An approach based on ontologies and multi-agent systems to hide heterogeneity: Application to port information system

Mehdi ABID, Benayad NSIRI and Yassine SERHANE

***Abstract*—**Nowadays everyone is conscious of the technology's evolution that plays a key role in the development of information technology. The optimization of this complex system is based on a collaborative sharing of information and knowledge between each port information system. Unfortunately, many problems arise during a collaboration between different heterogeneous ports information systems (Incorrect interpretation, semantic ambiguity, linguistic differences, overlapping information… etc.), in addition these problems could be technical, semantic or structural.

To solve those previous problems we came up with a solution that is relying on a distributed architecture based on a mediator, Multi Agent Systems that rely on each port information systems, local ontologies, global ontology and adapters that provide a unified interface which hides heterogeneity of the associated source between all information systems.

***Keywords*—-** Mediation, interoperability, port information system, global ontology, local ontology, multi agent system, heterogeneous system.

I. Introduction

For over thirty years, technology's components development plays a key role in the development of different information systems, this tendency has resulted in a creation of a mosaic of heterogeneous ports information systems. During a collaboration between the different port information systems, the necessity of using the resources of all collective networks remains indispensable, in order to share risks, reduce cost and reduce delays.

Due to obstacles, the exchange and management of various information in a collaboration between the ports information systems becomes increasingly complicated: automated exchange, overlapping information and presentation of data in different languages etc. These linguistic differences can be one of the main causes that generates name conflicts issues [1] where each business or system may designate an entity according to its local area, these naming conflicts can be synonyms conflicts, homonyms conflicts or polysemy conflicts.

Synonyms conflict [2]: different words expressing the same information, for example the word « lift» (American English) and « elevator» (British English) are two different words referring to the same information.

Homonyms conflict [3]: same word different origin with the same pronunciation that have different meaning, eg. if two different systems exchange data, while the data exchanged is « left », the probable conflict is presented in a sentence « left » can refer to the past of the verb « leave » or the opposite of « right ».

Polysemy conflict [4]: same word of the same origin changes meaning depending on the context. For example the word « Get » can be used as « we get money » in this sense the word « Get » means the amount obtained, whereas if we use « we get it » in this case the term « get » means to understand.

The table below Fig 1 sums up the following types of conflicts (synonyms, homonyms and polysemy conflict) we took in consideration different meaning, pronunciation and spelling of each word or data that we can have while exchanging data between different information systems.

Conflict	Meanin g	Pronunciatio n	Spelling
Synonyms	Same	Different	Differen t
Homonyms	Different	Same	Same
Polysemy	Different	Different	Same

Figure 1: Types of conflicts

The data can then be presented differently from one system to another, this heterogeneity involves communication and

Mehdi ABID Laboratoire d'informatique et d'aide à la décision. Faculté des Sciences Ain Chock, Université Hassan II. Casablanca Maroc (phone: 00212-677-115323; e-mail: 90.abidmehdi@gmail.com).

Benayad NSIRI Laboratoire d'informatique et d'aide à la décision. Faculté des Sciences Ain Chock, Université Hassan II. Casablanca Maroc

(E-mail : nsiri2000@yahoo.fr)

Yassine SERHANE Laboratoire d'informatique et d'aide à la décision. Faculté des Sciences Ain Chock, Université Hassan II. Casablanca Maroc (phone: 00212-677-115323; e-mail: serhane.y@gmail.com).

knowledge sharing between these systems which have not been designed to collaborate with each other from the beginning. Ontologies have been commonly used to solve two important and related problems that occur in large organizations: information integration and knowledge representation [5]. In order to solve these problems several approaches have been proposed such as « mapping ontologies » [6] « unification ontologies » [7] for sharing and reusing various information between these heterogeneous information systems' databases.

In this article we address the problem of providing a solution to solve semantic, technical and structural conflicts using ontologies, and an architecture based on multi-agent systems. The reason behind the use of multi-agent systems technology is to facilitate semantic interoperability [8] after any interaction task between existing ontologies and then enrich the knowledge of all the different agents of information systems, the use of agents in our system is designed so that any agent can satisfy our goals exploiting its own resources and skills, it may even be interacting in an environment where other existing agents interact on shared knowledge.

