

# Webdam – Kick-off meeting

Serge Abiteboul, INRIA Saclay

ENS Cachan, 01/07/2009

INSTITUT NATIONAL  
DE RECHERCHE  
EN INFORMATIQUE  
ET EN AUTOMATIQUE



centre de recherche  
**SACLAY - ÎLE-DE-FRANCE**

# Webdam

Topic:	Foundations of Web Data Management European Research Council Advanced Grant
Founding:	2.4 millions Euros
Timing:	2009 – 2013
Institution:	INRIA Saclay – Île-de-France
Hosts:	Univ. Paris Sud (Gemo) and ENS Cachan (Dahu)



# Organization

Technical content

The people

Meetings

Editing

Website

Discussion



# Technical content

INSTITUT NATIONAL  
DE RECHERCHE  
EN INFORMATIQUE  
ET EN AUTOMATIQUE



centre de recherche  
**SACLAY - ÎLE-DE-FRANCE**

# Webdam thesis in brief

## Management of local data

- Well-accepted model: relational model
- Well-accepted theory: FO & concurrency control & etc.
- Development of systems and theory in parallel

## Management of distributed/Web data – not the case

- Rapid software development & no well-accepted model/theory

## Illustration

- Relational database course: clean and neat course
- Distributed databases: a series of recipes

***The Web is too complex to continue hacking systems with no unifying underlying model***



# A first shift: distribution in autonomous systems

Web/distributed/peer-to-peer data management

*Focus is on **information** residing on (possibly very many) **autonomous systems***

Information

- Structured data (relations)
- Semistructured (XML, graphs)
- Unstructured (text)
- Metadata
- Knowledge
- Physical data: indices, views, cache...

Focus is more on data management than on semantic Web but border is unclear



# A model for Web data management

What should be expected for a such a model?

- Describe query processing, changes, monitoring, communication, service choreography, data integration/exchange, distributed applications
- Road map

What are the issues ?

First attempt – ActiveXML: XML with embedded service calls

- With “local” logic (tree pattern queries): a calculus for Web data
- An algebra
- Optimization
- Verification



# A second shift: deduction

Relational system: FO on everyone's desk

Hardwired logic

- Concurrency control with 2PL
- Integrity constraint enforcement
- Access control enforcement

On the Web or in distributed systems: this is not available

Verification of distributed applications

- Work on business processes
- Verification of ActiveXML applications





# What is the target?

Improve understanding of Web applications

Facilitate teaching Web data management technology

Improve Web applications: better performance, more reliable, better access control, etc.

Facilitate development of Web applications – programmers productivity



# The people

INSTITUT NATIONAL  
DE RECHERCHE  
EN INFORMATIQUE  
ET EN AUTOMATIQUE



centre de recherche  
**SACLAY - ÎLE-DE-FRANCE**

# Webdam members

**Already in:** Abiteboul, Segoufin, Vianu, Senellart, Galland, Marinoiu, Bourhis, Kharlamov

**Potentially:** The members of Gemo and Dahu who are interested

**Soon:**

1. Balder ten Cate - starting April 2009 for 6 months or more
2. Philippe Rigaux – seconding “détachement” september 2009
3. Marie-Christine Rousset – ½ seconding “délégation” september 2009

**Short visits:** Amélie Marian, Tova Milo, Alkis Polizotis

**Other ideas?**



# Researchers who are being considered

discussion



# Hiring season

Advertise there are positions

Do it primarily by contacting friends/good groups

Priority for Webdam:

researchers with some experience (more than PhD or Posdoc)



# Events

INSTITUT NATIONAL  
DE RECHERCHE  
EN INFORMATIQUE  
ET EN AUTOMATIQUE



centre de recherche  
**SACLAY - ÎLE-DE-FRANCE**

# Workshops with other projects - candidates

## Local

- STREP Fox (Luc Segoufin): Foundations of XML
- ANR DataRing (Patrick Valduriez): Distributed data management
- ANR Docflow (Anca Muscholl): Verification of Active XML

## ERC

- Stefano Ceri: New paradigms for Web search
- Anne-Marie Kermarrec: Web search in P2P
- Carlo Ghezzi: Self-Managing Situated Computing



# Workshops with other projects

**Como Lake**

**June 19-20, 2009**

**Joint with ERCs of Stefano Ceri and Carlo Ghezzi**

**Workshop in corsica or lumini: joint with Fox in 2010?**





# Prospective meeting

**Core: Advisory board of Webdam + Webdam + guests  
next to VLDB - end August 2009**

- In paris
- Deliverable:
  - Manifesto on Theory of Web data management
  - Roadmap

**Other ideas?**



# Editing

# Web Distributed Management of Data

Text book

Used in courses at Orsay, Dauphine & Telecom ParisTech

Pushed and completed

Rigaux, Abiteboul, Rousset, Senellart



# Second Foundations

Continuation of Foundations of databases AHV

## Foundations of Web data management

Web publication but high editorial standard

Topics: XML, tree automata, data integration/exchange, query processing...

**Lead? Participants?**



# Other ideas

## Tutorial and text on Active XML

- To be proposed to PODS or ICDDT?

## Book – long paper on ActiveXML

## Manifesto in favor of theory of Web data management

- With Fox?

## What else?



Website : <http://webdam.inria.fr>

to do: intranet & mailing list

# Merci

# Discussion