

PERSONAL INFORMATION **Carlos CASTILLO**

 <https://www.chato.cl/science>

SUMMARY Leads the Web Science and Social Computing research group at the Department of Engineering of Universitat Pompeu Fabra, and coordinates Horizon Europe project FINDHR (Grant Agreement 101070212). Researcher in algorithmic fairness and social computing: 120+ articles in top-tier journals and international conferences, receiving two test of time awards and five best paper awards.

CURRENT POSITION

2021– **ICREA Research Professor**

ICREA and Universitat Pompeu Fabra

Leads the Web Science and Social Computing (WSSC) group, which focuses on algorithmic fairness, responsible AI, and web data mining. The group and leads and participates in various EU-funded research projects.

PREVIOUS POSITIONS

2017–2021 **Distinguished Research Professor**

Universitat Pompeu Fabra.

Led the Web Science and Social Computing (WSSC); participated in two EU H2020 Projects.

2016–2017 **Director of Research for Data Science**

Eurecat Technology Centre of Catalonia.

Led a group of four scientists and two PhD students who participated in one FP7 and three H2020 projects from the EU. Researched on fairness in information retrieval and data mining, and on data anonymization methods enabling privacy-preserving data mining.

2012–2015 **Principal Scientist**

Qatar Computing Research Institute, Department of Social Computing.

Started in this position as a Senior Scientist. Created the crisis computing team, fully funded by Qatar Foundation. Recruited and managed two scientists, one postdoc, one engineer, and three research assistants; supervised six PhD students as interns. The researchers in the team obtained a best paper award at ISCRAM, a specialized conference on computing for crisis management. The team won the Grand Prize at the World Open Source Software Challenge with AIDR: Artificial Intelligence for Disaster Response.

2006–2012 **Senior Research Scientist**

Yahoo! Labs, Spain

Started as a postdoc partially funded by a Torres Quevedo grant, then as a Research Scientist fully funded by Yahoo. Worked on mining links, contents, and queries, obtaining a best paper award at Web Intelligence for work on query flow graphs and ten years later a test-of-time award for seminal work on social media quality. Co-supervised nine PhD students during internships, two of which obtained best student paper awards. Led the development of the Correlator semantic search demo, and of the query analysis platform Clues. First researcher to win the You Rock! Award at Yahoo, for work on mining real-time news content from television and web sites.

2005–2006 **Postdoctoral Fellow**

- 2006 Sapienza University of Rome, Italy
Worked on graph algorithms for web spam detection, which formed the basis for incorporating links into the assessment of content quality on the web.
- 2005 Universitat Pompeu Fabra, Spain
Led several of the first large-scale characterization studies of national web domains.

EDUCATION

- 2000-2004 **PhD in Computer Science**
University of Chile. PhD supervised by Dr. Ricardo Baeza-Yates.
- 1994-2000 **Computer Engineering**
University of Chile. Note: this was a 6-years degree.

BOOKS/MONOGRAPHS

- 2016 **Big Crisis Data**
Cambridge University Press. 224 pages. [394 citations](#)
This book introduces methods for dealing with social media messages during time-critical events, with a focus on natural and man-made disasters. The first part focuses on the technical aspects of data processing, and follows computing disciplines of databases, natural language processing, machine learning, network analysis, and online algorithms. The second part focuses on the context in which data processing occurs, and is oriented to studies of information sciences and human factors, including crowdsourcing, human-computer interaction, computer-supported collaborative work, and information visualization.
- 2013 **Information and Influence Propagation in Social Media (with W. Chen, L. Lakshmanan)**
Synthesis Lectures on Data Management, Morgan and Claypool. 178 pages. [557 citations](#)
This monograph studies information diffusion models, including their main properties and extensions, as well as selected problems on information propagation, particularly influence maximization. It includes important practical problems such as learning weight influence, sparsifying influence networks, and acquiring and publishing research datasets and software.
- 2011 **Adversarial Web Search (with B. Davison)**
Foundations and Trends in Information Retrieval, NOW Publishers. 109 pages. [135 citations](#)
This monograph surveys the adversarial relationship between search engines and content providers willing to manipulate them. We show how search engine spammers create contents and links to lure visitors, and how web search engines try to detect these deceptive practices.

PUBLICATIONS BY TOPIC

- Note** Citation counts in this list are included for publications with 100 citations or more, according to Google Scholar. When available, acceptance rates for conferences are noted in parenthesis.

Algorithmic Fairness – Risk Assessment

- 1 Manuel Portela, Carlos Castillo, Songül Tolan, Marzieh Karimi-Haghighi, Antonio Andres Pueyo: "A Comparative User Study of Human Predictions in Algorithm-Supported Recidivism Risk Assessment." *Journal of Artificial Intelligence and Law*. Springer, 2024.
- 2 Marzieh Karimi-Haghighi, Carlos Castillo, Songül Tolan, Kristian Lum: "Effect of Conditional Release on Violent and General Recidivism: A Causal Inference Study." *Journal of Experimental Criminology*. Springer, 2023.
- 3 Marzieh Karimi-Haghighi, Carlos Castillo: "Efficiency and Fairness in Recurring Data-Driven Risk Assessments of Violent Recidivism." In *Proc. of Symposium on Applied Computing (SAC) 2021*, pp. 994-1002. ACM Press (25%).

- 4 Marius Miron, Songül Tolan, Emilia Gomez, Carlos Castillo: "Evaluating causes of algorithmic bias in juvenile criminal recidivism." *Journal of Artificial Intelligence and Law*. Springer, 2020
- 5 Marius Miron, Songül Tolan, Emilia Gomez, Carlos Castillo: "Why Machine Learning May Lead to Unfairness: Evidence from Risk Assessment for Juvenile Justice in Catalonia." In *Proc. of International Conference on Artificial Intelligence and Law (ICAIL)* 2019, Montréal, Canada. ACM Press (18%). **Best paper award**. [144 citations](#).

