

# Curriculum Vitae

**Grace Hui Yang**

Georgetown University  
Department of Computer Science  
St. Mary's Hall Room 338  
3700 Reservoir Road, NW  
Washington, DC, USA, 20057

Telephone: +1 (202) 687-6355  
Email: [huiyang@cs.georgetown.edu](mailto:huiyang@cs.georgetown.edu)  
Web: <http://www.cs.georgetown.edu/~huiyang>

## Education

- 2011                    **Ph.D. in Language and Information Technology**  
                          **Carnegie Mellon University**, Pittsburgh, PA, USA  
                          Language Technologies Institute, School of Computer Science  
                          Thesis: Personalized Concept Hierarchy Construction  
                          Advisor: Jamie Callan  
                          Committee Members: Jaime Carbonell, Christos Faloutsos, and Eduard Hovy
- 2006                    **Master of Language Technologies**  
                          **Carnegie Mellon University**, Pittsburgh, PA, USA  
                          Language Technologies Institute, School of Computer Science
- 2004                    **Master of Computer Science**  
                          **National University of Singapore**, Singapore  
                          Department of Computer Science, School of Computing  
                          Thesis: Event-based Question Answering  
                          Advisor: Tat-Seng Chua  
                          Committee Members: Hwee-Tou Ng and Min-Yan Kan
- 2001                    **B.Sc. in Computer Science**  
                          **National University of Singapore**, Singapore  
                          Department of Computer Science, School of Computing  
                          Senior Thesis: Online Video Recording and Streaming  
                          Advisor: Kim-Teng Lua

## Professional Experience

- Aug 2011–present    **Assistant Professor**, Department of Computer Science, Georgetown University, USA  
Jun 2012–Aug 2012   **Visiting Researcher**, U. S. Patent and Trademark Office (USPTO), USA  
Jun 2009–Sep 2009   **Research Intern**, Microsoft, Redmond, WA, USA  
Aug 2004–Jul 2011   **Research Assistant**, Carnegie Mellon University, USA  
Aug 2001–Jul 2004   **Full-time Tutor**<sup>1</sup>, National University of Singapore, Singapore  
May 1999–Jan 2000   **Engineering Intern**, Singapore Telecommunication, Singapore

---

<sup>1</sup> Tutor in National University of Singapore, who then adopted a British higher education scheme, is different from teaching assistant in the American universities. Tutor in National University of Singapore is a full-time position, similar to a junior teaching staff. The major duty included 10 to 12 hour in-classroom instructions per week to a student body of 15 to 25.

## Awards & Honors

2015	National Science Foundation Faculty Early Career Development (CAREER) Award
2015	Best Paper Runner-up, IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM 2015)
2001	First Class Honors, Computer Science, National University of Singapore
1996-2001	Singapore Ministry of Education Scholarship

## Grants / Funding

1. National Science Foundation CAREER IIS-1453721. Statistical Information Retrieval Modeling for Complex Search. Hui Yang (PI). \$552,012 (expected; continuing grant renewable each year until Jan 2020, with \$290,261 currently awarded). Feb 2015-Jan 2020.
2. Defense Advanced Research Projects Agency FA8750-14-2-0226. Dynamic Information Retrieval Modeling for Domain-Specific Search. Hui Yang (PI). \$947,747 (expected; continuing grant renewable each year until Sep 2017, with \$677,974 currently awarded. Initially awarded \$697,989 in Sep 2014, with additional \$ 249,758 awarded in May 2016). Sep 2014 – Oct 2017.
3. National Science Foundation CNS-1223825: TWC: Small: Assessing Online Information Exposure Using Web Footprints. Lisa Singh (PI), Micah Sherr (co-PI), Hui Yang (co-PI), \$499,996. Jan 2013 – Dec 2016.
4. National Science Foundation III-1649045. III: Student Travel Fellowships for SIGIR Privacy-Preserving IR Workshop 2016. Hui Yang (PI). \$12,000. Sep 2016 – Aug 2017.
5. United States Patent and Trademark Office 12289075: Evaluating Search and Linguistics Tools for Patents. Hui Yang (PI). \$41,993. Jun 2012 – Aug 2012.

Total: \$2,053,748 in funding awarded to Georgetown University. \$1,553,752 as PI.

## Publications

*When possible, the acceptance rates of the conference papers are obtained directly from the publication venue, as well as from the Association for Computing Machinery (ACM) Digital Library. The Journal Impact Factors are obtained from the Thomson Reuter's Journal Citation Report in the publication year. In the field of Information Retrieval, conference full papers are usually 8 to 12 pages long, while short papers are 4 to 6 pages long.*

*For articles published while at Georgetown University (Fall 2011 and later), Georgetown University student authors or post-doc authors under my supervision are marked with \*.*

## Book

1. Grace Hui Yang, Marc Sloan, Jun Wang. Dynamic Information Retrieval Modeling. Synthesis Series of lectures/monographs on Information Concepts, Retrieval, and Services. Morgan & Claypool. ISBN-10: 1627055258. ISBN-13: 978-1627055253. June 1 2016.

