

Antoine Amarilli

✉ a3nm@a3nm.net

📧 a3nm.net

Né le 7 février 1990 à Colmar (68), France

Nationalité française

Ma thèse, soutenue en mars dans l'équipe DBWeb de Télécom ParisTech, porte sur les fondements de la **gestion de données incertaines**. Ses résultats ont été publiés (ou acceptés pour publication) dans les conférences majeures de **communautés variées** : théorie des bases de données (**ICDT'14**, **PODS'16**), logique informatique (**LICS'15**), automates et algorithmique (**ICALP'15**) et intelligence artificielle (**IJCAI'15**, **IJCAI'16**), ainsi que **neuf** autres publications avec comité de lecture.

J'ai par ailleurs une forte expérience de mobilité avec **plus d'un an de séjours à l'étranger** (dans quatre pays) avant et pendant ma thèse, **trois collaborations internationales** en cours (deux avec Tel Aviv, une avec Oxford), et **27 exposés dont 13 à l'étranger** donnés depuis le début de ma thèse. J'ai également enseigné pendant trois ans, y compris des **cours de niveau M2**, relu des articles pour 12 conférences et journaux, remporté des prix à des **concours de programmation informatique**, et fait l'expérience de l'industrie chez Google où j'ai publié un **brevet**.

Études

- 2013–2016 **Doctorat**, “*Tirer parti de la structure des données incertaines*”, Télécom ParisTech.
Encadré par Pierre Senellart, commencé le 1^{er} septembre 2013.
Soutenu le **14 mars 2016**, mention très honorable; allocation jusqu'au 31 août 2016.
- 2012–2013 **Stages de recherche à l'étranger**, *Tel Aviv (Israël) et Oxford (Royaume-Uni)*.
Obtention d'un financement de thèse par l'École normale supérieure (allocation spécifique).
- 2010–2012 **Master parisien de recherche en informatique (MPRI)**, *École normale supérieure*.
Mention très bien, classé 5^e sur 58 étudiants. Stage de M2 avec Pierre Senellart à Télécom ParisTech.
- 2009–2010 **Licence d'informatique**, *École normale supérieure*. *Mention très bien*.
- 2007–2009 **Classes préparatoires MPSI-MP***, *Lycée Kléber, Strasbourg*.
Admis à l'École normale supérieure par le concours INFO, classé 2^e.
- 2007 **Baccalauréat scientifique**, *Lycée des Pontonniers, Strasbourg*. *Mention très bien*.

Mobilité

- 2014 et 2015 Visites de recherche à **National University of Singapore** (2 mois)
Dans le cadre de mon travail avec Tang Ruiming et Stéphane Bressan [TAS+16; TAS+14] sur la tarification (*data pricing*) des documents XML, j'ai effectué **deux séjours d'un mois à Singapour** pendant ma thèse.
- 2013 Stage à l'**université d'Oxford** avec **Michael Benedikt** (5 mois)
Pendant ma quatrième année à l'École normale supérieure, j'ai effectué deux stages de recherche à l'étranger. J'ai travaillé pendant **5 mois à Oxford** avec Michael Benedikt sur la réponse aux requêtes en monde ouvert. Notre travail a été publié à **IJCAI'15** [AB15a] et **LICS'15** [AB15b]. J'ai poursuivi cette collaboration en séjournant **trois semaines à Oxford** pendant ma thèse, en 2014 et 2015. Cette collaboration se poursuit toujours, notamment avec un travail à **IJCAI'16** [ABB+16] sur la réponse aux requêtes sous relations d'ordre.
- 2012–2013 Stage à **Tel Aviv University** avec **Tova Milo** (3 mois)
Pendant ma quatrième année à l'ENS, j'ai également séjourné **3 mois à Tel Aviv** pour travailler avec Tova Milo et Yael Amsterdamer sur l'interrogation de la foule (*crowdsourcing*). Notre travail a été publié à **ICDT'14** [AAM14a]. Nous avons poursuivi notre collaboration en rédigeant un papier vision [AAM14b] que j'ai développé lors d'une **visite à Tel Aviv en 2014**; notre nouveau travail est actuellement en soumission.
- 2011 Stage à **Google New York** avec **Cyril Allauzen** (4 mois)
J'ai effectué mon stage de M1 à Google New York sur l'implémentation de techniques de reconnaissance vocale avec des transducteurs à états finis pondérés, pour des méthodes de décodage par minimisation du risque bayésien (*Minimum Bayes Risk decoding*). Ce travail a été valorisé sous la forme d'un **brevet** [AAM15].

