Mapping between the formulas XBRL and the Multidimensional Data Model.

Academic Track @ XBRL Week,

Frankfurt (Germany). June, 1st – 2nd 2016

European Central Bank.

Ignacio Santos, PhD

Carlos III University of Madrid, Computer Science Department

Summary

1. Summary.
2. Introduction.
3. XBRL Data Model and Calculation.
4. XBRL Formula Specification and the Multidimensional Data Model.
5. Semantic Questions.
6. Proof of Concept.
7. Validations.
8. Conclusions.
9. Future works
10. Questions.

Introduction I

This presentation is based in:

- XBRL Formula Specification in the Multidimensional Data Model of I. Santos et al. Journal: Information Systems (JCR, Q2), Volume 57, April 2016, pages 20-37.
- My Doctoral thesis: "Framework for multidimensional definition of the data model structure, taxonomies and rules of the XBRL specification". Carlos III University of Madrid, February 10th, 2016.

Information Systems 57 (2016) 20–37

Contents lists available at ScienceDirect

Information Systems

journal homepage: www.elsevier.com/locate/infosys

XBRL formula specification in the multidimensional data model

Ignacio Santos [a,*], Elena Castro [b], Manuel Velasco [b]

[a] Bank of Spain, Department of Information Systems, c/ Alcalá 522, 28027 Madrid, Spain
[b] Carlos III University of Madrid, Computer Science Department, Avda. de la Universidad No. 30, Madrid, 28911 Leganés, Spain

ARTICLE INFO

Article history:
Received 22 August 2012
Received in revised form
7 August 2015
Accepted 2 November 2015
Available online 17 November 2015

Keywords:
Extensible Business Reporting Language

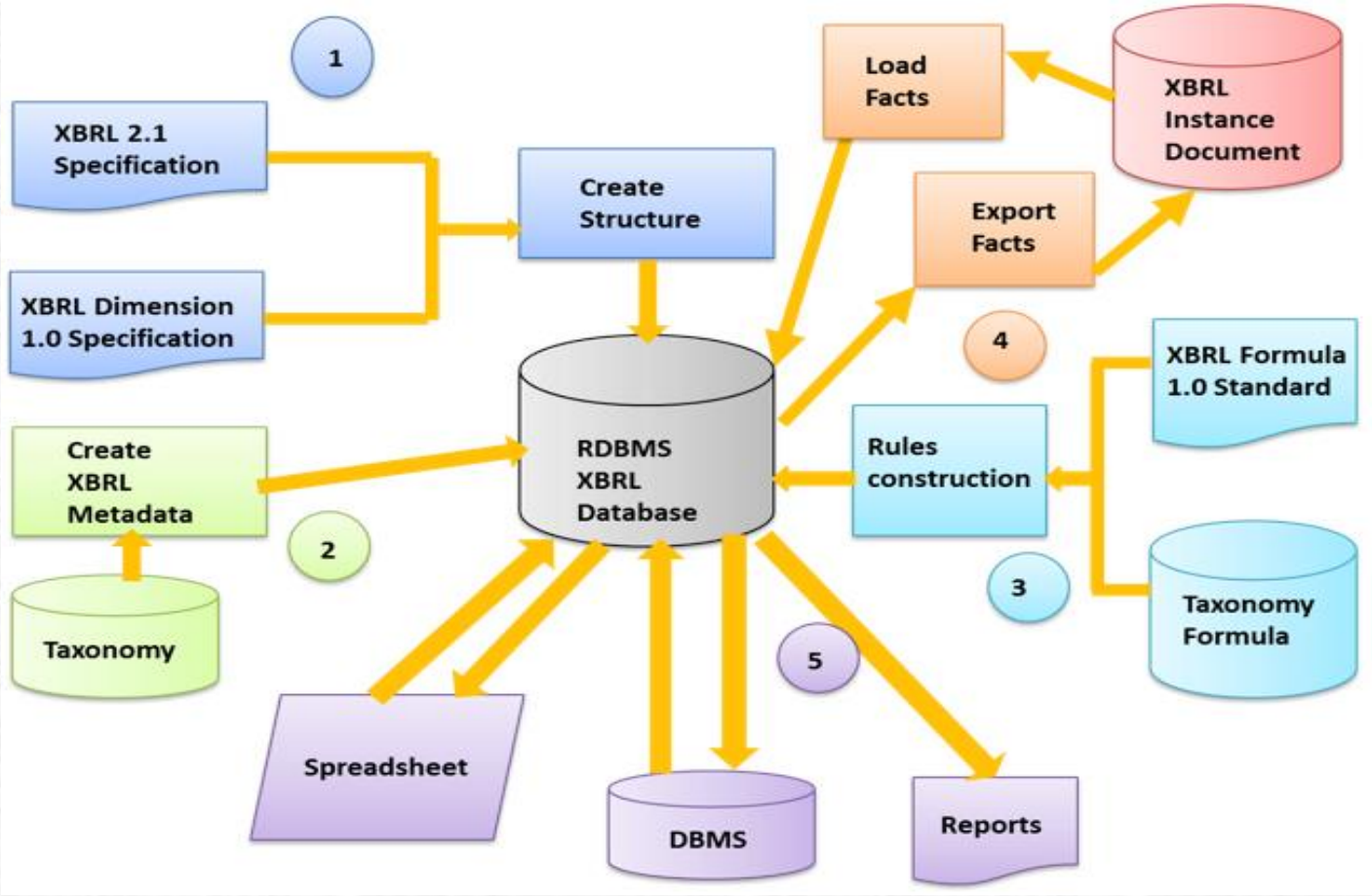
ABSTRACT

XBRL is a specification used to exchange financial/economic information. It is actively used by many international institutions and agencies. In the USA, Canada, Europe, China, etc. all financial entities and companies quoted on the stock market have to report compulsorily to the supervisory and regulatory authority using the XBRL specification. XBRL consists of a set of taxonomies defining different accounting regulations for a specific statement and the statement itself. Reports are generated from various sources and are validated at origin. XBRL displays business information which is multidimensional and whose logical destination for storage is a data warehouse. The proposal presented here focuses on the

Introduction II

- This presentation focuses on research into the mapping between the XBRL data model (XBRLDM) and the multidimensional data model (MDM), as well as its automation.
- The MDM is a straightforward model that combines objects, dimensions (hierarchies), measures and attributes for representing real work business problems. In addition, this model is the heart of *On-Line Analytical Processing* (OLAP), which requires complex queries that can be solved by the MDM in real time.
- Conversion between the aforementioned models will be made using the *Model Driven Architecture* (MDA) paradigm, which ensures interoperability and solves the problem of heterogeneity between systems.
- UML/MDA (*OMG, 2015*) is a powerful tool that has helped in different areas of *Information Technology* (IT) to model structured and robust systems. However, the techniques of verification and validation of the software is not supported in the standard MDA.

