

Curriculum Vitae – Bogdan Cautis

Professor, head of the LaHDAK research team

Computer Science Department, University of Paris-Sud 11 (I.U.T. Orsay)

Address: University Paris-Sud 11, bâtiment 650 (PCRI), 91405 Orsay, France / Office: 271
Phone: +33-172-92-59-84
Email: bogdan.cautis@u-psud.fr, cautis@lri.fr
Homepage: www.lri.fr/~cautis/
Research team page: <https://lahdak.lri.fr>
Birth: 26/09/1979, Constanta (Romania)
Citizenship: Romanian and French

Research interests: data mining, social networks and graph mining, information retrieval, database theory and systems, uncertain data, information extraction, knowledge extraction and knowledge management.

EDUCATION

Habilitation to Supervise Research (HdR), March 2012

University Pierre et Marie Curie, Paris
Dissertation: *Web Data Management: Weaving Theory and Applications*

Ph.D. in Computer Science, October 2004 - September 2007

INRIA & University of Paris South, Orsay
Dissertation: *Signing and Reasoning about Tree Updates*
Research advisors: Serge Abiteboul (INRIA) and Tova Milo (University of Tel Aviv)

M.Sc. in Computer Science – D.E.A. Algorithmique (Logics, Verification), October 2003 - September 2004

Ecole Polytechnique (Palaiseau, France)

International Program, October 2002 - September 2003

Ecole Polytechnique (Palaiseau, France)

Engineering Degree in Computer Science, October 1998 – February 2004

Politehnica University Bucharest (Romania) / Università Degli Studi di Brescia (Italy)

EXPERIENCE

Huawei Noah's Ark Lab, Hong Kong, September 2015 – August 2017

Senior Researcher, on extended leave of absence from University of Paris-Sud 11

University of Paris-Sud 11 and INRIA Saclay, France, September 2013 – present

Full Professor in “Data Management and Data Mining” (permanent position)

Telecom ParisTech (formerly known as ENST), France, October 2007 – August 2013

Associate Professor in “Web Data Management” (permanent position)

UC San Diego, Mai 2006 – July 2006, Sept. 2010, April 2013 / University of Tel Aviv, Mai 2005

Visiting researcher

University of Cergy-Pontoise October 2005 – August 2007

Teaching assistant (moniteur à l’enseignement)

RESEARCH WORK (recently funded projects)

- ALICIA: Adaptive learning for intelligent crowdsourcing and information access (ANR – PI, coordinator)
- KEYSTONE: Semantic keyword-based search on structured data sources (EU COST – WG co-chair)
- X-DATA: Data integration for analysis and prediction (funded by the French Ministry of Industry)
- ARCOMEM: Leveraging the wisdom of the crowds and the social Web for intelligent preservation (FP7)

RESEARCH COLLABORATIONS

UC San Diego, Yahoo Research, Pandora Research, Northeastern University, Renmin University, Free University of Bolzano, Hong Kong University, National University of Singapore, Chinese University of Hong Kong, University of Helsinki

PHD SUPERVISION

Nora Derouiche – “Targeted Information Extraction” – started in October 2008 – defended in March 2012

Silviu Maniu – “Data Management in Social Networks” – started in October 2009 – defended in September 2012

Georges Gouriten – “Knowledge-based Content Suggestions on the Social Web” – started in October 2012, dropped out

Ioana Ileana – “Extracting and Archiving Web Data: a holistic approach” – started in October 2011 - defended in October 2014

Raphael Bonaque – “Social-aware Top-k Keyword Search Over Rich Web Content” – started October 2013 - defended in October 2016

Paul Lagrée – “Adaptive Learning for User-Centric Applications” – started in October 2014 - defended in October 2017

Yangke Sun – “Influence and Diffusion-Aware Group Recommendation in Social Media” – to start on October 1st 2018

TEACHING ACTIVITIES

Teaching responsibility for courses (in French or English) in the broad area of databases, data mining, information retrieval and the Web, for 150+h per year, undergraduate / master level, at Telecom ParisTech, Ecole Polytechnique, U. Paris-Sud, Renmin U., HKU.

