



Activity Report 2014

Team LINKS

Linking Dynamic Data

RESEARCH CENTER
Lille - Nord Europe

THEME
Data and Knowledge Representation
and Processing

Table of contents

1. Members	1
2. Overall Objectives	1
3. Application Domains	2
3.1. Collective Intelligence	2
3.2. Linked Bibliographic Collections	2
4. New Software and Platforms	3
4.1. QuiX-Tool Suite	3
4.2. SmartHal	3
5. New Results	4
5.1. Highlights of the Year	4
5.2. Querying Heterogeneous Linked Data	4
5.3. Managing Dynamic Linked Data	5
5.4. Linking Data Graphs	5
6. Bilateral Contracts and Grants with Industry	5
7. Partnerships and Cooperations	5
7.1. Regional Initiatives	5
7.2. National Initiatives	6
7.2.1. ANR	6
7.2.2. Competitvity Clusters	6
7.3. European Initiatives	6
7.4. International Initiatives	6
7.5. International Research Visitors	6
7.5.1. Visits of International Scientists	6
7.5.2. Visits to International Teams	6
7.5.2.1. Sabbatical programme	6
7.5.2.2. Research stays abroad	6
8. Dissemination	7
8.1. Promoting Scientific Activities	7
8.1.1. Scientific events organisation	7
8.1.2. Scientific events selection	7
8.1.3. Journal	7
8.1.4. Scientific Responsibilities	7
8.2. Teaching - Supervision - Juries	7
8.2.1. Teaching	7
8.2.2. Supervision	8
8.2.3. Juries	8
8.3. Popularization	8
9. Bibliography	8

Team LINKS

Keywords: Databases, Tree Automata, Logics

Creation of the Team: 2013 January 01.

1. Members

Research Scientists

Joachim Niehren [Team leader, Inria, Senior Researcher, HDR]

Pierre Bourhis [CNRS, Researcher]

Faculty Members

Iovka Boneva [Univ. Lille I, Associate Professor]

Angela Bonifati [Univ. Lille I, Professor, HDR]

Aurélien Lemay [Univ. Lille III, Associate Professor]

Slawomir Staworko [Univ. Lille III, Associate Professor]

Sophie Tison [Univ. Lille I, Professor, HDR]

Engineers

Denis Debarbieux [Inria, until Aug 2014, granted by Région Nord-Pas-de-Calais]

Guillaume Bagan [Ingénieur CNRS]

PhD Students

Adrien Boiret [Allocation, Univ. Lille I]

Tom Sebastian [Employé Inovimax]

Vasile-Radu Ciucanu [Allocation, Univ. Lille I]

Antoine Mbaye Ndione [Univ. Lille I, until Feb 2014]

Grégoire Laurence [Univ. Lille I, until June 2014]

Post-Doctoral Fellow

Vincent Hugot [Inria]

Visiting Scientists

Yves Roos [Univ. Lille I, Associate Professor]

Anne-Cécile Caron [Univ. Lille I, Associate Professor]

Martin Musicante [Invited Professor from Universidade Federal do Rio Grande do Norte, from Dec 2014]

Administrative Assistants

Aurore Hermant [Inria, since February 2014]

Julie Jonas [Inria, until February 2014]

2. Overall Objectives

2.1. Presentation

We will develop algorithms for answering logical querying on heterogeneous linked data collections in hybrid formats, distributed programming languages for managing dynamic linked data collections and workflows based on queries and mappings, and symbolic machine learning algorithms that can link datasets by inferring appropriate queries and mappings.

The following three paragraphs summarize our main research objectives.

Querying Heterogeneous Linked Data We will develop new kinds of schema mappings for semi-structured datasets in hybrid formats including graph databases, RDF collections, and relational databases. These induce recursive queries on linked data collections for which we will investigate evaluation algorithms, containment problems, and concrete applications.

Managing Dynamic Linked Data In order to manage dynamic linked data collections and workflows, we will develop distributed data-centric programming languages with streams and parallelism, based on novel algorithms for incremental query answering, study the propagation of updates of dynamic data through schema mappings, and investigate static analysis methods for linked data workflows.

