

## **Nazli Goharian, Ph.D.**

Information Retrieval Lab  
Dept. of Computer Science  
Georgetown University  
Washington, DC 20057  
[nazli@ir.cs.georgetown.edu](mailto:nazli@ir.cs.georgetown.edu)

### **Areas of Interest**

Information Retrieval, Text Mining, Natural Language Processing, Health/ Medical Search & Mining

### **Education**

<b>Ph.D. in Computer Science</b>	Florida Institute of Technology	July 2001
<b>M.Sc. in Computer Science</b>	George Mason University	May 1995
<b>B.Sc. in Computer Science</b>	Dortmund University	May 1992

### **Employment**

<b>Clinical Professor of Computer Science</b>	Aug. 2013 - Present
<b>Clinical Associate Professor</b> Department of Computer Science Georgetown University Washington DC	Jan. 2010 – July 2013
<b>Clinical Associate Professor</b>	Aug. 2006 – Dec. 2009
<b>Clinical Assistant Professor</b>	Aug. 2001 – July 2006
<b>Senior Lecturer</b> Department of Computer Science Illinois Institute of Technology Chicago, Illinois	Aug. 2000 – July 2001
<b>Project Manager &amp; Technical Lead of Data Warehouse</b> MCI WorldCom Conferencing Chicago, Illinois	Mar. 1999 – Aug. 2000
<b>Instructor</b> Department of Computer Science Florida Institute of Technology Melbourne, Florida	Aug. 1997 – Dec. 1998
<b>Software Engineer</b> Unisys Corporation Reston, Virginia	May 1995 – Jan. 1997
<b>Software Engineer</b> XPAND Corporation Reston, Virginia	Dec. 1993 – Mar. 1995
<b>Software Engineer</b> Telemarc, Inc. Fairfax, Virginia	Dec. 1992 – Dec. 1993

## Honors & Recognitions

**Best Long Paper Award**, EMNLP 2017

**Best poster award**, Biomedical Informatics Symposium, Georgetown, 2012

**Illinois Institute of Technology's Julia Beveridge Award (Female Faculty of the Year)**, 2009

**Teacher of the Year Award**, Computer Science Department, Illinois Institute of Technology, 2007

**Dean's Excellence Award in Teaching**, College of Science and Letters, Illinois Institute of Technology, 2005

**Teacher of the Year Award**, Computer Science Department, Illinois Institute of Technology, 2003

**Teacher of the Year Award**, Computer Science Department, Illinois Institute of Technology, 2002

## Grants

- "Enhancing the CS Undergraduate Curriculum to Include Data Mining and Information Retrieval" NSF Award # 0119469, Total of 4 participants, Co-PI. (2001-2005) \$339,500
- "CRCD/EI: An Undergraduate CS Specialization in Knowledge Management Systems" NSF Award, Total of 5 participants, Co-PI. (2004-2007) \$281,999
- "Biomedical Literature Search" IIT Research Institute, Principle Investigator (PI). (2005-2006) \$28,763
- "Infrastructure for Change: From a Teaching Department to National Prominence" NSF Award # CNS-1204347, Co-PI. (2012-2015) \$460,000
- "Health Surveillance: Detecting Health Trends via Social Media" NSF Research Experiences for Undergraduates Supplement (PI), Supplement to Award# IIP-1362046, (2014 – 2019) \$8,000
- **Adobe Research Gift** (2017-2018) \$10,000

## Thesis Supervision

### PhD Supervision

- Saket Mengle [Dec. **2010**, IIT], *Categorizing Text and Detecting Passages and Category Relationships*. [currently employed by Dataxu, Boston]
- Andrew Yates [co-advisor: Ophir Frieder; Graduated, May **2016**, GU], *Identifying Real World Concepts in Social Media*. [currently postdoc at Max Planck Institute, Germany]
- Arman Cohan [started 2013; candidate]
- Luca Soldiani [started 2013; candidate]
- Sean Macavaney [started 2016; co-advisor: Ophir Frieder]

### Master Thesis Supervision

- Ankit Jain, [Graduated **2003**, IIT], *Parallel Implementation of a Sparse Matrix Information Retrieval System*.
  - **IIT Distinguished Master's Thesis Award**. Represented IIT at The 2004 Midwestern Association of Graduate Schools (MAGS) competition.

