

CANDIDATO

Marco Maria Santese

TITOLO

Progettazione e sviluppo di tecniche per l'integrazione di dati tradizionali e multimediali nel sistema MOMIS

RELATORE

Prof. Domenico Beneventano

CORRELATORE

Ing. Mirko Orsini (DataRiver S.r.l.)

Dott. Alberto Corni

Università degli Studi di Modena e Reggio Emilia
Laurea Magistrale in Ingegneria Informatica (270/04)
Anno Accademico 2013/2014

Introduzione

- **Tirocinio** universitario presso DataRiver S.r.l.
- **MOMIS**: strumento per l'integrazione di dati provenienti da sorgenti (semi)strutturate eterogenee.
- **Sorgenti Dati** multimediali e tradizionali.

Obiettivo

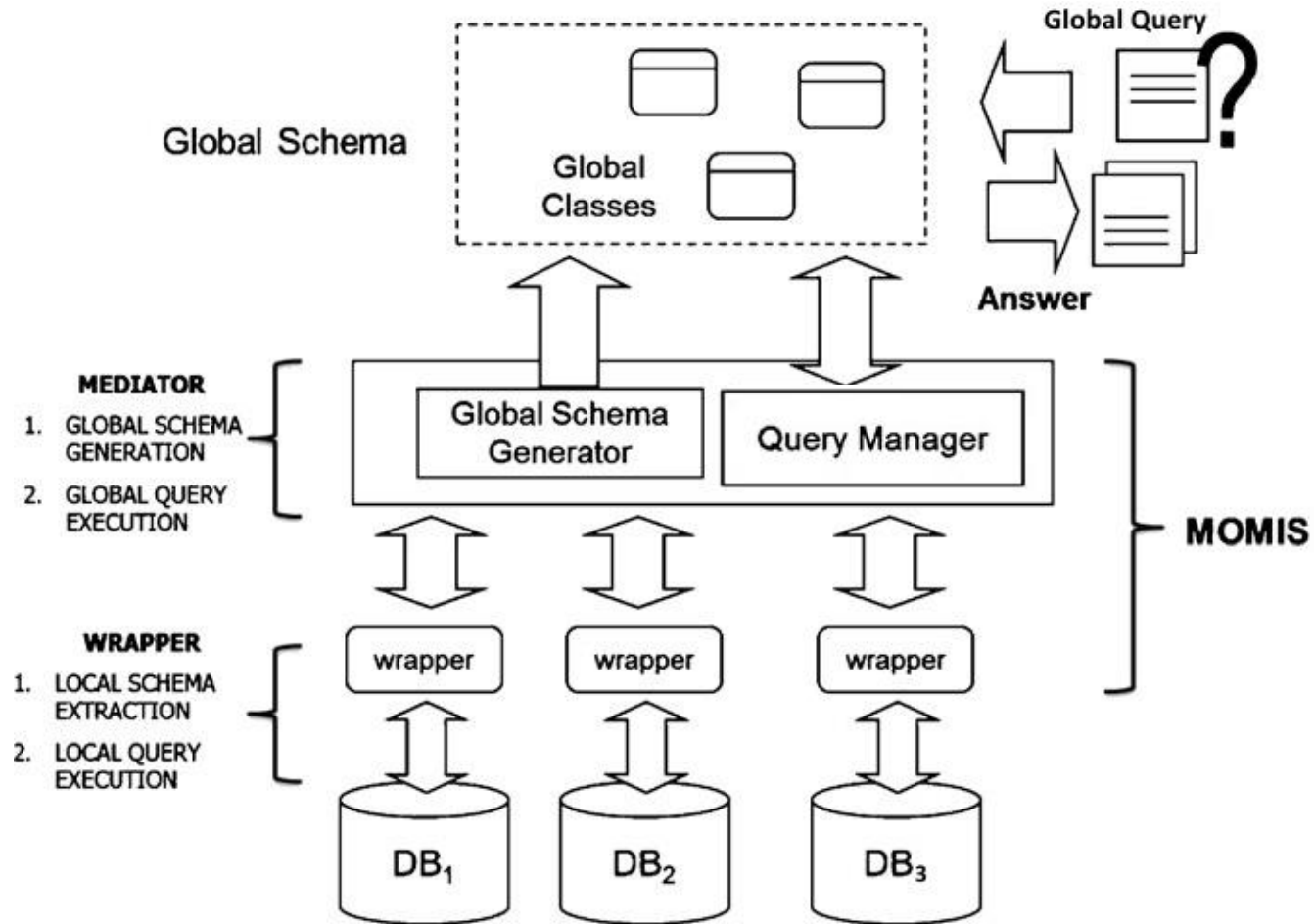
- **Sviluppo** di un'estensione per integrare ed interrogare sorgenti multimediali e tradizionali.
 - ❖ DataRiver: Finanziamento dal Fondo per l'Innovazione Tecnologica del Ministero dello Sviluppo Economico (FIT Start-Up)