Introduction III



XBRL Data Model and Calculations I

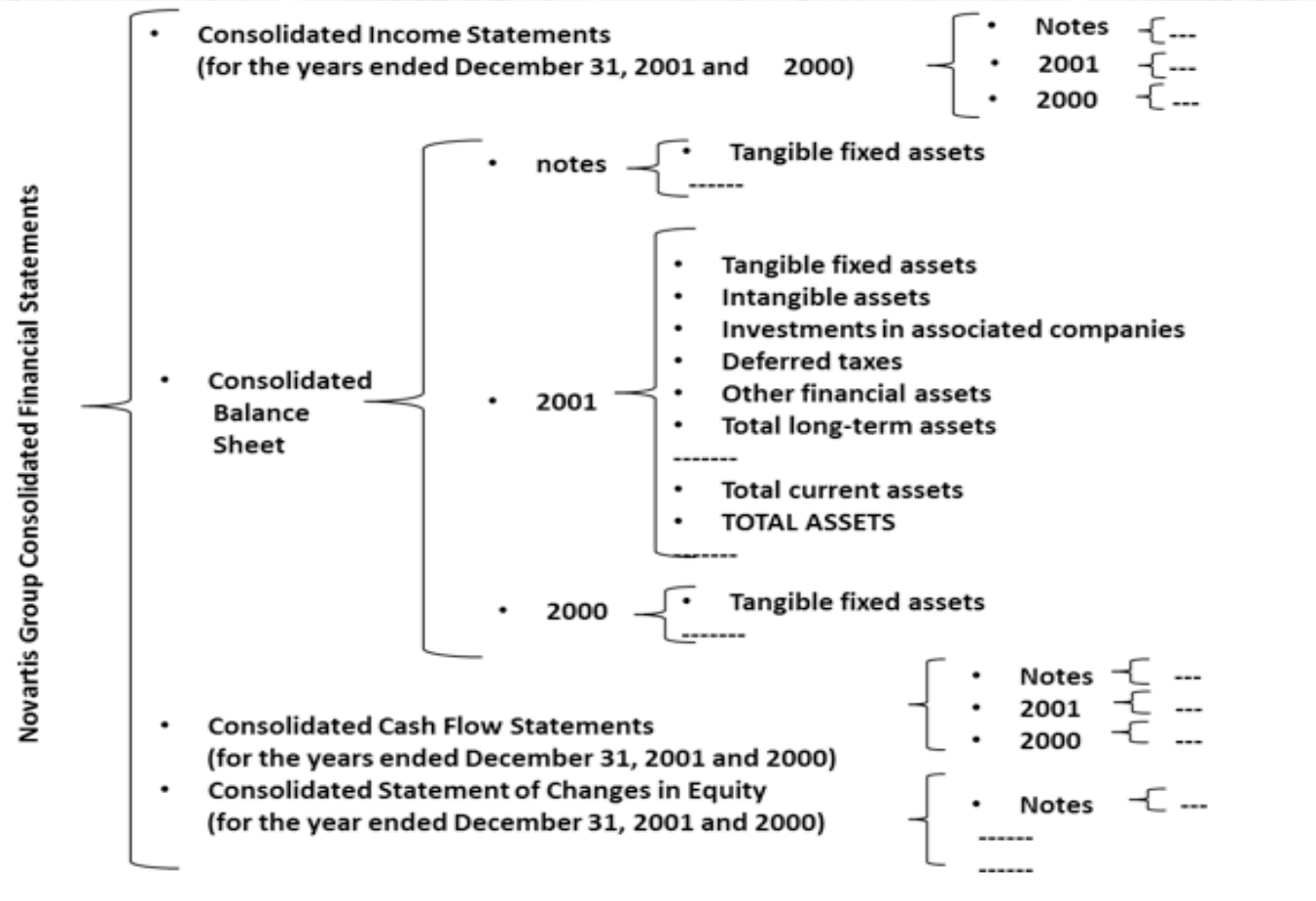
- XBRL semantic information, separated from application software, uses and extends the XML standard.
- A report or XBRL instance document references a set of XML or XBRL Schemas. This set of schemas in the XBRLDM, called a *Discoverable Taxonomy Set* (DTS) specifies the concepts, rules and constraints.
- The role, *Calculation* provides simple calculation relationships between different elements (basic concepts, dimensions and dimension attributes), but does not allow formulas or complex expressions.
- The previous version, XBRL 1.0, was based on hierarchies with a tree structure, since XML, the language on which it was based, is hierarchical.
- An example of a XBRL instance document generated under these limitations is presented in *next slide* (*XBRL International and Novartis International AG,* consisting of a set of basic concepts with a set of *dimension-dimension attribute* pairs.