## Patent Filled

- Systems and methods for identifying critical errors that can be used for targeted radiology resident training, January 20, 2016, under Serial No. 62/280,883

## Selected Refereed Publications

(Publications can be downloaded from IR Lab site (<http://ir.cs.georgetown.edu/publications>))

### 2017

1. Andrew Yates\*, Arman Cohan\*, Nazli Goharian, "Depression and Self-Harm Risk Assessment in Online Forums", Conference on Empirical Methods in Natural Language Processing (EMNLP), Sept 2017. **EMNLP'17 Best Long Paper Award**. \*Equal contribution

2. Luca Soldaini, Andrew Yates, Nazli Goharian, "[Denoising Clinical Notes for Medical Literature Retrieval with Convolutional Neural Model](#)", 26th ACM International Conference on Information and Knowledge Management (CIKM 2017). Nov 2017, (short).
3. Arman Cohan, Allen Fong, Raj Ratwani, Nazli Goharian, "Identifying Harm Events in Clinical Care through Medical Narratives", The 8th ACM Conference on Bioinformatics, Computational Biology, and Health Informatics (ACM BCB), Aug 2017.
4. Luca Soldaini, Nazli Goharian, Andrew Yates, "Learning to Reformulate Long Queries for Clinical Decision Support", Journal of the Association for Information Science and Technology (JASIST) Special Issue on Biomedical Information Retrieval, DOI: 10.1002/asi.23924, Sept 2017.
5. Arman Cohan, Sydney Young, Andrew Yates, and Nazli Goharian, "[Triaging Content Severity in Online Mental-Health Forums](#)", Journal of the Association for Information Science and Technology (JASIST) Special Issue on Biomedical Information Retrieval, DOI:10.1002/asi.23865, Sept 2017.
6. Arman Cohan, and Nazli Goharian, "[Contextualizing Citations for Scientific Summarization using Word Embeddings and Domain Knowledge](#)", ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR 2017), Aug 2017. (short).
7. Sean MacAvaney, Arman Cohan, and Nazli Goharian "[A Framework for Cross-Domain Clinical Temporal Information Extraction](#)", International Workshop on Semantic Evaluation (SemEval 2017), Aug 2017.
8. Arman Cohan, Nazli. Goharian, "[Scientific Document Summarization via Citation Contextualization and Scientific Discourse](#)", International Journal on Digital Libraries, Special Issue on Bibliometric-Enhanced Information Retrieval and Natural Language Processing, Springer. May 2017.
9. Arman Cohan, Alen Fong, Nazli Goharian, and Raj Ratwani. "[A Neural Attention Model for Categorizing Patient Safety Events](#)", In Proceedings of the 39th European Conference on Information Retrieval (ECIR '17), April 2017. (short).
10. Luca Soldaini and Nazli Goharian, "[Learning to Rank for Consumer Health Search: a Semantic Approach](#)", In Proceedings of the 39th European Conference on Information Retrieval (ECIR '17), April 2017. (short).

## **2016**

11. L. Soldaini and N. Goharian, [QuickUMLS: a fast, unsupervised approach for medical concept extraction](#), In Proceedings of the Medical Information Retrieval (MedIR) workshop at SIGIR 2016, July 2016.
12. A. Cohan, K. Meurer, N. Goharian, "[Temporal Information Processing for Clinical Narratives](#)", NAACL HLT 10<sup>th</sup> international workshop on semantic evaluation Semantic Analysis: Clinical TempEval (SemEval'16), June 2016.
13. A. Cohan, S. Young, N. Goharian, "[Triaging Mental Health Forum Posts](#)", NAACL HLT 3rd Computational Linguistics and Clinical Psychology - From Linguistic Signal to Clinical Reality Workshop (CLPsych'16), June 2016.
14. A. Cohan and N. Goharian, "[Revisiting Summarization Evaluation for Scientific Articles](#)", In Proceedings of the 11<sup>th</sup> International Conference on Language Resources and Evaluation (LREC'16), May 2016.

15. A. Yates, A. Kolcz, N. Goharian and O. Frieder, "[Effects of Sampling on Twitter Trend Detection](#)", In Proceedings of the 11<sup>th</sup> International Conference on Language Resources and Evaluation (LREC'16), May 2016.
16. A. Yates, J. Joselow, N. Goharian, "[The news cycle's influence on social media activity](#)", International AAAI Conference on Web and Social Media (ICWSM), May 2016. (short)
17. A. Yates, N. Goharian, O. Frieder, "[Learning the relationships between drug, symptom, and medical condition mentions in social media](#)", International AAAI Conference on Web and Social Media (ICWSM). May 2016. (short)
18. A. Cohan, L. Soldaini, N. Goharian, A. Fong, R. Filice, R. Ratwani, "[Identifying Significance of Discrepancies in Radiology Reports](#)", SIAM International Conference on Data Mining (SDM) Workshop on data Mining for Medicine and Healthcare, May 2016.