II. Background

In the port sector several studies have been developed to solve any interoperability problems during a collaboration or data exchange between many firms, where these ports information systems development organizations work autonomously. These studies cover also different conceptions, constructions, and exploitation of data acquired from different information systems during a cooperation between these firms. This process requires high costs due to software maintenance of each port information systems, in the early 2000s different port firms wished to impose the use of a common centralized storage for information (Docker #, date in, date out, broker, commodity-type, palette, register, stevedore … etc.) between each heterogeneous information system. In addition, this approach involves several financial, strategic and technical issues, like misinterpretations due to linguistic differences between the information systems and the delay of data updates due to data redundancy (same data stored in different databases); It is necessary to study different knowledge sharing methods, which will help understand what information systems need to be used to reach the goals set [9]. Also with the technology evolution many approaches have been proposed to establish a common interface between each port information system and mask the heterogeneity of these systems during a collaboration.

The use of multi-agent systems in the technology sector is particularly interesting in the field of interoperability between heterogeneous information systems.

As shown in Fig 2 an example based on the use of Multi Agent Systems in the search of information between several registered users in different information systems.

M. Miranda and Al [10] propose a multi-agent system based on ontologies that provide data access to various information systems, this approach can receive messages from multiple

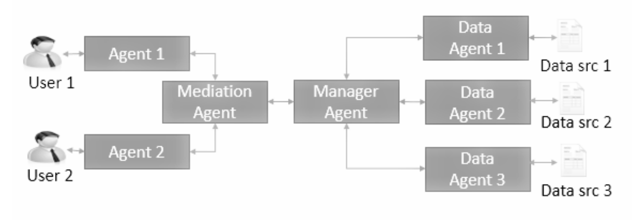

Figure 2: Search of information by MAS

clients, and specify the data that may be provided by a specific agent.

As shown in Fig 3, the system contains several parts based on agents that receive and process the various requests. Depending on the setting of multi-agent system, an agent acts only with the transmission of a specific message, messages are transmitted to agents in charge of processing the service requested.

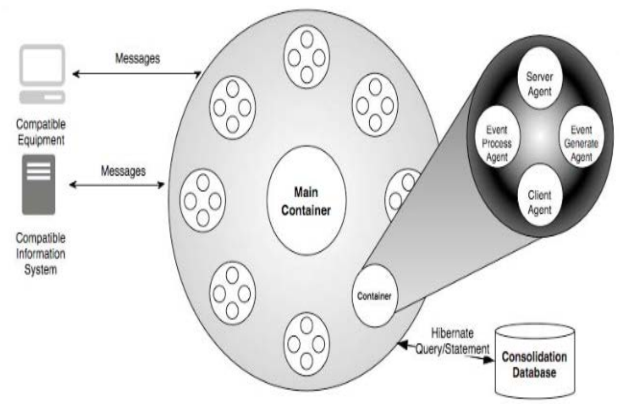

Figure 3: Communication using a multi-agent system

Thanks to the use of a global ontology which defines the communication between these agents, they become easier to find, so that the server agent can easily locate the specific agent to which the message will be sent systematically. This dynamic flow of information allows to have set of general processing agents as well as the ability to add and remove specific agents when needed for an interoperability service [10]. Due to unification and distribution of different data during a communication between these different systems, the exchange of data becomes necessary for optimizing existing resources within a firm to facilitate effective decision-making process.

III. Method

Our study is based on how to manage different requests between information systems during a collaboration, and the use of semantic concepts in the mediation process while exchanging data between these systems, it helps to rely on structured data such as OWL, in order to solve heterogeneity issues, however several misinterpretations can occur with different OWL versions because of linguistic differences [11] or interpretation differences that change from one domain to another.