XBRL Data Model and Calculations II

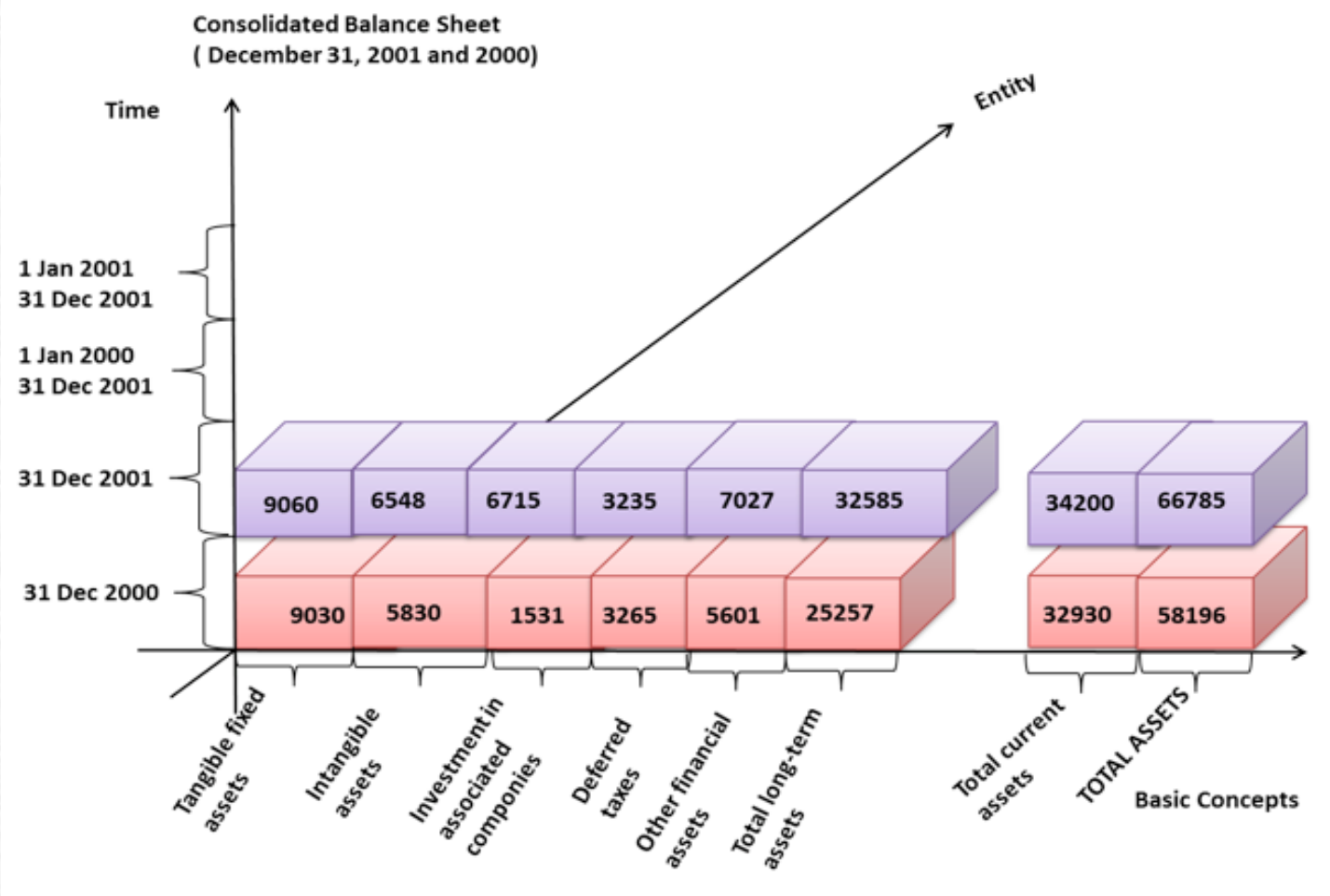
Consolidated Balance Sheet
(at December 31, 2001 and 2000)

Assets	Notes	2001 CHF millions	2000 CHF millions
Long-term assets			
Tangible fixed assets	8	9060	9030
Intangible assets	9	6548	5830
Investment in associated companies	11	6715	1531
---	---	---	---
Deferred taxes	12	3235	3265
Other financial assets	13	7027	5601
Total long-term assets		*32585*	*25257*
Current assets			
---		---	---
---		---	---
Total current assets		*34200*	*32939*
TOTAL *ASSETS*		*66785*	*58196*