Publications avec comité de lecture

- [AAM14a] Antoine AMARILLI, Yael AMSTERDAMER et Tova MILO. “On the Complexity of Mining Itemsets from the Crowd Using Taxonomies”. In : *Proc. ICDT*. 2014.
- [AAM14b] Antoine AMARILLI, Yael AMSTERDAMER et Tova MILO. “Uncertainty in Crowd Data Sourcing Under Structural Constraints”. In : *Proc. UnCrowd*. 2014.
- [AAM15] Antoine AMARILLI, Cyril ALLAUZEN et Mehryar MOHRI. *Minimum Bayesian Risk Methods for Automatic Speech Recognition*. United States Patent 9123333. 2015.
- [AB15a] Antoine AMARILLI et Michael BENEDIKT. “Combining Existential Rules and Description Logics”. In : *Proc. IJCAI*. 2015.
- [AB15b] Antoine AMARILLI et Michael BENEDIKT. “Finite Open-World Query Answering with Number Restrictions”. In : *Proc. LICS*. 2015.
- [ABB+16] Antoine AMARILLI, Michael BENEDIKT, Pierre BOURHIS et Michael VANDEN BOOM. “Query Answering with Transitive and Linear-Ordered Data”. In : *Proc. IJCAI*. To appear. 2016.
- [ABS15] Antoine AMARILLI, Pierre BOURHIS et Pierre SENELLART. “Provenance Circuits for Trees and Treelike Instances”. In : *Proc. ICALP*. 2015.
- [Ama14] Antoine AMARILLI. “The Possibility Problem for Probabilistic XML”. In : *Proc. AMW*. 2014.
- [Ama15a] Antoine AMARILLI. “Possibility for Probabilistic XML”. In : *Ingénierie des Systèmes d’Information* (2015).
- [Ama15b] Antoine AMARILLI. “Structurally Tractable Uncertain Data”. In : *Proc. PhD Symposium of SIGMOD/PODS*. 2015.
- [AS13] Antoine AMARILLI et Pierre SENELLART. “On the Connections between Relational and XML Probabilistic Data Models”. In : *Proc. BNCOD*. 2013.
- [OAS12] Marilena OITA, Antoine AMARILLI et Pierre SENELLART. “Cross-Fertilizing Deep Web Analysis and Ontology Enrichment”. In : *Proc. VLDS*. 2012.
- [TAS+14] Ruiming TANG, Antoine AMARILLI, Pierre SENELLART et Stéphane BRESSAN. “Get a Sample for a Discount”. In : *Proc. DEXA*. 2014.
- [TAS+16] Ruiming TANG, Antoine AMARILLI, Pierre SENELLART et Stéphane BRESSAN. “A Framework for Sampling-Based XML Data Pricing”. In : *Transactions on Large-Scale Data and Knowledge-Centered Systems* (2016).
- [TBA+15] Aliaksandr TALAIIKA, Joanna BIEGA, Antoine AMARILLI et Fabian M. SUCHANEK. “IBEX: Harvesting Entities from the Web Using Unique Identifiers”. In : *Proc. WebDB*. 2015.

Enseignement

J’ai effectué une mission d’enseignement complète de 64 heures par an pendant mes trois années de doctorat.

