بسم الله الرحمن الرحيم

SUDAN UNIVERSITY OF SCIENCE AND TECHNOLOGY

COLLEGE OF GRADUATES STUDY

Faculty of Computer Science and Information System

SURVEY ON MEDIATION SYSTEMS FOR INTEGRATION OF HETEROGENEOUS DATA

انظمة الوساطة لتكامل البيانات غير المتجانسة. دراسة مسحية

A thesis Submitted in Partial Fulfillment of the Requirements for the Master's Degree in Computer Science

Suad Mohammed Fadelamoula

SUPERVISOR:

Dr. Omer Ishag

January 2010

Outline

- Introduction
- Problem Statement
- Research Objectives
- Research Scope
- Research Methodology
- Result &Discussion
- Conclusion & Recommendation

2

Introduction

The goal of a data integration system consists in offering a uniform interface to provide access to a collection of distributed data sources, which can be heterogeneous, autonomous and dynamic. The most important advantage of a data integration system is that it enables users to specify what they want, rather thinking about how to obtain the answers.

3

Research Problem

- Several mediation systems were developed to resolve integration of heterogeneous information source this research is aimed to investigate these mediation systems and provide a comprehensive comparison between them, the comparison is based on a set of attributes including integration approach, Virtual or Materialized, integration model, query language, Mapping specification, Mapping generation, data sources, internal representation and external representation.

4

Research Objectives

- General Objectives:
  - To investigate mediation systems include TSIMMIS, MOMIS, MIX, SIMS, e-XML according to specific criteria.
- Specific Objectives:
  - To compare some of the mediation systems depending on the criteria such as approach to data source modeling ,

approach to data integration (Virtual or Materialized), integration model, query language, Mapping specification, Mapping Generation, data sources, internal representation and external representation

- To discuss the comparison in terms of characteristics that are corresponding and different

Research Scope

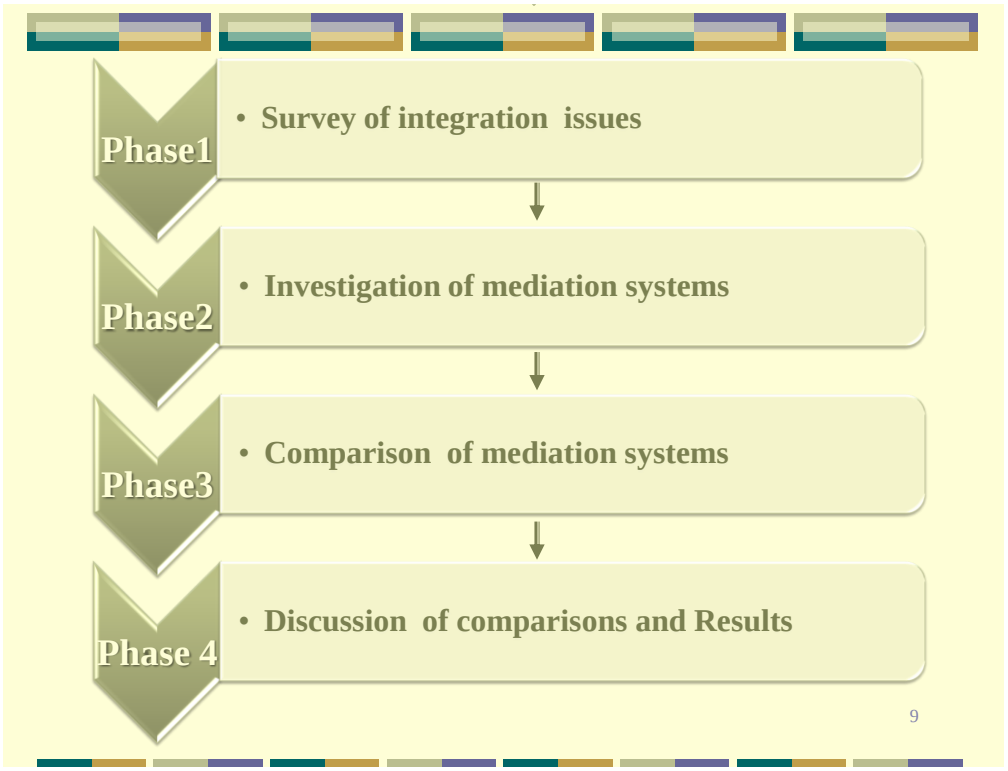
- This research is concerns mainly to the investigation and comparison of mediation systems Namely (TSIMMIS, MOMIS, MIX, SIMS, e-XML). The comparison includes integration approach, Virtual or Materialized, integration model, query language, Mapping specification, Mapping generation, data sources, internal representation and external representation.

7

Research Methodology

- In this research, well survey of database integration issues, investigating the mediation system theoretically and compared between them.
- The methods of research development will be conducting according to the workflow process as illustrated in next slide.

8



Results & Discussion

- Several mediation systems were compared namely: TSIMMIS, MOMIS, MIX, SIMS, e-XML. Different architectural aspects of these systems are taken into consideration namely: type of data sources, internal representation, external representation, Integration approach, etc. Most mediation systems overcome the problem of hard-wiring the mappings into the wrappers by using some specific mapping language e.g., there is the MIX uses *XMAS*, while TSIMMIS has MSL (Mediation Specification Language).

10

Comparison of Mediation System

Systems	Approaches to data sources modeling	Approaches for data integration	Integration Model	Query language	Mappings Specification	Mappings Generation	Data Sources	Internal Rep-resentation	External Representati
TSIMMIS	GAV	Virtual	OEM	Lorel	MSL	Provided by the mediation engineer	structured and unstructured	Relational	Semi-structured
MOMIS	GAV	Virtual	ODMI3 /OD_{LI3}	ODL_{I3}	Mapping Tables	Generated by Algorithm	Semi-structured and structured	Relational	Semi-structured
MIX	GAV	Virtual	XML	XMAS	XMAS	Provided by the mediation engineer	Xml	xml	structured
e-XML	GAV	Hybrid	XML	XQuery	XML	XML	Semi-structured and structured	Relational	semi-structured
SIMS	Hybrid	Virtual	LOOM	LOOM	LOOM	Generated by Algorithm	Semi-structured and structured	Relation	semi-structured

11

Recommendations

- There are several mediation system uncovered like Agora , AutoMed System
- the mediation systems presented are published between 1990 and 2002.
- There are other features uncovered like speed and security of each system.
- Download any source of mediation systems and implement.

12

Conclusion

- with regard the approach to data source modeling most of mediation used the GAV.
- all systems used the virtual approach in approach to data integration .
- XML has become the standard for represent ,convert and exchanging data on the Internet and on intranets in the enterprise.
- query language may differ from one data integration system to another .the researchers often choose a query language corresponding to the data model.

13

- the internal presentation most of mediation system used the relation database expect MIX system used XML.
- both GAV and LAV have some drawbacks that overcome by GLAV but most of mediation system used GAV .

14

سُبحَانَ رَبِّكَ رَبِّ العِزَّةِ عَمَّا يَصِفُونَ* وَسَلَامٌ عَلَى
المُرسَلِينَ *وَالحَمدُ لِلهِ رَبِّ العَالَمِينَ*

15