Hybrid Integration Approach for Heterogeneous Information Sources, Applied to Medical Resources and Patients Medical Documents

L. Cherrat [1], M. Ezziyyani [1,2], M. Ezziyyani [2], M. Essaaidi[1] , M. Bennouna [2]

[1] Abdelmalek Essaâdi University, Faculty of Sciences of Tetuan, Telecommunication and Information System Laboratory *(LaSIT)*
[2] Abdelmalek Essaâdi University, Faculty of Sciences and Technologies of Tangier, UFR Sciences Physique d'Ingénieurs *(SPI)*
[3] *Murcie University, Department of Biology*

►Abstract

The goal of this research is to promote the establishment of an information system to provide to the Internet users the quality information on health and medicine related to the Patient Medial Documents (PMD). This information are extracted from several certified heterogeneous sources of the health localized dynamically. In addition, all the data of patients and the information relating to any medical activity (operation, consultation, analysis, radiology, etc...) are stored in Data warehouse. These Data warehouses are distributed over several regions according to the geographical distribution location of the patients. Such a functional architecture will take us to deal with two integration methodologies of the heterogeneous information systems most recommended. The Data warehouse method used to store the information related to the PDM and the virtual integration method used to retrieve the health and medicine scientific information.

In this article, we propose a new hybrid approach of integrating the heterogeneous information systems between Data warehouse and virtual mediation. We develop a mechanism rewriting query and an adapted algorithm for updating Data warehouse based on the mobile agents.

►Introduction

The information sources on the Web are:

- **Heterogeneous**
- **Distributed**
- **Autonomous**

→ **Integration need**

Information systems Integration

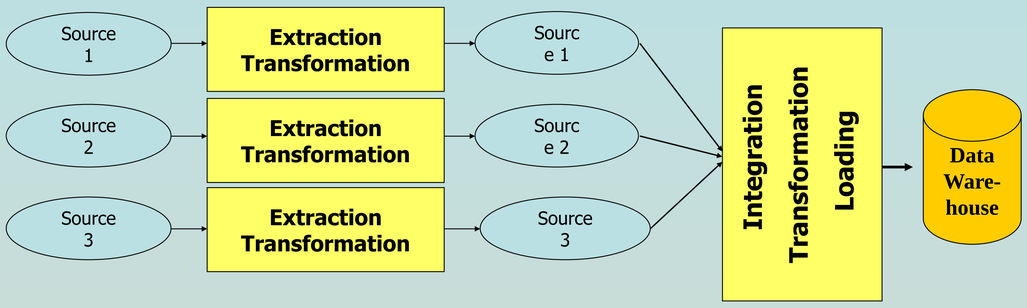
- **Provides a unified sight for multiple data sources.**
- **Ensures data availability and multiple updating**
- **Avoid overloading the network and allows access speed**
- **Facilitates the access to information.**

►Integration Methods for the information systems

There are several methods of integration, the most used are:

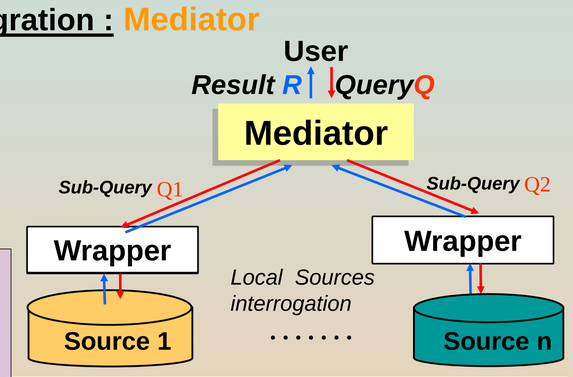
1.Materialized integration: Data warehouse

A Data warehouse is defined as a set of data from various sources, which vary in time and not volatile which are used in the process of decision support.



The process of building a data warehouse

2.Virtually Integration : Mediator



The virtually integration process pass through the following steps:

1. User sends the query to the mediator.
2. The mediator locates the sources.
3. The mediator breaks up the query into queries adapted for each source.
4. The wrapper translated the queries under a common language to the query language specific and compatible to the source.
4. Finally, the reverses process of recombining the different results coming from each sources.

