Dr. Nazli Goharian

Dept. of Computer Science
Georgetown University
St. Mary's Hall
Washington, DC 20057
nazli@ir.cs.georgetown.edu

Areas of Interest

Information Retrieval and Text Classification

Education

Ph.D. in Computer Science	Florida Institute of Technology	July 2001

Thesis: "A Sparse Matrix Approach for Information Retrieval"

M.Sc. in Computer Science	George Mason University	May 1995
B.Sc. in Computer Science	Dortmund University	May 1992

Employment

Clinical Associate Professor — Jan. 2010 – Present
Department of Computer Science
Georgetown University
Washington DC

Clinical Associate Professor — Aug. 2006 – Dec. 2009
Clinical Assistant Professor — Aug. 2001 – July 2006
Senior Lecturer — Aug. 2000 – July 2001
Department of Computer Science
Illinois Institute of Technology
Chicago, Illinois

Project Manager of Data Warehouse — Nov. 1999 – Aug. 2000
Project Technical Lead of Data Warehouse — Mar. 1999 – Nov. 1999
MCI WorldCom Conferencing
Chicago, Illinois

Instructor — Aug. 1997 – Dec. 1998
Department of Computer Science
Florida Institute of Technology
Melbourne, Florida

Software Engineer — May 1995 – Jan. 1997
Unisys Corporation
Reston, Virginia

Software Engineer — Dec. 1993 – Mar. 1995
XPAND Corporation
Reston, Virginia

Software Engineer — Dec. 1992 – Dec. 1993
Telemarc, Inc.
Fairfax, Virginia

<u>**Honors & Recognitions**</u>
Illinois Institute of Technology's Julia Beveridge Award (Faculty Category), 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

<u>**Professional Activities**</u>
    <u>**Conference Leadership Role**</u>
    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

    <u>**Conference Committee Participation Role**</u>
    Program Committee, ACM Conf. on Information & Knowledge Management (CIKM), 2001-2006, 2008, 2009
    Program Committee, IEEE Int'l Conference on Information Technology (ITCC), 2002, 2004, 2005
    Program Committee, IEEE/ACM Int'l Conference on Scalable Information Systems (INFOSCALE), 2006-2008
    Program Committee, IEEE conference on String Processing in Information Retrieval (SPIRE), 2006
    Program Committee, ACM SIGIR Workshop on Open Source Information Retrieval (OSIR), 2006
    Program Committee, IEEE Int'l Conference on Intelligence and Security Informatics (ISI), 2007-2010
    Program Committee, ACM SIGKDD Workshop on CyberSecurity and Intelligence Informatics , 2009

    <u>**Refereeing:**</u>
    IEEE Transactions of Computers, Journal of the American Society for Information Science and Technology,
    The Computer Journal, Information Processing Letters, Journal of Computer Systems Science & Engineering,
    Wiley Publisher: Database book, Addison Wesley: Introduction to Data Mining by P. Tan, M. Steinbach, V. Kumar

<u>**Professional Memberships**</u>
    Association for Computing Machinery (ACM)
    IEEE Computer Society

<u>**University Services**</u>
    <u>**Departmental Committees at Illinois Institute of Technology**</u>
    Member, ABET Accreditation Committee (2002-3 and 2008-9)
    Member, Undergraduate Curriculum Committee (2001- Present)
    Coordinator of course, Data Mining (2003- Present)
    Coordinator of course, Information Retrieval Systems (2002- Present)
    Coordinator of course, Database Management Systems (2001-Present)
    Member, Graduate Admissions Committee (2000- Present)

    <u>**Course Development at Georgetown University**</u>
    Undergraduate/Graduate course: Information Retrieval (COCS 416)

    <u>**Course Development at Illinois Institute of Technology**</u>
    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 course: Advanced Topics in Information Retrieval (CS595)

Teaching at Georgetown University
Information Retrieval Systems   (Spring 2010)

Teaching at Illinois Institute of Technology
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

PhD Supervision
- Saket Mengle [Graduated, Dec. 2009, IIT], *Categorizing Text and Detecting Passages and Category Relationships.*

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.

