



HAL
open science

Semantics enactment for interoperability assessment in Enterprise Information Systems

Esma Yahia, Mario Lezoche, Alexis Aubry, Hervé Panetto

► **To cite this version:**

Esma Yahia, Mario Lezoche, Alexis Aubry, Hervé Panetto. Semantics enactment for interoperability assessment in Enterprise Information Systems. *Annual Reviews in Control*, 2012, 36 (1), pp.101-117. 10.1016/j.arcontrol.2012.03.008 . hal-00671856

HAL Id: hal-00671856

<https://hal.science/hal-00671856v1>

Submitted on 19 Feb 2012

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Semantics enactment for interoperability assessment in Enterprise Information Systems

Esma Yahia^{*}, Mario Lezoche, Alexis Aubry, Hervé Panetto

*Research Centre for Automatic Control (CRAN), University of Lorraine, CNRS, Campus Scientifique, Faculté des Sciences et Techniques, BP 70239, 54506 Vandoeuvre-lès-Nancy Cedex, France
(e-mail: {Esma.Yahia, Mario.Lezoche, Alexis.Aubry, Herve.Panetto}@univ-lorraine.fr)*

Abstract: The grown complexity of the modern enterprise poses a series of challenges, among them keeping competitiveness in the fast changing environment in which the enterprise evolves. Addressing Enterprise Integration is considered as a key to achieve the goal of any enterprise either it is a single or a networked enterprise. Enterprise Modelling is a prerequisite to enable the common understanding of the enterprises and its various interactions in order to “provide the right information, at the right time, at the right place”. However, problems often emerge from a lack of understanding of the semantics of the elaborated models resulting from various modelling experience based on different methods and tools. This paper describes the challenges associated to semantics enactment in Information Systems models. To facilitate this enactment, it proposes an approach based on a fact-oriented modelling perspective. Then, it also provides an algorithm to automatically build semantic aggregates that help in highlighting Enterprise Models core embedded semantics. A case study on the field of B2M interoperability is performed in order to illustrate the application of the presented approach.

Keywords: Enterprise Models, Information Systems, Semantics Enactment

1. INTRODUCTION

When evolving in a competitive global market, enterprises are forced to become increasingly agile and flexible in order to manage the fast changing business conditions. Today’s challenges mainly concern Enterprise Integration (EI). Indeed, *EI deals with removing organisational barriers and/or improving interactions among people, systems, applications, departments, and companies (in terms of material, informational, decision and workflows)* (Vernadat, 2009).

Enterprise Modelling (EM) plays a critical role in this integration, enabling the capture of all the information and knowledge relevant for the enterprise operations and organisation ((Boudjlida and Panetto, 2008); (Vernadat, 1996)).

The produced Enterprise Models are mainly related to artefacts such as processes, behaviours, activities, information, resources, objects/material flows, goals, systems infrastructure and architectures....Those Enterprise Models must contain the necessary and sufficient semantics in order to be intelligible and then enabling the global Enterprise Integration. For instance, the business semantics of the process model is mainly brought along by languages such as the Business Process Modelling Notation (BPMN¹). Moreover, enriching this semantics is still an open issue; researches made by (Boudjlida and Panetto, 2008) could be quoted in terms of process models annotations.

¹ <http://www.bpmn.org>

Among all Enterprise Models, Information Systems (IS) models are considered as the core models of the enterprise. Concretely, the complexity of EI relies on the fact that an enterprise (a single or a networked enterprise) comprises numerous and heterogeneous Information Systems either at the business or manufacturing level such as ERP (Enterprise Resource Planning), MES (Manufacturing Execution System), SCM (Supply Chain Management), PDM (Product Data Management) and CRM (Customer Relationship Management). Those ISs need i) to share specified information and ii) to operate on that information according to a shared operational semantics iii) in order to realise a specified purpose in a given context. Achieving these actions is commonly called *interoperation* (Whitman, Santanu and Panetto, 2006). Fisher, in (Fisher, 2006), characterises also by the term *interoperation*, the relationship that consists on the exchange and the cooperative use of information.

Ensuring semantic interoperability between heterogeneous enterprise information systems, in a collaborative context, means analysing and manipulating their conceptual models. However, these conceptual models are non-exploitable for evaluation the interoperation. In fact their semantics, due to the heterogeneity of the models (tools, methods), is often tacit and maybe scrambled due to the implementation requirements or non-expressed as depending on the practices and expertise levels of the modeller (Castano and De Antonellis, 1998). Thus, when studying the interoperability between enterprise information systems, enacting the implicit semantics is necessary.

The section 2 highlights challenges associated with the enactment of semantics in enterprise information systems. One of these challenges is the enactment of the semantics structure. In section 3, the paper proposes to structure the semantics of a conceptual model (associated to an enterprise application). The proposed approach consists in identifying the “core” and “extended” semantics and, then, in recursively building some semantic aggregates of concepts (called the “semantic blocks”). Each one of these semantic aggregates identifies and emphasises the border of one sub-system model with its own core semantics. It focuses on “what is important” in the system without taking care on implementation artefacts. the paper shows in section 4, how the semantic blocks can be used for evaluating if two information systems can interoperate through their subsystems (identified by the semantic blocks).

In order to illustrate the proposed approach, a case study is presented in the section 5. This case study deals with B2M (Business to Manufacturing) interoperability requirements between an Enterprise Resource Planning (ERP) system and a Manufacturing Execution System (MES) applications and consists in applying our approach in order to extract the semantics embedded into those ISs.

Finally, this paper concludes with some remarks and perspectives for on-going research.

2. Enterprise information systems Conceptual Models

Information Systems are tools that contribute to the management, the treatment, the transport and the diffusion of the information through the enterprise or through several networked enterprises. Thus, ISs are critics as they contribute to the achievement of the enterprise mission. The collaborative context has led to the necessity of making interoperate a lot of information systems – that are autonomous and heterogeneous (architectures, functions, models,...) – for achieving the global collaboration objective.

For achieving this need of interoperation, information must be physically exchange (technical interoperability), must be understood (conceptual interoperability) and must be used for the purpose for which it has been produced (conceptual and organisational interoperability).

Castano and De Antonellis, in (Castano and De Antonellis, 1998), have shown that information systems conceptual models are generally defined and engineered by different experts (with their own experience of modeling). These experts use a lot of different enterprise models, which are necessarily heterogeneous. This heterogeneity can be syntactic or semantic.

- According to (Benhlila and Chiadmi, 2006), the syntactic heterogeneity is due to the different formats of storage (XML,...), the different query languages (XQuery, SQL,...), the different access protocols (HTTP,...)...
- According to (Park and Ram, 2004), the semantic heterogeneity is due to the different possible interpretations of the real world by modelers when building the information systems. For instance, (Kavouras, 2003) propose two possible causes that can lead to semantic heterogeneity: the covering difference (details level) and the classification difference (conceptualization).

Some standard of exchange (STEP, ISO...) have contributed to solve syntactic heterogeneity. However, the semantic conflicts are still an open issue.

2.1. Semantic conflicts in enterprise models

The semantic conflicts in enterprise conceptual models highlight problems for interpreting the models due to difficulties for making explicit the tacit semantics. Zouggar, in (Zouggar, Chen and Vallespir, 2009), proposed two types of semantic problems when studying enterprise models. The first type concerns the language used for modelling and the artefacts that define the semantics of the language. For instance, the concept of “function” has not the same semantics in CIMOSA² and in

² Computer Integrated Manufacturing Open System Architecture (<http://www.cimosa.de/>)

IDEF0³. In CIMOSA, a function represents the sequence of operations in production management whereas in IDEF0, a function treats the static aspects of a system. The second type concerns the conceptualisation. For instance, the concept “Employee” has not the same semantics in every enterprise. For a first enterprise, an employee is a person who works in the enterprise whereas for a second enterprise, an employee is a person who is not manager.

The difficulty of operating with the various Enterprise Models comes out from the fact that the majority of those models have been made by different experts with several modelling experiences. That has led, for instance, to various conceptual representations for the same semantics. Since the majority of conceptual models have been fulfilled a posteriori and not a priori, implementation-based functionalities and constraints can cause interferences in the semantics understanding of those models. Let us consider, for instance, the extract of two different conceptual models in figure 1. Intuitively, those classes carry the same semantics, but are modelled differently. For instance, the *WEIGHT* of a *PRODUCT* on the right side of the figure is represented by a class due to an implementation constraint; when other classes are related to it, this facilitates querying for specific values related to the weight for example. While, on the left side of the figure, the *WEIGHT* of a *PRODUCT* is modelled by two attributes (its value and its unit).

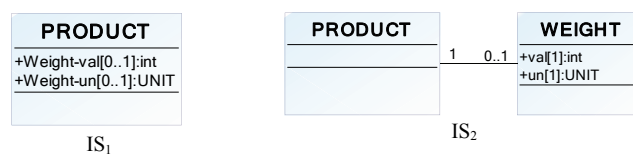


Figure 1 - Two extracts of conceptual models.

2.2 Fact-oriented modelling

While studying an implemented information system model, we observe that its semantics is tacit as it is scrambled due to implementation requirements. Interoperability assessment has then to take place at a conceptual level in order to enact any “grain” of semantics embedded in the application, allowing bringing out the tacit semantics and making it explicit (Yahia, Aubry and Panetto, 2012). Indeed, the starting point of our approach consists on a very detailed conceptual model where each “atom” of semantics embedded in any artefact (classes, attributes, associations, constraints ...) is presented with a unified concrete syntax.

³ Integration Definition for Function Modeling (<http://www.idef.com/idef0.htm>)

Fact-oriented modelling is a conceptual, natural language based approach avoiding such conflicting conceptual representations. It queries the information semantics of business domains in terms of the underlying facts of interest, where all facts and rules may be verbalised in a language readily understandable by users of those business domains (Halpin, 2007). Fact-oriented models are attribute-free, treating all elementary facts as relationships.

Object-Role Modelling (ORM) is the most popular fact-oriented approach. In fact, ORM makes no explicit use of attributes; instead it pictures the world in terms of lexical and non-lexical concepts that play roles (take part in relationships) (Halpin, 2006). This leads to a greater semantics stability and populatability, as well as facilitates natural verbalisation (Halpin, 2007).

In our work, ORM could be used as a modelling language. However, the existing conceptual models, in industrial context, are mainly represented with the UML notation. Hence getting a spread out of an attribute-free conceptualisation could be made using the UML notation but based on the ORM approach. Taking into account the ORM definitions, UML class diagram notation will be used and the UML concepts and the UML attributes are called respectively non-lexical concepts and lexical concepts.

When applying the fact-oriented modelling on the examples of the figure 1, the following models (figure 2), that eases the semantics enactment, are obtained.

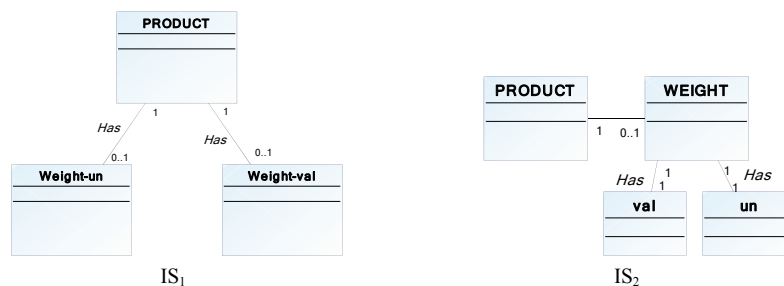


Figure 2 - The conceptual models of Figure 1 using the fact-oriented modelling perspective.

3. A SEMANTICS STRUCTURING PROCESS

After enacting finest-grained semantics embedded into CISs models, resulting with a normalised FOM, the semantics has to be structured into semantic aggregates (Yahia, Aubry and Panetto, 2012). Each of those identified aggregates represents a “semantic molecule”, composed of atomic concepts, with its own minimal mandatory semantics.

To build such aggregates, a recursive approach is proposed for analysing the detailed semantics of the IS conceptual models obtained by the conceptualisation approach presented in section 2. These models are considered to embed the whole explicit semantics of the associated IS.

The structuring approach starts by identifying core atomic concepts and it ends by computing the semantic aggregates (namely, the semantic blocks) according to algorithms based on graph theory.

3.1 Core and extended semantics

When considering an available fact-oriented conceptual model from one IS (outputs from section 2), we can distinguish the mandatory (constrained) and non-mandatory (non-constrained) association roles, which represent mandatory and non-mandatory concepts expressing semantics.

