

Juan Manuel Caicedo Carvajal

| | | |
|---------------------------------|--|--|
| CONTACT INFORMATION | Language Technologies Institute Carnegie Mellon University Pittsburgh, PA 15213. USA. | Website: http://cavorite.com E-mail: juan@cavorite.com jcaicedo@cs.cmu.edu |
| RESEARCH INTERESTS | Information Extraction: text mining, named entity recognition, vertical web crawling. Information Retrieval: vertical search engines. | |
| EDUCATION | Carnegie Mellon University , Pittsburgh, PA, USA <i>Master in Language Technologies</i> Expected graduation date: June 2013 August 2011 – present | |
| | Universidad de los Andes , Bogotá, Colombia <i>M.S. in Computer Science</i> Advisor: Professor José Abásolo January 2005 – March 2007 GPA: 4.6/5.0 | |
| | Institut National des Sciences Appliquées, INSA , Lyon, France <i>Exchange Student, Computer Science Department</i> August 2003 – June 2004 | |
| | Universidad Icesi , Cali, Colombia <i>B.S. Computer Systems Engineering</i> Graduated <i>Magna Cum Laude</i> August 1999 – February 2005 GPA: 4.51/5.0 | |
| PROFESSIONAL EXPERIENCE | CERN , Geneva, Switzerland <i>Fellow</i> January 2009 – September 2010 Member of the LHCb Online team, responsible for the administration and operations of the computing infrastructure used for the LHCb experiment. Responsible for the data storage system, including monitoring of the file system and profiling and benchmarking applications that use it intensively. | |
| | Casa Editorial El Tiempo , Bogotá, Colombia <i>Senior Software Engineer</i> February 2007 – November 2008 Member of the engineering team responsible for the development and operations of the online version of the largest newspaper in Colombia. Designed and implemented an information retrieval system for searching the news archive. Participated in the administration of the web infrastructure. The main responsibilities were load balancing and HTTP cache configuration, database replication configuration and monitoring, performance optimization and profiling, configuration management and software deployment. | |
| | Mazda, Compañía Colombia Automotriz , Bogotá, Colombia <i>Intern</i> January 2005 – December 2006. Designed and developed a web-based application for querying the data warehouse. Developed the tools to extract, transform and load data from the transactional systems to the data warehouse. | |
| | LIRIS Research Center , Lyon, France <i>Research Intern</i> February 2004 – June 2004 Developed a system for detecting changes on web documents and notifying users about those changes. | |
| REFEREED PUBLICATIONS | J. M. Caicedo Carvajal, J-C. Garnier, N. Neufeld, and R. Schwemmer, “A High-Performance Storage System for the LHCb Experiment”. <i>IEEE Transactions on Nuclear Science</i> , 2010. | |
| CONFERENCE AND JOURNAL ARTICLES | J.M Caicedo Carvajal, A. Gaviria, and J. Moreno, “Hechos y palabras: la realidad colombiana vista a través de la prensa escrita (Words and facts: an analysis of Colombian reality through the written news media)”. Documentos CEDE, Volume 46, Universidad de los Andes, 2011. J.M. Caicedo Carvajal, N. Neufeld, and R. Schwemmer, “Evaluation of Cluster File systems for the LHCb Experiment”. To appear in <i>Proceedings of the 18th International Conference on Computing in High Energy and Nuclear Physics (CHEP)</i> , Taipei, Taiwan, October 2010. | |

J.M. Caicedo Carvajal, N. Neufeld, and R. Schwemmer, "Optimising the HLT Farm at the LHCb Experiment". *Proceedings of the 17th IEEE NPSS Real Time Conference*, Lisbon, Portugal, May 2010.

HONOURS AND AWARDS

CEF Scholarship, 2003. Granted by the Embassy of France in Colombia to support a full academic year at a French university.

Andrés Bello Distinction, 1999. Granted to students with the top scores in the Colombian Integral High School Examinations.

COURSES

7th European Summer School in Information Retrieval (ESSIR 2009)

Padua, Italy

September 2009

This summer school's curriculum focuses on core areas of Information Retrieval, such as models, evaluation and machine learning.

32nd CERN School of Computing (CSC 2009)

Göttingen, Germany

August 2009

The CSC is a summer school on the computing technologies and tools used for physics computing.

LANGUAGES

Spanish: Native

English: Full professional proficiency

French: Professional working proficiency

SKILLS

Programming languages: Java, Python, Scala, JavaScript.

System administration: RHEL, CentOS and Ubuntu. Configuration of monitoring tools (Ganglia, Nagios and Munin) and configuration management systems (Quattor).

Web technologies: Configuration and administration of web servers, HTTP load balancers and caching systems. Web application development, particularly using the Django Application Framework.

Other technologies: Lucene, Solr, Hadoop, EC2, S3, R, MySQL, PostgreSQL, SVN, Mercurial.

REFEREES

Available on request.