## Journal Articles

*(Peer-reviewed)*

1. Hui Yang. Dynamic Browsing Hierarchy Construction by Minimum Evolution. ACM<sup>2</sup> Transactions on Information Systems (TOIS). Vol. 33, No. 3, March 2015. Article 13. Page 13:1--13:33  
DOI: <http://dx.doi.org/10.1145/2714574>.  
(Impact factor = 1.021)
2. Hui Yang, Dongyi Guan\*, and Sicong Zhang\*. The query change model: Modeling session search as a Markov decision process. ACM Transactions on Information Systems (TOIS). Vol. 33, No. 4, Article 20, May 2015, page 1-33. DOI: <http://dx.doi.org/10.1145/2747874>  
(Impact factor = 1.021)
3. Marc Sloan, Hui Yang, Jun Wang. A Term-Based Methodology for Query Reformulation Understanding. Information Retrieval Journal (IRJ). April 2015, Volume 18, Issue 2, page 145-165.  
<http://link.springer.com/article/10.1007/s10791-015-9251-5>  
(Impact factor = 0.917)

## Magazine Articles

*(Peer-reviewed)*

1. Hui Yang and Jamie Callan. OntoCop: Constructing Ontologies for Public Comments. IEEE Intelligent Systems, 24(5), page 70-75. IEEE Computer Society. 2009.  
(Impact factor = 3.56)

## Conference Full Papers

*(Peer-reviewed)*

1. Sicong Zhang\*, Grace Hui Yang, Lisa Singh, Li Xiong. Safelog: Anonymizing Query Logs with Differential Privacy for Web Search. To appear in the Association for the Advancement of Artificial Intelligence (AAAI) 2016 Fall Symposium on Privacy and Language Technologies. Virginia, VA. November 17–19, 2016.<sup>3</sup>
2. Jiyun Luo\*, Xuchu Dong\*, Hui Yang. Learning to Reinforce Search Effectiveness. In Proceedings of the ACM SIGIR International Conference on the Theory of Information Retrieval (ICTIR 2015). Northampton, Massachusetts. September 27-30, 2015.  
(Acceptance rate = 51%)
3. Jiyun Luo\*, Xuchu Dong\*, Hui Yang. Session Search by Direct Policy Learning. In Proceedings of the ACM SIGIR International Conference on the Theory of Information Retrieval (ICTIR 2015). Northampton, Massachusetts. September 27-30, 2015.  
(Acceptance rate = 51%)
4. Jiyun Luo\*, Sicong Zhang\*, Xuchu Dong\* and Hui Yang. Designing States, Actions, and Rewards for Using POMDP in Session Search. In Proceedings of The 37th the European Conference on Information Retrieval (ECIR 2015). Vienna, Austria. March 29-April 02, 2015.  
(Acceptance rate = 23%)
5. Lisa Singh, Grace Hui Yang, Micah Sherr, Andrew Hian-Cheong, Kevin Tian, Janet Zhu and Sicong

---

<sup>2</sup> ACM is the Association for Computing Machinery, the premier membership organization for computing professionals.

<sup>3</sup> The acceptance rate is not available yet.

