

# Pierre Senellart

Professeur en informatique

Gestion de données, Web, IA

32 villa de Lourcine  
75014 Paris, France

☎ (+33) 6-48-36-36-20

✉ pierre@senellart.com

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

Français, né le 19 juin 1981.



## Carrière académique

- septembre 2016 – **Professeur**, *École normale supérieure, Université PSL*, Paris, France.  
... *Département Informatique (ENS, CNRS, Inria)*. En délégation à l'Institut universitaire de France (2020-2025).
- décembre 2017 – **Professeur invité**, *Télécom Paris, IP Paris*, Paris, France.  
décembre 2020 *Département Informatique et Réseaux*.
- août 2014 – **Chercheur sénior**, *National University of Singapore*, Singapore, Singapour.  
juillet 2016 *Département Informatique, laboratoire IPAL*.
- janvier 2013 – **Professeur**, *Télécom ParisTech*, Paris, France.  
août 2016 *Département Informatique et Réseaux*.
- novembre 2012 – **Chercheur et conférencier invité**, *The University of Hong Kong*, Hong Kong, Chine.  
avril 2013 *Département d'informatique*
- juin 2008 – **Maître de conférences**, *Télécom ParisTech*, Paris, France.  
décembre 2012 *Département Informatique et Réseaux*.
- septembre 2008 – **Chargé de TP**, *École polytechnique*, Palaiseau, France.  
décembre 2008 *TP de bases de données*.
- février 2008 – **Chargé de TP**, *École normale supérieure de Cachan*, Cachan, France.  
mai 2008 *TP de bases de données*.
- février 2008 – **Chercheur post-doctorant**, *Max-Planck-Institut für Informatik*, Sarrebruck, Allemagne.  
mai 2008 *Département Databases and Information Systems*.
- septembre 2005 – **Doctorant et moniteur**, *Université Paris-Sud & INRIA Futurs*, Orsay, France.  
mai 2008 *Sujet : comprendre le Web caché. Directeur : Serge Abiteboul. 64h d'enseignement par an.*
- mars 2007 – **Étudiant chercheur**, *Oxford University, Department of computer science*, Oxford, Royaume-Uni.  
mai 2007 *Sujet : Échange de données. Directeur : Georg Gottlob.*
- 2003 – 2005 **Vacataire**, *École nationale des ponts et chaussées*, Marne-la-Vallée, France.  
*Enseignement d'un cours de C++ aux élèves de première année (2 semestres de 60 heures chacun).*
- mars 2003 – **Stagiaire de DEA**, *INRIA Futurs*, Orsay, France.  
août 2003 *Sujet : Identification de sites Web logique. Directeur : Serge Abiteboul.*
- février 2002 – **Stagiaire de recherche**, *Chalmers Tekhniska Högskola*, Göteborg, Suède.  
juin 2002 *Sujet : Vérification automatique des multiplicateurs. Directrice : Mary Sheeran.*
- juin 2001 – **Stagiaire de recherche**, *Université catholique de Louvain*, Louvain-la-neuve, Belgique.  
juillet 2001 *Sujet : Extraction automatique de mots similaires dans un dictionnaire. Directeur : Vincent Blondel.*

## Formation

- juin 2012 **Habilitation à diriger les recherches en Informatique**, *Université Pierre et Marie Curie*, Paris, France.  
*Sujet : XML probabiliste : un modèle de données pour le Web*

- décembre 2007 **Doctorat en Informatique**, *Université Paris–Sud*, Orsay, France.  
Sujet : Comprendre le Web caché. Mention Très Honorable.
- septembre 2003 **Magistère en Mathématiques fondamentales et appliquées et informatique**, *Université Paris–Sud & ENS*, Orsay & Paris, France.
- septembre 2003 **DEA en Informatique**, *Université Paris–Sud*, Orsay, France.  
Mention Très Bien.
- septembre 2001 **Maîtrise en Informatique**, *Université Paris–Sud & ENS*, Orsay & Paris, France.  
Mention Très Bien.
- juin 2001 **Licence en Informatique**, *Université Paris–Sud & ENS*, Orsay & Paris, France.  
Mention Très Bien.
- 2000 – 2005 **École normale supérieure (Paris)**.
- juin 1998 **Baccalauréat S**, *Lycée Mariette*, Boulogne-sur-mer, France.  
Mention Très Bien.

## Encadrement de jeunes chercheurs

- Doctorants
- Aryak Sen (avec Silviu Maniu, depuis 2024). *Scalabilité d'un système de gestion de provenance et des probabilités des données.*
- Emmanouil Sylligardos (avec Paul Boniol, depuis 2024). *Compromis entre précision et temps d'exécution dans l'assemblage et la sélection de modèles pour l'analyse des séries temporelles.*
- Antoine Gauquier (avec Ioana Manolescu, depuis 2023). *Construction intelligente d'un entrepôt de données multimodales et hétérogènes, avec traçabilité des données.*
- Pratik Karmakar (avec Stéphane Bressan, depuis 2023). *Qualité, incertitude et lignage des données.*
- Shrey Mishra (depuis 2021). *Vers une base de connaissance de théorèmes.*
- Baptiste Lafosse (avec Jean-Marie Lagniez, depuis 2021). *Compilateur dédié à l'évaluation de requêtes SQL.*
- Yann Ramusat (avec Silviu Maniu, thèse soutenue en 2022). *Provenance basée sur les semi-anneaux pour les bases de données orientées graphes.*
- Debabrota Basu (avec Stéphane Bressan, thèse soutenue en 2018). *Apprendre à faire des décisions en présence d'information incomplète: apprentissage par renforcement, géométrie de l'information, et applications au monde réel.*
- Mikaël Monet (avec Antoine Amarilli, thèse soutenue en 2018). *Complexité combinée de l'évaluation probabiliste de requêtes.* [Prix de thèse BDA 2019.](#)
- Miyong Han (thèse soutenue en 2018). *Approches d'apprentissage par renforcement dans les environnements dynamiques.*
- Quentin Lobbé (avec Dana Diminescu, thèse soutenue en 2018). *Archives, fragments Web et diasporas. Pour une exploration désagrégée de corpus d'archives Web liées aux représentations en ligne des diasporas.* [Prix de la recherche de l'Inathèque 2019.](#)
- David Montoya (avec Serge Abiteboul, thèse soutenue en 2017). *Une base de connaissance personnelle intégrant les données d'un utilisateur et une chronologie de ses activités.*
- Su Yang (avec Serge Abiteboul, 2016 – 2017). *Extraction sémantique d'information personnelle, intégration et interrogation.*
- Antoine Amarilli (avec Pierre Bourhis, thèse soutenue en 2016). *Tirer parti de la structure des données incertaines.* [Prix de thèse E. W. Beth 2017 en logique, langages et information.](#) [Prix de thèse Télécom ParisTech 2017.](#)
- M. Lamine Ba (avec Talal Abdesslem, thèse soutenue en 2015). *Exploitation de la structure des données incertaines.*