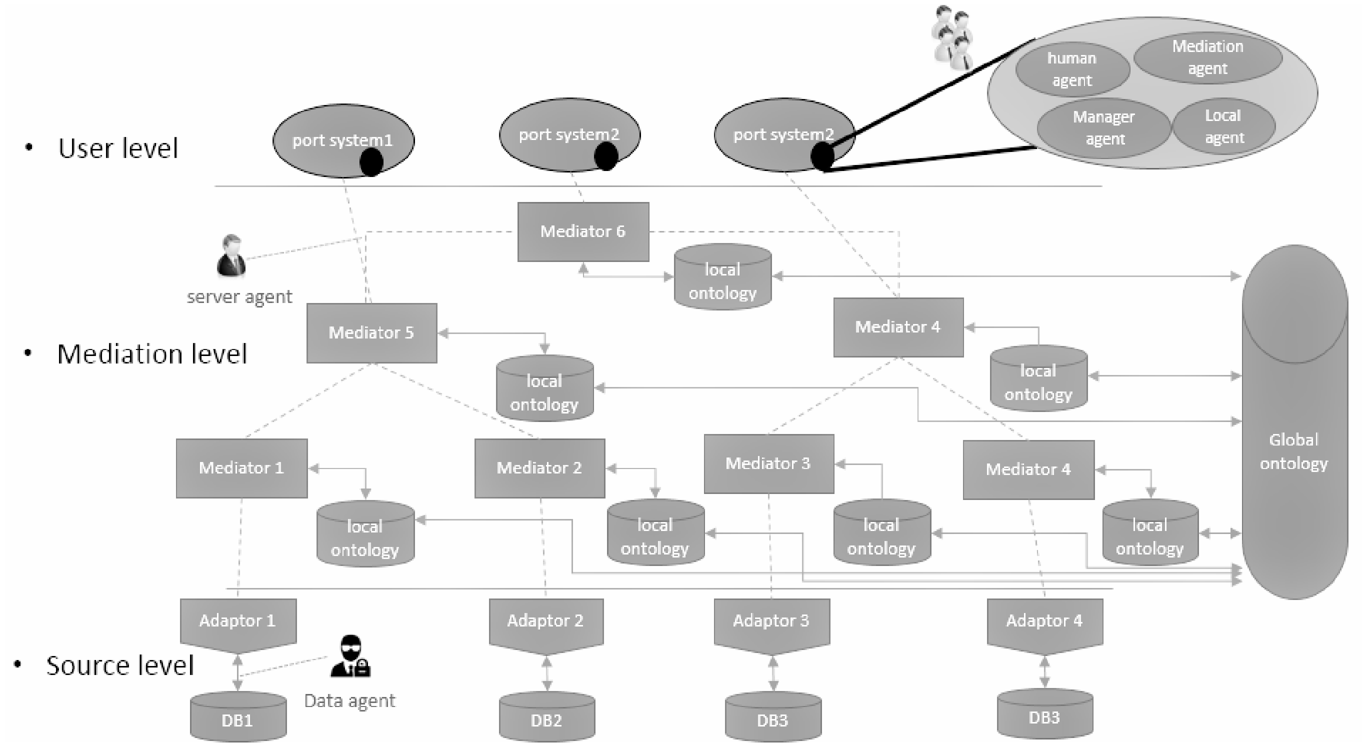

Figure 4: Our mediation architecture

To resolve semantic conflicts homogeneity, we decided to base on our study related to the use of multi-agents systems' concept, with 4 types of agents (local agent, mediation agent, manager agent and human agent) provided with an architecture relying on a concept of context mediation to specify values of the various structures that will allow us to grasp the semantics of the data.

Our approach is based on three levels: 1) user level 2) mediation level, 3) source level. Fig 4.

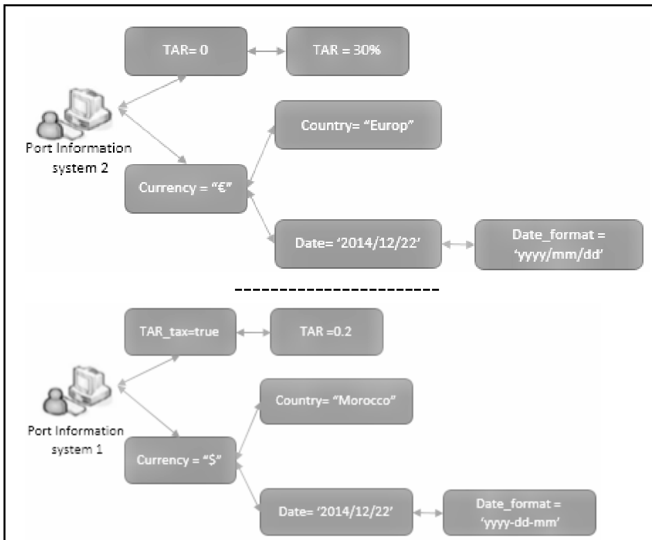

Figure 5: Model of two different structure in port information systems

The first level: User level, is composed of different port information systems equipped with an architecture that is based on multi-agent systems to manage the semantic heterogeneity between internal components and to ensure a good understanding of data exchanged during a collaboration with a heterogeneous port information system. This level contains four types of agents: local agent, mediation agent, manager agent, and human agent.