The set of mandatory concepts represents all the necessary and sufficient elements which make the conceptual model semantically coherent and understandable. It comprises of all the non-lexical and lexical concepts linked to constrained association roles with a multiplicity equal to 1 or 1..*. On the contrary, the non-mandatory concepts correspond to the non-mandatory roles (multiplicity equal to 0..1 or *) and are only enriching the semantics of those IS conceptual models.

To some extent, the set of mandatory concepts corresponds to the core semantics that is embedded into a given IS conceptual model: without this semantics, the IS cannot operate. The extended semantics is defined by the set of mandatory and non-mandatory concepts.

3.2 Some mathematical definitions

For each IS conceptual model, the following notations are defined.

Definition 1. C_{IS} is the set of the identified lexical and non-lexical concepts, formally defined by

$$C_{IS} = \{c_i | c_i \text{ is a lexical or a non-lexical concept from the IS conceptual model}\}$$

Moreover, two subsets of C_{IS} are defined as follows:

- NLC_{IS} is the subset of C_{IS} restricted to the non-lexical concepts and,
- LC_{IS} is the subset of C_{IS} restricted to the lexical concepts.

We can note that:

$$C_{IS} = NLC_{IS} \cup LC_{IS}$$

$$NLC_{IS} \cap LC_{IS} = \emptyset$$

Definition 2. Rel_{IS} is the set of the identified associations between concepts. Formally, it is defined by

$$Rel_{IS} = \{rel(c_i, c_j) | (c_i, c_j) \in (C_{IS})^2 \wedge c_i \text{ is associated to } c_j\}$$

Definition 3. $Mult(rel(c_i, c_j))$ is the multiplicity of the role of c_j when considering the association between c_i and c_j if it exists. For each $(c_i, c_j) \in (C_{IS})^2$, if $rel(c_i, c_j)$ exists then we have $Mult(rel(c_i, c_j)) \in \{*, 0..1, 1..*\}$ and it is read c_j is associated to c_i with a multiplicity equal to $Mult(rel(c_i, c_j))$.

Definition 4. MC_{IS} is the subset of C_{IS} restricted to mandatory concepts (the core semantics). It is formally defined by

$$MC_{IS} = \{c_i | \exists (c_i, rel(c_i, c_j)) \in C_{IS} \times Rel_{IS} \wedge Mult(rel(c_i, c_j)) \in \{1, 1..*\}\}$$

Moreover, we define two subsets of C_{IS} as follows:

- $MNLC_{IS}$ is the subset of C_{IS} restricted to the mandatory non-lexical concepts and,
- MLC_{IS} is the subset of C_{IS} restricted to the mandatory lexical concepts.

We can note that:

$$MC_{IS} = MNLC_{IS} \cup MLC_{IS}$$

$$MNLC_{IS} \cap MLC_{IS} = \emptyset$$

$$MNLC_{IS} = MC_{IS} \cap NLC_{IS}$$

$$MLC_{IS} = MC_{IS} \cap LC_{IS}$$

Definition 5. For each non-lexical concept c_j , we can define the set of its associated mandatory lexical concepts as follows:

$$MLC(c_j) = \{c_i \in LC_{IS} \mid (\exists rel(c_j, c_i) \in Rel_{IS} \mid Mult(rel(c_j, c_i)) \in \{1, 1..*\})\}$$

Definition 6. For each non-lexical concept c_j , we can define the set of its associated mandatory non-lexical concepts as follows:

$$MNL C(c_j) = \{c_i \in NLC_{IS} \mid (\exists rel(c_j, c_i) \in Rel_{IS} \mid Mult(rel(c_j, c_i)) \in \{1, 1..*\})\}$$

If we consider a concept defined in the context of the IS core semantics, we notice that, in order to be semantically effective in the studied domain, this concept needs to be associated on the one hand to its mandatory lexical concepts and on the other hand to other non-lexical concepts. This defines the notion of Semantic Block (SB).

Let us consider the conceptual model of the Figure 3 and its fact-oriented modelling transformation on Figure 4 according to section 2. For this example we have:

$$NLC_{IS} = \{C1; C2; C3; C4; C5; C6; C7; C8\}$$

$$LC_{IS} = \{A1C1; A2C1; A1C2; A2C2; A3C3; A1C3; A2C3; A1C4; A2C4; A1C5; A2C5; A1C6; A2C6; A1C7; A2C7; A1C8; A2C8\}$$

$$MNL C_{IS} = \{C2; C3; C4; C5; C6; C7; C8\}$$

$$MLC_{IS} = \{A1C1; A1C2; A2C2; A1C3; A1C4; A1C5; A1C6; A1C7; A1C8\}$$

$$MNL C(C2) = \{C5\}$$

$$MLC(C2) = \{A1C2; A2C2\}$$

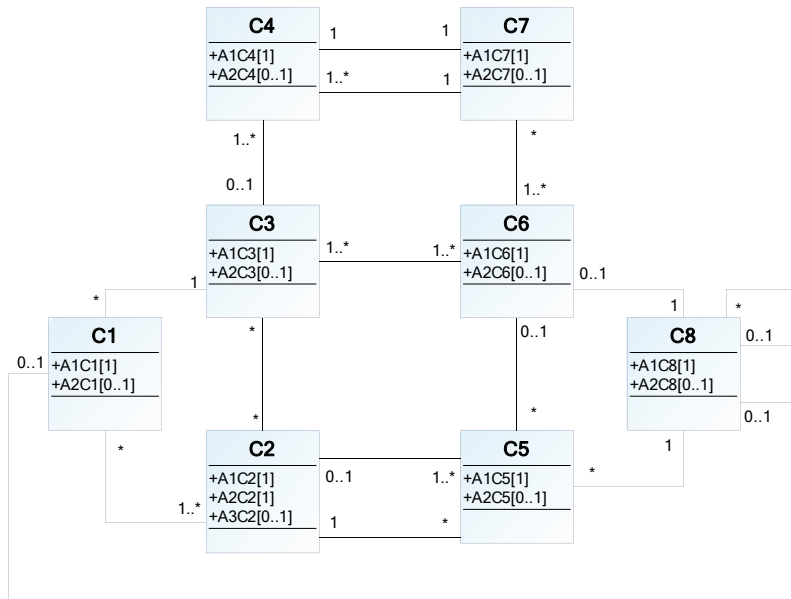


Figure 3 - An instance of conceptual model

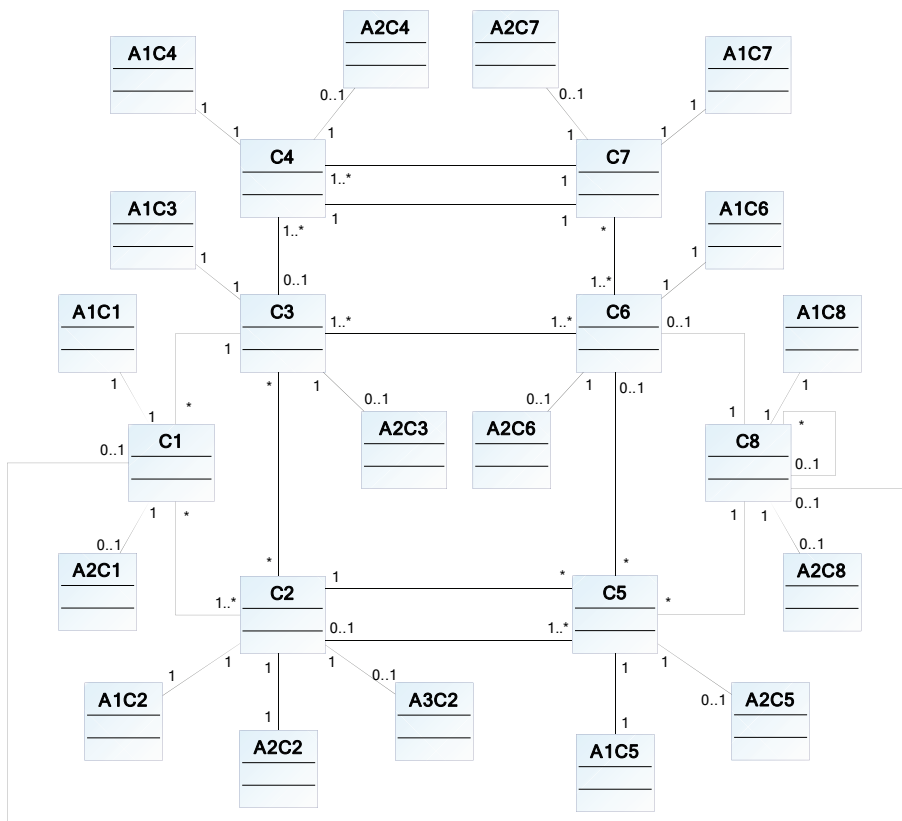


Figure 4 - "Fact-oriented modelling" transformation of the model of Figure 3

3.3. Semantic blocks identification

3.3.1 Definition

Considering a particular non-lexical concept c_i from NLC_{IS} , a semantic block, denoted as $SB(c_i)$ and associated with the concept c_i , represents the set of the concepts necessary for the minimal semantics definition of the non-lexical concept c_i given by the conceptual model.

Let us consider the example on Figure 4. Let us build the semantic block of the concept $C2$. The intrinsic mandatory semantics of the concept $C2$ is defined by the semantics of the mandatory lexical concepts that are associated to it, namely $A1C2$ and $A2C2$. Moreover, a given instance of the concept $C2$ exists only if it is associated to at least one instance of the concept $C5$. That means that $C5$ is mandatory for expressing the semantics of $C2$. Recursively, the intrinsic mandatory semantics of the concept $C5$ is defined by $SB(C5)$. Thus $SB(C5)$ is mandatory for expressing the semantics of $C2$. Moreover, considering the roles of $C1$ and $C3$ in their association with $C2$, the minimal multiplicity is, then, equal to 0. That means that the existence of any instance of $C2$ is not stipulated by the existence of one instance of $C1$ or $C3$ ($C1$ and $C3$ only give a supplementary non-mandatory semantics for understanding $C2$). Finally, the semantic block of the concept $C2$ is $SB(C2) = \{C2 \cup \{A1C2, A2C2\} \cup SB(C5)\} = \{C2\} \cup MLC(C2) \cup_{c_j \in MNLC(C2)} SB(c_j)$.

By generalising this result, $SB(c_i)$ can be formally defined as follows:

$$SB(c_i) = \{c_i\} \cup MLC(c_i) \cup_{c_j \in MNLC(c_i)} SB(c_j) \quad (1)$$

This definition, suggests that the notion of semantic block is recursive.

For the example, the semantic blocks of $C5$ and $C8$ could be defined by $SB(C5) = \{C5 \cup \{A1C5\} \cup SB(C2) \cup SB(C8)\}$ and that $SB(C8) = \{C8 \cup \{A1C8\}\}$. Finally, the semantic blocks of $C2$ is deduced: $SB(C2) = \{\{C2, C5, C8\} \cup \{A1C2, A2C2\} \cup \{A1C5\} \cup \{A1C8\}\}$.

3.3.2 How to compute the Semantic blocks?

To simplify the computation of the semantic block of one concept c_i , this paper proposes, first, to identify the set of non-lexical concepts that are included in the semantic block and, second, to add the associated mandatory lexical concepts. That means that $SB(c_i)$ is determined as follows: $SB(c_i) = SB_c(c_i) \cup SB_a(c_i)$ with

- $SB_c(c_i) = \{c_i\} \cup_{c_j \in MNL(c_i)} SB_c(c_j)$ and,
- $SB_a(c_i) = MLC(c_j) | c_j \in SB_c(c_i)$

For the example, $SB_c(C2) = \{C2, C5, C8\}$ and $SB_a(C2) = \{A1C2, A2C2, A1C5, A1C8\}$.

In the following, the graph theory is used to compute all the semantic blocks of a conceptual model.

3.4 Using graph theory for building $SB_c(c_i)$

To facilitate the building of the semantic blocks, it is needed to identify, for each c_i from NLC_{IS} , the associated set $SB_c(c_i)$ by using graph theory modelling and its associated mathematical tools.