Algorithmic Fairness – Ranking and Recommendation

- 6 Arpit Merchant, Carlos Castillo: "Disparity, Inequality, and Accuracy Tradeoffs in Graph Neural Networks for Node Classification." In *Proc. of ACM International Conference on Information and Knowledge Management (CIKM)* 2023. ACM Press (24%).
- 7 Marzieh Karimi-Haghighi, Carlos Castillo, Davinia Hernández-Leo: "A Causal Inference Study on the Effects of First Year Workload on the Dropout Rate of Undergraduates." In *Proc. of International Conference on Artificial Intelligence and Education (AIED)* 2022. Springer (20%).
- 8 Corinna Hertweck, Carlos Castillo, Michael Mathioudakis. "Designing Affirmative Action Policies under Uncertainty." *Journal of Learning Analytics*. Society for Learning Analytics Research, 2022.
- 9 Michael Mathioudakis, Carlos Castillo, Giorgio Barnabo, Sergio Celis: "Affirmative Action Policies for Top-k Candidates Selection, With an Application to the Design of Policies for University Admissions." In *Proc. of Symposium on Applied Computing (SAC)* 2020, Brno, Czech Republic. ACM Press (24%).
- 10 Francesco Fabbri, Marialuisa Croci, Francesco Bonchi, Carlos Castillo. "Exposure Inequality in People Recommender Systems: The Long-Term Effects." In *Proc. of AAAI International Conference on Web and Social Media (ICWSM)* 2022. AAAI Press (20%).
- 11 Francesco Fabbri, Yanhao Wang, Francesco Bonchi, Carlos Castillo and Michael Mathioudakis. "Rewiring What-to-Watch-Next Recommendations to Reduce Radicalization Pathways." In *Proc. of The Web Conference (WWW)* 2022. ACM Press (18%). **Best paper award**.
- 12 Francesco Fabbri, Francesco Bonchi, Ludovico Boratto, Carlos Castillo: "The Effect of Homophily on Disparate Visibility of Minorities in People Recommender Systems." In *Proc. of AAAI International Conference on Web and Social Media (ICWSM)* 2020. Springer (17%).
- 13 Meike Zehlike, Tom Sühr, Ricardo Baeza-Yates, Francesco Bonchi, Carlos Castillo, Sara Hajian: "Fair Top-k Ranking with Multiple Protected Groups." *Information Processing and Management* 59(1), Special Issue on Algorithmic Bias and Fairness in Search and Recommendation. Jan 2022.
- 14 Meike Zehlike, Francesco Bonchi, Carlos Castillo, Sara Hajian, Mohamed Megahed and Ricardo Baeza-Yates: "FA*IR: A Fair Top-k Ranking Algorithm." In *Proc. of ACM International Conference on Information and Knowledge Management (CIKM)* 2017, Singapore. ACM Press (20%). [583 citations](#).

Algorithmic Fairness – Other Topics

- 15 Ioannis Billionis, Ricardo C. Berrios, Luis Fernandez-Luque, Carlos Castillo: "Disparate Model Performance and Stability in Machine Learning Clinical Support for Diabetes and Heart Diseases." In *Proc. of American Medical Informatics Association Annual Symposium (AMIA)*, 2025.
- 16 Júlia Riera, David Solans, Marzieh Karimi-Haghighi, Carlos Castillo, Caterina Calsamiglia: "Gender Disparities in Child Custody Sentencing in Spain: a Data Driven Analysis." In *Proc. of International Conference on Artificial Intelligence and Law (ICAIL)* 2023.
- 17 Teodora Sandra Buda, Joao Guerreiro, Jesus Omana Iglesias, Carlos Castillo, Oliver Smith, Aleksandar Matic: "Foundations for Fairness in Digital Health Apps." *Frontiers in Digital Health*. 2022.
- 18 David Solans, Francesco Fabbri, Caterina Calsamiglia, Carlos Castillo, Francesco Bonchi: "Comparing equity and effectiveness of different algorithms in an application for the room rental market." In *Proc. of Conference on Artificial Intelligence, Ethics, and Society (AIES)* 2021, pp. 978-988. AAAI/ACM (37%).

- 19 David Solans, Christopher Tauchmann, Aideen Farrell, Karolin Kappler, Hans-Hendrik Huber, Carlos Castillo: "Learning to Classify Morals and Conventions: Artificial Intelligence in Terms of the Economics of Convention." In *Proc. of AAAI International Conference on Web and Social Media (ICWSM)* 2021. AAAI Press.
- 20 David Solans, Battista Biggio, Carlos Castillo: "Poisoning Attacks on Algorithmic Fairness." In *Proc. of European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases (ECML/PKDD)* 2020, pp. 162-177. AAAI Press (19%). [105 citations](#).
- 21 Gemma Galdon Clavell, Mariano Martín Zamorano, Carlos Castillo, Oliver Smith and Aleksandar Matic: "Auditing Algorithms: On Lessons Learned and the Risks of Data Minimization." In *Proc. of Conference on Artificial Intelligence, Ethics, and Society (AIES)* 2020. ACM Press (34%).
- 22 Alexandra Olteanu, Carlos Castillo, Fernando Diaz, Emre Kıcıman: "Social Data: Biases, Methodological Pitfalls, and Ethical Boundaries." *Frontiers in Big Data*, Volume 2, July 2019. [827 citations](#).

Discrimination, Sexism, and Hate Speech

- 23 Laura Casanova, Priscilla Álvarez-Cueva, Carlos Castillo: "Evolution Over 62 Years: An analysis of sexism in the lyrics of the most-listened-to songs in Spain." *Cogent Arts and Humanities*, Vol. 11, No 1. Taylor and Francis, 2024.
- 24 Francielle Marques, Davinia Hernandez-Leo, Carlos Castillo: "Measuring gender bias in student satisfaction in higher education: A cross-department study." *Cogent Education*, Vol. 11, Taylor and Francis, 2024.
- 25 Valerio Lorini, Javier Rando, Diego Saez-Trumper, Carlos Castillo: "Uneven Coverage of Natural Disasters in Wikipedia: the Case of Floods." In *Proc. of International Conference on Information Systems for Crisis Response and Management (ISCRAM)* 2020, pp. 688–703.
- 26 Ugur Kursuncu, Manas Gaur, Carlos Castillo, Amanuel Alambo, Krishnaprasad Thirunarayan, Valerie Shalin, Dilshod Achilov, I. Budak Arpinar, Amit Sheth: "Modeling Islamist Extremist Communications on Social Media using Contextual Dimensions: Religion, Ideology, and Hate." In *Proc. of ACM Conference on Computer-Supported Cooperative Work and Social Computing (CSCW)* November 2019, Austin, USA. ACM Press (31%).
- 27 Alexandra Olteanu, Carlos Castillo, Jeremy Boy and Kush Varshney: "The Effect of Extremist Violence on Hateful Speech Online." In *Proc. of AAAI International Conference on Web and Social Media (ICWSM)* 2018, Stanford, CA, USA. AAAI Press (16%). [211 citations](#).

Crisis Informatics

- 28 Hemant Purohit, Cody Buntain, Amanda Hughes, Steve Peterson, Valerio Lorini, Carlos Castillo: "Engage and Mobilize! Understanding Evolving Patterns of Social Media Usage in Emergency Management." In *Proc. of ACM Conference on Computer-Supported Cooperative Work and Social Computing (CSCW)* 2025. ACM Press.
- 29 Valerio Lorini, Paola Rufolo, Carlos Castillo: "Venice Was Flooding ... One Tweet at a Time." In *Proc. of ACM Conference on Computer-Supported Cooperative Work and Social Computing (CSCW)* 2022. ACM Press.
- 30 Fedor Vitiugin, Carlos Castillo: "Cross-Lingual Query-Based Summarization of Crisis-Related Social Media: An Abstractive Approach Using Transformers." In *Proc. of HyperText (HT)* 2022. ACM Press (36%). **Best newcomer paper award.**
- 31 Lan Li, Aisha Aldosery, Fedor Vitiugin, Naomi Nathan, David Novillo-Ortiz, Carlos Castillo, Patty Kostkova: "The Response of Governments and Public Health Agencies to COVID-19 Pandemics on Social Media: A Multi-country Analysis of Twitter Discourse." *Frontiers in Public Health*, Vol. 9, September 2021.
- 32 Pakhee Kumar, Ferda Ofli, Muhammad Imran, Carlos Castillo. "Detection of Disaster-Affected Cultural Heritage Sites from Social Media Images Using Deep Learning Techniques." *Journal on Computing and Cultural Heritage* 13(3):23, 2020. ACM Press.
- 33 Rahul Pandey, Carlos Castillo, and Hemant Purohit: "Ranking and Grouping Social Media Requests for Emergency Services Using Serviceability Model." *Social Network Analysis and Mining* 10 (22) 2020, Springer.