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*, **1st Place, 2007 IIT Research Day Comp – (Grad Students)**
- Alana Platt, [2008,IIT], *Detecting Off-Topic Search.*

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).
- 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.*
- Andrew Yates, [2009, IIT], Efficiency improvements of IRIS and SimpleDM educational tools.

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 Implementation.*
- Napawan Sawasdichai [Graduated May 2004, IIT], *Outside Advisor- Institute of Design, User-centered Web design.*
- Ahu Sieg [proposal passed 2005, Depaul], *Contextual Information Access with Ontologisms: Integrating Domain Knowledge* and User Profiles (School of Computer Science, Telecommunications and Information Systems).
- 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.*

Grants

"Enhancing the CS Undergraduate Curriculum to Include Data Mining and Information Retrieval" $339,500
    National Science Foundation Award # 0119469, Total of 4 participants, Co-PI. (2001-2005)
"CRCD/EI: An Undergraduate CS Specialization in Knowledge Management Systems" $281,999
    National Science Foundation, Oct 2004-Sept 2007, Total of 5 participants, Co-PI. (2004-2007)
"Biomedical Literature Search" $28,763
    IIT Research Institute, September 2005-August 2006, Principle Investigator (PI). (2005-2006)

Refereed Publications

(Publications can be downloaded from IIT IR Lab site (http://ir.iit.edu/publications/publications.html)

Journals

1. N. Goharian, A. Jain, Q. Sun, "Comparative Analysis of Sparse Matrix Algorithms for Information Retrieval", Journal *of Systemics, Cybernetics and Informatic,* Vol 1, No 1, 2003.

2. 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.

3. 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.

4. 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*, 55 (10), July 2004.

5. 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*, 59 (12), September 2008.

6. J. Urbain, O. Frieder, and N. Goharian, "A Dimensional Retrieval Model for Integrating Semantics and Statistical Evidence in Context for Genomics Literature Search," *Computers in Biology and Medicine* 39 (1), January 2009.

7. S. Mengle and N. Goharian, "Passage Detection Using Text Classification", *Journal of American Society for Information Science and Technology*, 60 (4), March 2009.

8. J. Urbain, O. Frieder, N. Goharian, "Passage relevance models for genomics search", *BMC Bioinformatics*, 10(Suppl 3):S3 (19 March 2009).

9. S. Mengle and N. Goharian, "Ambiguity Measure Feature Selection Algorithm", *Journal of American Society for Information Science and Technology*, 60 (5), April 2009.

Conferences
1998-2001

10. S. Alaoui-Mounir, N. Goharian, M. Mahoney, A. Salem, O. Frieder, "Fusion of Information Retrieval Engines (FIRE)", *International Conference on Parallel and Distributed Processing Techniques and Applications* , 1998.

11. N. Goharian, T. El-Ghazawi, D. Grossman, "On Implementation of Information Retrieval as Sparse Matrix

    Application", *International Conference on Parallel and Distributed Processing Techniques and Applications (PDPTA 99),* Las Vegas, Nevada, 1999.

12. N. Goharian, T. El-Ghazawi, D. Grossman, A. Chowdhury, "On the Enhancements of a Sparse Matrix Information Retrieval Approach", *International Conference on Parallel and Distributed Processing Techniques and Applications (PDPTA 2000),* Las Vegas, Nevada, 2000.

13. N. Goharian, T. El-Ghazawi, D. Grossman, "Enterprise Text Processing: A Sparse Matrix Approach", *IEEE International Conference on Information Technology: Coding & Computing,* Las Vegas, Nevada 2001.

2002

14. 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.

15. A. Jain and N. Goharian, "On Parallel Implementation of Sparse Matrix Information Retrieval Engine", *The 2002 International Multi-conferences in Computer Science: on Information and Knowledge Engineering (IKE'02),* Las Vegas, Nevada, June 2002.

16. 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.

17. E. Jensen, S. Beitzel, A. Pilotto, N. Goharian, O. Frieder "Parallelizing the Buckshot Algorithm for Efficient Document Clustering", *ACM 11th Conference on Information and Knowledge Management (CIKM)*, 2002.

2003

18. 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 18th Symposium on Applied Computing (SAC)*, March 2003.

