

# Pierre Senellart

Professor in computer science

Data management, Web, AI

32 villa de Lourcine  
75014 Paris, France

☎ (+33) 6-48-36-36-20

✉ pierre@senellart.com

<https://pierre.senellart.com/>

French, born 19 June 1981.



## Academic career

- September 2016 – Present  
**Professor**, *École normale supérieure, Université PSL*, Paris, France.  
*Computer Science Department (ENS, CNRS, Inria)*. On partial secondment to Institut universitaire de France (2020–2025).
- December 2017 – December 2020  
**Adjunct professor**, *Télécom Paris, IP Paris*, Paris, France.  
*Network and Computer Science* department.
- August 2014 – July 2016  
**Senior research fellow**, *National University of Singapore*, Singapore, Singapore.  
*Computer Science* department, IPAL laboratory.
- January 2013 – August 2016  
**Professor**, *Télécom ParisTech*, Paris, France.  
*Network and Computer Science* department.
- November 2012 – April 2013  
**Visiting researcher and lecturer**, *The University of Hong Kong*, Hong Kong, China.  
Computer science department
- June 2008 – December 2012  
**Associate professor**, *Télécom ParisTech*, Paris, France.  
*Network and Computer Science* department.
- September 2008 – December 2008  
**Teaching assistant**, *École polytechnique*, Palaiseau, France.  
Database labs.
- February 2008 – May 2008  
**Teaching assistant**, *École normale supérieure de Cachan*, Cachan, France.  
Database labs.
- February 2008 – May 2008  
**Post-doctoral researcher**, *Max-Planck-Institut für Informatik*, Saarbrücken, Germany.  
*Databases and Information Systems* department.
- September 2005 – May 2008  
**PhD candidate and teaching assistant**, *Université Paris–Sud & INRIA Futurs*, Orsay, France.  
Topic: Understanding the hidden Web. Advisor: Serge Abiteboul. 64hr of teaching a year.
- March 2007 – May 2007  
**Visiting research student**, *Oxford University, Department of computer science*, Oxford, United Kingdom.  
Topic: Data exchange Advisor: Georg Gottlob.
- 2003 – 2005  
**Part-time teaching assistant**, *École nationale des ponts et chaussées*, Marne-la-Vallée, France.  
C++ course for first-year students (2 semesters, 60 hour each).
- March 2003 – August 2003  
**Research intern (master)**, *INRIA Futurs*, Orsay, France.  
Topic: Identifying logical websites. Advisor: Serge Abiteboul.
- February 2002 – June 2002  
**Research intern**, *Chalmers Tekhniska Högskola*, Göteborg, Sweden.  
Topic: Automatic verification of multipliers. Advisor: Mary Sheeran.
- June 2001 – July 2001  
**Research intern**, *Université catholique de Louvain*, Louvain-la-neuve, Belgium.  
Topic: Automatic extraction of similar words in a dictionary. Advisor: Vincent Blondel.

## Education

- June 2012 **Habilitation to supervise research in Computer science**, *Université Pierre et Marie Curie*, Paris, France.  
Topic: Probabilistic XML: A Data Model for the Web
- December 2007 **PhD in Computer science**, *Université Paris–Sud*, Orsay, France.  
Topic: Understanding the Hidden Web. Mention Très Honorable.
- September 2003 **Magistère (postgraduate diploma) in Fundamental and applied mathematics and Computer science**, *Université Paris–Sud & ENS*, Orsay & Paris, France.
- September 2003 **DEA (M.Sc.) in Computer science**, *Université Paris–Sud*, Orsay, France.  
Mention Très Bien (summa cum laude).
- September 2001 **Maîtrise (undergraduate diploma) in Computer science**, *Université Paris–Sud & ENS*, Orsay & Paris, France.  
Mention Très Bien (summa cum laude).
- June 2001 **Licence (B.Sc.) in Computer science**, *Université Paris–Sud & ENS*, Orsay & Paris, France.  
Mention Très Bien (summa cum laude).
- 2000 – 2005 **École normale supérieure (Paris)**.
- June 1998 **Baccalauréat S (scientific high school diploma)**, *Lycée Mariette*, Boulogne-sur-mer, France.  
Mention Très Bien (summa cum laude).

## Young researcher supervision

- PhD students
- Aryak Sen (with Silviu Maniu, since 2024). *Scalability of a data provenance and probability management system*.
- Emmanouil Sylligardos (with Paul Boniol, since 2024). *Accuracy and execution time trade-off in ensembling and model selection for time series analytics*.
- Antoine Gauquier (with Ioana Manolescu, since 2023). *Intelligent construction of a multimodal and heterogeneous data warehouse, with data traceability*.
- Pratik Karmakar (with Stéphane Bressan, since 2023). *Quality, uncertainty, and lineage of data*.
- Shrey Mishra (graduated in 2024). *Multimodal Extraction of Proofs and Theorems from the Scientific Literature*.
- Baptiste Lafosse (with Jean-Marie Lagniez, 2021 – 2023). *Compiler dedicated to the evaluation of SQL queries*.
- Yann Ramusat (with Silviu Maniu, graduated in 2022). *The Semiring-Based Provenance Framework for Graph Databases*.
- Debabrota Basu (with Stéphane Bressan, graduated in 2018). *Learning to make decisions with incomplete information: reinforcement learning, information geometry, and real-life applications*.
- Mikaël Monet (with Antoine Amarilli, graduated in 2018). *Combined complexity of probabilistic query evaluation*. [BDA 2019 PhD award](#).
- Miyong Han (graduated in 2018). *Reinforcement learning approaches in dynamic environments*.
- Quentin Lobbé (with Dana Diminescu, graduated in 2018). *Archives, Web fragments, and diasporas. For a disaggregated exploration of web archives related to online representations of diasporas*. [Inathèque research award 2019](#).
- David Montoya (with Serge Abiteboul, graduated in 2017). *A personal knowledge base integrating user data and activity timeline*.
- Su Yang (with Serge Abiteboul, 2016 – 2017). *Semantic extraction of personal information, integration, and querying*.

