

# Fernando Diaz

## Curriculum Vitae

### Education

PhD	<b>University of Massachusetts Amherst</b> Computer Science “Regularizing Query-Based Retrieval Scores” James Allan (chair), W. Bruce Croft, Sridhar Mahadevan, John Staudenmayer	2008
MS	<b>University of Massachusetts Amherst</b> Computer Science “Browsing-Based User Language Models” James Allan (advisor)	2004
BS	<b>University of Michigan Ann Arbor</b> Computer Science	1998
BA	<b>University of Michigan Ann Arbor</b> Political Science	1998

### Research Experience

<b>Microsoft Research</b> Montréal, QC	Principal Research Manager June 2018-present
---	---

I lead a multidisciplinary research group focused on fairness, accountability, transparency, and ethics (FATE). The group externally publishes basic research and actively collaborates with engineering teams in Microsoft. In addition, I maintain an active research agenda in core information retrieval and its application to production systems such as bing.

<b>Spotify</b> New York, NY	Director of Research March 2017-May 2018
--------------------------------	---

I led the applied research organization of twenty researchers focused on search, recommendation, metrics, and evaluation. I helped define the company’s research approach and culture, including how it interacts internally with product teams and externally with the academic community. In addition, I was an active member of the recommendations leadership team, involved in strategic planning for a two year product roadmap.

<b>Microsoft Research</b> New York, NY	Senior Researcher May 2012-March 2017
---	--

I led and participated in information retrieval research projects covering crisis informatics, attention modeling, fairness in machine learning, text summarization, and deep learning for search. These themes included extensive collaboration internally with product groups and externally with academic institutions.

<b>Yahoo Research</b> New York, NY	Senior Research Scientist January 2008-March 2012
---------------------------------------	--

I led and participated in information retrieval research projects including federated search, time-sensitive web search, match-making systems, multi-document summarization, and modeling user mouse behavior. My role balanced developing production code, technology transfer from research to applied research and engineering groups, and presentation of research results to the wider research community.

**Institute for Pure and Applied Mathematics**  
Los Angeles, CA

Visiting Fellow  
September 2007–December 2007

I was a visiting fellow at UCLA for the program in “Mathematics of Knowledge and Search Engines”.

**University of Massachusetts**  
Amherst, MA

Research Assistant  
September 2001–August 2007

Dissertation research focused on information retrieval under the supervision of James Allan.

**Overture Research (now Yahoo Research)**  
Pasadena, CA

Intern  
June 2003–December 2003

I collaborated with Dr. Rosie Jones to develop tools for the analysis and classification of search queries into temporal categories.

**Sony Computer Science Laboratories**  
Tokyo, Japan

Intern  
June 2001–August 2001

As part of participation in the NSF Summer Institute in Japan, I worked with Dr. Hitoshi Iida and Dr. Koiti Hasida of Sony Computer Science Laboratories on the construction of a dialog-based information retrieval system.

**Multi-Agent Systems Laboratory**  
Amherst, MA

Research Assistant  
September 2000–June 2001

I collaborated with Professors Victor Lesser and Beverly Woolf on a project which leveraged information retrieval techniques in a distributed educational system. We developed a system to infer symbolic and planning primitives from statistical text analysis.

**Bell & Howell Information and Learning**  
Ann Arbor, MI

Research Assistant  
May 1999–August 2000

I worked on the implementation of information retrieval research in a real world search system. Tasks included system design and engineering and statistical evaluation of system robustness.

## Teaching Experience

### Web Search Engines

New York University (Courant)

Spring 2013, Fall 2014, Fall 2016

Co-instructed a graduate level course on web search engines. Classes consisted of lectures with evaluation based on homeworks, exams, and a final project.

### Web Search

Asian Summer School in Information Access (ASSIA 2013)

Summer 2013

Graduate level lecture on web search engines.

