

**CEN/TC XBRL**

Date: 2013-06

**TC XBRL WI XBRL003**

CEN/TC XBRL

Secretariat: NEN

## **Improving transparency in financial and business reporting — Harmonisation topics — Part 2: Guidelines for data point modeling**

*Einführendes Element — Haupt-Element — Teil 2: Ergänzendes Element*

*Élément introductif — Élément central — Partie 2 : Élément complémentaire*

ICS:

Descriptors:

Document type: CWA  
Document subtype:  
Document stage: CEN Enquiry  
Document language: E

## Contents

	Page
Foreword .....	3
Introduction.....	4
1 Scope .....	6
2 Terms and definitions.....	6
3 What is a data model .....	7
3.1 General .....	Error! Bookmark not defined.
3.2 The term “model”.....	8
3.3 Data-oriented process of modelling .....	8
3.4 The conceptual data model as a first step aiming for a database system .....	8
3.5 Description of data modelling approaches for supervisory purposes .....	9
3.5.1 Introduction.....	9
3.5.2 Using the “form centric” modelling approach.....	10
3.5.3 Using the “data centric” modelling approach .....	11
3.6 Description of dimensional modelling .....	13
3.7 The concept of normalisation .....	14
4 Why use a multidimensional data model .....	18
4.1 Introduction.....	18
4.2 Multidimensional data model.....	18
4.3 Operations that can be carried out on a multidimensional data model.....	19
5 Why data modelling is essential for collecting supervisory information .....	20
5.1 Introduction .....	20
5.2 Objective of Data Point modelling .....	20
5.3 TITLE NEEDED .....	Error! Bookmark not defined.
5.3.1 Increase of knowledge and understanding.....	21
5.3.2 Improvement of integration of changes .....	22
5.3.3 Reduction of risk of duplicate information .....	22
5.3.4 Higher harmonisation .....	25
5.4 Classification of Data Point modelling in the data modelling concept.....	26
5.5 Area of application.....	27
5.6 What are the technical constraints.....	28
6 How do you proceed in creating a Data Point Model .....	29
6.1 Introduction.....	29
6.2 Define dictionary elements.....	30
6.3 Specify hierarchies .....	31
6.4 Define Data Points .....	32
6.5 Define normalised tables and ensure quality of Data Point Model .....	32
6.6 Distribute Data Point Model .....	33
7 What the future holds for us .....	33
Bibliography.....	36

## Foreword

This document has been prepared by CEN/WS XBRL, the secretariat of which is held by NEN.

CWA XBRL 001 consists of the following parts, under the general title *Improving transparency in financial and business reporting — Harmonisation topics*:

- Part 1: European data point methodology for supervisory reporting
- Part 2: Guidelines for data point modeling
- Part 3: European XBRL Taxonomy Architecture
- Part 4: European Filing Rules

This document is currently submitted to a public consultation.

## Introduction

### General

The purpose of this document is to support supervisory experts in the creation of a Data Point Model (DPM). By definition of the European Banking Authority (EBA) a DPM “is a structured formal representation of the data [...] identifying all the business concepts and its relations, as well as validation rules, oriented to all kind of implementers.”<sup>1</sup>

The underlying rules for the creation of such methods were initially introduced by the Eurofiling Initiative and developed further by the European Insurance and Occupational Pensions Authority (EIOPA). The main objective of data point modelling, the process of creating a DPM; “[it] should help to produce a better understanding of the legal background to the prudential reporting data and make data analysis much easier for both the institutions and regulators”<sup>2</sup>.

Further goals are to prevent redundancies, lower maintenance efforts and, in general, to facilitate working with national extensions on the European agreed upon data set to facilitate the descriptions of requirements that are sharable across national legislations. It is a requirement to have all the information collected by the national supervisory agencies, particularly in Europe, transformed into the same data structure with the same quality in order to be able to carry out standardized analysis of the data across Europe. The current implementations are not able to meet these European requirements for supervision “to achieve higher quality and better comparability of data”<sup>3</sup>. The main reasons for this are the differences between the data definitions and the data formats of the various national supervisory agencies, making comparison of reported data virtually impossible.

### Objective

The aim to harmonise the European supervisory reporting is to be able to carry out more comprehensive analysis and an increase of comparability of data. The supervisory agencies are already acquainted with the representation of regulations specified in laws, this document is going to introduce the reader to the concept of Data Point modelling methodology as well as to its main terms and definitions that will enable you to create Data Point Models that contain “all the relevant technical specifications necessary for developing an IT reporting format” on your own.

### Target Audience

In general you as banking supervisor are responsible to communicate with Information Technology (IT) experts in order to support the transfer of the essence of regulatory reporting to IT systems. In 2009 the Eurofiling Initiative has published the concept of Data Point modelling. Structures of data represented in supervisory tables as well as underlying laws and guidelines were defined in order to enable the interpretation of the reporting information by IT applications. IT specialists are responsible for the development of software, however most of the time they do not have the special business knowledge needed to gather reporting requirements from various sources such as legal texts like Solvency Regulations and National Banking Acts for building a faultless system. Therefore the task of creating a DPM is assigned to you.

This document introduces basic principles deemed necessary in the modelling process. On the basis of the explanations given in this document you will be able to provide prerequisites for deriving data formats on the basis of a DPM as well as setting up a powerful data warehouse. This implies that the model is issued in a format that is understood by both parties, involved in transforming legislation into a model: business experts and IT specialists. The topics regarding supervisory reporting are kept short and limited to the content relevant for this paper. The idea is to convey the creation of the Data Point Model to you, as you are a supervisor with

---

<sup>1</sup> EBA (2011a), p.22

<sup>2</sup> EBA (2011a), p.30

<sup>3</sup> EBA (2011a), p.29

analytical capabilities and personal interest in this topic. No special IT knowledge is expected. The first sections will give you an overview on the required IT knowledge.

National banking supervisors have a mandate to evaluate the financial situation of financial institutions in their country. To be able to perform the necessary analytics, financial data is required from these institutions. The requirements are described in the form of texts and tables of data. To make a comprehensive model from these texts and tables a model is being created to enable IT support in communicating and storing the necessary data. A common problem with the NSA's is that IT staff has little financial background and financial specialists have little IT background. This makes data modelling a problematic area as both specialities are needed. This document is aimed at providing the tools and knowledge of creating a DPM by the financial specialists. The result, a model, can later in the process be perfected by IT staff.

## 1 Scope

This paper is a handbook for supervising experts. The main body consists of four sections. The interrogative form helps in choosing which section promises most answers to your problem.

After this first introductory section the main part starts to provide basic knowledge about different types of data models and data modelling approaches. The second and third section provide an overview of data models in general in contrast to the fourth section that highlights the necessity of data modelling for supervisory data. This fourth section derives the objectives based on the background information of the preceding sections. Furthermore one paragraph classifies the Data Point Model introduced by the Eurofiling Initiative and elaborated by EIOPA and EBA where many new terms related to DPM are introduced. A paragraph, which explains the areas of application for the DPM follows. The fourth section concludes with a paragraph introducing a subset of the technical constraints that need to be considered in the creation process of the DPM. The fifth section gives step by step instructions to create a DPM. The paper concludes with remarks on the progress achieved so far and provides an outlook on the software that is being developed at the moment to support you during the creation process. The last section also evaluates the DPM process to more traditional approaches. New terms are introduced throughout the text when they come up for the first time and can additionally be looked up in the glossary, which can be found in the appendix at the end of the paper.

## 2 Terms and definitions

For the purposes of this document, the following terms and definitions apply.

**NOTE** The terms definitions used in connection with Data Point modelling are inspired by vocabulary already known through their use for describing multidimensional databases and data warehouses. IT specialists originally introduced these terms. However, for an understanding and creation of Data Point Models they are now established in the language of business specialists as well.

### 2.1

#### **data point**

a Data Point can be compared to a cell in a table that holds reportable information and the row- and columnheaders characterising the Data Point can be regarded as the dimension and member combinations that apply to the Data Point

### 2.2

#### **default member**

a member in an enumerable dimension that will represent the dimension-member combination on a Data Point when that dimension is not explicitly associated

### 2.3

#### **dictionary element**

an abstract term for dimensioned elements, dimensions, domains and members

### 2.4

#### **dimension**

a dimension represents the “by” condition of a Data Point

Note 1 to entry: Dimensions literally describe the dimensioned element in order to limit the range of interpretation and thereby qualify the dimensioned element. One dimension either has a definite (i.e. countable) number of members, which is called an explicit dimension, or an infinite number of members represented as values, that follow a specific typing pattern, which is known as a typed dimension.

### 2.5

#### **dimensioned element**

a dimensioned element shows the nature of the data by typing it. It holds information about the underlying structure of the cell that is specified. In IT contexts a dimensioned element is referred to as metadata

**2.6****domain**

a domain is a classification system to categorize items that share a common semantic identity

Note 1 to entry: A Domain provides therefore an unambiguous collection of items in a value range. The items of a Domain can have a definite, and therefore countable, number of items, or an infinite number of elements that follow a specific (syntax) pattern.

**2.7****domain member**

each element that is part of a domain is called a domain member

Note 1 to entry: It is also possible to have members that do not belong to a domain; they can refer to a dimension directly.

Note 2 to entry: Domain members can either be explicitly named or defined by a type.

**2.8****enumerable dimension**

an enumerable dimension is a dimension that “specifies a finite number of members

**2.9****fact**

a fact describes the quantitative aspects of data reported

EXAMPLE An amount, a number, a string of text, a date.

**2.10****hierarchy**

nesting (setting relationships in a parent-child like architecture) of dictionary elements

**2.11****non enumerable dimension**

a non-enumerable dimension “specifies an undefined number of [members] [...] [it] defines syntactic constraints on the values of the members, i.e. a data type or a specific pattern

**2.12****sub-domain**

a sub-domain is a subset of the members of a domain

**2.13****taxonomy**

a taxonomy describes a valid Data Point Model

**2.14****templates**

graphical representation of a set of supervisory data

**3 What is a data model****3.1 Introduction**

Data models outline the relationships between data.<sup>4</sup> It is important that the person responsible for modelling takes time to capture all relations between data that can be shown in the model. It is essential that the model is reviewed by third parties involved. Thereby errors can be identified in advance. Furthermore it helps to get a clearly structured model that can save time and costs later.

---

<sup>4</sup> Cf. Gartner (2012)

### 3.2 The term “model”

The term model has its origin in the French noun “modèle”. In IT context a model pictures a target-oriented system instead of directly intervening in the complex system.<sup>5</sup> Specifically in terms of data models this means a real system, a system from the domain comprised of real components that are tangible and dynamic, is mapped to a model to reduce complexity.<sup>6</sup> This may help to find a suitable solution to an existing problem. The model needs to be created as close to reality as possible with attention to requirements regarding structure and behaviour. Nevertheless, in order to raise the comprehensibility, aspects irrelevant for the purpose of modelling may be left out. The importance of a single aspect and whether it is worth being specified in the model is depending on the decision of the domain experts. This strongly depends on the modeller’s understanding, creativity and capability to associate the object system with the model.

The challenge of data modelling is that a data model “must be simple enough to communicate [it] to the end user [...] [and] [...] detailed enough for the database design to use to create the physical structure”.<sup>7</sup> The same principle applies to message design and its physical representation.

In the following paragraph the procedure of data-oriented modelling is presented.

### 3.3 Data-oriented process of modelling

The data-oriented process focuses on describing the static structure of the reporting system in contrast to the function-oriented process, which begins with modelling the functions of the reporting system and adds the data in a later stage.

As data is the focuspoint of the banking supervisors the data-oriented process is applied. Additionally, in the course of time, data [objects] do not change as much as functions do. Functions are not being taken into account here.

Applying the data oriented process, data objects are specified first as well as the attributes that belong to each data object. The next step is to put the objects in relation to each other. Furthermore the data model can imply integrity conditions and define operations that can be carried out on the data.<sup>8</sup>

### 3.4 The conceptual data model as a first step aiming for a database system

The data-oriented modelling takes place on 3 different levels that are built upon one another.

