

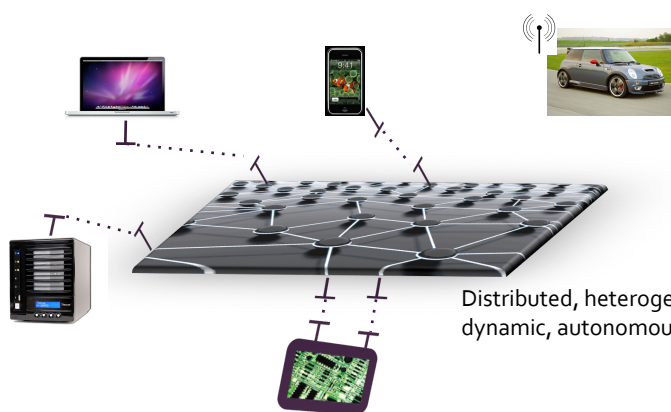


## HADAS Group Heterogeneous Autonomous Distributed Data Services

Christine Collet, [hadas@imag.fr](mailto:hadas@imag.fr)



## DIGITAL ENVIRONMENT



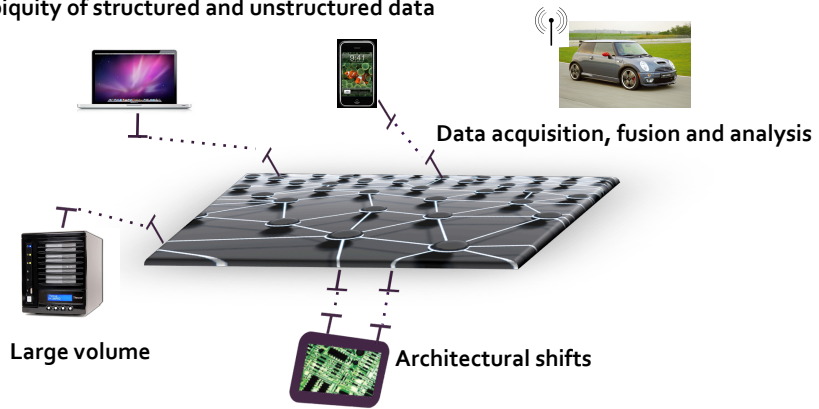
Distributed, heterogeneous, mobile,  
dynamic, autonomous objects

Plethora of heterogeneous data

2

## TURNING POINT IN DATA MANAGEMENT

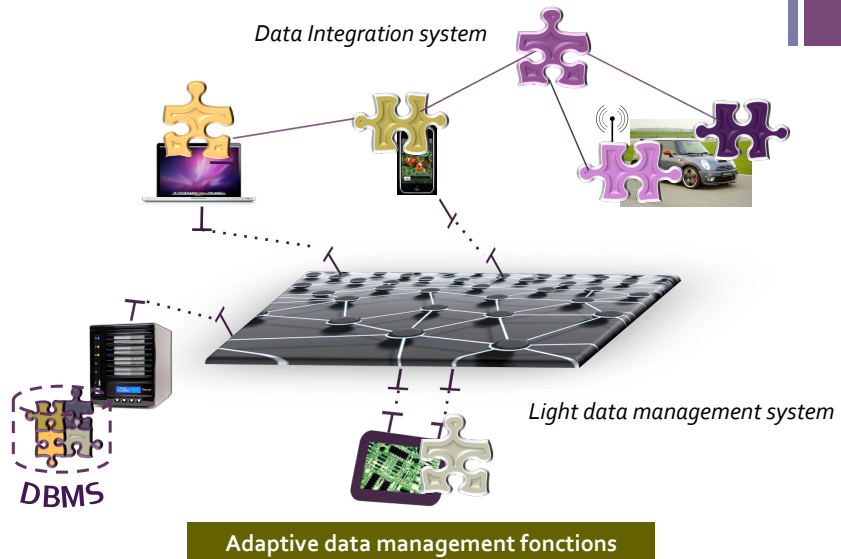
Ubiquity of structured and unstructured data