Sommario

- Architettura di MOMIS
 - Query Manager (Tradizionale)
- Il framework multimediale Windsurf (Università di Bologna)
 - *Range Query*
 - *Top-K Query*
 - *Next Nearest Neighbour*
- Framework MOMIS-Windsurf
 - Interrogazione di dati tradizionali e multimediali
- Demo

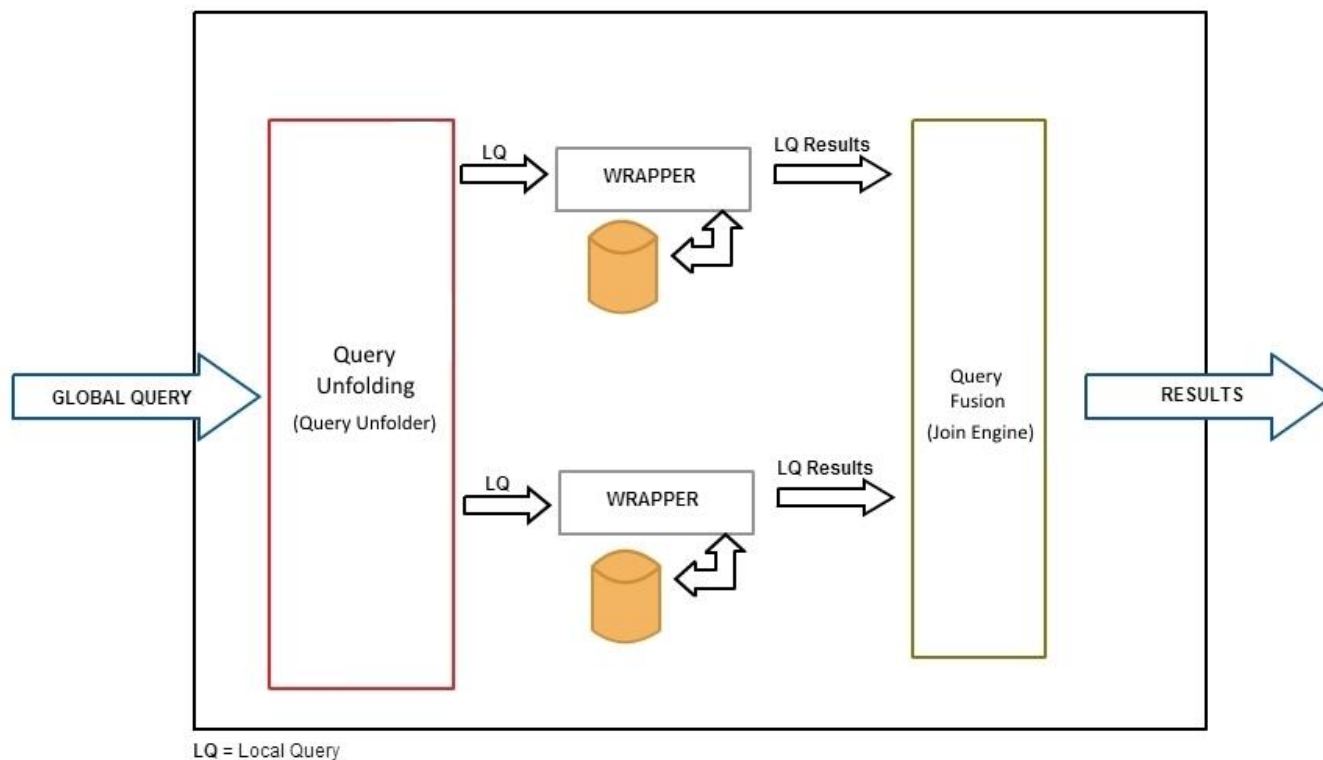
MOMIS: Architettura

- Architettura wrapper/mediator



Query Manager (Tradizionale)

- *Query Unfolding*:
 - Scomposizione della query globale in query locali per le singole sorgenti
- *Query Fusion*:
 - Fusione dei risultati locali → SQL-engine (DBMS)



Windsurf

- **Windsurf** (**W**avelet-based **IND**exing of Image**S** Using **R**egion **F**ragmentation)
 - Sviluppato dal DBGGroup dell'Università di Bologna, coordinato dal Prof. Paolo Ciaccia
 - Ricerca di dati in sorgenti multimediali complesse: Multimedia Query
 - Integrata in MOMIS come WLS (Windsurf Local Source)