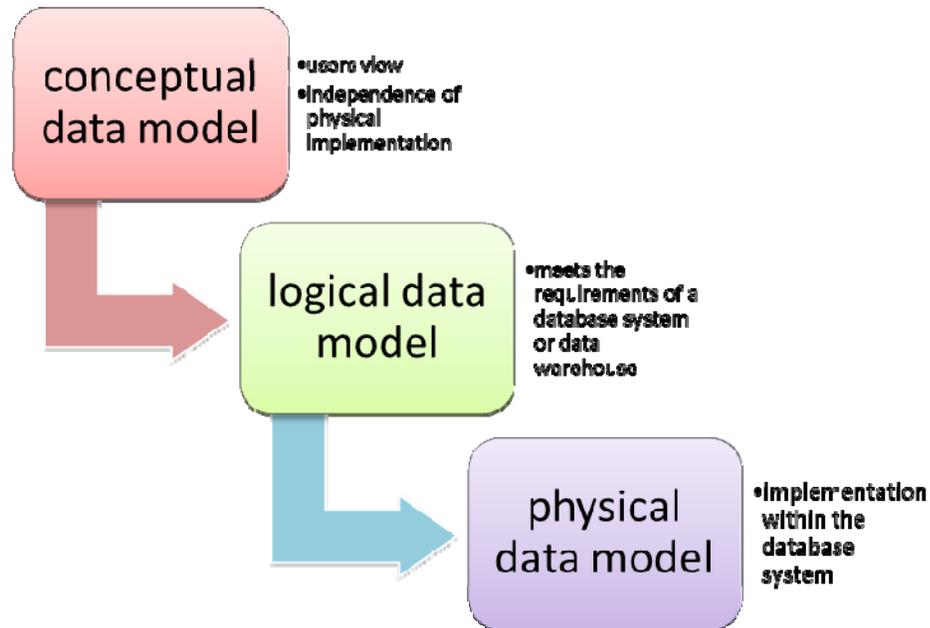
---

<sup>5</sup> Cf. Ferstl, O./ Sinz,E. (2013), p. 22

<sup>6</sup> Cf. ibidem, p. 20

<sup>7</sup> ZaZa Network (2007)

<sup>8</sup> Cf. Baeumle-Courth P./Nieland S./Schröder H. (2004), p.56



**Figure 1 — Levels of data-oriented modelling**

The conceptual data model reflects your reporting requirements. You are in the best position to know what pieces of information are requested. The conceptual model helps you in the communication with your IT specialists. This is an important step to avoid unpleasant surprises later when the model is implemented in the IT department. The model is built regardless of the database system or data warehouse to be used.<sup>9</sup> Relevant facts of the object system are to be specified without loss of information. However, you, as the creators of the conceptual model do not need to be technically skilled as the succeeding steps of data modelling are carried out by IT specialists. They should be concerned about the technical requirements. It is very important that this first step of preparing the conceptual data model is carefully elaborated before transferring the information to the IT. This can be ensured by early reviews, which include all parties concerned.

The logical data model as well as the physical data model are prepared by the IT specialists. In essence, the logical data model immediately follows the conceptual model (see Figure 1). When aimed at a database approach in contrast to the conceptual model it also takes the requirements of the database or the data warehouse into account.<sup>10</sup> The physical data model as a final step describes the actual implementation into an existing database system.<sup>11</sup>

### **3.5 Description of data modelling approaches for supervisory purposes**

#### **3.5.1 Introduction**

This paragraph deals with the methods that are used to disseminate data and identify all of its appropriate aspects. The two most appropriate methods of expressing regulatory data in a structure to determine the context this information is associated with, will be discussed here.

Both modelling approaches refer to metadata.

Definitions for data and metadata are given below:

<sup>9</sup> Cf. 1keydata (2013a)

<sup>10</sup> Cf. 1keydata (2013b)

<sup>11</sup> Cf. 1keydata (2013c)

Data is “information processed or stored by a computer. This information may be in the form of text documents, images, audio clips, software programs, or other types of data. Computer data may be processed by the computer’s CPU and is stored in files and folders on the computer’s hard disk.”<sup>12</sup>

Metadata “describes data. It provides information about a certain item’s content.”<sup>13</sup>

While data is a number like “50” the metadata adds qualifying information to the number. The explanation on the “form centric” and the “data centric” modelling approaches will clarify the difference.

**3.5.2 Using the “form centric” modelling approach**

The “form centric” approach is an ordinary table format with information held in a cell of a predefined table called a template. Here a template is understood as a graphical representation of a set of supervisory data. This approach identifies reporting data by their position in the templates. In this case each datum is defined by its coordinate in the table that is represented by the combination of columns and rows of a template. Each coordinate has a code that is based on the row code and the column code. This means that the data reported on basis of coordinate codes is meaningless without the context of the template. In the following example, each cell that represents a data requirement is described by a code combination of its column and its row of the table Market Risk: Standardised form for position risk in equities (MKR SA EQU) of the COREP framework. The form represents market risk equity positions of the institutions that are subject to mandatory reporting. Throughout the whole document this table serves as an example to introduce terms and concepts of Data Point modelling to you. The table with annotations can be found in the appendix in full size in order to deliver better clarity.

MKR SA EQU		MARKET RISK: STANDARDISED APPROACH FOR POSITION RISK IN EQUITIES									
		POSITIONS				RISK CAPITAL CHANGE (%)	OWN FUNDS REQUIREMENTS	TOTAL RISK EXPOSURE AMOUNT			
		GROSS POSITIONS		NET POSITIONS							
		LONG	SHORT	LONG	SHORT						
		010	020	030	040	050	060	070			
010	EQUITIES IN TRADING BOOK							Cell linked to CA			
020	General risk						8,00				
021	Derivatives										
022	Other assets and liabilities										
030	Exchange traded stock-index futures broadly diversified subject to particular approach										
040	Other equities than exchange traded stock-index futures broadly diversified										
050	Specific risk						8,00				
080	Particular approach for position risk in CIUS										
090	Additional requirements for options (non-delta risks)										

**Figure 2 — Table MKR SA EQU as an example of a form centric approach<sup>14</sup>**

The “form centric” approach is oriented at the visualization of the data. Dependencies between the codes of the data are only shown in the templates, i.e. by identifying the appropriate headlines or by the indents of the label rows. A report based on the “form centric” approach, which uses codes for the identification of data, is not able to incorporate the dependencies visible.

<sup>12</sup> TechTerms (2013a)

<sup>13</sup> TechTerms (2013b)

<sup>14</sup> EBA (2013)

MKR SA EQU		MARKET RISK: STANDARDISED APPROACH	
		GROSS POSITIONS	
		LONG	SHORT
		010	020
010	<b>EQUITIES IN TRADING BOOK</b>		
020	General risk		
021	Derivatives	r021c01050	
022	Other assets and liabilities		
030	Exchange traded stock-index futures broadly diversified subject to particular approach		

**Figure 3 — Close up of table MKR SA EQU for higher visibility on important aspects**

On the basis of the section of sample table MKR SA EQU shown in Figure 3 the “form centric” approach is explained. The value reported by the monetary institution in each cell is called a fact. Facts are classified as data. Let us say the oval circled cell defined by the row position r021 and the column position c010 holds the monetary value 50. The coordinate code r021c010 in the red circle is the combination of the row position followed by the column position. Taking the template into account we realise the number “50” represents a value for derivatives as a gross position. When we include additionally the headline above column c010 we can conclude that a long-term position is reported.

Looking at the excerpt it is not specified to which year this information belongs. Neither do we know whether 50 represents a value in thousands or millions nor can we conclude its currency. We can imagine that it would be really hard for a non-supervisor to correctly classify this information 50. Now if you think about the table shown in Figure 3 again, what would that numbers tell you if you would not have any headlines labelling the rows and the columns? Obviously the information would be useless.

As a conclusion we see that the “form centric” approach doesn’t include information about the data reported, which is assumed to be known (like all figures are in thousands). Moreover without the context of the row and column position of the datum the information content is essentially zero.

### 3.5.3 Using the “data centric” modelling approach

In the “data centric” approach, data is identified by a set of characteristics. It is considered independently of its graphical representation by adding information that unambiguously defines the datum. Therefore no positional alignment is needed in order to give the datum a specific meaning. Any datum is expressed in terms of the categories necessary for their identification.

Information available is divided into two groups:

- qualifying information;
- quantifying information<sup>15</sup>.

Qualifying information is represented by attributes to certain categories while quantifying information describes the object evaluated.

Figure 5 shows a dimensioned element which holds the information about the main character of the datum to be reported. A dimensioned element shows the nature of the data. It holds information about the underlying structure of the cell that is specified. In IT contexts a dimensioned element is referred to as metadata. In our

<sup>15</sup> Cf. Sapia, C. / et al (1999)

example the dimensioned element specifies the amount type of the datum as a gross value. The corresponding categories called dimensions contain further information on the datum and therefore increase the quality of the datum to be reported. The dimensioned element as well as the dimensions belongs to the group of qualifying information, i.e. metadata. The number itself, “50” in our example, is called a fact and represents the quantifying information of the datum.



**Figure 4 — Example of a dimensioned element with corresponding dimensions for the cell r021c010 marked in MKR SA EQU**

One Data Point is represented by one cell of the table in the “form centric” approach. Going back to the example above used to explain the “form centric” approach defining the cell by a combination of row and column codes (like r021c010) we have got a Data Point specified by a dimensioned element with its corresponding dimensions. One possible dimension for example that can be derived looking at the table in Figure 3 is the risk type dimension. Various types of risk are listed in the rows of this table. “general risk” and “specific risk” are reasonable attributes for the risk type dimension. To identify the risk types business knowledge is needed. We cannot rely on the nesting (tabs) in the table as they might be used differently

amongst table creators for presentation purposes. Each dimensioned element is characterised by a variable number of dimensions. Each dimension is linked to one attribute, called a member, to characterise the Data Point. The dimensions represent the “by” conditions. Dimensions literally describe the dimensioned elements in order to limit the range of interpretation and thereby qualify a dimensioned element. One dimension either has a definite (i.e. countable) number of elements, which is called an enumerable dimension, or an unknown list of members to the regulator, which is called a non enumerable dimension<sup>16</sup>.

Members are attributes that can be assigned to a dimension. As members are often used for various dimensions, domains are introduced in order to reduce redundancy. Each domain contains semantically correlated members that can be used throughout the whole of the reporting framework. The dimension represents the semantic relevance for the specific use on the dimensioned element. All members are added to at least one domain that can be reused by a variety of dimensions.

Returning to the difference between metadata and data, the definitions are transferred to the vivid example of MKR SA EQU. The Data Point identified by the row and column code combination r021c010 in the table format holding a fact “50”, which can be referred to as data. The metadata is described by the dimensioned element specifying 50 to be a gross value and the selected domains, one for each applied dimension.

It should be ensured that each Data Point is defined only once in a reporting framework, regardless of whether it is included in more than one table. One major benefit is that the information can be assembled in various ways based on the preference of the supervisory expert. Therefore the form of the tables can be aligned with the previously used “form centric” tables. This results in a minimum adaptation time for the filers.