Antoine Amarilli (with Pierre Bourhis, graduated in 2016). *Leveraging the structure of uncertain data*. 2017 E. W. Beth Dissertation Prize in Logic, Language, and Information. Télécom ParisTech 2017 PhD award.

M. Lamine Ba (with Talel Abdesslem, graduated in 2015). *Harnessing uncertain data structure*.

Muhammad Faheem (graduated in 2014). *Intelligent content acquisition in Web archiving*.

Ioana Ileana (with Bogdan Cautis, graduated in 2014). *Query rewriting using views: a theoretical and practical perspective*.

Georges Gouriten (with Bogdan Cautis, 2012 – 2013). *Knowledge-based content suggestions on the social Web*.

Asma Souihli (graduated in 2012). *Querying probabilistic XML*.

Marilena Oita (graduated in 2012). *Deriving semantic objects from the structured Web*.

Evgeny Kharlamov (with Werner Nutt, graduated in 2010). *A probabilistic approach to XML data management*.

Postdoctoral researchers Shufan Jiang (2022 – 2023).

Nofar Carmeli (2021 – 2022).

Ashish Dandekar (2019 – 2020).

Roxana Horincar (2013 – 2015).

Fabian M. Suchanek (with Serge Abiteboul, 2010 – 2011).

Other Supervision of 10 Master's students and 7 Engineering students on long-term projects

## Contract research projects

- 2021 – 2026 **DesCartes**, CNRS@CREATE Program on Intelligent Modelling for Decision-making in Critical Urban Systems.  
Investigator, as part of a 50 million SGD project funded by Singapore's NRF.
- 2021 – 2024 **Dissemin**, *Development of an open-science platform*.  
Principal investigator, with funding 124 kEUR.
- 2019 – 2024 **PRAIRIE**, *Paris Artificial Intelligence Research Institute*.  
Chair holder, with funding 600 kEUR.
- 2018 – 2021 **BioQOP**, *ANR collaborative grant on optimizing queries to video surveillance data under privacy constraints*.  
Responsible for ENS, with funding 69 kEUR.
- 2013 – 2015 **XData**, *Collaboration grant with industrials on the acquisition of social network data and integration with company data*.  
Responsible for Télécom, with funding 175 kEUR.
- 2013 – 2015 **CCIPX**, *STIC-Asia collaborative project between Institut Mines–Télécom, Hanoi University of Science and Technology, and National University of Singapore*.  
Principal co-investigator. Funded 26 kEUR.
- 2012 – 2017 **MoDaS**, *Tova Milo's ERC advanced grant on Mob Data Sourcing*.  
Collaboration on research papers, exchange of students.
- 2012 – 2013 **Private Recommendation System in Social Networks**, *Campus France PROCORE France–Hong Kong partnership*.  
Principal co-investigator. Funded 8 kEUR.
- 2010 – 2014 **Diadem**, *Georg Gottlob's ERC advanced grant on Domain-centric Intelligent Automated Data Extraction Methodology*.  
Collaboration on research papers, exchange of students.

- 2011 – 2013 **ARCOMEM**, *European Union FP7 project on the Archiving of the social Web*.  
Leader of the work package on content acquisition ; responsible for Télécom, with funding 600 kEUR.
- 2009 – 2013 **Webdam**, *Serge Abiteboul's ERC advanced grant on the Foundations of Web Data Management*.  
Leader of the research axis on uncertain data; management board (with S. Abiteboul and V. Vianu).
- 2010 – 2012 **PANIC**, *ANR CONTINT project on the Proactivity of audience and the digitization of cultural industries*.  
Collaborator.
- 2009 – 2011 **DataRing**, *ANR VERSO project on P2P data sharing for online communities*.  
Leader of the work package on data modeling; responsible for Télécom, with funding 120 kEUR.

## Teaching activities

- Created courses** Algorithms and applications (L2, Université PSL); Data wrangling and data quality (M2, Université PSL); Anonymization, privacy (M2, Université PSL); Databases (L3, ENS); Data Scientist (professional training, Télécom ParisTech); ACM-ICPC training (problem-solving course leading to a programming contest, Télécom ParisTech); Web search (intensive course over one or several weeks, Télécom ParisTech & Renmin University of China); Web technologies (high-level course for business and engineering students, Télécom ParisTech & HEC); L<sup>A</sup>T<sub>E</sub>X (two-day intensive course for graduate students, EDITE de Paris); Web programming (introductory course, Université Paris-Sud)
- Courses with strong involvement** Algorithms (L3, ENS); Formal languages (L3, Télécom ParisTech); Web engineering (M1, Télécom ParisTech); Web data management (M2, Université Paris-Sud, Télécom-ParisTech, MPRI, & The University of Hong Kong); Database application design (National University of Singapore); Information theory (National University of Singapore)
- Participation in other courses** C++ programming (École des Ponts ParisTech); Java programming (Télécom ParisTech); Databases (Université Paris-Sud, Télécom ParisTech, & ENS Cachan); Collective intelligence (Télécom ParisTech); Operating systems (Université Paris-Sud)
- Teaching administration** Creation and management team of the IASD Master 2 (since 2019); Representative of Télécom ParisTech and then of ENS in the teaching committee of the MPRI Master 2 (2013–2018); Creation and management (with F. d'Alché) of the Data Science study track at Télécom ParisTech (2014–2016); Creation (with S. Cléménçon) and management (with S. Cléménçon, then F. d'Alché) of the Certificat d'Étude Spécialisé (CES) Data Scientist at Télécom ParisTech (2014–2016)