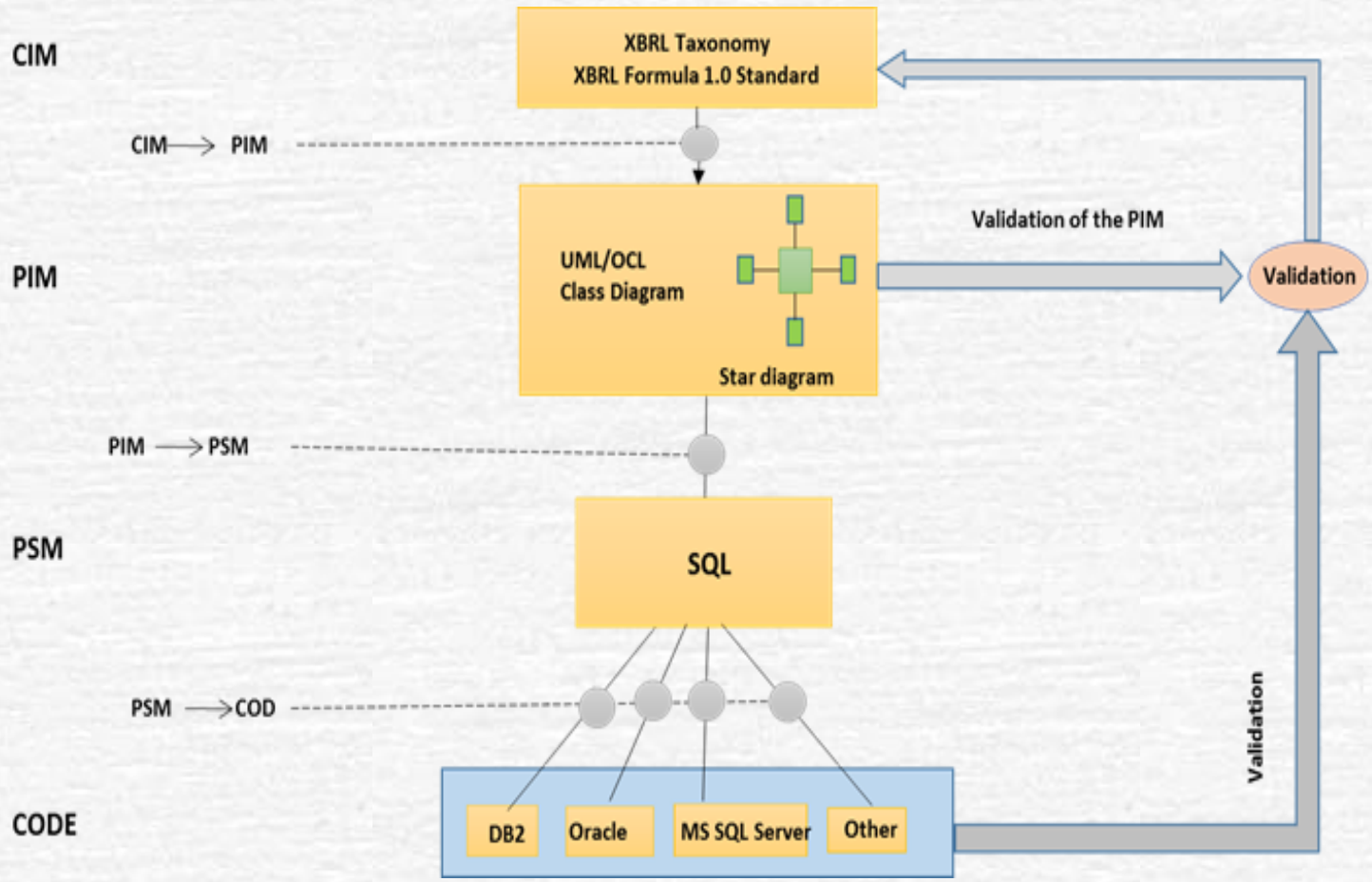
XBRL Data Model and Calculations III



XBRL Data Model and Calculation IV



XBRL formula Specification & the MDM I



XBRL formula Specification & the MDM II

- The CIM model in this research work is the set of XBRL instance documents, the taxonomies and XBRL Formula Specification.
- The rules and definitions from the CIM are obtained, through XBRL Formula Specification.
- The PIM used in this proposal is based on UML, which is a star model, the MDM.
- The set of constrains, dimensions and dimension attributes are collected in an automatic way from a taxonomy and its algorithm is shown.
- The PSM is a set of stored procedures or programmes in Cobol, C++, etc. The algorithm will show the mapping from the PIM to the PSM.
- The process of validation is divided in **two phases**.
  - The first phase is to test the UML star model / MDM (the PIM), from the XBRL taxonomies and the XBRL reports.
  - The second one is to validate the set of stored procedures using ROLAP technology.

XBRL Formula Specification & the MDM III

Def	XBRLDM, the CIM	MDM or star model (the PIM)
1	Fact	Fact
2	Assertion	Constraint
3	Filter	Set of pairs <dimension/attribute of dimension>
4	Implicit filter	Set of pairs <dimension/attribute of dimension>
5	Explicit filter	Set of pairs <dimension/attribute of dimension>
6	Basic concept	Pair <Dimension/attribute of dimension>
7	Constant	Constraint
8	Variable	Set of pairs <dimension/attribute of dimension>
9	FallbackValue	Constraint
10	Precondition	Constraint
11	Filter group	Set of pairs <dimension/attribute of dimension>

From the CIM to the PIM

XBRL Formula Specification & the MDM IV

Example using the Bank of Spain's consolidated public sector balance sheet 6610.

assertionSet

assertionSet

valueAssertion *val_6610-sh-2*

(($creditInstitutions + $InsuranceEntities + $otherEntities +

$settlementAdjustments) eq $TotalGroupEntities) and ($TotalGroupEntities eq 0))