Let us first define a semantic-dependency graph associated with a conceptual model. This semantic-dependency graph is a digraph $G = (V, E)$ where V is the set of nodes and E is the set of edges defined by a pair of nodes. Each node from V represents a non-lexical concept of the conceptual model. Each edge from E is built from the conceptual model as follows: the edge (c_i, c_j) exists if (i) there is an association between c_i and c_j in the conceptual model, and (ii) if the minimal multiplicity for the role of c_j is equal to 1 ($c_j \in MNL(c_i)$). That means that the existence of the edge (c_i, c_j) represents the fact that c_j is mandatory for expressing the semantics of c_i .

The Figure 5 shows the semantic-dependency graph associated with the conceptual model of the Figure 4.

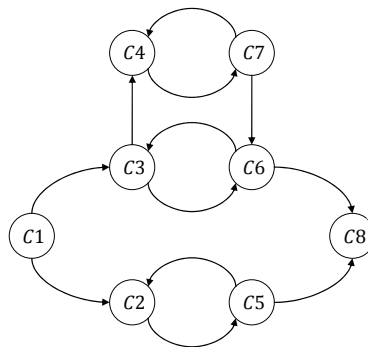


Figure 5 - Semantic-dependency graph associated with the conceptual model of Figure 4

Theorem 1. Given two particular concepts c_i and c_j , c_j belongs to $SB_c(c_i)$ if and only if there exists a directed path from c_i to c_j .

Proof. Let us consider a conceptual model and its associated semantic-dependency graph. To build the semantic block of the concept c_i , it is necessary to consider this concept as the starting point. This concept can thus be considered as the root in the semantic-dependency graph. Now we add in $SB_c(c_i)$ all the concepts c_k that must be instantiated to ensure the existence of a particular instance of c_i , i.e. all the concepts c_{1k} such that there is an association between c_i and c_{1k} in the conceptual model, and the minimal multiplicity for c_{1k} , considering this association, is equal to 1. This is the exact definition of all the successors of c_i in the semantic-dependency graph. Note that, by definition, there is a directed path from the concept c_i to these concepts c_{1k} . Iteratively, the only new concepts c_{2k} that can be added to $SB_c(c_i)$ are the successors of those first concepts c_{1k} . As successors of the concepts c_{1k} , there exists also a directed path from the concept c_i to the concepts c_{2k} (the path from c_i to c_{1k} plus the edge (c_{1k}, c_{2k})). Finally the semantic block of c_i contains exactly all the concepts c_j such that there exists a directed path from c_i to c_j . ■

Theorem 2. Given two particular concepts c_1 and c_2 , if c_2 belongs to $SB_c(c_1)$ then $SB_c(c_2)$ is included in $SB_c(c_1)$.

Proof. c_2 belongs to $SB_c(c_1)$ means that there exists a path from c_1 to c_2 (see theorem 1). Let us now consider a particular concept from $SB_c(c_2)$ denoted as c . By definition of $SB_c(c_2)$, there exists a path from c_2 to c and then a path from c_1 to c (the path from c_1 to c_2 plus the path from c_2 to c). That means that c is in $SB_c(c_1)$. Finally $SB_c(c_2) \subseteq SB_c(c_1)$. ■

Theorem 3. All the concepts that are in the same cycle in the semantic-dependency graph are associated with the same unique semantic block.

Proof. A cycle is a closed path. Let us consider two particular concepts, denoted as c_i and c_j , which belong to a cycle. In particular there is a path from c_i to c_j . That means that c_j is in $SB_c(c_i)$. Following the theorem 2, it is also possible to demonstrate that $SB_c(c_j) \subseteq SB_c(c_i)$. Moreover, there is a path from c_j to c_i . That means that c_i is in $SB_c(c_j)$. Following the theorem 2, that means that $SB_c(c_j) \supseteq SB_c(c_i)$. Finally, $SB_c(c_j) = SB_c(c_i)$. ■

The theorem 3 implies that there is one semantic block per strongly connected component⁴ of the semantic-dependency graph.

⁴ A strongly connected component of a directed graph is a maximal set of vertices such that for every pair of vertices u and v , there is a directed path from u to v and a directed path from v to u .

3.5 A semantic block meta-model

The semantic block architecture is formalised through the meta-model represented on Figure 6. This meta-model is based on the composite pattern (Gamma et al., 1995). This meta-model defines an arborescence of components representing hierarchies of objects and interprets the results given by the previous theorems.

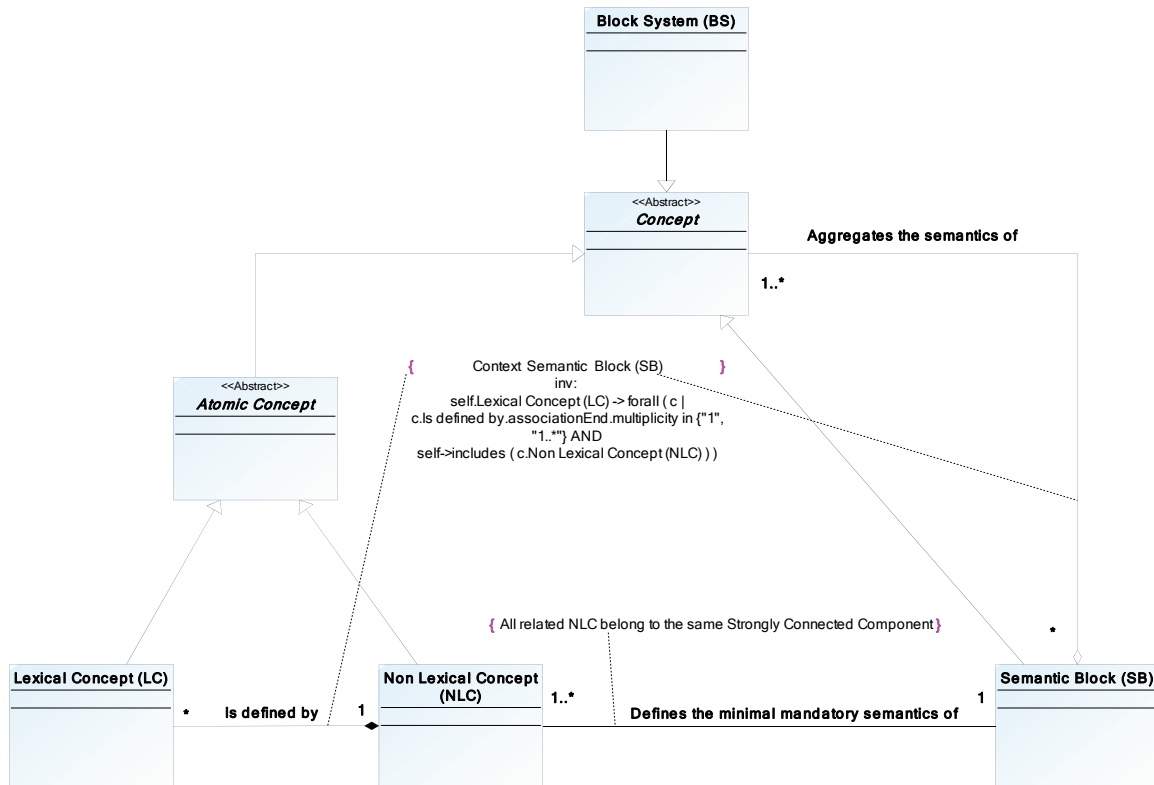


Figure 6 - Meta-model of the semantic block structure

A semantic block defines the minimal mandatory semantics of one or several non-lexical concepts such that these concepts are in the same strongly connected component (see theorem 3). Moreover, the semantics of one or several concepts can be aggregated into one or several semantic blocks. As the semantic block is a specialisation of the abstract class “Concept”, its semantics can be aggregated into one or several semantic blocks of higher levels (see theorem 2). The Block System represents the last level of aggregation and contains the minimal mandatory semantics of the studied IS conceptual model.

3.6 A procedure to compute the semantic blocks

Applying theorems 1 to 3, the following procedure is proposed to compute all the semantic blocks of a given conceptual model:

- i. Building the associated semantic-dependency graph.
- ii. Building the graph of the strongly connected components based on the semantic-dependency graph.
- iii. Computing the semantic blocks SB_c associated with each strongly connected component.
- iv. Computing, for each SB_c , the semantic block SB_a by adding all the mandatory lexical concepts associated to each non-lexical concept from SB_c .
- v. Computing $SB = SB_c \cup SB_a$.

These steps are detailed as follows.

3.6.1 Building the associated semantic-dependency graph

By definition of this graph, it can be easily obtained by considering each association between two concepts c_i and c_j and then building an edge from c_i to c_j if the minimal multiplicity for the role of c_j is equal to 1.

3.6.2 Building the graph of the strongly connected components

Theorem 3 implies that for building the semantic blocks, it is important to consider only one concept in a given strongly connected component (the other concepts share the same semantic block). That is the reason why the semantic-dependency graph could be simplified by considering only an equivalent graph where the nodes represent each strongly connected component of the former semantic-dependency graph, and where one of these nodes (e.g. SCC1) is connected to another node (e.g. SCC2) if there exists at least one edge from a concept from SCC1 to a concept from SCC2.

Identifying all the strongly connected components of a graph is a well-known problem in graph theory that can be solved with polynomial effort by using for instance Kosaraju-Sharir's algorithm (Kavouras, 2003).

The graph of the strongly connected components related to the semantic-dependency graph of Figure 5 is given on Figure 7. On this graph, the strongly connected components are defined as follows $SCC1 = \{C1\}$, $SCC2 = \{C2, C5\}$, $SCC3 = \{C3, C4, C6, C7\}$ and $SCC4 = \{C8\}$.

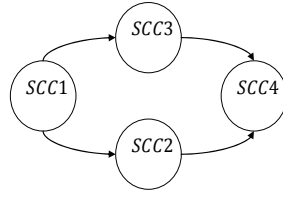


Figure 7 - Graph of the strongly connected components related to the graph of Figure 5

3.6.3 Computing SB_c associated with each strongly connected component

An algorithm is now proposed for computing all the semantic blocks SB_c associated with each strongly connected component (see Algorithm 1 that invokes Algorithm 2). The algorithm 1 BuildSemBlocks is applied on the graph of the strongly connected components (denoted as G_{SCC}).

Let us apply the algorithm BuildSemBlocks(G_{SCC}) on the graph of Figure 7. The following semantic blocks could be obtained:

- $SB_c(SCC1) = SCC1 \cup SCC2 \cup SCC3 \cup SCC4$,
- $SB_c(SCC2) = SCC2 \cup SCC4$,
- $SB_c(SCC3) = SCC3 \cup SCC4$ and
- $SB_c(SCC4) = SCC4$.

And finally replacing the strongly connected components by their content the following semantic blocks are obtained:

- $SB_c(C1) = \{C1, C2, C3, C4, C5, C6, C7, C8\}$,
- $SB_c(C2, C5) = \{C2, C5, C8\}$,
- $SB_c(C3, C4, C6, C7) = \{C3, C4, C6, C7, C8\}$ and
- $SB_c(C8) = \{C8\}$.

Algorithm *BuildSemBlocks*(G_{SCC})

[Initialisation]

L: List of the strongly connected components in G_{SCC}

For each $SCC \in L$ **Do**

$color(SCC) = -1$

[color is an indicator that defines if a node SCC has already been visited or not]

[-1 means not yet visited]

[0 means being visited]

[+1 means already visited]

```

Next SCC
For each SCC  $\in L$  Do
  If  $color(SCC) = -1$  Then
    [Building of the semantic block associated with
     SCC]
    BuildSB(SCC)[calling Algorithm 2]
  EndIf
Next SCC
Return

```

Algorithm 1. BuildSemBlocks algorithm

```

Algorithm BuildSB(SCC)


---


[Initialisation]
 $SB_c(SCC) = SCC$  [The semantic block associated with
 SCC initially contains all the concepts in the SCC]
 $color(SCC) = 0$  [SCC is being visited]
[Building]
[use of theorem 1]
For each SCC' successor from SCC in  $G_{SCC}$  Do
  If  $color(SCC') = -1$  Then
    [Building of the semantic block associated with
     SCC']
    BuildSB(SCC')
  EndIf
  [Use of theorem 2]
   $SB_c(SCC) = SB_c(SCC) \cup SB_c(SCC')$ 
Next SCC' successor from SCC in  $G_{SCC}$ 
Return  $SB_c(SCC)$ 

```

Algorithm 2. BuildSB algorithm

3.6.4 Computing, for each SB_c , the semantic block SB_a

Each semantic block SB_a contains the mandatory lexical concepts associated to the non-lexical concepts in SB_c . By applying the definition of SB_a ($SB_a(c_i) = MLC(c_j) | c_j \in SB_c(c_i)$) on the instance of Figure 4 we obtain:

- $SB_a(C1) = \{A1C1, A1C2, A2C2, A1C3, A1C4, A1C5, A1C6, A1C7, A1C8\}$,
- $SB_a(C2, C5) = \{A1C2, A2C2, A1C5, A1C8\}$,
- $SB_c(C3, C4, C6, C7) = \{A1C3, A1C4, A1C6, A1C7, A1C8\}$ and
- $SB_c(C8) = \{A1C8\}$.