## Administrative duties and responsibilities

- 2021 – 2026 **President of section 6 (computer science) of the national committee for scientific research**, *Comité national de la recherche scientifique*.  
Evaluation of CNRS researchers and laboratories, recruitment and promotion juries for CNRS researchers, other evaluation activities. Also member of the board of the CPCN.
- 2018 – 2020 **Research fellow**, *CERRE*.  
European think tank on the regulation of network industries
- 2018 – Present **Management of the entrance competition in computer science at ENS**, *Concours X/ENS*.  
Selection of examiners, of tests, in agreement with the representatives of the three other ENS schools and of École polytechnique, for the main ENS entrance competition. Management, selection, and examinations (with D. Naccache) for the alternative recruitment tracks (international selection, university track).
- 2017 – Present **Deputy director**, *DI ENS*.  
The DI ENS is both a teaching department and a mixed CNRS/ENS/Inria research unit; I am assisting the director (D. Pointcheval) in all matters related to the budget of the department, management of courses, recruitment of personnel, approval of expenses, etc.

- 2017 – 2021 **Board member, DIM RFSI.**  
The DIM RFSI is a regional alliance of computer science laboratories, allocating funding for research.
- 2017 – 2021 **Chief judge, co-organizer, secretary, ICPC SWERC.**  
SWERC is the Southwestern Europe regional competition of the International Collegiate Programming Competition, the largest programming competition organized by and for universities. I have been co-organizing it with colleagues at Télécom ParisTech, and was the chief judge and problem setter until 2020. We managed to attract record numbers of teams (76 in 2017, 89 in 2018, 98 in 2019–2020) from the participating countries (France, Italy, Spain, Portugal, Switzerland, Israel).
- 2016 – Present **Team leader, Valda.**  
I founded and am in charge of the Valda team, an ENS/CNRS/Inria team within DI ENS, growing the team from 2 to 7 permanent members in three years.
- 2016 – 2021 **Member of the board of section 6 (computer science) of the national committee for scientific research, Comité national de la recherche scientifique.**  
Evaluation of CNRS researchers and laboratories, recruitment and promotion juries for CNRS researchers, other evaluation activities.
- 2014 – 2016 **Designer of the entrance test in computer science, Concours commun Mines–Ponts.**  
Entrance competition for MP-Info classe préparatoire students.

## Service to the scientific community

- (Co-)chair of PC(Co-)président de track(Co-)chair of track  
BDA 2017; WebDB 2017; SUM 2016; UnCrowd 2014; DBCrowd 2013; BDA 2016 (demonstrations); VLDB 2015 (tutorials); DASFAA 2015 (tutorials); DEXA 2015 (tutorials); ICDE 2011 (tutorials); EDBT 2011 (industrial track); SIGMOD 2010 (programming contest)
- Program committees  
PODS 2012, 2016, 2019 & 2021; STACS 2020; EDBT 2013 (demonstrations), 2014 & 2018; ICDT 2011 & 2017; SIGMOD 2008 & 2009 (repeatability committee), 2013 & 2017 (regular PC member), 2015 (group leader), 2023; CIKM 2010 & 2014; VLDB 2013; WWW 2009 & 2010; ICDE 2010; ECML/PKDD 2008; BDA 2010, 2012, 2018, 2020, 2022; various workshops and minor conferences
- Editorial responsibilities  
Information director of the *Journal of the ACM* (2009–2015); Guest co-editor for a special issue of *Distributed and Parallel Databases* (2014)
- Journal refereeing activities  
*Journal of the ACM*; *Journal of Computer Science and Systems*; *VLDB Journal*; *Transactions on Database Systems*; *Information Systems*; *Transactions on Knowledge and Data Engineering*; *Data and Knowledge Engineering*; *Communications of the ACM*; ...
- Habilitation thesis committees  
Charles Paperman (reviewer, Université de Lille, 2024); Pierre Bourhis (president, Université de Lille, 2024); Stefan Mengel (reviewer, Université d'Artois, 2021); Federico Ulliana (reviewer, Université de Montpellier, 2021); Paolo Papotti (reviewer, Université Côte d'Azur, 2019); Raja Chiky (reviewer, Université Pierre-et-Marie-Curie, 2016)

PhD thesis committees	Edwige Cyffers (president, Université de Lille, 2024); Claire Soyez-Martin (président, Université de Lille, 2023); Pierre Faure–Giovagnoli (president, INSA Lyon, 2023); Rébecca Zucchini (reviewer, Université Paris-Saclay, 2023); Antonia Kormpa (reviewer, University of Oxford, 2023); Gaston Zanitti (reviewer & president, Université Paris-Saclay, 2023); Nicolas Crosetti (reviewer, Université de Lille, 2023); Sajad Nazari (reviewer, INSA Centre Val de Loire, 2022); Amit Kumar (president, Université Caen Normandie, 2021); Valentin Iovene (reviewer, Université Paris-Saclay, 2021); Wissam Mammam Kouadri (reviewer, Université Paris-Saclay, 2021); Julien Romero (president, IP Paris, 2020); Ugo Comignani (reviewer, Université Claude Bernard Lyon 1, 2019); Ngurah Agus Sanjaya Er (president, Télécom ParisTech, 2018); Govind (president, Université Caen Normandie, 2018); Paul Lagrée (reviewer, Université Paris-Saclay, 2017); Manuel Pozo (president, CNAM, 2016); Mira Sarkis (president, Télécom ParisTech, 2016); Mohammadreza Khelgati (reviewer, Universiteit Twente, 2016); Ahmad Assaf (Télécom ParisTech, 2015); François Rousseau (École polytechnique, 2015); Disheng Qiu (reviewer, Università degli Studi Roma Tre, 2015); Andrés Sanoja (Université Pierre-et-Marie-Curie, 2015); Tom Ameloot (Hasselt University, 2013); Zeynep Pehlivan (Université Pierre-et-Marie-Curie, 2013); Federico Ulliana (reviewer, Université Paris-Sud, 2012); Andrew Sellers (reviewer, University of Oxford, 2011)
Other committees	Gilles Kahn PhD Thesis Award of Société Informatique de France (2019, 2020, 2021); BDA PhD Thesis Award in Data Management (2019)