## **2015**

19. A. Cohan, and N. Goharian, "[Scientific Article Summarization Using Citation-Context and Article's Discourse Structure](#)", Empirical Methods in Natural Language Processing (EMNLP '15), 2015.
20. L. Soldaini, A. Yates, E. Yom-Tov, O. Frieder, and N. Goharian, "[Enhancing web search in the medical domain via query clarification](#)," Information Retrieval Journal, July 2015, [doi:10.1007/s10791-015-9258-y](https://doi.org/10.1007/s10791-015-9258-y).
21. A. Cohan, L. Soldaini, N. Goharian, "[Matching Citation Text and Cited Spans in Biomedical Literature: a Search-Oriented Approach](#)" North American Chapter of the Association for Computational Linguistics – Human Language Technologies (NAACL HLT 2015). (short)
22. L. Soldaini, A. Cohan, A. Yates, N. Goharian, O. Frieder, "[Retrieving Medical Literature for Clinical Decision Support](#)", 37th European Conference on Information Retrieval (ECIR 2015), 2015.
23. A. Yates, N. Goharian, O. Frieder, "[Extracting Adverse Drug Reactions from Social Media](#)", Proceedings of the Twenty-Ninth AAAI Conference on Artificial Intelligence (AAAI-15), 2015.
24. J. Parker, A. Yates, N. Goharian, and O. Frieder, "[Health Related Hypothesis Generation using Social Media Data](#)," *Journal of Social Network Analysis and Mining*, Springer, 2015.

## **2014**

25. A. Cohan, L. Soldaini, A. Yates, N. Goharian, O. Frieder, "[On Clinical Decision Support System](#)", 5th ACM conference on Bioinformatics, Computational Biology, and Health Informatics, 2014. (short paper)
26. A. Yates, J. Parker, N. Goharian, and O. Frieder, "[A Framework for Public Health Surveillance](#)", In Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC'14), May 2014.
27. A. Yates, N. Goharian, O. Frieder, "[Relevance-Ranked Domain-Specific Synonym Discovery](#)", in *36<sup>th</sup> European Conference on Information Retrieval (ECIR)*, April 2014.

## **2013**

28. A. Yates, N. Goharian, W. Yee, "[Semi-supervised Sentiment Analysis: Merging Labeled Sentences with Unlabeled Reviews to Identify Sentiment](#)", American Society for Information Science and Technology (ASIST), Nov. 2013.

29. E. W. Burger, H. Federoff, O. Frieder, N. Goharian, A. Yates, “**Social Media Communications Networks and Pharmacovigilance: SequelAE-2.0**”, IEEE 15th International Conference on e-Health Networking, Applications and Services (Healthcom), Oct 2013, (short).
30. J. Parker, Y. Wei, A. Yates, O. Frieder, N. Goharian “**[A Framework for Detecting Public Health Trends with Twitter](#)**”, The 2013 IEEE/ACM International Conference on Advances in Social Network Analysis and Mining, Aug. 2013.
31. A. Yates, N. Goharian, O. Frieder, “**[Extracting Adverse Drug Reactions from Forum Posts and Linking them to Drugs](#)**”, SIGIR Workshop on Health Search and Discovery, July-Aug 2013.
32. Y. Zhu and N. Goharian, “**[To Follow or Not to Follow: A Feature Evaluation](#)**”, 22<sup>nd</sup> International Conference on World Wide Web (WWW), May 2013 (short).
33. A. Yates, N. Goharian, and O. Frieder, “**[Graded Relevance Ranking for Synonym Discovery](#)**”, 22<sup>nd</sup> International Conference on World Wide Web (WWW), May 2013 (short).
34. A. Yates and N. Goharian, “**[ADRTrace: Detecting Expected and Unexpected Adverse Drug Reactions from User Reviews on Social Media Sites](#)**”, in *35<sup>th</sup> European Conference on Information Retrieval (ECIR)*, March 2013 (short).