3.6.5 Computing each semantic block SB

Each semantic block SB is the union of SB_c and SB_a . By applying this definition on the instance of Figure 4 we obtain:

- $SB(C1) = \{C1, C2, C3, C4, C5, C6, C7, C8\} \cup \{A1C1, A1C2, A2C2, A1C3, A1C4, A1C5, A1C6, A1C7, A1C8\}$,
- $SB(C2, C5) = \{C2, C5, C8\} \cup \{A1C2, A2C2, A1C5, A1C8\}$,
- $SB(C3, C4, C6, C7) = \{C3, C4, C6, C7, C8\} \cup \{A1C3, A1C4, A1C6, A1C7, A1C8\}$ and
- $SB(C8) = \{C8\} \cup \{A1C8\}$.

4. INTEROPERABILITY ASSESSMENT

Each semantic block obtained in section 3 identifies and emphasises the border of one sub-system model (of the studied IS) with its own core semantics. It focuses on “what is important” in the system without taking care on implementation artefacts. In fact, each semantic block embeds the whole minimal mandatory semantics for the associated subsystem to achieve its mission. Thus, this information is very useful when analysing, before putting in place a collaboration, if the collaborating information system can interoperate with a given information system or not.

Let us consider the following scenario. A given enterprise needs to collaborate with another enterprise through the information exchanges from its information system (IS_1) to the information system (IS_2) of the other enterprise. The contract that will be signed between the two enterprises mention what information the second enterprise must receive from the first enterprise. Based on the information exchange, IS_2 must continue to achieve its mission. The first enterprise would like to know if it can respect this contract or not (and thus if IS_1 can interoperate with IS_2 or not).

The information, which IS_1 can send to IS_2 , is conceptualised by C_{IS_1} . Let $C_{IS_2}^*$ be the conceptualisation of the information needed by IS_2 determined by the contract ($C_{IS_2}^* \subseteq C_{IS_2}$). For each $c_i^2 \in C_{IS_2}^*$, $SB(c_i^2)$ determines the minimal mandatory semantics associated to c_i^2 . Without, this knowledge, the application cannot achieve its mission anymore. Thus, it's necessary to evaluate, for each $c_i^2 \in C_{IS_2}^*$, if the semantics embedded by $SB(c_i^2)$ exists in C_{IS_1} or not. For doing that, we first define a semantic relationship that identifies if one concept from one IS is semantically subsumed by a set of concepts from the other IS. And then, we verify if this semantic relationship exists for each concept from $SB(c_i^2)$ or not.

4.1 Semantic relationship

If there exists a p-tuple $(c_{q_1}^1, \dots, c_{q_p}^1)$ of concepts in C_{IS_1} such that the concept $c_j^2 \in SB(c_i^2)$ is semantically subsumed by $(c_{q_1}^1, \dots, c_{q_p}^1)$, then the semantic relationship $R_{c_j}^2$ is identified and formally defined as $R_{c_j}^2 = \langle c_j^2, (c_{q_1}^1, \dots, c_{q_p}^1) \rangle$. This semantic relationship can be read as: c_j^2 is semantically subsumed by $(c_{q_1}^1, \dots, c_{q_p}^1)$.

The paper assume that an expert is able to identify if such a semantic relationship exists or not for each $c_j^2 \in SB(c_i^2)$. To achieve this, the expert can take profits from existing approaches (schema matching (Evermann, 2009), (Rahm and Bernstein, 2001)) and morphism (Goguen and Burstall, 1984), ontology mapping (Kalfoglou and Schorlemmer, 2003), (Choi, Song and Han, 2006)). The goal of this paper is not to develop such an approach.

Based on the semantics relationships, identified by the expert, the set of the retrieved semantic relationships denoted as R_c^2 could be formally defined as:

$$R_c^2 = \{ R_{c_j}^2 = \langle c_j^2, (c_{q_1}^1, \dots, c_{q_p}^1) \rangle \mid c_j^2 \in SB(c_i^2) \wedge (c_{q_1}^1, \dots, c_{q_p}^1) \in P(C_{IS_1}) \}$$

where $P(C_{IS_1})$ is the set of all p-tuples $(c_{q_1}^1, \dots, c_{q_p}^1)$ that can be made using the elements of C_{IS_1} , with $p \in [1, s_1]$ and $s_1 = |C_{IS_1}|$.

We also define R_c^{2e} , that is a subset of R_c^2 , restricted to the relationships $R_{c_j}^2 = \langle c_j^2, (c_{q_1}^1, \dots, c_{q_p}^1) \rangle$ where $(c_{q_1}^1, \dots, c_{q_p}^1)$ contains only mandatory concepts. Thus, R_c^{2e} can be formally denoted as:

$$R_c^{2e} = \{ R_{c_j}^2 = \langle c_j^2, (c_{q_1}^1, \dots, c_{q_p}^1) \rangle \mid c_j^2 \in SB(c_i^2) \wedge (c_{q_1}^1, \dots, c_{q_p}^1) \in P(MC_{IS_1}) \}$$

Two cases can then be considered, for interpreting the semantic relationship $R_{c_j}^2$:

- i. If there exists at least one $c_j^2 \in C_{IS_2}^*$ such that the semantic relationship $R_{c_j}^2$ cannot be identified, there exists a semantic loss when IS_1 interoperates with IS_2 .
- ii. If there exists a semantic relationship $R_{c_j}^2$ between c_j^2 and a unique or several concepts coming from C_{IS_1} . There is no semantic loss concerning c_j^2 .

4.2 Interoperability assessment

For evaluating the degree of interoperability between IS_1 and the subsystem of IS_2 defined by $SB(c_i^2)$ and to characterise the semantic gap, different measures need to be formalised (Yahia, Aubry and Panetto, 2012). These measures should be mainly

based on the comparison between an expected and a real state. The expected state is denoted as $R_{c_{expected}}^2$, and represents the set of the expected semantic relationships needed to ensure a full semantic interoperation: i.e. there exists a semantic relationship for each concept from $SB(c_i^2)$ ($|R_{c_{expected}}^2| = |SB(c_i^2)|$). The real state corresponds to R_c^2 that is the set of the semantic relationships identified by the expert.

The paper proposes two measures to evaluate respectively a potential and an effective interoperability.

4.2.1 Maximal Potential Interoperability measure

The potential interoperability measure between two information systems is computed by taking into consideration all the retrieved semantic relationships in R_c^2 .

To measure the potential interoperability, $\nu_{1 \rightarrow 2}$ (%) is defined as follows:

$$\nu_{1 \rightarrow 2} = \frac{|R_c^2|}{|R_{c_{expected}}^2|} = \frac{|R_c^2|}{|SB(c_i^2)|}$$

$\nu_{1 \rightarrow 2}$ measures the ratio of those semantic relationships over the total number of the expected semantic relationships to fully interoperate. When considering the entire semantics (mandatory and non-mandatory) provided by the concepts from C_{IS_1} to define the interoperability relationships R_c^2 , all these found relationships are expected to use **available** concepts, that means that all of their instances are filled by the user (even the non-mandatory concepts from NMC_{IS_1}). That is why this interoperability is qualified as the **maximal potential interoperability**. In fact, it is not possible to guaranty that the non-mandatory concepts are completely instantiated by the users of the information system IS_1 (as they are non-mandatory). The value $\nu_{1 \rightarrow 2}$ increases when the semantic gap decreases. The closer to 100% $\nu_{1 \rightarrow 2}$ is, the smaller the semantic gap is.

In fact, it is obvious achieving $\nu_{1 \rightarrow 2}$ equal to 100% when the expert retrieves the whole set of expected semantic relationships. This corresponds to the full interoperability. If it is less than 100%, it corresponds to a partial interoperability.

The **semantic gap** is valued by $1 - \nu_{1 \rightarrow 2}$. Moreover, the missing semantics is given by the semantics of the concepts c_j^2 from $SB(c_i^2)$ for which $R_{c_j}^2$ cannot be computed: these semantic relationships are identified in $R_{c_{expected}}^2 \setminus R_c^2$.

Therefore, $\nu_{1 \rightarrow 2}$ alone is not enough to evaluate the interoperability when considering the effectiveness of the returned semantic relationships. In fact, only the retrieved semantic relationships derived through the mandatory concepts from MC_{IS_1} are guaranteed to be *effective*: this corresponds to R_c^{2e} . Thus, we need to measure the **effectiveness** of the retrieved semantic relationships. Thus, effectiveness measure $\varepsilon_{1 \rightarrow 2}$ (%) is formalised as follows:

$$\varepsilon_{1 \rightarrow 2} = \frac{|R_c^{2e}|}{|R_c^2|}$$

The value $(1 - \varepsilon_{1 \rightarrow 2})$ represents the **risk** to retrieve non effective semantic relationships (because involving non-mandatory concepts).

$\nu_{1 \rightarrow 2}$ and $\varepsilon_{1 \rightarrow 2}$ as they are defined represent respectively the **potential interoperability** and the **effectiveness measures**.

Typically, IS_1 may interoperate with $SB(c_i^2)$ with a potential that equals to $\nu_{1 \rightarrow 2}$ (%), but with a related effectiveness that equals to $\varepsilon_{1 \rightarrow 2}$ (%). That means that among the semantic relationships used to compute R_c^2 , only $\varepsilon_{1 \rightarrow 2}$ (%) of them are guaranteed to be effective (without any risk of non-availability of the information).

Besides, to promote an **effective interoperability evaluation**, it is important to consider the interoperation effectiveness property to drive new measures. Indeed, in the following section, the effective interoperability measure is formalised.

4.2.2 Minimal Effective Interoperability measure

The effective interoperability evaluation for two information systems could be computed by taking into consideration the minimal necessary amount of semantics that represents the effective mandatory knowledge and that is given through R_c^{2e} . This allowed us to compute measures such as $\nu_{1 \rightarrow 2}^e$ (%) and $\varepsilon_{1 \rightarrow 2}^e$ (%) to evaluate **qualitatively** the interoperability.

For this, we define:

$$\nu_{1 \rightarrow 2}^e = \frac{|R_c^{2e}|}{|R_{c_{expected}}^2|} = \frac{|R_c^2|}{|R_{c_{expected}}^2|} \times \frac{|R_c^{2e}|}{|R_c^2|} = \nu_{1 \rightarrow 2} \times \varepsilon_{1 \rightarrow 2}$$

$\nu_{1 \rightarrow 2}^e$ measures the ratio of the returned effective semantic relationships over the total number of the expected semantic relationships to fully interoperate.

When considering all the effective returned semantic relationships, the **minimal effective interoperability** between the studied information systems is guaranteed, but with an effectiveness equal to 100%. The associated effectiveness is defined as follows:

$$\varepsilon_{1 \rightarrow 2}^e = \frac{|R_c^{2e}|}{|R_c^{2e}|} = 100\%$$

We can note that $\varepsilon_{1 \rightarrow 2}^e$ is always equal to 100%. It is not a surprise as $\nu_{1 \rightarrow 2}^e$ measures an effective interoperability.

4.2.3 Interoperability measures map

When evaluating the interoperability (potential and effective), both measures could be computed: an interoperability map can be pictured as in Figure 8. The points $MPI = (\nu_{1 \rightarrow 2}, \varepsilon_{1 \rightarrow 2})$ and $MEI = (\nu_{1 \rightarrow 2}^e, \varepsilon_{1 \rightarrow 2}^e)$ represent respectively the **maximal potential interoperability** and the **minimal effective interoperability**.