## Awards

- Junior member of Institut universitaire de France (2020–2025)
- Distinguished SIGMOD 2017 PC Member
- ACM HyperText 2014 Douglas Engelbart best paper award
- Adviser of Clément Genzmer, winner of the SIGMOD Programming Contest 2009

## Publications

We mention here international publications only; for a complete list (including national conferences and journals, technical reports, unrefereed material, preprints), see <http://pierre.senellart.com/publications/>.

### Book

Serge Abiteboul, Ioana Manolescu, Marie-Christine Rousset, Philippe Rigaux, and Pierre Senellart. *Web Data Management*. Cambridge University Press, New York, USA, January 2012.

### Book chapters (peer-reviewed)

Thomas Risse, Wim Peters, Pierre Senellart, and Diana Maynard. Documenting Contemporary Society by Preserving Relevant Information from Twitter. In Katrin Weller, Axel Bruns, Jean Burgess, Merja Mahrt, and Cornelius Puschmann, editors, *Twitter and Society*, pages 207–219. Peter Lang International Academic Publishers, November 2013.

Evgeny Kharlamov and Pierre Senellart. Modeling, Querying, and Mining Uncertain XML Data. In Andrea Tagarelli, editor, *XML Data Mining: Models, Methods, and Applications*, pages 29–52. IGI Global, November 2011.

### Book chapters (without peer review)

Pierre Senellart. On the Impact of Provenance Semiring Theory on the Design of a Provenance-Aware Database System. In Antoine Amarilli and Alin Deutsch, editors, *The Provenance of Elegance in Computation — Essays dedicated to Val Tannen*, pages 9:1–9:10. Schloss Dagstuhl — Leibniz-Zentrum für Informatik, March 2024.

Benny Kimelfeld and Pierre Senellart. Probabilistic XML: Models and Complexity. In Zongmin Ma and Li Yan, editors, *Advances in Probabilistic Databases for Uncertain Information Management*, pages 39–66. Springer-Verlag, May 2013.

Michael Benedikt and Pierre Senellart. Databases. In Edward K. Blum and Alfred V. Aho, editors, *Computer Science. The Hardware, Software and Heart of It*, pages 169–229. Springer-Verlag, January 2012.

Pierre Senellart and Vincent D. Blondel. Automatic discovery of similar words. In Michael W. Berry and Malu Castellanos, editors, *Survey of Text Mining II: Clustering, Classification and Retrieval*, pages 25–44. Springer-Verlag, January 2008.

Pierre Senellart and Vincent D. Blondel. Automatic discovery of similar words. In Michael W. Berry, editor, *Survey of Text Mining: Clustering, Classification and Retrieval*, pages 25–43. Springer-Verlag, August 2003.

#### Journal articles (peer-reviewed)

Pratik Karmakar, Mikaël Monet, Pierre Senellart, and Stéphane Bressan. Expected Shapley-Like Scores of Boolean Functions: Complexity and Applications to Probabilistic Databases. *Proceedings of the ACM on Management of Data*, 2(2), June 2024. Presented at the PODS 2024 conference in Santiago, Chile. Also CoRR abs/2401.06493.

Sihem Amer-Yahia, Senjuti Basu Roy, Lei Chen, Atsuyuki Morishima, James Abello Monedero, Pierre Bourhis, François Charoy, Marina Danilevsky, Gautam Das, Gianluca Demartini, Shady Elbassuoni, David Gross-Amblard, Emilie Hoareau, Munenari Inoguchi, Jared Kenworthy, Itaru Kitahara, Dongwon Lee, Yunyao Li, Ria Mae Borrromeo, Paolo Papotti, Raghav Rao, Sudeepa Roy, Pierre Senellart, Keishi Tajima, Saravanan Thirumuruganathan, Marion Tommasi, Kazutoshi Umemoto, Andrea Wiggins, and Koichiro Yoshida. Making AI Machines Work for Humans in FoW. *SIGMOD Record*, 49(2):30–35, December 2020.

Antoine Amarilli, Florent Capelli, Mikaël Monet, and Pierre Senellart. Connecting Knowledge Compilation Classes and Width Parameters. *Theory of Computing Systems*, 64(5):861–914, 2020. Also CoRR abs/1811.02944.

Michael Benedikt, Pierre Bourhis, Georg Gottlob, and Pierre Senellart. Monadic Datalog, Tree Validity, and Limited Access Containment. *ACM Transactions on Computational Logic*, 21(1):6:1–6:45, 2020.

Antoine Amarilli, Pierre Bourhis, Mikaël Monet, and Pierre Senellart. Evaluating Datalog via Tree Automata and Cycluits. *Theory of Computing Systems*, 63(7):1620–1678, 2019. Also CoRR abs/1808.04663.

Antoine Amarilli, M. Lamine Ba, Daniel Deutch, and Pierre Senellart. Computing Possible and Certain Answers over Order-Incomplete Data. *Theoretical Computer Science*, 797:42–76, 2019. Also CoRR abs/1801.06396.

Silviu Maniu, Reynold Cheng, and Pierre Senellart. An Indexing Framework for Queries on Probabilistic Graphs. *ACM Transactions on Database Systems*, 42(2):13:1–13:34, May 2017.