Linking Data Graphs Finally, we will develop symbolic machine learning algorithms, for inferring queries and mappings between linked data collections in various graphs formats from annotated examples.

3. Application Domains

3.1. Collective Intelligence

¹ Links represented in the data are important for web users, who try to locate relevant information. They typically want to pose their queries locally and obtain the answers from both local and remote repositories. With the concept of linked data collections, the users are provided with a virtual collection of data and links. The answers to a query need to follow both explicit and implicit links to external repositories. Nevertheless, we argue that the benefits of links are not limited to casual users. In this paragraph, we briefly discuss two applications in which linked data collections need to be created. In the past decade, most of the enterprise data was proprietary, thus residing within the enterprise repository, along with the knowledge derived from that data. Today's enterprises need to face the problem of information explosion, due to the Internet's instability to rapidly convey large amounts of information throughout the world via end-user applications and tools. A linked data collection thus represents a virtual knowledge repository, in which relevant data is collected and meaningful mappings between this data and external world are inferred. A linked data collection would ease the task of expert users to (i) process data, metadata and knowledge that derives from machine to machine processing and eventually make sense of it; (ii) assemble and disassemble pieces of data, metadata and knowledge to create aggregate opinions or to disaggregate opinions in a way that turns to be useful in decision making; (iii) continuously learn from user feedback to produce better knowledge and to enhance the semantics of data processing.

3.2. Linked Bibliographic Collections

The second example concerns scientists who want to quickly inspect relevant literature and datasets. In such a case, local knowledge that comes from a local repository of publications belonging to a research institute (e.g. HAL) need to be integrated with other Web-based repositories, such as DBLP, Google Scholar, ResearchGate and even Wikipedia. Indeed, the local repository may be incomplete or contain semantic ambiguities, such as mistaken or missing conference venues, mistaken long names for the publication venues and journals, missing explanation of research keywords, and opaque keywords. A linked data collection would lead to build both implicit and explicit links by (i) cleaning the errors with links to correct data e.g. via mappings from HAL to DBLP for the publications errors, and via mappings from HAL to Wikipedia for opaque keywords, (ii) thoroughly complete the list of publications of the research institute, and (iii) support complex queries on the corrected data combined with external links. Links are thus useful in all scenarios, in which massive data need to be understood, analyzed and processed. Finally, they can be processed in highly distributed scenarios, such as in the cloud, where their evaluation can be run on many different sites.

¹ Collective intelligence is a shared or group intelligence that emerges from the collaboration and competition of many individuals and appears in consensus decision making" (from Wikipedia - Collective Intelligence)

4. New Software and Platforms

4.1. QuiX-Tool Suite

Participants: Joachim Niehren [correspondant], Denis Debarbieux, Tom Sebastian.

The QuiX-Tool Suite provides tools to process XML streams and documents. The QuiX-Tool Suite is based on early algorithms: query answers are delivered as soon as possible and in all practical cases at the earliest time point. The QuiX-Tool Suite provides an implementation of the main XML standart over streams. XPath, XSLT, XQuery and XProc are W3C standarts while Schematron is an ISO one. The QuiX-Tool suite is developed in the Inria transfer project QuiXProc in cooperation with Innovimax. It includes among the others existing tools such as FXP and QuixPath, along with new tools, namely X-Fun. Both, a free and a professional version are available. The ownership of QuiX-Tool Suite is shared between Inria and Innovimax. The main application of QuiX-Tool Suite is its usage in QuiXProc, an professional implementation of the W3C pipeline language XProc owned by Innovimax.

The QuiXPath language is a large fragment of XPath with full support for the XML data model. The QuiXPath library provides a compiler from QuiXPath to FXP, which is a library for querying XML streams with a fragment of temporal logic.