## **2012**

35. Z. Tan, N. Goharian, M. Sherr, “**[\\$100,000 Prize Jackpot. Call Now! Identifying the Pertinent Features of SMS Spam](#)**”, *In proceedings of ACM 35<sup>th</sup> Conference on Research and Development in Information Retrieval (SIGIR)*, August 2012. (short)
36. J. Parker, A. Yates, N. Goharian, W. Yee, “**[Efficient Estimation of Aspect Weights](#)**”, *In proceedings of ACM 35<sup>th</sup> Conference on Research and Development in Information Retrieval (SIGIR)*, August 2012. (short)

## **2011**

37. N. Goharian, S. Mengle “**[Networked Hierarchies for Web Directories](#)**”, *20<sup>th</sup> International World Wide Web conference (WWW)*, March 2011. (short)

## **2010**

38. N. Goharian, S. Mengle, “**[Context Aware Query Classification Using Dynamic Query Window and Relationship Net](#)**”, *In proceedings of ACM 33<sup>rd</sup> Conference on Research and Development in Information Retrieval (SIGIR)*, July 2010. (short)
39. S. Mengle and N. Goharian, “**[Detecting Relationships among Categories using Text Classification](#)**”, *Journal of American Society for Information Science and Technology (JASIST)*, 61 (5), May 2010.
40. N. Goharian, O. Frieder, W. G. Yee, J. Mudrawala, “**[Enriching Peer-to-Peer File Descriptors Using Association Rules on Query Logs](#)**”, *32<sup>nd</sup> European Conference On Information Retrieval (ECIR)*, March 2010. (short)
41. S. Mengle, N. Goharian, “**Mining Temporal Relationships Among Categories**”, *ACM 25<sup>th</sup> Symposium on Applied Computing (SAC)*, March 2010.

## **2009**

42. S. Mengle and N. Goharian, “**[Ambiguity Measure Feature Selection Algorithm](#)**”, *Journal of American Society for Information Science and Technology (JASIST)*, 60 (5), April 2009.

43. J. Urbain, O. Frieder, N. Goharian, "[Passage relevance models for genomics search](#)", (Journal of) *BMC Bioinformatics*, 10(Suppl 3):S3 (19 March 2009).
44. S. Mengle and N. Goharian, "[Passage Detection Using Text Classification](#)", *Journal of American Society for Information Science and Technology (JASIST)*, 60 (4), March 2009.
45. S. Liu, Y. Mehrav, W. G. Yee, N. Goharian, and O. Frieder, "**A Sentence Level Probabilistic Model for Evolutionary Theme Pattern Mining for News Corpora**", *ACM 24<sup>th</sup> Symposium on Applied Computing (SAC)*, March 2009.
46. Platt, S. Mengle, N. Goharian, "**Improving Classification Based Off-topic Search Detection via Category Relationships**", *ACM 24<sup>th</sup> Symposium on Applied Computing (SAC)*, March 2009.
47. J. Urbain, O. Frieder, and N. Goharian, "**A Dimensional Retrieval Model for Integrating Semantics and Statistical Evidence in Context for Genomics Literature Search**", (Journal of) *Computers in Biology and Medicine* 39 (1), January 2009.

## 2008

48. J. Urbain, O. Frieder, and N. Goharian, "**Passage Relevance Models for Genomics Search**", *ACM Second International Workshop on Data and Text Mining in Bioinformatics*, Napa Valley, California, October 2008.
49. J. Urbain, N. Goharian, and O. Frieder, "[Probabilistic Passage Models for Semantic Search of Genomics Literature](#)," *Journal of the American Society of Information Science and Technology (JASIST)*, 59 (12), September 2008.
50. N. Goharian, S. Mengle, "[On Document Splitting in Passage Detection](#)", *In proceedings of ACM 31st Conference on Research and Development in Information Retrieval (SIGIR)*, July 2008. (short)
51. S. Mengle, N. Goharian, "**Detecting Hidden Passages from Documents**", *In proceedings of SIAM Conference on Data Mining (SIAM - SDM) Workshop*, April 2008.
52. S. Mengle, N. Goharian, A. Platt "**Discovering Relationships among Categories using Misclassification Information**", *ACM 23<sup>rd</sup> Symposium on Applied Computing (SAC)*, March 2008.
53. S. Mengle, N. Goharian, "**Using Ambiguity Measure Feature Selection Algorithm for Support Vector Machine Classifier**", *ACM 23<sup>rd</sup> Symposium on Applied Computing (SAC)*, March 2008.