PROFESSIONAL SERVICE

PC member: EDBT 2011, 2014, ICDT 2012, VLDB 2012, SIGMOD 2013-2014, 2017, ICDE 2014, 2017, 2019, CIKM 2014, 2016;
Journal reviewer: ACM TODS, VLDB Journal, IEEE TKDE, Information Systems.

CONFERENCE AND WORKSHOP PUBLICATIONS

- *Large-Scale Online Influence Maximization*. with Paul Lagrée, Olivier Cappé, and Silviu Maniu, IEEE International Conference on Data Mining (ICDM), 2017, New Orleans, USA.
- *KB-Enabled Query Recommendation for Long-Tail Queries*. with Zhipeng Huang, Reynold Cheng, and Yudian Zheng, the ACM International Conference on Information and Knowledge Management (CIKM) 2016 Indianapolis, USA.
- *Social, Structured and Semantic Search*. with Raphaël Bonaque, François Goasdoué, and Ioana Manolescu, International Conference on Extending Database Technology, EDBT 2016, Bordeaux, France.
- *A Network-Aware Approach for Searching As-You-Type in Social Media*. with Paul Lagrée and Hossein Vahabi, the ACM International Conference on Information and Knowledge Management (CIKM) 2015, Melbourne, Australia.
- *Snooping Wikipedia vandals with MapReduce*. with Michele Spina, Dario Rossi, Mauro Sozio, Silviu Maniu, ICC 2015.
- *Quantifying Trust Dynamics in Signed Graphs, the S-Cores Approach*. with Christos Giatsidis, Silviu Maniu, Dimitrios M. Thilikos, and Michalis Vazirgiannis, SIAM International Conference on Data Mining (SDM) 2014.
- *Complete yet Practical Search for Minimal Query Reformulations under Constraints*. with Ioana Ileana, Alin Deutsch, and Yannis Katsis, ACM SIGMOD International Conference on Management of Data (SIGMOD) 2014, Salt Lake City, USA.
- *Context Aware Top-k Processing using Views*. with Silviu Maniu, ACM International Conference on Information and Knowledge Management (CIKM) 2013, San Francisco, USA.
- *Network-aware Search in Social Tagging Applications: Instance Optimality versus Efficiency*. with Silviu Maniu, ACM International Conference on Information and Knowledge Management (CIKM) 2013, San Francisco, USA.
- *Automatic Extraction of Structured Web Data with Domain Knowledge*. with Talel Abdessalem, and Nora Derouiche, 28th International Conference on Data Engineering (IEEE ICDE) 2012 Washington DC, USA.
- *Casting a Web of Trust over Wikipedia: an Interaction-based Approach*. with Silviu Maniu and Talel Abdessalem, 20th International Conference on World Wide Web, WWW 2011, Hyderabad, India.
- *Challenges for View-Based Query Answering over Probabilistic XML*. with Evgeny Kharlamov, Proceedings of the 5th Alberto Mendelzon International Workshop on Foundations of Data Management (AMW) 2011, Santiago, Chile.
- *Building a Signed Network from Interactions in Wikipedia*. with Silviu Maniu and Talel Abdessalem Proceedings of the First ACM SIGMOD Workshop on Databases and Social Networks, DBSocial 2011, Athens, Greece, 2011.
- *Querying Data Sources that Export Infinite Sets of Views*. with Alin Deutsch and Nicola Onose, 12th International Conference on Database Theory (ICDT), St. Petersburg, Russia, 2009.
- *Pruning Nested XQuery Queries*. with Bilel Gueni, Talel Abdessalem and Emmanuel Waller, The 17th ACM Conference on Information and Knowledge Management, CIKM 2008, Napa Valley, USA.
- *XPath Rewriting Using Multiple Views: Achieving Completeness and Efficiency*. with Alin Deutsch and Nicola Onose, 11th International Workshop on the Web and Databases, WebDB 2008, Vancouver, Canada.
- *Reasoning about XML Update Constraints*. with Serge Abiteboul and Tova Milo, The ACM SIGMOD-SIGACT-SIGART Symposium on Principle of Databases (PODS), Pekin, China, 2007.
- *Distributed Access Control: A Privacy-conscious Approach*. ACM Symposium on Access Control Models and Technologies (SACMAT 07).
- *Digital Signatures for Modifiable Collections*. with Serge Abiteboul, Amos Fiat and Tova Milo, The International Conference on Availability, Reliability and Security (ARES), Vienne, Austria, 2006.
- *Lazy Query Evaluation for Active XML*. with Serge Abiteboul, Omar Benjelloun, Ioana Manolescu, Tova Milo and Nicoleta Preda, The ACM SIGMOD International Conference on Management of Data (SIGMOD), Paris, 2004.

