

ERC Grant
Number: 226513
Acronym: Webdam

Title: Foundations of Web Data Management

Activity Report part 3

Period covered: 31/12/2011-31/05/2013
Start date of project: 01/12/2008
PI: Serge Abiteboul
Host: INRIA Saclay-Île-de-France

Goals The Webdam ERC grant (S. Abiteboul) started in December 2008. The goal is to develop a formal model for Web data management that would open new horizons for the development of the Web in a well-principled way, enhancing its functionality, performance, and reliability. More specifically, the aim is to create a universally accepted formal framework for describing complex and flexible interacting Web applications featuring notably data exchange, sharing, integration, querying and updating. It is also to develop formal foundations enabling peers to concurrently reason about global data management activities and cooperate in solving specific tasks with desired quality of service.

1 Webdam Project Achievements

The project goes on according to plan. The main activities during this period were :

1. The implementation of the Webdamlog engine (demonstrated at SIGMOD 2013), a distributed datalog system, realized primarily by Emilien Antoine.

2. The development of a framework for distributed access control (article in preparation) with the participation of Gerome Miklau from U. Massachusetts (one year visit in Webdam).
3. Works on collaborative workflows (article in PODS 2013) with the participation of Victor Vianu from UCSD (six month visit in Webdam).
4. Works on probabilistic data led by Pierre Senellart (démonstration at SIGMOD 2012, article at ICDE 2013).

Major results were achieved in these different areas. Several members of the project have obtained prestigious positions, notably professor positions for Daniel Deutch in Tel Aviv U. and Balder ten Cate, UC Santa Cruz, and a research position for Fabian Suchanek at Max Plank Institute in Saarbrücken. Serge Abiteboul has been appointed to the Conseil National du Numérique (government council).

2 Publications 2013

The publications are all reachable from the Webdam publications page: http://webdam.inria.fr/wordpress/?page_id=8.

Conferences

- [1] Serge Abiteboul, Pierre Bourhis, Anca Muscholl, and Zhilin Wu. Recursive queries on trees and data trees. In *ICDT*, pages 93–104, 2013.
- [2] Serge Abiteboul and Victor Vianu. Collaborative data-driven workflows: Think global, act local. In *PODS*, 2013.
- [3] Diego Calvanese, Evgeny Kharlamov, Marco Montali, Ario Santoso, and Dmitriy Zheleznyakov. Verification of inconsistency-aware knowledge and action bases (extended version). *CoRR*, abs/1304.6442, 2013.
- [4] Chao Li, Daniel Yang Li, Gerome Miklau, and Dan Suciu. A theory of pricing private data. In *ICDT*, pages 33–44, 2013.
- [5] Chao Li and Gerome Miklau. Optimal error of query sets under the differentially-private matrix mechanism. In *Proceedings of the 16th International Conference on Database Theory, ICDT '13*, pages 272–283, New York, NY, USA, 2013.

Demonstrations (in proceedings of international conferences)

- [6] Serge Abiteboul, Émilien Antoine, Gerome Miklau, and Julia Stoyanovich. [demo] rule-based application development using webdamlog. In *SIGMOD*, 2013.

3 Publications 2012

Journals

- [7] Serge Abiteboul, Pierre Bourhis, and Victor Vianu. Comparing workflow specification languages: A matter of views. *ACM Trans. Database Syst.*, 37(2):10, 2012.
- [8] Serge Abiteboul, T.-H. Hubert Chan, Evgeny Kharlamov, Werner Nutt, and Pierre Senellart. Capturing continuous data and answering aggregate queries in probabilistic XML. *ACM Trans. Database Syst.*, 36(4):25, 2011.
- [9] Serge Abiteboul, Georg Gottlob, and Marco Manna. Distributed XML design. *J. Comput. Syst. Sci.*, 77(6):936–964, 2011.
- [10] Bogdan Cautis and Evgeny Kharlamov. Answering queries using views over probabilistic XML: complexity and tractability. *PVLDB*, 5(11):1148–1159, 2012.
- [11] Fabian M. Suchanek, Serge Abiteboul, and Pierre Senellart. PARIS: Probabilistic alignment of relations, instances, and schema. *PVLDB*, 5(3):157–168, 2011.

Conferences

- [12] Serge Abiteboul, Yael Amsterdamer, Daniel Deutch, Tova Milo, and Pierre Senellart. Finding optimal probabilistic generators for XML collections. In *ICDT*, pages 127–139, 2012.
- [13] Serge Abiteboul, Emilien Antoine, and Julia Stoyanovich. Viewing the Web as a distributed knowledge base. In *ICDE*, pages 1–4, 2012.
- [14] Serge Abiteboul, Pierre Bourhis, and Victor Vianu. Highly expressive query languages for unordered data trees. In *ICDT*, pages 46–60, 2012.
- [15] Serge Abiteboul, Pierre Senellart, and Victor Vianu. The ERC Webdam on foundations of Web data management. In *WWW (Companion Volume)*, pages 211–214, 2012.
- [16] Michael Benedikt, Pierre Bourhis, and Pierre Senellart. Monadic datalog containment. In *ICALP*, pages 79–91, 2012.
- [17] Bernardo Cuenca Grau, Ernesto Jiménez-Ruiz, Evgeny Kharlamov, and Dmitriy Zheleznyakov. Ontology evolution under semantic constraints. In *KR*, 2012.
- [18] Fabian M. Suchanek, David Gross-Amblard, and Serge Abiteboul. Watermarking for ontologies. In *ISWC*, pages 697–713, 2011.

Workshops

Demonstrations (in proceedings of international conferences)

- [19] Serge Abiteboul, Yael Amsterdamer, Tova Milo, and Pierre Senellart. Auto-completion learning for XML. In *SIGMOD Conference*, pages 669–672, 2012.
- [20] Michael Benedikt, Tim Furche, Andreas Savvides, and Pierre Senellart. ProFoUnd: program-analysis-based form understanding. In *WWW*, pages 313–316. ACM, 2012.
- [21] Luying Chen, Michael Benedikt, and Evgeny Kharlamov. QUASAR: querying annotation, structure, and reasoning. In *EDBT*, pages 618–621. ACM, 2012.
- [22] Asma Souihli and Pierre Senellart. Demonstrating ProApproX 2.0: a predictive query engine for probabilistic XML. In *CIKM*, pages 2734–2736, 2012.
- [23] Fabian M. Suchanek and David Gross-Amblard. Adding fake facts to ontologies. In *WWW*, pages 421–424, 2012.

Book

- [24] Serge Abiteboul. *Sciences des données: de la Logique du premier ordre à la Toile*. Fayard, 2012. ack open <http://abiteboul.com/College/lecon.htm>.
- [25] Serge Abiteboul, Ioana Manolescu, Philippe Rigaux, Marie-Christine Rousset, and Pierre Senellart. *Web Data Management*. Cambridge University Press, 2011. ack open <http://webdam.inria.fr/Jorge>.

Others

- [26] Serge Abiteboul. Sharing distributed knowledge on the Web (invited talk). In *CSL*, pages 6–8, 2012.
- [27] Serge Abiteboul. Viewing the web as a distributed knowledge base. In *Description Logics*, 2012.
- [28] Alban Galland. *Distributed Data Management with Access Control*. PhD thesis, Université Paris-Sud, 2011.
- [29] Marilena Oita. *Deriving Semantic Objects from the Structured Web*. PhD thesis, Télécom ParisTech, October 2012.

- [30] Pierre Senellart. *Probabilistic XML: A Data Model for the Web*, June 2012. Habilitation to supervise research, Université Pierre et Marie Curie.