The X-Fun language is a functional language for defining transformations between XML data trees, while providing shredding instructions. X-Fun can be understood as an extension of Frisch's XStream language with output shredding, while pattern matching is replaced by tree navigation with XPath expressions. The QuiX-Tool suite includes QuiXSLT, which is a compiler from XSLT into a fragment of X-Fun, which can be considered as the core of XSLT. It also provides QuiXSchematron, which is a compiler from Schematron to X-Fun, and QuiXQuery, which is a compiler from XQuery to X-Fun.

QuixPath now covers 100 per cent of the XPathMark, a W3C benchmark for the language Xpath (querying XML trees). In particular, it includes aggregation operators, joins and arithmetics operations.

See also the web page <https://project.inria.fr/quix-tool-suite/>.

- Version: QuixPath v2.0.3
- Version: X-Fun v0.5.0
- Version: QuiXSLT v0.5.0
- Version: QuiXSchematron v1.0.2

4.2. SmartHal

Participants: Joachim Niehren [correspondant], Antoine Mbaye Ndione, Guillaume Bagan.

SmartHal is a better tool for querying the HAL bibliography database, while is based on Haltool queries. The idea is that a Haltool query returns an XML document that can be queried further. In order to do so, SmartHal provides a new query language. Its queries are conjunctions of Haltool queries (for a list of laboratories or authors) with expressive Boolean queries by which answers of Haltool queries can be refined. These Boolean refinement queries are automatically translated to XQuery and executed by Saxon. A java application for extraction from the command line is available. On top of this, we have build a tool for producing the citation lists for the evaluation report of the LIFL, which can be easily adapter to other Labs.

See also the web page <http://smarthal.lille.inria.fr/>.

This year, Smathal has been adapted for querying the version 3.0 of Hal. Moreover, maintenance and optimization has been proceded all over the year.

- Version: SmartHal v1.0.0

5. New Results

5.1. Highlights of the Year

In the objective Querying Heterogeneous Linked Data, Slawomir Staworko and Iovka Boneva have developed new ways to define schema for Graph Database and RDF [19]. This work has been influencing a group work of W3C on defining a schema for the DF format. This work is a continuation of [3] (by Iovka Boneva, Radu Ciucanu and Slawomir Staworko) developing a new schema for unordered trees over XML. Due to these works, Boneva is now a member of the Data Shapes Working Group which mission is to produce a language for defining structural constraints on RDF graphs. <http://www.w3.org/2014/data-shapes/charter>

In the objective Managing Dynamic Linked Data, the main break through is the development of QuixPath that now covers 100 per cent of the XPathMark, a W3C benchmark for the language Xpath (querying XML trees). In particular, it includes aggregation operators, joins and arithmetics operations. The core of QuixPath is based on techniques presented in [6] (by Tom Sebastian, Denis Debardieux and Joachim Niehren).

In the objective Linking Data Graphs, different methods have been developed to learn queries over graph. More precisely, the queries learned are conjunctive queries with joins. These techniques have been presented in [13] and demonstrated in [4] at the conference VLDB.

5.2. Querying Heterogeneous Linked Data

Angela Bonifati, Gianvito Summa, Esther Pacitt (U Montpellier 2) and Fady Draid (U Montpellier 2) [5] consider peer-to-peer data management systems (PDMS), where each peer maintains mappings between its schema and some acquaintances, along with social links with peer friends. In this context, the goal is reformulating conjunctive queries from a peer's schema into other peer's schemas. Precisely, queries against a peer node are rewritten into queries against other nodes using schema mappings thus obtaining query rewritings. They propose a new notion of 'relevance' of a query with respect to a mapping that encompasses both a local relevance (the relevance of the query w.r.t. the mapping) and a global relevance (the relevance of the query w.r.t. the entire network). Based on this notion, they conceived a new query reformulation approach for social PDMS which achieves great accuracy and flexibility. This has been implemented and experimented in a prototype.