- Zhang\*. Public Information Exposure Detection: Helping Users Understand Their Web Footprints. In Proceedings of the 2015 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM 2015). Paris, France. August 25-28, 2015. (Best Paper Runner-up)  
(Acceptance rate = 18%)
6. Jiyun Luo\*, Sicong Zhang\*, Hui Yang. Win-Win Search: Dual-Agent Stochastic Game in Session Search. In Proceedings of the 37th Annual ACM SIGIR Conference (SIGIR 2014). Gold Coast, Australia. July 6-11, 2014.  
(Acceptance rate = 21%)
  7. Dongyi Guan\*, Hui Yang. Is the First Query the Most Important: An Evaluation of Query Aggregation Schemes in Session Search. In Proceedings of the Asia Information Retrieval Society Conference 2014 (AIRS 2014). Kuching, Sarawak, Malaysia. Dec 3-5, 2014.<sup>4</sup>
  8. Dongyi Guan\*, Sicong Zhang\*, and Hui Yang. Utilizing Query Change for Session Search. In Proceedings of the 36th Annual ACM SIGIR Conference (SIGIR 2013). Dublin, Ireland. July 28 - August 1, 2013.  
(Acceptance rate = 20%)
  9. Dongyi Guan\* and Hui Yang. Increasing Stability of Result Organization for Session Search. In Proceedings of the 34th European Conference on Information Retrieval (ECIR 2013). Moscow, Russia. March 24-27, 2013.  
(Acceptance rate = 16%, oral presentation)
  10. Hui Yang. Constructing Task-Specific Taxonomies for Document Collection Browsing. In Proceedings of Joint Meeting of Conference on Empirical Methods in Natural Language Processing and Conference on Computational Natural Language Learning (EMNLP-CoNLL2012). Jeju, Korea. July 12-14, 2012.  
(Acceptance rate = 24%)
  11. Hui Yang, Anton Mityagin, Krysta Svore, and Sergey Markov. Collecting High Quality Overlapping Labels at Low Cost. In Proceedings of the 33th Annual ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR 2010), Geneva, Switzerland. July 19-23, 2010.  
(Acceptance rate = 16.7%)
  12. Hui Yang and Jamie Callan. A Metric-based Framework for Automatic Taxonomy Induction. In Proceedings of the 47th Annual Meeting of the Association for Computational Linguistics (ACL 2009), Singapore. Aug 2-7, 2009.  
(Acceptance rate = 21%)
  13. Hui Yang and Jamie Callan. Ontology generation for large email collections. In Proceedings of the 8th National Conference on Digital Government Research (Dg.O 2008). Montreal, Canada. May, 2008.<sup>5</sup>
  14. Hui Yang and Jamie Callan, Near-Duplicate Detection by Instance-level Constrained Clustering, In Proceedings of the 29th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR 2006), Seattle, WA. Aug 6-11 2006.  
(Acceptance rate = 18.5%)
  15. Hui Yang, Jamie Callan, Stuart Shulman, Next Steps in Near-Duplicate Detection for eRulemaking. In Proceedings of the 6th National Conference on Digital Government Research (Dg.O 2006), San Diego, CA. May 21-24 2006.  
(Acceptance rate = 39.2%)

---

<sup>4</sup> The acceptance rate has not been provided by the conference organization.

<sup>5</sup> Acceptance rate is not available. The same conference in 2006 has an acceptance rate = 39.2%.

16. Hui Yang, Jamie Callan. Near-Duplicate Detection for eRulemaking, In Proceedings of the 5th National Conference on Digital Government Research (Dg.O 2005), Atlanta, GA. May 15-18 2005.  
(Acceptance rate = 37.5%)
17. Hui Yang, Tat-Seng Chua. Web-based List Question Answering. In Proceedings of the 20th International Conference on Computational Linguistics (COLING 2004), Geneva, Switzerland. Aug 23-27 2004.<sup>6</sup>
18. Hui Yang, Tat-Seng Chua, Shuguang Wang, Chun-Keat Koh. Structured Use of External Knowledge for Event-based Open Domain Question Answering. In Proceedings of the 26th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR 2003), Toronto, Canada. July 28-Aug 3 2003.  
(Acceptance rate = 17%)
19. Hui Yang, Lekha Chaison, Yunlong Zhao, Shi-Yong Neo, Tat-Seng Chua. VideoQA: Question Answering on News Video. In Proceedings of the 11th Annual ACM International Conference on Multimedia (ACM MM 2003), Berkeley, CA. Nov 2-8, 2003.  
(Acceptance rate = 17%)
20. Hui Yang, Tat-Seng Chua. QUALIFIER: Question Answering by Lexical Fabric and External Resource. In Proceedings of the 10th Conference of the European Chapter of the Association for Computational Linguistics (EACL 2003), Budapest, Hungary. April 12-17, 2003.  
(Acceptance rate = 26.5%)

#### Conference Short Papers

*(Peer-reviewed)*

1. Janet Zhu, Sicong Zhang\*, Lisa Singh, Grace Hui Yang and Micah Sherr. Generating Risk Reduction Recommendations to Decrease Vulnerability of Public Online Profiles. The 2016 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining. San Francisco, CA, USA. August 18-21, 2016.<sup>7</sup>
2. Sicong Zhang\*, Grace Hui Yang, Lisa Singh. Anonymizing Query Logs by Differential Privacy. In Proceedings of the 39th Annual ACM SIGIR Conference (SIGIR 2016). Pisa, Italy. July 17-21, 2016.  
(Acceptance rate = 40%)
3. Hui Yang, Jiyun Luo\*, Christopher Wing\*. Detecting the Eureka Effect in Complex Search Tasks. In Proceedings of The 37th European Conference on Information Retrieval (ECIR 2015). Vienna, Austria. March 29-April 02, 2015.  
(Acceptance rate = 38%)
4. Sicong Zhang\*, Jiyun Luo\*, Hui Yang. A POMDP Model for Content-Free Document Re-ranking. In Proceedings of the 37th Annual ACM SIGIR Conference (SIGIR 2014). Gold Coast, Australia. July 6-11, 2014.  
(Acceptance rate = 40%)
5. Jiyun Luo\*, Christopher Wing\*, Hui Yang, and Marti Hearst. The Water Filling Model and The Cube Test: Multi-Dimensional Evaluation for Professional Search. The 22nd International Conference on Information and Knowledge Management (CIKM 2013), Burlingame, CA, USA. October 27 - Nov 1, 2013.<sup>8</sup>
6. Sicong Zhang\*, Dongyi Guan\*, and Hui Yang. Query Change as Relevance Feedback in Session

<sup>6</sup> Acceptance rate is not available. The same conference in 2006 has an acceptance rate = 23%.