**Experimental Design for Information Systems**

University of Trento

Summer 2012

Co-instructed a graduate level course on information retrieval and data mining. Classes included a combination of lectures and paper discussion.

**Advanced Information Retrieval and Databases**

New York University (Polytechnic)

Spring 2011

Co-instructed a graduate level seminar on advanced information retrieval. Classes included a combination of lectures and paper discussion.

**Information Retrieval**

University of Massachusetts Amherst

Fall 2006

Prepared slides and provided lectures for a graduate-level information retrieval course. Instructor: Professor James Allan.

**Applied Information Theory**

University of Massachusetts Amherst

Fall 2006

Provided grading support for a graduate-level information theory course. Instructor: Professor Erik Learned-Miller.

**Databases**

University of Massachusetts Amherst

Spring 2006

Provided teaching assistance for a graduate-level database course. This included grading home works, holding office hours, and maintaining a course web page. Instructors: Professors Yanlei Diao and Gerome Miklau.

**Supervision  
Experience**

**Former Interns**

Jaime Arguello (2009, 2010), Associate Professor, University of North Carolina.

Ahmed Hassan (2009), Research Manager, Microsoft Research.

Jangwon Seo (2010), Software Engineer, Google.

Qi Guo (2011), Software Engineer, Google.

Matthew Ekstrand-Abeug (2013), Software Engineer, Google.

Pavel Metrikov (2013), Data Scientist, Microsoft.

Teresa Bracamonte (2013), Software Engineer, Equifax.

Rishabh Mehrotra (2016), Research Scientist, Spotify.

Ioannis Paparrizos (2014), PhD Student, Columbia University.

Chris Kedzie (2015), PhD Student, Columbia University.

David Abel (2015), PhD Student, Brown University.

Cristina Garbacea (2016), PhD Student, University of Michigan.

Jesse Anderton (2017), PhD Student, Northeastern University.

**PhD Examiner**

Jaime Arguello, Carnegie Mellon University, 2011.

Maria-Hendrike Peetz, University of Amsterdam, 2015.

Matthew Ekstrand-Abeug, Northeastern University, 2017.

Jesse Anderton, Northeastern University, *in progress*.  
Rodrigo Nogueira, New York University, *in progress*.

## Publications

### Metrics

	<i>ACM Digital Library</i>	<i>Scopus</i>	<i>Google Scholar</i>
articles	58	70	89
citations	1059	1667	3691
citations/article	18.26	27.11	41.47
h-index	17	24	31

### Thesis

**Fernando Diaz**. *Autocorrelation and Regularization of Query-Based Retrieval Scores*. PhD thesis, University of Massachusetts Amherst, February 2008

### Chapter

Jaime Arguello and **Fernando Diaz**. *Relevance Ranking of Vertical Search Engines*, chapter Vertical Selection and Aggregation. 2013

### Journal

J. Shane Culpepper, **Fernando Diaz**, and Mark D. Smucker. Research frontiers in information retrieval: Report from the third strategic workshop on information retrieval in lorne (swirl 2018). *SIGIR Forum*, 52(1), August 2018

Ryen W. White, **Fernando Diaz**, and Qi Guo. Search result prefetching on desktop and mobile. *ACM Trans. Inf. Syst.*, 35(3), May 2017

**Fernando Diaz**. Worst practices for designing production information access systems. *SIGIR Forum*, 50(1), June 2016

**Fernando Diaz**, Michael Gamon, Jake Hofman, Emre Kiciman, and David Rothschild. Online and social media data as an imperfect continuous panel survey. *PLoS ONE*, 11(1), 2016

Muhammad Imran, Carlos Castillo, **Fernando Diaz**, and Sarah Vieweg. Processing social media messages in mass emergency: A survey. *ACM Comput. Surv.*, 47(4), July 2015