New system components for data management

3

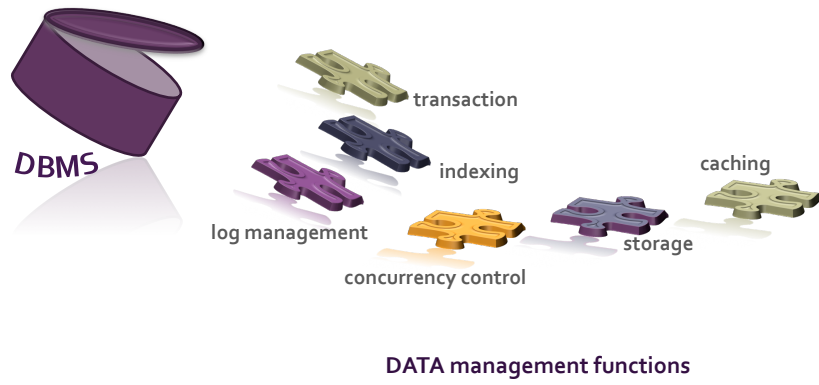
## DATA MANAGEMENT REQUIREMENTS



Adaptive data management functions

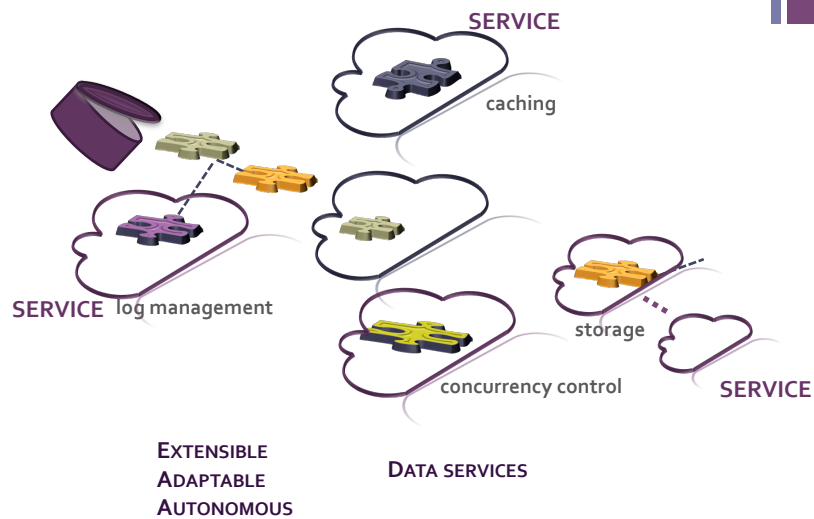
4

## DECONSTRUCTING DBMS



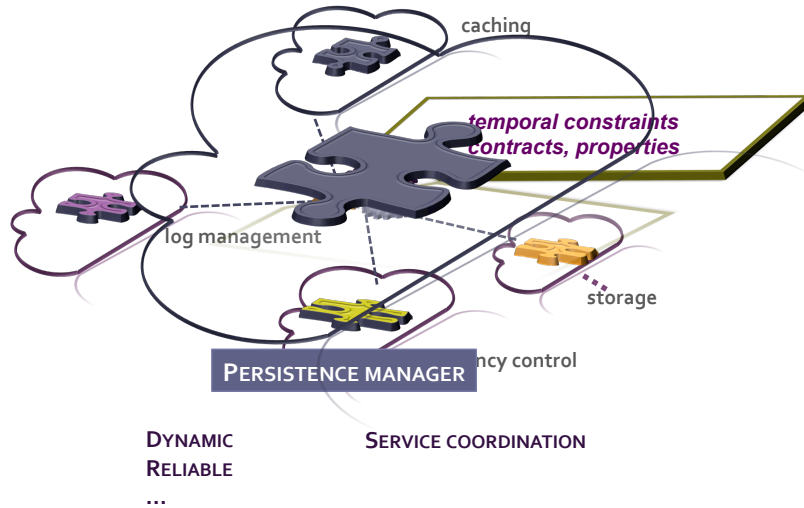
5

## From FUNCTIONS to SERVICES



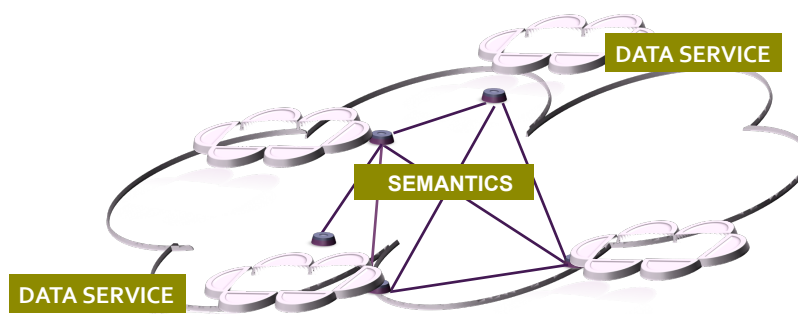
6

## COMPOSING DATA MANAGEMENT FUNCTIONS



7

## GENERAL OBJECTIVE : Semantic-based infrastructure for managing data



**HADAS: The right data services you need ...**

8



## THE GROUP

### ■ RESEARCHER (1)

- Geneveva Vargas-Solar (CR CNRS)

### ■ PROFESSORS AND ASSOCIATE PROFESSORS(5)

- Christophe Bobineau (MC Grenoble INP)
- Christine Collet (Pr Grenoble INP)
- Fabrice Jouanot (MC UJF)
- Marie-Christine Rousset(Pr UJF)
- Alexandre Termier (MCF UJF)

### ■ POSTDOC (1)

- Noha Ibrahim (ANR Optimacs)

■ *Departure (2008): C. Roncancio, C. Labbé*

■ *Retirement(2007): M. Adiba*

