

ALBERTO BARRÓN CEDENO

*PhD student*

*Department of Information Systems and Computation  
Technical University of Valencia  
Camino de Vera s/n. Valencia (Spain)  
lbarron@dsic.upv.es*

---

## Education

---

- 2008- TECHNICAL UNIVERSITY OF VALENCIA, SPAIN  
*Ph.D. (Computer Science), current*
- 2008 TECHNICAL UNIVERSITY OF VALENCIA, SPAIN  
*M. Sc. (Artificial Intelligence, Pattern Recognition and Digital Image), December 2008*
- 2007 NATIONAL UNIVERSITY OF MEXICO, MEXICO  
*M.Sc. (Computer Science), June 2007*
- 2004 NATIONAL UNIVERSITY OF MEXICO, MEXICO  
*Bachelor (Computer Engineering), October 2004*

---

## Academic Positions

---

- 2005–2007 Professor (Computers and Programming), National University of Mexico
- 2003–2004 Professor Assistant, National University of Mexico

---

## Research Interests

---

Information Retrieval and Extraction; Natural Language Processing;  
Text Similarity Analysis; Plagiarism Detection; Terms Extraction

---

## Publications

---

### Journals

- 2010 Potthast, M., Barrón-Cedeño, A., Stein, B., Rosso, P. Cross-Language Plagiarism Detection. Language Resources and Evaluation, Special Issue on Plagiarism and Authorship Analysis (2010), pp. 1–18. Springer Netherlands.  
(impact factor: 0.634)
- 2009 Pinto D., Civera J., Barrón-Cedeño A., Juan A., and Rosso P. A statistical approach to crosslingual natural language tasks. Journal of Algorithms 64(1), pp. 51–60. Academic Press.  
(impact factor: 0.444)
- 2009 Barrón-Cedeño A., Rosso, P. On the Relevance of Search Space Reduction in Automatic Plagiarism Detection. Procesamiento del Lenguaje Natural, 43 (2009), pp. 141–149. SEPLN.

---

## International Conferences

- 2010 Potthast, M., Barrón-Cedeño, A., Eiselt, A., Stein, B., and Rosso, P. Overview of the 2nd International Competition on Plagiarism Detection. In: Notebook Papers of CLEF 2010 LABs and Workshops. Padua, Italy, 2010.