conceptName *ifrs-gp:InvestmentsInSubsidiariesAtCost*

factVariable $creditInstitutions fallbackValue =0

explicitDimension CreditInstitutionConsolidatedGroup

sp-bs-d-FR-dist:DistributionDimension

factVariable $InsuranceEntities fallbackValue =0

explicitDimension InsuranceEntities

sp-bs-d-FR-dist:DistributionDimension

factVariable $otherEntities fallbackValue =0

explicitDimension otherEntities

sp-bs-d-FR-dist:DistributionDimension

factVariable $settlementAdjustments fallbackValue =0

explicitDimension settlementAdjustments

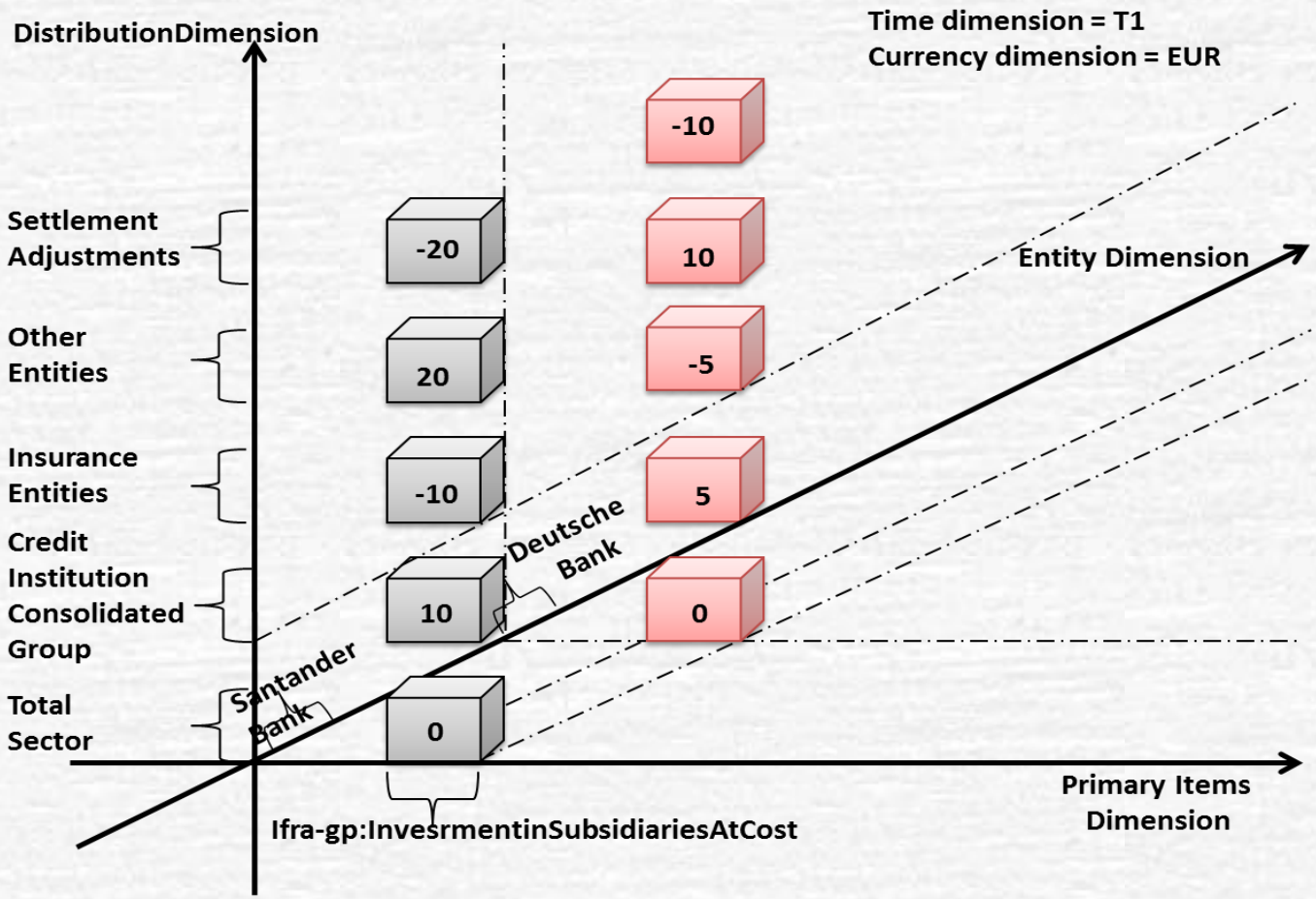
sp-bs-d-FR-dist:DistributionDimension

factVariable $TotalGroupEntities

explicitDimension TotalSectorial

sp-bs-d-FR-dist:DistributionDimension

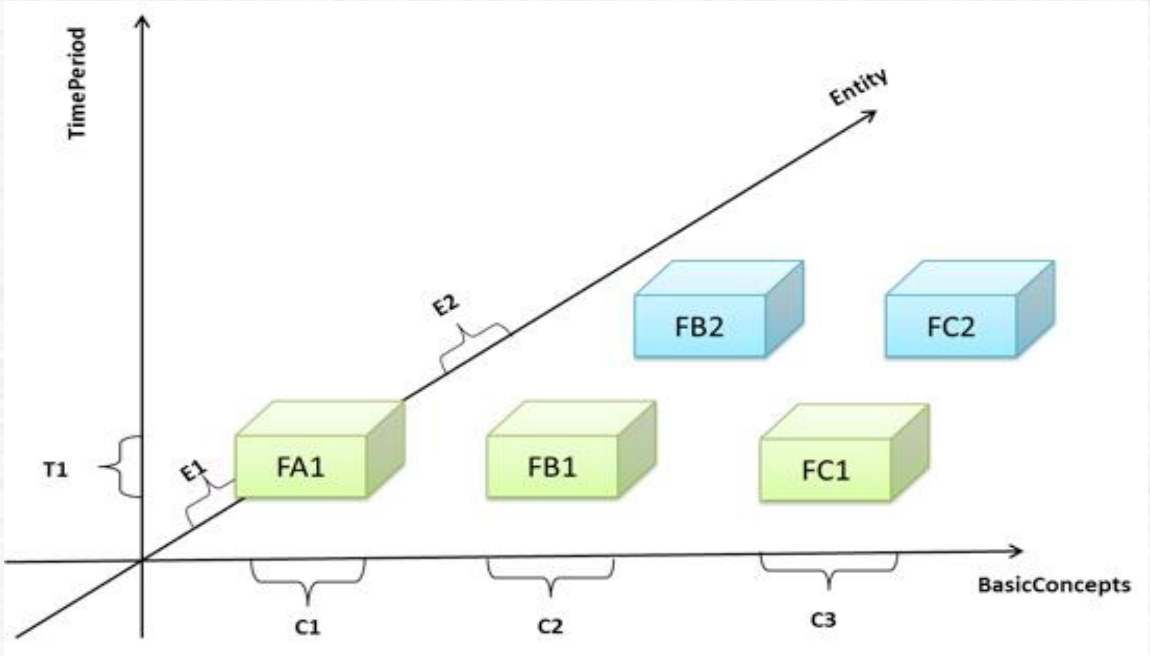
XBRL Formula Specification & the MDM V



Example of assertion presented graphically

Semantic Questions I

- One particular problem arises when the instance document uses implicit filters and a fact is null.