JOURNAL PUBLICATIONS

- *Model-free Inference of Diffusion Networks using Reproducing Kernel Hilbert Space Embeddings*. with Shoubo Hu, Zhitang Chen, Laiwan Chan, Yanhui Geng, and Xiuqiang He, to appear in the Springer Data Mining and Knowledge Discovery Journal (DAMI) 2019.
- *Entity Based Query Recommendation for Long-Tail Queries*. with Zhipeng Huang, Reynold Cheng, Yudian Zheng, Nikos Mamoulis, and Jing Yan, to appear in ACM Transactions on Knowledge Discovery from Data (ACM TKDD) 2019.
- *Algorithms for Online Influencer Marketing*. with Paul Lagrée, Olivier Cappé, and Silviu Maniu, to appear in ACM Transactions on Knowledge Discovery from Data (ACM TKDD) 2019.
- *As-You-Type Social Aware Search*. with Paul Lagrée, Hossein Vahabi, ACM Transactions on Intelligent System and Technology (TIST), 2017.
- *Answering Queries using Views over Probabilistic XML: Complexity and Tractability*. with Evgeny Kharlamov, Proceedings of the VLDB Endowment (PVLDB), Volume 5, 2012.
- *Querying XML Data Sources that Export Very Large Sets of Views*. with Alin Deutsch, Nicola Onose and Vasilis Vassalos, ACM Transactions on Database Systems (TODS), 36(1):5 (2011).
- *Efficient Rewriting of XPath Queries Using Query Set Specification*. with Alin Deutsch, Nicola Onose, and Vasilis Vassalos, Proceedings of the VLDB Endowment (PVLDB), Volume 2, 2009.
- *Querying Data Sources that Export Infinite Sets of Views*. with Alin Deutsch and Nicola Onose, Theory of Computing Systems (TCS) 49(2) : 367-428 (2011), Special Issue “Best papers of ICDT 2009”.
- *Reasoning about XML Update Constraints*. with Serge Abiteboul and Tova Milo, Journal of Computer and System Sciences (JCSS) 2009.

DEMONSTRATIONS AND TUTORIALS

- *Multi-model Databases and Tightly Integrated Polystores: Current Practices, Comparisons, and Open Challenges*. With Jiaheng Lu and Irena Holubova, 3-hour tutorial at the ACM International Conference on Information and Knowledge Management (CIKM), 2018, Turin, Italy.
- *Mixed-instance Querying: Lightweight Integration Architecture for Data Journalism*. with R. Bonaque, T. D. Cao, F. Goasdoué, J. Letelier, I. Manolescu, O. Mendoza, S. Ribeiro, X. Tannier, M. Thomazo. Demonstration in the Proceedings of the VLDB Endowment (PVLDB), 2016.
- *Taagle: Efficient, Personalized Search in Collaborative Tagging Networks*. with Silviu Maniu, Demonstration in the ACM SIGMOD International Conference on Management of Data (SIGMOD), Scottsdale, USA, 2012.
- *LotusX: A Position-Aware XML Graphical Search System with Auto-Completion*. with Chunbin Lin, Jiaheng Lu and Tok Wang Ling, Demonstration in the 28th International Conference on Data Engineering (IEEE ICDE) 2012, Washington DC, USA.
- *ObjectRunner: Lightweight, Targeted Extraction and Querying of Structured Web Data*. with Talel Abdessalem and Nora Derouiche, Demonstration in the Proceedings of the VLDB Endowment, Volume 3, 2010, Singapore.
- *An Electronic Patient Record on Steroids: Distributed, Peer-to-Peer, Secure and Privacy-conscious*. with Serge Abiteboul, Bogdan Alexe, Omar Benjelloun, Irini Fundulaki, Tova Milo and Arnaud Sahuguet, Demonstration in the International Conference on Very Large Data Bases (VLDB), Toronto, Canada, 2004.