**Fernando Diaz**. Experimentation standards for crisis informatics. *SIGIR Forum*, 48(2), December 2014

Daniel G. Goldstein, Siddharth Suri, R. Preston McAfee, Matthew Ekstrand-Abueg, and **Fernando Diaz**. The cost of annoying online ads. *Journal of Marketing Research*, 51(6), December 2014. **Finalist: Paul E. Green Award**

Hemant Purohit, Carlos Castillo, **Fernando Diaz**, Amit Sheth, and Patrick Meier. Emergency-relief coordination on social media: Automatically matching resource requests and offers. *First Monday*, 19(1), 2014

Elad Yom-Tov and **Fernando Diaz**. The effect of social and physical detachment on information need. *ACM Trans. Inf. Syst.*, 31(1), January 2013

Yi Chang, Anlei Dong, Pranam Kolari, Ruiqiang Zhang, Yoshiyuki Inagaki, **Fernando Diaz**, Hongyuan Zha, and Yan Liu. Improving recency ranking using twitter data. *ACM Trans. Intell. Syst. Technol.*, 4(1), February 2013

**Fernando Diaz**. Regularizing query-based retrieval scores. *Information Retrieval*, 10(6), December 2007

Rosie Jones and **Fernando Diaz**. Temporal profiles of queries. *ACM Trans. Inf. Syst.*, 25(3), July 2007

## Preprints

Alexandra Olteanu, Carlos Castillo, **Fernando Diaz**, and Emre Kiciman. Social data: Biases, methodological pitfalls, and ethical boundaries. Available at SSRN, 2017

## Conference

Rishabh Mehrotra, James McInerney, Hugues Bouchard, Mounia Lalmas, and **Fernando Diaz**. Towards a fair marketplace: Counterfactual evaluation of the trade-off between relevance, fairness and satisfaction in recommendation systems. In *Proceedings of the 27th ACM conference on Information and knowledge management*, 2018

Jean Garcia-Gathright, Brian St. Thomas, Christine Hosey, Zahra Nazari, and **Fernando Diaz**. Understanding and evaluating user satisfaction with music discovery. In *Proceedings of the 41st International ACM SIGIR Conference on Research and Development in Information Retrieval*, 2018

Rishabh Mehrotra, Amit Sharma, Ashton Anderson, **Fernando Diaz**, Hanna Wallach, and Emine Yilmaz. Auditing search engines for differential satisfaction across demographics. In *Proceedings of the 26th International Conference on World Wide Web*, 2017

Omar Alonso, Serge-Eric Tremblay, and **Fernando Diaz**. Automatic generation of event timelines from social data. In *Proceedings of the 2017 ACM on Web Science Conference*, 2017

Bhaskar Mitra, **Fernando Diaz**, and Nick Craswell. Learning to Match Using Local and Distributed Representations of Text for Web Search. In *Proceedings of the 26th International Conference on World Wide Web*, 2017

Matthew Ekstrand-Abueg, Richard McCreddie, Virgil Pavlu, and **Fernando Diaz**. A study of realtime summarization metrics. In *Proceedings of the 25th ACM International Conference on Information and Knowledge Management*, 2016

Rahul Goel, Sandeep Soni, Naman Goyal, John Paparrizos, Hanna Wallach, **Fernando Diaz**, and Jacob Eisenstein. *The Social Dynamics of Language Change in Online Networks*. 2016

**Fernando Diaz**. Learning to rank with labeled features. In *Proceedings of the 2016 ACM International Conference on the Theory of Information Retrieval*, 2016

**Fernando Diaz**, Bhaskar Mitra, and Nick Craswell. Query expansion with locally-trained word embeddings. In *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics*, 2016

**Fernando Diaz**, Qi Guo, and Ryen W. White. Search result prefetching using cursor movement. In *Proceedings of the 39th International ACM SIGIR Conference on Research and Development in Information Retrieval*, 2016