### ■ CURRENT DOCTORAL STUDENTS (9)

- Anis Benyelloul ANR Continuum
- Victor Cuevas-Vicenttin ANR WebContent
- Javier A. Espinosa-Oviedo Mexican Grant
- Benjamin Negrevergne MENRT
- Mohamed Othman Abdallah MENRT
- Alberto Portilla-Flores Mexican Grant
- Rémi Tournaire Territorial grant
- *Lourdes Martinez-Medina ANR Blanc*
- *Carlos Manuel Lopez Enriquez ANR*

2-3 PHD students per year

9

## OUTLINE

- Research context and objectives

- HADAS scientific themes and projects

- Achievements: collaborations, numbers and dynamics

- Management

- Self-assessment

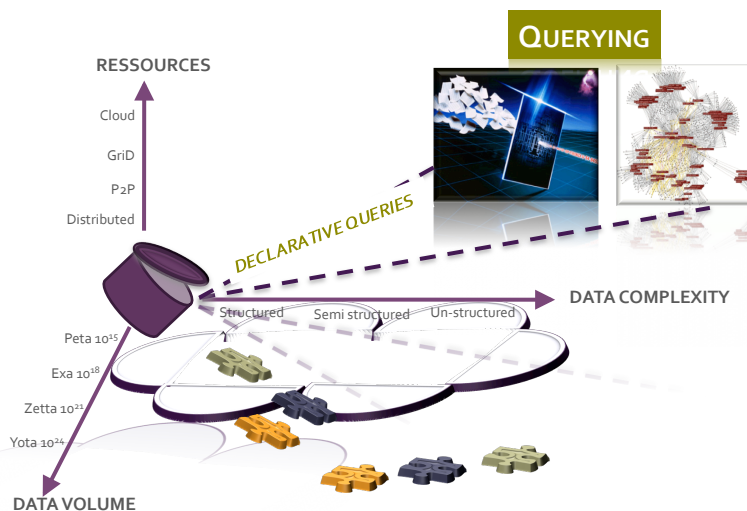


10

## RESEARCH THEMES

- ACCESSING DATA IN THE LARGE
- COMPOSING DATA SERVICES ON THE FLY
- REASONING ON DATA SEMANTICS

## ACCESSING DATA IN THE LARGE



## FOCUS

### Query evaluation

- Machine-learning-based adaptive query optimization
- Combine data and network management in dynamic ad-hoc networks (ANR blanc UBIQUEST)