### 3.6 Description of dimensional modelling

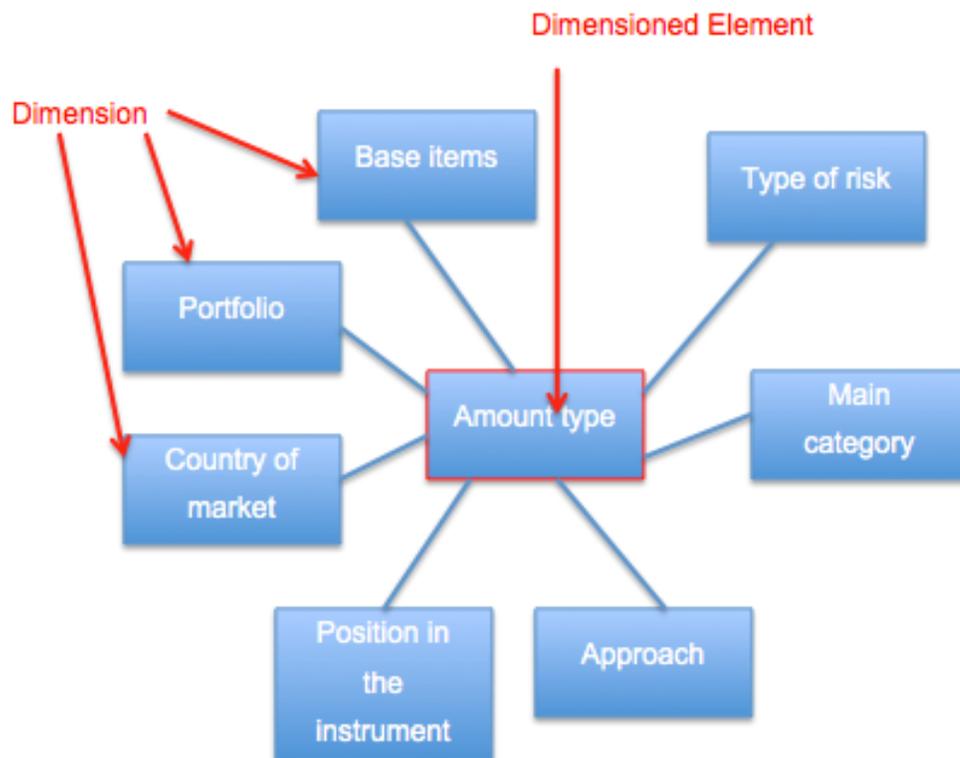
Dimensional modelling is the innovative modelling type to create multidimensional data models. Depending on the conditions, the dimensional model may be “simpler, more expressive and easier to understand”<sup>17</sup> than divergent modelling techniques. Dimensional modelling is used by the data centric approach introducing dimensions to qualify the information that consists of numeric data at the forefront like values, counts, weights, balances and occurrences. The main information about the datum, i.e. the data type of the fact, is held in the dimensioned element, which is verified here by the amount type dimension as it contains crucial information about the Data Point to be specified. Further qualifying information that is associated with the Data Point is specified by the members of the applied dimensions.<sup>18</sup>

---

<sup>16</sup> Cf. Declerck, T./ Hommes, R./ Heinze, K. (2013)

<sup>17</sup> Ballard, C./et al (1998), p. 42

<sup>18</sup> Cf. Ballard, C./et al (1998), p. 42



**Figure 5 — Dimensional model for MKR SA EQU**

The term ‘metrics’ is used as a synonym for ‘dimensioned element’ in other sources.<sup>19</sup> However for the rest of this paper the term dimensioned element is used. Taken literally it is the one that is defined by the application of dimension-member combinations.

### 3.7 The concept of normalisation

As stated before the redundancy is to be reduced by the use of the Data Point Model. The most popular approach to achieve this is the process of normalisation. As this is an IT specific proven concept it will be introduced to you in this paragraph.

Figure 6 shows what a typical table created by business users looks like. The values are reported in order to store them in a database and carry out analysis.

<sup>19</sup> Declerck, T./ Hommes, R./ Heinze, K. (2013)

Figure 6 — Sample table for products sold on certain days

Scanning the table many questions remain unanswered for the untrained reader. Hereinafter is a list of questions that shall serve as a set for thoughts:

- Unit of measure: What does “50” mean? Units? Currencies?
- Reporting entity: Are the values of a single country or institution?
- Definition of the used members: What is considered as derivatives?
- ...

This set of questions was developed in very short time only. It is obvious that it is important that the reporting entity and the supervisor share the same vision. In order to avoid discrepancies in the interpretation of the figures the table must be unambiguous.

In order to leave no room the questions above need to be answered. The information held in the figures of this table must be made explicit to all users on both ends of the communication process.

Another way to express the same facts in order to answer some of the questions raised, could be in plain text.

The cell r021c010 of MKR SA EQU holds the, for you as banking supervisors obvious, following information:

50,000 € worth of derivatives were held by a certain institution at a certain date.

All the cells in the table are reported by one institution and each Data Point in that table is to be sent for one reporting date.

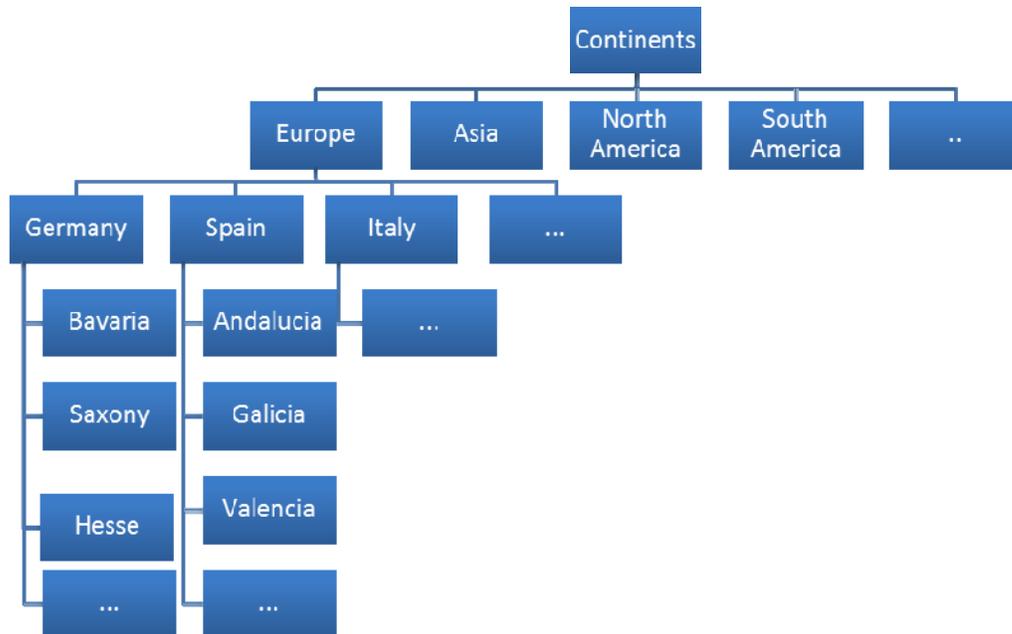
It is obvious in this method of representation that for all facts stored in the example table MKR SA EQU are

- of monetary value;
- in one common currency;
- reported in the precision of thousands.

It is still not yet known who reported the figures. Furthermore there is no definition of the axes’ members. The members that add qualified information about a single value need to be specified in order to prevent

discrepancies in the interpretation of readers. The task now is to check what level of detail is required for the facts reported in order to carry out the demanded analysis at a later stage. On the basis of this decision abstract categories are created. It is advised to carry out this task in a team of experts.

For example if we want to analyse the credit risks taken, it might be interesting to not only obtain knowledge about the countries where the risk was taken but also about the different regions within the countries as this might reveal a difference in the risk aversion of the various regions. Therefore it is not sufficient to name a category “country” and list below all countries. Referring to the mentioned example a further breakdown is needed that lists the regions of each country. For these different levels of detail a hierarchy can be defined in order to derive aggregated information about one country or one continent later. A sample breakdown with selected continents, countries and regions is shown below.



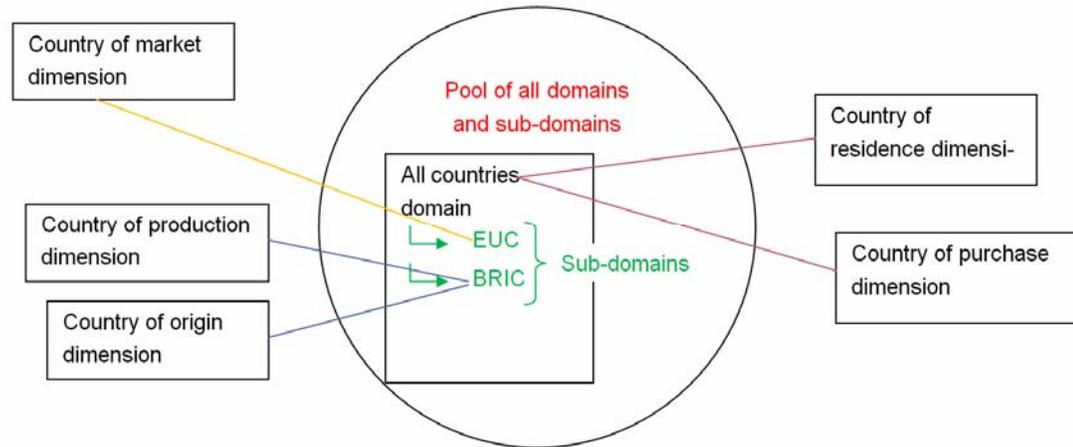
**Figure 7 — Hierarchy of countries to show different levels of detail**

The country category is just an example to make you aware of the level of abstraction you may choose for the categories identified.

A list of the identified categories to the facts reported in the table above (Figure 6) follows.

- A monetary value: some numeral data type.
- In a currency: closed list of currencies allowed.
- In thousands: closed list of precision types allowed.
- A reporting period or a point in time: A closed list of periods as all reports are required to cover predetermined periods.
- If the figure was reported by a single bank, a closed list of all banks that report to the national supervisor may be a good way to categorise the fact.
- An explanation document of the axes’ members is needed to be referred to, where each member of each dimension applicable for MKR SA EQU is unambiguously defined.

Each member must be created only once and allocated to one domain. The members must be created consistent and without doubling of logically same elements under different labels. The domains can be assigned to dimensions. Suppose we created the full hierarchy like visualised above, we could assign a (sub)domain called 'European countries' to a dimension named 'country of market'. In this domain all the European countries would be listed. Also there could be another (sub)domain called 'BRIC' containing the countries Brazil, Russia, China and India. This BRIC (sub)domain could be assigned to two dimensions, the 'country of origin' dimension and the 'country of production' dimension. Last but not least we could build another domain called 'all countries' where all the members that are already assigned to other (sub)domains as well as remaining countries are included. This domain can again be assigned to multiple dimensions. The figure below pictures this scenario.



**Figure 8 — Pool of shared domains**

Once domains are created, domains can be assigned to a variety of dimensions. That prevents redundancy of members and defines them uniquely for satisfying the requirements of communication via computers. This step is called normalisation. A technical definition for normalisation is as follows:

Normalisation is the transfer of a data model to a certain state. The various states are differed by levels of the 'normal form' and achieved by applying them on the data model. The third normal form is enough to prevent redundancies and inconsistencies. Therefore the maintenance of the held data is facilitated applying the third normal form.<sup>20</sup>

To achieve this, the two main aims are:

- arranging data into logical groups such that each group describes a small part of the whole<sup>21</sup>;
- restrict to the level of detail needed<sup>22</sup>.

In order to bring your data model into the third normalised form you need to group members in domains and make sure that the domains do not overlap. It must be possible to unambiguously assign the members to a single domain. Therefore it is important to use meaningful names for members, domains and dimensions. It is also advised to prepare a handbook where the names are differentiated. Following these rules, consistency throughout the model can be achieved.

<sup>20</sup> Cf. Minhorst, A. (2005), p. 49

<sup>21</sup> databasedev (2013)

<sup>22</sup> Heinze, K. (2013)

## 4 Why use a multidimensional data model

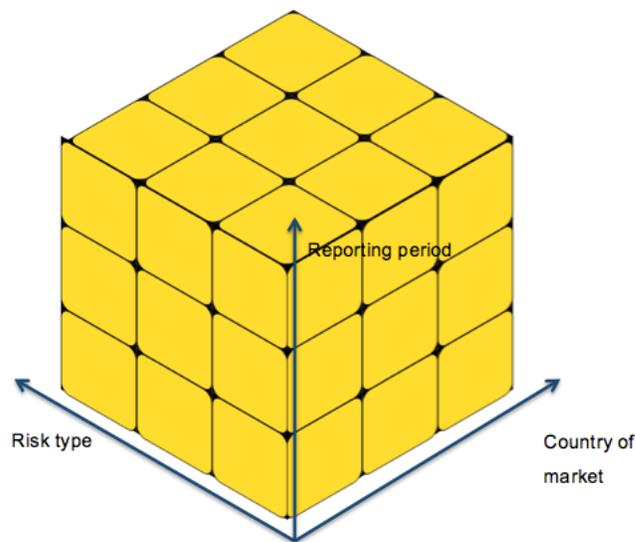
### 4.1 Introduction

The data in the conceptual model can be modelled dimensionally as well as hierarchically<sup>23</sup>. The reason it is advised to create a multidimensional data model, is that it is closer to the presentation form that the user is accustomed to and therefore easier to understand for him.

### 4.2 Multidimensional data model

The multidimensional data model supports the “data centric” approach with its two groups: qualifying and quantifying data.

In order to make it clear we go ahead with the example of MKR SA EQU that you are already familiar with. We simplify the model below to show three categories only in order to improve the clarity displaying it on paper.