- Christopher Kedzie, **Fernando Diaz**, and Kathleen McKeown. Real-time web scale event summarization using sequential decision making. In *Proceedings of the 25th International Joint Conference on Artificial Intelligence*, 2016
- Fernando Diaz**. Pseudo-query reformulation. In *Proceedings of the 38th European Conference on IR Research*, 2016
- Jaime Arguello, Sandeep Avula, and **Fernando Diaz**. Using query performance predictors to improve spoken queries. In *Proceedings of the 38th European Conference on IR Research*, 2016
- Fernando Diaz**. Condensed list relevance models. In *Proceedings of the 2015 International Conference on The Theory of Information Retrieval*, May 2015
- Christopher Kedzie, Kathleen McKeown, and **Fernando Diaz**. Predicting salient updates for disaster summarization. In *Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*, July 2015
- Pavel Metrikov, **Fernando Diaz**, Sébastien Lahaie, and Justin Rao. Whole page optimization: How page elements interact with the position auction. In *EC '14: Proceedings of the fifteenth ACM conference on Electronic commerce*, 2014
- Alexandra Olteanu, Carlos Castillo, **Fernando Diaz**, and Sarah Vieweg. Crisislex: A lexicon for collecting and filtering microblogged communications in crises. In *Proceedings of The 8th International AAAI Conference on Weblogs and Social Media*, 2014
- Milad Shokouhi, Rosie Jones, Umut Ozertem, Karthik Raghunathan, and **Fernando Diaz**. Mobile query reformulations. In *Proceedings of the 37th Annual ACM SIGIR Conference*, 2014
- Peter B. Golbus, Imed Zitouni, Jin Young Kim, Ahmed Hassan, and **Fernando Diaz**. Contextual and dimensional relevance judgments for reusable serp-level evaluation. In *Proceedings of the 23rd International Conference on World Wide Web*, 2014
- Fernando Diaz**, Ryan W. White, Georg Buscher, and Dan Liebling. Robust models of mouse movement on dynamic web search results pages. In *Proceedings of the 22nd ACM conference on Information and knowledge management (CIKM 2013)*, 2013
- Muhammad Imran, Shady Elbassuoni, Carlos Castillo, **Fernando Diaz**, and Patrick Meier. Extracting information nuggets from disaster-related messages in social media. In *10th International Conference on Information Systems for Crisis Response and Management*, 2013. **Best Paper**
- Qi Guo, **Fernando Diaz**, and Elad Yom-Tov. Updating users about time critical events. In Pavel Serdyukov, Pavel Braslavski, Sergei O. Kuznetsov, Jaap Kamps, Stefan Rüger, Eugene Agichtein, Ilya Segalovich, and Emine Yilmaz, editors, *Advances in Information Retrieval*, volume 7814. 2013
- Jaime Arguello, **Fernando Diaz**, and Jamie Callan. Learning to aggregate vertical results into web search results. In *Proceedings of the 20th ACM international conference on Information and knowledge management*, 2011
- Elad Yom-Tov and **Fernando Diaz**. Out of sight, not out of mind: on the effect of social and physical detachment on information need. In *Proceedings of the 34th international ACM SIGIR conference on Research and development in Information Retrieval*, 2011. **Best Paper Nomination**
- Elad Yom-Tov and **Fernando Diaz**. Location and timeliness of information sources during news events. In *Proceedings of the 34th international ACM SIGIR conference on Research and development in Information Retrieval*, 2011