Once the local agent faces a problem of ambiguous term, it reports to the mediation agent by sending the list of incomprehensible data.

The mediation agent treats any technical issue of ambiguous data and ensure that they are consistent with the data structure by solving structural conflicts. Fig 5.

The Manager Agent treats all semantic problem by checking the semantic consistency of data, it also has information caching enabled which is a list of all the ambiguous data previously processed in order to transmit to the human agent only new cases where a misunderstanding cannot be resolved without human intervention. Fig 6.

The second level: Mediation level, consists of server agents, mediator agents, and a global ontology that includes local ontologies.

The mediator simplifies, abstracts, combines and describes the data [12] using to the server agent to process and transmit different data issued by the various requests of the ports information systems. Mediators can be structured in an organized and structured hierarchy between any information systems [1].

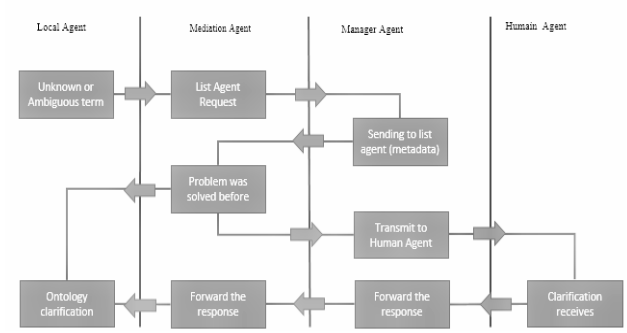

Figure 6: Prototype of mediation approach between agents

The local ontology contains a repository of different data knowledge of each mediator, and describes the context of the internal data of each port information system, it also contains generic knowledge of each port system, less abstract than a global ontology one in order to facilitate internal data exchanges between users of a particular information system.

The global ontology is the result of a merger between all the different local ontologies to ensure semantic interoperability and cooperation between them, and share various data information between any knowledge base of each port information system, this approach integrates ontology selection, mapping, and merging processes in order to minimize human mediation [13]. This ontology (global ontology) replaces the local ontology when it comes to data exchange between the different port systems, to exploit the internal resources and the collective resources [1].

The third level: Source level, is composed of data agents, adapters and various databases.

The adapter is positioned between the mediators and the database, it is thus responsible for providing the results in a unified interface in order to hide the heterogeneity of the associated source, and the role of the data agent is to control the access privileges assigned to various external information systems users.

IV. RESULTS

We will present in this section different techniques that have been previously proposed in the approaches presented previously, to understand the concept of semantic interoperability when exchanging data during a collaborating between heterogeneous port information systems where each system shares information according to its own unique data model different from the others at the level of databases, design and modeling.

There are two main concepts: mapping ontology and fusion. The mapping purpose is to represent bonding between the different ontologies where each ontology's concept has an equivalent in the other ontologies [14], it will ease access to knowledge bases of other systems. The aim of fusion which resulted in the creation of the global ontology is to unify all the ontologies in order to create a common vocabulary and enable interoperability of data exchanged by resolving any semantic conflict.

Based on our study, we introduced an architecture which relies on 3 levels.

User level: Has 4 types of agents (local agent, mediation agent, manager agent and human agent) and different port information systems (application) as well as their different users.

Mediation level which contains a set of server agents, mediator agents and a global ontology that includes all local ontologies in order to process requests from different users and thus solve heterogeneity issues.