Starting from these points, the different improvements that could take place can follow two phases:

1. Firstly, reaching the *maximal intrinsic interoperability*: this corresponds to the point $MII = (v_{1 \rightarrow 2}, \epsilon_{1 \rightarrow 2}^e)$ in Figure 8. This effort could be conducted by transforming all the non-mandatory concepts (taking part into R_C^2) into mandatory ones.
2. Secondly, reaching the *full interoperability* (100% of interoperability with 100% of effectiveness). This corresponds to a substantial effort as there is a need to add the missing semantics to reach (100%, 100%). This missing semantics is identified by the semantics of the concepts taking part into $R_{C_{expected}}^2 \setminus R_C^2$.

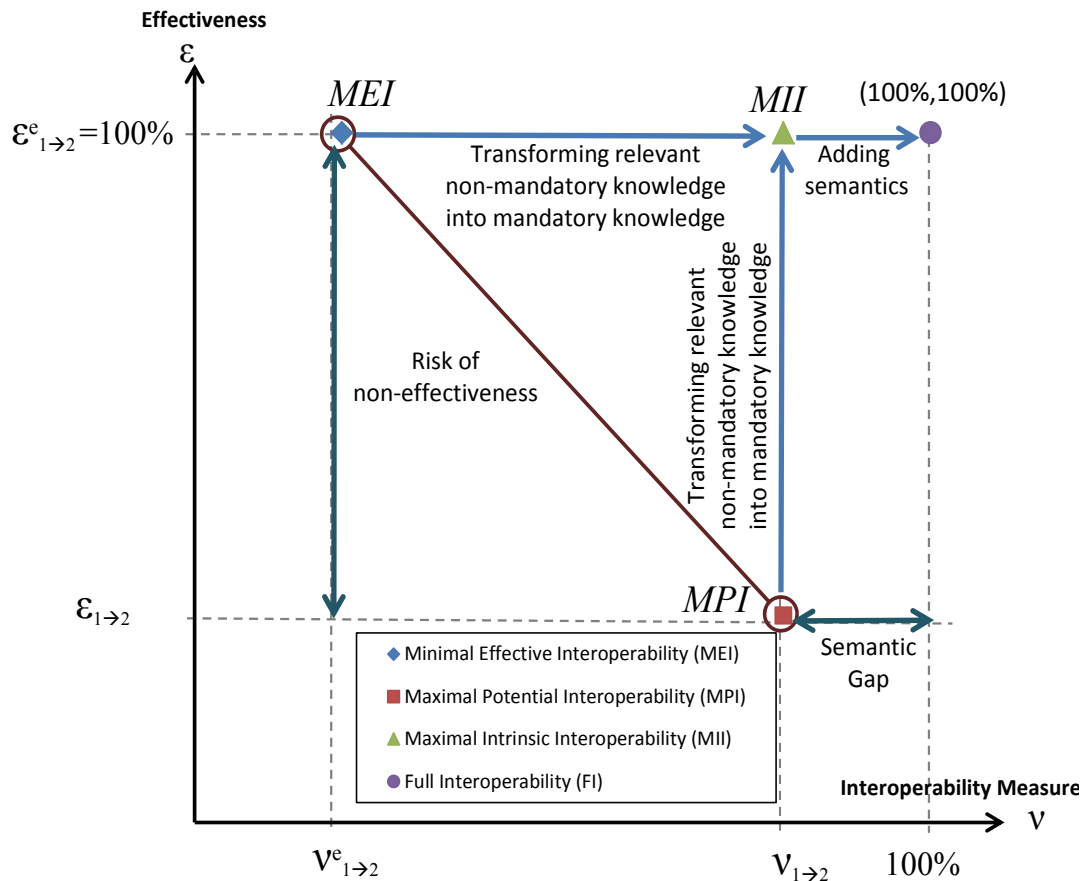


Figure 8 - The Interoperability Measures Map

5. CASE STUDY: RAW MATERIAL PURCHASE

The proposed approach of ISs semantics enactment and interoperability assessment is illustrated by the following case study that consists of two ISs dealing with B2M interoperability requirement. These ISs have been provided by a local technical centre: the AIPL-PRIMECA⁵ (Atelier Inter-établissements de Productique Lorraine) in which the ERP Sage X3 application is cooperating with the MES Flexnet application in order to insure the manufacturing of a certain family of products. In such

⁵ AIPL-PRIMECA, www.aip-primeca.net/lorraine/

industrial large scale Enterprise Information Systems, applications comprise a multitude of tables and relations. Flexnet (a MES application) has around 800 tables with 300 relations. Once its model are conceptualised, we get about 600 concepts and 500 associations. SAGE X3 has around 1600 tables with 900 relations, and when it is conceptualised, 1200 concepts and 1000 associations can be highlighted.

Actually, a specified process has been chosen to support our research; it consists of the Raw Material Purchase. For instance, Figure 9 represents the conceptual model for the purchase order process related to Flexnet.

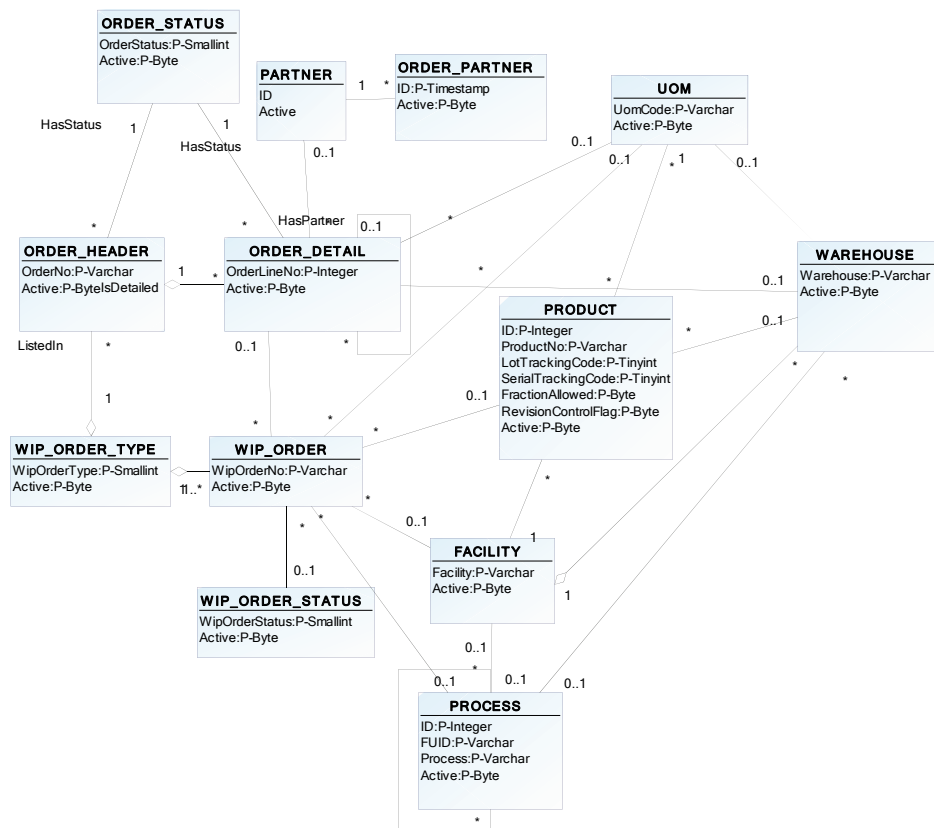


Figure 9 – Conceptual model for the purchase order process from Flexnet

5.1 Semantics structuring of Flexnet MES model

Manufacturing Execution Systems (MES) are information technology systems that manage manufacturing operations in factories. Actually, a specific process implemented in Flexnet application has been chosen to support our validation process; it consists of the purchase order process. Figure 10 represents the fact-oriented model of this process. Note that, in this model, classes with capital letters represent the non-lexical concepts. In order to compute the semantic blocks for structuring the model semantics of this process, the procedure presented in section 3.6 is applied.

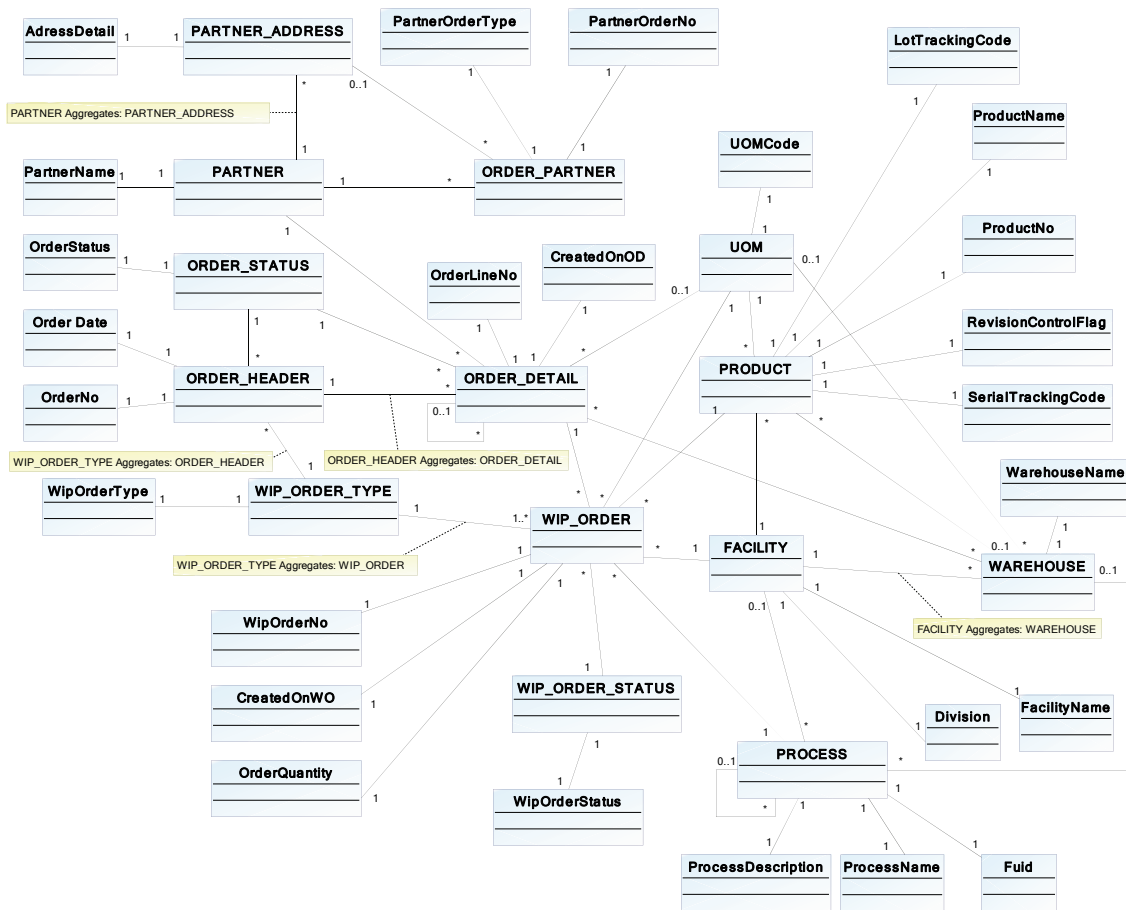


Figure 10 – Fact-oriented model of the purchase order process in Flexnet application

5.1.1 Building the associated semantic-dependency graph

The semantic-dependency graph related to the conceptual model of Flexnet is given on Figure 11.

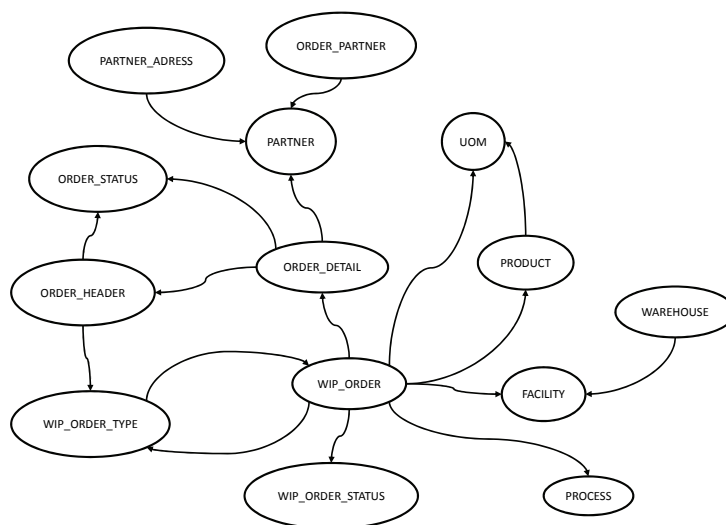


Figure 11 - Semantic-dependency graph related to the conceptual model of Flexnet MES

5.1.2 Building the graph of the strongly connected components based on the semantic-dependency graph

The graph of the strongly connected components related to the semantic-dependency graph of Flexnet MES is given on Figure 12. We can note that only one merged node has been built (namely SCC1) representing the strongly connected components of the concepts: WIP_ORDER, WIP_ORDER_TYPE, ORDER_DETAIL and ORDER_HEADER. All the other strongly connected components consist of only one concept.