Debabrota Basu, Qian Lin, Weidong Chen, Hoang Tam Vo, Zihong Yuan, Pierre Senellart, and Stéphane Bressan. Regularized Cost-Model Oblivious Database Tuning with Reinforcement Learning. *Transactions on Large-Scale Data and Knowledge-Centered Systems*, 28:96–132, 2016.

Ruiming Tang, Antoine Amarilli, Pierre Senellart, and Stéphane Bressan. A Framework for Sampling-Based XML Data Pricing. *Transactions on Large-Scale Data and Knowledge-Centered Systems*, 24:116–138, 2016.

Serge Abiteboul, Yael Amsterdamer, Daniel Deutch, Tova Milo, and Pierre Senellart. Optimal Probabilistic Generation of XML Documents. *Theory of Computing Systems*, 57(4):806–842, November 2015.

Thomas Risse, Elena Demidova, Stefan Dietze, Wim Peters, Nikos Papailiou, Katerina Doka, Yannis Stavarakas, Vassilis Plachouras, Pierre Senellart, Florent Carpentier, Amin Mantrach, Bogdan Cautis, Patrick Siehndel, and Dimitris Spiliotopoulos. The ARCOMEM Architecture for Social and Semantic Driven Web Archiving. *Future Internet*, 6(4):688–716, November 2014.

Vassilis Plachouras, Florent Carpentier, Muhammad Faheem, Julien Masanès, Thomas Risse, Pierre Senellart, Patrick Siehndel, and Yannis Stavarakas. ARCOMEM Crawling Architecture. *Future Internet*, 6(3):518–541, August 2014.

Fabian M. Suchanek, Serge Abiteboul, and Pierre Senellart. PARIS: Probabilistic Alignment of Relations, Instances, and Schema. *Proceedings of the VLDB Endowment*, 5(3):157–168, December 2011. Presented at the VLDB 2012 conference, Istanbul, Turkey. Also CoRR abs/1111.7164.

Benjamin Nguyen, Antoine Vion, François-Xavier Dudouet, Dario Colazzo, Ioana Manolescu, and Pierre Senellart. XML Content Warehousing: Improving Sociological Studies of Mailing Lists and Web Data. *Bulletin of Sociological Methodology*, 112(1):5–31, October 2011. Also CoRR abs/1108.4596.

Serge Abiteboul, T-H. Hubert Chan, Evgeny Kharlamov, Werner Nutt, and Pierre Senellart. Capturing Continuous Data and Answering Aggregate Queries in Probabilistic XML. *ACM Transactions on Database Systems*, 36(4), 2011.

Michael Benedikt, Evgeny Kharlamov, Dan Olteanu, and Pierre Senellart. Probabilistic XML via Markov Chains. *Proceedings of the VLDB Endowment*, 3(1):770–781, September 2010. Presented at the VLDB 2010 conference, Singapore.

Georg Gottlob and Pierre Senellart. Schema Mapping Discovery from Data Instances. *Journal of the ACM*, 57(2), January 2010.

Serge Abiteboul, Benny Kimelfeld, Yehoshua Sagiv, and Pierre Senellart. On the Expressiveness of Probabilistic XML Models. *VLDB Journal*, 18(5):1041–1064, October 2009.

Vincent D. Blondel, Anahí Gajardo, Maureen Heymans, Pierre Senellart, and Paul Van Dooren. A measure of similarity between graph vertices: applications to synonym extraction and Web searching. *SIAM Review*, 46(4):647–666, 2004. Also CoRR cs.IR/0407061.

#### Conference articles (peer-reviewed)

Shrey Mishra, Antoine Gauquier, and Pierre Senellart. Modular Multimodal Machine Learning for Extraction of Theorems and Proofs in Long Scientific Documents. In *Proc. JCDL*, Hong Kong, China, December 2024. Late-breaking paper.

Angelo Saadeh, Pierre Senellart, and Stéphane Bressan. Confidential Truth Finding with Multi-Party Computation. In *Proc. DEXA*, pages 332–337, Penang, Malaysia, August 2023.

Antoine Gauquier and Pierre Senellart. Automatically Inferring the Document Class of a Scientific Article. In *Proc. DocEng*, pages 11:1–11:10, Limerick, Ireland, August 2023.

Shrey Mishra, Lucas Pluinage, and Pierre Senellart. Towards Extraction of Theorems and Proofs in Scholarly Articles. In *Proc. DocEng*, pages 25:1–25:4, Limerick, Ireland, August 2021.

Yann Ramusat, Silviu Maniu, and Pierre Senellart. Provenance-Based Algorithms for Rich Queries over Graph Databases. In *Proc. EDBT*, pages 73–84, Nicosia, Cyprus, March 2021.

Pierre Senellart. Provenance in Databases: Principles and Applications. In *Proc. RW*, pages 104–109, Bolzano, Italy, September 2019.

Debabrota Basu, Pierre Senellart, and Stéphane Bressan. BelMan: Bayesian Bandits on the Belief–Reward Manifold. In *Proc. ECML/PKDD*, pages 167–183, Würzburg, Germany, September 2019.

Silviu Maniu, Pierre Senellart, and Suraj Jog. An Experimental Study of the Treewidth of Real-World Graph Data. In *Proc. ICDT*, pages 12:1–12:18, Lisbon, Portugal, March 2019.

Miyoungh Han, Pierre-Henri Wuillemin, and Pierre Senellart. Focused Crawling through Reinforcement Learning. In *Proc. ICWE*, pages 261–278, Cáceres, Spain, June 2018.

Ben Spencer, Michael Benedikt, and Pierre Senellart. Form Filling based on Constraint Solving. In *Proc. ICWE*, pages 95–113, Cáceres, Spain, June 2018.