**Figure 9 — Multidimensional model**

The multidimensional data model visualised by a cube in Figure 9 is specified by three categories: risk type, reporting period and country of market. These categories are referred to as dimensions and, as stated before, serve as examples for qualifying information. The single cells that make up the cube carry quantifying information. Most of the time Data Points hold values that can be summed upon demand.

The dimensions risk type, reporting period and country of market that show a semantic relationship between them are used to specify an orthogonal<sup>24</sup> structure to the data space.

It is possible to carry out arithmetic operations on the numeric values in each cell.

Two major advantages with this modelling technique are:

- firstly, the collected figures are each represented once in the model, and
- secondly, the ratios on a higher level of aggregation can be computed by means of the existing values.

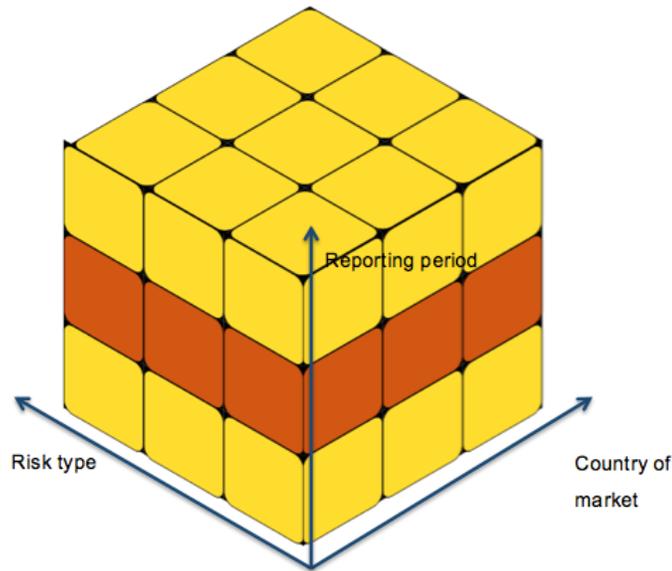
---

<sup>23</sup> Collins, J. (2013)

<sup>24</sup> meeting at a right angle

### 4.3 Operations that can be carried out on a multidimensional data model

It is possible to create individual views on the present extensive multidimensional data model. One approach is to look at slices of the large whole. This is often visualised by referring to a single selected domain of one of the dimensions and therefore receiving figuratively a slice of the cube. Actually one might say that one dimension is not taken into account with this view on the cube.<sup>25</sup>

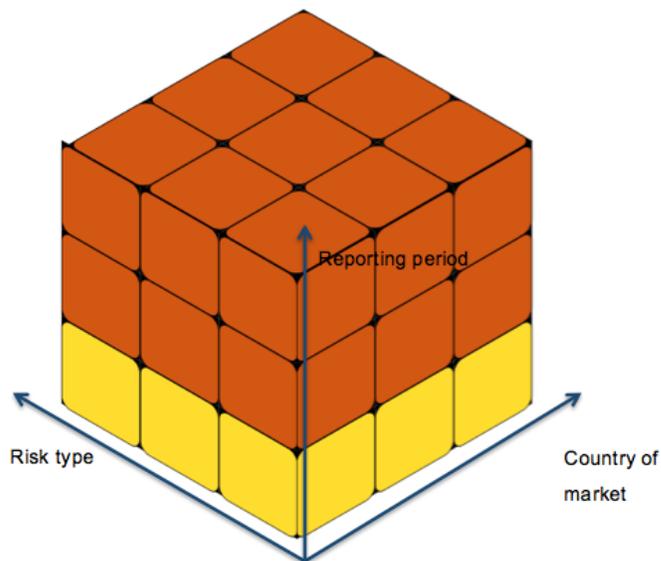


**Figure 10 — Slicing visualised**

Referring to the example cube shown in Figure 10 we focus on the orange highlighted part. By slicing we get all reported risk types of all countries of market at a certain reporting period. Whether the reporting period situated on this dimension is a domain describing days, months, quarters of the year or even whole years remains to be seen.

With dicing in contrast to slicing all dimensions remain considered. The process of dicing figuratively cuts a hexahedron out of the big cube.

<sup>25</sup> Cf. Verma, R. (2009a)



**Figure 11 — Dicing visualised**

Adhering to the same example Figure 11 pictures the effect of dicing. According to the model cube one attribute on the reporting period dimension is excluded for the analysis. Therefore dicing results in a new hexahedron smaller than the original cube.<sup>26</sup>

Figure 11 represents the idea of looking at the more recent reporting periods leaving out the figures of reporting periods long ago. As the exemplary Figure 11 is much larger in reality it is also representative for analyses that are carried out to compare the figures of a given period of years, like certain decades. The difference to the slicing is visualised in Figure 11. By having multiple attributes of each dimension coloured in orange, the dicing process takes multiple characteristics of all dimensions into consideration.

## **5 Why data modelling is essential for collecting supervisory information**

### **5.1 Introduction**

The massive amount of information reported and the request to analyse this data in many different ways appears to be problematic if the data is not structured in any way. A new type of data modelling was introduced by the Eurofiling Initiative called Data Point modelling. It is meant to combine the advantages of the various data modelling types in regard to supervisory reporting. Data modelling is essential for all participants to enable the communication of clear and unambiguous definitions of terms used in the reporting framework.

### **5.2 Objective of Data Point modelling**

The Eurofiling Initiative is about to set a syntax standard for collecting information for supervisory and statistical reporting. The aim is to benefit from international solutions instead of proprietary ones. E.g. validation software for data received, mapping software for transforming the received data into databases and rendering software to make the exchanged data visible to parties that are not directly involved in the communication like accountants and actuaries. The data format to which a DPM can be transferred later is variable. At present the preferred standard syntax is a format called Extensible Business Reporting Language (XBRL).<sup>27</sup> It was chosen because of its characteristics being adapted to the requirements of the financial sector. The use of XBRL does not imply an enforced standardisation of business reporting. On the contrary, the syntax is a flexible one which is intended to support all current aspects of reporting in different countries

<sup>26</sup> Cf. Verma, R. (2009b)

<sup>27</sup> Cf. Piechocki, M. (2012)

and industries. Its extensible nature means that it can be adjusted to meet particular business requirements, even at the individual organization level.

Moreover the EBA has given signals that XBRL will be the format that it will require to receive the data collected by national authorities in.<sup>28</sup>

The four main reasons for modelling Data Points (whether using XBRL or not remains to be seen) are illustrated in the following paragraphs.

The DPM is a multidimensional model. As an example the figure below represents the cell r021c010 of Figure 12 of the table MKR SA EQU.

The dimensions are coloured in dark red. The members of the domains that are assigned to the dimensions are coloured in a light red colour. The applicable domain members for each of the dimensions are made visible in the centre of the figure in green colours.

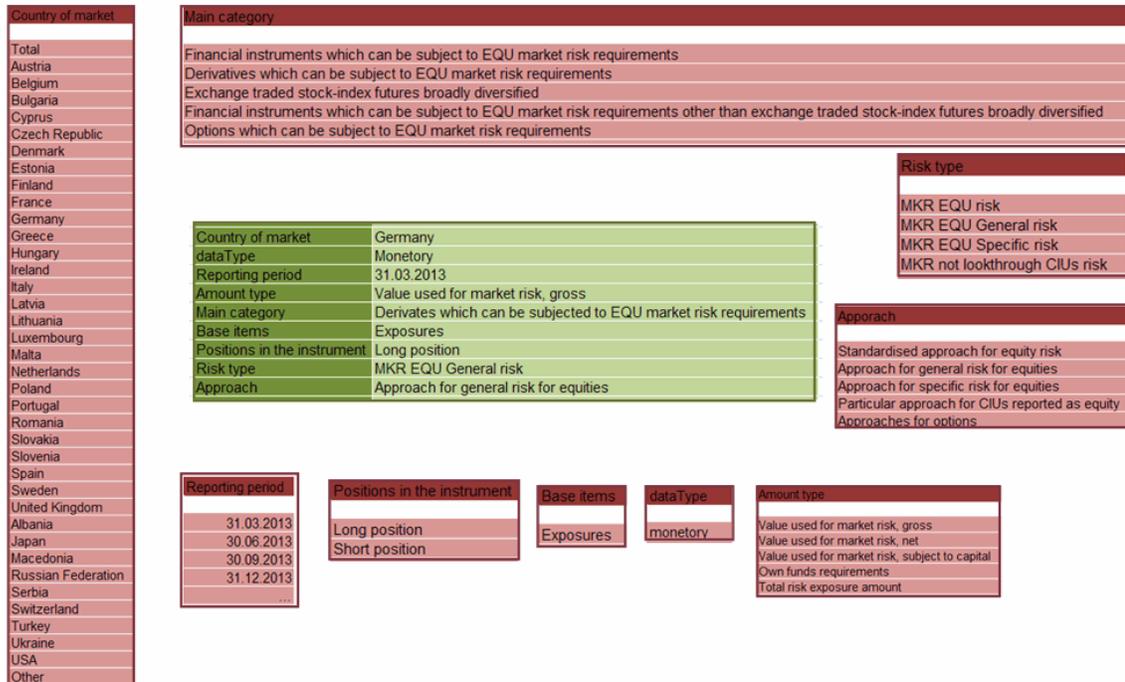


Figure 12 — Example of Data Point Model visualised

## 5.3 Main features

### 5.3.1 Increase of knowledge and understanding

As the Data Point Model is built by you, the supervising experts, it is ensured that the know-how is transferred in a data model which shows data required in the appropriate detail. In order to create a sustainable system it is important to gather not only the information needed at present, but all details to the collected data that can be specified and might gain in importance in the future. Using the concept of Data Point methodology it is ensured that the data is arranged comprehensible for the supervisory department. It is not only the data that business specialists are most familiar with. The relations within the information is another reason for the transfer of the task of building a Data Point Model to you as supervisory experts. The creation of the Data

<sup>28</sup> Cf. EBA (2011a)

Point Model underpins the already existing knowledge held by you and makes the transformation of the information to the IT specialists possible.

### 5.3.2 Improvement of integration of changes

With a well designed Data Point Model it can be ensured that the data structure is defined explicitly and without redundancies. This means that no single fact is described in two different ways. Therefore every single piece of information is unique. If more information is required qualifying aspects may be added to the fact in conjunction with the construction of a new dimension as needed. Figure 13 shows this case.<sup>29</sup>

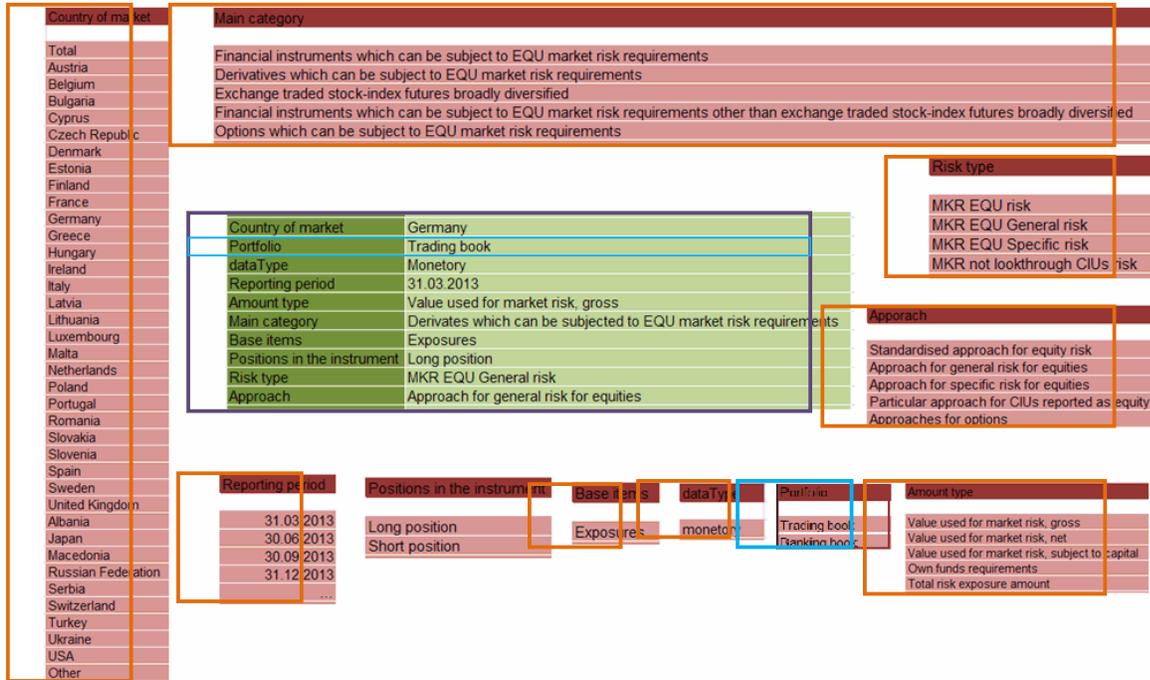


Figure 13 — Extensibility of Data Point Model is shown by adding a portfolio-dimension

The portfolio dimension (framed a light blue) was added as requirements in relation to distinct trading book and banking book have to be applied. It is unproblematic to add new dimensions when they are requested. This is very important for analysis by the data warehouse later as well as slicing and dicing, which is explained in Section 4.3. The out-dated requests do not have to be modified. They are still showing the same results on an expanded Data Point Model. This makes integration of changes very easy.