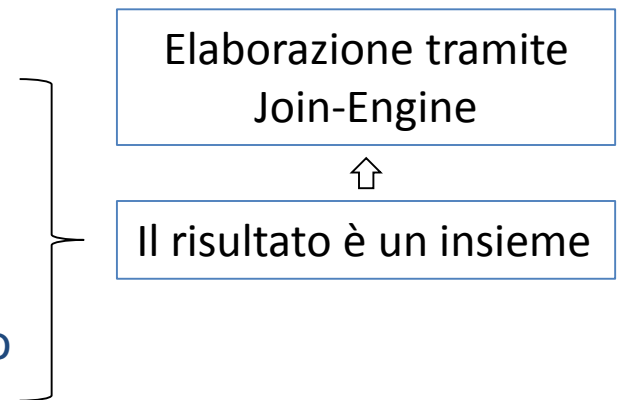
- **Multimedia Query:**
 - Range Query
 - Top-*k* Query
 - Next Nearest Neighbour

- Versione open-source limitata alla gestione di immagini

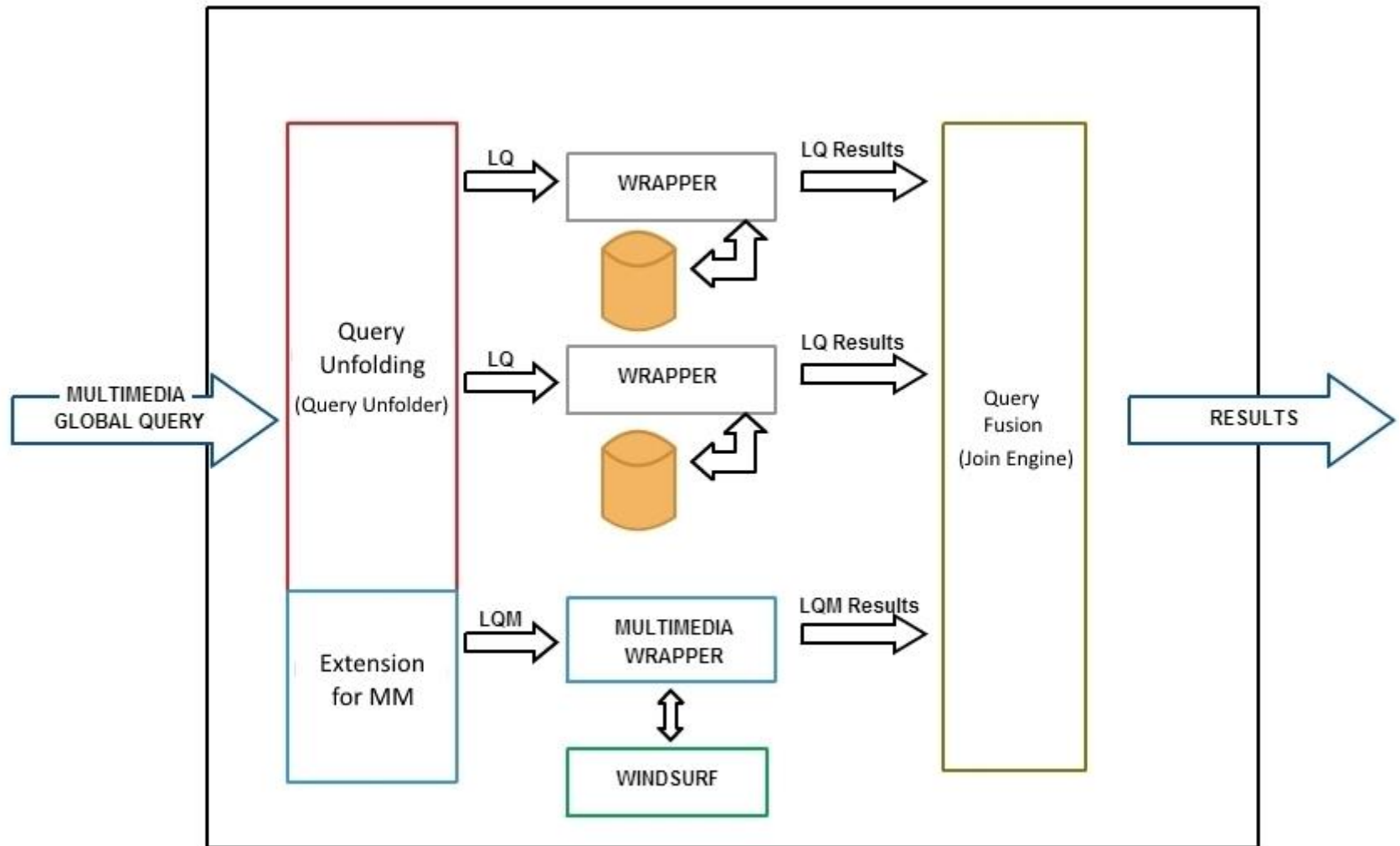
Framework MOMIS-Windsurf : Obiettivo

- MOMIS con una Sorgente Multimediale Windsurf
 - Data and Multimedia Global Schema
 - Multimedia Global Query
 - Parte tradizionale
 - Parte Multimediale (con immagine di riferimento)
- Quali *Multimedia Global Query* sono possibili considerando la sorgente multimediale Windsurf e mantenendo il Query Manager di MOMIS concettualmente *inalterato* ?

- Range Query
 - Risultato corretto e completo
- Top-*k* Query
 - Risultato corretto ma *non* completo



Framework MOMIS-Windsurf : Architettura



LQ = Local Query
 LQM = Local MM Query

Esempio (1/2)

- Sorgente Tradizionale

PERSONA(ID,NOME,ETA)

- Sorgente Multimediale

WLS(ID,IMAGE,DISTANCE)

- Schema Globale

GLOBAL(ID,NOME,ETA,IMAGE,DISTANCE)

GLOBAL	WLS	PERSONA
ID	ID	ID
NOME		NOME
ETA		ETA
IMAGE	IMAGE	
DISTANCE	DISTANCE	

- Multimedia Global Query (Top-k Query)

- “Data un’immagine di riferimento (casa) individuare le k=5 persone con casa simile ed età >18”

```

SELECT NOME,ETA,IMAGE
FROM GLOBAL
WHERE ETA > 18
ORDER BY DISTANCE(<target image >) TOP K
    
```

Esempio (2/2)

1. *Query Unfolding* (Scomposizione nelle Query Locali)

- **LQT**: SELECT * from PERSONA where ETA > 18
- **LQM**: SELECT * from WLS ORDER BY DISTANCE(<target image >) TOP-K

- Esecuzione delle Query Locali sulle rispettive sorgenti
 - ❖ **LQM** restituisce K risultati (immagini)

2. *Query Fusion*

- SELECT *
FROM LQT **JOIN** LQM USING(ID)

- ❖ I K risultati di **LQM** possono non avere corrispondenza in quelli di **LQT**

- ❖ Per le Range Query questo problema non sussiste

Top-K Query Corrette e Complete

1. Utilizzare $K' > K$ in LQM