Pierre Bourhis, Andreas Morak and Andreas Pieris [14] investigated classes of queries for which the problem of open query answering of disjunctive guarded TGDs a decent complexity (e, g exptime). The complete picture of the complexity of answering (unions of) conjunctive queries under the main guarded-based classes of disjunctive existential rules has been recently settled. It has been shown that the problem is very hard, namely 2ExpTime-complete, even for fixed sets of rules expressed in lightweight formalisms. The central objective of the present paper is to understand whether simpler query languages (bounded tree width and acyclic queries) have a positive impact on the complexity of query answering under the main guarded-based classes of disjunctive existential rules.

In [3], a new formalism for schema for unordered trees have been developed. It is based on a notion of regular expressions of multisets of labels. Different problems of static analysis like emptiness and containment are studied and their complexity. Different simpler schema are studied leading to interesting complexity for the different studied problems. Finally, they study the expressive power of the proposed schema languages and compare them with yardstick languages of unordered trees (FO, MSO, and Presburger constraints) and DTDs under commutative closure. The results show that the proposed schema languages are capable of expressing many practical languages of unordered trees and enjoy desirable computational properties.

In [7], Adrian Boiret, Vincent Hugot and Joachim Niehren and Ralf Treinen (University Paris 7) proposes a notion deterministic tree automata for unordered trees. While the existing notions are well-investigated concerning expressiveness, they all lack a proper notion of determinism, which makes it difficult to distinguish subclasses of automata for which problems such as inclusion, equivalence, and minimization can be solved efficiently. In this paper, the authors propose and investigate different notions of "horizontal determinism",

starting from automata for unranked trees in which the horizontal evaluation is performed by finite state automata.

5.3. Managing Dynamic Linked Data

Tom Sebastian, Denis Debarbieux, Olivier Gauwin (U Bordeaux), Joachim Niehren, Mohamed Zergaoui (Innovimax) [6] present new techniques to evaluate XPath queries on trees received in a streaming way. It introduces early nested word automata in order to approximate earliest query answering algorithms for nested word automata. The notion of early query answering algorithm is based on stack-and-state sharing for running early nested word automata on all answer candidates with on-the-fly determinization. These techniques allow to implement a more important part of XPath and outperform all the previous tools in coverage of XPathMark benchmark.

5.4. Linking Data Graphs

Angela Bonifati, Radu Ciucanu, Slawomir Staworko developed techniques to learn conjunctive queries from example given by a user. The main part is to infer joins between relations from the positive and negative tuples. Different techniques to deduce informative examples are presented and interestingly they can be done in polynomial time. The techniques are published in [13] and demonstrated in [4].

Grégoire Laurence, Aurélien Lemay, Joachim Niehren, Slawek Staworko, Marc Tommasi [16] explain how to learn sequential top-down tree-to-word transducers (STWs). First, they present a Myhill-Nerode characterization of the corresponding class of sequential tree-to-word transformations (STW). Next, they investigate what learning of STWs means, identify fundamental obstacles, and propose a learning model with abstain. Finally, they present a polynomial learning algorithm.

6. Bilateral Contracts and Grants with Industry

6.1. Bilateral Contracts with Industry

Innovimax, Cifre and Engineer (2010-2014) The PhD thesis of Tom SEBASTIAN within the QUIXPROC project is supervised by J. NIEHREN in cooperation with M. ZERGAOUI the head of the INNOVIMAX company. The software development in this context is supported by D. DEBARBIEUX, a senior engineer co-funded by INNOVIMAX and Inria.

7. Partnerships and Cooperations

7.1. Regional Initiatives

7.1.1. FUI Région AAP 14 Hermes(2013-2015)

Participants: Angela Bonifati [correspondent], Joachim Niehren, Iovka Boneva Denis Debarbieux

The Hermes projecton “Relation Client Personnalisée et Contextualisée” is coordinated by Bonifati from Links. Our partners are the Université Lille 1, Logos Keyneosoft, Cylande, Norsys, Numsght, Leroy Merlin, Kiabi and Auchan. The project addresses the problem of enriching the client communication within the marketing process. Starting from heterogeneous data sources (connected devices, social networks and traditional marketing channels), one has to extract the necessary information at hand. The data sources can be seen in a streaming fashion as they produce continuous data.

