

Using Fact-Based Modelling to Develop a Common Language

A Use Case

Inge Lemmens¹(✉), Jan Mark Pleijsant², and Rob Arntz³

¹ CogNIAM Finance, Amsterdam, The Netherlands
Inge.Lemmens@cogniamfinance.com

² ABN AMRO, Amsterdam, The Netherlands
Jan.Mark.Pleijasant@nl.abnamro.com

³ I-refact, 's Hertogenbosch, The Netherlands
Rob.Arntz@i-refact.com

Abstract. In today's business environment, the challenges an organization has to face have increased in amount and complexity. Not only competition has become tougher, organizations, and in particular financial institutions have to fulfil an increasing amount of regulations imposed by external organizations. To fulfil these legal obligations, a common understanding is required to remove ambiguities within the organizations and to ensure correct reporting.

A common understanding is achieved through the use of a common language in which each relevant term is foreseen of a single Definition that contains no ambiguity such that the risk of misinterpretation is reduced drastically and the time spent on research in case of a new reporting query coming from a (change of) legislation decreases. In this paper, it is explained how fact-based modelling is used to develop the common understanding, by using the fact types as the basic building blocks for the Definitions.

Keywords: Fact-Based Model · Common language · Concepts and definitions

1 Introduction

In today's business environments, the challenges an organization has to face have increased in amount and complexity. Not only competition has become tougher, organizations, and in particular Financial Institutions, have to fulfil an increasing amount of regulations imposed by external organizations like the "National Central Bank", the "National Financial Authority", and the "European Central Bank".

For an organization to be compliant to ever changing regulation it is of importance to understand the regulation and to have an unambiguous translation of the language of the regulation into the language used within the organization. This is a challenge if there is no common language that is used in an organization. And even if the common language exists, without a common understanding of the terms used in the common language, there are no guarantees that compliance is achieved.

Developing a common language is often thought of as developing a lexicon (an alphabetic list of words with information about used in a given field). However, if not done properly, the risk of misinterpretation due to ambiguities and inconsistencies in the lexicon remains. All too often, a lexicon is developed by defining the terms in a specific context, isolated from the other terms and their Definitions. That is, developing a lexicon is tackled as a “writing exercise”. In this paper, a modelling approach to the development of a lexicon is introduced. This approach is currently in use within a large Financial Institution as the means to come to a common understanding of the terminology used within the organization.

The Semantic Information Model plays a central role in the Financial Institution’s goal of model-driven development. This role of the SIM, as conceived by the organization, is explained in section 2. In section 3 we introduce the Semantic Information Model (SIM) and its elements. To develop this Semantic Information Model, a development approach is developed to cater the specific requirements of the organization. This development approach is explained in section 4. To assure that the resulting lexicon is of the correct quality, quality requirements with respect to the terms and Definitions are identified, which are based on ISO/IEC 11179-1 [1]. This leads to a four-level qualification schema, which is introduced in section 5. In section 6, the future development of the SIM and associated relationships is explained.

2 The Role of the Semantic Information Model

The development of an organization-specific lexicon that is the single source of reference for the semantics associated with the Business Terms that are used in communication with the stakeholders is realized by developing a *Semantic Information Model (SIM)*, which is an information model that consists of Business Concepts, their associated Business Terms and Definitions as well as the Relationships between the Business Concepts such that the information structure of the organization is represented at semantic level. The Semantic Information Model aims to be the trusted source of information for all Business Concepts, associated terms and Definitions and Relationships. The Financial Institution also intends to use the Semantic Information Model to map all the data sources containing the actual data regarding the Business Concepts to the associated Business Terms such that lineage from Business Terms to data elements in all sources is achieved.

In the frame of model-driven development, the Financial Institution associates the following purposes to the Semantic Information Model:

1. It serves as the basis for the lexicon of Business Terms and associated Definitions, describing the common language of the organization.
2. It provides insight in the Relationships between the Business Concepts that are represented through the Business Terms such that dependencies between Business Concepts become insightful.
3. It serves as the central point of reference whereby external terms provided by e.g. legislative organizations, are matched against.