<sup>7</sup> Acceptance rate is not available.

<sup>8</sup> Acceptance rate is not available. The same conference for long paper in the same year has an acceptance rate = 16.86%.

Search. In Proceedings of the 36th Annual ACM SIGIR Conference (SIGIR 2013). Dublin, Ireland. July 28 - August 1, 2013.

(Acceptance rate = 34%)

7. W. Brad Moore, Yifang Wei, Adam Orshefsky, Micah Sherr, Lisa Singh, and Hui Yang. Understand Site-Based Inference Potential for Identifying Hidden Attributes. The 5th ASE/IEEE International Conference on Information Privacy, Security, Risk and Trust (PASSAT 2013), Washington, DC, USA. September 8-14, 2013.<sup>9</sup>

### **Poster /Extended Abstracts**

*(Peer-reviewed)*

1. Andrew Jie Zhou\*, Hui Yang, Hongkai Wu\*. Minerva II: A Novel Entity Discovery Tool. (Late Breaking Work) In Proceedings of 35th Conference on Human Factors in Computing Systems (CHI 2016). San Jose, CA, USA. May 7-12, 2016.

(Acceptance rate = 43.4%)

2. Lisa Singh, Hui Yang, Micah Sherr, Yifang Wei, Andrew Hian-Cheong, Kevin Tian, Janet Zhu, Sicong Zhang\*, Tavish Vaidya, Elchin Asgarli. Helping Users Understand Their Web Footprints (poster). In Proceedings of the 24th International World Wide Web Conference (WWW 2015). Florence, Italy. May 18-22, 2015.

(Acceptance rate = 30%)

3. Hui Yang, Tat-Seng Chua. FADA: Find All Distinct Answers (poster). In Proceedings of the 13th International World Wide Web Conference (WWW 2004). New York, USA. May 17-22, 2014.<sup>10</sup>

4. Hui Yang, Tat-Seng Chua. Effectiveness of Web Page Classification on Finding List Answers (poster). In Proceedings of the 27th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR 2004), Sheffield, South York Shire, UK. July 25–29, 2004.

(Acceptance rate = 50%)

5. Hui Yang, Tat-Seng Chua, Shuguang Wang. Modeling Web Knowledge for Answering Event-based Questions (poster). In Proceedings of the 12th International World Wide Web Conference (WWW 2003), Budapest, Hungary, May 2003.<sup>11</sup>

### **Demonstration papers**

*(Peer-reviewed)*

1. Andrew Jie Zhou\*, Jiyun Luo\*, Hui Yang. DUMPLING: A Novel Dynamic Search Engine (demo paper). In proceedings of the 38th Annual ACM SIGIR Conference 2015 (SIGIR 2015). Santiago, Chile. August 9-13, 2015.

(Acceptance rate = 50%)

2. Christopher Wing\* and Hui Yang. FitYou: Integrating Health Profiles to Real-Time Contextual Suggestion (demo paper). In Proceedings of the 37th Annual ACM SIGIR Conference (SIGIR 2014). Gold Coast, Australia. July 6-11, 2014.

(Acceptance rate = 47%)

3. Jiyun Luo\*, Dongyi Guan\*, and Hui Yang. InfoLand: Information Lay-of-Land for Session Search

---

<sup>9</sup> Acceptance rate is not available. The same conference for long paper in the same year has an acceptance rate = 9.6%.

<sup>10</sup> Acceptance rate is not available. The same conference for posters usually has an acceptance rate = 50%.

<sup>11</sup> Acceptance rate is not available. The same conference for posters usually has an acceptance rate = 50%.

(demo paper). In Proceedings of the 36th Annual ACM SIGIR Conference (SIGIR 2013). Dublin, Ireland. July 28 - August 1, 2013.

