Juan Manuel Caicedo Carvajal

Contact Information

Website: http://cavorite.com
E-mail: me@cavorite.com

Professional Experience

Airbnb, San Francisco, California, USA
Staff Software Engineer. Relevance Team **May 2019 – Present**
Member of the team responsible for the overall relevance and the user experience in the search flow. Major contributions:

- **Page Structure Optimization** Built the framework to test different layouts of search results. Ran online and offline experiments for evaluating their impact in conversion rates and user-experience metrics.
- **Internally Promoted Content** Built the system for ranking internally promoted sections on the search results pages, subject to constraints on the user experience and on top level metrics.
- **Infrastructure Optimization** Reduced latency of the online runtime for serving search results and query autocomplete suggestions.

Twitter, San Francisco, California, USA
Staff Software Engineer. Technical lead, Search Quality Team **June 2013 – May 2019**
Designed, developed and deployed multiple components of the Twitter search engine. Ran several offline and online experiments for improving the quality of the search results. Major contributions:

- **Tweet Search** Built the runtime component of the first stage ranking model. The system is part of the ranking pipelines of the search engine and the home timeline.
- **Query Understanding** Built components for query segmentation, annotation and rewriting. Built the pipeline for spell checking and query suggestions.
- **User Search** Built the pipeline for expanding the index, including structed information from multiple sources. Developed features for improving the relevance of user results.
- **Query Autocompletion** Built the pipeline of the ranking model for the autocomplete suggestions and developed features for personalizing the results.
- **Overall Relevance** Implemented an embeddings model based on search sessions. This model was used for generating related searches and personalization features. Built tools for running offline experiments and monitoring the quality of the search results.

Carnegie Mellon University, Pittsburgh, Pennsylvania, USA
Graduate Research Assistant **October 2012 – May 2013**
Implemented tools and conducted experiments related web crawling and adaptive information filtering.
Advisor: Professor James Callan.

LinkedIn, Mountain View, California, USA
Software Engineer Intern. Search Relevance Team **June 2012 – August 2012**
Contributed to the evaluation and improvement of a component for search query segmentation and classification. Extended the existing models, and developed tools to evaluate and monitor their results.

CERN, Geneva, Switzerland
Fellow. LHCb Online team **January 2009 – September 2010**
Contributed to the administration and operations of the computing infrastructure of the LHCb experiment. Managed for the data storage system, including monitoring of the file system and profiling and benchmarking of high throughput applications.

Casa Editorial El Tiempo Bogotá, Colombia
Senior Software Engineer. Web Applications Team **February 2007 – November 2008**
Built the system for the newspaper online search engine, including the pipeline for indexing the news archive. Participated in the administration of the web infrastructure, including load balancing, HTTP caching configuration, database replication, performance optimization and profiling, and configuration management.

Education

Carnegie Mellon University, Pittsburgh, PA, USA
Master in Language Technologies **August 2011 – May 2013**
Advisor: Prof. James Callan.

Universidad de los Andes, Bogotá, Colombia
M.S. in Computer Science **January 2005 – March 2007**
Advisor: Professor José Abásolo.

Institut National des Sciences Appliquées, INSA, Lyon, France
Exchange Student, Computer Science Department **August 2003 – June 2004**

Universidad Icesi, Cali, Colombia
B.S. Computer Systems Engineering **August 1999 – February 2005**
Graduated *Magna Cum Laude*.

Refereed Publications

J. M. Caicedo Carvajal, J-C. Garnier, N. Neufeld, and R. Schwemmer, "A High-Performance Storage System for the LHCb Experiment". *IEEE Transactions on Nuclear Science*, 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". *Proceedings of the 18th International Conference on Computing in High Energy and Nuclear Physics (CHEP)*, 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.

Honors

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 retrieval models, search results evaluation and machine learning applications.

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.

Skills

Programming languages: Java, Scala, Python, Javascript.

Information Retrieval and Data Processing: Lucene, Scalding, Hadoop, Heron, Pig, Pandas.

Last modified: June 2020.