7.2. National Initiatives

7.2.1. ANR

7.2.1.1. ANR Aggreg

Participants: Joachim Niehren [correspondent], Pierre Bourhis, Aurelien Lemay, Adrien Boiret This project has been accepted this year and it is in collaboration with University Paris 7, University of Marseille and University of Caen. The main goal of the Aggreg project is to develop efficient algorithms for answering aggregate queries for databases and data streams of various kinds.

7.2.2. Competitivity Clusters

We participate to the following <http://www.picom.fr/> (Pôle de compétitivité PICOM - regional research cluster on commerce industries). In particular, the Hermes project has been conceived within the cluster.

7.3. European Initiatives

7.3.1. Collaborations with Major European Organizations

Partner 1: University of Oxford, Departement of Computer Science Database Group.

This collaboration is related the Inria North-European Lab Lille-Oxford. It is related to managing linked data and its exchange. If the Database Group has deep roots with Joachim Niehren and Angela Bonifati, new topics have been recently developed by younger researchers as Slawek Staworko and Pierre Bourhis.

7.4. International Initiatives

7.4.1. Inria International Partners

7.4.1.1. Declared Inria International Partners

Links is in Inria North-European Lab team with University of Oxford. The main people involved are Joachim Niehren [correspondent], Pierre Bourhis and Angela Bonifati, but the cooperation is equally relevant for Iovka Boneva, Aurélien Lemay, Slawek Staworko, Sophie Tison, Radu Ciucanu (PhD student). The Oxford database group (<http://www.cs.ox.ac.uk/isg/db>) is one of the top database groups world wide. The main persons involved will be Michael Benedikt [correspondent], Dan Olteanu, Andreas Pieris (postdoc). Further promising cooperation opportunities are to be explored with members of Georg Gottlob's ERCproject DiaDem(<http://www.cs.ox.ac.uk/projects/DIADEM/index.html>) on semantics-based information extraction.

7.5. International Research Visitors

7.5.1. Visits of International Scientists

Martin Musicante from Universidade Federal do Rio Grande do Norte has been an invited Professor since December 2014.

7.5.2. Visits to International Teams

7.5.2.1. Sabbatical programme

Slawomir Stawork has been in sabbatical at University of Edinburg for a year.

7.5.2.2. Research stays abroad

Pierre Bourhis has visited University of Oxford for more than a month in different visits over the year.

8. Dissemination

8.1. Promoting Scientific Activities

8.1.1. Scientific events organisation

8.1.1.1. General chair, scientific chair

Sophie Tison is in the scientific committee of RTA (Rewriting Techniques and Applications) and STACS (Symposium on Theoretical Aspects of Computer Science). She is also in the editorial committee of RAIRO-ITA (Theoretical Informatics and Applications).

8.1.2. Scientific events selection

8.1.2.1. Member of the conference program committee

Joachim Niehren was member of the program committees of LATA (International Conference on Language and Automata Theory and Applications) 2015, WPTE (Rewriting Techniques for Program Transformations and Evaluation) 2015, STACS (Symposium on Theoretical Aspects of Computer Science) 2014 and WPTE (Rewriting Techniques for Program Transformations and Evaluation) 2015.

Iovka Boneva was member of program committee ABDIS (East-European Conference on Advances in Databases and Information Systems) 2014

Angela Bonifati was member of the program committees of VLDB (Very large Database Conference) 2014 and 2015, EDBT (International Conference on Extending Database Technology) 2014, SIGMOD (International conference in Data Management) 2014, ICDT (International Conference on Database Theory) 2014 and 2015 and ICDE (International Conference of Data Engineering) 2015.

Slawomir Stawork is member of the program committees of WebDB (International Workshop on the Web and Databases) 2014.

8.1.3. Journal

8.1.3.1. Reviewer

The members of the team reviewed papers in numerous journal papers such as Journal of ACM, TODS, TKDE, PVLDB, Information System, Information and Computation.