Figure 7: Screenshots of protégé

We also implemented a knowledge base that depends on various ports information systems designed with the open source Framework « Protégé » (see Fig 7) in order to define logical characteristics for classes as OWL expressions shown in Fig 8 (example of ontology creation with OWL) , we decided to create this port ontology which is based on EDIFACT-ONU standards (UN rules for the exchange of administration, commerce and transport computerized data), a set of international standards, directories and manuals for computerized data exchange [15].

```
<rdf:RDF
        xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
        xmlns:owl="http://www.w3.org/2002/07/owl#"
        xmlns:xsd="http://www.w3.org/2001/XMLSchema#"
        xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#">
        <owl:Ontology rdf:about="http://www.ontologie.fr/MonOntologie"/>
        <owl:versionInfo>Version 1.00</owl>
        <rdfs:comment>this ontology contains classes and properties of various
          Port Information System</rdfs:comment>
        <dc:creator> Mehdi ABID </dc:creator>
</rdf:RDF>
```

Figure 8: Ontology creation with OWL

Source level: Composed of various types of databases system management (Postgres SQL, Informix, Oracle Database, SQL…) belonging to different ports information systems, contains different adapters that provide a unified interface to hide the heterogeneity between these information systems.

Fig 9 shows an example of a request processing sent by a port information system to an external user through mediators, different local ontologies and global ontology.

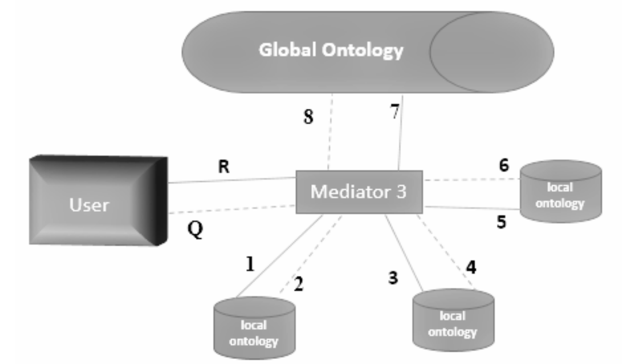

Figure 9: Example of a request processing

Once a user sends a request to an external system, this request will be automatically processed by mediators at mediation level, these mediators send their requests to their appropriate local ontology, if in this phase a null response is returned, the mediator sends the request to the global ontology resulting from the coalition of all local ontologies in order to ensure the a response for the request issued by the user.

The objective of our fusion process is to create a global ontology which includes each local ontology, and present a new coherent framework in order to enhance the vocabulary of the exchanged data and to ensure interoperability between these port information systems, and to share different existing data created by the information systems during a collaboration among them.

This process based on three levels (selection level, standardization level, sharing level) involves the task of grouping all local ontologies.

Selection level: each system selects the ontologies to unify (see Fig 10).

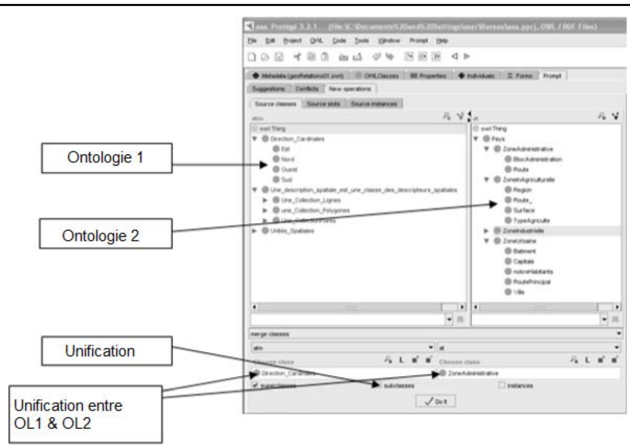

Figure 10: Process of unification between two ontologies.

Standardization level: is based on the determination of instances or subclasses of each local ontology to unify in order to create a global ontology (sharing level), this ontology (global ontology) collects information from various data and creates a shared vocabulary between all the ports information systems (see Fig 11).

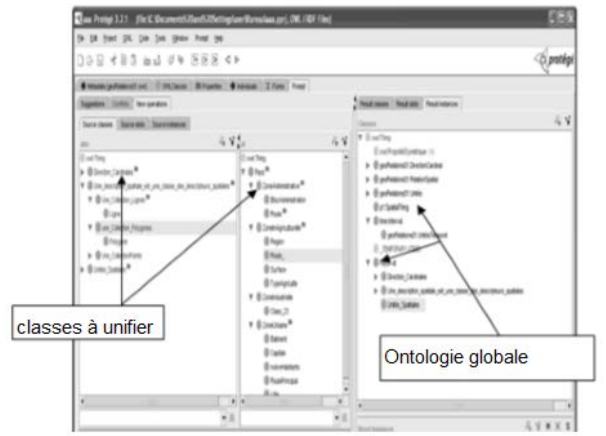

Figure 11: The global ontology

In our context, SPARQL was used as a query language and protocol to access the files describing our semantic web services as shown in Fig 12, and to implement the technology used in our proposed fusion and mapping among the different ontologies. SPARQL is a W3C Candidate Recommendation towards a standard query language for the Semantic Web [9], in fact this language is one of the pillars of the Semantic Web, It helps define the syntax and the semantics of the query in order to fetch suitable data from the database.