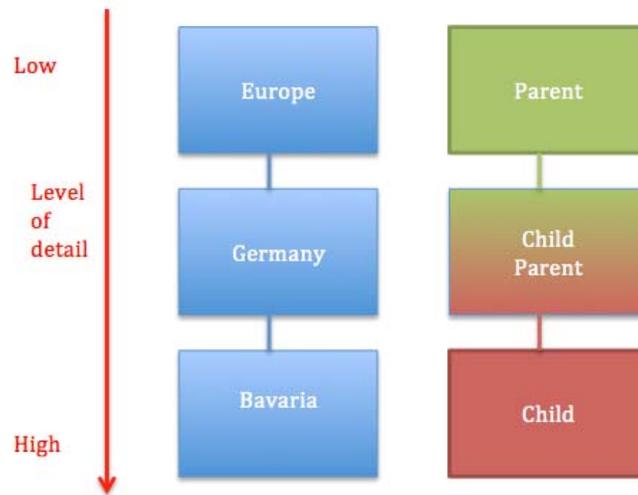
### 5.3.3 Reduction of risk of duplicate information

This goal refers to the avoidance of duplicate information. With normalization on Modelling Data Points, dimensions and members can be reused. As explained in previous sections, it is advised to combine members in a domain possibly also sub-domains, which can then be associated to a dimension. Hierarchies are defined to group sub-domains of already existing domains.

Most of the time we can identify different levels of detail for members of one domain. This means a kind of natural hierarchy is formed. You can represent these members of different levels of detail by sub-domains. We try to picture these relationships as hierarchies as this information can be reused for the definition of rules for calculations (total has individual facts). Hierarchical presentation and understanding how members are

<sup>29</sup> Heinze, K. (2012), p. 79

interrelated are further purposes of defining hierarchies. In hierarchical modelling this is called a parent-child relationship, which is figuratively shown below.<sup>30</sup>



**Figure 14 — Shows the relations of the parent-child relationships with Germany in the focus**

With Germany as an example for one country at the forefront of our thinking we can identify each one of the 16 German states like Bavaria, Saxony and Hesse as children of the country Germany. However, Germany can also take the place of a child if we add the continents to our context.

This means that one continent consists of several countries. Each single country may be composed of states. The advantage that can be derived from hierarchies is better explained by another explicit example. If we store the data at a level of detail that represents every state, the figures for country as well as continent can be computed. It is possible to aggregate the states of each country simultaneously. If required we can also aggregate the countries of one continent in order to receive the information on continental basis.

As it is possible to compute the lower levels of detail from the higher levels of detail it is advised to store the information in the highest level of detail accessible.

In order to build a Data Point Model, which can be used and maintained in the future, hierarchies should be built. The information about the nesting of members in a hierarchy improves its understanding by humans and helps to arrange potential new supervising criteria. Another use for hierarchies is to express the possible mathematical relationships between members if they are assigned to numerical dimensioned elements. A 'total' dimensioned elements can be comprised from multiple 'detail' dimensioned elements all carrying a different member. The validation rules shown below in the Excel file provide a basis for possible hierarchies.

<sup>30</sup> Cf. IBM (w.y)

Table number	TableBasedFormula
MKR SA EQU	{r010;c060}={r020;c060}+{r050;c060}+{r080;c060}+{r090;c060}
MKR SA EQU	{r020}={r021}+{r022}
MKR SA EQU	{r020}={r030}+{r040}
MKR SA EQU	{r020;c050}*0,08={r020;c060}
MKR SA EQU	{r050;c050}*0,08={r050;c060}
MKR SA EQU	{r090;c060}={r100;c060}+{r110;c060}+{r120;c060}+{r130;c060}
MKR SA EQU	{r010;c060}*12,5={r010;c070}

Figure 15 — Validation rules for MKR SA EQU

Figure 16 shows a hierarchy for the risk type domain. Having the excerpt from an Excel file above as well as the belonging table MKR SA EQU with its row and column positions listed, we are able to derive a clearer view of the hierarchy of the members contained in the risk type dimension.

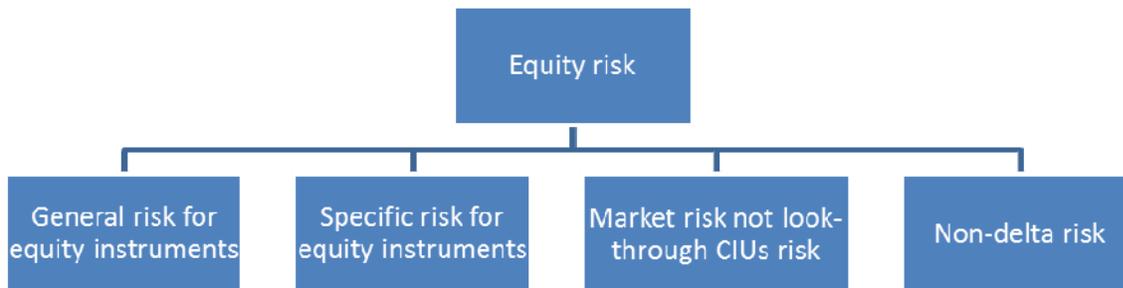
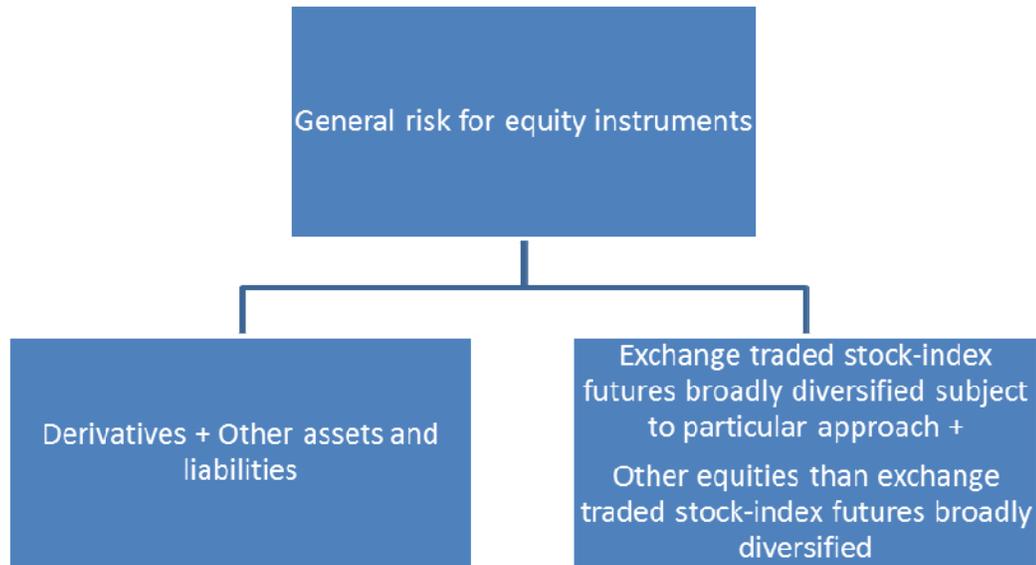


Figure 16 — Hierarchies of risk type domain depicted

Moreover from the second and third row of the validation rules depicted Figure 15 we can derive further information about the composition of the general risk listed in row 020 of table MKR SA EQU.

With a new risk to be reported the decision is to be made whether the risk is at top level contrary to the equity risk or below the equity risk member and therefore in the same level as the four types of risks depicted above in Figure 16, which builds up the equity risk. It is also possible that there is a change in regulation that requires splitting up one of the lower level risks if additional consideration is demanded. According to this scenario a third level of equity risks will be introduced further breaking down one of the second level equity risks like the example visualised in Figure 17.



**Figure 17 — Further breakdowns for general risk for equity instruments**

Furthermore sub-domains can clarify relations between members. A sub-domain is a subset of the domain containing a part of the whole. A sub-domain, just like a domain, can be assigned to a dimension. If we want to restrict the choice of members of a given domain to be assigned to a dimension, we can build a sub-domain containing selected members of the whole in order to reduce redundancy. One conceivable sub-domain for the country of market dimension can be labelled “European countries”, represented by the domain 'EUC', which is an acronym for the whole name. Its members would be all countries in European Union. Spain, Portugal, Germany as well as France and all other countries that belong to the European Union in a political point of view would be members of this sub-domain. Other domain keys contain different countries or additional ones, or parts of the ones in the example. Any, until now, non-existent combination of countries can be expressed by a new domain or sub-domain. However there might be another dimension like for example country of production. This dimension logically needs countries as members as well. It is possible to use any domain or sub-domain defined for any dimension. Figuratively, a pool of domains and sub-domains is created, which contains the domains and sub-domains to be chosen from for the dimension to specify.

#### 5.3.4 Higher harmonisation

Especially, the use of the “data centric” as well as the multidimensional approach, it is possible to carry out extensive queries in a data warehouse. The sharing of Data Points, equals in various reporting frameworks like COREP and FINREP, support the harmonisation process.

Based on the reporting frameworks COREP and FINREP as well as some other, smaller ones, common dimensions between these frameworks were identified to reach a higher degree of harmonisation by sharing dimensions and members across frameworks. The following figure shows the interlinkage between common dimensions across the universe of European reporting frameworks.

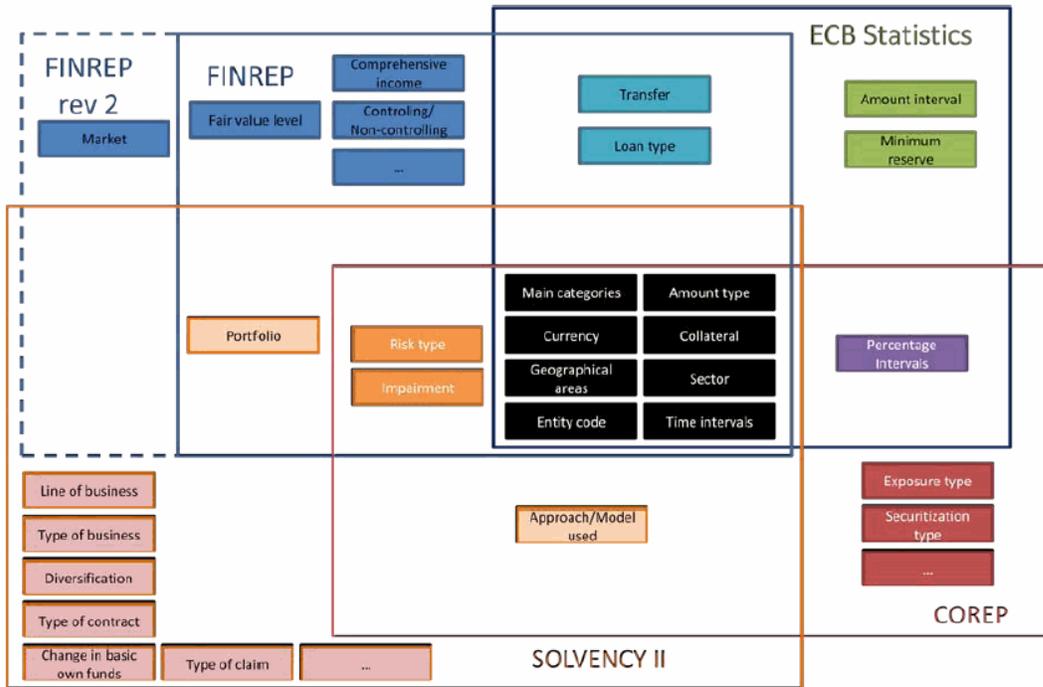


Figure 18 — Dovetail connection between different common reporting frameworks<sup>31</sup>