- 2010 Barrón-Cedeño, A., Rosso, P., Agirre, E., and Labaka G. Plagiarism Detection across Distant Language Pairs. In: Proceedings of the 23rd International Conference on Computational Linguistics (COLING 2010). Beijing, China, 2010. ACL
- 2010 Potthast, M., Stein, B., Barrón-Cedeño, A., Rosso, P. An Evaluation Framework for Plagiarism Detection. In: Proceedings of the 23rd International Conference on Computational Linguistics (COLING 2010). Beijing, China, 2010. ACL
- 2010 Barrón-Cedeño, A. On the Mono- and Cross-Language Detection of Text Reuse and Plagiarism. In: Proceedings of the 33rd Annual International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR 2010), p. 914. Geneva, Switzerland, 2010. ACM
- 2010 Barrón-Cedeño, A. and Rosso, P. Towards the 2nd International Competition on Plagiarism Detection and Beyond. In: Proceedings of the 4th International Plagiarism Conference. Newcastle upon Tyne, UK, 2010
- 2010 Barrón-Cedeño, A., Potthast, M., Rosso, P., Stein, B., and Eiselt, Andreas. Corpus and Evaluation Measures for Automatic Plagiarism Detection. In: Proceedings of the Seventh Conference on International Language Resources and Evaluation (LREC'10). Valletta, Malta, 2010. European Language Resources Association (ELRA)
- 2010 Sidorov, G., Barrón-Cedeño, A., and Rosso, P. English-Spanish Large Statistical Dictionary of Inflectional Forms. In: Proceedings of the Seventh Conference on International Language Resources and Evaluation (LREC'10). Valletta, Malta, 2010. European Language Resources Association (ELRA)
- 2010 Barrón-Cedeño, A., Basile, C., Degli Esposti, M., Rosso, P. Word Length  $n$ -grams for Text Re-Use Detection. In: Gelbukh A. (ed.) CICLing 2010, LNCS (6008), pp. 687–699, 2010. Springer-Verlag Berlin Heidelberg.
- 2009 Barrón-Cedeño A., Eiselt, A., and Rosso P. Monolingual Text Similarity Measures: A Comparison of Models over Wikipedia Articles Revisions. In: Natural Language Processing (ICON-2009), pp. 29-38. Hyderabad, India. Macmillan Publishers.
- 2009 Potthast, M., Stein, B., Eiselt, A., Barrón-Cedeño, A., and Rosso, P. Overview of the 1st International Competition on Plagiarism Detection. In: Proceedings of the SEPLN 2009 Workshop on Uncovering Plagiarism, Authorship and Social Software Misuse (PAN 09), pp. 1–9, CEUR-WS.org
- 2009 Barrón-Cedeño A. and Rosso P. On Automatic Plagiarism Detection based on  $n$ -grams Comparison. In: Boughanem et al. (eds.) ECIR 2009, LNCS (5478), pp. 696–700, 2009. Springer-Verlag Berlin Heidelberg.
- 2009 Barrón-Cedeño A., Rosso P., and Benedí, J.M. Reducing the Plagiarism Detection Search Space on the Basis of the Kullback-Leibler Distance. In: Gelbukh A. (ed.) CICLing 2009, LNCS (5449), pp. 523–534, 2009. Springer-Verlag Berlin Heidelberg.
- 2009 Barrón-Cedeño A., Sierra, G., Drouin, P., and Ananiadou, S. An Improved Automatic Term Recognition Method for Spanish. In: Gelbukh A. (ed.) CICLing 2009, LNCS (5449), pp. 125–136, 2009. Springer-Verlag Berlin Heidelberg.
- 2008 Pinto D., Civera J., Juan A, Rosso P., and Barrón-Cedeño A. A statistical approach to crosslingual natural language tasks. In: Fourth Latin American Workshop on Non-Monotonic Reasoning, Puebla, Mexico.

- 
- 2008 Barrón-Cedeño A., Rosso P., Pinto, D., and Juan A. On Cross-lingual Plagiarism Analysis using a Statistical Model. In: Proceedings of the ECAI'08 PAN Workshop Uncovering Plagiarism, Authorship and Social Software Misuse, pp. 9-13, Patras, Greece.
- 2008 Barrón-Cedeño, A. and Rosso, P. Towards the Exploitation of Statistical Language Models for Plagiarism Detection with Reference. In: Proceedings of the ECAI'08 PAN Workshop Uncovering Plagiarism, Authorship and Social Software Misuse, pp. 15-19, Patras, Greece
- 2008 Barrón-Cedeño, A., Sierra, G., and Kemper N. Can TF-IDF and fuzzy Logic Improve Onomasiological Inference Ranking? Or Keywords Frequency is Good Enough? In: Advances on Applied Computer and Applied Computational Science. Proceedings of the 7th WSEAS International Conference of Applied Computer and Applied Computational Science. Hangzhou, China.
- 2006 Barrón A., Sierra G., and Villaseñor E. C-value aplicado a la extracción de términos multi-palabra en documentos técnicos y científicos en español. In: 7th Mexican International Conference on Computer Science (ENC 2006), IEEE Computer Press. San Luis Potosi, Mexico (in Spanish).
- 2006 Sierra G., Alarcón R., Aguilar C., and Barrón A. Towards the Elaboration of a Corpus on Definitional Contexts. In: Proceedings of the 12th EURALEX International Congress.
- 2006 Sierra G., Alarcón R., Aguilar C., Barrón A., Benítez V., Baca I. Corpus de contextos definitorios: una herramienta para la lexicografía y la terminología. In: X Simposio Iberoamericano de Terminología (RITERM 2006). Montevideo, Uruguay.