- Jangwon Seo, **Fernando Diaz**, Evgeniy Gabrilovich, Vanja Josifovski, and Bo Pang. Generalized link suggestions via web site clustering. In *Proceedings of the 20th International Conference on World Wide Web*, 2011
- Jaime Arguello, **Fernando Diaz**, Jamie Callan, and Ben Carterette. A methodology for evaluating aggregated search results. In *Proceedings of the 33rd European conference on Advances in information retrieval*, 2011. **Best Student Paper**
- Jing Bai, **Fernando Diaz**, Yi Chang, Zhaohui Zheng, and Keke Chen. Cross-market model adaptation with pairwise preference data for web search ranking. In *Proceedings of the 23rd International Conference on Computational Linguistics: Posters*, 2010
- Fernando Diaz**, Donald Metzler, and Sihem Amer-Yahia. Relevance and ranking in online dating systems. In *SIGIR '10: Proceeding of the 33rd international ACM SIGIR conference on Research and development in information retrieval*, 2010. **ICML 2011 Invited Cross-Conference Presentation**
- Jaime Arguello, **Fernando Diaz**, and Jean-François Paiement. Vertical selection in the presence of unlabeled verticals. In *Proceedings of the 33rd international ACM SIGIR conference on Research and development in information retrieval*, 2010
- Anlei Dong, Ruiqiang Zhang, Pranam Kolari, Jing Bai, **Fernando Diaz**, Yi Chang, Zhaohui Zheng, and Hongyuan Zha. Time is of the essence: improving recency ranking using twitter data. In *WWW '10: Proceedings of the 19th international conference on World wide web*, 2010
- Anlei Dong, Yi Chang, Zhaohui Zheng, Gilad Mishne, Jing Bai, Ruiqiang Zhang, Karolina Buchner, Ciya Liao, and **Fernando Diaz**. Towards recency ranking in web search. In *WSDM '10: Proceedings of the third ACM international conference on Web search and data mining*, 2010
- Jaime Arguello, Jamie Callan, and **Fernando Diaz**. Classification-based resource selection. In *CIKM '09: Proceeding of the 18th ACM conference on Information and knowledge management*, 2009
- Ahmed Hassan, Rosie Jones, and **Fernando Diaz**. A case study of using geographic cues to predict query news intent. In *Proceedings of the 17th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*, 2009
- Jaime Arguello, **Fernando Diaz**, Jamie Callan, and Jean-François Crespo. Sources of evidence for vertical selection. In *Proceedings of the 32nd international ACM SIGIR conference on Research and development in information retrieval*, 2009. **Best Paper**
- Fernando Diaz** and Jaime Arguello. Adaptation of offline vertical selection predictions in the presence of user feedback. In *SIGIR 2009*, 2009
- Fernando Diaz**. Integration of news content into web results. In *Proceedings of the Second ACM International Conference on Web Search and Data Mining*, 2009. **Best Paper**
- Fernando Diaz**. Improving relevance feedback in language modeling with score regularization. In *SIGIR '08: Proceedings of the 31st annual international ACM SIGIR conference on Research and development in information retrieval*, 2008
- Fernando Diaz**. A method for transferring retrieval scores between collections with non-overlapping vocabularies. In *SIGIR '08: Proceedings of the 31st annual international ACM SIGIR conference on Research and development in information retrieval*, 2008
- Fernando Diaz**. Theoretical bounds on and empirical robustness of score regularization to different similarity measures. In *SIGIR '08: Proceedings of the 31st annual international ACM SIGIR conference on Research and development in information retrieval*, 2008

- Fernando Diaz**. Performance prediction using spatial autocorrelation. In *SIGIR '07: Proceedings of the 30th annual international ACM SIGIR conference on Research and development in information retrieval*, 2007
- Fernando Diaz** and Donald Metzler. Pseudo-aligned multilingual corpora. In Manuela M. Veloso, editor, *IJCAI 2007, Proceedings of the 20th International Joint Conference on Artificial Intelligence*, 2007
- Fernando Diaz** and Donald Metzler. Improving the estimation of relevance models using large external corpora. In *SIGIR '06: Proceedings of the 29th annual international ACM SIGIR conference on Research and development in information retrieval*, 2006
- Fernando Diaz**. Regularizing ad hoc retrieval scores. In *CIKM '05: Proceedings of the 14th ACM international conference on Information and knowledge management*, 2005
- Fernando Diaz** and Rosie Jones. Using temporal profiles of queries for precision prediction. In *SIGIR '04: Proceedings of the 27th annual international ACM SIGIR conference on Research and development in information retrieval*, 2004
- Diane Kelly, **Fernando Diaz**, Nicholas J. Belkin, and James Allan. A user-centered approach to evaluating topic models. In *26th European Conference on Information Retrieval Research*, 2004
- Fernando Diaz**. Using wearable computers to construct semantic representations of physical spaces. In *Proceedings of the Sixth International Symposium on Wearable Computers*, 2002