- 34 Yara Rizk, Hadi Samer Jomaa, Mariette Awad, Carlos Castillo: "A Computationally Efficient Multi-modal Classification Approach of Disaster-related Twitter Images." In *Proc. of Symposium on Applied Computing (SAC)* 2019, Limassol, Cyprus. ACM Press (24%).
- 35 Valerio Lorini, Carlos Castillo, Francesco Dottori, Milan Kalas, Domenico Nappo, Pater Salamoun: "Integrating Social Media into a Pan-European Flood Awareness System: A Multilingual Approach." In *Proc. of International Conference on Information Systems for Crisis Response and Management (ISCRAM)* 2019, Valencia, Spain. **Best paper award.**
- 36 Hemant Purohit, Carlos Castillo, Muhammad Imran and Rahul Pandey: "Social-EOC: Serviceability Model to Rank Social Media Requests for Emergency Operation Centers." In *Proc. of Advances in Social Network Analysis and Mining (ASONAM)* August 2018, Barcelona, Spain. IEEE/ACM (15%).
- 37 Muhammad Imran, Prasenjit Mitra, Carlos Castillo: "Twitter as a Lifeline: Human-annotated Twitter Corpora for NLP of Crisis-related Messages." In *Proc. of International Conference on Language Resources and Evaluation (LREC)* May 2016, Portorož, Slovenia, pp. 1638-1643. [440 citations.](#)
- 38 Muhammad Imran, Carlos Castillo, Fernando Diaz, Sarah Vieweg: "Processing Social Media Messages in Mass Emergency: A Survey." *ACM Computing Surveys*, Vol. 47, No. 4, 2015. [1,021 citations.](#)
- 39 Alexandra Olteanu, Sarah Vieweg and Carlos Castillo: "What to Expect When the Unexpected Happens: Social Media Communications Across Crises." In *Proc. of ACM Conference on Computer-Supported Cooperative Work and Social Computing (CSCW)* March 2015, Vancouver, Canada, pp. 994-1009. ACM Press (28%). [618 citations.](#)
- 40 Anubhav Jain, Julius Adebayo, Eduardo de Leon, Weihua Li, Lalana Kagal, Patrick Meier, Carlos Castillo: "Mobile Application Development for Crisis Data." *Procedia Engineering*, Vol. 107, pp. 255-262 (HumTech 2015, 12-14 May in Boston, USA).
- 41 Sarah Vieweg, Carlos Castillo and Muhammad Imran: "Integrating Social Media Communications into the Rapid Assessment of Sudden Onset Disasters." In *Proc. of International Conference on Social Informatics (SocInfo)* 2014, Barcelona, Spain, pp. 444-461. Springer (23%). [143 citations.](#)
- 42 Hemant Purohit, Amit Sheth, Carlos Castillo, Patrick Meier, Fernando Diaz: "Emergency-Relief Coordination on Social Media: Automatically Matching Resource Requests and Offers." *First Monday* Vol. 19 No. 1, January 2014. [167 citations.](#)
- 43 Alexandra Olteanu, Carlos Castillo, Fernando Diaz and Sarah Vieweg: "CrisisLex: A Lexicon for Collecting and Filtering Microblogged Communications in Crises." In *Proc. of AAAI International Conference on Web and Social Media (ICWSM)* 2014, Ann Arbor, MI, USA. AAAI Press (20%). [503 citations.](#)
- 44 Muhammad Imran, Shady Elbassuoni, Carlos Castillo, Fernando Diaz and Patrick Meier: "Extracting Information Nuggets from Disaster-Related Messages in Social Media." In *Proc. of International Conference on Information Systems for Crisis Response and Management (ISCRAM)* 2013, Baden-Baden, Germany. [566 citations.](#) **Best paper award.**

Crisis Informatics – Hybrid Intelligence

- 45 Rahul Pandey, Hemant Purohit, Carlos Castillo, Valerie L. Shalin. "Modeling and Mitigating Human Annotation Errors to Design Efficient Stream Processing Systems with Human-in-the-Loop Machine Learning." *International Journal of Human-Computer Studies* 160. April 2022. Elsevier.
- 46 Hemant Purohit, Carlos Castillo, Muhammad Imran and Rahul Pandey: "Ranking of Social Media Alerts with Workload Bounds in Emergency Operation Centers." In *Proc. of IEEE/WIC/ACM International Conference on Web Intelligence (WI)* December 2018, Santiago, Chile. IEEE/ACM.
- 47 Ferda Ofli, Patrick Meier, Muhammad Imran, Carlos Castillo, Devis Tuia, Nicolas Rey, Julien Briant, Pauline Millet, Friedrich Reinhard, Matthew Parkan, Stéphane Joost: "Combining Human Computing and Machine Learning to Make Sense of Big (Aerial) Data for Disaster Response." *Big Data* Vol. 4 No. 1, pp. 47-59. March 2016. [208 citations.](#)

- 48 Muhammad Imran, Carlos Castillo, Ji Lucas, Patrick Meier, Sarah Vieweg: "AIDR: Artificial Intelligence for Disaster Response." In *Proc. of International Conference on World Wide Web (WWW) Demonstrators Track 2014*, Seoul, Korea, ACM Press [695 citations](#)

News and Time-Sensitive Information

- 49 Julia Proskurnia, Ruslan Mavlyutov, Carlos Castillo, Karl Aberer and Philippe Cudre-Mauroux: "Efficient Document Filtering Using Vector Space Topic Expansion and Pattern-Mining: The Case of Event Detection in Microposts." In *Proc. of ACM International Conference on Information and Knowledge Management (CIKM) 2017*, Singapore. ACM Press (21%).
- 50 Janette Lehmann, Carlos Castillo, Mounia Lalmas, Ricardo Baeza-Yates: "Story-focused Reading in Online News and its Potential for User Engagement." *Journal of the Association for Information Science and Technology (JASIST)* Vol. 68 No. 4, pp. 869-883. 2016.
- 51 Alexandra Olteanu, Carlos Castillo, Nicholas Diakopoulos, and Karl Aberer: "Comparing Events Coverage in Online News and Social Media: The Case of Climate Change." In *Proc. of AAAI International Conference on Web and Social Media (ICWSM) 2015*, Oxford, UK, pp. 288-297. AAAI Press (19%). [122 citations](#).
- 52 Carlos Castillo, Mohammed El-Haddad, Jürgen Pfeffer and Matt Stempeck: "Characterizing the Life Cycle of Online News Stories Using Social Media Reactions." In *Proc. of ACM Conference on Computer-Supported Cooperative Work and Social Computing (CSCW) 2014*, Baltimore, USA. ACM Press (27%). [274 citations](#).
- 53 Janette Lehmann, Carlos Castillo, Mounia Lalmas and Ethan Zuckerman: "Transient News Crowds in Social Media." In *Proc. of AAAI International Conference on Web and Social Media (ICWSM) 2013*, Boston, USA. AAAI Press (20%).
- 54 Eduardo Ruiz, Vagelis Hristidis, Carlos Castillo and Aristides Gionis: "Measuring and Summarizing Movement in Microblog Postings." In *Proc. of AAAI International Conference on Web and Social Media (ICWSM) 2013*, Boston, USA. AAAI Press (20%).
- 55 Eduardo Ruiz, Vagelis Hristidis, Carlos Castillo, Aristides Gionis and Alejandro Jaimes: "Correlating Financial Time Series with Micro-Blogging Data." In *Proc. of ACM International Conference on Web Search and Data Mining (WSDM) 2012*, Seattle, Washington, pp. 513-522. ACM Press. (21%). [418 citations](#).