#### 5.4 Classification of Data Point modelling in the data modelling concept

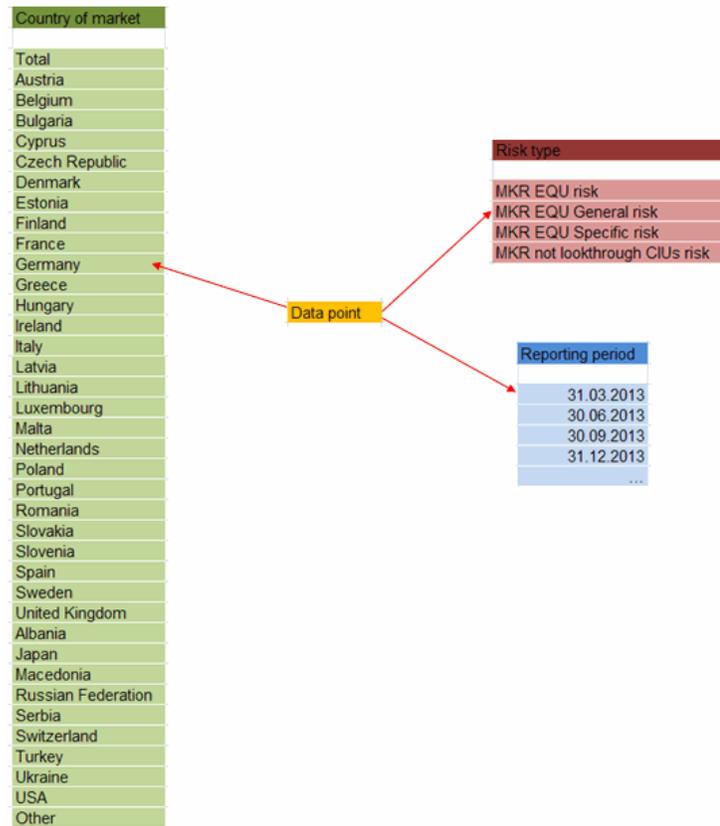
With the knowledge about data modelling gained in the previous sections we are now able to describe the characteristics of a Data Point Model. The concept of Data Point modelling is based on the “data centric” approach described in Section 3.5.3. This data structure eases the comprehensibility by business experts responsible for the creation of the Data Point Models. The “data centric” approach has further advantages. Moreover the gain in uncomplicated extensibility and the reduction of risk of duplicate information plead for the data centric design. Associated to these gains also is the next characteristic of Data Point Models that can be stated.

Without doubt supervisory reporting focuses on the data collected of the monetary institutions obligated to report.

The modelling of Data Points is part of the creation of the conceptual data model. The logical data model and the physical data model rely on a well-designed Data Point Model in the conceptual modelling stage. This is visualised above in Figure 1. The DPM is therefore to be created, well thought out and reviewed by interested parties.

A greatly simplified view of the Data Point representing the cell r021c010 of MKR SA EQU with only 3 associated dimensions is visualised in the following figure. Possible combinations of members of the three chosen dimensions country of market, risk type and reporting period as visualised simplified in Figure 9 are presented below.

<sup>31</sup> EBA (2011b), p.50



**Figure 19 — Shows Data Point and three applicable dimensions: country of market, risk type and reporting period**

A Data Point is a combination of dimensions with each dimension pointing to one of its domain members. In a table a Data Point is represented by a cell. For example we gain knowledge of the MKR EQU General risk taken by all monetary institutions belonging to the German market in the reporting period reported by 30<sup>th</sup> of March 2013. Information can be filtered in many ways. Also, the information about any other risk type applicable for the table MKR SA EQU that was taken by the German monetary institutions by the reporting period of the 30<sup>th</sup> of March 2013 is available to us. Moreover we can find out the risk aversion for the different risk types of each countries' monetary institutions by the reporting period 30<sup>th</sup> of March. According to this scheme the information is clearly identified and therefore leaves less room for interpretation.

## 5.5 Area of application

The advantages of Data Point Models in the segment of supervisory reporting are especially recognised by the visualisation of reporting data in different views by using pivot tables. The tables can be aggregated, which allows compressed analysis.

MKR SA EQU				MARKET RISK: STANDARDISED APPROACH FOR POSITION RISK IN EQ			
				POSITIONS			
				GROSS POSITIONS		NET POSITIONS	
				LONG	SHORT	LONG	SHORT
				010	020	030	040
010	<b>EQUITIES IN TRADING BOOK</b>						
020	General risk						
021	Derivatives						
022	Other assets and liabilities						
030	Exchange traded stock-index futures broadly diversified subject to particular approach						

Figure 20 — Excerpt from the reporting table MKR SA EQU

A fact, in most cases, is a numeric value accompanied by dimensional properties in the form of dimension member combinations. The assignment of a Data Point to a cell may not be allowed. The cells coloured in dark grey show this case. For instance, when a Data Point would not make sense, as the type of content does not exist in reality, the cell is greyed out. Another reason for the regulator to not allow to report values for cells and therefore grey them out is if the regulator is just not interested in the value or is able to cumulate it.

The views enabled by pivot tables omit unnecessary detailed information for the analysis. The very detailed facts are aggregated in order to provide an overview for the user. Nevertheless the numbers represented in the table are of high quality as the facts that are once reported are broken down into their smallest possible units and can be cumulated subsequently if desired. Moreover the data and its metadata reported are in machine readable form, which leads to a benefit as the gathering of data only takes place once.

### 5.6 What are the technical constraints

Attention should be paid to some rules of which an excerpt is listed below. The source of these constraints is a Wiki that started off as a joint venture of XBRL Spain and the University of Bucaramanga.<sup>32</sup> As its aim is to develop a standard that is adopted by all parties, anyone interested is welcome to contribute ideas to the wiki. Amendments and additions to the content of the wiki are still possible and therefore the rules listed below are not final. It is assumed that additional constraints will evolve in the future as more and more people determine points of contact with the concepts of Data Point modelling and XBRL. It is strongly recommended that you follow these rules described below as well as the ones in the wiki.

For the DPM a couple of constraints, which are considered very important, were specified in connection with hierarchies.

- 1) All members must be part of some hierarchy being built by a domain and its members.
- 2) Any single member can only appear once in any single hierarchy.
- 3) The hierarchy is built upon rules that are defined in a set of hierarchy relationships.
- 4) Each hierarchy has to be built from exactly one root element.<sup>33</sup>

Moreover when using XBRL, additional rules to the ones defined for the DPM must be considered. Especially working with domains is further specified.

- 5) Each member has to be referenced by a domain.

<sup>32</sup> Cf. XBRL Spain (2012)

<sup>33</sup> Cf. Declerck, T./ Hommes, R./ Heinze, K. (2013)

- 6) For each domain one member is set as default.
- 7) One dimension has to point to at least one domain or sub-domain.
- 8) Each member must be unique.<sup>34</sup>

The most current and complete list of all constraints can be found at the wiki, which is "regularly updated with the help of the Eurofiling initiative and XBRL Spain"<sup>35</sup>. The filing rules in particular are updated by a CEN (European Committee for Standardisation) workshop.<sup>36</sup>

## 6 How do you proceed in creating a Data Point Model

### 6.1 Introduction

As it is likely that the reporting requirements will increase in the future, the Data Point Model has to be extended continually. This section gives you an understanding of an iterative process for modelling a Data Point Model for a delimited supervisory reporting area mostly represented by one or more templates.

The process flowchart is pictured below.

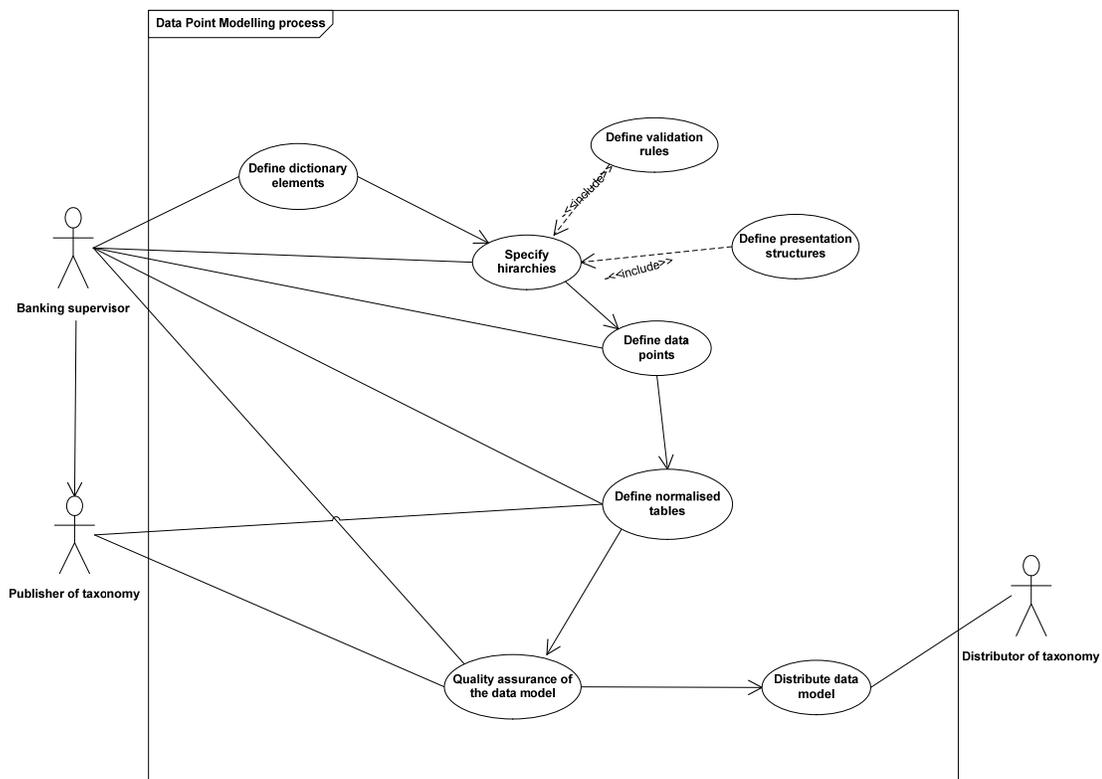


Figure 21 — Process of creating a Data Point Model

<sup>34</sup> Cf. ibidem

<sup>35</sup> XBRL Spain (2012)

<sup>36</sup> Cf. CEN (2009)

Your aim is to transfer the reporting data into the data model with regard to new analysis capabilities. An IT expert may contribute to the normalisation of tables and might carry out the quality assurance of the data model as he needs a complete and consistent data model in order to derive the taxonomy from it.

Moreover, data modellers must have the knowledge on how to create DPMs. We use an example again to explain the essential process.

### 6.2 Define dictionary elements

First of all we need to define dimensioned elements, dimensions as well as domains and their members. They form the dictionary elements of the model.

We start off with one business template. As we are already familiar with MKR SA EQU, we stay with this template.

MKR SA EQU		MARKET RISK: STANDARDISED APPROACH FOR POSITION RISK IN EQUITIES					RISK CAPITAL CHANGE (€)	OWN FUNDS REQUIREMENTS	TOTAL RISK EXPOSURE AMOUNT
		POSITIONS							
		GROSS POSITIONS		NET POSITIONS					
		LONG	SHORT	LONG	SHORT				
		010	020	030	040	050	060	070	
<b>EQUITIES IN TRADING BOOK</b>								Cell linked to CA	
010	General risk								
020	Derivatives		50				8,00		
021	Other assets and liabilities								
022	Exchange traded stock-index futures broadly diversified subject to particular approach								
030	Other equities than exchange traded stock-index futures broadly diversified								
040	Specific risk						8,00		
050	Particular approach for position risk in CItIs								
080	Additional requirements for options (non-delta risks)								
090									

Figure 22 — MKR SA EQU template

#### Distinction between quantitative and qualitative aspects

Having chosen a template, we have to distinguish between quantitative and qualitative aspects for each Data Point.

