

# FERNANDO R. FARFÁN

E-mail: ffarfan@umich.edu

Tel.: (734) 335-0744

Database Systems • Data Management Systems • Information Retrieval • Web Applications

Domain-specific Information Retrieval • Storage Systems • Software Engineering

Bioinformatics • Electronic Health Records • Software Design and Development

---

## EDUCATION

---

### Florida International University

Ph.D. in Computer Science

Miami, Florida

December 2009

- *Overall Outstanding Graduate Student FIU-School of Computing and Information Sciences, 2009*

### Florida International University

M.S. in Computer Science

Miami, Florida

December 2005

- *Sponsored by Fulbright Grant*

### Universidad del Valle de Guatemala

B.S. in Computer Science

Guatemala City, Guatemala

August 2001

- *Cum Laude*
- *Honors Thesis "Analysis, Design and Development of a Learning Integrator System"*

---

## PROFESSIONAL EXPERIENCE

---

### University of Michigan

*Postdoctoral Research Fellow*

Ann Arbor, Michigan

Jan 2010 – Present

Integration of known genetic data, as found in databases, to perform a more effective analysis of gene expression data from microarray experiments and next generation sequencing data. Our current tools and techniques provide scientists with new, refined biological hypotheses based on prior knowledge, optimizing the time previously required to derive these. I manage and mentor undergraduate and graduate students, and lead the development of systems and tools to be used by our team of biologists and bioinformaticians.

### Florida International University

*Research Assistant*

Miami, Florida

Jan 2006 – Dec 2009

Research in databases and storage systems, XML parsing and processing, domain-specific Information Retrieval on semi-structured and structured documents. Developed Information Retrieval systems for large corpuses of documents of various domains (e.g. electronic medical records, patents) with an emphasis on graph-structured, large datasets.

### Banco De Desarrollo Rural BANRURAL

*Senior Analyst*

Guatemala City, Guatemala

Jan 2000 – Jul 2004

Team leader for the in-house software development division of the bank. I managed a team of four developers in charge of several banking modules, in particular the Internet banking platform, online payment reception for utilities, and online tax payment system. Our online tax payment system was the first one to go live in the country, even though our team had limited human resources and started developing several months after all other competing banks started. This system was also chosen as the preferred system by the National Tax Office in Guatemala, due to its reliability and easy-to-use Web interface.

### Grupo Polaris

*Junior Programmer*

Guatemala City, Guatemala

May 1998 – Dec 1999

Developed n-tier applications using Microsoft Visual Basic and Microsoft SQL Server. Developing credit card printing software that reduced the printing process from eight hours to 20 minutes.

---

## PUBLICATIONS

---

### Conferences and Workshops

- F. Farfán, J. Ma, M.A. Sartor, G. Michailidis and H.V. Jagadish. **THINK-Back: KNowledge-based Interpretation of High Throughput data.** *BMC Bioinformatics* 2012, 13(Supl 2):S4.
- V. Hristidis, E. Ruiz, A. Hernandez, F. Farfán and R. Varadarajan. **PatentsSearcher: A Novel Portal to Search and Explore Patents.** *3<sup>rd</sup> International Workshop on Patent Information Retrieval, PaIR 2010.*
- F. Farfán, V. Hristidis, A. Ranganathan, and M. Weiner. **XOntoRank: Ontology-Aware Search of Electronic Medical Records.** *25<sup>th</sup> IEEE International Conference on Data Engineering, ICDE 2009.*
- F. Farfán, V. Hristidis, A. Ranganathan, and R. P. Burke. **Ontology-Aware Search on XML-based Electronic Medical Records.** *24<sup>th</sup> IEEE International Conference on Data Engineering, ICDE 2008*
- V. Hristidis, F. Farfán, R. P. Burke, A. F. Rossi, and J. A. White. **Information Discovery on Electronic Medical Records.** *National Science Foundation Symposium on Next Generation of Data Mining and Cyber-Enabled Discovery for Innovation, NGDM 2007.*
- F. Farfán, V. Hristidis, and R. Rangaswami. **Beyond Lazy XML Parsing.** *18<sup>th</sup> International Conference on Database and Expert System Applications, DEXA 2007.*
- M. Bhadkamkar, F. Farfán, V. Hristidis, and R. Rangaswami. **Storing Trees on Disk Drives.** *3<sup>rd</sup> USENIX Conference on File and Storage Technologies, FAST Work In Progress 2005 (Short Paper).*

### Journal Papers

- M. Bhadkamkar, F. Farfán, V. Hristidis, and R. Rangaswami. **Storing Semi-structured Data on Disk Drives.** *ACM Transactions on Storage* Vol. 5, Issue # 2, May 2009.
- F. Farfán, V. Hristidis, and R. Rangaswami. **2LP: A double-lazy XML parser.** *Elsevier Information Systems* 34 (2009), pp. 145-163.

### Book Chapters

- F. Farfán, V. Hristidis. **Overview of XML.** *Information Discovery on Electronic Health Records.* CRC - Taylor & Francis. Editor: V. Hristidis. 2010.
- F. Farfán, R. Varadarajan, V. Hristidis. **Electronic Health Records.** *Information Discovery on Electronic Health Records.* CRC - Taylor & Francis. Editor: V. Hristidis. 2010.
- V. Hristidis, R. Varadarajan, F. Farfán. **Searching Electronic Health Records.** *Information Discovery on Electronic Health Records.* CRC - Taylor & Francis. Editor: V. Hristidis. 2010.
- V. Hristidis, F. Farfán, R. P. Burke, A. F. Rossi, J. A. White. **Challenges for Information Discovery on Electronic Medical Records.** *Next Generation of Data Mining.* CRC - Taylor & Francis. Editors: H. Kargupta, J. Han, P. Yu, R. Motwani, V. Kumar. 2008.