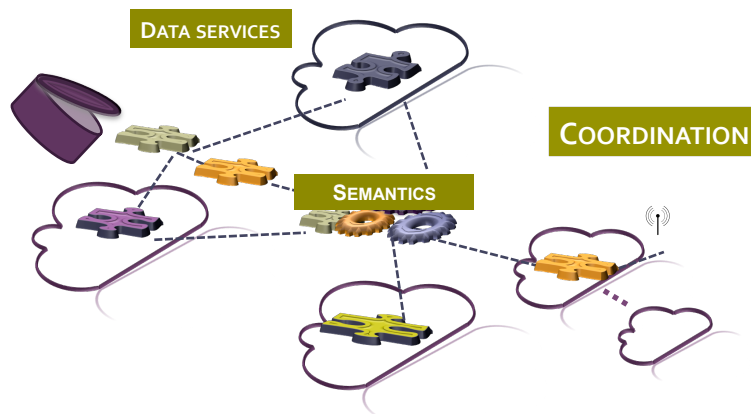
### Mining

- Pattern mining on multicore processors (with MESCAL group)
- Data mining on the chips (TIMA, MSTIC DAMOCLES project)

OBE, ACS, DryadeParent, DigDag

13

## ADAPTABLE SERVICES COMPOSITION



14

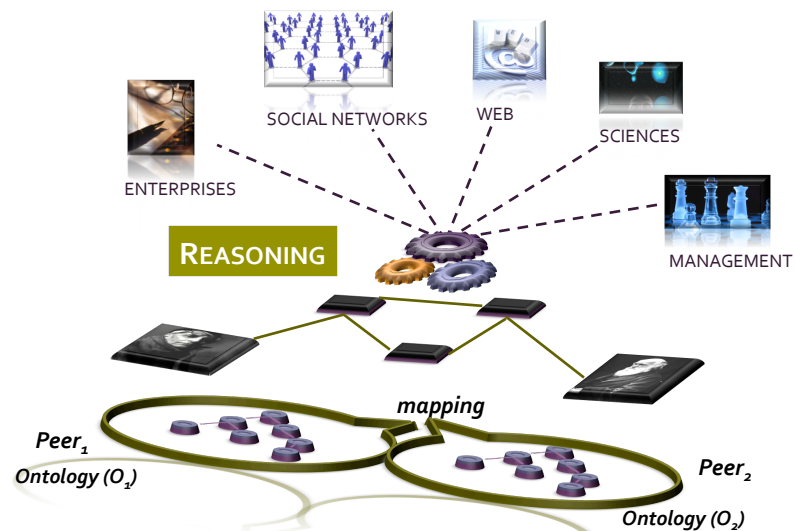
## FOCUS

- **Reliable services coordination**
  - Contract based secure, atomic and adaptable services coordination (ECOS ORCHESTRA)
  - Mashup language for integrating data on the web (E-CLOUDS, INP BQR REDSHINE)
  
- **Query processing using services (ANR OPTIMACS)**
  - Services coordination for querying data
  - Multi-dimensional optimization using operational research techniques

PYROS, MEOBI, REKHA, SEBAS

15

## REASONING ON DATA SEMANTICS



16

## FOCUS

- **Querying in P2P Data Management Systems**
  - Automatic discovery of probabilistic mappings between taxonomies of classes (SomeWhere, ANR WebContent, Cluster ISLE : WEB INTELLIGENCE)
  - Decentralized handling uncertainty and trust (ANR DATARING, ERC WebDam)
- **Handling online contextual data for adapting applications to evolving environments (ANR CONTINUUM)**
- **Services / coordination properties description and checking**
  - Atomic, secure and adaptable services coordination (ORCHESTRA)
  - Mashing up web data (E-CLOUDSS, RED-SHINE)