Quantitative are the figures reported. Like “50” for the cell identified by the row label “derivates” and the column “gross positions; long” (r021c010). We could also say the data, as defined in Section 3.5, belongs to the quantitative aspects.

Qualitative aspects are pieces of information given in order to reduce interpretability of the datum reported. Characteristics that specify the datum belong to this information, which are also called metadata.

#### Summary of quantitative aspects

The measurement to the dimensioned element needs to be added. There are two different types of time to be distinguished: “stock” and “flow”. Flows, in contrast to stocks, are representing durations, i.e. measures reported for a period like cash flows, revenue and costs. Stocks are, for example, assets and liabilities representing an instant for stocks. Therefore the measure is of a certain date.

The quantitative aspects in this template have the property of stock values as the numbers represent the market risk at a certain date.

#### Classification of the qualitative aspects in categories

Now we figure out domains by which the data can be grouped. We have for example different risk types, which categorise the data. General risk for equity instruments, specific risk for equity instruments, market risk not look-through CIUs risk and non-delta risk are the risk types that can be identified in the table.

#### *Creation of domains*

In order to prevent redundancies, domains are being created. Members that share the same semantic aspect are assigned to a domain, expressing this aspect.

The different risk types can be assigned to one common domain, as they consist of the same semantic identity. We call the domain “risk types for market risks for equity instruments” in order to give it a meaningful name. Moreover a domain that includes all countries should be created. To ease recognition we call the domain containing all countries “all countries”. The domains can be directly and indirectly derived from the template. A lot of information is obvious to you as banking supervisors, however the topic of defining domains should not be kept out here. One further example is given by using Euros for identifying the currency of figures. We may also add US-Dollars, Pound and names for other currencies that come to our mind to add them to a domain named “all currencies”. We could also introduce a domain that holds information about the multiplier the figure is related to. We see that most of the information can only be derived by supervisory experts, especially those pieces of information that are not explicitly given in the template. This step is successfully completed once all members for the description of the Data Points in a template are part of one domain.

#### *Definition of dimensions*

The next step is to define dimensions that refer to at least one domain. They provide a specific meaning for a domain when linked to a Data Point. A domain member and its corresponding dimension form one qualitative aspect of a Data Point.

The dimension for our MKR SA EQU template that refers to the “all countries” domain is called “country of market”. We give the dimension for risk types the same name as given to the domain. Finally we want all domains applicable in the MKR SA EQU template to refer to one dimension.

#### *Definition of a default member for each explicit domain*

For explicit dimensions (dimensions that carry a closed list of members) a default member must be defined. The default member is implicitly applied when a dimension is not explicitly associated to a Data Point. This is the case when a Data Point that has a dimensional context of 9 dimensions but only 6 dimensions are explicitly associated with according members, then the three further dimensions are implicitly included with their members set as default.

### **6.3 Specify hierarchies**

The next step is the specification of hierarchies regarding to a set of members as well as the definition of calculation rules and concepts for presentation purposes.

#### *Definition of hierarchies between domain members*

The relations between domain members must be specified by building hierarchical relationships. Three types of hierarchies are foreseen:

- parent-child relationships for presentational purposes (presentation relationship)
- summation-item relationships for aggregation purposes (rule relationship), and
- domain-member relationships that explain the semantics amongst members (basic relationship).

These can all be added in later stages.

Now the difference between risk types of the lower level of detail and risk types of the higher level of detail is specified. As shown in figure Figure 16 the members of the “risk types” dimension can be formed in a hierarchy with supervisory knowledge in order to allow the aggregation of members adding up to “general risk” or even “equity risk”. Furthermore in this step of the DPM creation process the sub-domains are defined. A good example, which was introduced before in S

Section 5.3.3 (see Figure 9) is the “all countries” domain. Sub-domains are the EUR sub-domain containing all European countries, as well as the BRIC sub-domain that include the Brazil, Russia, India and China.

**6.4 Define Data Points**

The third step is the creation of Data Points by building relations between one dimensioned element and its associated dimensions.

In our case the dimensioned element is based on the dimension amount type. For our sample in cell r021c010 the dimensioned element specifies a “value used for market risk, gross”. The applicable dimensions of the Data Point pictured in **Error! Reference source not found.** are as follows: type of risk, country of market, position in the instrument, main category, portfolio, base items and approach. When a Data Point is reported as fact it holds additional information about reporting entity and the period type. Also information about the unit, the decimals or precision and the language, if the fact is string based.. The identifier is a string of characters representing one reporting entity. The reporting entity is represented by an identifier. The period type gives information about the validity of the value reported. A value can be reported for a specific point in time or for a period in time.

**6.5 Define normalised tables and ensure quality of Data Point Model**

The fourth and the fifth step are carried out with the help of the publisher of the taxonomy.

National market:		MARKET RISK: STANDARDISED APPROACH FOR POSITION RISK IN EQUITIES				RISK CAPITAL CHARGE (N)		OWN FUNDS REQUIREMENTS		TOTAL RISK EXPOSURE AMOUNT		Dimensions:						
MKR SA EQU		POSITIONS				00	00	00	00	00	00	Total Austria Belgium Bulgaria Canada Czech Republic Denmark Estonia France Germany Greece Hungary Ireland Italy Latvia Lithuania Luxembourg Malta Netherlands Poland Portugal Romania Slovakia Slovenia Spain Sweden Switzerland United Kingdom Austria Japan Macedonia Pakistan Russia Switzerland Turkey USA Other						
		GROSS POSITIONS		NET POSITIONS														
		LONG	SHORT	LONG	SHORT													
		00	00	00	00													
000	EQUITIES IN TRADING BOOK											Call indicator CA						
001	General risk											MC/MCF financial instruments which are for subject to EQR market risk requirements						
002	Derivatives											MC/MCF financial instruments which are for subject to EQR market risk requirements						
003	Other assets and liabilities											MC/MCF other assets and liabilities which can be subject to EQR market risk requirements						
004	Exchange traded stock index futures broadly diversified subject to particular approach											MC/MCF exchange traded stock index futures broadly diversified requirements						
005	Other equities than exchange traded stock index futures broadly diversified											MC/MCF financial instruments which are for subject to EQR market risk requirements other than exchange traded stock index futures broadly diversified						
006	Specific risk											MC/MCF financial instruments which are for subject to EQR market risk requirements						
007	Particular approach for position risk in CDOs											MC/MCF CDOs requirements						
008	Additional requirements for equities (non-derivative)											MC/MCF equities which can be subject to EQR market risk requirements						

Figure 23 — Annotated template MKR SA EQU

The task now is to define normalised tables derived from templates and with regard to the dimensional possibilities within the table.

The table above that can be found in the appendix in full size, was created by supervisory experts and is now available for the taxonomy publisher to check the quality requirements. All specified dimensions can be found in the table. The taxonomy publisher is not perfectly acquainted with the business requirements derived from the new legislation however he checks the table for comprehensibility and the technical constraints needed to be met in order to deduct the taxonomy from the DPM. The business requirements need to be reviewed by supervisory experts.

In the table shown in Figure 23 redundancies can be recognised. Looking at the annotations on the right hand side of the table, we detect the redundancy of MKR EQU in two dimensions: MC and RT, which stands for main category and risk type. The risk type differs in some cases between MKR EQU risk, MKR EQU general risk, MKR EQU specific risk and MKR not look-through CIUs risk. The information that the members of the domain “risk type” refer to, the approach “market risk for equities” is repeated on each member. If those members are combined in one Data Point with the member “market risk” of the domain “approach”, then the information is redundant in both domains. It needs to be ensured where the approach dimension is stored, i.e. together with the risk type to reduce the number of domains or in separate domains, one for risk type and one for approach.

If a taxonomy publisher detects such inconsistency, he should get in contact with you to put his concerns forward and ask for a justification of the different domains and respective dimensions used for this table. After the data model is finalised it should be checked that it fully reflects all requirements for the generation of the corresponding taxonomy.

## 6.6 Distribute Data Point Model

Finally, the DPM can be forwarded to the appropriate department for creating a taxonomy and initiating the following process steps. The creation of the taxonomy will be followed by a quality assurance process before its storage and publication. If the quality assurance for the taxonomy fails due to an erroneous DPM, the process of DPM modelling will be iterated until the taxonomy is approved for publication.

## 7 What the future holds for us

In order to help you in your task to create and review Data Point Models, software was programmed. As the market realises the possibility of new levels of sales, new applications for the creation of XBRL taxonomies will be introduced soon. One program that is considered as user-friendly for the purpose of creating a DPM is DPM Architect for XBRL, developed by the Banco de España and introduced firstly at the XBRL Week in May 2012.<sup>37</sup> The software can not only help you to build up and review the DPM, it is also intended to generate XBRL taxonomies, which is the next step in the process. The MKR SA EQU template is used again as an example to show some excerpts of the implementation of the process of creating a DPM with DPM Architect.

The amount type dimension was selected to serve as dimensioned element. Applicable characteristics of the amount type of the MKR SA EQU framework are shown below.

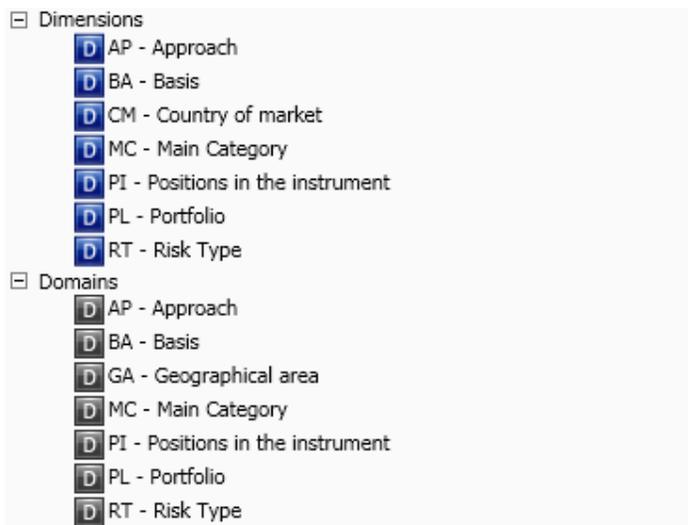
		Data type	Period type	
1 - Gross value	1	Monetary	Stock	Gross value
2 - Net value	2	Monetary	Stock	Net value
3 - Subject to capital value	3	Monetary	Stock	Subject to capital value
4 - Own funds requirements	4	Monetary	Stock	Own funds requirements
5 - Total risk exposure amount	5	Monetary	Stock	Total risk exposure amount

**Figure 24 — The attribute for amount type and period type of the dimensioned element of MKR SA EQU**

For each member of the dimensioned element, also known as metric, an amount type as well as a period type have to be defined. The period types are stock and their data types are monetary. Meaningful names were chosen for the metrics (net value, subject to capital value, own funds requirements and total risk exposure amount).

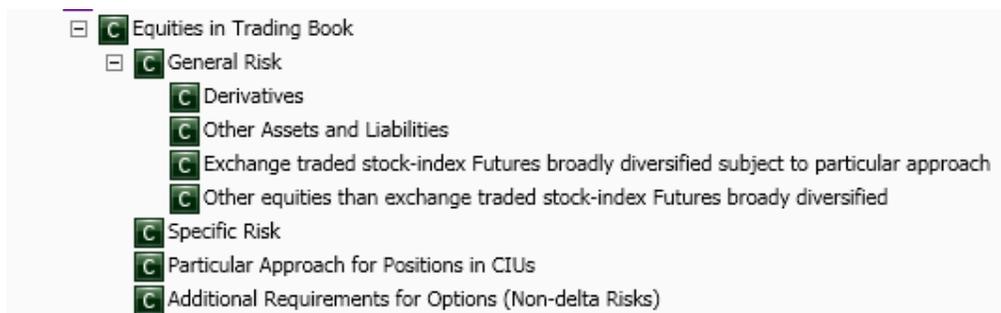
Moreover the list of dimensions and domains specified can be retrieved.

<sup>37</sup> Banco de España (2012)



**Figure 25 — View of dimensions and domains specified for MKR SA EQU**

Figure 25 shows a helpful view to check the completeness of the DPM. Also, the informative value of the naming of the dimensions and domains can be examined. An example of a presentation hierarchy is given by the next screencapture (Figure 26).