To meet the interoperability needs of a growing number of hardware and software available today on the Web site, the design and development of a system, both flexible and efficient is needed. The aim of such system is to intercept users queries and find data and services most appropriate to user's queries, to define the parameters, to invoke the service and to return the result the manner transparently to users. To manage these problems of integration and to ensure an effective interoperability between different services on the Web and the availability of heterogeneous resources on the Web, we propose the development of a tool that facilitates the exploitation of these resources. This software component acts as a bridge between the user and a set of heterogeneous resources whose architecture is based on the XML model. This system based generally on two basic methods of integration, Virtually Integration and Data warehouse.

Comparative study:

	Data warehouse	Mediator
Advantages	+Interrogation of only one local source. +Allows to release itself from the constraints of networks. +Allows a historisation of data . +Allows an annotation, modification or filtering of data.	+ Facility of Addition /Suppression of sources (GAV). + Facility Rewriting of Query (LAV) + Consultation in real-time.
Disadvantages	- Data warehouse size. - Data warehouse Coherence. - Data update problems. - Data refresh . - Security management.	-Number of treatments to be carried out is more important. - Search time becomes longer.

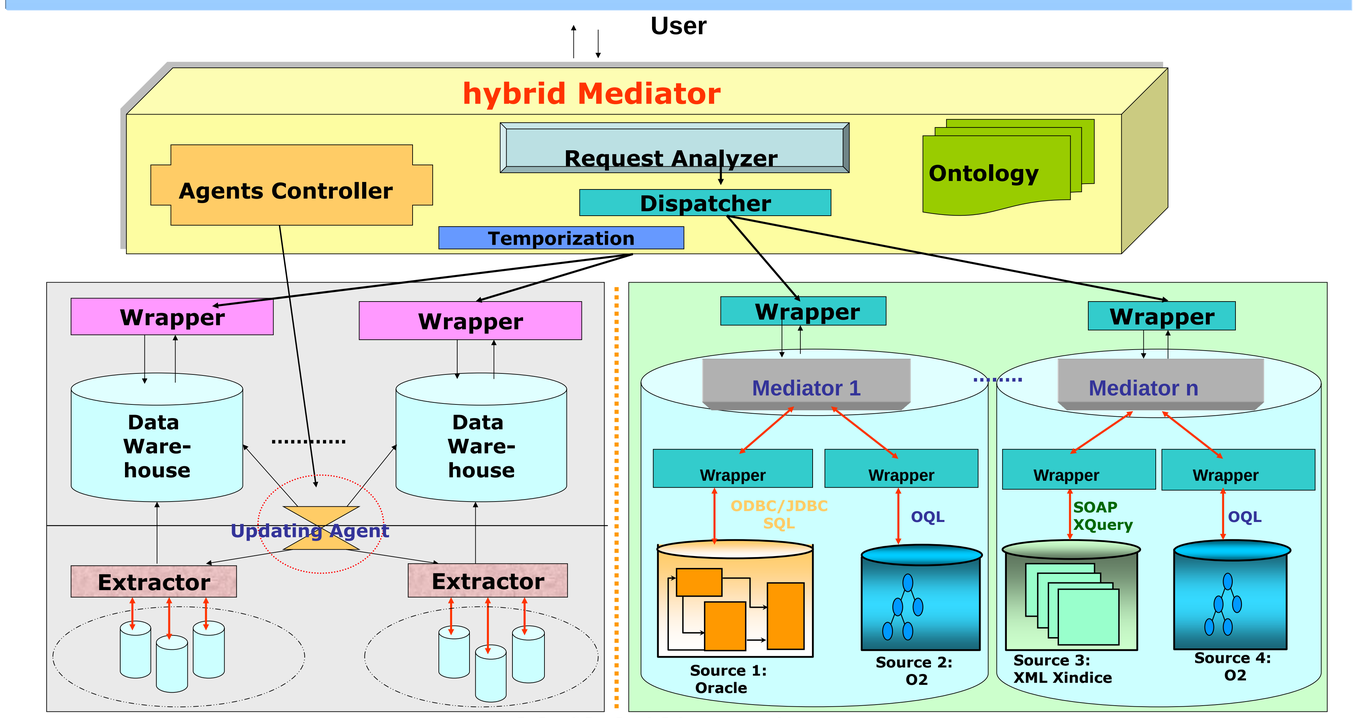
►General context of research:

We propose a new approach of integration:

A **Hybrid integration** between the virtual mediation « **Mediator**» and the materialized mediation « **data Warehouses**» .