### Technical Reports

- M. Bhadkamkar, F. Farfán, V. Hristidis, and R. Rangaswami. **Efficient Native Storage Systems for Semistructured Data.** *Florida International University Technical Report FIU-SCIS-TR-2006-09-01.* September 2006.

---

## PRESENTATIONS

---

- **Knowledge-based Interpretation of High Throughput Data.**  
Tutorial at the Great Lakes Bioinformatics Conference GLBIO 2012. Ann Arbor, Michigan.
- **THINK-Back: KNowledge-based Interpretation of High Throughput data.**  
Presentation at the Great Lakes Bioinformatics Conference GLBIO 2011. Athens, Ohio.
- **XOntoRank: Ontology-Aware Search of Electronic Medical Records.**  
Presentation at the 25<sup>th</sup> International Conference on Data Engineering, ICDE 2009. Shanghai, China.
- **XOntoRank: Ontology-Aware Search of Electronic Medical Records.**  
Student Lecture Series at the School of Computing and Information Sciences, Florida International University, March 2009. Miami, Florida.

- **Ontology-Aware Search on XML-based Electronic Medical Records.**  
Poster Session at the 24<sup>th</sup> International Conference on Data Engineering, ICDE 2008. Cancún, México.
- **Beyond Lazy XML Parsing.**  
Presentation at the 18<sup>th</sup> International Conference on Databases and Expert Systems, DEXA 2007.  
Regensburg, Germany.

---

## SYSTEMS DEVELOPED AND DEMONSTRATIONS

---

- **THINK Back (<http://www.eecs.umich.edu/db/think/>)**  
The KNowledge-based Interpretation of High Throughput data project looks back at prior biological knowledge to provide novel interpretations to new systems biology experiments. These tools are currently being deployed as web services to be used by biologists and bioinformaticians worldwide.
- **PatentsSearcher (<http://www.patentssearcher.com/>)**  
The *PatentsSearcher* search engine applies cutting edge research to efficiently search and rank a corpus of U.S. and international patent documents.
- **XOntoRank**  
The *XOntoRank* system addresses the problem of ontology-aware keyword search of XML documents with a particular focus on XML-based Electronic Medical Records.
- **Double-Lazy XML Parser**  
The 2LP framework for efficient XML parsing places internal physical pointers within the XML documents to allow the navigation process to skip large portions of the document during parsing.
- **Semi-structured Storage Systems**  
The SSS system is a semi-structured storage simulator to study the feasibility of novel in-disk placement layouts that exploit the physical features of the hard disks to improve access time.

---

## HONORS AND AWARDS

---

- **Overall Outstanding Graduate Student 2009.**  
FIU School of Computing and Information Sciences, December 2009.
- **Dissertation Year Fellowship 2009.**  
FIU University Graduate School
- **FIU Graduate Engineering Tuition Scholarship, 2005.**  
College of Engineering and Computing, FIU.
- **Fulbright Scholar, 2004-2005.**  
Department of State of the United States of America.
- **BARSA Scholarship, 2004.**  
Institute of International Education.
- **Distinguished Student, 1996, 1998-2000.**  
Universidad del Valle de Guatemala

---

## TECHNOLOGY SUMMARY

---

**Programming Languages:** Java • JSP • PHP • Visual Basic • ASP • C/C++ • C# • Groovy

**Relational Database Management Systems:** Microsoft SQL Server • Sybase • Oracle • MySQL

**Operating Systems:** Windows • Linux • Unix System V Release 4 • Solaris

**Other technologies:** XML • XPath • XQuery • Lucene • Google Web Toolkit • Google AJAX Search API • HTML • CSS • Smooks • Hibernate

**Software:** MS Word • MS Excel • MS Visio • MS Project • Image editors • LaTeX • Adobe CS3 Suite

---

## LANGUAGES

---

- Spanish (Native)
- English (Fluent)

## REFERENCES

---

▪ **Dr. H. V. Jagadish**

*Bernard A Galler Collegiate Professor of Electrical Engineering and Computer Science.*

*University of Michigan*

*2260 Hayward St Office: 4601 CSE Building*

*Ann Arbor, MI 48109*

*+1(734) 763-4079*

*jag@eecs.umich.edu*

▪ **Dr. Vagelis Hristidis (Major Advisor)**

*Associate Professor*

*Department of Computer Science and Engineering*

*University of California, Riverside*

*900 University Avenue*

*Riverside, CA 92521*

*+1(951) 827-1012*

*vagelis@cs.ucr.edu*

▪ **Dr. Anand Ranganathan**

*Research Staff Member*

*IBM T.J. Watson Research Center*

*1101 Kitchawan Road, Route 134,*

*Yorktown Heights, N.Y. 10598*

*+1(914)784-6257*

*arangana@us.ibm.com*

▪ **Dr. Raju Rangaswami**

*Associate Professor*

*School of Computing and Information Sciences*

*Florida International University*

*11200 SW 8<sup>th</sup> St. ECS-386*

*Miami, FL 33199*

*+1(305)348-6230*

*raju@cis.fiu.edu*