Misinformation, Credibility, Web Quality, and Spam

- 56 Giorgio Barnabo, Federico Siciliano, Carlos Castillo, Stefano Leonardi, Preslav Nakov, Giovanni Da San Martino, Fabrizio Silvestri: "Deep Active Learning For Misinformation Detection Using Geometric Deep Learning." *Online Social Networks and Media* 33, January 2023.
- 57 Giorgio Barnabò, Federico Siciliano, Carlos Castillo, Stefano Leonardi, Preslav Nakov, Giovanni Da San Martino, and Fabrizio Silvestri: "FbMultiLing-Misinfo: Challenging Large-Scale Multilingual Benchmark for Misinformation Detection." In *Proc. of IEEE International Joint Conference on Neural Networks (IJCNN) 2022*. IEEE.
- 58 Maurício Gruppi, Panayiotis Smeros, Sibel Adalı, Carlos Castillo, Karl Aberer: "SciLander: Mapping the Scientific News Landscape." In *Proc. of AAAI International Conference on Web and Social Media (ICWSM)*. AAAI.
- 59 Panayiotis Smeros, Carlos Castillo, Karl Aberer: "SciClops: Detecting and Contextualizing Scientific Claims for Assisting Manual Fact-Checking." In *Proc. of ACM International Conference on Information and Knowledge Management (CIKM) 2021*, pp. 1692-1702. ACM Press.
- 60 Panayiotis Smeros, Carlos Castillo, Karl Aberer: SciLens: Evaluating the Quality of Scientific News Articles Using Social Media and Scientific Literature Indicators. In *Proc. of International Conference on World Wide Web (WWW) 2019*, San Francisco, USA. ACM Press (18%).
- 61 Aditi Gupta, Ponnurangam Kumaraguru, Carlos Castillo and Patrick Meier: "TweetCred: A Real-time Web-based System for Assessing Credibility of Content on Twitter." In *Proc. of International Conference on Social Informatics (SoInfo)* November 2014. Springer (23%). [510 citations](#).
- 62 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. [489 citations](#).

- 63 Carlos Castillo, Marcelo Mendoza, Barbara Poblete: "Information Credibility on Twitter." In *Proc. of International Conference on World Wide Web (WWW)* 2011, Hyderabad, India, pp. 675-684. (12%). [3,270 citations](#). **Test of Time Award (10 years) in 2021**
- 64 Marcelo Mendoza, Barbara Poblete, Carlos Castillo: "Twitter Under Crisis: Can we trust what we RT?." In *SOMA 2010: KDD Workshop on Social Media Analytics*, Washington, DC. July 2010. [1,517 citations](#).
- 65 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.
- 66 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.
- 67 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. [175 citations](#).
- 68 Carlos Castillo, Debora Donato, Aristides Gionis, Vanessa Murdock, Fabrizio Silvestri: "Know your Neighbors: Web Spam Detection using the Web Topology". In *Proc. of ACM Conference on Research and Development in Information Retrieval (SIGIR)* 2007, Amsterdam, Netherlands, pp. 423-430. ACM Press (17%). [488 citations](#).
- 69 Carlos Castillo, Debora Donato, Luca Becchetti, Paolo Boldi, Massimo Santini, Sebastiano Vigna: "A Reference Collection for Web Spam." *SIGIR Forum*, Vol. 40, No. 2, December 2006. [311 citations](#).
- 70 Luca Becchetti, Carlos Castillo, Debora Donato, Stefano Leonardi, Ricardo Baeza-Yates: "Using Rank Propagation and Probabilistic Counting for Link-Based Spam Detection." Workshop on Web Mining and Web Usage Analysis (WebKDD) August 2006, Philadelphia, USA. ACM Press. [142 citations](#).

Networks and Influence

- 71 Julia Proskurnia, Przemyslaw A. Grabowicz, Ryota Kobayashi, Carlos Castillo, Philippe Cudré-Mauroux, and Karl Aberer: "Predicting the Success of Online Petitions Leveraging Multidimensional Time-Series." In *Proc. of International Conference on World Wide Web (WWW)* 2017, Perth, Australia. ACM Press (17%).
- 72 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." In *Proc. of ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD)* 2013, New York, USA, pp. 356-364. ACM Press (17%). [216 citations](#).
- 73 Dino Ienco, Francesco Bonchi, Carlos Castillo: "Meme Ranking to Maximize Posts Virality in Microblogging Platforms." *Journal of Intelligent Information Systems*, Vol. 40 No. 2, pp. 211-239. April 2013, Springer.
- 74 Michael Mathioudakis, Francesco Bonchi, Carlos Castillo, Aristides Gionis, Antti Ukkonen: "Sparsification of Influence Networks." In *Proc. of ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD)* 2011, San Diego, USA, pp. 529-537. ACM Press (8%). [219 citations](#).
- 75 Michalis Potamias, Francesco Bonchi, Carlos Castillo, Aristides Gionis: "Fast Shortest Path Distance Estimation in Large Networks." In *Proc. of ACM International Conference on Information and Knowledge Management (CIKM)* 2009, pp. 867-876. ACM Press (15%). [404 citations](#). **Best student paper award.**
- 76 Luca Becchetti, Paolo Boldi, Carlos Castillo, Aristides Gionis: "Efficient Algorithms for Large-Scale Local Triangle Counting." *Transactions on Knowledge Discovery from Data*, Vol. 4, No. 3, pp. 1-28. 2010. ACM Press. [106 citations](#).
- 77 Luca Becchetti, Paolo Boldi, Carlos Castillo, Aristides Gionis: "Efficient Semi-Streaming Algorithms for Local Triangle Counting in Massive Graphs". In *Proc. of ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD)* 2008, Las Vegas, USA, pp. 16-24. ACM Press. (19%). [416 citations](#).

