

# Florian Boudin

Last update: August 21, 2015

---

CONTACT INFORMATION	LINA - Université de Nantes 2 rue de la Houssinière 44322 Nantes, France	+33 2 51 12 59 98 florian.boudin@univ-nantes.fr <a href="http://florianboudin.org">http://florianboudin.org</a>
EMPLOYMENT HISTORY	<b>Associate Professor</b> Université de Nantes, France <b>Assistant Professor</b> Université d'Avignon et des Pays de Vaucluse, France <b>Postdoctoral Fellow</b> Université de Montréal, Canada <b>Research Assistant</b> Université d'Avignon et des Pays de Vaucluse, France	2011 to present 2010-2011 2009-2010 2006-2009
EDUCATION	<b>Ph.D. in Computer Science</b> Université d'Avignon et des Pays de Vaucluse, France Advisors: Juan-Manuel Torres-Moreno, Marc El-Bèze <b>M.S. in Computer Science</b> Université d'Avignon et des Pays de Vaucluse, France <b>B.S. in Mathematics and Computer Science</b> Université d'Avignon et des Pays de Vaucluse, France University of Birmingham, UK (ERASMUS Programme)	2006-2008 2004-2006 2000-2004
STUDENTS	<b>Ph.D. Students</b> ➤ Adrien Bougouin (expected graduation ~ 2015) <b>M.S. Students</b> ➤ Rémi Bois (2014; now Ph.D. student at Université de Rennes)	
PUBLICATIONS	<b>Journal Papers</b>  [1] Adrien Bougouin and Florian Boudin. TopicRank : ordonnancement de sujets pour l'extraction automatique de termes-clés. <i>Traitement Automatique des Langues</i> , 55:1. 2014. [2] Romain Deveaud and Florian Boudin. De quoi parle ce Tweet? Résumer Wikipédia pour contextualiser des microblogs. <i>The Information - Intelligence - Interaction (I3) Journal</i> . 2014. [3] Florian Boudin, Jian-Yun Nie, Joan Bartlett, Roland Grad, Pierre Pluye, and Martin Dawes. Combining classifiers for robust PICO element detection. <i>BMC Medical Informatics and Decision Making</i> , 10:29. 2010.  <b>Conference Papers</b>  [4] Florian Boudin, Hugo Mougard, and Benoit Favre. Concept-based Summarization using Integer Linear Programming: From Concept Pruning to Multiple Optimal Solutions. <i>Conference on Empirical Methods in Natural Language Processing (EMNLP)</i> . 2015. [5] Ophélie Lacroix, Denis Béchet, and Florian Boudin. Label Pre-annotation for Building Non-projective Dependency Treebanks for French. <i>Conference on Intelligent Text Processing and Computational Linguistics (CICLing)</i> . 2014. [6] Adrien Bougouin, Florian Boudin, and Béatrice Daille. TopicRank: Graph-Based Topic Ranking for Keyphrase Extraction. <i>International Joint Conference on Natural Language Processing (IJCNLP)</i> . 2013. [7] Florian Boudin. A Comparison of Centrality Measures for Graph-Based Keyphrase Extraction. <i>International Joint Conference on Natural Language Processing (IJCNLP)</i> . 2013.	

- [8] Florian Boudin and Emmanuel Morin. Keyphrase Extraction for N-best Reranking in Multi-Sentence Compression. *Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL-HLT)*. 2013.
- [9] Florian Boudin, Jian-Yun Nie, and Martin Dawes. Using a Medical Thesaurus to Predict Query Difficulty. *European Conference on Information Retrieval (ECIR)*. 2012.
- [10] Florian Boudin, Stéphane Huet, and Juan-Manuel Torres-Moreno. A Graph-based Approach to Cross-language Multi-document Summarization. *Conference on Intelligent Text Processing and Computational Linguistics (CICLing)*. 2011.
- [11] Florian Boudin, Jian-Yun Nie, and Martin Dawes. Positional Language Models for Clinical Information Retrieval. *Conference on Empirical Methods in Natural Language Processing (EMNLP)*. 2010.
- [12] Florian Boudin, Jian-Yun Nie, and Martin Dawes. Clinical Information Retrieval using Document and PICO Structure. *Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL-HLT)*. 2010.
- [13] Florian Boudin, Lixin Shi, and Jian-Yun Nie. Improving Medical Information Retrieval with PICO Element Detection. *European Conference on Information Retrieval (ECIR)*. 2010.
- [14] Florian Boudin, Marc El-Bèze, and Juan-Manuel Torres-Moreno. A Scalable MMR Approach to Sentence Scoring for Multi-Document Update Summarization. *International Conference on Computational Linguistics (COLING)*. 2008.
- [15] Florian Boudin, Juan Torres-Moreno, and Marc El-Bèze. Mixing Statistical and Symbolic Approaches for Chemical Names Recognition. *Conference on Intelligent Text Processing and Computational Linguistics (CICLing)*. 2008.
- [16] Florian Boudin, Juan-Manuel Torres-Moreno, and Patricia Velázquez-Morales. An Efficient Statistical Approach for Automatic Organic Chemistry Summarization. *International Conference on Natural Language Processing (GoTAL)*. 2008.
- [17] Florian Boudin and Juan-Manuel Torres-Moreno. A Cosine Maximization Minimization approach for User Oriented Multi-Document Update Summarization. *Recent Advances in Natural Language Processing (RANLP)*. 2007.
- [18] Florian Boudin and Juan Torres Moreno. NEO-CORTEX: A Performant User-Oriented Multi-Document Summarization System. *Conference on Intelligent Text Processing and Computational Linguistics (CICLing)*. 2007.