8.1.4. Scientific Responsibilities

S. TISON is head of the computer science lab in Lille (LIFL). She is elected member of the “Comité National de la Recherche Scientifique (CoNRS)” (Section 6). She was president of the scientific and technical council of the cluster “Pôle de Compétitivité Industries du commerce” until June 2014. and she is member of the administrative board of the cluster Pictanovo. She have been elected member of the Administrative committee of University of Lille 1 since September 2014 She was member of the evaluation committee AERES IRIT.

Boneva is a member of the Data Shapes Working Group which mission is to produce a language for defining structural constraints on RDF graphs. <http://www.w3.org/2014/data-shapes/charter>

8.2. Teaching - Supervision - Juries

8.2.1. Teaching

License: Angela Bonifati, Introduction to Databases, 54h, L3, Université Lille 1, France
Master (Mocad): Pierre Bourhis Information extraction, 18h, M2, Université Lille 1, France
Master (Mocad): A. Bonifati Information extraction, 18h, M2, Université Lille 1, France
Master : S. Tison Advanced algorithms and complexity, M1, 57h, Université Lille 1, France
Master : S. Tison, Fouille de données, M1, 30h, Université Lille 1, France
Master: A. LEMAY, XML Technologies, 16h, M2, Université Lille 3, France
DUT : Iovka Boneva, 100h, Université Lille 1, France

Angela Bonifati is the co-responsible of the Master MOCAD of the University of Lille1.

Aurélien Lemay is pedagogical responsible for Computer Science and numeric correspondent for UFR LEA Lille 3.

8.2.2. Supervision

PhD : A. M. NDIONE, Probabilistic algorithms for tree automata and transducers. Since Sept. 2010. 16th April 2014. Supervised by Niehren and Lemay

PhD: G. LAURENCE, Normalisation et Apprentissage de Transformations d' Arbres en Mots. Since Sept 2010 4th June 2014 Supervised by Niehren and Lemay and Staworko.

PhD in progress : R.CIUCANU, Cross-model queries and schema mappings: evaluation and learning. Since Sept. 2012. Supervised by Bonifati.

PhD in progress : T. SEBASTIAN, Streaming algorithms for XSLT. Since May 2011. Supervised by Niehren.

PhD in progress: A. BOIRET, Top-down tree transformations with look-ahead : foundations and learning. Since Sept. 2011. Supervised by Niehren and Lemay.

8.2.3. Juries

HDR: Joachim Niehren was reviewer of the HDR of Pierre Senellart (Institut Télécom Paris-Tech) and comity member of Agata Savary (University of Blois).

Sophie Tison was reviewer for the HDR of E. Contejean (University South Paris) and M. Morvan (University East Paris). She was a member of the committy of B. Ben Amor.

PhD:

Joachim Niheren was president of the jury of Evguenia Kopylova and Antoine Thomas.

Angela Bonifati was member of the jury of Ioana Ileana.

Sophie Tison was reviewer of the PhD of J. Amavi and member of the jury of A. Ndione, R. Auguste, A. Aissaoui and O. Rihawi.

Slawomir Staworko was member of the Jury of Jean Decoster.

Selection Comity :

Iovka Boneva was member of the selection comity for Junior Researcher (CR2) Inria Saclay

Sophie Tison was member of Selection comity for Professor at Lille1, and Associate Professor at Lens and Marseille. She was also in the Evaluation Comity of for the pre-proposition of ANR and member of the jury IUF senior 2014.

Pierre Bourhis was member of a comity selection for Associate Professor at Rennes.

8.3. Popularization

The team has participated at “Journée RIC “ of University Lille 1. In particular, Angela Bonifati, Radu Ciucianu and Tom Sebastian presented two demonstrations related to and X-Fun.