### **Book Chapters**

- 2010 Barrón-Cedeño, A., Vila, M., and Rosso P. Detección automática de plagio: De la copia exacta a la paráfrasis. (Automatic Plagiarism Detection: from exact copy to paraphrase). In: Panorama actual de la lingüística forense en el ámbito legal y policial: Teoría y práctica. Madrid. Euphonia Ediciones SL. 2010 (in Spanish)

### **Honors and Awards**

---

- 2009 Winner of the MAVIR prize for the best M.Sc. degree thesis on Language Technologies and Scientific Communication (Spain)
- 2007– National Council of Science and Technology (Mexico) scholarship (Ph.D. student)
- 2007 M.Sc degree, cum laude (National University of Mexico)
- 2004–2006 National Council of Science and Technology (Mexico) scholarship (M.Sc. student)
- 2004–2006 Postgraduate Studies Direction (National University of Mexico) scholarship (M.Sc. student)
- 2004 Bachelor degree, cum laude (National University of Mexico)

## Research Internships

---

- 2010 Department of Information Studies / Information School, University of Sheffield, UK.  
Advisor: Paul Clough (Senior Lecturer)  
Main topic: analysis of text reuse and plagiarism  
Period: June–October
- 2009 Department of Mathematics, University of Bologna, Italy.  
Advisor: Mirko Degli Esposti (Department Head)  
Main topics: plagiarism detection  
Period: November.
- 2006-2007 Department of Linguistics and Translation, University of Montreal, Canada.  
Advisor: Patrick Drouin (Associate Professor)  
Main topics: automatic extraction of terms  
Period: September 2006–January 2007

## Program Committees

---

- 2011 34th ACM International Conference on Research and Development in Information Retrieval (SIGIR 2011)
- 2011 8th International Workshop on Text-based Information Retrieval (TIR 2011)
- 2010 CLEF Lab PAN. Uncovering Plagiarism, Authorship and Social Software Misuse
- 2010 7th International Workshop on Text-based Information Retrieval (TIR 2010)
- 2010 IADIS International Conference WWW/Internet
- 2009 Journal of Language Resources and Evaluation  
Special Issue on Plagiarism and Authorship Analysis
- 2009 SEPLN'09 Workshop PAN. Uncovering Plagiarism, Authorship and Social Software Misuse
- 2009 IADIS International Conference WWW/Internet
- 2008 IADIS International Conference WWW/Internet

## Organization of Scientific Events

---

- 2010 2nd International Competition on Plagiarism Detection (at CLEF)
- 2009 1st International Competition on Plagiarism Detection (at SEPLN)
- 2006 3rd Seminary on Language Engineering (Mexico)

## Major Invited Talks

---

- 2010 Tutorial: “Detection of Plagiarism and Text Reuse” (in conjunction with Paolo Rosso).  
8th International Conference on Natural Language Processing (ICON 2010), IIT Kharagpur,  
India, December 11th.
- 2010 “Detection of Plagiarism”. MIT Campus of Anna University, Chennai, India, December 6th.

- 2010 "Detection of Plagiarism and Text Reuse". Tsinghua University, Beijing, China. August 30th
- 2010 "The Quest of a Model for Cross-Language Plagiarism Detection". IR Talk, University of Sheffield, UK. July 16th.
- 2009 "Automatic Plagiarism Detection". Seminary on Physics and Mathematics. University of Bologna, Italy. November 30th.
- 2009 "Plagiarism Detection". University of Guanajuato, Mexico.
- 2007 "Métodos para la obtención automática de términos en un área de especialidad" (*Methods for Automatic Term Extraction*). In: Third Colloquium on Computational Linguistics, National University of Mexico, Mexico, Mexico.
- 2006 "Diccionario onomasiológico difuso" (*Fuzzy Onomasiological Dictionary*). 3rd Seminary on Language Engineering, National University of Mexico, Mexico.

### Relevant Computer Skills

---

Programming languages: Java, Python, C

Operative Systems: GNU Linux/UNIX (advanced); MS Windows

### Language skills

---

Spanish (native language)

English (high level)

French (basic knowledge)