4. It serves as the central point of reference for all mappings to IT-related data models such that traceability throughout the complete chain of development (i.e. from Term as used in external glossaries like e.g. laws and regulations, to realization in IT application) is achieved.
5. It serves as communication mechanism from Business to Business, from Business to IT and from IT to Business.

3 The Elements of the Meta Model of the Semantic Information Model

One of the main purposes of the Semantic Information Model is to form the basis for the lexicon of the Financial Institution. That is, in early phases, the Semantic Information Model and the lexicon were considered to be the same thing. Therefore, the focus was on the Business Concepts, their associated Terms and Definitions. Later, the Relationships between Business Concepts were introduced as an important element, differentiating between the lexicon and the Semantic Information Model.

In the frame of “practicing what you preach”, the meta model Semantic Information Model is defined together with the Subject Matter Experts, on the basis of the protocol that is developed and which is explained in section 4. As said, central elements of the meta model of the Semantic Information Model are Business Concept, Business Term, Definition and Relationship.

3.1 Business Concepts and Business Facts

In order to determine the elements of the meta model of the Semantic Information Model, inspiration is taken from the CogNIAM framework [2, 3]. Thereby, a selection is made by looking at the FIBO Foundations [4], the Financial Industry Business Ontology Foundations meta model.

The central element in the meta model of the Semantic Information Model is the Business Concept. A Business Concept is defined as “*a Thing that is important enough to the Business that Business Facts about the Thing are necessary to run the Business*”. This definition implies that Business Concepts are only limited to those Things that are used to run the Business. This aids in the classification of Business Concepts versus not-so-relevant concepts. For example, in the case of the Financial Institution, through the definition of Business Concept, the Thing denoted by the Term “Credit risk” would be denoting a Business Concept while the Thing denoted by the Term “Housecat” would not be denoting a Business Concept.

In the definition of Business Concept, there is a reference to “Thing” and “Business Fact”. For defining “Thing”, the definition as stated in FIBO is taken, namely: “*a Thing is a set of elements which are defined according to the facts given for that kind of things.*”. In fact-based modelling terminology, a “Thing” as defined above represents an Object Type.

A Business Fact is defined as a “*fact that describes a Business Concept*”. For example, “first name” would be a “business fact” about the Business Concept

“Individual”. It should be noted that the term Business Fact might be confusing since it represents a *type of fact*, not a specific fact. However, the term “fact type” is not accepted by the Business users while the term “fact” is acceptable.

The Financial Institution has a need to distinguish between “Characteristics” and “Relationships”, whereby a business fact either represents a Characteristic or a Relationship. This distinction is introduced because not only “Characteristic” and “Relationships” are terms that the business is acquainted with, it is also used for the mapping to the underlying technical models.

A Characteristic is defined as: “a Business Fact that represents an aspect of a Business Concept”, while a Relationship is defined as: “a Business Fact that represents a meaningful link between two or more Business Concepts”.

The model fragment associated with the elements of the meta model of the Semantic Information Model introduced above, are given in Figure 1.

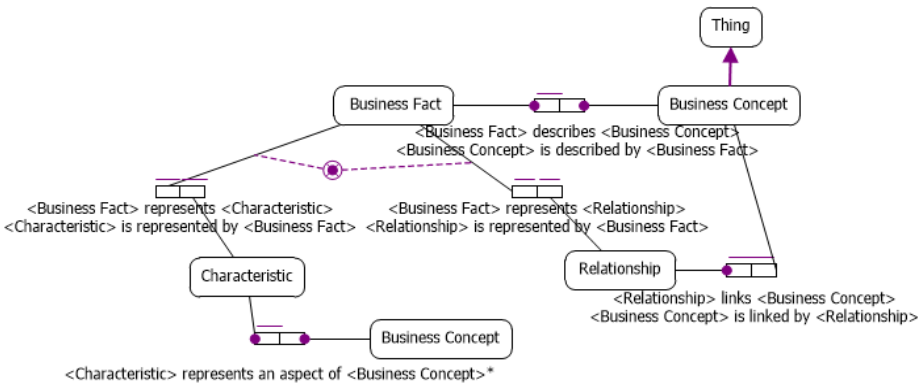


Fig. 1. Model fragment for Business Concept and Business Fact.

