

Jonathan Elsas

Language Technologies Institute
Carnegie Mellon University
5000 Forbes Avenue
Pittsburgh, PA 15213

jelsas@cs.cmu.edu
www.cs.cmu.edu/~jelsas

RESEARCH INTERESTS

Information Retrieval: retrieval in structured document collections; passage retrieval; web retrieval

Social Media: information access social environments, such as blogs, message boards and online forums

Machine Learning: machine learning for document ranking; online algorithms

EDUCATION

Carnegie Mellon University, Pittsburgh, PA — **PhD**, 2010 (Expected)

Advisor: Jaime Carbonell

University of North Carolina, Chapel Hill, NC — **Master's in Information Science**, Spring 2005

Honors: *Burmeister Information Science Fellowship* for the 2003 – 2004 academic year, awarded by the School of Information and Library Science.

Advisor: Robert Losee, Gary Marchionini

Carleton College, Northfield, MN — **BA in Mathematics**, June 1998

Honors: *sigma xi, cum laude*

REFEREED PUBLICATIONS

- Jonathan L. Elsas, Susan T. Dumais. "Leveraging Temporal Dynamics of Document Content in Relevance Ranking," Proceedings of the **Third ACM International Conference on Web Search and Data Mining (WSDM 2010)**, 2010. (to appear)
- Jonathan L. Elsas, Jaime G. Carbonell. "It Pays to be Picky: An Evaluation of Thread Retrieval in Online Forums," Proceedings of the **32nd Annual ACM Conference on Research & Development in Information Retrieval (SIGIR 2009)**, 2009.
- Eytan Adar, Jaime Teevan, Susan Dumais, Jonathan L. Elsas. "The Web Changes Everything: Understanding the Dynamics of Web Content," Proceedings of the **Second ACM International Conference on Web Search and Data Mining (WSDM 2009)**, 2009.
Best Student Paper Award
- Vitor R. Carvalho, Jonathan L. Elsas, William W. Cohen, Jaime G. Carbonell. "Suppressing Outliers in Pairwise Preference Learning," Proceedings of the **17th Annual ACM Conference on Information and Knowledge Management (CIKM 2008)**, 2008.
- Vitor R. Carvalho, Jonathan L. Elsas, William W. Cohen, Jaime G. Carbonell. "A Meta-Learning Approach to Robust Rank Learning," Proceedings of the **SIGIR 2008 Workshop on Learning to Rank for Information Retrieval (LR4IR 2008)**, 2008.
- Jonathan L. Elsas, Jaime Arguello, Jamie Callan, Jaime G. Carbonell. "Retrieval and Query Expansion Models for Blog Feed Search," Proceedings of the **31st Annual International ACM SIGIR Conference on Research & Development on Information Retrieval (SIGIR 2008)**, 2008.

- Jaime Arguello, Jonathan L. Elsas, Jamie Callan, Jaime G. Carbonell. "Document Representation and Query Expansion Models for Blog Recommendation," Proceedings of the **Second International Conference on Weblogs and Social Media (ICWSM 2008)**, 2008.
Best Paper Award Nominee
- Jonathan L. Elsas, Vitor R. Carvalho, Jaime G. Carbonell. "Fast Learning of Document Ranking Functions with the Committee Perceptron," Proceedings of the **First ACM International Conference on Web Search and Data Mining (WSDM 2008)**, 2008.
- Efron, M., Elsas, J., Marchionini, G., and Zhang, J. 2004. Machine learning for information architecture in a large governmental website. In Proceedings of the **4th ACM/IEEE-CS Joint Conference on Digital Libraries (JCDL '04)**. ACM Press, New York, NY, 151-159.
- Junliang Zhang, Tim Shearer, Gary Marchionini, Miles Efron, [Jonathan] Elsas. *RelationBrowser++: an information exploration and searching tool.*, **DG.O** 2004.