Muhammad Faheem (thèse soutenue en 2014). *Acquisition intelligente de contenu dans l'archivage du Web.*

Ioana Ileana (avec Bogdan Cautis, thèse soutenue en 2014). *Réécriture de requêtes avec des vues : une perspective théorique et pratique.*

Georges Gouriten (avec Bogdan Cautis, 2012 – 2013). *Suggestions de contenu sur le Web social fondées sur la connaissance.*

Asma Souihli (thèse soutenue en 2012). *Interrogation de données XML probabilistes.*

Marilena Oita (thèse soutenue en 2012). *Inférer des objets sémantiques du Web structuré.*

Evgeny Kharlamov (avec Werner Nutt, thèse soutenue en 2010). *Une approche probabiliste de la gestion de données XML.*

Post-doctorants Shufan Jiang (2022 – 2023).

Nofar Carmeli (2021 – 2022).

Ashish Dandekar (2019 – 2020).

Roxana Horincar (2013 – 2015).

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

Autres Encadrement de 9 Étudiants de Master et 7 Élèves ingénieur sur des projets à long terme

## Recherche contractuelle

2021 – 2026 **DesCartes**, Programme CNRS@CREATE sur la modélisation intelligente pour la prise de décision dans les systèmes urbains critiques.

Participant, faisant partie d'un projet de 50 millions de SGD financé par la NRF de Singapour.

2021 – 2024 **Dissemin**, Développement d'une plate-forme de science ouverte.

Porteur, financement de 124 kEUR.

2019 – 2024 **PRAIRIE**, Institut de recherche en intelligence artificielle de Paris.

Titulaire d'une chaire, financement de 600 kEUR.

2018 – 2021 **BioQOP**, Projet collaboratif ANR sur l'optimisation des requêtes à des données de vidéo-surveillance sous contraintes de confidentialité.

Responsable pour l'ENS, financement de 69 kEUR.

2013 – 2015 **XData**, Projet de collaboration avec industriels sur l'acquisition de données sociales et l'intégration avec des données d'entreprise.

Responsable pour Télécom, financement de 175 kEUR.

2013 – 2015 **CCIPX**, Projet collaboratif STIC-Asie entre l'Institut Mines-Télécom, Hanoi University of Science and Technology et National University of Singapore.

Co-investigateur principal. Financement de 26 kEUR.

2012 – 2017 **MoDaS**, ERC advanced grant de Tova Milo sur le sourçage de données par la foule.

Collaboration sur des articles de recherche, échange d'étudiants.

2012 – 2013 **Private Recommendation System in Social Networks**, Partenariat France-Hong Kong Campus France PROCORE.

Co-investigateur principal. Financement de 8 kEUR.

2010 – 2014 **Diadem**, ERC advanced grant de Georg Gottlob sur l'extraction de données automatisée intelligente centrée sur un domaine.

Collaboration sur des articles de recherche, échange d'étudiants.

2011 – 2013 **ARCOMEM**, Projet FP7 de l'union européenne sur l'archivage du Web social.

Responsable de l'activité sur l'acquisition de contenu ; responsable pour Télécom, financement de 600 kEUR.

- 2009 – 2013 **Webdam**, *ERC advanced grant de Serge Abiteboul sur les fondements de la gestion de données du Web*.  
Responsable de l'axe de recherche sur les données incertaines; comité de direction (avec S. Abiteboul et V. Vianu).
- 2010 – 2012 **PANIC**, *Projet ANR CONTINT sur la proactivité des audiences et la numérisation des industries culturelles*.  
Collaborateur.
- 2009 – 2011 **DataRing**, *Projet ANR VERSO sur le partage de données en pair-à-pair pour les communautés en ligne*.  
Dirige l'activité sur la modélisation des données; responsable pour Télécom, financement de 120 kEUR.

