



Florian Boudin

Mis à jour - last update: 2015-03-18

Formation - Education

- 2006-2008 **Doctorat en Informatique - Ph.D. in Computer Science.**
Université d'Avignon et des Pays de Vaucluse, France
Directeurs - *Advisors* : Juan-Manuel Torres-Moreno, Marc El-Bèze
- 2004-2006 **Master Recherche Informatique - M.S. in Computer Science.**
Université d'Avignon et des Pays de Vaucluse, France
- 2000-2004 **Licence Mathématique et Informatique - B.S. in Mathematics and Computer Science.**
Université d'Avignon et des Pays de Vaucluse, France
University of Birmingham, UK (ERASMUS Programme)

Parcours professionnel - Employment history

- 2011-Présent **Maître de conférences - Associate professor.**
Université de Nantes, France
- 2010-2011 **ATER - Assistant professor.**
Université d'Avignon et des Pays de Vaucluse, France
- 2009-2010 **Stagiaire post-doctoral - Postdoctoral fellow.**
Université de Montréal, Canada
- 2006-2009 **Assistant de recherche - Research assistant.**
Université d'Avignon et des Pays de Vaucluse, France

Intérêts de recherche - Research interests

- **Traitement Automatique de la Langue - Natural Language Processing**
 - Résumé automatique - *summarization*, extraction de termes-clés - *keyphrase extraction*

Compétences - Skills

- **Programmation - Programming** : Python, C, Java, HTML/CSS, PHP, SQL, \LaTeX
- **Langues - Languages** : Français - *French* (natif - *native speaker*), Anglais - *English* (bon - *good*)

Encadrement - Students

- **Etudiants en Doctorat - Ph.D. Students**
 - Adrien Bougouin (expected graduation \sim 2015)
- **Etudiants en Master - M.S. Students**
 - Rémi Bois (2014; now Ph.D. student at Université de Nantes)

Service – Service

- Conference Program Chair : RECITAL 2013
- Conference PCs : COLING 2012, LACL 2012, RECITAL (2010, 2014, 2015)
- Conference Reviews: AAAI 2015, TALN (2013, 2014, 2015) JAD'12, KDIR (2009, 2010, 2011, 2012), RECITAL 2011, CIKM 2011
- Journal Reviews: Traitement Automatique des Langues, Information Retrieval

Enseignement - Teaching experience

- **Université de Nantes, France** 800h+
 - Programmation, bases de données, apprentissage automatique, technologies web
- **Université d'Avignon et des Pays de Vaucluse, France** 96h
 - Programmation, informatique décisionnelle
- **Université de Montréal, Canada** 54h
 - Atelier en informatique
- **IUT d'Avignon, France** 100h
 - Programmation, programmation web, bases de données

Projets - Funding

2015 **GOLEM - Approche par Optimisation pour l'Extraction de Mots-clés.**
CNRS - PEPS INS2I/INSMI 2015, 6k€ over one year. Sole PI.

Publications - Publications

Journal Articles

- [1] **Adrien Bougouin, Florian Boudin**, TopicRank : ordonnancement de sujets pour l'extraction automatique de termes-clés, *Traitement Automatique des Langues (TAL 55-1)*, 2014.
- [2] **Romain Deveaud, Florian Boudin**, De quoi parle ce Tweet? Résumer Wikipédia pour contextualiser des microblogs, *The Information - Intelligence - Interaction (I3) Journal*, 2014.
- [3] **Florian Boudin, Jian-Yun Nie, Joan Bartlett, Roland Grad, Pierre Pluye, Martin Dawes**, Combining classifiers for robust PICO element detection, *BMC Medical Informatics and Decision Making*, 2010.

Conference Papers

- [4] **Ophélie Lacroix, Denis Béchet, Florian Boudin**, Label Pre-annotation for Building Non-projective Dependency Treebanks for French, *Conference on Intelligent Text Processing and Computational Linguistics (CICLing)*, 2014.
- [5] **Adrien Bougouin, Florian Boudin, Béatrice Daille**, TopicRank: Graph-Based Topic Ranking for Keyphrase Extraction, *International Joint Conference on Natural Language Processing (IJCNLP)*, 2013.
- [6] **Florian Boudin**, A Comparison of Centrality Measures for Graph-Based Keyphrase Extraction, *International Joint Conference on Natural Language Processing (IJCNLP)*, 2013, short paper.
- [7] **Florian Boudin, Emmanuel Morin**, Keyphrase Extraction for N-best Reranking in Multi-Sentence Compression, *North American Chapter of the Association for Computational Linguistics (NAACL)*, 2013.
- [8] **Florian Boudin, Jian-Yun Nie, Martin Dawes**, Using a Medical Thesaurus to Predict Query Difficulty, *European Conference on Information Retrieval (ECIR)*, 2012, short paper.
- [9] **Florian Boudin, Stéphane Huet, Juan-Manuel Torres-Moreno**, A Graph-based Approach to Cross-language Multi-document Summarization, *Conference on Intelligent Text Processing and Computational Linguistics (CICLing)*, 2011.
- [10] **Florian Boudin, Jian-Yun Nie, Martin Dawes**, Positional Language Models for Clinical Information Retrieval, *Empirical Methods in Natural Language Processing (EMNLP)*, 2010.
- [11] **Florian Boudin, Lixin Shi, Jian-Yun Nie**, Improving Medical Information Retrieval with PICO Element Detection, *European Conference on Information Retrieval (ECIR)*, 2010.
- [12] **Florian Boudin, Jian-Yun Nie, Martin Dawes**, Clinical Information Retrieval using Document and PICO Structure, *North American Chapter of the Association for Computational Linguistics (NAACL)*, 2010.
- [13] **Florian Boudin, Juan-Manuel Torres-Moreno, Marc El-Bèze**, Mixing Statistical and Symbolic Approaches for Chemical Names Recognition, *Conference on Intelligent Text Processing and Computational Linguistics (CICLing)*, 2008.