Communities and Collaboration

- 78 Virginia Morini, Maria Sansoni, Giulio Rossetti, Dino Pedreschi and Carlos Castillo: "Participant Behavior and Community Response in Online Mental Health Communities: Insights from Reddit." *Computers in Human Behavior*, Elsevier, 2025.
- 79 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. [106 citations](#).
- 80 Aris Anagnostopoulos, Carlos Castillo, Adriano Fazzone, Stefano Leonardi, and Evimaria Terzi: "Algorithms for Hiring and Outsourcing in the Online Labor Market." In *Proc. of ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD) 2018*, London, UK. ACM Press. (11%)
- 81 Aris Anagnostopoulos, Carlos Castillo, Aristides Gionis, Luca Becchetti, Stefano Leonardi: "Online Team Formation in Social Networks." In *Proc. of International Conference on World Wide Web (WWW) 2012*, Lyon, France, pp. 839-848. ACM Press. (12%). [415 citations](#).
- 82 Aris Anagnostopoulos, Carlos Castillo, Aristides Gionis, Luca Becchetti, Stefano Leonardi: "Power in Unity: Forming Teams in Large-Scale Community Systems." In *Proc. of ACM International Conference on Information and Knowledge Management (CIKM) 2010*, Toronto, Canada, pp. 599-608. 2010. ACM Press. (18%) [171 citations](#).
- 83 Paolo Boldi, Francesco Bonchi, Carlos Castillo, and Sebastiano Vigna: "Viscous Democracy for Social Networks." *Communications of ACM*, No. 6, June 2011. ACM Press. [105 citations](#).

Mining Search Logs and Query Reformulations

- 84 Ilija Subasic, Carlos Castillo: "Investigating query bursts in a web search engine." *Web Intelligence and Agent Systems*, Vol. 11, pp. 107-124. 2013. IOS Press.
- 85 Paolo Boldi, Francesco Bonchi, Carlos Castillo, Sebastiano Vigna: "Query Reformulation Mining: Models, Patterns and Applications." *Information Retrieval* Vol. 13 No. 13, pp. 257-298. Springer. 2011.
- 86 Ilija Subasic, Carlos Castillo: "The Effects of Query Bursts on Web Search Results." In *Proc. of IEEE/WIC/ACM International Conference on Web Intelligence (WI) 2010*, Toronto, Canada, pp. 374-381 (17%). **Best student paper award**.
- 87 Aristides Anagnostopoulos, Luca Becchetti, Carlos Castillo and Aristides Gionis: "An Optimization Framework for Query Recommendation." In *Proc. of ACM International Conference on Web Search and Data Mining (WSDM) 2010*, New York, USA, pp. 161-170. (16%) [109 citations](#)
- 88 Ilaria Bordino, Carlos Castillo, Debora Donato and Aristides Gionis: "Query Similarity by Projections on the Query-Flow Graph." In *Proc. of ACM Conference on Research and Development in Information Retrieval (SIGIR) 2010*, Geneva, Switzerland, pp. 515-522. ACM Press (17%).
- 89 Ingmar Weber, Carlos Castillo: "The Demographics of Web Search." In *Proc. of ACM Conference on Research and Development in Information Retrieval (SIGIR) 2010*, Geneva, Switzerland, pp. 523-530. ACM Press (17%). [209 citations](#).
- 90 Paolo Boldi, Francesco Bonchi, Carlos Castillo, Sebastiano Vigna: "From 'dango' to 'japanese cakes': Query Reformulation Models and Patterns." In *Proc. of IEEE/WIC/ACM International Conference on Web Intelligence (WI) September 2009*, Milan, Italy. IEEE CS Press (18%). **Best paper award**. [107 citations](#).
- 91 Paolo Boldi, Francesco Bonchi, Carlos Castillo, Debora Donato, Sebastiano Vigna: "Query Suggestions Using Query-Flow Graphs." *Workshop on Web Search Click Data (WSCD)*, pp. 56-63, 2009. [204 citations](#).
- 92 Ricardo Baeza-Yates, Christian Middleton, Carlos Castillo: "The Geographical Life of Information." In *Proc. of IEEE/WIC/ACM International Conference on Web Intelligence (WI) September 2009*, Milan, Italy. IEEE CS Press (18%).
- 93 Paolo Boldi, Francesco Bonchi, Carlos Castillo, Debora Donato, Aristides Gionis, Sebastiano Vigna: "The Query-Flow Graph: Model and Applications". In *Proc. of ACM International Conference on Information and Knowledge Management (CIKM) 2008*. Napa Valley, CA, pp. 52-60. ACM Press (17%). [449 citations](#).

- 94 Francesco Bonchi, Carlos Castillo, Debora Donato, Aristides Gionis: "Topical Query Decomposition". In *Proc. of ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD)* 2008, Las Vegas, USA, pp. 609-618. ACM Press (19%).
- 95 Barbara Poblete, Aristides Gionis, Carlos Castillo: "Dr. Searcher and Mr. Browser: a Unified Hyperlink-Click Graph". In *Proc. of ACM International Conference on Information and Knowledge Management (CIKM)* 2008, Napa Valley, CA, USA, pp. 1123-1132. ACM Press (17%).
- 96 Ricardo Baeza-Yates, Carlos Castillo, Flavio Junqueira, Vassilis Plachouras, Fabrizio Silvestri: "Challenges on Distributed Information Retrieval" (Invited Paper). In *Proc. of IEEE International Conference on Data Engineering (ICDE)* April 2007, Istanbul, Turkey. IEEE CS Press. [149 citations](#).

Ranking on the Web

- 97 Sofiane Abbar, Habibur Rahman, Saravanan Thirumuruganathan, Carlos Castillo and Gautam Das: "Ranking Item Features by Mining Online User-Item Interactions." In *Proc. of IEEE International Conference on Data Engineering (ICDE)* 2014, Chicago, USA, pp. 460-471. ACM Press (20%).
- 98 Sihem Amer-Yahia, Francesco Bonchi, Carlos Castillo, Esteban Feuerstein, Isabel Mendez-Diaz and Paula Zabala: "Composite Retrieval of Diverse and Complementary Bundles." *IEEE Transactions on Knowledge and Data Engineering*, Feb. 2014.
- 99 Eugene Agichtein, Carlos Castillo, Debora Donato, Aristides Gionis, Gilad Mishne : "Finding High-Quality Content in Social Media". In *Proc. of ACM International Conference on Web Search and Data Mining (WSDM)* 2008, Stanford, USA, pp. 183-194. ACM Press (16%). [2,162 citations](#) **Test of Time Award (10 years) in 2018**
- 100 Carlos Castillo, Debora Donato, Aristides Gionis: "Estimating the Number of Citations using Author Reputation". In *Proc. of Symposium on String Processing and Information Retrieval (SPIRE)* 2007, Santiago, Chile, pp. 107-117. Springer (35%). [118 citations](#).
- 101 Ricardo Baeza-Yates, Paolo Boldi, Carlos Castillo: "Generic Damping Functions for Propagating Importance in Link-Based Ranking." *Journal of Internet Mathematics*, Vol. 3, No. 4, pp. 445-478, 2006. A K Peters.
- 102 Ricardo Baeza-Yates, Paolo Boldi and Carlos Castillo: "Generalizing PageRank: Damping Functions for Link-Based Ranking Algorithms". In *Proc. of ACM Conference on Research and Development in Information Retrieval (SIGIR)* August 2006, Seattle, Washington, USA, pp. 308-315. ACM Press (19%). [144 citations](#).

