

CONTACT Jason James Soo *Phone:* (773) 442-2225
INFORMATION Alexandria, VA *E-mail:* jason@cs.georgetown.edu

OBJECTIVE A position solving *real* problems requiring creative solutions. Focuses include information retrieval, data mining, and an interest in security. Research positions are desired.

EDUCATION **Georgetown University**, Washington D.C.

- Ph.D., Computer Science (*post quals.*) **August 2011 to Present**

Illinois Institute of Technology, Chicago, IL

- M.S., Computer Science **June 2009 to May 2010**
- B.S., Computer Science **Aug. 2005 to May 2009**

PUBLICATIONS Primary and secondary author:

- J. Soo, *A Non-Learning Approach to Spelling Correction in Web Queries*, 22nd International World Wide Web Conference (WWW), Rio de Janeiro, Brazil, May 2013. (to appear).
- J. Soo, O. Frieder, *On Foreign Name Search*, 32nd European Conference on Information Retrieval (ECIR), Milton Keynes, UK, March 2010.
- Cathey, Rebecca and Soo, Jason and Frieder, Ophir and Amir, Michlean and Frieder, Gideon, *Worldwide Accessibility to Yizkor Books*, Next Generation Information Technologies and Systems, Springer Berlin / Heidelberg, Vol. 5831, 2009.
- J. Soo, R. Cathey, O. Frieder, M. Amir, and G. Frieder, *Yizkor Books: A Voice for the Silent Past*, ACM Seventeenth Conference on Information and Knowledge Management (CIKM), Napa Valley, California, October 2008.
- J. Soo, O. Frieder, *Segments: A sub-string matching system*, JASIST, *pending*.

PROFESSIONAL **MITRE Corporation**¹ - Data Analytics and Knowledge Department, McLean, VA

Software Systems Engineer, Sr. **August 2011 to Present**

- Award recipient of the Young Researcher Fellowship. Awarded to 4 people at MITRE. Award includes funding for one year to research a topic of interest.
- Lead two research efforts: forensics analysis on large data collections and protocol reverse engineering
- Experience working within the intelligence community
- Rapid prototyping to help meet urgent customer needs
- Task lead: managed customers, deadlines, and deliverables
- Working towards my EnCase Forensics certificate
- Lead an effort to develop educational data mining videos

Orbitz Worldwide² - Hotel Connectivity Team, Chicago, IL

Software Engineer II **July 2010 to May 2011**

- Use Hadoop to mine cached market availability for hotel supply
- Assist in envisioning and deploying data capture projects for data mining use

¹Details of projects and performance are limited due to the sensitive nature of the work in regard to national security.

²Took time off between Orbitz and MITRE to travel and climb Mt. Kilimanjaro

- Use Weka to design and compare various clustering algorithms for hotel recommendations
- Collaborate with various teams to design a custom cache eviction algorithm
- Implement a new framework for custom cache eviction policies
- Develop large scale, high capacity production systems
- Design, coordinate, and analyze the result of several production feature tests
- Purposed and prototyped a data-visualization tool, later adopted by Orbitz
- Orbitz “Best Idea Contest” winner

ViViD Sounds - Chicago, IL

Founder, Co-Owner and Lead Software Developer **Aug. 2008 to Jan. 2010**

- Designed and developed a results guaranteed geo-local search engine
- Designed, developed, and deployed ViViD Sounds³ in Ruby on Rails
- Managed the company’s software and organization
- Promoted the company and built intra-business agreements

MojoTech - Chicago, IL

Founder, Co-Owner, and Software Developer **Dec. 2007 to Dec. 2008**

- Implemented large scale online survey systems for Fortune 500 companies
- Provided time/cost estimates, and monitored deadlines
- Managed a team of 4 developers

Information Retrieval Laboratory - IIT, Chicago, IL

Research Assistant **Nov. 2005 to May 2010**

- Lead developer of USHMM’s⁴ Locality Searcher⁵ and Yizkor Book Search
- Worked on SQLGenerator⁶ - A scalable XML retrieval engine
- Developed a querying interface for the SQLGenerator engine
- Developed a language independent spell checking algorithm used by the USHMM
- Designed and developed a semi-real-time Twitter summarization search engine
- Designed and ran tests to determine indexing/retrieval effectiveness and efficiency
- Developed a query builder capable of querying multiple search engines with a single query which was used by the government
- Reviewed and edited an introductory Ruby programming book
- Assisted in developing an insider misuse detection system

TECHNICAL	Focus in data mining, information retrieval, forensics, large data, and web development Programming Languages: BASH, Objective-C, Java, PHP, Python, Ruby/RoR Operating Systems: Linux, OS X, and Windows Applications: Apache, git, Hadoop, Lucene, Make/Rake, MongoDB, MySQL, SSH, Vim, Weka
INTERESTS	Alternative energy, cyber war/security, entrepreneurship
HOBBIES	Cycling, DJing, frisbee, iPhone development, photography, reading, scuba diving, skydiving, snorkeling, soccer, volleyball

³<http://www.vividsounds.com>
⁴United States Holocaust Memorial Museum
⁵<http://zchor.ir.iit.edu/locations>
⁶<http://ir.iit.edu/projects/SQLGenerator.html>