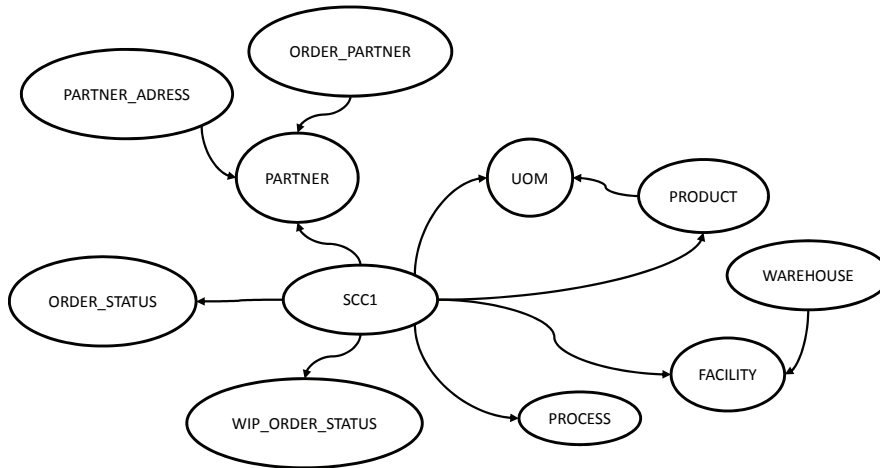


Figure 12 - Graph of the strongly connected components related to the semantic-dependency graph of Flexnet MES

5.1.3 Computing the semantic blocks SB_c associated with each strongly connected component

Table 1 lists the different semantic blocks SB_c related to Flexnet application after applying algorithm 1 (BuildSemBlocks) to the graph of the strongly connected components on Figure 12.

SB_c	<i>Concepts</i>
$SB_c(\text{WAREHOUSE})$	WAREHOUSE, FACILITY
$SB_c(\text{ORDER_PARTNER})$	ORDER_PARTNER, PARTNER
$SB_c(\text{PARTNER_ADDRESS})$	PARTNER_ADDRESS, PARTNER
$SB_c(\text{PARTNER})$	PARTNER
$SB_c(\text{SCC1}) = SB_c \left(\begin{array}{c} \text{WIP_ORDER,} \\ \text{ORDER_DETAIL,} \\ \text{ORDER_HEADER,} \\ \text{WIP_ORDER_TYPE} \end{array} \right)$	WIP_ORDER, WIP_ORDER_TYPE, ORDER_DETAIL, ORDER_HEADER, WIP_ORDER_STATUS, PROCESS, FACILITY, PRODUCT, UOM, ORDER_STATUS, PARTNER
$SB_c(\text{PROCESS})$	PROCESS
$SB_c(\text{PRODUCT})$	PRODUCT, UOM, FACILITY
$SB_c(\text{UOM})$	UOM
$SB_c(\text{WIP_ORDER_STATUS})$	WIP_ORDER_STATUS

$SB_c(\text{FACILITY})$	FACILITY
$SB_c(\text{ORDER_STATUS})$	ORDER_STATUS

Table 1 - Semantic Blocks (SB_c) of Flexnet MES

5.1.4 Computing, for each SB_c , the related semantic block SB_a

Table 2 lists the different semantic blocks SB_a related to Flexnet application.

SB_a	<i>Concepts</i>
$SB_a(\text{WAREHOUSE})$	WarehouseName, FacilityName, Division
$SB_a(\text{ORDER_PARTNER})$	PartnerOrderNo, PartnerOrderType, PartnerName
$SB_a(\text{PARTNER_ADDRESS})$	AdressDetail, PartnerName
$SB_a(\text{PARTNER})$	PartnerName
$SB_a \left(\begin{array}{l} \text{WIP_ORDER,} \\ \text{ORDER_DETAIL,} \\ \text{ORDER_HEADER,} \\ \text{WIP_ORDER_TYPE} \end{array} \right)$	WipOrderNo, CreatedOnWO, OrderQuantity, WipOrderType , OrderLineNo, CreatedOnOD, OrderDate, OrderNo, WipOrderStatus, ProcessName, ProcessDescription, Fuid, FacilityName, Division, LotTrackingCode, ProductName, ProductNo, RevisionControlFlag, SerialTrackingCode, UOMCode, OrderStatus, PartnerName
$SB_a(\text{PROCESS})$	ProcessName, ProcessDescription, Fuid
$SB_a(\text{PRODUCT})$	LotTrackingCode, ProductName, ProductNo, RevisionControlFlag, SerialTrackingCode , UOMCode, FacilityName, Division
$SB_a(\text{UOM})$	UOMCode
$SB_a(\text{WIP_ORDER_STATUS})$	WipOrderStatus
$SB_a(\text{FACILITY})$	FacilityName, Division
$SB_a(\text{ORDER_STATUS})$	OrderStatus

Table 2 - Semantic Blocks (SB_a) of Flexnet MES

5.1.5 Computing $SB = SB_c \cup SB_a$

Table 3 lists the different semantic blocks SB related to Flexnet application.

SB	<i>Concepts</i>
$SB(\text{WAREHOUSE})$	WAREHOUSE, WarehouseName, FACILITY, FacilityName, Division
$SB(\text{ORDER_PARTNER})$	ORDER_PARTNER, PartnerOrderNo, PartnerOrderType, PARTNER, PartnerName
$SB(\text{PARTNER_ADDRESS})$	PARTNER_ADDRESS, AdressDetail, PARTNER, PartnerName

<i>SB</i> (PARTNER)	PARTNER, PartnerName
<i>SB</i> (WIP_ORDER, ORDER_DETAIL, ORDER_HEADER, WIP_ORDER_TYPE)	WIP_ORDER, WipOrderNo, CreatedOnWO, OrderQuantity, WIP_ORDER_TYPE, WipOrderType, ORDER_DETAIL, OrderLineNo, CreatedOnOD, ORDER_HEADER, OrderDate, OrderNo, WIP_ORDER_STATUS, WipOrderStatus, PROCESS, ProcessId, ProcessDescription, Fuid, FACILITY, FacilityName, Division, PRODUCT, LotTrackingCode, ProductName, ProductNo, RevisionControlFlag, SerialTrackingCode, UOM, UOMCode, ORDER_STATUS, OrderStatus, PARTNER, PartnerName
<i>SB</i> (PROCESS)	PROCESS, ProcessName, ProcessDescription, Fuid
<i>SB</i> (PRODUCT)	PRODUCT, LotTrackingCode, ProductName, ProductNo, RevisionControlFlag, SerialTrackingCode, UOM, UOMCode, FACILITY, FacilityName, Division
<i>SB</i> (UOM)	UOM, UOMCode
<i>SB</i> (WIP_ORDER_STATUS)	WIP_ORDER_STATUS, WipOrderStatus
<i>SB</i> (FACILITY)	FACILITY, FacilityName, Division
<i>SB</i> (ORDER_STATUS)	ORDER_STATUS, OrderStatus

Table 3 - Semantic Blocks (SB) of Flexnet MES

The procedure presented in section 3.6 has been implemented in the MEGA Suite environment. MEGA Suite supports UML notations and allows building our own meta-model based on its ad-hoc MOF6 meta-model. The meta-model presented on Figure 6 has been implemented in the MEGA Suite. In this implementation, the semantic block is conceptualised as a UML package and the lexical and non-lexical concepts are conceptualised as UML classes. The procedure presented in section 3.6 has been implemented taking advantage of MEGA programming facilities.

Figure 13 provides a model representing all the semantic blocks related to the Flexnet purchase order process and their inclusion relationships. Figure 14 provides the conceptual model associated to the semantic block *SB*(PRODUCT), and including all the mandatory concepts required to obtain the full semantics for the concept PRODUCT.

⁶ OMG's MetaObject Facility: <http://www.omg.org/mof/>

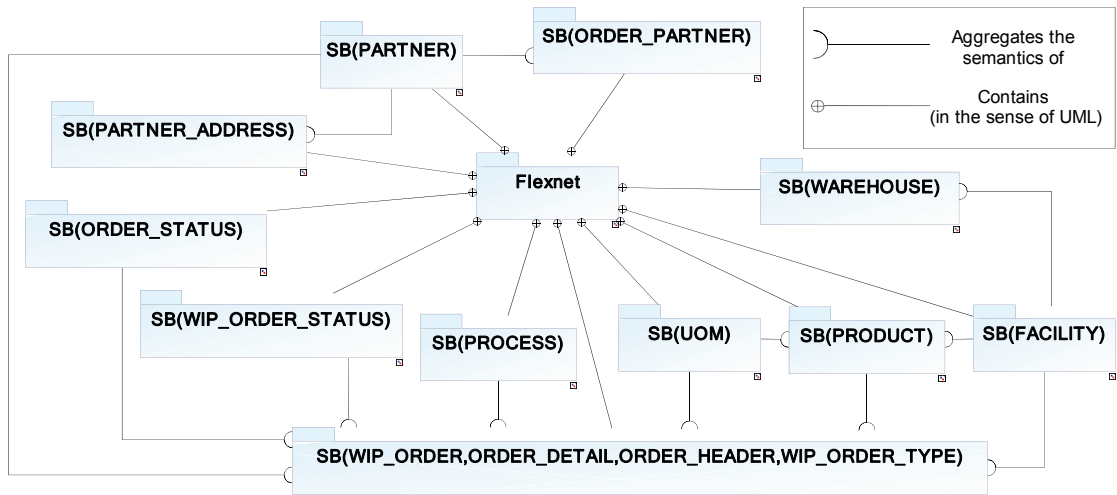


Figure 13 - The computed semantic blocks related to Flexnet MES

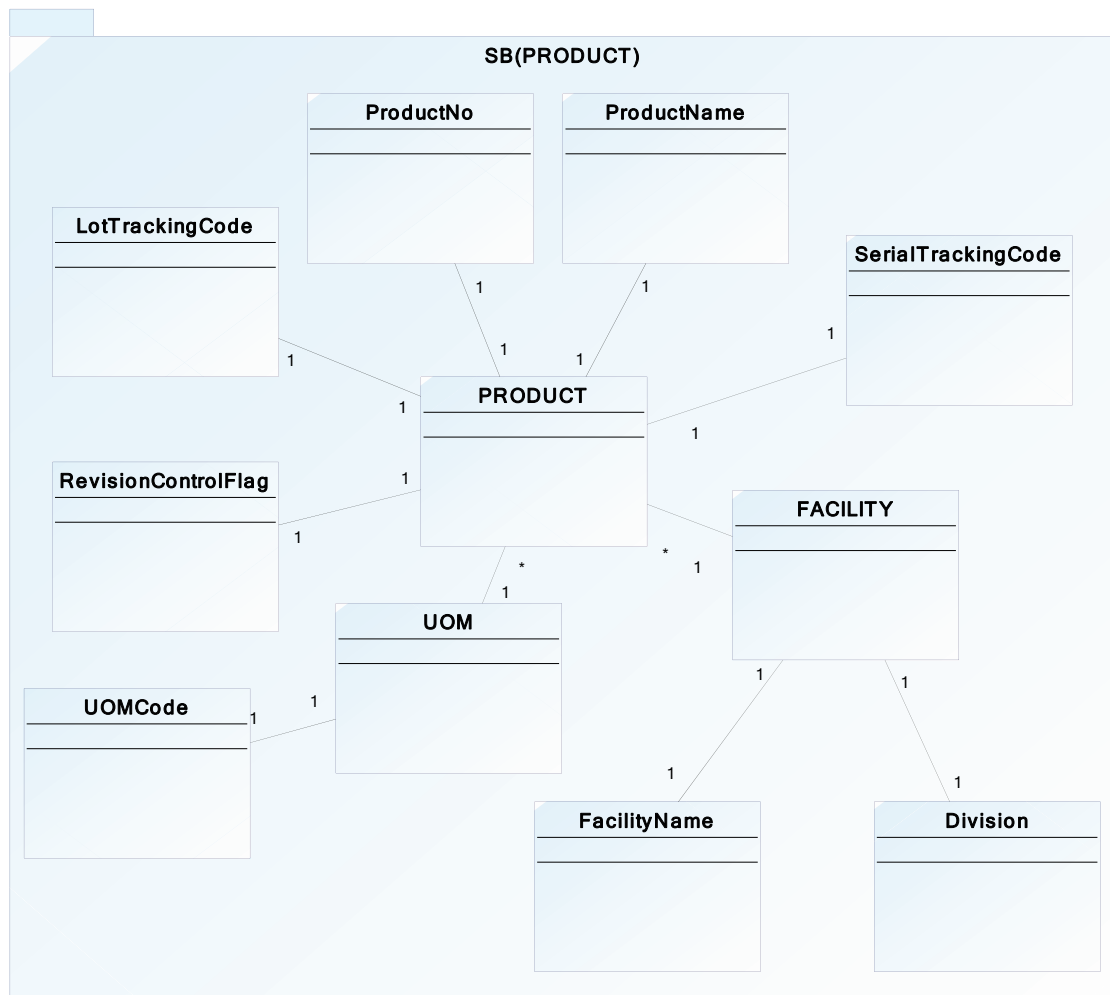


Figure 14 - The conceptual model associated to the semantic block SB(PRODUCT)

