

Carlos Alberto Alejandro CASTILLO Ocaranza

Diagonal 177
Barcelona 08018
Catalunya, Spain

<http://www.chato.cl/research/>
chato@chato.cl
+34 93 183 8842

Education UNIVERSITY OF CHILE 2000–2004
Ph.D. in Computer Science; advisor: Ricardo Baeza-Yates.

UNIVERSITY OF CHILE 1994–2000
Computing Engineer. This is a 6.5-years-degree.
Graduated *summa cum laude*.

Post-doctoral fellow at Universitat Pompeu Fabra in Barcelona, Catalunya-Spain with Vicente López (2005) and Università di Roma “La Sapienza” in Rome, Italy with Stefano Leonardi (2006).

Publications A full list is available at <http://www.chato.cl/research/>
The following is a selection of publications by topic.

WEB USAGE MINING

Paolo Boldi, Francesco Bonchi, Carlos Castillo, Sebastiano Vigna: “From ‘dango’ to ‘japanese cakes’: Query Reformulation Models and Patterns”. To appear in IEEE/ACM Web Intelligence. Milan, Italy, September 2009. IEEE CS Press. (18%).

Francesco Bonchi, Carlos Castillo, Debora Donato, Aristides Gionis: “Topical Query Decomposition”. ACM SIGKDD, pp. 609-618. Las Vegas, USA, 2008. ACM Press. (19% acceptance rate).

Paolo Boldi, Francesco Bonchi, Carlos Castillo, Debora Donato, Aristides Gionis, Sebastiano Vigna: “The Query-Flow Graph: Model and Applications”. ACM CIKM, pp. 52-60. Napa Valley, CA, 2008. ACM Press. (17%).

Barbara Poblete, Aristides Gionis, Carlos Castillo: “Dr. Searcher and Mr. Browser: a Unified Hyperlink-Click Graph”. ACM CIKM, pp. 1123-1132. Napa Valley, CA, USA, 2008. ACM Press. (17%).

ADVERSARIAL IR ON THE WEB

Luca Becchetti, Carlos Castillo, Debora Donato, Ricardo Baeza-Yates, Stefano Leonardi: “Link Analysis for Web Spam Detection”. ACM Transactions on the Web, Vol. 2, No. 1, Art. 2, 2008. ACM Press.

Carlos Castillo, Debora Donato, Aristides Gionis, Vanessa Murdock, Fabrizio Silvestri: “Know your Neighbors: Web Spam Detection using the

Web Topology”. ACM SIGIR, pp. 423-430. Amsterdam, Netherlands, 2007. ACM Press. (17%).

Josiane Xavier-Parreira, Carlos Castillo, Debora Donato, Sebastian Michel, Gerhard Weikum: “The JXP Method for Robust PageRank Approximation in a Peer-to-Peer Web Search Network”. The VLDB Journal, Vol. 17, No. 2, pp. 291-313, 2008. Springer.

WEB CHARACTERIZATION

Ricardo Baeza-Yates, Christian Middleton, Carlos Castillo: “The Geographical Life of Information”. To appear in IEEE/ACM Web Intelligence. Milan, Italy, September 2009. IEEE CS Press. (18%).

Ricardo Baeza-Yates, Carlos Castillo and Efthimis N. Efthimiadis: “Characterization of National Web Domains”. ACM Transactions on Internet Technology, Vol. 7, No. 2, Art. 9. May 2007. ACM Press.

Luciana Buriol, Carlos Castillo, Debora Donato, Stefano Leonardi and Stefano Millozzi: “Temporal Evolution of the Wikigraph”. ACM/WIC/IEEE Web Intelligence, pp. 45-51. Hong Kong, December 2006. IEEE CS Press. (18%).

Ricardo Baeza-Yates, Carlos Castillo and Felipe Saint-Jean: “Web Dynamics, Structure and Page Quality”. In M. Levene and A. Poulouvasilis (eds.): “Web Dynamics”, pp. 93–109, 2004. Springer.

WEB CRAWLING

Carlos Castillo, Alberto Nelli and Alessandro Panconesi: “A Memory-Efficient Strategy for Exploring the Web”. Web Intelligence Conference, pp. 680-686. Hong Kong, December 2006. IEEE CS Press. (18%).

Ricardo Baeza-Yates, Carlos Castillo, Mauricio Marin, Andrea Rodriguez: “Crawling a Country: Better Strategies than Breadth-First for Web Page Ordering”. WWW, pp. 864-872. Chiba, Japan, 2005. ACM Press. (20%).

Ricardo Baeza-Yates and Carlos Castillo: “Crawling the Infinite Web: Five Levels are Enough”. Workshop of Algorithms on Web Graphs (WAW), pp. 156-167. Rome, Italy, 2004. Springer. (45%).

RANKING ON THE WEB

Eugene Agichtein, Carlos Castillo, Debora Donato, Aristides Gionis, Gilad Mishne : “Finding High-Quality Content in Social Media”. Web Search and Data Mining (WSDM), pp. 183-194. Stanford, CA, USA, 2008. ACM Press. (16%).

