

Carlos Alberto Alejandro CASTILLO Ocaranza

<http://www.chato.cl/research/>
chato@acm.org

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 researcher at Universitat Pompeu Fabra in Barcelona with Vicente López (2005) and Sapienza University of Rome with Stefano Leonardi (2006).

Books Carlos Castillo: “Big Crisis Data.” 2016 (Forthcoming), Cambridge University Press.

Wei Chen, Laks V.S. Lakshmanan, and Carlos Castillo: “Information and Influence Propagation in Social Networks”. Monograph in: Synthesis Lectures on Data Management. October 2013, Morgan and Claypool Publishers.

Carlos Castillo and Brian D. Davison: “Adversarial Web Search”. Monograph in Foundations and Trends in Information Retrieval, Vol. 4, No 5. 2011. Now Publishers.

Papers Selected publications by topic including acceptance rates in parenthesis. Full list available at <http://www.chato.cl/research/>

SOCIAL MEDIA – CRISES AND HUMANITARIAN COMPUTING

Muhammad Imran, Carlos Castillo, Fernando Diaz, Sarah Vieweg: “Processing Social Media Messages in Mass Emergency: A Survey.” In ACM Computing Surveys, Volume 47, Issue 4, 2015.

Alexandra Olteanu, Sarah Vieweg and Carlos Castillo: “What to Expect When the Unexpected Happens: Social Media Communications Across Crises.” Computer Supported Collaborative Work and Social Media (CSCW). 2015. ACM Press (28%).

Sarah Vieweg, Carlos Castillo and Muhammad Imran: “Integrating Social Media Communications into the Rapid Assessment of Sudden Onset Disasters.” International Conference on Social Informatics (SocInfo). November 2014. Springer (23%).

Hemant Purohit, Amit Sheth, Carlos Castillo, Patrick Meier, Fernando Diaz: “Emergency-Relief Coordination on Social Media: Automatically Matching Resource Requests and Offers”. *First Monday* 19 (1), January 2014.

Alexandra Olteanu, Carlos Castillo, Fernando Diaz and Sarah Vieweg: *CrisisLex: A Lexicon for Collecting and Filtering Microblogged Communications in Crises*. *Weblogs and Social Media (ICWSM)*. Ann Arbor, MI, USA. June 2014. AAAI Press (20%).

Muhammad Imran, Shady Elbassuoni, Carlos Castillo, Fernando Diaz and Patrick Meier: “Extracting Information Nuggets from Disaster-Related Messages in Social Media”. *International Conference on Information Systems for Crisis Response and Management (ISCRAM)*. Baden-Baden, Germany, 2013. *Best paper award*.

SOCIAL MEDIA – ONLINE NEWS

Janette Lehmann, Carlos Castillo, Mounia Lalmas, Ricardo Baeza-Yates: “Story-focused Reading in Online News and its Potential for User Engagement.” To appear in the *Journal of the Association for Information Science and Technology (JASIST)*.

Alexandra Olteanu, Carlos Castillo, Nicholas Diakopoulos, and Karl Aberer: “Comparing Events Coverage in Online News and Social Media: The Case of Climate Change.” *International Conference on Web and Social Media (ICWSM)*. 2015. AAAI Press (19%).

Carlos Castillo, Mohammed El-Haddad, Jürgen Pfeffer and Matt Stempeck: “Characterizing the Life Cycle of Online News Stories Using Social Media Reactions”. *Computer Supported Collaborative Work and Social Media (CSCW)*. Baltimore, USA, 2014. ACM Press (27%).

Janette Lehmann, Carlos Castillo, Mounia Lalmas and Ethan Zuckerman: “Transient News Crowds in Social Media”. *Weblogs and Social Media (ICWSM)*. Boston, USA, 2013. AAAI Press (20%).

SOCIAL MEDIA – NETWORKS AND INFLUENCE

Lilian Weng, Jacob Ratkiewicz, Nicola Perra, Bruno Gonçalves, Carlos Castillo, Francesco Bonchi, Rossano Schifanella, Filippo Menczer, Alessandro Flammini: “The Role of Information Diffusion in the Evolution of Social Networks”. *Knowledge Discovery and Data Mining (KDD)*, pp. 356-364. New York, USA. August 2013, ACM Press. (17%).

Dino Ienco, Francesco Bonchi, Carlos Castillo: “Meme Ranking to Maximize Posts Virality in Microblogging Platforms”. *Journal of Intelligent Information Systems*. April 2013, Springer.

Michael Mathioudakis, Francesco Bonchi, Carlos Castillo, Aristides Gionis, Antti Ukkonen: “Sparsification of Influence Networks”. Knowledge Discovery and Data Mining (KDD), pp. 529-537. San Diego, USA, 2011. ACM Press (8%).