## 2007

54. J. Urbain, N. Goharian O. Frieder, "**Combining Semantics, Context, and Statistical Evidence in Genomics Literature Search**", *The 7th IEEE Int. Conference on Bioinformatics and Bioengineering (BIBE)*, 2007.
55. A. Platt and N. Goharian "**Short Query Sequences in Misuse Detection**" *IEEE 5<sup>th</sup> International Conference on Intelligence and Security Informatics (ISI)*, May 2007.
56. N. Goharian and A. Platt "**DOTS: Detection of Off-Topic Search Via Result Clustering**" *IEEE 5<sup>th</sup> International Conference on Intelligence and Security Informatics (ISI)*, May 2007.
57. A. Platt, N. Goharian, S. Mengle "**Using User Query Sequence to Detect Off-Topic Search**", *ACM 22<sup>nd</sup> Symposium on Applied Computing (SAC)*, March 2007.

## 2006

58. N. Goharian, A. Platt "**Detection Using Clustering Query Results**" *IEEE 4<sup>th</sup> International Conference on Intelligence and Security Informatics (ISI)*, May 2006. (short)

## 2005

59. N. Goharian, L. Ma, "**Off-Topic Access Detection In Information Systems**", *ACM 14<sup>th</sup> Conference on Information and Knowledge Management (CIKM)*, November 2005. (short)

60. N. Goharian, L. Ma, C. Meyer, “**Detecting Misuse of Information Retrieval Systems Using Data Mining Techniques**” *IEEE 3<sup>rd</sup> International Conference on Intelligence and Security Informatics (ISI)*, May 2005.
61. S. Argamon, N. Goharian, D. Grossman, O. Frieder, N. Raju “**A Specialization in Information and Knowledge Management Systems for the Undergraduate Computer Science Curriculum**”, *IEEE International Conference on Information Technology: Coding & Computing (ITCC)*, April 2005.
62. L. Ma and N. Goharian “**Query Length Impact on Misuse Detection in Information Retrieval Systems**”, *ACM 20<sup>th</sup> Symposium on Applied Computing (SAC)*, March 2005.

#### 2004

63. N. Goharian, L. Ma “**Using Relevance Feedback to Detect Misuse in Information Retrieval Systems**” *ACM 13<sup>th</sup> Conference on Information and Knowledge Management (CIKM)*, November 2004. (short)
64. S. Beitzel, E. Jensen, A. Chowdhury, D. Grossman, O. Frieder, N. Goharian,” **On Fusion of Effective Retrieval Strategies in the Same Information Retrieval System**”, *Journal of American Society for Information Science and Technology (JASIST)*, 55 (10), July 2004.
65. N. Goharian, D. Grossman, O. Frieder, N. Raju “**Migrating Information Retrieval from the Graduate to the Undergraduate Curriculum**”, *Journal of Information Systems Education*, 15 (1), April 2004.
66. N. Goharian, D. Grossman, N. Raju “**Extending the Undergraduate Computer Science Curriculum To Include Data Mining**”, *IEEE International Conference on Information Technology: Coding & Computing (ITCC)*, Las Vegas, Nevada, April 2004.
67. P. Jain, N. Goharian, A. Weiser, S. Kimm, S. Kim, J. Stern, J. Pazona, C. Wambi, R. Yap, L. Blunt, and R. Nadler, “**Efficiency and Safety of the Healthtronics LithoTron® Lithotripter**”, *Journal of EndoUrology*, 18 (1), January/February 2004.

#### 2003

68. Cathey, L. Ma, N. Goharian, D. Grossman, “**Misuse Detection for Information Retrieval Systems**”, *ACM 12<sup>th</sup> Conference on Information and Knowledge Management (CIKM)*, November 2003.
69. L. Ma, N. Goharian, A. Chowdhury, M. Chung, “**Extracting Unstructured Data From Template Generated Web Documents**”, *ACM 12<sup>th</sup> Conference on Information and Knowledge Management (CIKM)*, November 2003.
70. S. Beitzel, E. Jensen, A. Chowdhury, O. Frieder, D. Grossman, N. Goharian, “**Disproving the Fusion Hypothesis: An Analysis of Data Fusion via Effective Information Retrieval Strategies**”, *ACM 18<sup>th</sup> Symposium on Applied Computing (SAC)*, March 2003.