### **Workshop Papers**

- [19] Florian Boudin. Reducing Over-generation Errors for Automatic Keyphrase Extraction using Integer Linear Programming. *Workshop on Novel Computational Approaches to Keyphrase Extraction*. 2015.
- [20] Emmanuel Morin, Amir Hazem, Florian Boudin, and Elizaveta Loginova-Clouet. LINA: Identifying Comparable Documents from Wikipedia. *Eighth Workshop on Building and Using Comparable Corpora*. 2015.
- [21] Romain Deveaud and Florian Boudin. Effective Tweet Contextualization with Hashtags Performance Prediction and Multi-Document Summarization. *INitiative for the Evaluation of XML Retrieval (INEX)*. 2013.
- [22] Romain Deveaud and Florian Boudin. LIA/LINA at the INEX 2012 Tweet Contextualization track. *INitiative for the Evaluation of XML Retrieval (INEX)*. 2012.
- [23] Florian Boudin, Amir Hazem, Nicolas Hernandez, and Prajol Shrestha. Participation du LINA à DEFT2012 (LINA at DEFT2012). *Défi Fouille de Textes (DEFT)*. 2012.
- [24] Romain Deveaud, Florian Boudin, and Patrice Bellot. LIA at INEX 2010 Book Track. *INitiative for the Evaluation of XML Retrieval (INEX)*. 2011.
- [25] Florian Boudin, Jian-Yun Nie, and Martin Dawes. Deriving a test collection for clinical information retrieval from systematic reviews. *Data and Text Mining in Biomedical Informatics (DTMBIO)*. 2010.

- [26] Florian Boudin, Marc El-Bèze, and Juan-Manuel Torres-Moreno. The LIA Update Summarization system at TAC-2008. *Text Analysis Conference (TAC)*. 2008.
- [27] Florian Boudin, Benoit Favre, Frederic Béchet, Marc El-Bèze, Laurent Gillard, and Juan-Manuel Torres-Moreno. The LIA-Thales summarization system at DUC-2007. *Document Understanding Conference (DUC)*. 2007.
- [28] B. Favre, F. Béchet, P. Bellot, F. Boudin, M. El-Beze, L. Gillard, G. Lapalme, and J.M. Torres-Moreno. The LIA-Thales summarization system at DUC-2006. *Document Understanding Conference (DUC)*. 2006.

#### National Conference Papers

- [29] Adrien Bougouin, Florian Boudin, and Béatrice Daille. Influence des domaines de spécialité dans l'extraction de termes-clés. *Traitement Automatique des Langues Naturelles (TALN)*. 2014.
- [30] 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.
- [31] Nicolas Hernandez and Florian Boudin. Construction d'un large corpus écrit libre annoté morpho-syntaxiquement en français. *Traitement Automatique des Langues Naturelles (TALN)*. 2013.
- [32] Romain Deveaud and Florian Boudin. Contextualisation automatique de Tweets à partir de Wikipédia. *Conférence en Recherche d'Information et Applications (CORIA)*. 2013.
- [33] Florian Boudin and Nicolas Hernandez. Détection et correction automatique d'erreurs d'annotation morpho-syntaxique du French TreeBank. *Traitement Automatique des Langues Naturelles (TALN)*. 2012.
- [34] Stéphane Huet, Florian Boudin, and 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.
- [35] Romain Deveaud, Florian Boudin, Eric SanJuan, and Patrice Bellot. Correction de césures et enrichissement de requêtes pour la recherche de livres. *Conférence en Recherche d'Information et Applications (CORIA)*. 2011.
- [36] Florian Boudin and 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.

#### Book Chapters

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

#### Ph.D. Thesis

- [38] Florian Boudin. Exploration d'approches statistiques pour le résumé automatique de texte. PhD thesis. Laboratoire Informatique d'Avignon – Université d'Avignon, 2008.

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

SERVICE	<b>Conference Program Chair:</b> RECITAL 2013 <b>Conference PCs:</b> LACL 2012, RECITAL (2010, 2014, 2015) <b>Conference Reviews:</b> AAAI 2015, TALN (2013, 2014, 2015), JAD'12 COLING 2012, KDIR (2009, 2010, 2011, 2012), CIKM 2011, RECITAL 2011 <b>Workshop Reviews:</b> BUCC 2015
---------	--

**Journal Reviews:** Traitement Automatique des Langues, Information Retrieval, Information Processing & Management, Journal of Cheminformatics, Language Resources and Evaluation

TEACHING	<b>Université de Nantes, France</b>	800h+
	Programming, Databases, Machine Learning	
	<b>Université d'Avignon et des Pays de Vaucluse, France</b>	96h
	Programming, Business Intelligence	
	<b>Université de Montréal, Canada</b>	54h
	Atelier en Informatique	
	<b>IUT d'Avignon, France</b>	100h
	Programming, Databases	