UNREFEREED PUBLICATIONS

- Jonathan Elsas, Pinar Donmez, Jamie Callan, Jaime G. Carbonell. "Pairwise Document Classification for Relevance Feedback," Proceedings of the **2008 Text REtrieval Conference (TREC 2008)** (to appear)
- Jaime Arguello, Jonathan Elsas, Changkuk Yoo, Jamie Callan, Jaime G. Carbonell. "Document and Query Expansion Models for Blog Distillation," Proceedings of the **2008 Text REtrieval Conference (TREC 2008)**
Second Best performing group in the TREC 2008 Blog Distillation task
- Jonathan Elsas, Jaime Arguello, Jamie Callan, Jaime G. Carbonell. "Retrieval and Feedback Models for Blog Distillation," Proceedings of the **2007 Text REtrieval Conference (TREC 2007)**
Best performing group in the TREC 2007 Blog Distillation task
- Jonathan Elsas. "An Evaluation of Projection Techniques for Document Clustering: Latent Semantic Analysis and Independent Component Analysis," *Master's Thesis, UNC-CH School of Information and Library Science*, July, 2005.
- Gary Marchionini, Stephanie W. Haas, Junliang Zhang, Jonathan Elsas. "Accessing Government Statistical Information," **Computer**, vol. 38, no. 12, pp. 52-61, December, 2005.

PROFESSIONAL EXPERIENCE

Engineering Intern,

Google Product Search, Google, Pittsburgh, PA — September-December, 2009

Worked with **Natalie Glance**, on Google Product Search Reviews.

Research Intern,

Microsoft Research, Microsoft, Redmond, WA — July-September, 2008

With **Susan Dumais**, investigated leveraging dynamic content characteristics in web search.

Research Assistant,

Language Technologies Institute, Carnegie Mellon University, Pgh., PA — 2005-Present

With **Jaime Carbonell**, researched document and passage retrieval using structured information needs over large, heterogeneous document collections. Researched techniques for applying fast machine learning algorithms to document ranking for optimization of high-precision retrieval evaluation measures.

Research Assistant,

School of Information and Library Science, UNC, Chapel Hill, NC — 2003-2005

Under the guidance of **Dr. Gary Marchionini**, led development of a web document clustering toolkit. Developed and extended several information visualization and exploration tools for link analysis and faceted exploration of web documents.

**Senior Consultant,
Covansys, Cary, NC — 2003**

Extended and maintained enterprise-class VoiceXML interpreter. Responsibilities included integrating Nuance speech recognition engine, performance analysis and tuning, and evaluating and developing automated testing mechanisms.

**Java Developer and Performance Engineer,
Lucent Technologies, Cary, NC — 2000-2002**

Duties included: Multi-threaded server-side Java programming; performance and stability modeling for existing and proposed software; performance test plans, evaluation tools and analysis. Projects included: next-generation prototype for an internet call waiting service; e-business proactive web help system; VXML voice gateway; CORBA-based telecom-network access gateway based on the PARLAY standard.

**Automation Analyst,
Acadia Insurance, Westbrook, ME — 1999-2000**

Developed, enhanced, tested, and implemented software for entry, processing, and statistical reporting of insurance policy information using a COBOL based programming language. Compiled and analyzed data for management reports using MS Excel and MS Access.

PROFESSIONAL SERVICE

Program Committee, NAACL-HLT 2010, SIGIR 2009, CIKM 2008 Workshop on Search in Social Media (SSM 2008)

Reviewer, WWW 2008, TOIS 2008

Co-Organizer, Information Retrieval Discussion Series at CMU, 2007-present

System Administrator, ACM SIGIR conference paper review system for SIGIR 2005

REFERENCES

Jaime Carbonell, jgc@cs.cmu.edu

Jamie Callan, callan+@cs.cmu.edu

Gary Marchionini, march@ils.unc.edu

Robert Losee, losee@unc.edu