3.2 Business Terms

A Business Concept can only be talked about if there is a Business Term that denotes the Business Concept. Therefore, it is stated that each Business Concept is denoted by at least one Business Term. A Business Term is defined as: “a word or phrase that designates a Business Concept”. Moreover, the Financial Institutions has also realized the potential need for using a term to denote a Business Fact.

From an enterprise-wide common language perspective, the Financial Institution has decided that for each Business Concept, there is exactly one Preferred Business Term. Synonyms are allowed, but they are only allowed as reference to the Preferred Business Term associated with the Business Concept. For example, if the Business Term “Customer” is the preferred term to denote the Business Concept, then the Definition of the synonym “Client” is a mere reference to the Business Term “Customer”, by stating in the Definition of “Client”: “see Customer”. Figure 2 depicts the model fragment of the Semantic information Model that captures the above.

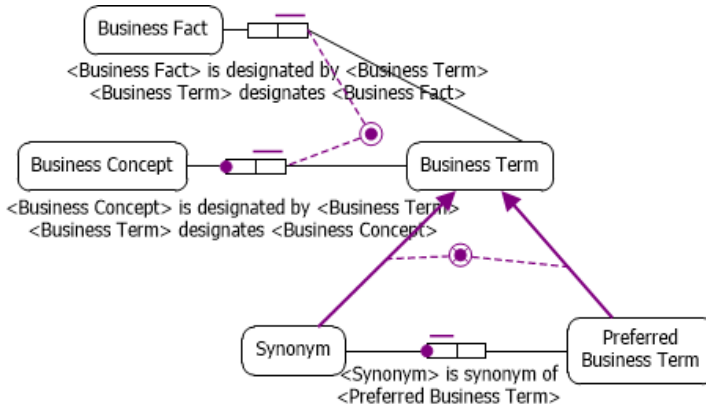


Fig. 2. Model fragment associated with Business Term.

3.3 Definitions

Associated with a Business Concept is a Definition. That is, for each Business Concept, there is exactly one Definition. However, in order to tailor for the option to specify in the Definition of a synonym the reference to the preferred Business Term, the choice has been made to state that with each Business Term exactly one Definition is associated. A Definition is thereby defined as: “*a phrase that states the exact meaning of a Business Term*”.

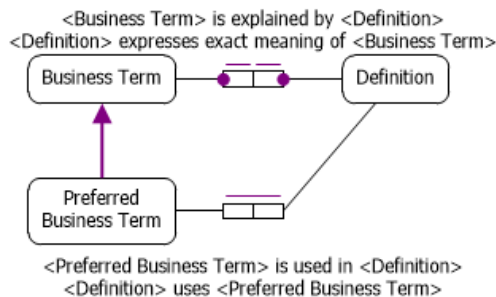


Fig. 3. Model fragment associated with Definition.

As shown in Figure 3, in a Definition, preferred Business Terms are used. This to fulfil one of the quality criteria for Definitions that are described in section 5.

4 The Protocol

To “populate” the meta model of the Semantic Information Model structure as defined above (i.e. to develop the Semantic Information Model), the Financial Institution has developed a protocol that aids in developing correct and consistent Definitions. As the observant reader can deduce from above, the identification of the Relationships between

the concepts are determining the Definitions. That is, the Definitions are derived from the fact-based model that is developed. By doing so, the consistency between Definitions is guaranteed, and no Business Concept is defined in isolation.