Antoine Amarilli, Mikaël Monet, and Pierre Senellart. Connecting Width and Structure in Knowledge Compilation. In *Proc. ICDT*, Vienna, Austria, March 2018.

Antoine Amarilli, M. Lamine Ba, Daniel Deutch, and Pierre Senellart. Possible and Certain Answers for Queries over Order-Incomplete Data. In *Proc. TIME*, pages 4:1–4:19, Mons, Belgium, October 2017.

Antoine Amarilli, Mikaël Monet, and Pierre Senellart. Conjunctive Queries on Probabilistic Graphs: Combined Complexity. In *Proc. PODS*, pages 217–232, Chicago, USA, May 2017. Also CoRR abs/1703.03201.

Antoine Amarilli, Pierre Bourhis, Mikaël Monet, and Pierre Senellart. Combined Tractability of Query Evaluation via Tree Automata and Cycluits. In *Proc. ICDT*, pages 6:1–6:19, Venice, Italy, March 2017.

Antoine Amarilli, Yael Amsterdamer, Tova Milo, and Pierre Senellart. Top-k Queries on Unknown Values under Order Constraints. In *Proc. ICDT*, pages 5:1–5:18, Venice, Italy, March 2017.

Miyoungh Han, Pierre Senellart, Stéphane Bressan, and Huayu Wu. Routing an Autonomous Taxi with Reinforcement Learning. In *Proc. CIKM*, Indianapolis, USA, October 2016. Industry track, short paper.

Antoine Amarilli, Pierre Bourhis, and Pierre Senellart. Tractable Lineages on Treelike Instances: Limits and Extensions. In *Proc. PODS*, pages 355–370, San Francisco, USA, June 2016.

Muhammad Faheem and Pierre Senellart. Adaptive Web Crawling through Structure-Based Link Classification. In *Proc. ICADL*, pages 39–51, Seoul, South Korea, December 2015.

David Montoya, Serge Abiteboul, and Pierre Senellart. Hup-Me: Inferring and Reconciling a Timeline of User Activity with Smartphone and Personal Data. In *Proc. SIGSPATIAL*, Seattle, USA, November 2015. Short paper.



Debabrota Basu, Qian Lin, Weidong Chen, Hoang Tam Vo, Zihong Yuan, Pierre Senellart, and Stéphane Bressan. Cost-Model Oblivious Database Tuning with Reinforcement Learning. In *Proc. DEXA*, pages 253–268, Valencia, Spain, September 2015.

Siyu Lei, Silviu Maniu, Luyi Mo, Reynold Cheng, and Pierre Senellart. Online Influence Maximization. In *Proc. KDD*, pages 645–654, Sydney, Australia, August 2015.

Antoine Amarilli, Pierre Bourhis, and Pierre Senellart. Provenance Circuits for Trees and Treelike Instances. In *Proc. ICALP*, pages 56–68, Kyoto, Japan, July 2015.

Changping Meng, Reynold Cheng, Silviu Maniu, Pierre Senellart, and Wangda Zhang. Discovering Meta-Paths in Large Heterogeneous Information Networks. In *Proc. WWW*, pages 754–764, Florence, Italy, May 2015.

Georges Gouriten, Silviu Maniu, and Pierre Senellart. Scalable, Generic, and Adaptive Systems for Focused Crawling. In *Proc. Hypertext*, pages 35–45, Santiago, Chile, September 2014. Douglas Engelbart Best Paper Award.

Ruiming Tang, Antoine Amarilli, Pierre Senellart, and Stéphane Bressan. Get a Sample for a Discount: Sampling-Based XML Data Pricing. In *Proc. DEXA*, pages 20–34, Munich, Germany, September 2014.

M. Lamine Ba, Talel Abdesslem, and Pierre Senellart. Uncertain Version Control in Open Collaborative Editing of Tree-Structured Documents. In *Proc. DocEng*, pages 27–36, Florence, Italy, September 2013.

Muhammad Faheem and Pierre Senellart. Intelligent and Adaptive Crawling of Web Applications for Web Archiving. In *Proc. ICWE*, pages 306–322, Aalborg, Denmark, July 2013.

Yael Amsterdamer, Yael Grossman, Tova Milo, and Pierre Senellart. Crowd Mining. In *Proc. SIGMOD*, pages 241–252, New York, USA, June 2013.

Asma Souihli and Pierre Senellart. Optimizing Approximations of DNF Query Lineage in Probabilistic XML. In *Proc. ICDE*, pages 721–732, Brisbane, Australia, April 2013.

Michael Benedikt, Pierre Bourhis, and Pierre Senellart. Monadic Datalog Containment. In *Proc. ICALP*, pages 79–91, Warwick, United Kingdom, July 2012.

Jiaheng Lu, Pierre Senellart, Chunbin Lin, Xiaoyong Du, Shan Wang, and Xinxing Chen. Optimal Top-k Generation of Attribute Combinations based on Ranked Lists. In *Proc. SIGMOD*, pages 409–420, Scottsdale, USA, May 2012.

Georges Gouriten and Pierre Senellart. API Blender: A Uniform Interface to Social Platform APIs. In *Proc. WWW*, Lyon, France, April 2012. Developer track. Also CoRR abs/1301.2086.

Serge Abiteboul, Pierre Senellart, and Victor Vianu. The ERC Webdam on Foundations of Web Data Management. In *Proc. WWW*, pages 211–214, Lyon, France, April 2012. European project track.

Serge Abiteboul, Yael Amsterdamer, Daniel Deutch, Tova Milo, and Pierre Senellart. Finding Optimal Probabilistic Generators for XML Collections. In *Proc. ICDT*, pages 127–139, Berlin, Germany, March 2012.

Michael Benedikt, Georg Gottlob, and Pierre Senellart. Determining Relevance of Accesses at Runtime. In *Proc. PODS*, pages 211–222, Athens, Greece, June 2011.