To stick to a security policy of information systems, which mainly relates to data exchange, as well as meeting the standards of a collaboration between different port firms, the security of XML data exchanged remains essential.

```
PREFIX vcard: http://www.w3.org/2001/vcard-rdf/3.0#
SELECT ?source ?destination
FROM 192.168.1.10/ratisseur/admin/oa/historique/xml/rdf/manutention.rdf

Exécution : sparql --query vcard1.rq | source | | destination |
                    ============
        | "Le_havre" | | "Casablanca" |
        | "Casablanca" | | "Tunisie" |
        | "Marseille" | | "Tanger" |
        | "Tanger" | | "Marseille"
        | "Agadir" | | "Casablanca" |
                    ------------
```

Figure. 12 Extract of SPARQL Query

In this study, several security approaches of data exchanged between different systems have been developed such as the use of XACML that provides a way to standardize the different access control decisions for XML documents.

The XACML used in our approach As in Fig 13, allows protection of data shared between different systems, and only systems with a specific key can decrypt the parts concerning the encrypted data [1].

```
<SIP>
    <?xacml version= 'TODO'>
        <sparql xmlns="http://www.w3.org/2005/sparql-results#":
            <source>Casablanca</source>
            <destination>Le Havre</ destination>
                <EncryptedData Id='ED1' xmlns='http://www.w3.org/2001/04/xmlenc#'
                 Type='http://www.w3.org/2001/04/xmlenc#Element'>
                    <CipherData>
                        <CipherValue>B423ZA2</cipherValue>
                    <CipherData>
                 </EncryptedData>
                 <destination>Marseille </ destination>
                <EncryptedData Id='ED1' xmlns='http://www.w3.org/2001/04/xmlenc#'
                 Type='http://www.w3.org/2001/04/xmlenc#Element'>
                    <CipherData>
```


Figure 13: Extract of XACML data

V. Conclusion

The technological evolution has established a patchwork of information systems based on different languages and many types of database, it has made it more difficult for individual firms to perform a collaboration between each other, due to their system is limited to the level of utilization of semantics, erroneous interpretations of data in a collaboration can be triggered.

In this paper we have proposed an approach based on the resolution of semantic heterogeneity in the mediation process that improved the inter-exchange of data within different port companies in collaborative level. This process is based on a 3-tier architecture (user level, mediation level, source level), and the use of the concept of multi-agent system, these agents who cooperate with each other to solve many ambiguous data problems without human intervention.

References

[1] Mehdi.ABID, Benayad NSIRI, Yassine SERHANE *"Interoperability, between Different Port Information Systems"* in *International Journal of Mathematics and Computers in Simulation*, vol. 8, 2014. pp. 156–161.

[2] Saravanan Muthaiyah and Larry Kerschberg chieving "*interoperability in e-government services with two modes of semantic bridging: SRS and SWRL"* in *Journal of Theoretical and Applied Electronic Commerce Research* ISSN 0718-1876 Electronic Versión VOL 3 ISSUE 3 DECEMBER 2008 / 52-63.

[3] Rakhi Tripathi, M.P. Gupta, Jaijit Bhattacharya, "*Interoperability adoption among government and corporate portals in India: a study"* in *Journal of Enterprise Information Management*, Vol. 25 Iss: 2, pp.98 - 122.

[4] Chrisa Tsinaraki, Yannis Velegrakis, Nadzeya Kiyavitskaya, and John Mylopoulos *"A Context-Based Model for the Interpretation of Polysemous Terms in Lecture Notes"* in *Computer Science* Volume 6427, 2010, pp 939-956.

[5] César Bravo, Jose Aguilar, Addison Ríos-Bolívar, Joseph Aguilar-Martin, Francklin Rivas-Echeverría "*A Generalized Data Meta-Model for Production Companies Ontology Definition*" in *International Journal of Systems Applications, Engineering & Development*, Issue 4 vol. 2, 2008.