## Activités d'enseignement

- Cours créés** Algorithmique et applications (L2, Université PSL); Data wrangling et qualité des données (M2, Université PSL); Anonymisation, confidentialité (M2, Université PSL); Bases de données (L3, ENS); Sciences des données (formation professionnelle, Télécom ParisTech); Entraînement ACM-ICPC (cours de résolution de problèmes conduisant à un concours de programmation, Télécom ParisTech); Recherche sur le Web (cours intensif sur une ou plusieurs semaines, Télécom ParisTech & Renmin University of China); Technologies du Web (vision d'ensemble pour des étudiants commerciaux et ingénieurs, Télécom ParisTech & HEC); L<sup>A</sup>T<sub>E</sub>X (cours intensif de deux jours pour doctorants, EDITE de Paris); Programmation Web (cours introductif, Université Paris-Sud)
- Cours avec une forte implication** Algorithmique (L3, ENS); Théorie des langages (L3, Télécom ParisTech); Ingénierie du Web (M1, Télécom ParisTech); Gestion des données du Web (M2, Université Paris-Sud, Télécom-ParisTech, MPRI, & The University of Hong Kong); Conception d'application de gestion de données (National University of Singapore); Théorie de l'information (National University of Singapore)
- Participation dans d'autres cours** Programmation C++ (École des Ponts ParisTech); Programmation Java (Télécom ParisTech); Bases de données (Université Paris-Sud, Télécom ParisTech & ENS Cachan); Intelligence collective (Télécom ParisTech); Systèmes d'exploitation (Université Paris-Sud)
- Administration de l'enseignement** Création et équipe de direction du Master 2 IASD (depuis 2019); Représentant de Télécom ParisTech puis de l'ENS à la commission des études du Master 2 MPRI (2013-2018); Création et gestion (avec F. d'Alché) de la filière Science des données à Télécom ParisTech (2014-2016); Création (avec S. Cléménçon) et gestion (avec S. Cléménçon, puis F. d'Alché) du certificat d'étude spécialisé (CES) Data Scientist à Télécom ParisTech (2014-2016)

## Fonctions et responsabilités administratives

- 2021 – 2026 **Président de la section 6 (informatique) du comité national de la recherche scientifique**, *Comité national de la recherche scientifique*.  
Évaluation des chercheurs et unités CNRS, jurys de recrutement et de promotion de chercheurs CNRS, autres activités d'évaluation. Également membre du bureau de la CPCN.
- 2018 – 2020 **Chercheur associé**, *CERRE*.  
Think tank européen sur la régulation des industries de réseau
- 2018 – ... **Gestion du concours d'entrée en informatique à l'ENS**, *Concours X/ENS*.  
Sélection des examinateurs, des sujets, en accord avec les représentants des trois autres ENS et de l'École polytechnique, pour le concours d'entrée principal. Gestion, sélection et auditions (avec D. Naccache) pour les voies alternatives de recrutement (sélection internationale, voie universitaire).

- 2017 – ... **Directeur adjoint, DI ENS.**  
Le DI ENS est à la fois un département d'enseignement et une unité mixte de recherche CNRS/ENS/Inria; j'assiste le directeur (D. Pointcheval) dans tous les domaines liés au budget du département, à la gestion des enseignements, au recrutement du personnel, à l'approbation des dépenses, etc.
- 2017 – 2021 **Membre du bureau, DIM RFSI.**  
Le DIM RFSI est une alliance régionale de laboratoires d'informatique, qui alloue du financement à la recherche.
- 2017 – 2021 **Juge en chef, co-organisateur, secrétaire, ICPC SWERC.**  
Le SWERC est la compétition régionale pour l'Europe du Sud-Ouest de l'International Collegiate Programming Competition, la plus grande compétition de programmation organisée par et pour les universités. Je co-organise SWERC avec des collègues de Télécom ParisTech depuis 2017, et été le juge et concepteur de problème en chef jusqu'en 2019–2020. Nous avons réussi à attirer un nombre record d'équipes (76 en 2017, 89 en 2018, 98 en 2019–2020) des pays participants (France, Italie, Espagne, Portugal, Suisse, Israël).
- 2016 – ... **Responsable d'équipe, Valda.**  
J'ai fondé et je suis responsable de l'équipe Valda, une équipe ENS/CNRS/Inria au sein du DI ENS.
- 2016 – 2021 **Membre du bureau de la section 6 (informatique) du comité national de la recherche scientifique, Comité national de la recherche scientifique.**  
Évaluation des chercheurs et unités CNRS, jurys de recrutement et de promotion de chercheurs CNRS, autres activités d'évaluation.
- 2014 – 2016 **Concepteur du sujet du concours en informatique, Concours commun Mines–Ponts.**  
Concours d'informatique pour les étudiants de classe préparatoire MP option Info.

## Service à la communauté scientifique

- (Co-)président de PC BDA 2017; WebDB 2017; SUM 2016; UnCrowd 2014; DBCrowd 2013; BDA 2016 (démonstrations); VLDB 2015 (tutoriels); DASFAA 2015 (tutoriels); DEXA 2015 (tutoriels); ICDE 2011 (tutoriels); EDBT 2011 (industrial track); SIGMOD 2010 (programming contest)
- Comités de programme PODS 2012, 2016, 2019 & 2021; STACS 2020; EDBT 2013 (démonstrations), 2014 & 2018; ICDT 2011 & 2017; SIGMOD 2008 & 2009 (comité de répétabilité), 2013 & 2017 (membre du comité de programme), 2015 (responsable de groupe), 2023; CIKM 2010 & 2014; VLDB 2013; WWW 2009 & 2010; ICDE 2010; ECML/PKDD 2008; BDA 2010, 2012, 2018, 2020, 2022; divers workshops et petites conférences
- Responsabilités éditoriales Directeur de l'information du *Journal of the ACM* (2009-2015); Co-éditeur d'un numéro spécial de *Distributed and Parallel Databases* (2014)
- Révisions d'articles de journaux *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*; ...
- Jurys d'HdR Stefan Mengel (rapporteur, Université d'Artois, 2021); Federico Ulliana (rapporteur, Université de Montpellier, 2021); Paolo Papotti (rapporteur, Université Côte d'Azur, 2019); Raja Chiky (rapporteur, Université Pierre-et-Marie-Curie, 2016)