Our research focuses on:

- A Query rewriting mechanism.
- An adapted algorithm for updating data warehouses based on the mobile agents.



New model of hybrid integration System

►Case study: Health sources Integration

Generally they are two basic problems in medical information sources:

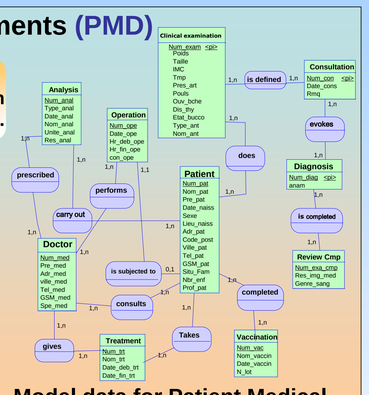
1. Heterogeneity of the Medical resources on the Web (Files, HTML, XML, Web Site, Structured Database…). So, the users are faced to sources whose structure, semantics and language query are different.
2. Difficulty of evaluating the reliability of medical and health information published on the web.

The establishment of an information system is necessary to provide to the surfer on the web to choose the certificated information on health and medicine.

Patient Medial Documents (PMD)

PMD is a computerized file that includes information on the care that you were given.

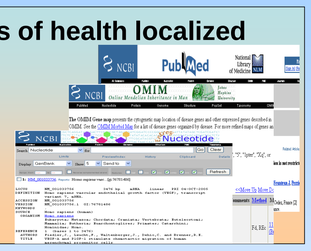
- ✓Gives to health professionals a useful information to your care.
- ✓Avoid the risk of the useless repetitions of redundant acts and the undesirable interactions of medications.
- ✓Help each patient to better control its fitness.



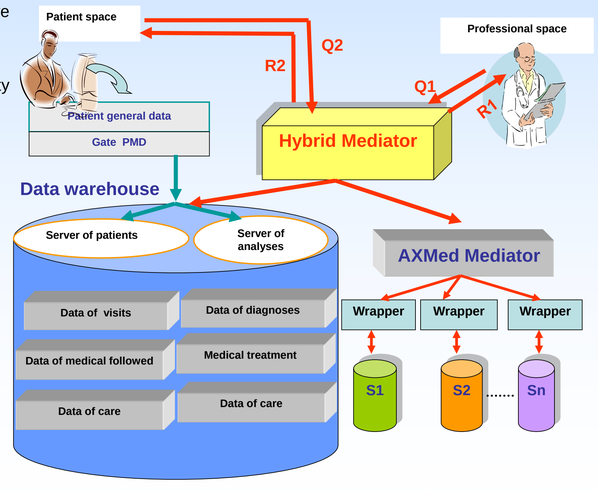
Model data for Patient Medical Documents (PMD)

Certified heterogeneous sources of health localized dynamically

They are some evaluate websites via the Internet by certification organizations such as **Health On the Net** (HON) through the development of a international websites quality code as following the standards accorded by these organizations.



Being given a patient PMD, the system locates the most adapted medical Web sites to the type of patient disease described in the PMD and carries out a selective selection according to the quality of the sites.



Functional Architecture of hybrid mediation system

Functionalities : We propose the realization of a system able to federate several data sources specific to the patients.These data can be exploited by patients, administrators, specialists and doctors. The architecture that we propose is distributed over several servers. Each server handles the data of a specific region. (analysis, consultation Radiology, Etc…). The data of each server are extracted from the existing health systems on the region (the concept of data warehouse). And for the exploitation of these data on different servers, we will use an interface for the virtual data integration via a mediator using the concept of ontology.

►Conclusion

The hybrid mediation system based on the AX^{Med} mediator appears to be an answer to the integration problem of distributed and heterogeneous health data involved in different diseases offering profitable and operable results. However there are still a number of improvements to the system. A possible improvement for future work would be the incorporation of the ontology concept to avoid information redundancy and the integration of an increasingly important quantity of data sources including highly specialized sources.

Second IFIP International Conference on new Technologies, Mobility and Security
NTM'S 2008 -Tangier- Morocco (5-7 November 2008)