5.2 Need of interoperation and interoperability assessment

When considering the long term planning, the ERP computes, for a given period, its needs in term of raw materials and then launches some purchase orders. Hence, those purchase orders have to be exported from the ERP to the MES that have to bring backward the ERP with the stock state and the purchase order status. Here, the need of interoperation between the two ISS clearly appears. We propose here to evaluate the interoperability levels from the ERP to MES regarding the concept $WIP_ORDER (C_{IS_2}^* = \{WIP_ORDER\})$ in the MES. That means that the semantic need for the MES coming from the ERP is defined by $SB(WIP_ORDER, ORDER_DETAIL, ORDER_HEADER, WIP_ORDER_TYPE)$ (see Table 3).

According to section 4.2, it is necessary, at first, to identify, if it exists, the semantic relationship for each concept in $SB(WIP_ORDER, ORDER_DETAIL, ORDER_HEADER, WIP_ORDER_TYPE)$.

The Table 4 gives the semantic relationships identified by the experts concerning each concept of $SB(WIP_ORDER, ORDER_DETAIL, ORDER_HEADER, WIP_ORDER_TYPE)$ and relatively to the concepts embedded in the conceptual model of Sage X3 given on Figure 15 (the mandatory concepts are represented in bold).

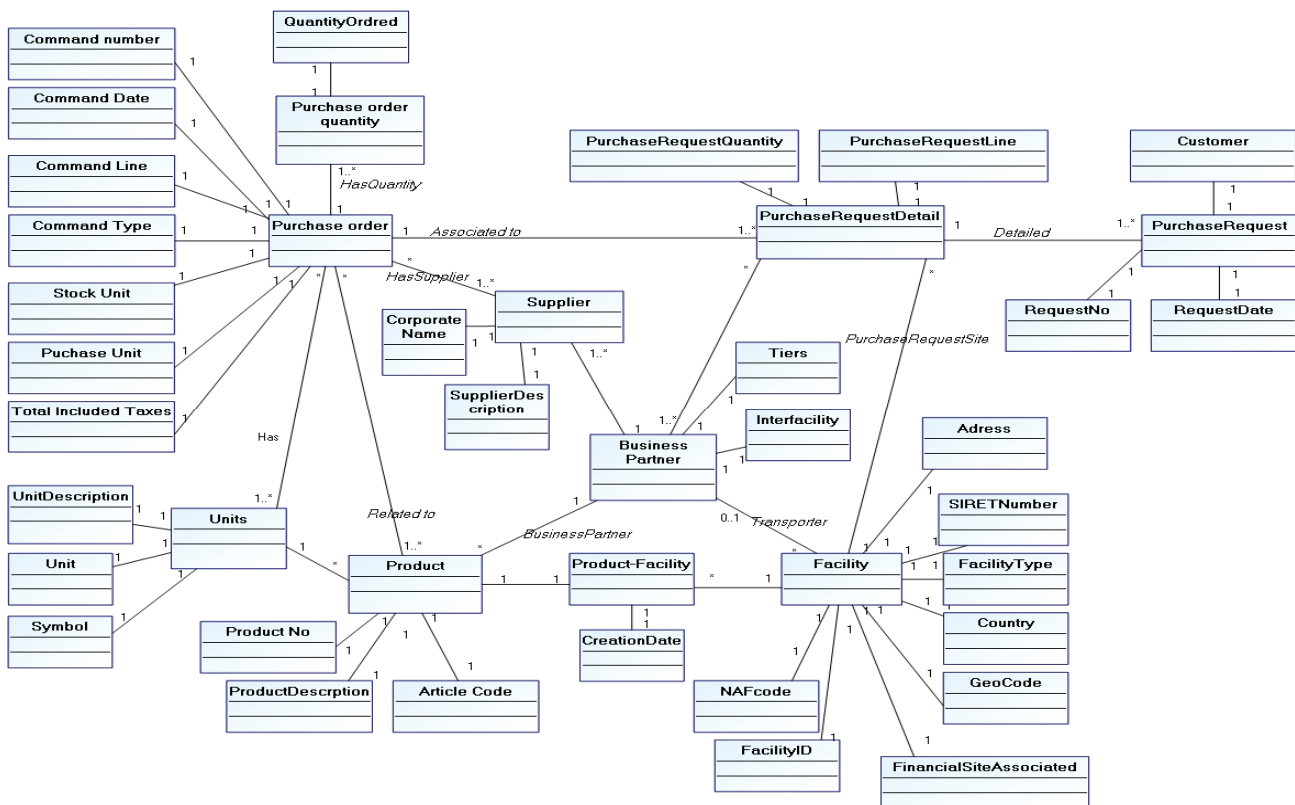


Figure 15 – Fact-oriented model of the purchase order process in SAGE X3 application

$SB \left(\begin{matrix} \text{WIP_ORDER,} \\ \text{ORDER_DETAIL,} \\ \text{ORDER_HEADER,} \\ \text{WIP_ORDER_TYPE} \end{matrix} \right)$	R_c^2	R_c^{2e}
WIP_ORDER	$\langle \text{WIP_ORDER, (PurchaseRequestDetail, PurchaseRequest)} \rangle$	$\langle \text{WIP_ORDER, (PurchaseRequestDetail, PurchaseRequest)} \rangle$
WipOrderNo	$\langle \text{WipOrderNo, (RequestNo)} \rangle$	$\langle \text{WipOrderNo, (RequestNo)} \rangle$
CreatedOnWO	$\langle \text{CreatedOnWO, (RequestDate)} \rangle$	$\langle \text{CreatedOnWO, (RequestDate)} \rangle$
OrderQuantity	$\langle \text{OrderQuantity, (PurchaseOrder)} \rangle$	$\langle \text{OrderQuantity, (PurchaseOrder)} \rangle$
WIP_ORDER_TYPE	\emptyset	\emptyset
WipOrderType	\emptyset	\emptyset
ORDER_DETAIL	$\langle \text{ORDER – DETAIL, (PurchaseRequestQuantity)} \rangle$	$\langle \text{ORDER – DETAIL, (PurchaseRequestQuantity)} \rangle$
OrderLineNo	$\langle \text{OrderLineNo, (CommandLineNo)} \rangle$	$\langle \text{OrderLineNo, (CommandLineNo)} \rangle$
CreatedOnOD	\emptyset	\emptyset
ORDER_HEADER	$\langle \text{ORDER – HEADER, (PurchaseOrder)} \rangle$	$\langle \text{ORDER – HEADER, (PurchaseOrder)} \rangle$
OrderDate	$\langle \text{OrderDate, (CommandDate)} \rangle$	$\langle \text{OrderDate, (CommandDate)} \rangle$
OrderNo	$\langle \text{OrderNo, (CommandNumber)} \rangle$	$\langle \text{OrderNo, (CommandNumber)} \rangle$
WIP_ORDER_STATUS	\emptyset	\emptyset
WipOrderStatus	\emptyset	\emptyset
PROCESS	\emptyset	\emptyset
ProcessId	\emptyset	\emptyset
ProcessDescription	\emptyset	\emptyset
Fuid	\emptyset	\emptyset
FACILITY	$\langle \text{FACILITY, (Facility)} \rangle$	$\langle \text{FACILITY, (Facility)} \rangle$
FacilityName	\emptyset	\emptyset
Division	$\langle \text{Division, (FacilityType)} \rangle$	$\langle \text{Division, (FacilityType)} \rangle$
PRODUCT	$\langle \text{PRODUCT, (Product)} \rangle$	$\langle \text{PRODUCT, (Product)} \rangle$
LotTrackingCode	\emptyset	\emptyset
ProductName	\emptyset	\emptyset
ProductNo	$\langle \text{ProductNo, (Product No)} \rangle$	$\langle \text{ProductNo, (Product No)} \rangle$
RevisionControlFlag	\emptyset	\emptyset
SerialTrackingCode	\emptyset	\emptyset
UOM	$\langle \text{UOM, (Units)} \rangle$	$\langle \text{UOM, (Units)} \rangle$
UOMCode	$\langle \text{UOMcode, (Unit)} \rangle$	$\langle \text{UOMcode, (Unit)} \rangle$
ORDER_STATUS	$\langle \text{ORDER – STATUS, (PurchaseOrderStatus)} \rangle$	\emptyset
OrderStatus	$\langle \text{OrderStatus, (PurchaseOrderStatus)} \rangle$	\emptyset
PARTNER	$\langle \text{PARTNER, (BusinessPartner, Supplier)} \rangle$	$\langle \text{PARTNER, (BusinessPartner, Supplier)} \rangle$
PartnerName	\emptyset	\emptyset

Table 4 – Semantic relationships for $SB \left(\begin{matrix} \text{WIP_ORDER,} \\ \text{ORDER_DETAIL,} \\ \text{ORDER_HEADER,} \\ \text{WIP_ORDER_TYPE} \end{matrix} \right)$

The maximal potential interoperability $v_{1 \rightarrow 2}$ is computed as follows:

$$v_{1 \rightarrow 2} = \frac{|R_c^2|}{\left| SB \left(\begin{array}{c} \text{WIP_ORDER,} \\ \text{ORDER_DETAIL,} \\ \text{ORDER_HEADER,} \\ \text{WIP_ORDER_TYPE} \end{array} \right) \right|} = \frac{18}{33} = 55\%$$

Moreover its associated effectiveness $\epsilon_{1 \rightarrow 2}$ is computed:

$$\epsilon_{1 \rightarrow 2} = \frac{|R_c^{2e}|}{|R_c^2|} = \frac{16}{18} = 89\%$$

The semantic gap is valued by 45% and formally defined by all the concept in $SB(\text{WIP_ORDER}, \text{ORDER_DETAIL}, \text{ORDER_HEADER}, \text{WIP_ORDER_TYPE})$.

The minimal effective interoperability $v_{1 \rightarrow 2}^e$ is computed:

$$v_{1 \rightarrow 2}^e = v_{1 \rightarrow 2} \times \epsilon_{1 \rightarrow 2} = 48\%$$

Finally, the interoperability measures map is built on the figure 16 bellow.