Michalis Potamias, Francesco Bonchi, Carlos Castillo, Aristides Gionis: “Fast shortest path distance estimation in large networks”. ACM CIKM, pp. 867-876, 2009. ACM Press (15%). *Best student paper award*

Luca Becchetti, Paolo Boldi, Carlos Castillo, Aristides Gionis: “Efficient Algorithms for Large-Scale Local Triangle Counting”. In Transactions on Knowledge Discovery from Data, Vol. 4, No. 3, pp. 1-28. 2010. ACM Press.

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%).

SOCIAL MEDIA – COMMUNITIES AND COLLABORATION

Daniela Iosub, David Laniado, Carlos Castillo, Mayo Fuster Morell, Andreas Kaltenbrunner: “Emotions under Discussion: Gender, Status and Communication in Online Collaboration”. PLOS One, August 2014.

Aris Anagnostopoulos, Carlos Castillo, Aristides Gionis, Luca Becchetti, Stefano Leonardi: “Online Team Formation in Social Networks”. WWW, pp. 839-848. Lyon, France. 2012. ACM Press. (12%)

Aris Anagnostopoulos, Carlos Castillo, Aristides Gionis, Luca Becchetti, Stefano Leonardi: “Power in Unity: Forming Teams in Large-Scale Community Systems”. ACM CIKM, pp. 599-608. Toronto, Canada. 2010. ACM Press. (18%)

Paolo Boldi, Francesco Bonchi, Carlos Castillo, and Sebastiano Vigna: “Viscous Democracy for Social Networks”. In Communications of ACM, No 6, June 2011. ACM Press.

SOCIAL MEDIA – OTHER APPLICATIONS

Eduardo Ruiz, Vagelis Hristidis, Carlos Castillo and Aristides Gionis: “Measuring and Summarizing Movement in Microblog Postings”. Weblogs and Social Media (ICWSM). Boston, USA, 2013. AAAI Press (20%).

Eduardo Ruiz, Vagelis Hristidis, Carlos Castillo, Aristides Gionis and Alejandro Jaimes: “Correlating Financial Time Series with Micro-Blogging Data”. Web Search and Data Mining (WSDM), pp. 513-522. Seattle, Washington, 2012. ACM Press. (21%).

Francesco Bonchi, Carlos Castillo, Aristides Gionis and Alejandro Jaimes: “Social Network Analysis and Mining for Business Applications”. ACM Transactions on Intelligent Systems and Technology (TIST), Vol. 2 Issue 3, April 2011.

WEB QUALITY – CREDIBILITY ISSUES AND SPAM

Aditi Gupta, Ponnurangam Kumaraguru, Carlos Castillo and Patrick Meier: “TweetCred: A Real-time Web-based System for Assessing Credibility of Content on Twitter”. International Conference on Social Informatics (SocInfo). November 2014. Springer (23%).

Carlos Castillo, Marcelo Mendoza, Barbara Poblete: “Predicting Information Credibility in Time-Sensitive Social Media”. Internet Research, special issue on The Predictive Power of Social Media. 2013.

Carlos Castillo, Marcelo Mendoza, Barbara Poblete: “Information Credibility on Twitter”. World Wide Web Conference (WWW), pp. 675-684. Hyderabad, India. 2011. (12%)

Jacob Abernethy, Olivier Chapelle, Carlos Castillo: “Graph Regularization Methods for Web Spam Detection”. Machine Learning Journal, Vol. 81, No. 2, pp. 207-225, 2010. Springer.

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.

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.

WEB QUALITY – RANKING

Sihem Amer-Yahia, Francesco Bonchi, Carlos Castillo, Esteban Feuerstein, Isabel Mendez-Diaz and Paula Zabala: “Composite Retrieval of Diverse and Complementary Bundles”. IEEE TKDE, Feb. 2014.

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%).

USAGE MINING - SEARCH LOGS

Ilija Subasic, Carlos Castillo: “The Effects of Query Bursts on Web Search Results”. ACM/IEEE Web Intelligence 2010, pp. 374-381. Toronto, Canada. (17%). *Best student paper award*.

Ingmar Weber, Carlos Castillo: “Demographics of Web Search”. ACM SIGIR, pp. 523-530. Geneva, Switzerland, 2010. ACM Press. (17%)

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

Francesco Bonchi, Carlos Castillo, Debora Donato, Aristides Gionis: “Topical Query Decomposition”. Knowledge Discovery and Data Mining (KDD), pp. 609-618. Las Vegas, USA, 2008. ACM Press. (19%).

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%).