[6] Hongzhe Liu, Hong Bao , Junkang Feng in "*Instance Assisted Ontology Alignment for Digital Museums*", Proceedings of the 7th *WSEAS International Conference on Simulation, Modelling and Optimization*, Beijing, China, September 15-17, 2007.

[7] J. de Brunji, M.Ehrig, C.Feier, F.Martin-Recuerda, F.Scharffe, and M.Weiten "*Ontology mediation, merging and aligning. In semantic Web T Technologies*", published by Wiley, 2006, *available online at: http://www.debruijn.net/publication/mediation-chapter.pdf.*

[8] Brenda J. Powers "*A Multi-Agent Architecture for NATO Network Enabled Capabilities: Enabling Semantic Interoperability in Dynamic Environments* (NC3A RD-2376)" In *Service-Oriented Computing: Agents, Semantics, and Engineering* Volume 5006, 2008, pp 93-103.

[9] ANDRA JAKOBSONE , SARMA CAKULA " *Information Flow Modeling to Provide Sustainable Cooperation between Educational Institutions and Entrepreneurs*" in WSEAS *Recent Advances in Educational Technologies* – Cambridge USA conferences 2013 ISBN: 978-1-61804-155-5.

[10] Miguel Miranda, Maria Salazar, Filipe Portela, Manuel Santos, António Abelha, José Neves and José Machado. "*Multi-agent systems for HL7 interoperability services*", in *International Conference on Health and Social Care Information Systems and Technologies*, 2012. Procedia Technology 5 (2012) 725-733.

[11] Miller AG 1995 wordNett: *A lexical data base for English. Communication of ACM*, pp 38-41.

[12] Wiederhold, M. Genesereth. "*the basis for mediation, proceedingé* of the third *international conference on cooperative information systems* p.140-159 may 1995.

[13] Jungmin Kim, Pankoo Kim, Hyunsook Chung "*Ontology construction using online ontologies based on selection, mapping and merging"* in *International Journal of Web and Grid Services* (IJWGS), Vol. 7, No. 2, 2011.

[14] Sarma Cakula, Abdel-Badeeh M. Salem "*E-Learning Developing Using Ontological Engineering*" *WSEAS TRANSACTIONS on INFORMATION SCIENCE and APPLICATIONS*.

[15] Recommendation on the use of standards, codes and procedures for the exchange of electronic data interchange (EDI) in cefact information navigation http://www.danubecommission.org/uploads/doc/publication/CEFACT/FR/Recommandation_EDI_ CEFACT.pdf.

Mehdi ABID was born in Casablanca, Morocco in 19-Decembre 1990. He received the M.S Degree in in science: computer and internet engineering in 2013 from faculty of science university Hassan 2 Casablanca and he is currently completing a PhD in "Laboratoire d'informatique et d'aide à la décision" in Information System Integrated Logistics, he has authored and co-authored several publications in journals and international conferences (Scopus).

Benayad NSIRI received in 2000 his D.E.A (French equivalent of M.Sc. degree) in electronics from the Occidental Bretagne University in Brest, France; his Ph.D. degree from Telecom Bretagne in 2004. While in 2005, he received a MBI degree in computer sciences from Telecom Bretagne, and in 2010 he received HDR degree from Hassan II University, Casablanca, Morocco. Currently, he is a Professor in Ain chock faculty of sciences, Hassan II University at Casablanca, Morocco; a member in LIAD laboratory, Hassan II University and a member associate in Lab-STICC laboratory at Telecom Bretagne, Brest, France. Professor Benayad NSIRI has advised and co-advised more than 7 PhD theses, contributed to more than 60 articles in regional and international conferences and journals. His research interests include but not restricted to computer science, communication, signal and image processing, adaptive techniques, blind deconvolution, MCMC methods, seismic data and higher order statistics.

Yassine SERHANE was born in Essaouira, Morocco in 15th February 1990. He received his M.S Degree in science: Computer and Network Engineering in 2013 from faculty of science University of Hassan 2 Casablanca and he is currently completing a PhD in "Laboratoire d'informatique et d'aide à la décision" in Machine Learning, he has co-authored several publications in journals and international conferences (Scopus).