Web Crawling and Characterization

- 103 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.
- 104 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. [179 citations](#).
- 105 Gabriel H. Tolosa, Fernando R. A. Bordignon, Ricardo Baeza-Yates, Carlos Castillo: "Distinctive Features of the Argentinian Web." In *Proc. of Latin American Web Conference (LA-WEB)* 2007, Santiago, Chile. IEEE CS Press.
- 106 Luciana Buriol, Carlos Castillo, Debora Donato, Stefano Leonardi and Stefano Millozzi: "Temporal Analysis of the Wikigraph". In *Proc. of IEEE/WIC/ACM International Conference on Web Intelligence (WI)* 2006, Hong Kong, pp. 45-51. IEEE CS Press (18%). [218 citations](#).
- 107 Carlos Castillo, Alberto Nelli and Alessandro Panconesi: "A Memory-Efficient Strategy for Exploring the Web". In *Proc. of IEEE/WIC/ACM International Conference on Web Intelligence (WI)* December 2006, Hong Kong, pp. 680-686. IEEE CS Press (18%).
- 108 Ricardo Baeza-Yates, Carlos Castillo, Mauricio Marin, Andrea Rodriguez: "Crawling a Country: Better Strategies than Breadth-First for Web Page Ordering". In *Proc. of International Conference on World Wide Web (WWW)* 2005, Chiba, Japan, pp. 864-872. ACM Press (20%). [192 citations](#).

- 109 Carlos Castillo: "Effective Web Crawling" (Abstract from PhD Thesis). *ACM SIGIR Forum*, Vol. 39 No. 1, pp. 55-56. June 2005. [373 citations](#).
- 110 Ricardo Baeza-Yates, Carlos Castillo and Vicente López: "Characteristics of the Web of Spain." *Cybermetrics*, Vol. 9, No. 1, pp. 1-41. 2005. [103 citations](#).
- 111 Ricardo Baeza-Yates and Carlos Castillo: "Crawling the Infinite Web: Five Levels are Enough". In *Proc. of Workshop of Algorithms on Web Graphs (WAW) 2004*, Rome, Italy, pp. 156-167. Springer (45%).
- 112 Ricardo Baeza-Yates, Carlos Castillo and Felipe Saint-Jean: "Web Dynamics, Structure and Page Quality". In M. Levene and A. Poullovassilis (eds.): "Web Dynamics", pp. 93–109, 2004. Springer.
- 113 A. Jaimes, J. Ruiz-del-Solar, R. Verschae, R. Baeza-Yates, C. Castillo, D. Yaksic and E. Davis: "On the Image Content of a Web Segment: Chile as a Case Study." *Journal of Web Engineering*, Vol. 3 No. 2, pp. 153-168. 2004.
- 114 A. Jaimes, J. Ruiz-del-Solar, R. Verschae, D. Yaksic, R. Baeza-Yates, E. Davis and C. Castillo: "On the Image Content of the Chilean Web." In *Proc. of Latin American Web Conference (LA-WEB) 2003*, Santiago, Chile, pp. 72-83. IEEE Cs. Press.
- 115 Ricardo Baeza-Yates and Carlos Castillo: "Balancing Volume, Quality and Freshness in Web Crawling." in A. Abraham, J. Ruiz-del-solar, M. Köppen (Eds.), *Soft-Computing Systems: Design, Management and Applications*, Frontiers in Artificial Intelligence and Applications 97, IOS Press, pp. 565-572, 2002.
- 116 Ricardo Baeza-Yates, Felipe Saint-Jean and Carlos Castillo: "Web Structure, Age, and Page Quality." In *Proc. of Symposium on String Processing and Information Retrieval (SPIRE) 2002*, pp. 113-131. Springer LNCS. [137 citations](#).
- 117 Ricardo Baeza-Yates and Carlos Castillo: "Relating Web characteristics with link based Web page ranking." In *Proc. of Symposium on String Processing and Information Retrieval (SPIRE) 2001*, Laguna San Rafael, Chile, pp. 21-32. IEEE Cs. Press.

Diversity in Recommendation

- 118 Lorenzo Porcaro, Emilia Gómez, Carlos Castillo: "Assessing the Impact of Music Recommendation Diversity on Listeners: A Longitudinal Study." *ACM Transactions on Recommender Systems*, Vol. 2, No. 1, 2024. ACM Press.
- 119 Lorenzo Porcaro, Emilia Gómez, Carlos Castillo: "Perceptions of Diversity in Electronic Music: the Impact of Listener, Artist, and Track Characteristics." In *Proc. of ACM Conference on Computer-Supported Cooperative Work and Social Computing (CSCW) 2022*, Vol. 6, Issue CSCW1, Article 109, pp. 1-26. ACM Press.
- 120 Lorenzo Porcaro, Carlos Castillo, Emilia Gómez: Diversity by Design in Music Recommender Systems. *Transactions of the International Society for Music Information Retrieval (TISMIR)*, Vol 4, No 1, pp. 114-126. 2021.

Other Topics

- 121 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. [479 citations](#).
- 122 Francesco Bonchi, Carlos Castillo, Debora Donato, Aristides Gionis: "Taxonomy driven lumping for sequence mining." *Data Mining and Knowledge Discovery Journal*, ECML/PKDD special issue, Vol. 9, No. 2, pp. 227-244. 2009. Springer.
- 123 Patrizia Andronico, Marina Buzzi, Carlos Castillo and Barbara Leporini: "Evaluating a Modified Google User Interface Via Screen Reader." *Journal of Universal Access in the Information Society*, Vol. 7, No, 3, pp. 155-177. 2008. Springer.
- 124 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.

RESEARCH GRANTS

Note: from my graduation in 2004 and until 2015 I carried a scientific career at research labs that were not dependent on external competitive funding.

2022-2025 **Horizon Europe – FINDHR**

Coordinator in “Fairness and Intersectional Non-Discrimination in Human Recommendation.” (FINDHR) under grant agreement 101070212. Total grant amount 3.3M€.

2020-2023 **EU H2020 – SoBigData++**

PI from Partner UPF in “European Integrated Infrastructure for Social Mining and Big Data Analytics.” (SoBigData++) under grant agreement 871042. Total grant amount 10M€.

2018-2021 **EU H2020 – ALOHA**

PI from Partner UPF in “Adaptive and secure deep Learning On Heterogeneous Architectures.” (ALOHA) under Grant Agreement 780788. Total grant amount 6M€.

2019-2020 **VW Foundation – AIVC**

PI from Partner UPF in “Artificial Intelligences as Virtual Citizens” (AIVC) of the Volkswagen Foundation’s planning grants. Total grant amount 150K€.

2018-2019 **Facebook Research Awards – Big Picture**

Winner and PI of the research award “Capturing the Big Picture from Social Media for Global Flood Awareness” from Facebook’s Crisis Informatics Grants. Unrestricted gift 50K\$.

2017–2018 **Data Transparency Lab – DTL Grant**

Co-PI of the research award “A tool for fair rankings in search” from the DTL, with PhD advisee Meike Zehlike. Unrestricted gift 50K€.

2016-2017 **EU H2020 – DECODE**

PI from Partner Eurecat in “Decentralised Citizens Owned Data Ecosystem” (DECODE) under Grant Agreement 732546. Co-write proposal but transferred to UPF shortly after the project started. Total grant amount 5M€.