Jurys de thèses de doctorat	Claire Soyez-Martin (président, Université de Lille, 2023); Pierre Faure–Giovagnoli (président, INSA Lyon, 2023); Rébecca Zucchini (rapporteur, Université Paris-Saclay, 2023); Antonia Kormpa (rapporteur, Université d’Oxford, 2023); Gaston Zanitti (rapporteur & président, Université Paris-Saclay, 2023); Nicolas Crosetti (rapporteur, Université de Lille, 2023); Sajad Nazari (rapporteur, INSA Centre Val de Loire, 2022); Amit Kumar (président, Université Caen Normandie, 2021); Valentin Iovene (rapporteur, Université Paris-Saclay, 2021); Wissam Mammam Kouadri (rapporteur, Université Paris-Saclay, 2021); Julien Romero (président, IP Paris, 2020); Ugo Comignani (rapporteur, Université Claude Bernard Lyon 1, 2019); Ngurah Agus Sanjaya Er (président, Télécom ParisTech, 2018); Govind (président, Université Caen Normandie, 2018); Paul Lagrée (rapporteur, Université Paris-Saclay, 2017); Manuel Pozo (président, CNAM, 2016); Mira Sarkis (président, Télécom ParisTech, 2016); Mohammadreza Khelgati (rapporteur, Universiteit Twente, 2016); Ahmad Assaf (Télécom ParisTech, 2015); François Rousseau (École polytechnique, 2015); Disheng Qiu (rapporteur, 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 (rapporteur, Université Paris-Sud, 2012); Andrew Sellers (rapporteur, Université d’Oxford, 2011)
Autres jurys	Prix de thèse Gilles Kahn de la Société Informatique de France (2019, 2020, 2021); Prix BDA de la meilleure thèse en gestion de données (2019)

## Prix et récompenses

- Membre junior de l’Institut universitaire de France (2020-2025)
- Membre distingué du comité de programme de SIGMOD 2017
- Prix Douglas Engelbart du meilleur article à ACM HyperText 2014
- Supervision de Clément Genzmer, vainqueur du concours de programmation SIGMOD 2009

## Publications

Ne sont mentionnées que les publications internationales; pour une liste complète (incluant les conférences et revues nationales, rapports techniques, travaux non relus, prépublications), voir <http://pierre.senellart.com/publications/>.

### Livre

Serge ABITEBOUL, Ioana MANOLESCU, Marie-Christine ROUSSET, Philippe RIGAUX et Pierre SENELLART : *Web Data Management*. Cambridge University Press, New York, USA, janvier 2012.

### Chapitres de livre (avec revue par les pairs)

Thomas RISSE, Wim PETERS, Pierre SENELLART et Diana MAYNARD : Documenting Contemporary Society by Preserving Relevant Information from Twitter. In Katrin WELLER, Axel BRUNS, Jean BURGESS, Merja MAHRT et Cornelius PUSCHMANN, éditeurs : *Twitter and Society*, pages 207–219. Peter Lang International Academic Publishers, novembre 2013.

Evgeny KHARLAMOV et Pierre SENELLART : Modeling, Querying, and Mining Uncertain XML Data. In Andrea TAGARELLI, éditeur : *XML Data Mining: Models, Methods, and Applications*, pages 29–52. IGI Global, novembre 2011.

### Chapitres de livre (sans revue par les pairs)

Pierre SENELLART : On the Impact of Provenance Semiring Theory on the Design of a Provenance-Aware Database System. In Antoine AMARILLI et Alin DEUTSCH, éditeurs : *The Provenance of Elegance in Computation — Essays dedicated to Val Tannen*, pages 12 :1–12 :10. Schloss Dagstuhl — Leibniz-Zentrum für Informatik, mars 2024.

Benny KIMELFELD et Pierre SENELLART : Probabilistic XML: Models and Complexity. In Zongmin MA et Li YAN, éditeurs : *Advances in Probabilistic Databases for Uncertain Information Management*, pages 39–66. Springer-Verlag, mai 2013.

Michael BENEDIKT et Pierre SENELLART : Databases. In Edward K. BLUM et Alfred V. AHO, éditeurs : *Computer Science. The Hardware, Software and Heart of It*, pages 169–229. Springer-Verlag, janvier 2012.

Pierre SENELLART et Vincent D. BLONDEL : Automatic discovery of similar words. In Michael W. BERRY et Malu

CASTELLANOS, éditeurs : *Survey of Text Mining II: Clustering, Classification and Retrieval*, pages 25–44. Springer-Verlag, janvier 2008.

Pierre SENELLART et Vincent D. BLONDEL : Automatic discovery of similar words. In Michael W. BERRY, éditeur : *Survey of Text Mining: Clustering, Classification and Retrieval*, pages 25–43. Springer-Verlag, août 2003.

#### Articles de journal (avec revue par les pairs)

Pratik KARMAKAR, Mikaël MONET, Pierre SENELLART et Stéphane BRESSAN : Expected Shapley-Like Scores of Boolean Functions: Complexity and Applications to Probabilistic Databases. *Proceedings of the ACM on Management of Data*, 49(2), juin 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 BORROMEO, Paolo PAPOTTI, Raghav RAO, Sudeepa ROY, Pierre SENELLART, Keishi TAJIMA, Saravanan THIRUMURUGANATHAN, Marion TOMMASI, Kazutoshi UMEMOTO, Andrea WIGGINS et Koichiro YOSHIDA : Making AI Machines Work for Humans in FoW. *SIGMOD Record*, 49(2):30–35, décembre 2020.

Antoine AMARILLI, Florent CAPELLI, Mikaël MONET et Pierre SENELLART : Connecting Knowledge Compilation Classes and Width Parameters. *Theory of Computing Systems*, 64(5):861–914, 2020. Également CoRR abs/1811.02944.

Michael BENEDIKT, Pierre BOURHIS, Georg GOTTLÖB et 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 et Pierre SENELLART : Evaluating Datalog via Tree Automata and Cycluits. *Theory of Computing Systems*, 63(7):1620–1678, 2019. Également CoRR abs/1808.04663.