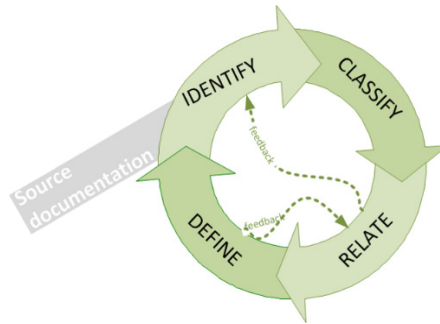


Fig. 4. The 4 phases of the overall development process.

As can be seen in Figure 4, the overall process for the development of the Semantic Information Model can be split in 4 main phases, namely:

1. The *identify* phase in which Business Concepts are identified and foreseen of a preliminary meaning and context.
2. The *classify* phase in which Business Concepts are formally placed in an overall hierarchy.
3. The *relate* phase in which the Business Concepts are formally placed in relationship to each other.
4. The *define* phase in which the Definition is created based on the outcome of the three previous phases.

Each of the phases identified consists of a set of activities that have to be performed in order to result in a concise, correct and understandable Semantic Information Model. These activities are inspired by the CogNIAM protocol as described in [2]. The use of concrete examples to illustrate the business concepts and Relationships has shown to be of a great advantage to come to a common understanding of the Business Terms.

4.1 Definition Development and Validation Process

To implement the protocol, the cooperation of the subject matter experts (SMEs) is important. Therefore, the protocol is supported by a development and validation process that is implemented in the organization to ensure the correct quality of Definitions. This process consists of:

1. A development phase, in which the Semantic Information Model is developed in accordance to the protocol. In this phase, 3 workshops with the SMEs are identified, namely:

- (a) The clarification workshop in which clarification of context, meaning and purpose of the Business Concept is given by the subject matter experts. In this workshop, a sketch of the Relationships between the Business Concepts is the major outcome.
 - (b) The Definition workshop in which the Business Concepts are classified, the Relationships identified in the previous workshop are fine-tuned and the Business Concepts are defined in accordance to the identified Relationships.
 - (c) The confirmation workshop in which the Definitions are fine-tuned and confirmed by the subject matter experts.
2. A challenge phase, in which the Definitions of the Business Concepts and their Relationships are challenged by other subject matters experts to ensure that the intended meaning is correct. This is considered to be a quality check on usefulness, comprehensiveness and consistency of the Business Concepts.
 3. A review phase in which the common understanding is reviewed by a broad community throughout the different business domains that use the Business Terms.
 4. An approval phase in which the bank-wide Data Definition Board is responsible for approving the identified Business Concepts and their associated Definitions and Relationships.

This overall development and validation process is implemented to ensure that the developed Semantical Information Model is a consistent and coherent model, whereby the Definitions constructed fulfil the quality requirements listed in the next section.

5 Quality Criteria

In [1], ISO has identified several quality criteria for data definitions. These quality criteria have been taken as the basis for developing quality criteria for the Definitions of the Business Concepts. Thereby, the criteria are organized to differentiate between different levels of quality. The aim of the Financial Institution is to get the Definitions to level 3.

The quality criteria themselves are defined as follows:

1. Level I criteria – from incoherent to ambiguous:
 - (a) A Definition is a descriptive phrase or sentence.
 - (b) A Definition does not contain a reference to itself.
 - (c) A Definition follows the following pattern: “a <Business Term> is a <more general Business Term in hierarchy> that/which/what/who <discriminating reasons>.”.
2. Level II criteria – from ambiguous to unambiguous
 - (a) A Definition states what the concept is, not only what it is not.
 - (b) A Definition states the essential meaning.
 - (c) A Definition is unambiguous.
 - (d) A Definition does not introduce a second-order circular reference.
 - (e) A Definition does not introduce a contradiction with or between other Definitions.