(Acceptance rate = 50%)

4. Hui Yang, Jamie Callan, Stuart Shulman. DURIAN: A demo for near-duplicate detection (demo description). In Proceedings of the 6th National Conference on Digital Government Research (Dg.O 2006), San Diego, CA. May 21-24, 2006.<sup>12</sup>

## Workshop Papers

*(Peer-reviewed)*

1. Chris A. Mattmann, Grace Hui Yang, Harshavardhan Manjunatha, Thamme Gowda N, Andrew Jie Zhou\*, Jiyun Luo\*, Lewis John McGibbney. Multimedia Metadata-based Forensics in Human Trafficking Web Data. The Search and Exploration of X-Rated Information (SEXI) workshop on the 9th ACM International Conference on Web Search and Data Mining WSDM 2016. San Francisco, USA. Feb 22-26, 2016.
2. Sicong Zhang\*, Hui Yang, Lisa Singh. Applying Epsilon-Differential Private Query Log Releasing Scheme to Document Retrieval. In the 2nd International Workshop on Privacy-Preserving Information Retrieval Workshop (PIR15) in conjunction with SIGIR 2015. Santiago, Chile. August 9-13, 2015.
3. Dongyi Guan\*, Hui Yang. Query Aggregation in Session Search. In the 3rd Workshop on Data-driven User Behavior Modeling and Mining from Social Media (DUBMOD'14) in conjunction with CIKM 2014. Shanghai, China. November 3-7, 2014.
4. Hui Yang. A Fragment-Based Similarity Measure for Concept Hierarchies and Ontologies. In the 7th International Workshop on Exploiting Semantic Annotations in Information Retrieval (ESAIR'14) in conjunction with CIKM 2014. Shanghai, China. November 3-7, 2014.
5. Sicong Zhang\*, Hui Yang, Lisa Singh. Increased Information Leakage from Text. In Proceedings of the 1st International Workshop on Privacy-Preserving IR (PIR 2014) in SIGIR 2014. Gold Coast, Australia.
6. Hui Yang. Resolving Path Inconsistency and Task Specification in Taxonomy Construction. In the People's Web Meets NLP workshop of the 50th Annual Meeting of the Association for Computational Linguistics. Jeju, Korea. July 8-14, 2012.
7. Hui Yang. Automatic and Interactive Browsing Hierarchy Construction for Scientific Publication Collections. In Workshop on Mining Scientific Publications of the ACM/IEEE Joint Conference on Digital Libraries. Washington, DC, USA. June 10-14, 2012.
8. Hui Yang and Jamie Callan. Metric-Based Ontology Learning. Workshop on Ontologies and Information Systems for the Semantic Web of 17th Conference on Information and Knowledge Management. Napa Valley, CA. Oct 26-30, 2008.
9. Hui Yang and Jamie Callan. Learning the Distance Metric in a Personal Ontology. Workshop on Ontologies and Information Systems for the Semantic Web of 17th Conference on Information and Knowledge Management. Napa Valley, CA. Oct 26-30, 2008.
10. Hui Yang and Jamie Callan. Human-Guided Ontology Learning. In Second Workshop on Human-Computer Interaction and Information Retrieval (HCIR 2008). Microsoft Research, Redmond. Oct 23, 2008.

## Non-referred publications (TREC, TRECVID, CLEF)<sup>13</sup>

---

<sup>12</sup> Acceptance rate is not available.

1. Grace Hui Yang and Ian Soboroff. TREC 2016 Dynamic Domain Track Overview. In Proceedings of the Text REtrieval Conference 2016 (TREC 2016). Gaithersburg, MD, USA. Nov 15-18, 2016.
2. Raziieh Rahimi and Grace Hui Yang. An Investigation of Basic Retrieval Models for the Dynamic Domain Task. In Proceedings of the Text REtrieval Conference 2016 (TREC 2016). Gaithersburg, MD, USA. Nov 15-18, 2016.
3. Grace Hui Yang, John Frank, Ian Soboroff. TREC 2015 Dynamic Domain Track Overview. In Proceedings of the Text REtrieval Conference 2015 (TREC 2015). Gaithersburg, MD, USA. Nov 18-20, 2015.
4. Jiyun Luo\* and Hui Yang. Re-ranking via User Feedback: Georgetown University at TREC 2015 DD Track. In Proceedings of the Text REtrieval Conference 2015 (TREC 2015). Gaithersburg, MD, USA. Nov 18-20, 2015.
5. Jiyun Luo\*, Xuchu Dong\* and Hui Yang. Modeling Rich Interactions in Session Search - Georgetown University at TREC 2014 Session Track. In Proceedings of the Text REtrieval Conference 2014 (TREC 2014). Gaithersburg, MD, USA. Nov 19-21, 2014.
6. Christopher Wing\* and Hui Yang. Query Refinement: Negation Detection and Proximity Learning - Georgetown at TREC 2014 Clinical Decision Support Track. In Proceedings of the Text REtrieval Conference 2014 (TREC 2014). Gaithersburg, MD, USA. Nov 19-21, 2014.
7. Jiyun Luo\* and Hui Yang. Boosting Venue Page Rankings for Contextual Retrieval - Georgetown at TREC 2013 Contextual Suggestion Track. In Proceedings of the Text Retrieval Conference 2013 (TREC 2013). Gaithersburg, MD, USA.
8. Sicong Zhang\* and Hui Yang. Applying the Query Change Retrieval Model on Session Search - Georgetown at TREC 2013 Session Track. In Proceedings of the Text Retrieval Conference 2013 (TREC 2013). Gaithersburg, MD, USA.
9. Jiyun Luo\* and Hui Yang. Query Formulation for Prior Art Search - Georgetown University at CLEF-IP 2013. In Proceedings of the CLEF 2013 Conferences and Labs of the Evaluation Forum (CLEF 2013), Sep 23-26, 2013, Valencia, Spain.
10. Dongyi Guan\*, Hui Yang, Nazli Goharian. Effective Structured Query Formulation for Session Search. In Proceedings of the Text Retrieval Conference 2012 (TREC 2012). Nov 6-9. Gaithersburg, MD, USA.
11. Andrew Yates, Dave DeBoer, Hui Yang, Nazli Goharian, Steve Kunath, Ophir Frieder. (Not Too) Personalized Learning to Rank for Contextual Suggestion. In Proceedings of the Text Retrieval Conference 2012 (TREC 2012). Nov 6-9. Gaithersburg, MD, USA.
12. Hui Yang, Luo Si, Jamie Callan. Knowledge Transfer and Opinion Detection in the TREC2006 Blog Track. In Proceedings of the 15th Text REtrieval Conference 2006 (TREC2006), Gaithersburg, MD. Nov 14-17 2006.
13. Alex Hauptmann, Ming-Yu Chen, Mike Christel, C. Huang, W.-H. Lin, Toubin Ng, N. Papernick, A. Velivelli, Jun Yang, Rong Yan, Hui Yang, and H.D. Wactlar. Confounded Expectations: Informedia at TRECVID 2004. In Proceedings of the 13th Text REtrieval Conference Video (TRECVID2004),