**Figure 26 — Summary of hierarchies specified for MKR SA EQU**

The domain member hierarchies can be modified in a more detailed view which is pictured below. The tool also provides the possibility to define aggregations for calculation purposes.

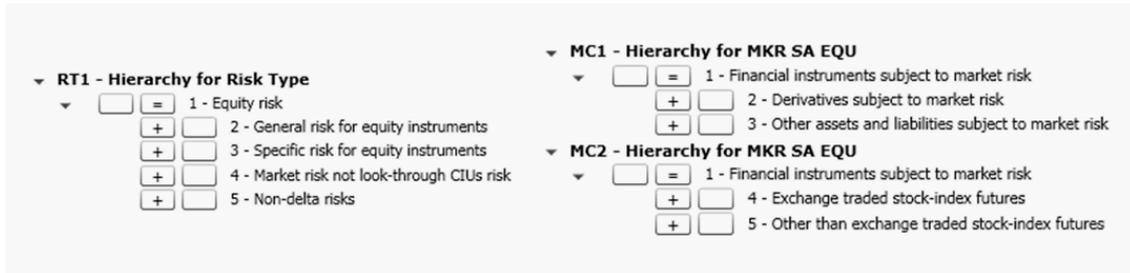


Figure 27 — Hierarchies for selected domains of MKR SA EQU

Finally, a table can be visualised. Figure 28 shows the row column codes of each cell which correspond to a Data Point defined with the help of the DPM Architect. Not existing combinations are greyed out so they cannot be reported. If the table shown by the tool corresponds to the template originally defined by you, you have done a great job creating a faultless DPM.

		All Positions				Own Funds Req	Total Risk Expos
		Gross Positions		Net Positions		Positions Subject to Capital Charge	
		Long (010)	Short (020)	Long (030)	Short (040)		
Equities in Trading Book (010)	General Risk (020)	r20c10	r20c20	r20c30	r20c40	r20c5	r20c6
	Derivatives (021)	r21c10	r21c20				
	Other Assets and	r22c10	r22c20				
	Exchange traded	r30c10	r30c20	r30c30	r30c40		
	Other equities t	r40c10	r40c20	r40c30	r40c40		
	Specific Risk (050)	r50c10	r50c20	r50c30	r50c40	r50c5	r50c6
	Particular Approach for Positions	r60c10	r60c20	r60c30	r60c40	r60c5	r60c6
	Additional Requirements for Opt					r70c6	

Figure 28 — Table generated by DPM Architect to summarise the information given during the creation process of the DPM

The tool is already available for testers and is used to produce taxonomies in production in Banco de España. DPM Architect will be published on Banco de España’s website this year. Currently Banco de España is providing the tool only upon request.<sup>38</sup>

Personally I expect more software support in the future. In order to transfer knowledge about the DPM to IT specialists the CEN workshop published a Data Point Meta Model in order to show the “components for the construction of a formal model that describes sets of Data Points relevant to European supervisory frameworks”<sup>39</sup> that can be viewed via their wiki.

I hope this paper helped you to gain an overview on the preparation of Data Point Models. The methodology is very new and needs training and experience. I would be delighted to receive positive feedback that the thesis provides an efficient guidance for the task of creating and understanding DPMs. The reasonable training time for IT specialists as well as you, banking supervisors, needed to gain knowledge about the new methodology is worthwhile, as the DPM eases the communication about the required data by you. This offers additional opportunities for better analysis by slicing and dicing through the pool of supervisory data.

<sup>38</sup> Morilla, V. (2013)

<sup>39</sup> Declerck, T./ Hommes, R./ Heinze, K. (2013)

## Bibliography

### List of literature

- [1] Baeumle-Courth P., Nieland S., Schröder H. (2004): Wirtschaftsinformatik, München, Oldenbourg Verlag
- [2] Ballard, C./ et al (1998), Data Modeling Techniques for Data Warehousing, Springville: Vervante
- [3] Cordts, S./ Blakowski, G./ Brosius, G (2011): Datenbanken für Wirtschaftsinformatiker, Wiesbaden: Vieweg+Teubner Verlag
- [4] Ferstl, O., Sinz, E. (2013), Grundlagen der Wirtschaftsinformatik, 7th Edn., München: Oldenbourg Wissenschaftsverlag
- [5] Groth, R./ et al (1983), Projektmanagement in Mittelbetrieben, Planung und Durchführung einmaliger großer Vorhaben, Köln: Deutscher Instituts-Verlag
- [6] Heinze, K. (2012), Modernisierung der Datenformate, in: Handbuch Bankenaufsichtliches Meldewesen, Heidelberg: FinanzColloquium Heidelberg, p. 68-92
- [7] Kummer, W./ Spühler, R./ Wyssen R. (1986), Projekt-Management, Leitfaden zu Methode und Teamführung in der Praxis, 2nd Edn., Zürich: Verlag Industrielle Organisation
- [8] Minhorst, A. (2005), Das Access 2003 Entwicklerbuch, München: Addison-Wesley Verlag
- [9] Piechocki, M. (2012), Supervising Models: XBRL and Data Point modelling, in: iBR interactive business reporting, Vol. 02, No. 3, p. 26-29
- [10] Platz, J./ Schmelzer, H. (1986), Projektmanagement in der industriellen Forschung und Entwicklung, Einführung anhand von Beispielen aus der Informationstechnik, Berlin/Heidelberg/New York: Springer Verlag
- [11] w.a. (1982), Systems Engineering. Ein Leitfaden zur methodischen Durchführung umfangreicher Planungsvorhaben, Daenzer (Ed.), 3rd Edn., Zürich: Verlag Industrielle Organisation

### List of Internet and intranet sources

- [12] 1keydata (2013a), Conceptual Data Model, <http://www.1keydata.com/datawarehousing/conceptual-data-model.html>, retrieval, 22.02.2013
- [13] 1keydata (2013b), Logical Data Model, <http://www.1keydata.com/datawarehousing/logical-data-model.html>, retrieval, 22.02.2013
- [14] 1keydata (2013c), Physical Data Model, <http://www.1keydata.com/datawarehousing/physical-data-model.html>, retrieval, 22.02.2013
- [15] Banco de España (2012), DPM Architect for XBRL, Update & Demo, [www.eurofiling.info/201212/presentations/DPM\\_Morilla\\_20121213.ppsx](http://www.eurofiling.info/201212/presentations/DPM_Morilla_20121213.ppsx), retrieval, 10.04.2013
- [16] Bundesbank (2013), Marktrisikomeldung Aktiennetttoposition, [http://www.bundesbank.de/Redaktion/DE/Downloads/Service/Meldewesen/Bankenaufsicht/PDF/mkrak\\_al t.pdf?\\_\\_blob=publicationFile](http://www.bundesbank.de/Redaktion/DE/Downloads/Service/Meldewesen/Bankenaufsicht/PDF/mkrak_al t.pdf?__blob=publicationFile), retrieval, 25.02.2013

- [17] CEN (2009), CEN Workshops,  
<http://www.cen.eu/CEN/Sectors/TechnicalCommitteesWorkshops/Workshops/Pages/default.aspx>,  
 retrieval, 29.04.2013
- [18] Collins, J. (2013), Comparison of Relational and Multi-Dimensional Database Structures,  
<http://www.alphadevx.com/a/36-Comparison-of-Relational-and-Multi-Dimensional-Database-Structures>,  
 retrieval, 28.02.2013
- [19] CSB Advocates (2013) , An update on the forthcoming CRD IV Rules - Malta,  
<http://www.hg.org/article.asp?id=30273>, retrieval, 24.4.2013
- [20] databasedev (2013), Database Design & Normalization,  
[http://www.databasedev.co.uk/database\\_normalization\\_process.html](http://www.databasedev.co.uk/database_normalization_process.html), retrieval, 13.03.2013
- [21] Declerck, T./ Hommes, R./ Heinze, K. (2013), CEN Workshop Agreement,  
[http://wikixbrl.info/index.php?title=European\\_Data\\_Point\\_Methodology](http://wikixbrl.info/index.php?title=European_Data_Point_Methodology), retrieval, 08.04.2013
- [22] EBA (2011a), EBA Consultation Paper on Draft Implementing Technical Standards on Supervisory reporting requirements for institutions,  
<http://www.eba.europa.eu/cebs/media/Publications/Consultation%20Papers/2011/CP50/CP50-ITS-on-reporting.pdf>, retrieval, 22.2.2013
- [23] EBA (2011b), Eurofiling, Data modelling and ExcelXBRLGen, <http://www.openfiling.info/wp-content/uploads/data/EurofilingWebinar20110713.pdf>, retrieval, 28.2.2013
- [24] EBA (w.y.a), About us, <http://eba.europa.eu/Aboutus.aspx>, retrieval, 11.03.2013
- [25] EBA (w.y.b), Supervisory Reporting Introduction, <http://www.eba.europa.eu/Supervisory-Reporting/Introduction.aspx>, retrieval, 02.03.2013
- [26] ECB (2013), Number of monetary financial institutions (MFIs), February 2013,  
[http://www.ecb.int/stats/money/mfi/general/html/mfis\\_list\\_2013-02.en.html](http://www.ecb.int/stats/money/mfi/general/html/mfis_list_2013-02.en.html), retrieval, 02.03.2013
- [27] Gartner (2012), Semantic Data Model, <http://www.gartner.com/it-glossary/semantic-data-model/>, retrieval, 08.03.2013
- [28] IBM (w.y.), Hierarchien über- und untergeordneter Elemente,  
[http://pic.dhe.ibm.com/infocenter/rdahelp/v7r5/index.jsp?topic=%2Fcom.ibm.datatools.dimensionsal.ui.doc%2Ftopics%2Fc\\_dm\\_pc\\_hierarchies.html](http://pic.dhe.ibm.com/infocenter/rdahelp/v7r5/index.jsp?topic=%2Fcom.ibm.datatools.dimensionsal.ui.doc%2Ftopics%2Fc_dm_pc_hierarchies.html), retrieval, 10.03.2013
- [29] Janssen, C. (w.y.), Pivot Table Techopedia explains Pivot Table,  
<http://www.techopedia.com/definition/14649/pivot-table>, retrieval, 05.03.2012
- [30] Oxford University Press (2013), Oxford Dictionaries,  
<http://oxforddictionaries.com/definition/english/model?q=model>, retrieval, 28.02.2013
- [31] Rouse M. (2010), Definition data modelling, <http://searchdatamanagement.techtarget.com/definition/data-modeling>, retrieval, 14.02.2013
- [32] Sapia, C./ et al (1999), Extending the E/R Model for the Multidimensional Paradigm,  
[http://link.springer.com/chapter/10.1007%2F978-3-540-49121-7\\_9?LI=true](http://link.springer.com/chapter/10.1007%2F978-3-540-49121-7_9?LI=true), retrieval, 10.03.2013
- [33] TechTerms (2013a), Data, <http://www.techterms.com/definition/data>, retrieval, 09.03.2012
- [34] TechTerms (2013b), Metadata, <http://www.techterms.com/definition/metadata>, retrieval, 09.03.2012
- [35] Verma, R. (2009a), Slicing,  
[http://www.hypertextbookshop.com/dataminingbook/public\\_version/contents/chapters/chapter003/section004/blue/page004.html](http://www.hypertextbookshop.com/dataminingbook/public_version/contents/chapters/chapter003/section004/blue/page004.html), retrieval, 10.03.2013

- [36] Verma, R. (2009b), Dicing,  
[http://www.hypertextbookshop.com/dataminingbook/public\\_version/contents/chapters/chapter003/section004/blue/page004.html](http://www.hypertextbookshop.com/dataminingbook/public_version/contents/chapters/chapter003/section004/blue/page004.html), retrieval, 10.03.2013
- [37] Wayne State University (2013), S.M.A.R.T. Objectives, <http://wayne.edu/hr/leads/phase1/smart-objectives.php>, retrieval, 15.03.2013
- [38] XBRL Spain (2012), Main Page, Note to the readers, [http://wikixbrl.info/index.php?title=Main\\_Page](http://wikixbrl.info/index.php?title=Main_Page),  
retrieval, 10.04.2013
- [39] ZaZa Network (2007), Database Development Overview,  
[http://www.zazanetwork.com/resources\\_services/articles/databases/database\\_development.aspx#top](http://www.zazanetwork.com/resources_services/articles/databases/database_development.aspx#top),  
retrieval, 28.02.2013