ADEMS, SOMEWHERE, CMF, SEBAS

17

## HADAS POSITION / LIG THEMES

### Infrastructures : from network to data

Optimization and evaluation of queries on heterogeneous and multi-scale data sets (service-based, machine-learning based, combining with networks)

### Software : from foundation to engineering

Adaptive and autonomous data service composition (cloud computing)  
Parallel data mining algorithms (multicore processors, chips)

### Interaction : from perception, to action and dialog

### Knowledge : from extraction and transformation to usage

Models and Algorithms for reasoning on distributed semantics of data and services

18

## HADAS POSITION / LIG CHALLENGES

### Social challenges

- **Intelligent building:** distributed management of data and services
- **Open enterprise:** data mining, resources composition (data and services)

### Technological challenges

- **Sensor networks, embedded systems:** integrated management of data and networks. Learning based query optimization
- **Virtualization and performance:** parallel data mining algorithms on multicore architectures

### Conceptual challenges

- **Access to information:** semantics and data filtering; coordination of operators or services for accessing data
- **Emerging computing models:** reasoning on data semantics. Algorithms and data management services for cloud computing

19

## OUTLINE

- Research context and objectives
- HADAS scientific themes and project
- Achievements: collaborations, numbers and dynamics
- Management
- Self-assessment



20

## INTERNATIONAL COLLABORATIONS

**MEXICO**

- 20 years of collaboration in TICS
- LAFMI (coordination 2002 – 2006)
- LAFMIA (UMI 3175 2008)

**CHINA**

- LIAMA (LIA)

**VIETNAM**

- MICA (UMI ??? 2006)
- ASIA-LINK CONE

**JAPAN**

- University of Hokkaido
- University of Osaka
- NII

21

## INTERNATIONAL COLLABORATIONS

- **Hokkaido University**
  - Peer to peer composition of resources for the Semantic Web
  - PhD student S. Tandabany (one year in Japan)
- **Osaka University** (Pr. Takashi Washio of I.S.I.R)
  - Graph mining algorithms
  - Japanese contract with Fujitsu and Japan Airlines
- **NII** (Prof. Takeaki Uno)
  - Parallelisation of "LCM" algorithm
  - PhD student B. Négrevérigne (three months in Japan)

22

## INTERNATIONAL COLLABORATIONS

### ■ VIETNAM

- MICA laboratory in Ha Noi
- Participation to the Asian projects (4 Ph.D. students)
- ASIA-LINK CONE project
- Lectures and research tutorials



### ■ CHINA

- Joint developments and publications with LIAMA
- ANR Blanc project (UBIQUEST)

23

## INTERNATIONAL COLLABORATIONS

### ■ Scientific results

- Publications and prototypes
- Education of graduate students through co-advising contracts
- Organisation of thematic schools (LAFMI/LAFMIA schools)



### ■ Current research themes

- Services based infrastructures for managing distributed data with reliability, e-cloud computing
- 5 PhD students A. Portilla, V. Cuevas and J. A. Espinosa-Oviedo, C. Lopez-Enriquez, L. A. Martinez-Medina
- ORCHESTRA (ECOS-ANUIES), REDSHINE (BQR Grenoble INP), e-CLOUDSS (Microsoft)

24



## LAFMIA (UMI 3175)

### FRENCH MEXICAN LABORATORY OF INFORMATICS AND AUTOMATIC CONTROL

#### ■ History

- LAFMI - LAFMAA (2002 – 2006)
  - Virtual labs on Informatics and Automatic Control
- Coordination LAFMI 2002 – 2005 M. Adiba, 2005 -2006 G. Vargas-Solar