#### 2002

71. E. Jensen, S. Beitzel, A. Pilotto, N. Goharian, O. Frieder “**Parallelizing the Buckshot Algorithm for Efficient Document Clustering**”, *ACM 11<sup>th</sup> Conference on Information and Knowledge Management (CIKM)*, 2002.
72. N. Goharian, P. Jain, G. Kora, A. Jain “**A Web-Based Medical Diagnosis and Treatment System for Urinary Tract Infections**”, *The 2002 International Conference on Mathematics and Engineering Techniques in Medicine and Biological Sciences (METMBS'02)*, Las Vegas, Nevada, June 2002.
73. D. Grossman, N. Goharian, N. Raju, O. Frieder “**Extending the Undergraduate Computer Science Curriculum To Include Information Retrieval and Data Mining**”, *IASTED International Conference on Computer and Advanced Technology in Education (CATE 2002)*, Cancun, Mexico, 2002.

#### TREC, TAC, and CLEF:

- Luca Soldaini, Will Edman, Nazli Goharian. “Team GU-IRLAB at CLEF eHealth 2016: Task 3.” Conference and Labs of the Evaluation Forum (CLEF). 2016.

- A. Cohan, L. Soldaini, S. Mengle, N. Goharian, "[Towards Citation-Based Summarization of Biomedical Literature](#)", Proc. of Text Analysis Conference (TAC – **Summarization Track**), 2014.
- L. Soldaini, A. Cohan, A. Yates, N. Goharian, O. Frieder, "[Query Reformulation for Clinical Decision Support Search](#)", Proc of the 23rd Text REtrieval Conference Proceedings (TREC - **Medical Track**), 2014.
- A. Yates , D. DeBoer , H. Yang , N. Goharian , S. Kunath , O. Frieder, "[\(Not Too\) Personalized Learning to Rank for Contextual Suggestion](#)", TREC 2012 – **Contextual Track**.
- D. Guan, H. Yang, N. Goharian, "[Effective Structured Query Formulation for Session Search](#)", TREC 2012. **Query Session Track**.
- J. Urbain, N. Goharian, O. Frieder, "IIT TREC-2007 **Genomics Track: Using Concept-based Semantics in Context for Genomics Literature Passage Retrieval**", TREC 2007.
- J. Urbain, N. Goharian, O. Frieder, "IIT TREC-2006: **Genomics Track**", *Proceedings of the Fifth Text REtrieval Conference*, November 2006.
- J. Urbain, N. Goharian, O. Frieder, "IIT TREC-2005: **Genomics Track**", *Proceedings of the Fourteenth Text REtrieval Conference*, November 2005.

#### **Book Chapter**

- S. Beitzel, E. Jensen, A. Chowdhury, D. Grossman, N. Goharian, O. Frieder "**Recent Results on Fusion of Effective Retrieval Strategies in the Same Information Retrieval System**", *Lecture Notes in Computer Science*, Springer-Verlag, ISSN: 0302-9743, Vol 2924 / 2003 ,Chapter: pp. 101 – 111, 2005.

#### **University Service @Georgetown University**

Member of the Main Campus Faculty Senate, Spring 2017, and (July 2017 – June 2020)

#### **Departmental Committees at Georgetown University (2010-Present)**

Member, Graduate Committee (2010-2011, 2017-2018)

Chair, Seminar Series (2010-Present)

Founder and organizer of Semi-Annual CS Graduate Research Presentation Days ( Nov 2012- Present)

#### **Course Development at Georgetown University (2010-Present)**

Under/Grad course:

Information Retrieval (COCS488)

non-CS-majors Information Systems (COSC 016)

Graduate courses:

Text Mining & Analysis (COSC 586)

Health Search and Mining (COSC592)

Topics in Information Systems (COSC 799)

Doctoral Seminar in Information Search and Mining (COSC 886)

#### **Teaching at Georgetown University (2010-Present)**

Cosc886-Doctoral Seminar in Information Systems: (Information Search & Mining, 2013; Text mining & NLP, 2016)

Cosc799-Topics in Information Systems: (Spring 2011, Fall 2014)

Cosc592-Health Search and Mining: (Spring 2017)

Cosc586-Text Mining & Analysis: (Fall 2011, 2012, 2013, 2015, 2017)

Cosc488-Information Retrieval: (Spring 2010, 2011, 2012, Fall 2013, 2014, 2015, 2016)