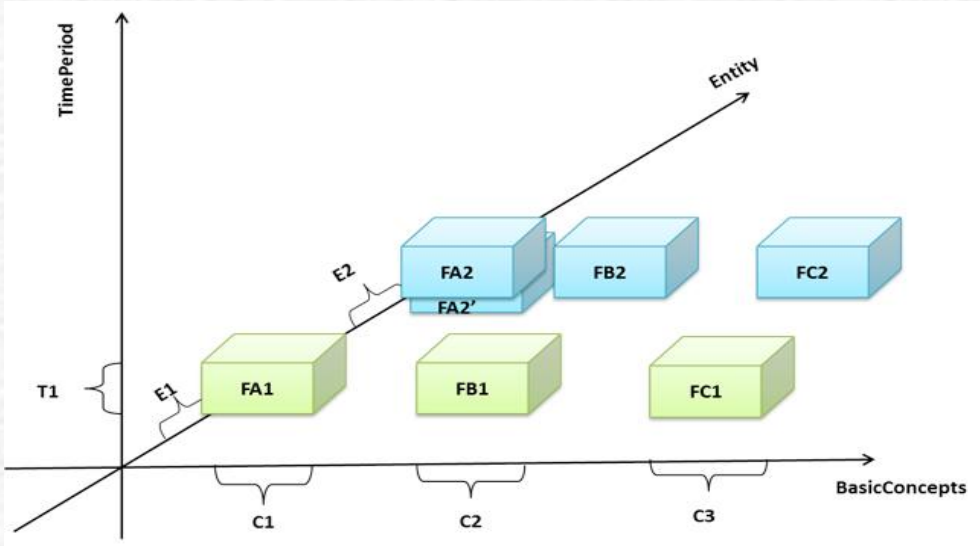
The XBRL processor or DBMS must validate:

1. $FA1 = FB1 + FC1$
2. $FA2 = FB2 + FC2$

Since *FA2* is null according to the assertion *FA2=FB2+FC2* is not validated.

Semantic Questions II

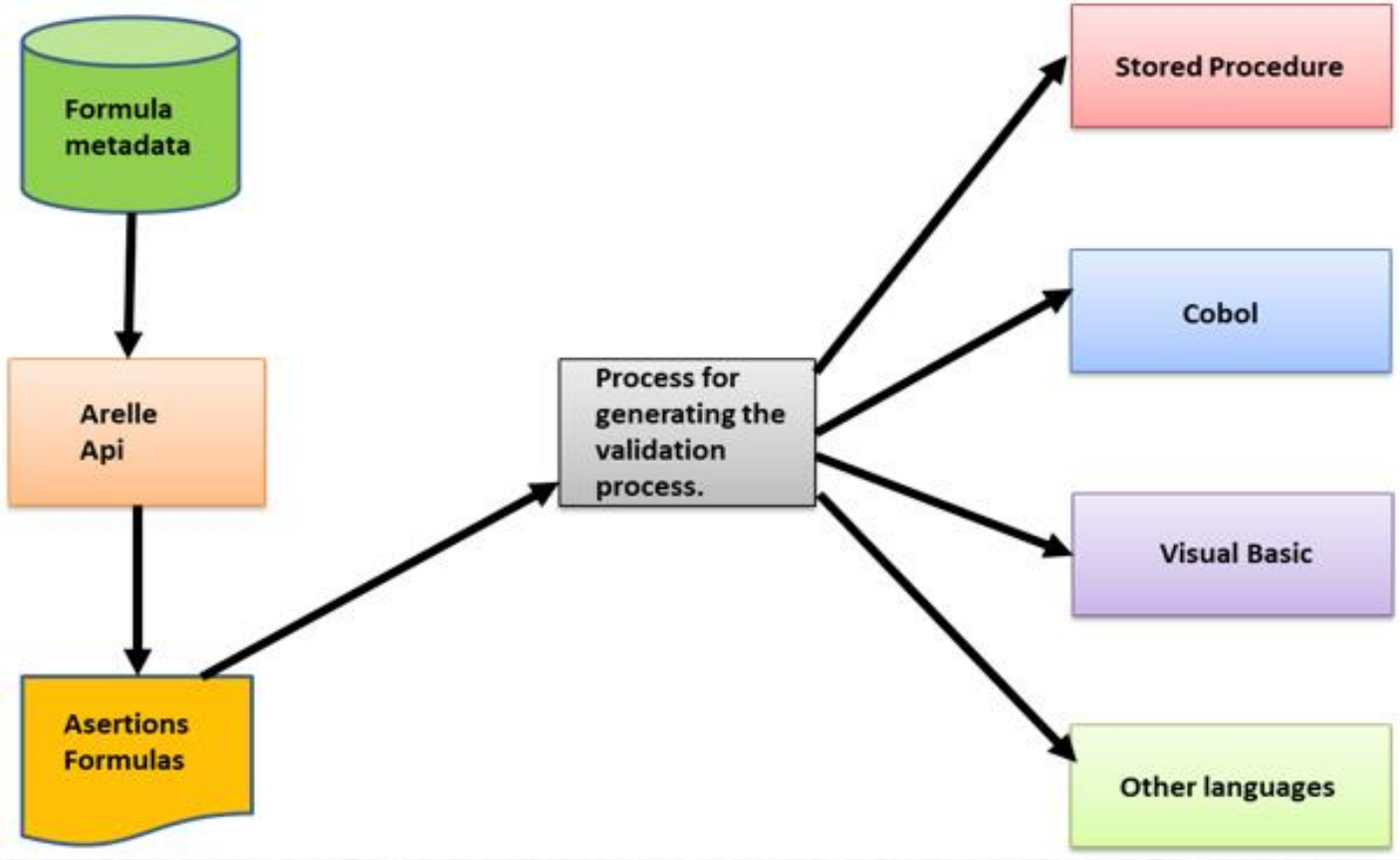
- Another potentially problematic case is when the instance document uses implicit filters and a variable has two facts, since it contains an additional dimension (i.e., explicit filter) compared to the other variables.