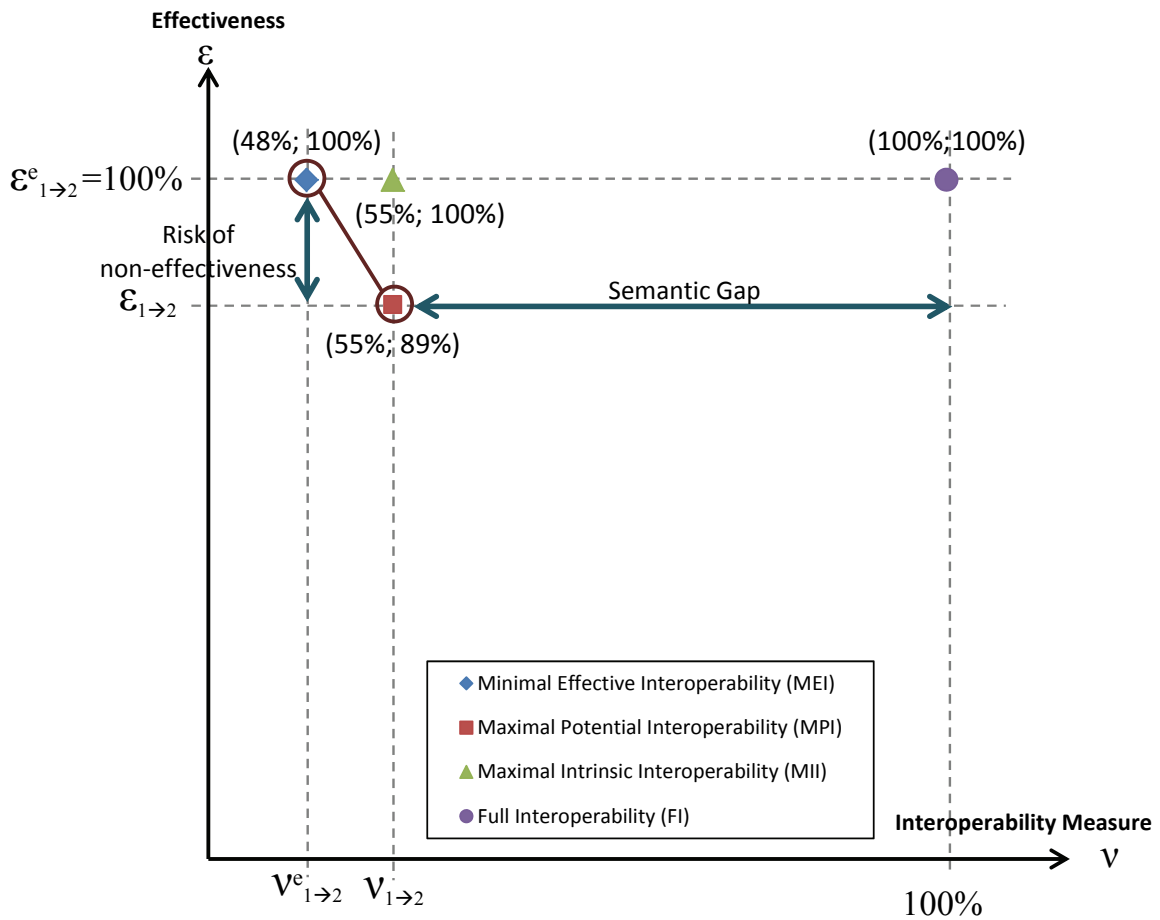


Figure 16 – Interoperability Measures Map from SAGE X3 ERP to FlexNet MES

6. DISCUSSION

This paper tackled the problem of information systems interoperability through a deep analysis of the semantics structure embedded into the related conceptual models and through a qualitative and quantitative assessment of the semantic interoperability between those systems. The previous case study demonstrated that, even if the technical complexity is high due to the large amount of data, the major issue is still the analysis of the hidden semantics of the concepts and the identification of semantic relationships between these concepts or between aggregates of concepts denoted semantic blocks. Each semantic block identifies and emphasises the border of one sub-system model with its own core semantics. It focuses on “what is important” in the system without taking care on implementation artefacts. The proposed approach is applying graph theory on existing conceptual models in order to compute automatically semantic blocks and then enacting the minimal and necessary semantics for the information system to fulfil its own mission. In a context of interoperation, that means that these aggregates of semantics must be fully exchanged between interoperating systems. In most cases, the exchange is only partial and leads to a semantic gap. It is then essential to evaluate, a priori, this semantic gap and thus assessing the future interoperation process.

The Interoperability Measures Map (Figure 8) is a tool that addresses this issue, pointing out the risk of non-effectiveness of the interoperation process, and the lack of semantics in the source system that can cause malfunctioning in the destination system. Taking into account the identified semantic blocks, such assessment may also be recursively conducted when sub-systems interoperate. Indeed, if two systems partially interoperate, some of their sub-systems might be fully interoperable.

Based on this analysis, a strategy of improvement of the involved systems may be suggested to the Information Systems Architects (Yahia, Aubry and Panetto, 2012).

The proposed approach assumes three strong hypotheses:

1. A domain expert is in charge of the analysis of the conceptual models. He is responsible of enriching those models by taking into account many sources of tacit or hidden knowledge (such as implemented triggers, software behaviours, systems parameters, non-functional aspects of the related applications, non-formalised users' practices, etc.).
2. The semantic relationships are identified by a group of domain experts.
3. The enriched conceptual models and the semantic relationships are supposed to be semantically sound, complete, and context-coherent.

Removing these hypotheses is an open issue because of the strong involvement of skill-based knowledge that are generally not formalised. The research community working on knowledge extraction and discovery (Yangarber et al., 2000); (Polkowski,

2000)) is trying to automatize this process but there is no generic solution (Maimon and Rokach, 2005). At least, from our perspective, a domain expert will have to validate the resulting model.

7. CONCLUSION

Semantics enactment among *ISs* conceptual models is a critical issue in the context of Enterprise Models. Indeed, extracting these semantics has the advantage to ease the understanding and then the use of the exchanged information among heterogeneous information systems (In single or distributed Enterprises)

In this paper, the fact-oriented modelling is proposed to get a spread out representation for *ISs* conceptual models. This has allowed us to identify the Core and the extended Semantics for a given *IS*, respectively composed by the mandatory and non-mandatory concepts.

The originality of this paper lies on the elaboration of the semantic blocks for enacting Enterprise Models semantics embedded and, often hidden, in complex Information Systems models. Moreover, each semantic block identifies and emphasises the border of one sub-system model with its own core semantics.

Based on this identification, a set of metrics is proposed to assess the interoperability level between two information systems.

The semantics blocks identification and interoperability assessment is then illustrated in a use case based on existing B2M applications: the ERP Sage X3 and the MES Flexnet enterprise software applications, which have to interoperate in order to achieve a global process performance. This case study demonstrates the feasibility and the relevance of the proposed approach in a real enterprise context.

As discussed in the previous section, different research perspectives are pointed out. We are currently working on analysing the correlation that may exist between sub-systems interoperability and systems interoperability. Indeed, the semantic recursive structure imposed by the hierarchy of semantic blocs may affect the individual interoperability measures and the global one.

Moreover, the current Interoperability Measures Map is dealing with the relationship between two systems that must interoperate when, in a general case, the interoperation process involves more than two systems that cooperate all together to fulfil a global mission. We are investigating this more generic configuration that may have an impact on the proposed measuring system.

As previously discussed, removing the three strong hypotheses is still an open issue that will need complementary and multi-disciplinary knowledge.

REFERENCES

Benhlima, L. and Chiadmi, D. (2006) 'Vers l'interopérabilité des systèmes d'information hétérogènes', la revue électronique des technologies de l'information e-TI.

Boudjlida, N. and Panetto, H. (2008) 'Annotation of enterprise models for interoperability purposes ', Proceedings of the International Workshop on Advanced Information Systems for Enterprises IWAISE'08, Constantine, Algeria.

Castano, S. and De Antonellis, V. (1998) 'A framework for expressing semantic relationships between multiple information systems for cooperation', Information Systems, vol. 23, no. 3-4, May-June, pp. 253-277, Available: issn 0306-4379.

Chapurlat, V. (2007) Vérification et validation de modèles de systèmes complexe : application à la modélisation d'entreprise, Habilitation à Diriger les Recherches de Université de Montpellier 2.

Chiang, R.H.L., Barron, T.M. and Storey, V.C. (1994) 'Reverse engineering of relational databases: Extraction of an EER model from a relational database', Data and Knowledge Engineering, vol. 12, no. 2, pp. 107--142.

Choi, N., Song, I.Y. and Han, H. (2006) 'A survey on ontology mapping', ACM Sigmod Record, vol. 35, no. 3, pp. 34-41.

De Bo, J., Spyns, P. and Meersman, R. (2003) 'Creating a "DOGMAtic" multilingual ontology infrastructure to support a semantic portal', In proceedings of the OTM confederated international workshops HCI-SWWA, IPW, JTRES, WORM, WMS, and MRSM, Catania , Italie, 253-266.

Engelbart, D.C. (1962) Augmenting human intellect: A conceptual framework, Menlo Park, CA, USA: Stanford Research Institute.

Evermann, J. (2009) 'Theories of meaning in schema matching: An exploratory study', Information Systems, vol. 34, no. 1, pp. 28-44, Available: issn 0306-4379.

Fisher, D.A. (2006) An Emergent Perspective on Interoperation in Systems of Systems, Pittsburgh, PA: Carnegie Mellon University, Software Engineering Institute.

- Fonkam, M. and Gray, W. (1992) 'An Approach to Eliciting the Semantics of Relational Databases', In proceeding of CAiSE'92, Lecture Notes in Computer Science 593 Springer, Manchester, UK, 463-480.
- Gamma, E., Helm, R., Johnson, R. and Vlissides, J. (1995) Design patterns: elements of reusable object-oriented software, Addison-wesley Reading, MA.
- Goguen, J. and Burstall, R. (1984) 'Introducing institutions', Logics of Programs, pp. 221-256.
- Guarino, N. (1998) 'Formal Ontology and Information Systems', Proceedings of FOIS'98, Trento, Italy , 3-15.
- Halpin, T. (2006) 'Object-role modeling (ORM/NIAM)', Handbook on Architectures of Information Systems, pp. 81-103.
- Halpin, T. (2007) 'Fact-oriented modeling: Past, present and future', Conceptual Modelling in Information Systems Engineering, pp. 19-38.
- Kalfoglou, Y. and Schorlemmer, M. (2003) 'Ontology mapping: the state of the art', The knowledge engineering review, vol. 18, no. 1, pp. 1-31.
- Kavouras, M. (2003) 'A unified ontological framework for semantic integration', International Workshop on Next Generation Geospatial Information, Boston MA, USA.
- Lezoche, M., Panetto, H. and Aubry, A. (2011) 'Conceptualisation approach for cooperative information systems interoperability', In Proceedings of 13th International Conference on Enterprise Information Systems, Beijing, China. Extended version to appear in Lecture Notes in Business Information Processing LNBIP 102 (2012).
- Maimon, O.Z. and Rokach, L. (2005) Data mining and knowledge discovery handbook, Springer Science & Business edition.
- Nijssen, G.M. and Halpin, T.A. (1989) Conceptual Schema and Relational Database Design: a fact oriented approach, Upper Saddle River, NJ, USA: Prentice-Hall.
- Obrst, L. (2003) 'Ontologies for semantically interoperable systems', Proceedings of the 12 th international conference on Information and knowledge management, New Orleans, USA, 366--369.
- Park, J. and Ram, S. (2004) 'Information systems interoperability: What lies beneath?', ACM Transactions on Information Systems (TOIS, vol. 22, no. 4, pp. 595--632.

Polkowski, L. (2000) 'Rough set methods and applications: new developments in knowledge discovery in information systems'.

Rahm, E. and Bernstein, P.A. (2001) 'A survey of approaches to automatic schema matching', the VLDB Journal, vol. 10, no. 4, pp. 334-350.

Vernadat, F.B. (1996) Enterprise Modelling and Integration: principles and applications, Chapman and Hall.

Vernadat, F.B. (2009) 'Enterprise Integration and Interoperability', Springer Handbook of Automation, pp. 1529-1538.

Whitman, L., Santanu, D. and Panetto, H. (2006) 'An Enterprise Model of Interoperability ', In Proceeding of the 12th IFAC Symposium on Information Control Problems in Manufacturing (INCOM'2006, Saint Etienne, France.

Yahia, E., Aubry, A. and Panetto, H. (2012) 'Formal measures for semantic interoperability assessment in cooperative enterprise information systems', Computers in Industry. DOI : 10.1016/j.compind.2012.01.010.

Yangarber, R., Grishman, R., Tapanainen, P. and Huttunen, S. (2000) 'Automatic acquisition of domain knowledge for Information Extraction', Proceedings of the 18th conference on Computational linguistics (COLING '00), Stroudsburg, PA, USA, 940--946.

Zouggar N., Chen D., and Vallespir B. (2009) 'Semantic Enrichment of Enterprise Modelling'. In Proceedings of the International Conference on Interoperability for Enterprise Software and Applications (I-ESA'2009), Beijing, P.R. China