2014-2016 **EU H2020 – TYPES**

PI from Partner Eurecat in “Towards transparency and privacy in the online advertising business” (TYPES) under Grant Agreement 653449. Total grant amount 4M€.

2013–2015 **Qatar Foundation – CSAIL-DM**

PI from Qatar Computing of “Data Management for Social Computing” with Lalana Kagal from MIT-CSAIL, financed by Qatar Foundation. Total grant amount 2M€.

RESEARCH AGREEMENTS

Note: only recent agreements are listed.

2020– **Generalitat de Catalunya, Department of Justice**

Formal data sharing agreement to study criminal recidivism. This has resulted in a collaboration with their research division, the Center of Juridical Studies and Specialized Training.

2019– **Universitat de Barcelona, Group of Advanced Violence Studies**

Formal research agreement with the Department of Psychology, for joint research on criminal recidivism prediction. This has resulted in multiple publications.

2019–2020, 2020–2021 **Joint Research Center in Seville, Spain**

Two expert contracts for research activities within the HUMAINT project on human and machine intelligence, resulting in multiple joint publications.

2018, 2019–2020 **Joint Research Center in Ispra, Italy**

Two expert contracts for exploratory research activities on social media for disaster risk reduction, resulting on extensions to an Emergency Management System and joint publications.

2018–2021 **ETICAS Research and Consulting**

Formal collaboration agreement to perform joint algorithmic impact assessments. To date, four such assessments for clients in South America, the US, and Spain, have been conducted.

2018–2019 **Generalitat de Catalunya, Audiovisual Council**

Contract for creating a framework to monitor ethical and legal compliance of inclusion guidelines, addressing video-on-demand platforms. Report was adopted on December 2019.

GRADUATE STUDENTS

PhD Supervision

2017– Universitat Pompeu Fabra.

I am co-advising Marina Estevez (UPF), Ioannis Billionis (UPF and AdheraHealth), Karma Peiró (UPF, Dept. of Communication).

I advised PhD graduates Marzieh Karimi (UPF, 2023), Fedor Vitiugin (UPF, 2023), Valerio Lorini (UPF, 2022), David Solans (UPF, 2022), and co-advised Francesco Fabbri (UPF and Eurecat, 2022), Lorenzo Porcaro (UPF, 2022), Panayiotis Smeros (EPFL, 2022), Meike Zehlike (HU Berlin, 2022).

PhD Supervision During Internships

2012–2015 Social Computing Group, Qatar Computing Research Institute, Qatar.

I supervised 6 PhD students during internships: Aditi Gupta (IIIT Delhi), Muhammad Imran (Trento Univ. – also as a Postdoc), Alexandra Olteanu (EPFL), Hemant Purohit (Wright State Univ.), Jakob Rogstadius (Madeira Univ.), and Diego Saez-Trumper (Pompeu Fabra Univ.). M. Imran obtained a best paper award at ISCRAM (2013), A. Gupta was a runner-up for best paper at SocInfo (2014). I continued mentoring H. Purohit, D. Saez-Trumper, M. Imran, and A. Olteanu until their graduation.

2006–2012 Web Mining Group, Yahoo! Research, Spain.

I co-supervised 9 PhD students during internships: Ludovico Boratto (Univ. Cagliari), Claudio Corsi (Univ. Pisa), Dino Ienco (Univ. Turin), Michael Mathioudakis (Toronto Univ.), Michael Potamias (Boston Univ.), Jacob Ratkiewicz (Indiana Univ.), Eduardo Ruiz (Univ. California Riverside), Ilija Subasic (KU Leuven), and Robert West (Stanford). M. Potamias obtained a best student paper award at CIKM (2009), and I. Subasic obtained a best student paper award at WI (2010). I continued collaborating with D. Ienco and E. Ruiz until their graduation.

MSc Supervision

2019– Supervision for the Master in Intelligent Information Systems (UPF)

To date, I have supervised 3 MSc theses from Aideen Farrell (2020), Clothilde Breger (2020), and Sara Galindo (2019).

TEACHING

2019– **Fairness, Accountability, Transparency, and Ethics in Data Processing**

Interdisciplinary course coordinated by me and done in collaboration with faculty from law and philosophy, for undergraduate and graduate students of the Information and Communication Technologies Department at UPF. Scope: personal data processing, data-driven decision support, automated decision making.

2019– **Mining of Massive Datasets**

Part of the degree on Mathematical Engineering in Data Science at UPF. Topics: similarity metrics, association rule mining, recommender systems, outlier analysis, data streams, time series.

2018–2020 **Introduction to Network Science**

Part of the degree on Mathematical Engineering in Data Science at UPF. Topics: graph theory, random networks, scale-free networks, homophily, assortativity, link-based centrality, graph clustering, spectral methods, spreading phenomena.

2015–2016 **Algorithmic Methods of Data Science**

Part of the Master in Data Science, Sapienza University of Rome, Italy. Topics: vector space model, clustering, k-means algorithm, hierarchical clustering, summarization, indexing, text indexing, link-based ranking, graph evolution models, dense sub-graphs, spectral graph clustering, recommender systems, mining social media.

2007–2008 **Fundamentals of Databases**

Part of the degree in Telecommunications Network Engineering. Universitat Pompeu Fabra, Spain. Topics: entity-relationship model, relational algebra, relational database management systems, XML retrieval.

TUTORIALS

Data Biases and Limitations

2016-2018 Tutorial on the Limits of Social Data (with A. Olteanu, F. Diaz, and E. Kiciman)
ICWSM 2016, KDD 2017, WSDM 2018, SDM 2018, WWW 2018

2016 Tutorial on Algorithmic Biases and Discrimination-Aware Data Mining (with S. Hajian and F. Bonchi). ACM Knowledge Discovery and Data Mining (KDD)

Crisis Informatics

2014 Tutorial on Crisis Informatics (with H. Purohit and F. Diaz).
SIAM International Conference on Data Mining (SDM)

2013 Tutorial on Citizen Sensing and Crises (with H. Purohit, P. Meier, and A. Sheth).
AAAI Web and Social Media Conference (ICWSM)

Information Propagation

2012 Tutorial on Information and Influence Spread in Social Networks (with W. Chen and L. Lakshmanan). ACM Knowledge Discovery and Data Mining (KDD)

KEYNOTES AND INVITED TALKS

Algorithmic Fairness

2024 Invited talks at (i) ECIR in Glasgow, (ii) KDD Responsible AI Day in Barcelona, (iii) Discovery Science in Pisa.

2023 Invited talks at (i) WSDM Integrity Workshop (remotely), (ii) Responsible AI at Nokia (remotely), (iii) Canodrom Ateneu Innovacio, (iv) CCCB.