Serge Abiteboul, T-H. Hubert Chan, Evgeny Kharlamov, Werner Nutt, and Pierre Senellart. Aggregate Queries for Discrete and Continuous Probabilistic XML. In *Proc. ICDT*, pages 50–61, Lausanne, Switzerland, March 2010.

Alban Galland, Serge Abiteboul, Amélie Marian, and Pierre Senellart. Corroborating Information from Disagreeing Views. In *Proc. WSDM*, pages 131–140, New York, USA, February 2010.

Pierre Senellart and Georg Gottlob. On the Complexity of Deriving Schema Mappings from Database Instances. In *Proc. PODS*, pages 23–32, Vancouver, Canada, June 2008.

Yann Ollivier and Pierre Senellart. Finding Related Pages Using Green Measures: An Illustration with Wikipedia. In *Proc. AAAI*, pages 1427–1433, Vancouver, Canada, July 2007.

Pierre Senellart and Serge Abiteboul. On the Complexity of Managing Probabilistic XML Data. In *Proc. PODS*, pages 283–292, Beijing, China, June 2007.

Serge Abiteboul and Pierre Senellart. Querying and Updating Probabilistic Information in XML. In *Proc. EDBT*, pages 1059–1068, Munich, Germany, March 2006. Short paper.

François-Xavier Dudouet, Ioana Manolescu, Benjamin Nguyen, and Pierre Senellart. XML Warehousing Meets Sociology. In *Proc. IADIS ICWI*, pages 170–174, Lisbon, Portugal, October 2005.

Mats Attnäs, Pierre Senellart, and Jean Senellart. Integration of SYSTRAN MT systems in an open workflow. In *Proc. MT Summit*, Phuket, Thailand, September 2005.

Pierre Senellart. Identifying Websites with Flow Simulation. In *Proc. ICWE*, pages 124–129, Sydney, Australia, July 2005.

#### Workshop articles (peer-reviewed)

Shrey Mishra, Yacine Brihrouche, Théo Delemazure, Antoine Gauquier, and Pierre Senellart. First Steps in Building a Knowledge Base of Mathematical Results. In *Proc. SDP*, pages 165–174, Bangkok, Thailand, August 2024.

Shufan Jiang and Pierre Senellart. Extracting Definienda in Mathematical Scholarly Articles with Transformers. In *Proc. WIESP*, online, Indonesia, November 2023.

Yann Ramusat, Silviu Maniu, and Pierre Senellart. Efficient Provenance-Aware Querying of Graph Databases with Datalog. In *Proc. GRADES-NDA*, pages 4:1–4:9, Philadelphia, USA, June 2022.

Yann Ramusat, Silviu Maniu, and Pierre Senellart. Semiring Provenance over Graph Databases. In *Proc. TaPP*, London, United Kingdom, July 2018.

David Montoya, Thomas Pellissier Tanon, Serge Abiteboul, Pierre Senellart, and Fabian M. Suchanek. A Knowledge Base for Personal Information Management. In *Proc. LDOW*, Lyon, France, April 2018.

Ognjen Savkovic, Evgeny Kharlamov, Werner Nutt, and Pierre Senellart. Towards Approximating Incomplete Queries over Partially Complete Databases. In *Proc. AMW*, Montevideo, Uruguay, June 2017.

M. Lamine Ba, Roxana Horincar, Pierre Senellart, and Huayu Wu. Truth Finding with Attribute Partitioning. In *Proc. WebDB*, pages 27–33, Melbourne, Australia, May 2015.

Marilena Oita and Pierre Senellart. FOREST: Focused Object Retrieval by Exploiting Significant Tag Paths. In *Proc. WebDB*, pages 55–61, Melbourne, Australia, May 2015.

Silviu Maniu, Reynold Cheng, and Pierre Senellart. ProbTree: A Query-Efficient Representation of Probabilistic Graphs. In *Proc. BUDA*, Snowbird, USA, June 2014. Workshop without formal proceedings.

M. Lamine Ba, Talel Abdesslem, and Pierre Senellart. Merging Uncertain Multi-Version XML Documents. In *Proc. DChanges*, Florence, Italy, September 2013.

Thomas Risse, Wim Peters, and Pierre Senellart. Social and Semantic Driven Web Harvesting. In *Proc. Building Web Observatories*, Paris, France, May 2013.

Marilena Oita, Antoine Amarilli, and Pierre Senellart. Cross-Fertilizing Deep Web Analysis and Ontology Enrichment. In *Proc. VLDS*, pages 5–8, Istanbul, Turkey, August 2012. Vision article.

Meghyn Bienvenu, Daniel Deutch, Davide Martinenghi, Pierre Senellart, and Fabian M. Suchanek. Dealing with the Deep Web and all its Quirks. In *Proc. VLDS*, pages 21–24, Istanbul, Turkey, August 2012. Vision article.

M. Lamine Ba, Talel Abdesslem, and Pierre Senellart. Towards a Version Control Model with Uncertain Data. In *Proc. PIKM*, Glasgow, United Kingdom, October 2011.

Marilena Oita and Pierre Senellart. Deriving Dynamics of Web Pages: A Survey. In *Proc. TAWW*, pages 25–32, Hyderabad, India, March 2011.

Evgeny Kharlamov, Werner Nutt, and Pierre Senellart. Value Joins are Expensive over (Probabilistic) XML. In *Proc. LID*, pages 41–48, Uppsala, Sweden, March 2011.

Marilena Oita and Pierre Senellart. Archiving Data Objects Using Web Feeds. In *Proc. IWAW*, pages 31–41, Vienna, Austria, September 2010.

Evgeny Kharlamov, Werner Nutt, and Pierre Senellart. Updating Probabilistic XML. In *Proc. Updates in XML*, Lausanne, Switzerland, March 2010.