- 2015–2016 **Uncertain Data Management**, *M2 Data and Knowledge, Université Paris-Saclay*, Conception et enseignement du cours avec Silviu Maniu : cours magistraux, TP (12 heures).
- 2012–2016 **Projet de programmation : problèmes pratiques et concours**, *Télécom ParisTech*, Niveau master ; réalisation de supports, responsable d’un groupe (24 heures/an, trois ans).
- 2012–2016 **Théorie des langages**, *Télécom ParisTech*, enseignement de tronc commun de 1^e année, Responsable d’un groupe : cours magistraux, TDs, TPs (16.5 heures/an, trois ans).
- 2012–2015 **Technologies du Web**, *Master COMASIC (Polytechnique, Télécom, ENSTA ParisTech)*, Conception du cours et enseignement : cours magistraux et TP (12 heures/an, deux ans).
- 2014 **Coach pour le concours ACM ICPC SWERC**, *Télécom ParisTech*, Encadrement des élèves représentant Télécom ParisTech au concours SWERC (Porto, Portugal).

Exposés à l'étranger

Cette section recense les exposés que j'ai donnés à l'étranger. Elle n'inclut pas 14 exposés donnés en France (IRIF, BDA'15, BDA'14, Highlights'14, INRIA Lille, LIRMM, Télécom ParisTech, ENS Cachan).

- 2015 Combining Existential Rules and Description Logics *IJCAI, Buenos Aires*
- 2015 Provenance Circuits for Trees and Treelike Instances *ICALP, Kyoto*
- 2015 Finite Open-World Query Answering with Number Restrictions *LICS, Kyoto*
- 2015 Structurally Tractable Uncertain Data *SIGMOD/PODS PhD Symposium, Melbourne*
- 2015 Harvesting Entities from the Web Using Unique Identifiers *WebDB, Melbourne*
- 2015 Top-k Querying of Incomplete Data under Order Constraints *Réunion du projet CCIPX, Hanoi*
- 2015 Provenance Circuits for Trees and Treelike Instances *Séminaire d'équipe, NUS, Singapour*

- 2014 Open-World Finite Query Answering Under Number Restrictions *Dagstuhl*
Invité au séminaire Dagstuhl « *Querying and Reasoning Under Expressive Constraints* »
- 2014 The Possibility Problem for Probabilistic XML *AMW, Cartagena, Colombie*
- 2014 Uncertainty in Crowd Data Sourcing Under Structural Constraints *UnCrowd, Bali*
- 2014 On the Complexity of Mining Itemsets from the Crowd Using Taxonomies
Séminaire « Web Data Management » à Tel Aviv University ; puis conférence ICDDT à Athènes

- 2013 On the Connections between Relational and XML Probabilistic Data Models *BNCOD, Oxford*

Relecture pour des journaux et conférences

- 2018 International Conference on Database Theory (ICDT), membre du comité de programme
- 2017 International Conference on Database Theory (ICDT)
- 2016 Bases de Données Avancées (BDA), membre du comité de programme des démos
- 2016 Scalable Uncertainty Management (SUM), membre du comité de programme
- 2016 *VLDB Journal*
- 2016 Principles of Database Systems (PODS)
- 2015 *Journal of Logic and Computation*
- 2015 *Review of Symbolic Logic*
- 2015 International Conference on Data Engineering (ICDE)
- 2014 *Distributed and Parallel Databases*
- 2014 Latin American Theoretical Informatics Symposium (LATIN)
- 2013 ACM Special Interest Group on Management of Data (SIGMOD)

Langues

- Français Langue maternelle *Deuxième place à la Dictée ECE 2012.*
- Anglais Courant *Diplôme IGCSE, Option internationale au baccalauréat.*
- Allemand Notions *Étudié au collège et au lycée.*

Distinctions

- 2015 Première place au **concours de programmation** Google Hash Code.
- 2010 et 2011 Médaille d'argent au **concours de programmation** ACM ICPC SWERC.
- 2010 Deuxième place au **concours national d'informatique** Prolog.in.
- 2008 Première place au **concours national d'informatique** Prolog.in.
- 2007 Premier accessit au **concours général de philosophie**.