9. Bibliography

Publications of the year

Doctoral Dissertations and Habilitation Theses

- [1] G. LAURENCE. *Normalization and Learning of Tree to Words Transductions*, Université des Sciences et Technologie de Lille - Lille I, June 2014, <https://tel.archives-ouvertes.fr/tel-01053084>

- [2] A. M. NDIONE. *Approximate Membership for Words and Trees*, Université Lille 1, April 2014, <https://tel.archives-ouvertes.fr/tel-01092549>

Articles in International Peer-Reviewed Journals

- [3] I. BONEVA, R. CIUCANU, S. STAWORKO. *Schemas for Unordered XML on a DIME*, in "Theory of Computing Systems", October 2014, 40 p. [DOI : 10.1007/s00224-014-9593-1], <https://hal.inria.fr/hal-01076329>
- [4] A. BONIFATI, R. CIUCANU, S. STAWORKO. *Interactive Join Query Inference with JIM*, in "Very Large Database Endowment - Demo Paper: PVLDB", June 2014, vol. 7, n^o 13, pp. 1541-1544, System demo, <https://hal.inria.fr/hal-01006126>
- [5] A. BONIFATI, G. SUMMA, E. PACITTI, F. DRAIDI. *Query Reformulation in PDMS Based on Social Relevance*, in "Transactions on Large-Scale Data and Knowledge-Centered Systems", March 2014, vol. LNCS, pp. 59-90 [DOI : 10.1007/978-3-642-54426-2_3], <http://hal-lirmm.crs.cnr.fr/lirmm-01088153>
- [6] T. SEBASTIAN, D. DEBARBIEUX, O. GAUWIN, J. NIEHREN, M. ZERGAOUI. *Early Nested Word Automata for XPath Query Answering on XML Streams*, in "Theoretical Computer Science", March 2015, 42 p. , <https://hal.inria.fr/hal-00966625>

International Conferences with Proceedings

- [7] A. BOIRET, V. HUGOT, J. NIEHREN, R. TREINEN. *Deterministic Automata for Unordered Trees*, in "Fifth International Symposium on Games, Automata, Logics and Formal Verification (Gandalf)", Verona, Italy, EPTCS, September 2014, vol. 161, pp. 189-202 [DOI : 10.4204/EPTCS.161.17], <https://hal.inria.fr/hal-01020236>
- [8] A. BOIRET, V. HUGOT, J. NIEHREN, R. TREINEN. *Logics for Unordered Trees with Data Constraints on Siblings*, in "LATA : 9th International Conference on Language and Automata Theory and Applications", Nice, France, March 2015, <https://hal.inria.fr/hal-01088761>
- [9] I. BONEVA, A. BONIFATI, R. CIUCANU. *Graph Data Exchange with Target Constraints*, in "EDBT/ICDT Workshops - Querying Graph Structured Data (GraphQ)", Brussels, Belgium, March 2015, <https://hal.inria.fr/hal-01095838>
- [10] A. BONIFATI, R. CIUCANU, A. LEMAY. *Interactive Path Query Specification on Graph Databases*, in "18th International Conference on Extending Database Technology (EDBT). System Demo", Brussels, Belgium, March 2015, <https://hal.inria.fr/hal-01097771>
- [11] A. BONIFATI, R. CIUCANU, A. LEMAY. *Learning Path Queries on Graph Databases*, in "18th International Conference on Extending Database Technology (EDBT)", Brussels, Belgium, March 2015, <https://hal.inria.fr/hal-01068055>
- [12] A. BONIFATI, R. CIUCANU, A. LEMAY, S. STAWORKO. *A Paradigm for Learning Queries on Big Data*, in "First International Workshop on Bringing the Value of "Big Data" to Users (Data4U)", Hangzhou, China, September 2014 [DOI : 10.1145/2658840.2658842], <https://hal.inria.fr/hal-01052676>
- [13] A. BONIFATI, R. CIUCANU, S. STAWORKO. *Interactive Inference of Join Queries*, in "17th International Conference on Extending Database Technology (EDBT)", Athens, Greece, March 2014, pp. 451-462 [DOI : 10.5441/002/EDBT.2014.41], <https://hal.inria.fr/hal-00875680>