## Workshop

- Sarah Bird, Solon Barocas, Kate Crawford, **Fernando Diaz**, and Hanna Wallach. Exploring or Exploiting? Social and Ethical Implications of Autonomous Experimentation in AI. In *Workshop on Fairness, Accountability, and Transparency in Machine Learning*, 2016
- David Abel, Alekh Agarwal, **Fernando Diaz**, Akshay Krishnamurthy, and Robert E. Schapire. Exploratory gradient boosting for reinforcement learning in complex domains. In *ICML 2016 Workshop on Abstraction in Reinforcement Learning*, 2016
- Fernando Diaz**. Experimentation standards for crisis informatics. In *KDD 2014 Workshop on Data Science for Social Good*, 2014
- Muhammad Imran, Shady Elbassuoni, Carlos Castillo, **Fernando Diaz**, and Patrick Meier. Practical extraction of disaster-relevant information from social media. In *Proceedings of the 22nd international conference on World Wide Web companion*, 2013
- Annie Louis, Eric Crestan, Youssef Billawala, Rao Shen, **Fernando Diaz**, and Jean-François Crespo. Use of query similarity for improving presentation of news verticals. In *Proceedings of the First International Workshop on Searching and Integrating New Web Data Sources - Very Large Data Search, Seattle, WA, USA, September 2, 2011*, 2011
- Rosie Jones, Ahmed Hassan, and **Fernando Diaz**. Geographic features in web search retrieval. In *Proceedings of the 2Nd International Workshop on Geographic Information Retrieval*, 2008
- Donald Metzler, **Fernando Diaz**, Trevor Strohman, and W. B. Croft. Umass at robust 2005: Using mixtures of relevance models for query expansion. In *The Fourteenth Text REtrieval Conference (TREC 2005) Notebook*, 2005



**Fernando Diaz** and James Allan. When less is more: Relevance feedback falls short and term expansion succeeds at hard 2005 (notebook version). In *The Fourteenth Text REtrieval Conference (TREC 2005) Notebook*, 2005

Nasreen Abdul-Jaleel, James Allan, W. Bruce Croft, **Fernando Diaz**, Leah Larkey, Xiaoyan Li, Donald Metzler, Mark D. Smucker, Trevor Strohman, Howard Turtle, and Courtney Wade. Umass at trec 2004: Novelty and hard. In *Online Proceedings of 2004 Text REtrieval Conference*, 2004

## **Tutorial**

Jean Garcia-Gathright, Christine Hosey, Brian St. Thomas, Ben Carterette, and **Fernando Diaz**. Mixed methods for evaluating user satisfaction. In *Proceedings of the 12th ACM Conference on Recommender Systems*, 2018

Carlos Castillo, **Fernando Diaz**, and Hemant Purohit. Leveraging social media and web of data to assist crisis response coordination. In *SDM '14: Proceedings of the 14th SIAM International Conference on Data Mining*, 2014

Kira Radinsky, **Fernando Diaz**, Susan Dumais, Milad Shokouhi, Anlei Dong, and Yi Chang. Temporal web dynamics and its application to information retrieval. In *Proceedings of the sixth ACM international conference on Web search and data mining*, 2013

**Fernando Diaz**, Jaime Arguello, and Milad Shokouhi. Integrating and ranking aggregated content on the web. In *WWW '12: Proceedings of the 21th international conference on World wide web*, 2012

**Fernando Diaz**, Mounia Lalmas, and Milad Shokouhi. From federated to aggregated search. In *SIGIR '10: Proceedings of the 33rd annual international ACM SIGIR conference on Research and development in information retrieval*, 2010