Cosc285-Data Mining: (Fall 2010, 2011, Spring 2013, 2014, 2015, 2016, 2017, Fall 2017)

Cosc280-Intro to Database: (Fall 2010, Spring 2015)

Cosc051-Computer Science I: (Fall 2016)

Cosc016-Information Systems for non-CS-majors: (Spring 2012, Fall 2012, Spring 2014)



## University Service@ Illinois Institute of Technology

### Departmental Committees at Illinois Institute of Technology (2000-2009)

Member, ABET Accreditation Committee (2002-2009)  
Member, Undergraduate Curriculum Committee (2001- 2009)  
Course supervising manager of Data Mining (2003- 2009)  
Course supervising manager of Information Retrieval Systems (2002- 2009)  
Course supervising manager of Database Management Systems (2001-2009)  
Member, Graduate Admissions Committee (2000- 2009)

### Course Development at Illinois Institute of Technology (2000-2009)

Undergraduate/Graduate course: Information Retrieval Systems (CS429)  
Undergraduate/Graduate course: Data Mining (CS422)  
Undergraduate/Graduate course: Database Organization (CS425)  
Undergraduate course: Information and Knowledge Management (CS495)  
Graduate/Doctoral Seminar: Advanced Topics in Information Retrieval (CS595)

### Teaching at Illinois Institute of Technology (2000-2009)

Database Management Systems (Spring 1998 – Fall 2009, both Fall and Spring semesters)  
Information Retrieval Systems (Fall 2002 – Fall 2009, Every Fall semester)  
Data Mining (Spring 2003 – Spring 2009, Every Spring semester)  
Information and Knowledge Management (Spring 2006)  
Advanced Topics in Information Retrieval (Fall 2007, Spring 2008)

## Student Supervision

### Master Project Supervision

- Qian Sun, [2002, IIT], *Impl. and Eval. of the Sparse Matrix Compression Algorithms for IR System.*
- Steven Beitzel, [2002, IIT], *An Analysis of Data Fusion via Effective Information Retrieval Strategies.*
- Eric Jensen, [2002, IIT], *Parallelizing the Buckshot Algorithm for Efficient Document Clustering.*
- Mi Sun Chung, [2003, IIT], *Development of an Educational Data Mining Tool.*
- Ramya Krishnamurthy, [2003, IIT], *Bioinformatics: Dynamic Generation of Biological Paths.*
- Rebecca Cathey, [2003, IIT], *Misuse Detection for Information Retrieval Systems.*
- Saket Mengle, [2007, IIT], *Text Classification, 1<sup>st</sup> Place, 2007 IIT Research Day Comp – (Grad Students)*
- Alana Platt, [2008, IIT], *Detecting Off-Topic Search.*
- Jordan Wilberding, [2009, IIT], *Encrypted Index Performance in Information Retrieval Systems*
- Jason Soo, [2009, IIT], *Search for Named Entities in Foreign Languages*

### Independent Research/Projects Supervision

- Guruprasad Kora, [2002, IIT], *Medical Informatics: Computer Assisted Diagnosis and Treatment System.*
- Sterling Stein, [2002, IIT], *Mapping of Index Compression Techniques on CSR Information Retrieval.*
- Jeff Heard, [2004, IIT], *Design and implementation of an Information Retrieval Instructional System (IRIS).*
- Ling Ma (multiple semesters), *Misuse Detection (Detecting Off-Topic Search)*
- Alana Platt, [2005, IIT], *Design and implementation of the project IRIS from Java in C++.*
- Levi Citrin, [2005, IIT], *Design and implementation of the project IRIS from Java in C++.*
- Chris Meyer, [2005, IIT], *Classifier based Misuse Detection System.*
- Paul Brar, [2006, IIT], *Web interface development for Managing Misuse Detection System.*
- Jay Munrawala, [2008-2009, IIT] *Enriching Peer-to-Peer File Descriptors using Association rules on Query Logs*
- Andrew Yates, [2009, IIT], *Efficiency improvements of IRIS and SimpleDM educational tools.*
- Jon Parker, [2011-2012, GU], *Opinion Mining and Sentiment Analysis of User Reviews*
- Henry Tan, [2011-2012, GU], *Short text (SMS) spam detection*
- Andrew Yates, [2012, 2013, GU], *Opinion Mining and Sentiment Analysis of User Reviews*
- David DeBoer, [2012, GU], *Contextual Suggestion Trec [with two other faculties & student: A. Yates]*