USAGE MINING – QUERY REFORMULATIONS

Aristides Anagnostopoulos, Luca Becchetti, Carlos Castillo and Aristides Gionis: “An Optimization Framework for Query Recommendation”. Web Search and Data Mining (WSDM), pp. 161-170, New York, USA. 2010. (16%)

Ilaria Bordino, Carlos Castillo, Debora Donato and Aristides Gionis: “Query Similarity by Projections on the Query-Flow Graph”. SIGIR 2010. ACM Press. (17%)

Paolo Boldi, Francesco Bonchi, Carlos Castillo, Sebastiano Vigna: “Query Reformulation Mining: Models, Patterns and Applications”. Information Retrieval, Springer. 2010.

Paolo Boldi, Francesco Bonchi, Carlos Castillo, Sebastiano Vigna: “From ‘dango’ to ‘japanese cakes’: Query Reformulation Models and Patterns”.

In IEEE/ACM Web Intelligence. Milan, Italy, September 2009. IEEE CS Press. (18%) *Best paper award*.

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%)

WEB CRAWLING AND CHARACTERIZATION

Carlos Castillo, Ricardo Baeza-Yates, Berthier Ribeiro-Neto: “Web Crawling”. Chapter 12 in Ricardo Baeza-Yates and Berthier Ribeiro-Neto, “Modern Information Retrieval, Second Edition”. 2011. Addison-Wesley Professional.

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, 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 Analysis of the Wikigraph”. ACM/WIC/IEEE Web Intelligence, pp. 45-51. Hong Kong, December 2006. IEEE CS Press. (18%)

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 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%).

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.

OTHER TOPICS

Sofiane Abbar, Habibur Rahman, Saravanan Thirumuruganathan, Carlos Castillo and Gautam Das: “Ranking Item Features by Mining Online User-Item Interactions”. International Conference on Data Engineering (ICDE). Chicago, USA, 2014. ACM Press. (20%)

Francesco Bonchi, Carlos Castillo, Debora Donato, Aristides Gionis: “Taxonomy driven lumping for sequence mining”. In the Data Mining and

Knowledge Discovery Journal, ECML/PKDD special issue, Vol. 9, No. 2, pp. 227-244. 2009. Springer.

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

SAPIENZA UNIVERSITY OF ROME 2015
Lecturer of the course on algorithms for data science, with Aris Anagnostopoulos. Sapienza University of Rome, Italy.

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.

SAPIENZA UNIVERSITY OF ROME 2006
Teaching assistant in the course on web information retrieval at Sapienza University of Rome, 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).

Service

Program co-chair of ACM Digital Health 2016. Program co-chair of Web Search and Data Mining (WSDM) 2014. Track co-chair of Crowdsourcing Systems and Social Media (in WWW) 2015.

Co-chair of the Workshop on Social Web for Disaster Management (SWDM) in 2015 (co-located with WWW). Co-founder and co-chair of the Workshop on Credibility and Quality Issues on the Web (WebQuality) from 2011 to 2014 (co-located with WWW). Co-chair of the Workshop on Adversarial Information Retrieval on the Web (AIRWeb) in 2008 and 2007 (co-located with WWW). Co-chair of the Predictive Analytics Workshop in 2014 and the Graph Labeling Workshop (GraphLab) in 2007 (co-located with ECML/PKDD).

Member of the editorial board for the Journal on Internet Research (2012-present). Associate editor of ACM Transactions on the Web (2015-present). PC and/or SPC of WWW, WSDM, SIGIR, KDD, CIKM, ICWSM, AAI, etc. Reviewer for TOIT, TOIS, JASIST, etc.

ACM Senior Member, IEEE Senior Member.

Professional experience

EURECAT 2016–present
Director of Research for Data Mining.

SAPIENZA UNIVERSITY OF ROME 2015
Visiting professor of the groups on algorithms. Research on algorithms for online team formation.

QATAR COMPUTING RESEARCH INSTITUTE 2012–2015
Principal Scientist with the Social Computing group. Research focused on the interaction between traditional news and social media, and applications of social media mining to support emergency response during humanitarian crises. Senior Scientist (2012–2015).

YAHOO! RESEARCH 2006–2012
Senior Research Scientist at the Barcelona lab. Researched on Web spam detection, ranking on social media, and other topics. Postdoctoral Researcher (2006–2008), Research Scientist (2008–2011), Senior Research Scientist (2011–2012).

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.

Grants

CENIT SOCIAL MEDIA 2010–2013
Participated in a coordination/co-leading role (until 2012) on a 25M Euro grant to study social media in Spain.

TORRES QUEVEDO GRANT 2006–2008, 2010–2012
“Torres Quevedo” grant from the Ministry of Science and Innovation, Spain.

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.

Languages

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