Antoine AMARILLI, M. Lamine BA, Daniel DEUTCH et Pierre SENELLART : Computing Possible and Certain Answers over Order-Incomplete Data. *Theoretical Computer Science*, 797:42–76, 2019. Également CoRR abs/1801.06396.

Silviu MANIU, Reynold CHENG et Pierre SENELLART : An Indexing Framework for Queries on Probabilistic Graphs. *ACM Transactions on Database Systems*, 42(2):13 :1–13 :34, mai 2017.

Debabrota BASU, Qian LIN, Weidong CHEN, Hoang Tam VO, Zihong YUAN, Pierre SENELLART et 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 et 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 et Pierre SENELLART : Optimal Probabilistic Generation of XML Documents. *Theory of Computing Systems*, 57(4):806–842, novembre 2015.

Thomas RISSE, Elena DEMIDOVA, Stefan DIETZE, Wim PETERS, Nikos PAPAILIOU, Katerina DOKA, Yannis STAVRAKAS, Vassilis PLACHOURAS, Pierre SENELLART, Florent CARPENTIER, Amin MANTRACH, Bogdan CAUTIS, Patrick SIEHNDEL et Dimitris SPILIOTOPOULOS : The ARCOMEM Architecture for Social and Semantic Driven Web Archiving. *Future Internet*, 6(4):688–716, novembre 2014.

Vassilis PLACHOURAS, Florent CARPENTIER, Muhammad FAHEEM, Julien MASANÈS, Thomas RISSE, Pierre SENELLART, Patrick SIEHNDEL et Yannis STAVRAKAS : ARCOMEM Crawling Architecture. *Future Internet*, 6(3):518–541, août 2014.

Fabian M. SUCHANEK, Serge ABITEBOUL et Pierre SENELLART : PARIS: Probabilistic Alignment of Relations, Instances, and Schema. *Proceedings of the VLDB Endowment*, 5(3):157–168, décembre 2011. Présenté à la conférence VLDB 2012, Istanbul, Turquie. Également CoRR abs/1111.7164.

Benjamin NGUYEN, Antoine VION, François-Xavier DUDOUET, Dario COLAZZO, Ioana MANOLESCU et Pierre SENELLART : XML Content Warehousing: Improving Sociological Studies of Mailing Lists and Web Data. *Bulletin of Sociological Methodology*, 112(1):5–31, octobre 2011. Également CoRR abs/1108.4596.

Serge ABITEBOUL, T-H. Hubert CHAN, Evgeny KHARLAMOV, Werner NUTT et 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 et Pierre SENELLART : Probabilistic XML via Markov Chains. *Proceedings of the VLDB Endowment*, 3(1):770–781, septembre 2010. Présenté à la conférence VLDB 2010, Singapour.

Georg GOTTLÖB et Pierre SENELLART : Schema Mapping Discovery from Data Instances. *Journal of the ACM*, 57(2), janvier 2010.

Serge ABITEBOUL, Benny KIMELFELD, Yehoshua SAGIV et Pierre SENELLART : On the Expressiveness of Probabilistic XML Models. *VLDB Journal*, 18(5):1041–1064, octobre 2009.

Vincent D. BLONDEL, Anahí GAJARDO, Maureen HEYMANS, Pierre SENELLART et Paul Van DOOREN : A measure of similarity between graph vertices: applications to synonym extraction and Web searching. *SIAM Review*, 46(4):647–666, 2004. Également CoRR cs.IR/0407061.