- [14] **Florian Boudin, Juan-Manuel Torres-Moreno, Patricia Velazquez-Morales**, An Efficient Statistical Approach for Automatic Organic Chemistry Summarization, *International Conference on Natural Language Processing (GoTAL)*, 2008.
- [15] **Florian Boudin, Juan-Manuel Torres-Moreno, Marc El-Bèze**, A Scalable MMR Approach to Sentence Scoring for Multi-Document Update Summarization, *International Conference on Computational Linguistics (COLING)*, 2008, *short paper*.
- [16] **Florian Boudin, Juan-Manuel Torres-Moreno**, A Cosine Maximization Minimization approach for User Oriented Multi-Document Update Summarization, *Recent Advances in Natural Language Processing (RANLP)*, 2007.
- [17] **Florian Boudin, Juan-Manuel Torres-Moreno**, NEO-CORTEX: a performant user-oriented multi document summarization system, *Conference on Intelligent Text Processing and Computational Linguistics (CICLing)*, 2007.

National Conference Papers

- [18] **Adrien Bougouin, Florian Boudin, Béatrice Daille**, Influence des domaines de spécialité dans l'extraction de termes-clés, *Traitement Automatique des Langues Naturelles (TALN)*, 2014.
- [19] **Florian Boudin**, TALN Archives : une archive numérique francophone des articles de recherche en Traitement Automatique de la Langue, *Traitement Automatique des Langues Naturelles (TALN)*, 2013, *papier court*.
- [20] **Romain Deveaud, Florian Boudin**, Contextualisation automatique de Tweets à partir de Wikipédia, *Conférence en Recherche d'Information et Applications (CORIA)*, 2013.
- [21] **Nicolas Hernandez, Florian Boudin**, Construction d'un large corpus libre annoté morpho-syntaxiquement en français, *Traitement Automatique des Langues Naturelles (TALN)*, 2013.
- [22] **Florian Boudin, Nicolas Hernandez**, Détection et correction automatique d'erreurs d'annotation morpho-syntaxique du French TreeBank, *Traitement Automatique des Langues Naturelles (TALN)*, 2012.
- [23] **Stéphane Huet, Florian Boudin, Juan-Manuel Torres-Moreno**, Utilisation d'un score de qualité de traduction pour le résumé multi-document cross-lingue, *Traitement Automatique des Langues Naturelles (TALN)*, 2011.
- [24] **Romain Deveaud, Florian Boudin, Eric SanJuan, Patrice Bellot**, Correction de césures et enrichissement de requêtes pour la recherche de livres, *Conférence en Recherche d'Information et Applications (CORIA)*, 2011.
- [25] **Florian Boudin, Juan-Manuel Torres-Moreno**, Résumé automatique multi-document et indépendance de la langue: une première évaluation en français, *Traitement Automatique des Langues Naturelles (TALN)*, 2009, *papier court*.

Workshop Papers

- [26] **Romain Deveaud, Florian Boudin**, Effective Tweet Contextualization with Hashtags Performance Prediction and Multi-Document Summarization, *INitiative for the Evaluation of XML Retrieval (INEX)*, 2013.
- [27] **Florian Boudin, Amir Hazem, Nicolas Hernandez, Prajol Shrestha**, Participation du LINA à DEFT 2012, *Défi Fouille de Textes (DEFT)*, 2012.
- [28] **Romain Deveaud, Florian Boudin**, LIA/LINA at the INEX 2012 Tweet Contextualization track, *INitiative for the Evaluation of XML Retrieval (INEX)*, 2012.
- [29] **Florian Boudin, Jian-Yun Nie, Martin Dawes**, Deriving a Test Collection for Clinical Information Retrieval from Systematic Reviews, *International Workshop on Data and Text Mining in Biomedical Informatics (DTMBIO)*, 2010.
- [30] **Romain Deveaud, Florian Boudin, Eric SanJuan, Patrice Bellot**, LIA at INEX 2010: Ad Hoc, Book and Question Answering Tracks, *INitiative for the Evaluation of XML Retrieval (INEX)*, 2010.

- [31] **Florian Boudin, Marc El-Bèze, Juan-Manuel Torres-Moreno**, The LIA Update Summarization system at TAC-2008, *Text Analysis Conference (TAC)*, 2008.
- [32] **Florian Boudin, Benoît Favre, Frédéric Béchet, Marc El-Bèze, Laurent Gillard, Juan-Manuel Torres-Moreno**, The LIA-Thales summarization system at DUC-2007, *Document Understanding Conference (DUC)*, 2007.
- [33] **Benoît Favre, Frédéric Béchet, Patrice Bellot, Florian Boudin, Marc El-Bèze, Laurent Gillard, Guy Lapalme, Juan-Manuel Torres-Moreno**, The LIA-Thales summarization system at DUC-2006, *Document Understanding Conference (DUC)*, 2006.

Book Chapters

- [34] **Florian Boudin, Juan-Manuel Torres-Moreno**, A Maximization-Minimization Approach for Update Summarization, *Book chapter in Current Issues in Linguistic Theory: Recent Advances in Natural Language Processing*, 2009.

Ph.D. Thesis

- [35] **Florian Boudin**, Exploration d'approches statistiques pour le résumé automatique de texte, *Thèse de doctorat (Ph.D. Thesis)*, 2008.