---

<sup>13</sup> TREC stands for Text REtrieval Conference, which is an annual evaluation to advance Information Retrieval (IR) research conducted by the National Institute of Standards and Technology (NIST). As an experimental computer science field, it is important for IR research to be tested through standardized evaluations. Participating in TREC evaluations was thus one of the major research efforts from my students and me. The TREC publications are research reports after we have submitted our systems. TRECVID is similar evaluation series on multimedia content retrieval and CLEF is a European version of TREC.



Gaithersburg, MD. Nov 15-16 2004.

14. Tat-Seng Chua, Yunlong Zhao, Lekha Chaisorn, Chun-Keat Koh, Hui Yang, Huaxin Xu, Qi Tian. TREC 2003 Video Retrieval and Story Segmentation Task at NUS PRIS. In Proceedings of the 12th Text REtrieval Conference Video (TRECVID2003), Gaithersburg, MD. Nov 2003.
15. Hui Yang, Hang Cui, Min-Yen Kan, Mstislav Maslennikov, Long Qiu, Tat-Seng Chua. QUALIFIER in TREC-12 QA Main Task. In Proceedings of the 12th Text REtrieval Conference (TREC 2003), Gaithersburg, MD. Nov 2003.
16. Hui Yang, Tat-Seng Chua. The Integration of Lexical Knowledge and External Resources for Question Answering. In Proceedings of the 11th Text REtrieval Conference (TREC2002), Gaithersburg, MD. Nov 19-22 2002.

### **Non-refereed Magazine Articles, Newsletters, and Other Publications**

1. Hui Yang and Ian Soboroff. Privacy Preserving IR 2015: A SIGIR 2015 Workshop. ACM SIGIR Forum Volume 49, Issue 2, Jan 2016, Pages 98-101 (SIGIR Forum). ACM New York, USA. 2016.
2. Luo Si and Hui Yang. PIR 2014 the first international workshop on privacy-preserving IR: When information retrieval meets privacy and security. ACM SIGIR Forum Volume 48 Issue 2, Dec 2014, Pages 83-88, (SIGIR Forum). ACM New York, USA. 2014.
3. Hui Yang. Personal Concept Hierarchy Construction. ACM SIGIR Forum Volume 46 Issue 1, June 2012 (SIGIR Forum). ACM New York, USA. 2012.
4. Hui Yang. Automatic and Interactive Browsing Hierarchy Construction for Scientific Publication Collections. The Magazine of Digital Library Research. July/August 2012 (DLIB). Volume 18, Number 7/8. ISSN 1082-9873. 2012.

### **Proceedings (Editor)**

1. Luo Si and Hui Yang. Proceedings of the First Privacy-Preserving Information Retrieval Workshop (PIR 2014) co-located with the 37th Annual ACM SIGIR Conference 2014. Gold Coast, Australia. 2014.

### **Tutorials and Invited Talks**

*(excludes conference and workshop paper presentations)*

#### **Tutorials**

*(Peer-reviewed)*

1. Hui Yang, Marc Sloan, Jun Wang. Dynamic Information Retrieval Modeling. Tutorial in The Eighth International Conference on Web Search and Data Mining (WSDM 2015). Shanghai, China.
2. Hui Yang, Marc Sloan, Jun Wang. Dynamic Information Retrieval Modeling. Tutorial in the 37th Annual ACM SIGIR Conference 2014. (SIGIR 2014) Gold Coast, Australia. 2014.