## Articles de conférence (avec revue par les pairs)

- Angelo SAADEH, Pierre SENELLART et Stéphane BRESSAN : Confidential Truth Finding with Multi-Party Computation. *In Proc. DEXA*, Penang, Malaisie, août 2023.
- Antoine GAUQUIER et Pierre SENELLART : Automatically Inferring the Document Class of a Scientific Article. *In Proc. DocEng*, Limerick, Irlande, août 2023.
- Shrey MISHRA, Lucas PLUVINAGE et Pierre SENELLART : Towards Extraction of Theorems and Proofs in Scholarly Articles. *In Proc. DocEng*, Limerick, Irlande, août 2021.
- Yann RAMUSAT, Silviu MANIU et Pierre SENELLART : Provenance-Based Algorithms for Rich Queries over Graph Databases. *In Proc. EDBT*, pages 73–84, Nicosie, Chypre, mars 2021.
- Pierre SENELLART : Provenance in Databases: Principles and Applications. *In Proc. RW*, pages 104–109, Bolzano, Italie, septembre 2019.
- Debabrota BASU, Pierre SENELLART et Stéphane BRESSAN : BelMan: Bayesian Bandits on the Belief–Reward Manifold. *In Proc. ECML/PKDD*, pages 167–183, Würzburg, Allemagne, septembre 2019.
- Silviu MANIU, Pierre SENELLART et Suraj JOG : An Experimental Study of the Treewidth of Real-World Graph Data. *In Proc. ICDT*, pages 12 :1–12 :18, Lisbonne, Portugal, mars 2019.
- Miyoung HAN, Pierre-Henri WUILLEMIN et Pierre SENELLART : Focused Crawling through Reinforcement Learning. *In Proc. ICWE*, pages 261–278, Cáceres, Espagne, juin 2018.
- Ben SPENCER, Michael BENEDIKT et Pierre SENELLART : Form Filling based on Constraint Solving. *In Proc. ICWE*, pages 95–113, Cáceres, Espagne, juin 2018.
- Antoine AMARILLI, Mikaël MONET et Pierre SENELLART : Connecting Width and Structure in Knowledge Compilation. *In Proc. ICDT*, Vienne, Autriche, mars 2018.
- Antoine AMARILLI, M. Lamine BA, Daniel DEUTCH et Pierre SENELLART : Possible and Certain Answers for Queries over Order-Incomplete Data. *In Proc. TIME*, pages 4 :1–4 :19, Mons, Belgique, octobre 2017.
- Antoine AMARILLI, Mikaël MONET et Pierre SENELLART : Conjunctive Queries on Probabilistic Graphs: Combined Complexity. *In Proc. PODS*, pages 217–232, Chicago, USA, mai 2017. Également CoRR abs/1703.03201.
- Antoine AMARILLI, Pierre BOURHIS, Mikaël MONET et Pierre SENELLART : Combined Tractability of Query Evaluation via Tree Automata and Cycluits. *In Proc. ICDT*, pages 6 :1–6 :19, Venise, Italie, mars 2017.
- Antoine AMARILLI, Yael AMSTERDAMER, Tova MILO et Pierre SENELLART : Top-k Queries on Unknown Values under Order Constraints. *In Proc. ICDT*, pages 5 :1–5 :18, Venise, Italie, mars 2017.
- Miyoung HAN, Pierre SENELLART, Stéphane BRESSAN et Huayu WU : Routing an Autonomous Taxi with Reinforcement Learning. *In Proc. CIKM*, Indianapolis, USA, octobre 2016. Session industrielle, article court.
- Antoine AMARILLI, Pierre BOURHIS et Pierre SENELLART : Tractable Lineages on Treelike Instances: Limits and Extensions. *In Proc. PODS*, pages 355–370, San Francisco, USA, juin 2016.
- Muhammad FAHEEM et Pierre SENELLART : Adaptive Web Crawling through Structure-Based Link Classification. *In Proc. ICADL*, pages 39–51, Seoul, Corée du Sud, décembre 2015.
- David MONTROYA, Serge ABITEBOUL et Pierre SENELLART : Hup-Me: Inferring and Reconciling a Timeline of User Activity with Smartphone and Personal Data. *In Proc. SIGSPATIAL*, Seattle, USA, novembre 2015. Article court.
- Debabrota BASU, Qian LIN, Weidong CHEN, Hoang Tam VO, Zihong YUAN, Pierre SENELLART et Stéphane BRESSAN : Cost-Model Oblivious Database Tuning with Reinforcement Learning. *In Proc. DEXA*, pages 253–268, Valence, Espagne, septembre 2015.
- Siyu LEI, Silviu MANIU, Luyi MO, Reynold CHENG et Pierre SENELLART : Online Influence Maximization. *In Proc. KDD*, pages 645–654, Sydney, Australie, août 2015.
- Antoine AMARILLI, Pierre BOURHIS et Pierre SENELLART : Provenance Circuits for Trees and Treelike Instances. *In Proc. ICALP*, pages 56–68, Kyoto, Japon, juillet 2015.
- Changping MENG, Reynold CHENG, Silviu MANIU, Pierre SENELLART et Wangda ZHANG : Discovering Meta-Paths in Large Heterogeneous Information Networks. *In Proc. WWW*, pages 754–764, Florence, Italie, mai 2015.
- Georges GOURITEN, Silviu MANIU et Pierre SENELLART : Scalable, Generic, and Adaptive Systems for Focused Crawling. *In Proc. Hypertext*, pages 35–45, Santiago, Chili, septembre 2014. Prix Douglas Engelbart du meilleur article.
- Ruiming TANG, Antoine AMARILLI, Pierre SENELLART et Stéphane BRESSAN : Get a Sample for a Discount: Sampling-Based XML Data Pricing. *In Proc. DEXA*, pages 20–34, Munich, Allemagne, septembre 2014.
- M. Lamine BA, Talel ABDESSALEM et Pierre SENELLART : Uncertain Version Control in Open Collaborative Editing of Tree-Structured Documents. *In Proc. DocEng*, pages 27–36, Florence, Italie, septembre 2013.
- Muhammad FAHEEM et Pierre SENELLART : Intelligent and Adaptive Crawling of Web Applications for Web Archiving. *In Proc. ICWE*, pages 306–322, Aalborg, Danemark, juillet 2013.



Yael AMSTERDAMER, Yael GROSSMAN, Tova MILO et Pierre SENELLART : Crowd Mining. *In Proc. SIGMOD*, pages 241–252, New York, USA, juin 2013.

Asma SOUHLI et Pierre SENELLART : Optimizing Approximations of DNF Query Lineage in Probabilistic XML. *In Proc. ICDE*, pages 721–732, Brisbane, Australie, avril 2013.

Michael BENEDIKT, Pierre BOURHIS et Pierre SENELLART : Monadic Datalog Containment. *In Proc. ICALP*, pages 79–91, Warwick, Royaume-Uni, juillet 2012.

Jiaheng LU, Pierre SENELLART, Chunbin LIN, Xiaoyong DU, Shan WANG et Xinxing CHEN : Optimal Top-k Generation of Attribute Combinations based on Ranked Lists. *In Proc. SIGMOD*, pages 409–420, Scottsdale, USA, mai 2012.

Georges GOURITEN et Pierre SENELLART : API Blender: A Uniform Interface to Social Platform APIs. *In Proc. WWW*, Lyon, France, avril 2012. Session développeurs. Également CoRR abs/1301.2086.

Serge ABITEBOUL, Pierre SENELLART et Victor VIANU : The ERC Webdam on Foundations of Web Data Management. *In Proc. WWW*, pages 211–214, Lyon, France, avril 2012. Session Projets européens.

Serge ABITEBOUL, Yael AMSTERDAMER, Daniel DEUTCH, Tova MILO et Pierre SENELLART : Finding Optimal Probabilistic Generators for XML Collections. *In Proc. ICDT*, pages 127–139, Berlin, Allemagne, mars 2012.

Michael BENEDIKT, Georg GOTTLÖB et Pierre SENELLART : Determining Relevance of Accesses at Runtime. *In Proc. PODS*, pages 211–222, Athens, Grèce, juin 2011.