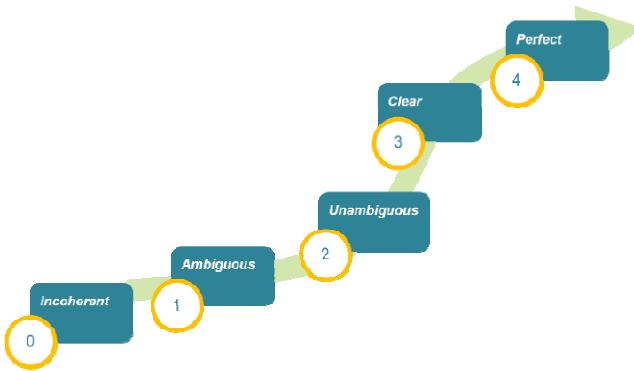


Fig. 5. The different levels of quality of the Definitions.

3. Level III criteria – from unambiguous to clear:

- (a) A Definition is stated in the singular.
- (b) A Definition contains only common understood abbreviations and terms.
- (c) The form of a Definition is such that it can replace the term in context.
- (d) A Definition uses the same terminology and consistent logical structure as related Definitions.
- (e) A Definition does not require additional explanations or other references to understand the meaning.
- (f) A Definition follows the writing guidelines of the Financial Institution.

4. Level IV criteria – from clear to perfect:

- (a) A Definition is expressed without embedding Definitions of other data or underlying concepts.
- (b) A Definition is expressed without embedding rationale, functional usage or procedural information.
- (c) A Definition consists of a single phrase specifying the concept.

6 Conclusions and Future Work

The development of a common understanding is gaining more and more interest from organizations, in particular from organizations that have to comply to all kinds of external regulations. This in turn has attracted interest of those organizations in semantic modelling. What distinguishes the Financial Institution whose way of working is described in this paper from other organizations is their well-defined way of working.

This way of working has not come about in a week’s time. It has taken the Financial Institution many months to come to this structure for the Semantic Information Model and the associated protocol. Earlier versions of the protocol did not emphasize the use of Relationships to define the Business Concepts. Only by practicing the way of working together with CogNIAM experts has given them the insight to use the Relationships between the concepts (the so-called Business Facts) as the basis for the

Definition. Also, the identification of quality criteria and associated quality levels has come about only in the last month but has already proven to be very useful in the governance process.

The work has not finished. In the process of building the Semantic Information Model, the Financial Institution realizes that defining Business Concepts and their Relationships is not enough. For a good understanding, it is also required to identify data integrity rules, like uniqueness, mandatory and referential integrity, as well as derivation rules. The latter are in particular of relevance for defining reporting terms that are derivations of data that conforms to the Business Concepts.

In the near future, the current version of the Semantic Information Model will be extended by adding rules that ensure the integrity of the data associated with the Business Concepts. This as a first step to use the Semantic Information Model as a means to aid in solving data quality issues.

The Semantic Information Model is only one of the business models and does not stand in isolation. As part of business architecture, the Relationships between the Business Information Model and other types of business models, like the business process model and business rules model will be further developed. By doing so, an integrated way of working can be developed and true model-driven development can be achieved.

References

1. ISO/IEC.: ISO/IEC 11179-1 Information technology – Metadata registries (MDR) – Part 1: Framework (2004)
2. Nijssen, S., Valera, S.: An Architecture Ecosystem for the Whole Systems Perspective, Including System Dynamics, Based on Logic & Set Theory and Controlled Natural Languages. Working paper for the OMG Architecture Ecosystem SIG (2012)
3. Lemmens, I., Valera, S.: Achieving interoperability at semantic level. In: PV 2013, Ensuring Long-Term Preservation and Adding Value to Scientific and Technical Data (2013)
4. Object Management Group: Financial Industry Business Ontology Foundations Version 1.0 – Beta 1 (2014). <http://www.omg.org/spec/EDMC-FIBO/FND/1.0/Beta1/>