#### ■ Joint international Unit on Informatics and Automatic Control (28.03.09)

- Institutions: CNRS, UJF, Grenoble INP, UTC, CONACyT
- Administrative location: CINVESTAV
- Antenas: Central valley of Mexico (Puebla, UDLAP), Northern region of Mexico (Baja California, CICESE; Monterrey, ITESM)

#### ■ Joint Projects and Co-supervisions of graduate students, double diploma programs

- Genoveva Vargas-Solar, deputy director

25

## ACHIEVEMENTS

#### ■ Publications

- 2,89 publi per Ne.year
- Paper awards
- 100% EC & C publishing

	2005	2006	2007	2008	2009	Par an
Total	33	29	28	45	16	
ACL	2	2	2	3	1	
ACT	14	15	9	22	6	
ACL+ACT	16	17	11	25	7	15,2
Ne = 2/3EC+1C+1/2IR	5,67	5,67	5,00	5,67	4,33	5,27
ACL+ACT per Ne	2,82	3,00	2,20	4,41	1,62	2,89

#### ■ Software licenses : Somewhere, DigDag

#### ■ 10 Doctoral dissertations

- average of 4 years duration
- Grants /BDI(4), Contracts/CIFRE(6)
- After thesis: Associate Professor, Research Engineers, Post-doc

#### ■ Contracts

- International : 2
- Industrial :
- ANR : 5
- BQR/MSTIC : 3

Microsoft, Orange labs, Japan Airlines, Fujitsu, GMALTO, Lyonnaise des eaux, SUEZ env., ... , CITI, INSA, INRIA Univ. Valenciennes, Univ. de Nice,

220 à 250K Euros per year

40K Euros per year / Ne

26

## SCIENTIFIC VISIBILITY

- Tutorials, keynotes (12) and communications
- Implication in scientific organisations
  - **BDA committee** member: Ch Collet (97-), *FLIPPER Study coordination*(2010-)
  - Member of the Extended Database Technology(**EDBT**) **endowment**: Ch. Collet (2004 - ) – EDBT schools management
  - Member of the IJCAI-09 advisory committee: M.-C. Rousset
  - **Conseil Scientifique du Département STII** (membre nommée, septembre 2006): M.-C. Rousset
  - **ERC starting grant evaluation** (2008, 2009): M.-C. Rousset
  - **Presidence of the Mexican Society of Computer Science**: G. Vargas-Solar (2007-2009)
- Program committee member (5 /year/person)
  - of the major conferences and journals in DB and IA
- Conferences, Workshops, Schools organization

27

## RESPONSIBILITIES

- **VP adjointe recherche groupe Grenoble INP** since 2007: Ch. Collet (April 2007-)
- **Chargée de Mission** pour
  - le Pôle information et Communication, Grenoble INP: Ch. Collet (2002-2007)
  - Relations Internationales auprès du collège Doctoral de l'UJF since 2007 : M.-C. Rousset
  - Auprès du LIG pour la prospective scientifique since 2007: M.-C. Rousset
- **Deputy director of the UMI LAFMIA** since 2008: G. Vargas-Solar

28

## OUTLINE

- Research context and objectives
- HADAS scientific themes and projects
- Achievements: collaborations, numbers and dynamics
- Management
- Self-assessment



29

## GROUP MANAGEMENT

- Meetings : bi-monthly meetings for permanent people + seminars for all people
- Social events and Outside white/green working Days
  - "Un Grenelle" de HADAS => 3 themes
- Shared resources and autonomous management of contracts
- Reserved resources for: schools participation, international actions, projects submissions

30

## Self-assessment (SWOT)

Solid Background,  
Animation and training  
Good recognition & visibility  
Numerous International  
collaborations

Few  
Publications in  
journals  
Few HDR

DB Research  
programs  
and networks  
Local Dynamics

Work overload  
Dispersion

31

## THANK YOU

More on <http://hadas.imag.fr>



32