#### **Invited Talks**

1. Statistical Modeling of Information Seeking. University of Southern California. Nov 30, 2016.
2. Dynamic Information Retrieval Modeling. Carnegie Mellon University. Sep 9, 2016.
3. Dynamic Information Retrieval Modeling. University of Amsterdam. Mar 18, 2016.
4. Dynamic Information Retrieval Modeling. York University. May 27, 2015.

5. Dynamic Information Retrieval Modeling. UC Berkeley. Apr 9, 2015.
6. Water Filling Modeling and The Cube Test, Tsinghua University. Oct 18, 2014
7. Dynamic Information Retrieval Modeling, Peking University. Oct 19, 2014
8. Dynamic Information Retrieval Modeling, Tsinghua University. Oct 18, 2014
9. Dynamic Information Retrieval Modeling. University of Maryland College Park. Oct 1, 2014
10. Deriving Personalized Browsing Hierarchy from Text. Tsinghua University. Mar 7, 2012.
11. Build Your Own Ontology: A Human-Guided Machine Learning Approach. University of Delaware. Oct 31, 2011.
12. Build Your Own Ontology: A Human-Guided Machine Learning Approach. Virginia Tech. Oct 14, 2011.
13. Personal Ontology Learning. Boston University. Apr 25, 2011.
14. Personal Ontology Learning. Media Lab. Massachusetts Institute of Technology. Mar 16, 2011.
15. Personal Ontology Learning. Georgetown University. Mar 4, 2011.
16. Introduction to Ontology Learning. Xi'an Jiaotong University. Sep 5, 2008.
17. Metric-based Ontology Learning. Chinese Academy of Science. Aug 15, 2008.
18. Question Answering and VideoQA. Georgia Institute of Technology. Nov 10, 2003.

## Teaching Experience

*New class development is marked by \*. The instructor rating is obtained via course evaluation for the question of "What is your overall evaluation of the instructor?"*

Assistant Professor	<b>Computer Science Department, Georgetown University, USA</b>	
	COSC 589 Web Search and Sense-making (Graduate)	S16
	Instructor rating: 4.86	
	COSC 282 Big Data Analytics (Undergraduate)*	F15
	Instructor rating: 4.14	
	COSC 878 Large Scale Statistical Machine Learning (Graduate) *	S15
	Instructor rating: 4.43	
	COSC 589 Web Search and Sense-making (Graduate)	S14
	Instructor rating: 5.0	
	COSC 289 Multimedia Information Processing (Undergraduate) *	F13
	Instructor rating: 3.86	
	COSC 051 Computer Science I: C++ Programming (Undergraduate)	S13
	Instructor rating: 3.42	
	COSC 488 Information Retrieval (Undergraduate)	F12
	Instructor rating: 5.0	
	COSC 688 Experimental Information Retrieval (Graduate) *	S12
	Instructor rating: 4.25	
	COSC 589 Web Search and Sense-making (Graduate) *	F11
	Instructor rating: 4.0	
	COSC 301/302/901 Independent Study	F12, S13, S14, S15, S16, F16

Teaching Assistant	<b>Carnegie Mellon University, Pittsburgh, PA, USA</b> Search Engine and Web Mining (Undergraduate) Information Retrieval (Graduate)	F10 S08
Tutor	<b>National University of Singapore, Singapore</b> Software Engineering (Undergraduate) Artificial Intelligence (Undergraduate) Multimedia Technologies (Undergraduate)	F01, F02, S03, F03, S04 F03 S02
Teaching Assistant	<b>National University of Singapore, Singapore</b> Programming in Java (Undergraduate)	F00, S01

## Student Mentoring

### *Thesis Supervision*

PhD, 2012–present	Sicong Zhang (Computer Science, Georgetown University) Thesis: Privacy-Preserving Information Retrieval. Expected graduation, May 2017.
PhD, 2014–present	Jiyun Luo (Computer Science, Georgetown University) Thesis: Dynamic Information Retrieval Modeling. Expected graduation, May 2017.
PhD, 2016–present	Zhiwen Tang (Computer Science, Georgetown University) Thesis: Retrieval Methodology for Intent Hierarchies. Expected graduation, May 2021.
MS, Graduated, 2013	Dongyi Guan (Computer Science, Georgetown University) Thesis: Structured Query Formulation and Result Organization for Session Search. Currently works at Microsoft.
MS, Graduated, 2014	Jiyun Luo (Computer Science, Georgetown University) Thesis: Prior Art Search and Its Evaluation Currently PhD student in Department of Computer Science, Georgetown University
MS, Graduated, 2016	Hongkai Wu (Computer Science, Georgetown University) Thesis: Modeling Search Engine's Explorations in Dynamic Search: An Ontological Perspective. Georgetown University. August 2016.
Undergraduate Senior Thesis, 2016–present	Angela Yang (Linguistics, Georgetown University) Thesis: Bandit algorithms for dynamic search. Expected graduation, May 2018

