also called entry point); not explicitly reported data in such instance documents will be assumed to be the default value (zero for numeric data)[5].

Like tables, table groups and modules are public concepts and may have names, references and descriptions in several languages.

---

[5] With the exception of incremental reporting processes, where only modified data is reported.
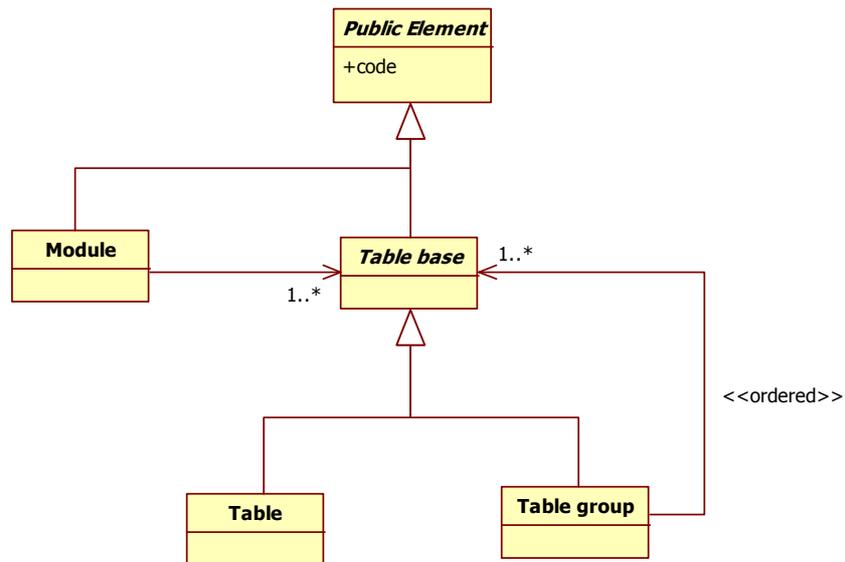
The concept of table base defined in this diagram is just an abstraction that represents either a table or a table group.

4.5 Taxonomies and frameworks

A taxonomy represents a set of reporting requirements enforced by a certain legal document or set of documents. A taxonomy is defined in terms of modules, tables and table groups. In addition, taxonomies include attributes to identify the legal normative, its version date[6] and an optional currency period.

In order to reduce the cost of maintenance, tables from previously released taxonomies that have not suffered any modification can be referred from a new taxonomy.

Frameworks define groups of taxonomies following some functional requirements (solvency information, financial information…). Taxonomies, in relation to frameworks, can be considered a specific version in time of the data requirements defined by a framework.

---

[6] The first release usually corresponds to the publishing date of its corresponding legal normative.

```
            ┌─────────────────┐
            │ Public Element  │
            ├─────────────────┤
            │ +code           │
            └─────────────────┘
                     △
         ┌───────────┴───────────┐
  ┌──────────────┐      ┌──────────────────┐
  │  Framework   │◆─────│    Taxonomy      │
  ├──────────────┤      ├──────────────────┤
  │              │      │ +normative       │
  └──────────────┘      │ +versionDate     │
                        │ +currencyPeriod  │
                        └──────────────────┘
                          ◆      ◆      imports
                ┌─────────┘      └────┐      └──────────┐
          ┌──────────┐      ┌──────────────┐      ┌──────────────┐
          │  Module  │      │  Table base  │      │  Table base  │
          ├──────────┤      ├──────────────┤      ├──────────────┤
          │          │      │              │      │              │
          └──────────┘      └──────────────┘      └──────────────┘
```
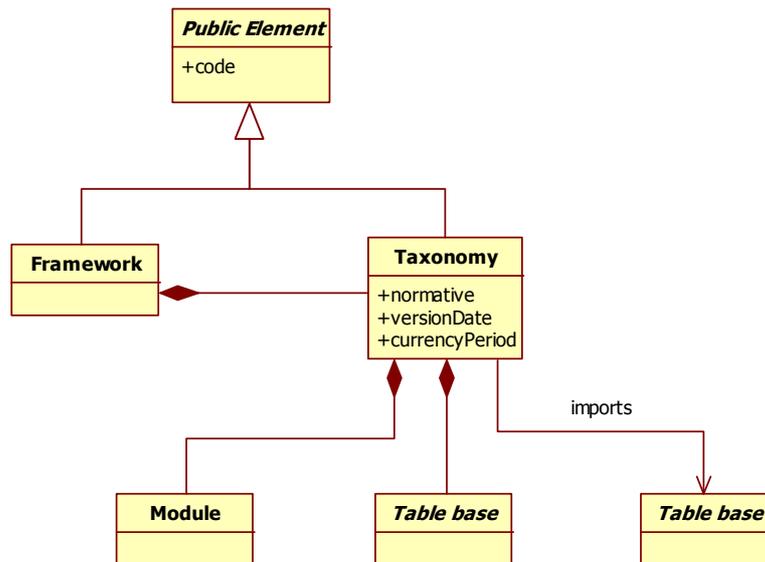
4.6 Validation rules

Validation rules are tests to be applied to reported data in order to check its consistency. If the result of a validation rule to a set of data is true, then the data reported is consistent according to that rule. If the result is false, the reported information presents an inconsistency that should be checked or corrected.

Validation rules are defined in terms of a set of fact variables that are bound to reported data during the validation process. The data that a fact variable should be bound to is defined in terms of filters. There are filters available to check any of the properties of the concepts defined in the dictionary. Given the high number of filters available in the XBRL specification, their details have not been included in this document.

A special set of variables (general variables) enables the definition of more advanced validation rules. Variables (both general and fact variables) can be bound during the evaluation process to a single value or to a set of them. This is controlled through the "asSequence" attribute in the model.

The set of data where a validation rule is to be applied can be restricted using two mechanisms:

- Filters: filters associated to validation rules narrow the set of data where a validation applies.

- Preconditions: logical expressions that verify a condition on the data where the assertion is potentially applied. If the result of the expression is false, that data is skipped.

Two types of validation rules are described by this model:

- Validation rules that check the content of the data given a certain logical expression.

- Validation rules that check whether some data has been reported.

Validation rules are associated to the sets of tables it applies to. Not reported information that is part of the content of a table where that rule applies will be assumed to have a default value. This default value is provided in the definition of fact variables. By default, numeric variable will be assumed to have a zero as default value.