2022 Invited talks (remotely) at (i) ACMW-GR Winter School, (ii) TheWebConf, (iii) ICAI Netherlands, and in-person at (i) CIREP Research Ethics Day (Barcelona), (ii) MILA (Montreal),

- 2021 Invited talks (remotely) at (i) Universiteit van Amsterdam, (ii) Vlaamse AI Academie, (iii) Fairness and Bias in IR, (iv) BIAS Workshop.
- 2020 Invited talks (remotely) at (i) U. Helsinki, Finland; (ii) U. of Chile; (iii) U. Michoacana, Mexico; (iv) RMIT, Australia; (v) Red Mutis, Colombia.
- 2019 Invited talks at (i) Joint Research Center Seville, Spain; (ii) University of Oviedo, Spain; (iii) Barcelona Center for Human Rights; (iv) Big Data and AI Congress, Barcelona; (v) Facebook, Menlo Park, USA; (vi) Google, Mountain View, USA; (vii) Center for Legal Studies and Training CEJFE, Barcelona; (viii) Univ. of Minho, Portugal. Keynote at the Latin American Web Conference in San Francisco, USA.
- 2018 Invited talks at (i) Facebook, UK; (ii) Spotify, UK; (iii) University of Ioannina, Greece; (iv) University of Wuppertal, Germany. Keynotes at: (i) Russian Summer School on Information Retrieval (RuSSIR), Kazan, Russia; (ii) CIKM Workshop on Data and Bias, Turin, Italy.
- 2017 Invited talks at (i) Center for Semantic Web Research. Santiago, Chile; (ii) École Polytechnique Fédérale de Lausanne, Switzerland.
- 2016 Keynote at Dutch-Belgian Information Retrieval Workshop (DIR). Delft, Netherlands.

Crisis Informatics

- 2018-2021 Invited lectures at University College London, UK.
- 2018 Invited talk at Keynote at the Russian Summer School on Information Retrieval (RuSSIR) in Kazan, Russia.
- 2017 Invited talks at: (i) Universitat Pompeu Fabra in Barcelona, Catalonia-Spain; (ii) Autonomous University of Bucaramanga, Colombia (videoconf.); (iii) Glasgow University, Scotland-UK.
- 2016 Invited talks at: (i) Workshop on Management of Large Forest Fires in Toledo, Spain; (ii) Workshop on Data for Good in Turin, Italy; (iii) Institute for Public Security of Catalonia, (iv) Machine Learning Study Group, (v) and Data Beers in Barcelona; (vi) Uppsala University, Sweden.
- 2015 Invited talks at (i) Global Humanitarian Policy Forum at United Nations, New York, USA; (ii) École Polytechnique Fédérale de Lausanne, Switzerland; (iii) Observational Studies through Social Media and Other Human-Generated Content in Stanford, CA, USA.
Keynote at WebMedia in Manaus, Brazil.
- 2014 Invited talk at Crisis Mappers Conference. Nairobi, Kenya.
Keynote at International Conference on Big Data Analytics. Delhi, India.
- 2013 Invited talk at Conference on Big Data Systems, Applications, and Privacy. Abu Dhabi, UAE.

Social Media and Online News

- 2016 Invited talk at Consell de l'Audiovisual de Catalunya.
- 2014 Invited lectures at summer schools: (i) European Summer School in Information Retrieval (ESSIR) in Thessaloniki, Greece; (ii) Artificial Intelligence Summer School in Beirut, Lebanon.
- 2013 Invited lecture at Tow Center, Journalism School, Columbia University. New York, USA.
Keynote at the Social News on the Web (SNOW) workshop. Rio de Janeiro, Brazil.
- 2012 Invited talk at International Conference on Extending Database Technology (EDBT). Berlin, Germany.

ORGANIZATION OF SCIENTIFIC MEETINGS

- 2020 **Fairness, Accountability, and Transparency (ACM FAT*) in Barcelona**
General Co-Chair with Mireille Hildebrandt.

- 2015, 2016, 2018 **Social Web for Disaster Management Workshop (SWDM)**
Chair. Co-located with WSDM in Los Angeles and Indianapolis in the US and Pisa, Italy.

- 2011–2014 **Web Quality Workshop (WebQuality)**

Chair and co-founder; this workshop was born from the fusion of AIRWeb and WICOW. Co-located with WWW in India, France, Brazil, and Korea.

2007–2008 **Adversarial Information Retrieval on the Web (AIRWeb)**

Chair. Co-located with WWW in Beijing, China and Banff, Canada.

ACADEMIC SERVICE

Editorial Boards and Executive Committee

2020–2023 Fairness Accountability and Transparency (FAcCT) – Executive Committee.

2015– Transactions on the Web, ACM.

2018–2022 Transactions on Social Computing, ACM.

2011–2017 Journal of Internet Research, Emerald.

Program Co-Chair

2023 The Web Conference (WWW). US.

2016–2017 ACM Digital Health Conference (DH). Canada-UK.

2014 Web Search and Data Mining Conference (WSDM) Research Track. US.

Program Committee (PC) Member, Senior PC Member, Track Chair

2008– PC/SPC Member of all major conferences in my area, including KDD (2010-2011, 2013-2016 and 2018 as PC, 2019, 2021, 2022 as Senior PC), WSDM (2008-2013, 2015 as PC, 2021, 2022 as Senior PC), WWW (2008-2014, 2016, 2020 as PC, 2015, 2018, and 2019 as Track Co-Chair, 2021, 2024 as Senior PC), SIGIR (2009-2011 as PC, 2012-2017 and 2019, 2021 as Senior PC, 2018 as Track Co-Chair), ICWSM (2010-2014 as PC, 2015-2016 and 2017 as Senior PC), CIKM (2009, 2014 as PC, 2015-2017, 2019-2022 as Senior PC)

PRESS COVERAGE

2016– **Algorithmic Discrimination**

International Science Daily (Jun 2022), Forbes (Aug 2019), ZDNet (Sep 2017).

Spain El Confidencial (Apr 2024), Diari Ara (Mar 2024), El Periódico (Apr 2020, Nov 2021), TV3 (Feb 2019, Jul 2017, May 2017), El País (Nov 2018), RTVE (Dec 2017, Jan 2017), ABC (Jul 2017), La Vanguardia (May 2019, Jul 2017, Feb 2017), La 2 (Jan 2017).

2013–2020 **Humanitarian Technologies**

International GRIST (Feb 2020), Venture Beat (Apr 2019), Engadget (Apr 2019), The Economist (Jan 2018), New Scientist (Apr 2016), Washington Post (Apr 2015), The Guardian (Dec 2014), CNN (Sep 2014), and specialized publications such as IRIN News (Nov 2013), UN Dispatch (Jan 2015), Emergency Management (Feb 2015).

Spain La Sexta (May 2019), La Vanguardia (Nov 2018).

2011–2017 **Information Credibility**

International The New Yorker (May 2014), The Washington Post (May 2014), The Huffington Post (Dec 2012), Slate Magazine (Dec 2012), The Wall Street Journal (Oct 2011), and specialized publications such as New Scientist (Apr 2013), The Verge (Dec 2012), Gizmodo (Dec 2012), MIT Technology Review (Apr 2013 and Apr 2015).

Spain El País (Apr 2017).

ACCREDITATION AND SCIENTIFIC SOCIETIES

2020 ACM Distinguished Member (Senior Member since 2015, Member since 2011)

2019 ELLIS Fellow

2016 Accredited as full professor (“*Catedrático*”) by AQU, the Agency for the Quality of the University System in Catalonia.

2015 IEEE Senior Member (Member since 2011).