## **Patents**

### **Granted**

**Fernando Diaz**, D. Metzler, and S. Amer-Yahia. System for determining and optimizing for relevance in match-making systems, 2016. US Patent 9,449,282

**Fernando Diaz**. Predicting audience response for scripting, 2015. US Patent 9,159,031

Y. Chang, Z. Zheng, **Fernando Diaz**, and J. Bai. Cross-market model adaptation with pairwise preference data, 2013. US Patent 8,489,590

R. Jones, **Fernando Diaz**, and A.H. Awadallah. System and method of geo-based prediction in search result selection, 2013. US Patent 8,352,466

**Fernando Diaz**. System for integrating content from external corpora into primary search engine results, 2012. US Patent 8,150,874

Rosie Jones and **Fernando Diaz**. System and method for providing temporal search results in response to a search query, 2009. US Patent 7,577,651

### **Applied**

V. Josifovski, E. Gabrilovich, B. Pang, **Fernando Diaz**, and J. Seo. Method or system for identifying website link suggestions, 2013. US Patent Application 13/339,142

E. Yom-Tov and **Fernando Diaz**. Using user's social connection and information in web searching, 2012. US Patent Application 13/111,656

## Recognition

### Awards

British Computer Society Karen Spärck Jones Award, 2017.  
Finalist for Paul E. Green Award, Journal of Marketing Research, 2015.  
Best Paper Award, ISCRAM 2013.  
Best Paper Nomination, SIGIR 2011.  
Best Student Paper Award, ECIR 2011.  
Best Paper Award, SIGIR 2009.  
Best Paper Award, WSDM 2009.

### Fellowships

Institute for Pure and Mathematics, UCLA, Postdoctoral Fellow, 2006.  
University of Massachusetts Opportunity Fellowship, 2000–2001.  
National Science Foundation Summer Institute in Japan, 2001.

### Invited Meetings and Presentations

New York University Text as Data Seminar, 2017.  
Columbia University Data Science Industry Summit, 2015.  
Asian Summer School in Information Access, 2013.  
NSF Task-Based Search Workshop, 2013.  
Strategic Workshop on Information Retrieval in Lorne, 2012.  
International Conference on Machine Learning, Invited Cross-Conference Session, 2011  
Networks and Network Analysis for the Humanities: An NEH Institute for Advanced Topics in Digital Humanities, 2010.

## Service

### Organization

SIGIR, General Co-Chair, 2021.  
FAT\*, Program Co-Chair, 2019.  
WSDM, General Co-Chair, 2014.  
TREC Web Track, Co-Organizer, 2013–2014.  
TREC Temporal Summarization Track, Co-Organizer, 2013–2014.  
NTCIR Recipe Search Track, Co-Organizer, 2014.  
SIGIR Workshop on Time-aware Information Access, Co-Organizer, 2012–2014.  
SIGIR Workshop on Social Web Search and Mining: Analysis of User Generated Content Under Crisis, Co-Organizer, 2011.

### Reviewing

SIGIR Awards Chair, Editor, 2014–2016  
SIGIR Forum, Editor, 2012–2014  
SIGIR PC 2008–2010  
SIGIR SPC 2011–present  
CIKM Vice Chair 2009  
CIKM PC 2010–2011, 2014  
CIKM SPC 2012  
WSDM PC 2010–2011  
WSDM SPC 2012–present  
WWW PC 2011–2014  
WWW Track Chair 2017  
ICML 2010–2012  
EMNLP PC 2010  
WISE 2010

## Skills

Extensive programming C/C++ experience on Unix platforms. Extensive experience with map/reduce computing (Hadoop). Experience with GNU Scientific Library, Matlab/Octave, R, Python, Perl, Java, and scientific computing in a clustered environment. English and Spanish fluency.

**References**

*Available on request.*