If $C1 \cup E2 \cup T1$ *is* $FA2$ *and* $FA2'$, then it is necessary to decide amongst:

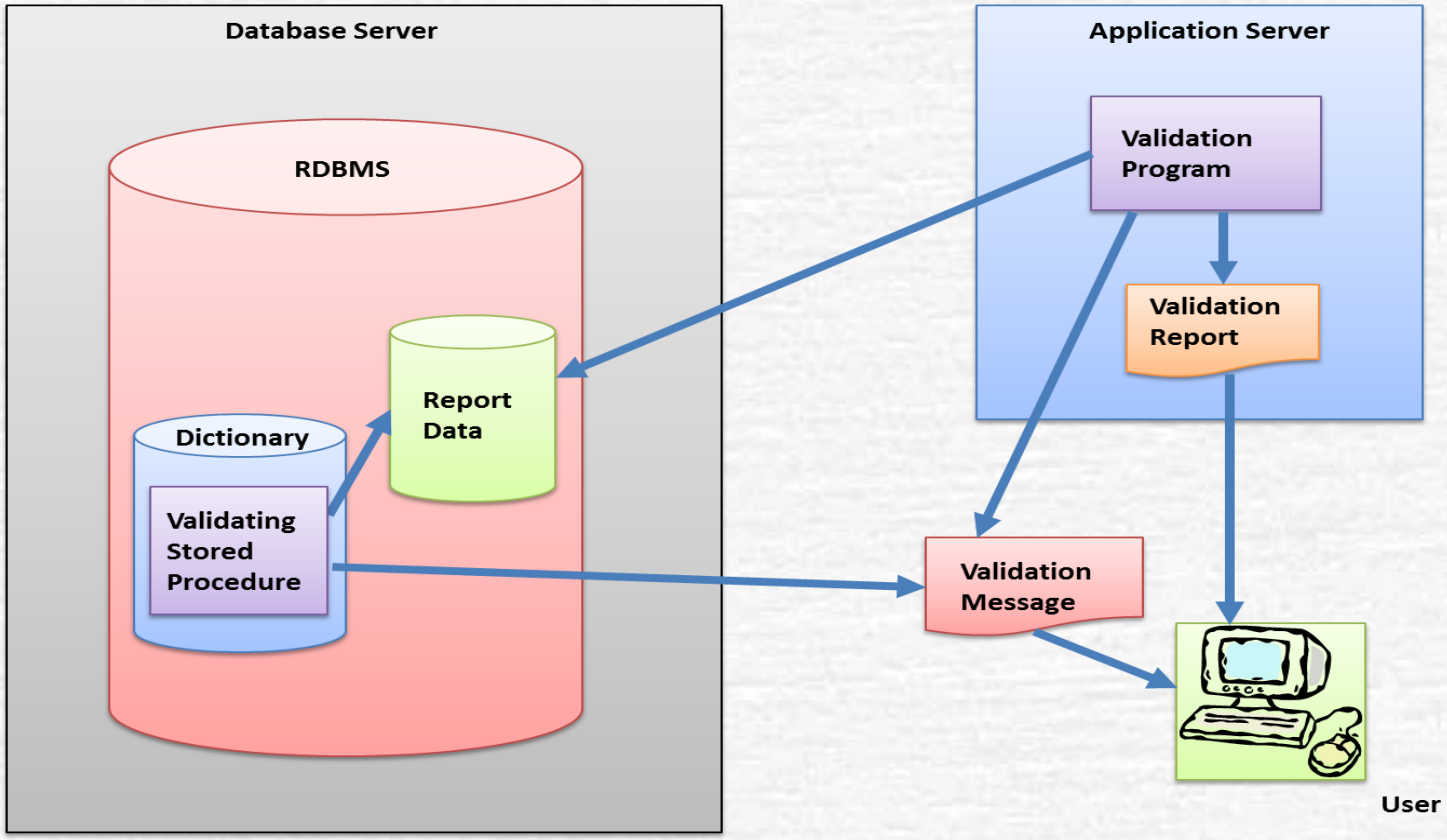
1. $FA2 = FB2 + FC2$ *y* $FA2' = FB2 + FC2$
2. $FA2 = FB2 + FC2$ *o* $FA2' = FB2 + FC2$
3. None

Proof of Concept I



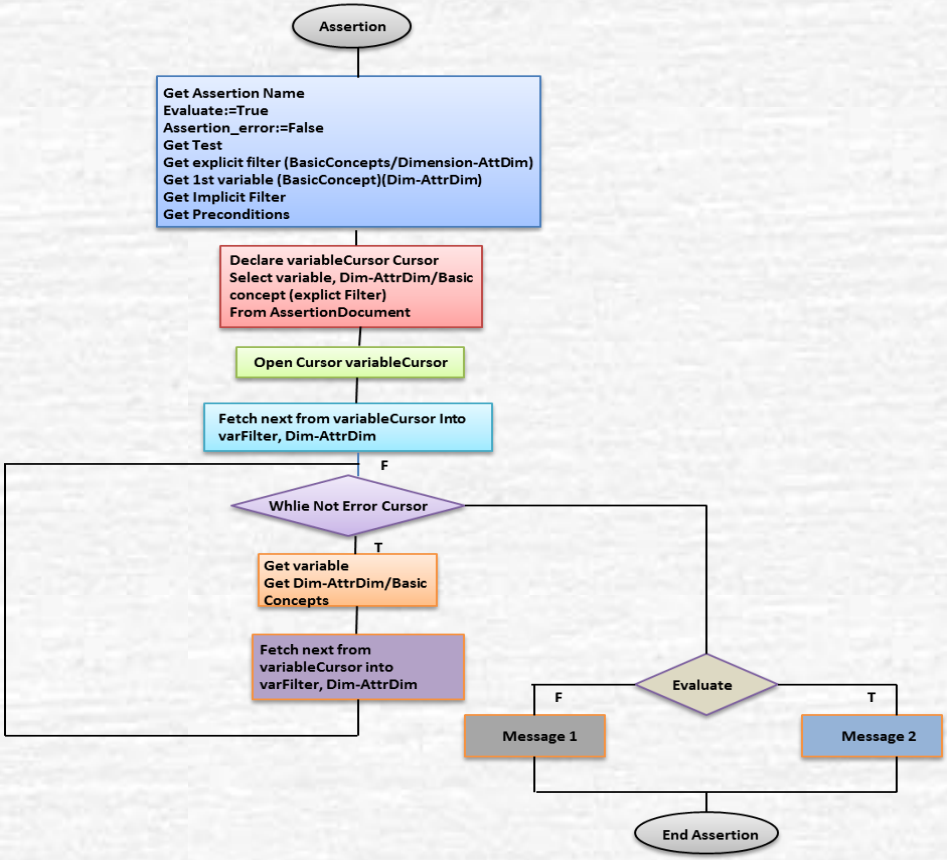
Obtaining the validation API

Proof of Concept II



Validation process

Proof of Concept III



Assertion process of validation-process-generating algorithm

Proof of Concept IV


```
Start
    begin
      if there are preconditions
        then if there are variables
          then if explicit filter has basic concepts  and not dimensions
            then template SQL V
            else if explicit filter does not have basic concepts and has dimensions
              then template SQL VI
              else template SQL VII
          else template SQL IV
        else if explicit filter has basic concepts and not dimensions
          then template SQL I
          else if explicit filter does not have basic concepts
            then template SQL II
            else template SQL III
    end
End
```