- [14] P. BOURHIS, M. MORAK, A. PIERIS. *Towards Efficient Reasoning under Guarded-based Disjunctive Existential Rules*, in "MFCS", Budapest, Hungary, August 2014, <https://hal.inria.fr/hal-01053179>
- [15] P. LABATH, J. NIEHREN. *A Uniform Programming Language for Implementing XML Standards*, in "41st International Conference on Current Trends in Theory and Practice of Computer Science", Pec pod Sněžkou, Czech Republic, Lecture Notes in Computer Science, Springer, January 2015, <https://hal.inria.fr/hal-00954692>
- [16] G. LAURENCE, A. LEMAY, J. NIEHREN, S. STAWORKO, M. TOMMASI. *Learning Sequential Tree-to-Word Transducers*, in "8th International Conference on Language and Automata Theory and Applications", Madrid, Spain, Springer, March 2014, <https://hal.inria.fr/hal-00912969>
- [17] G. MADELAINE, C. LHOSSAINE, J. NIEHREN. *Attractor Equivalence: An Observational Semantics for Reaction Networks*, in "First International Conference on Formal Methods in Macro-Biology", Nouméa, New Caledonia, Lecture Notes in Bioinformatics, Springer-Verlag, September 2014, <https://hal.archives-ouvertes.fr/hal-00990924>
- [18] A. M. NDIONE, A. LEMAY, J. NIEHREN. *Sublinear DTD Validity*, in "9th International Conference on Language and Automata Theory and Applications", Nice, France, March 2015, <https://hal.inria.fr/hal-00803696>
- [19] S. STAWORKO, I. BONEVA, J. E. LABRA GAYO, S. HYM, E. G. PRUD'HOMMEAUX, H. SOLBRIG. *Complexity and Expressiveness of ShEx for RDF*, in "International Conference on Database Theory", Brussels, Belgium, International Conference on Database Theory (ICDT), March 2015, <https://hal.archives-ouvertes.fr/hal-01094557>

National Conferences with Proceedings

- [20] A. BONIFATI, R. CIUCANU, S. STAWORKO. *Interactive Inference of Join Queries*, in "Gestion de Données - Principes, Technologies et Applications (BDA)", Grenoble-Autrans, France, October 2014, <https://hal.inria.fr/hal-01052764>
- [21] A. BONIFATI, R. CIUCANU, S. STAWORKO. *Interactive Join Query Inference with JIM*, in "Gestion de Données - Principes, Technologies et Applications (BDA)", Grenoble-Autrans, France, October 2014, <https://hal.inria.fr/hal-01052789>

Conferences without Proceedings

- [22] P. BOURHIS, M. MORAK, A. PIERIS. *Acyclic Query Answering Under Guarded Disjunctive Existential Rules and Consequences to DLs*, in "Informal Proceedings of the 27th International Workshop on Description Logics", Vienne, Austria, July 2014, <https://hal.inria.fr/hal-01098983>

Research Reports

- [23] P.-C. HÉAM, V. HUGOT, O. KOUCHNARENKO. *The Emptiness Problem for Tree Automata with at Least One Disequality Constraint is NP-hard*, FEMTO-ST, December 2014, <https://hal.inria.fr/hal-01089711>
- [24] M. SCHMIDT-SCHAUSS, D. SABEL, J. NIEHREN, J. SCHWINGHAMMER. *Observational Program Calculi and the Correctness of Translations*, May 2014, <https://hal.inria.fr/hal-00824349>

Scientific Popularization

-
- [25] P. BOURHIS, M. KRÖTZSCH, S. RUDOLPH. *How to Best Nest Regular Path Queries*, in "Informal Proceedings of the 27th International Workshop on Description Logics", Vienne, Austria, July 2014, <https://hal.inria.fr/hal-01098979>

Other Publications

- [26] P. BOURHIS, M. KRÖTZSCH, S. RUDOLPH. *Query Containment for Highly Expressive Datalog Fragments*, July 2014, <https://hal.inria.fr/hal-01098974>