Serge ABITEBOUL, T-H. Hubert CHAN, Evgeny KHARLAMOV, Werner NUTT et Pierre SENELLART : Aggregate Queries for Discrete and Continuous Probabilistic XML. *In Proc. ICDT*, pages 50–61, Lausanne, Suisse, mars 2010.

Alban GALLAND, Serge ABITEBOUL, Amélie MARIAN et Pierre SENELLART : Corroborating Information from Disagreeing Views. *In Proc. WSDM*, pages 131–140, New York, USA, février 2010.

Pierre SENELLART et Georg GOTTLÖB : On the Complexity of Deriving Schema Mappings from Database Instances. *In Proc. PODS*, pages 23–32, Vancouver, Canada, juin 2008.

Yann OLLIVIER et Pierre SENELLART : Finding Related Pages Using Green Measures: An Illustration with Wikipedia. *In Proc. AAAI*, pages 1427–1433, Vancouver, Canada, juillet 2007.

Pierre SENELLART et Serge ABITEBOUL : On the Complexity of Managing Probabilistic XML Data. *In Proc. PODS*, pages 283–292, Pékin, Chine, juin 2007.

Serge ABITEBOUL et Pierre SENELLART : Querying and Updating Probabilistic Information in XML. *In Proc. EDBT*, pages 1059–1068, Munich, Allemagne, mars 2006. Article court.

François-Xavier DUDOUET, Ioana MANOLESCU, Benjamin NGUYEN et Pierre SENELLART : XML Warehousing Meets Sociology. *In Proc. IADIS ICWI*, pages 170–174, Lisbonne, Portugal, octobre 2005.

Mats ATTNÄS, Pierre SENELLART et Jean SENELLART : Integration of SYSTRAN MT systems in an open workflow. *In Proc. MT Summit*, Phuket, Thaïlande, septembre 2005.

Pierre SENELLART : Identifying Websites with Flow Simulation. *In Proc. ICWE*, pages 124–129, Sydney, Australie, juillet 2005.

#### Articles de workshop (avec revue par les pairs)

Shrey MISHRA, Yacine BRIHMOUCHE, Théo DELEMAZURE, Antoine GAUQUIER et Pierre SENELLART : First Steps in Building a Knowledge Base of Mathematical Results. *In Proc. SDP*, Bangkok, Thaïlande, août 2024.

Shufan JIANG et Pierre SENELLART : Extracting Definienda in Mathematical Scholarly Articles with Transformers. *In Proc. WIESP*, en ligne, Indonésie, novembre 2023.

Yann RAMUSAT, Silviu MANIU et Pierre SENELLART : Efficient Provenance-Aware Querying of Graph Databases with Datalog. *In Proc. GRADES-NDA*, Philadelphie, USA, juin 2022.

Yann RAMUSAT, Silviu MANIU et Pierre SENELLART : Semiring Provenance over Graph Databases. *In Proc. TaPP*, Londres, Royaume-Uni, juillet 2018.

David MONTOYA, Thomas Pellissier TANON, Serge ABITEBOUL, Pierre SENELLART et Fabian M. SUCHANEK : A Knowledge Base for Personal Information Management. *In Proc. LDOW*, Lyon, France, avril 2018.

Ognjen SAVKOVIC, Evgeny KHARLAMOV, Werner NUTT et Pierre SENELLART : Towards Approximating Incomplete Queries over Partially Complete Databases. *In Proc. AMW*, Montevideo, Uruguay, juin 2017.

M. Lamine BA, Roxana HORINCAR, Pierre SENELLART et Huayu WU : Truth Finding with Attribute Partitioning. *In Proc. WebDB*, pages 27–33, Melbourne, Australie, mai 2015.

Marilena OITA et Pierre SENELLART : FOREST: Focused Object Retrieval by Exploiting Significant Tag Paths. *In Proc. WebDB*, pages 55–61, Melbourne, Australie, mai 2015.

Silviu MANIU, Reynold CHENG et Pierre SENELLART : ProbTree: A Query-Efficient Representation of Probabilistic Graphs. *In Proc. BUDA*, Snowbird, USA, juin 2014. Workshop sans actes formels.

- M. Lamine BA, Talel ABDESSALEM et Pierre SENELLART : Merging Uncertain Multi-Version XML Documents. *In Proc. DChanges*, Florence, Italie, septembre 2013.
- Thomas RISSE, Wim PETERS et Pierre SENELLART : Social and Semantic Driven Web Harvesting. *In Proc. Building Web Observatories*, Paris, France, mai 2013.
- Marilena OITA, Antoine AMARILLI et Pierre SENELLART : Cross-Fertilizing Deep Web Analysis and Ontology Enrichment. *In Proc. VLDS*, pages 5–8, Istanbul, Turquie, août 2012. Article décrivant une vision.
- Meghyn BIENVENU, Daniel DEUTCH, Davide MARTINENGI, Pierre SENELLART et Fabian M. SUCHANEK : Dealing with the Deep Web and all its Quirks. *In Proc. VLDS*, pages 21–24, Istanbul, Turquie, août 2012. Article décrivant une vision.
- M. Lamine BA, Talel ABDESSALEM et Pierre SENELLART : Towards a Version Control Model with Uncertain Data. *In Proc. PIKM*, Glasgow, Royaume-Uni, octobre 2011.
- Marilena OITA et Pierre SENELLART : Deriving Dynamics of Web Pages: A Survey. *In Proc. TFAW*, pages 25–32, Hyderabad, Inde, mars 2011.
- Evgeny KHARLAMOV, Werner NUTT et Pierre SENELLART : Value Joins are Expensive over (Probabilistic) XML. *In Proc. LID*, pages 41–48, Uppsala, Suède, mars 2011.
- Marilena OITA et Pierre SENELLART : Archiving Data Objects Using Web Feeds. *In Proc. IFAW*, pages 31–41, Vienne, Autriche, septembre 2010.
- Evgeny KHARLAMOV, Werner NUTT et Pierre SENELLART : Updating Probabilistic XML. *In Proc. Updates in XML*, Lausanne, Suisse, mars 2010.
- Serge ABITEBOUL, Sihem AMER-YAHIA, Amélie MARIAN, Alban GALLAND et Pierre SENELLART : Birds of a tag flock together. *In Proc. SSM*, New York, USA, février 2010. Workshop sans actes formels.
- Marc SPANIOL, Dimitar DENEV, Arturas MAZEIKA, Pierre SENELLART et Gerhard WEIKUM : Data Quality in Web Archiving. *In Proc. WICOW*, pages 19–26, Madrid, Espagne, avril 2009.
- Pierre SENELLART, Avin MITTAL, Daniel MUSCHICK, Rémi GILLERON et Marc TOMMASI : Automatic Wrapper Induction from Hidden-Web Sources with Domain Knowledge. *In Proc. WIDM*, pages 9–16, Napa, USA, octobre 2008.
- Vincent D. BLONDEL et Pierre SENELLART : Automatic extraction of synonyms in a dictionary. *In Proc. TMW*, pages 7–13, Arlington, USA, avril 2002. Également rapport technique n° 89 (2001), Université catholique de Louvain, Louvain-la-neuve.