Serge Abiteboul, Sihem Amer-Yahia, Amélie Marian, Alban Galland, and Pierre Senellart. Birds of a tag flock together. In *Proc. SSM*, New York, USA, February 2010. Workshop without formal proceedings.

Marc Spaniol, Dimitar Denev, Arturas Mazeika, Pierre Senellart, and Gerhard Weikum. Data Quality in Web Archiving. In *Proc. WICOW*, pages 19–26, Madrid, Spain, April 2009.

Pierre Senellart, Avin Mittal, Daniel Muschick, Rémi Gilleron, and Marc Tommasi. Automatic Wrapper Induction from Hidden-Web Sources with Domain Knowledge. In *Proc. WIDM*, pages 9–16, Napa, USA, October 2008.

Vincent D. Blondel and Pierre Senellart. Automatic extraction of synonyms in a dictionary. In *Proc. TMW*, pages 7–13, Arlington, USA, April 2002. Also registered as Technical Report 89 (2001), Université catholique de Louvain, Louvain-la-neuve.

## Demonstrations (peer-reviewed)

- Shrey Mishra, Neil Sharma, Antoine Gauquier, and Pierre Senellart. TheoremView: A Framework for Extracting Theorem-Like Environments from Raw PDFs. In *Proc. ECIR*, Lucca, Italy, April 2025. Demonstration.
- Fajrian Yunus, Pratik Karmakar, Pierre Senellart, Talel Abdesslem, and Stéphane Bressan. Using a Probabilistic Database in an Image Retrieval Application. In *Proc. EDBT*, Barcelona, Spain, March 2025. Demonstration.
- Pierre Senellart, Louis Jachiet, Silviu Maniu, and Yann Ramusat. ProvSQL: Provenance and Probability Management in PostgreSQL. In *Proc. VLDB*, pages 2034–2037, Rio de Janeiro, Brazil, August 2018. Demonstration.
- M. Lamine Ba, Sébastien Montenez, Talel Abdesslem, and Pierre Senellart. Monitoring moving objects using uncertain Web data. In *Proc. SIGSPATIAL*, pages 565–568, Dallas, USA, November 2014. Demonstration.
- Muhammad Faheem and Pierre Senellart. Demonstrating Intelligent Crawling and Archiving of Web Applications. In *Proc. CIKM*, pages 2481–2484, San Francisco, USA, October 2013. Demonstration.
- Yael Amsterdamer, Yael Grossman, Tova Milo, and Pierre Senellart. Crowd Miner: Mining association rules from the crowd. In *Proc. VLDB*, pages 1250–1253, Riva del Garda, Italy, August 2013. Demonstration.
- Asma Souihli and Pierre Senellart. Demonstrating ProApproX 2.0: A Predictive Query Engine for Probabilistic XML. In *Proc. CIKM*, pages 2734–2736, Maui, Hawaii, USA, October 2012. Demonstration.
- Serge Abiteboul, Yael Amsterdamer, Tova Milo, and Pierre Senellart. Auto-Completion Learning for XML. In *Proc. SIGMOD*, pages 669–672, Scottsdale, USA, May 2012. Demonstration.
- Michael Benedikt, Tim Furche, Andreas Savvides, and Pierre Senellart. ProFoUnd: Program-analysis-based Form Understanding. In *Proc. WWW*, pages 313–316, Lyon, France, April 2012. Demonstration.
- Pierre Senellart and Asma Souihli. ProApproX: A Lightweight Approximation Query Processor over Probabilistic Trees. In *Proc. SIGMOD*, pages 1295–1298, Athens, Greece, June 2011. Demonstration.
- Talel Abdesslem, M. Lamine Ba, and Pierre Senellart. A Probabilistic XML Merging Tool. In *Proc. EDBT*, pages 538–541, Uppsala, Sweden, March 2011. Demonstration.

## Posters (peer-reviewed)

- Thomas Risse, Stefan Dietze, Wim Peters, Katerina Doka, Yannis Stavarakas, and Pierre Senellart. Exploiting the Social and Semantic Web for guided Web Archiving. In *Proc. TPD*, pages 426–432, Paphos, Cyprus, September 2012. Poster.
- Benjamin Nguyen, Antoine Vion, François-Xavier Dudouet, Ioana Manolescu, Dario Colazzo, and Pierre Senellart. The WebStand Project. In *Proc. WebSci*, Athens, Greece, March 2009. Also CoRR abs/1002.0971 (Poster).
- Michalis Vazirgiannis, Dimitris Drosos, Pierre Senellart, and Akrivi Vlachou. Web Page Rank Prediction with Markov Models. In *Proc. WWW*, pages 1075–1076, Beijing, China, April 2008. Poster.

## Industrial experience

- November 2004 – February 2005 **Software author**, SYSTRAN S.A., La Défense, France.  
SYSTRAN Mobile for Pocket PC.
- September 2003 – September 2004 **R&D engineer**, SYSTRAN S.A., La Défense, France.  
SYSTRAN version 5.
- September 2002 – October 2002 **Software engineer**, *Le Robert*, Paris, France.  
Analysis of the dictionary graph and automatic extraction of similar words.

## Information technology

Programming languages	C++; Perl; Java; Scala; C; Python; PHP	OS	Linux; Unix; Windows
XML technologies	DOM; SAX; DTD; XML Schema; XPath; XSLT; XQuery; XSL-FO	Web	(X)HTML; CSS; JavaScript; PHP; JSP; SVG
DBMS	PostgreSQL; MySQL; Oracle	Miscellaneous	LaTeX; Unicode; Hadoop

## Languages

French native language  
German proficient  
Malay basics

English fluent  
Classical Arabic basics  
Ancient Greek high-school level