Carlos Castillo, Debora Donato, Aristides Gionis: “Estimating the Number of Citations using Author Reputation”. String Processing and Information Retrieval Symposium (SPIRE), pp. 107-117. Santiago, Chile, 2007. Springer. (35%).

Ricardo Baeza-Yates, Paolo Boldi and Carlos Castillo: “Generalizing PageRank: Damping Functions for Link-Based Ranking Algorithms”. ACM SIGIR, pp. 308-315. Seattle, Washington, USA, August 2006. ACM Press. (19%).

OTHER TOPICS

Francesco Bonchi, Carlos Castillo, Debora Donato, Aristides Gionis: ”Taxonomy-driven lumping for sequence mining”. To appear in the Data Mining and Knowledge Discovery Journal, ECML/PKDD special issue, 2009. Springer.

Luca Becchetti, Paolo Boldi, Carlos Castillo, Aristides Gionis: “Efficient Semi-Streaming Algorithms for Local Triangle Counting in Massive Graphs”. ACM SIGKDD, pp. 16-24. Las Vegas, USA, 2008. ACM Press. (19%).

Patrizia Andronico, Marina Buzzi, Carlos Castillo and Barbara Leporini: “Improving Search Engine Interfaces for Blind Users: a Case Study”. Journal of Universal Access in the Information Society, special issue on Information Systems Accessibility. Vol. 5, No.1, pp. 23-40, June 2006. Springer.

Teaching

UNIVERSITAT POMPEU FABRA 2005,2007-2008
Lecturer of the introductory course on databases (2005, 2007, 2008) and of the course on information retrieval (2005), the latter with Ricardo Baeza-Yates. Universitat Pompeu Fabra, Barcelona, Catalunya-Spain.

UNIVERSITÀ LA SAPIENZA 2006
Teaching assistant in the course on web information retrieval at Università di Roma “La Sapienza”, Italy.

UNIVERSITY OF CHILE 1998–2004
Lecturer on the course on Web Applications (2003), Information Retrieval with Ricardo Baeza-Yates (2003–2004), Development of Web Applications (2001). Teaching assistant of Programming Languages, Analysis and Design of Algorithms, and Introduction to Computer Science (1998–2000).

Other research activities

Organizer of the Adversarial Information Retrieval Workshop (AIRWeb) in 2008 with Kumar Chellapilla and Dennis Fetterly, co-located with WWW 2008 in Beijing, China.

Organizer of the Adversarial Information Retrieval Workshop (AIRWeb) in 2007 with Brian Davison and Kumar Chellapilla, co-located with WWW 2007 in Banff, Canada.

Organizer of the Graph Labeling Workshop (GraphLab) in 2007 with Ludovic Denoyer and Brian Davison, co-located with ECML/PKDD 2007 in Warsaw, Poland.

PC member of ECML/PKDD 2008, WWW 2008, WSDM 2008&2009, SPIRE 2008, ECIR 2007&2008, CIKM 2007, INFOSCALE 2007&2008, LAWEB 2007, AIRWEB 2006.

Reviewer for SIGIR, WWW, SPIRE, VLDB, CIKM, PODS, ACM TWEB, ACM TOIT, ACM TOIS, JWE, SP&E, JASIST, IP&M.

Professional experience

YAHOO! RESEARCH 2006–present
Researcher at the Barcelona lab. Researched on Web spam detection, ranking on social media, and other topics.

NEWTENBERG LTD. 2001–2004 (CTO) 2001–present (partner)
Founder, partner and CTO of this company, aimed at the development of applications that support both communication and coordination of people's doings. Directed the development of an XML-based content management system that is currently the flagship product of the company; licensees include the World Bank, the ministries of education of Colombia and Chile, and several newspapers.

TECNONAUTICA S.A. 1995–1998
Worked as software developer in the development of applications for the Web, including a Web site generator, an application for generating charts and graphs for the Web, a system for on-line access counters, and an application for managing advertising campaigns on the Web.

Awards and grants

INSTITUTE OF ENGINEERS OF CHILE 2001
“Ismael Valdes” award, for a graduated engineer with achievements in leadership and a good student record at each of the nine main universities.

MILLENNIUM SCHOLARSHIP 2002–2004
From the Center of Web Research (CWR), funding for postgraduate studies for three years. The CWR is funded by the Millennium Program of the Ministry of Planning and Cooperation, Government of Chile.

OTHER GRANTS 1994–2004
Grant from the Computer Science Department, University of Chile, fund-

ing for postgraduate studies for 4 years (2000-2004). Grant from the Faculty of Physics and Mathematics, University of Chile (1994,2001).

OTHER 1993, 1995
Iberoamerican mathematic olympics for students, Bronze Medal in Brazil (1995). Best score: 100%, in physics at the national SAT exam (1993).

Languages Proficient in English and Spanish. Basic knowledge of Italian and Catalan.