### *Thesis Committee*

PhD, 2012–14, Graduated 2014	Raimundo F. Dos Santos (Computer Science and Applications Department, Virginia Tech) Thesis: Similarity Measures under Spatial Context
Undergraduate Senior Thesis, 2015	Andrew Hian-Cheon (Computer Science, Georgetown University) Thesis: Building Web Footprints using Semantic Similarity
Undergraduate Senior	Janet Zhu (Computer Science, Georgetown University)

Thesis, 2015                    Thesis: Generating Recommendations to Reduce Identifiability of Public Online Profiles

Undergraduate Senior    Kevin Tian (Computer Science, Georgetown University)  
Interdisciplinary            Thesis: Common Knowledge Attack: Applying Population Norms to Infer Hidden  
Thesis, 2016                Attributes

## **Professional Service**

Editorial Board              Information Retrieval Journal  
2014 – present

Conference                  Workshop co-chair, ACM SIGIR  
Organization                ACM Special Interest Group on Information Retrieval  
2017

Workshop co-chair, ACM WSDM  
ACM International Conference on Web Search & Data Mining  
2017

Tutorial chair, ACM ICTIR  
ACM International Conference on Theory of Information Retrieval  
2016

Tutorial co-chair, ACM CIKM  
Conference on Information and Knowledge Management  
2015

Doctorial Consortium co-chair, ACM SIGIR  
ACM Special Interest Group on Information Retrieval  
2013-2014

Student Research Program chair, Dg.O  
Digital Government  
2009

Co-organizer, Carnegie Mellon University Information Retrieval Discussion Series  
2006 – 2011

Workshop                    Privacy-Preserving Information Retrieval Workshop, 2014 – 2016  
organization                TREC (Text REtrieval Conference) Dynamic Domain Track, 2015 – present

Journal Reviews            ACM Transactions on Information Systems (TOIS), 2012 – present  
Information Processing & Management (IPM), 2013  
Information Sciences (INS), 2013, 2014  
Natural Language Engineering, 2013, 2014

Grant Proposal              National Science Foundation IIS (2013, 2015, 2016)

Review Panels	National Science Foundation SaTC (2015)
Technical Program Committees	<p><b>Short Paper Program Committee Co-chair</b>, ICTIR (ACM International Conference on Theory of Information Retrieval) 2018</p> <p><b>Senior Program Committee/Area Chair</b>, Special Interest Group on Information Retrieval (SIGIR) 2014-present</p> <p><b>Senior Program Committee/Area Chair</b>, Association for Computational Linguistics (ACL) 2016</p> <p><b>Senior Program Committee</b>, European Conference on Information Retrieval (ECIR) 2017</p> <p>Program Committee, Association for Computational Linguistics (ACL) 2017</p> <p>Program Committee, ACM Special Interest Group on Information Retrieval (SIGIR) 2012-2013</p> <p>Program Committee, International World Wide Web Conference (WWW) 2016, 2017</p> <p>Program Committee, ACM International Conference on Information and Knowledge Management (CIKM) 2013, 2015</p> <p>Program Committee, ACM International Conference on Theory of Information Retrieval (ICTIR) 2015</p> <p>Program Committee, ACM International Conference on Web Search and Data Mining (WSDM) 2013-2016</p> <p>Program Committee, Annual Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL-HLT) 2012-2013</p> <p>Program Committee, ACM Special Interest Group on Information Retrieval (SIGIR) short paper 2011-2013</p> <p>Program Committee, The European Chapter of the Association for Computational Linguistics (EACL) 2014, 2017</p> <p>Program Committee, ACM Special Interest Group on Knowledge Discovery and Data Mining (KDD) 2013</p> <p>Program Committee, The Conference on Empirical Methods on Natural Language Processing (EMNLP) 2010, 2012</p> <p>Program Committee, Annual International Conference on Digital Government Research (Dg.O) 2009</p> <p>Mentor, ACM International Conference on Web Search and Data Mining (WSDM) Doctorial Consortium 2015, 2017</p>
<b>Departmental Service</b>	<p>Computational Linguistics Search Committee, 2016</p> <p>Data Analytics Program Graduate Admission Committee, 2016</p> <p>Computer Science Departmental Graduate Committee, 2011-2013, 2016-2017</p>

## **Entrepreneurship**

2010-2011

Telltale Information, Pittsburgh, PA, USA.

A spin-off company with Dr. Jamie Callan based on my Master's research on near-duplicate detection.