- Stima selettività dei predicati presenti in LQT per determinare K'

2. Modifica Metodo di Fusione

- Esecuzione LQT e uso di “Next Nearest Neighbour” Query su WLS
- Query Manager ancora basato su SQL-engine ma nuovi metodi di join

✧ Position Paper al SEBD 2015 (Symposium on Advanced Database Systems)




❖ Lavoro precedente su MOMIS & Multimedia:

[Beneventano,D.,Gennaro,C.,Bergamaschi,S.,Rabitti,F.: A mediator-based approach for integrating heterogeneous multimedia sources. Multimedia Tools and Applications, 2013]

- ❖ Basato su MILOS (CNR Pisa, Dott. Rabitti)

Demo

The screenshot shows a software interface with a main window titled "Momis - NewProject_Tesi.prj". The interface includes a "Project Glo" sidebar, a "Quick Access" search bar, and a toolbar with icons for "Clear MM", "Show MM", and "Show MM Query". A SQL query is displayed in a text area: `SELECT * FROM persone where eta > 18`. Below the query, there are buttons for "Add Multimedia" and "Run Query". A table displays the results of the query, with columns for "IMAGE" and "DISTANCE". The table contains three rows of data. A red circle highlights the "Limit 5" input field. An "Info" dialog box is open in the foreground, displaying the text "DISTANCE= 0.5697487954458229" and a "Copy to ClipBoard!" button. Another "Info" dialog box is visible in the background, showing a street scene image and a "Copy to ClipBoard!" button.

IMAGE	DISTANCE
	0.5697487954458229
	0.6060999140681845
	0.6084104276970637

Conclusioni e Sviluppi Futuri

- ✓ Progetto ed implementazione del framework MOMIS-Windsurf
 - Il sistema MOMIS è risultato essere modulare ed estensibile

- ✓ Case Study (Campo Medico)
 - Analisi integrata di esami PET (*Positron Emission Tomography*)

- Studio ed implementazione di nuovi metodi di fusione per le TOP-K Query

- Gestione di altri elementi Multimediali oltre alle immagini

FINE

Grazie per l'attenzione