Creation of SQL templates

Validation

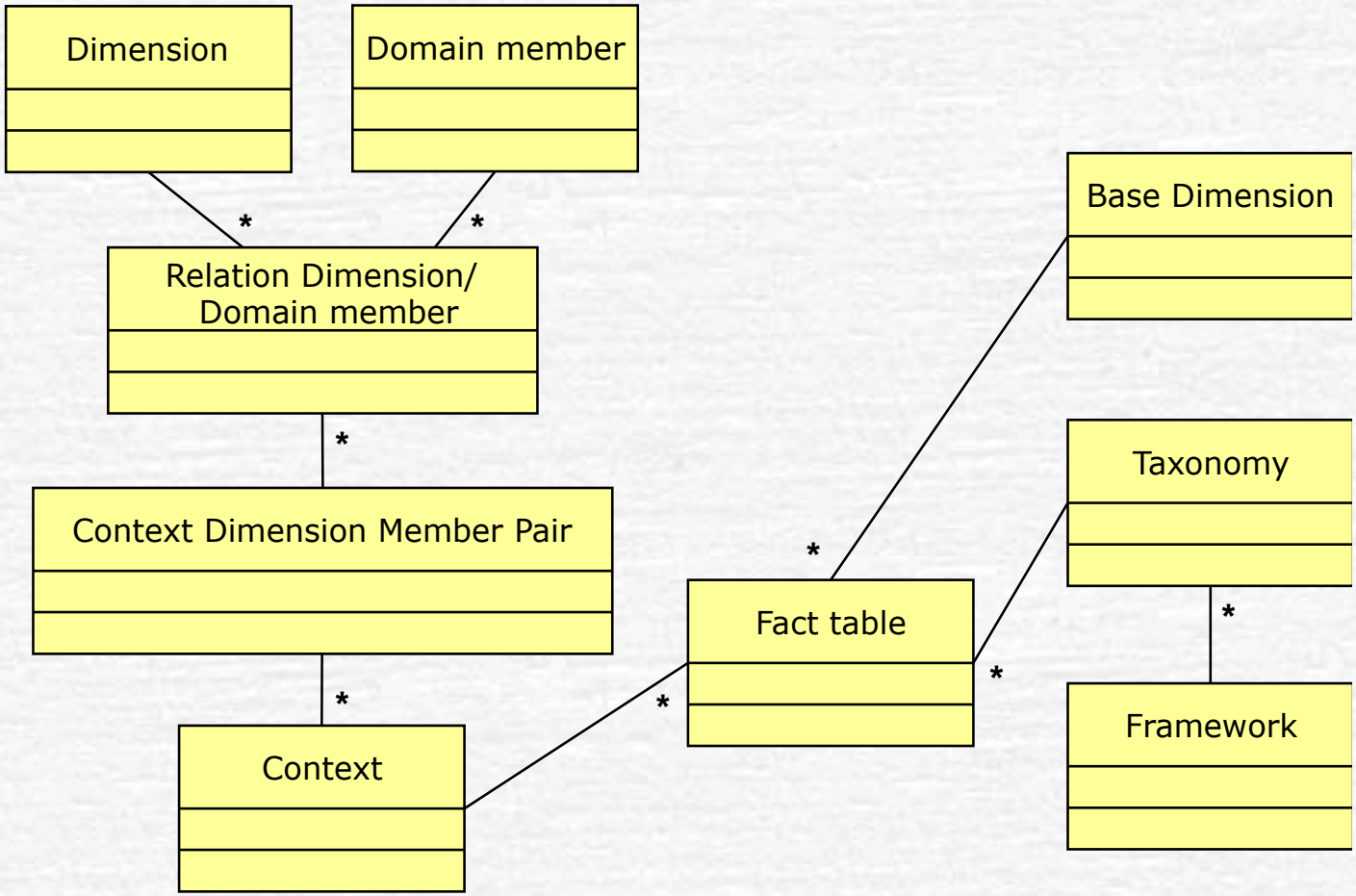


Diagram with the context in the star model.

Conclusions

- The development of the entire metadata model's life cycle using a robust architecture technology as the MDA, non-existent to date.

- The establishment of a data model design life cycle ensures fewer errors in the design, since it has been proven in concept testing and validation, and gives the possibility of making a set of test cases for analysing anomalies and other semantic questions.

- This paper has validated the interoperability of this technology by studying its design. Moreover, it shows how this model can be implemented in different databases of different vendors and even enables mapping to other platforms.

Future works

- Comparing the complexity of both types of algorithms, based on a star model with or without context, as well as the performance of both.
- Checking the performance of formula validation through stored procedures and comparing it with the performance using XML validation.
- Expand the use of this specification to other environments that are not Supervision and Regulation.
- Incorporate into this research the *Data Point Model* (DPM), and its implementation in XBRL.
- Through this development life cycle, facilitate the creation of public test games.
- Not only generate templates for SQL stored procedures, but also in other languages such as COBOL, .NET, Java, etc.

Questions

Mapping between the formulas XBRL and the Multidimensional Data Model.

Academic Track @ XBRL Week,

Frankfurt (Germany). June, 1st – 2nd 2016

European Central Bank.

Ignacio Santos, PhD

Carlos III University of Madrid, Computer Science Department