- Andrew Yates, [2012, GU], *Mining Social Media for Healthcare, Best Poster Award at ICBI Biomedical Informatics Symposium at Georgetown University (1 out of 28); Synonym Discovery.*
- Yanan Zhu, [2013, GU], *Whom to follow or not follow friends on Twitter*
- Jonah Joselow [2015, GU], *News cycle's influence on social media activity*
- Kevin Meurer [2015, 2016, GU], *Temporal extraction of events and relationships from clinical notes*
- Sydney Young [2016, GU], *Automatic Triaging of User Posts on a Mental Health Forum*
- William Edman [2016, GU], *Searching Web pages in response to user medical/health queries*

#### **Student Ph.D. Committee**

- Mohammed Aljlayl, [Graduated Aug. 2002, IIT], *on Arabic Search: The effectiveness of Monolingual and Bi-directional Information retrieval.*
- Vincent Nguyen, [Graduated May 2003, IIT], *Channel Alternation and Rotation for Sectorized Wireless Systems.*
- Kagan Agun, [Graduated Dec. 2004, IIT], *Hardware Assisted Pruned Inverted Index Component Design and Impl.*
- Napawan Sawasdichai [Graduated May 2004, IIT], *Outside Advisor- Institute of Design, User-centered Web design.*
- Steve Beitzel [Graduated May 2006, IIT], *On Understanding and Classifying Web Queries.*
- Eric Jensen [Graduated May 2006, IIT], *Repeatable Evaluation of Info. Retr. Effectiveness in Dynamic Environments.*
- Eric Burger [Graduated May 2006, IIT], *Ubiquitous Reach and Remote Control of Devices.*
- Rebecca Cathey, [Graduated Dec. 2007, IIT], *On Scalable Search.*
- Jay Urbain, [Graduated, 2008, IIT], *Multievidentiary Models for Genomics Search.*
- Dongmei Jia, [Graduated, Dec. 2008, IIT], *Improving Search Quality in Peer-to-Peer File Sharing Systems.*
- Linh Nguyen, [Graduated, May 2009, IIT], *Peer to Peer Information Retrieval.*
- Ahu Sieg [Graduated, June 2010, DePaul], *Contextual Information Access with Ontologies: Integrating Domain Knowledge and User Profiles.*
- Jon Parker [Graduated, May 2016, GU], *Effective and Efficient Binarization of Degraded Document Images*
- Jason Soo [Graduated May 2016, GU], *Search in adverse environments*

#### **Professional Activities**

##### **Conference Leadership Role**

**Area Chair (Information Extraction and Text Mining)**, ACL 2018

**Program Chair**, Information and Knowledge Engineering (IKE), 2003

**Conference Track Organizer**, IEEE International Conf. on Information Technology (ITCC) 2002, 2004, 2005

**Conference Track Organizer**, Information Systems Engineering, 2001 Int'l Multi-Conference, 2001

**Proceedings Chair**, ACM Conference on Information & Knowledge Management (CIKM), 2002, 2003

**Conference Session Organizer**, International Federation of Classification Societies (IFCS), 2004

##### **Conference Committee Participation Role**

Program Committee, International World Wide Web Conference; (WWW), 2012, 2018

Program Committee, International Conference on Language Resources and Evaluation (LREC), 2018

Program Committee, ACM Special Interest Group in Information Retrieval (SIGIR), 2013 - 2017

Program Committee, International Conference on Web Search and Data Mining (WSDM), 2013, 2014, 2016

Program Committee, IEEE conference on String Processing in Information Retrieval (SPIRE), 2006, 2011

Program Committee, ACM Conf. on Information & Knowledge Management (CIKM), 2001-2006, 2008, 2009

Program Committee, ACM SIGKDD Workshop on CyberSecurity and Intelligence Informatics, 2009 – 2011

Program Committee, 11th International Conference on Language Resources and Evaluation, 2018.

Program Committee, IEEE Int'l Conference on Intelligence and Security Informatics (ISI), 2007- 2011

Program Committee, IEEE/ACM Int'l Conference on Scalable Information Systems (INFOSCALE), 2006 - 2008

Program Committee, IEEE Int'l Conference on Information Technology (ITCC), 2002, 2004, 2005

Program Committee, ACM SIGIR Workshop on Open Source Information Retrieval, 2006

Program Committee, Digital Health, 2015, 2016, 2017