19. S. Stein, N. Goharian, ""On the Mapping of Index Compression Techniques on CSR Information Retrieval", *IEEE International Conference on Information Technology: Coding & Computing (ITCC 2003),* April 2003.

20. L. Ma, N. Goharian, A. Chowdhury, M. Chung, "Extracting Unstructured Data From Template Generated Web Documents", *ACM 12th Conference on Information and Knowledge Management (CIKM)*, November 2003.

21. R. Cathey, L. Ma, N. Goharian, D. Grossman, "Misuse Detection for Information Retrieval Systems", *ACM 12th Conference on Information and Knowledge Management (CIKM)*, November 2003.

2004

22. 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 2004),* Las Vegas, Nevada, April 2004.

23. N. Goharian, L. Ma "Using Relevance Feedback to Detect Misuse in Information Retrieval Systems" *ACM 13th Conference on Information and Knowledge Management (CIKM)*, November 2004.

2005

24. N. Goharian, L. Ma "Query Length Impact on Misuse Detection in Information Retrieval Systems", *ACM 20th Symposium on Applied Computing (SAC)*, March 2005.

25. 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 2005),* April 2005.

26. N. Goharian, L. Ma, C. Meyer, "Detecting Misuse of Information Retrieval Systems Using Data Mining Techniques" *IEEE 3rd International Conference on Intelligence and Security Informatics (ISI),* May 2005.

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

2006

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

29. N. Goharian, A. Platt, O. Frieder "On Off-Topic Web Browsing" *IEEE 4ᵗʰ International Conference on Intelligence and Security Informatics* (ISI*),* May 2006.

2007

30. A. Platt, N. Goharian, S. Mengle "Using User Query Sequence to Detect Off-Topic Search", *ACM 22ⁿᵈ Symposium on Applied Computing (SAC)*, March 2007.

31. N. Goharian and A. Platt "DOTS: Detection of Off-Topic Search Via Result Clustering" *IEEE 5ᵗʰ International Conference on Intelligence and Security Informatics* (ISI), May 2007.

32. A. Platt and N. Goharian "Short Query Sequences in Misuse Detection" *IEEE 5ᵗʰ International Conference on Intelligence and Security Informatics* (ISI), May 2007.

33. S. Mengle, N. Goharian, A. Platt "FACT: Fast Algorithm for Categorizing Text" *IEEE 5ᵗʰ International Conference on Intelligence and Security Informatics (ISI)*, May 2007.

34. 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.

2008

35. S. Mengle, N. Goharian, "Using Ambiguity Measure Feature Selection Algorithm for Support Vector Machine Classifier", *ACM 23ʳᵈ Symposium on Applied Computing (SAC)*, March 2008.

36. S. Mengle, N. Goharian, A. Platt "Discovering Relationships among Categories using Misclassification Information", *ACM 23ʳᵈ Symposium on Applied Computing (SAC)*, March 2008.

37. S. Mengle, N. Goharian, "Detecting Hidden Passages from Documents", *In proceedings of SIAM Conference on Data Mining (SDM 2008) Workshop*, April 2008.

38. N. Goharian, S. Mengle, "On Document Splitting in Passage Detection", *In proceedings of 31st Conference on Research and Development in Information Retrieval (ACM SIGIR),* July 2008.

39. 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.

2009

40. A. Platt, S. Mengle, N. Goharian, "Improving Classification Based Off-topic Search Detection via Category Relationships", *ACM 24ᵗʰ Symposium on Applied Computing (SAC)*, March 2009.

41. 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 24ᵗʰ Symposium on Applied Computing (SAC)*, March 2009.

TREC

42. J. Urbain, N. Goharian, O. Frieder, "IIT TREC-2005: Genomics Track", *Proceedings of the Fourteenth Text REtrieval Conference*, November 2005.

43. J. Urbain, N. Goharian, O. Frieder, "IIT TREC-2006: Genomics Track", *Proceedings of the Fifth Text REtrieval Conference*, November 2006.

44. J. Urbain, N. Goharian, O. Frieder, "IIT TREC-2007 Genomics Track: Using Concept-based Semantics in Context for Genomics Literature Passage Retrieval", *Proceedings of the Fourteenth Text REtrieval Conference*,November 2007.

Book Chapter

45. 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.