#### Démonstrations (avec revue par les pairs)

- Pierre SENELLART, Louis JACHET, Silviu MANIU et Yann RAMUSAT : ProvSQL: Provenance and Probability Management in PostgreSQL. *In Proc. VLDB*, pages 2034–2037, Rio de Janeiro, Brésil, août 2018. Démonstration.
- M. Lamine BA, Sébastien MONTENEZ, Talel ABDESSALEM et Pierre SENELLART : Monitoring moving objects using uncertain Web data. *In Proc. SIGSPATIAL*, pages 565–568, Dallas, USA, novembre 2014. Démonstration.
- Muhammad FAHEEM et Pierre SENELLART : Demonstrating Intelligent Crawling and Archiving of Web Applications. *In Proc. CIKM*, pages 2481–2484, San Francisco, USA, octobre 2013. Démonstration.
- Yael AMSTERDAMER, Yael GROSSMAN, Tova MILO et Pierre SENELLART : Crowd Miner: Mining association rules from the crowd. *In Proc. VLDB*, pages 1250–1253, Riva del Garda, Italie, août 2013. Démonstration.
- Asma SOUHLI et Pierre SENELLART : Demonstrating ProApproX 2.0: A Predictive Query Engine for Probabilistic XML. *In Proc. CIKM*, pages 2734–2736, Maui, Hawaii, USA, octobre 2012. Démonstration.
- Serge ABITEBOUL, Yael AMSTERDAMER, Tova MILO et Pierre SENELLART : Auto-Completion Learning for XML. *In Proc. SIGMOD*, pages 669–672, Scottsdale, USA, mai 2012. Démonstration.
- Michael BENEDIKT, Tim FURCHE, Andreas SAVVIDES et Pierre SENELLART : ProFoUnd: Program-analysis-based Form Understanding. *In Proc. WWW*, pages 313–316, Lyon, France, avril 2012. Démonstration.
- Pierre SENELLART et Asma SOUHLI : ProApproX: A Lightweight Approximation Query Processor over Probabilistic Trees. *In Proc. SIGMOD*, pages 1295–1298, Athens, Grèce, juin 2011. Démonstration.
- Talel ABDESSALEM, M. Lamine BA et Pierre SENELLART : A Probabilistic XML Merging Tool. *In Proc. EDBT*, pages 538–541, Uppsala, Suède, mars 2011. Démonstration.

#### Posters (avec revue par les pairs)

- Thomas RISSE, Stefan DIETZE, Wim PETERS, Katerina DOKA, Yannis STAVRAKAS et Pierre SENELLART : Exploiting the Social and Semantic Web for guided Web Archiving. *In Proc. TPDFL*, pages 426–432, Paphos, Chypre, septembre 2012. Affiche.
- Benjamin NGUYEN, Antoine VION, François-Xavier DUDOUET, Ioana MANOLESCU, Dario COLAZZO et Pierre SENELLART : The WebStand Project. *In Proc. WebSci*, Athènes, Grèce, mars 2009. Également CoRR abs/1002.0971 (Affiche).

Michalis VAZIRGIANNIS, Dimitris DROSOS, Pierre SENELLART et Akrivi VLACHOU : Web Page Rank Prediction with Markov Models. *In Proc. WWW*, pages 1075–1076, Pékin, Chine, avril 2008. Affiche.

## Expérience dans l'industrie

- novembre 2004 – **Auteur logiciel**, *SYSTRAN S.A.*, La Défense, France.  
février 2005 SYSTRAN Mobile pour Pocket PC.
- septembre 2003 – **Ingénieur R&D**, *SYSTRAN S.A.*, La Défense, France.  
septembre 2004 SYSTRAN version 5.
- septembre 2002 – **Ingénieur logiciel**, *Le Robert*, Paris, France.  
octobre 2002 Analyse du graphe du dictionnaire et extraction automatique de mots similaires.

## Informatique pratique

Langages de programmation	C++; Perl; Java; Scala; C; Python; PHP	Systèmes d'exploitation	Linux; Unix; Windows
Technologies XML	DOM; SAX; DTD; XML Schema; XPath; XSLT; XQuery; XSL-FO	Web	(X)HTML; CSS; JavaScript; PHP; JSP; SVG
SGBD	PostgreSQL; MySQL; Oracle	Divers	L <sup>A</sup> T <sub>E</sub> X; Unicode; Hadoop

## Langues

français	langue maternelle	anglais	courant
allemand	lu, écrit, parlé	arabe littéraire	notions
malais	notions	grec ancien	niveau lycée