

# Michalis Petropoulos

Computer Science & Engineering Department  
University at Buffalo, SUNY  
201 Bell Hall  
Box 602000  
Buffalo, NY 14260-2000

mpetro@buffalo.edu  
<http://www.cse.buffalo.edu/~mpetro>  
Phone: (716) 645-4733  
Fax: (716) 645-3464

---

## RESEARCH INTERESTS

---

Databases: user-oriented and agent-oriented web-based interfaces; semistructured and XML data; publishing and integration scenarios; rewriting queries using limited interfaces; schema integration.

---

## EDUCATION

---

<b>Ph.D., Computer Science</b> <i>University of California, San Diego</i> Thesis: Exporting and Utilizing Database Interfaces on the Web Advisor: Yannis Papakonstantinou	Dec 2005
<b>M.S., Computer Science</b> <i>University of California, San Diego</i>	Jun 2000
<b>Diploma, Computer Engineering</b> <i>Technical University of Crete, Chania, Greece</i>	Aug 1998

---

## WORK EXPERIENCE

---

<b>University at Buffalo, SUNY, Buffalo, NY</b> <i>Assistant Professor, Computer Science &amp; Engineering Department</i>	Jan 2006 – present
<b>app2you, Inc., La Jolla, CA</b> <i>Consultant</i> app2you.com is a hosted web service enabling users to sketch the pages of their database-driven application and the flow of information across pages. Contributed to the design and implementation of algorithms that automatically infer the database design and logic behind the web pages.	Jan – Aug 2007
<b>University of California, San Diego, La Jolla, CA</b> <i>Research Assistant, Computer Science &amp; Engineering Department</i> Under supervision of <i>Prof. Yannis Papakonstantinou</i>	Sep 1999 – Dec 2005
<b>University of California, San Diego, La Jolla, CA</b> <i>Consultant, Computer Science &amp; Engineering Department</i> Under supervision of <i>Prof. Alex Orailoglu</i> and <i>Prof. Yannis Papakonstantinou</i> Designed and implemented a web-based review system for the Graduate Admissions Committee. The project involves the integration of three relational databases created from three independent campus entities, implements a reviewing workflow for the committee members and applicants, and offers administration functionality to committee chairs and staff members.	Jan – Dec 2004
<b>IBM Almaden Research Center, San Jose, CA</b> <i>Research Intern, Advanced Database Services Group</i> Under supervision of <i>Dr. Jussi Myllymaki</i> Implemented two instance-based schema matching heuristics. Contributed to the development of a library that implements various graph algorithms that the proposed heuristics depend on.	Jun - Sept 2004
<b>Microsoft Research, Redmond, WA</b> <i>Research Intern, Database Group</i> Under supervision of <i>Dr. Phil Bernstein</i> Designed and implemented Protoplasm: an industrial-strength schema matching platform that focuses on extensibility and customizability. Implemented several linguistic and structural matching operators.	Jul – Nov 2003

**Bell Laboratories, Murray Hill, NJ**

Jun - Sep 2002

*Research Intern*, Network Data and Services Research Department  
Under supervision of *Dr. Juliana Freire* and *Dr. Prasan Roy*

Worked on relaxation techniques for queries against semi-structured data based on schema statistics.

**Enosys Software Inc., San Diego, CA**

Jan 2000 – Sep 2001

*Senior Software Engineer*

Under supervision of *Clifton McLellan (Director of Engineering)*

Led a team of engineers that designed and implemented the Application Builder component of the *Enosys Design Suite*. Application Builder is a tool for rapidly developing custom web interfaces for data integration applications, and features easy-to-use forms, summarization and navigation tables, and templates to assist in query and interface building.

**San Diego Supercomputer Center, San Diego, CA**

Jun 1999 - Sep 1999

*Research Intern*

Under supervision of *Dr. Chaitan Baru*

Designed and implemented heuristics for automatic specification of XML Schemas out of relational schemas for the purpose of visual display. The heuristics converted relational schema graphs into XML Schema trees in a way the user conceives as clearly as possible the content and semantics of the relational source. This work was part of the Mediation of Information using XML (MIX) project.

**Multimedia Systems Center AE, Chania, Greece**

Sep 1995 – Aug 1998

*Software Engineer*

Managed a desktop publishing team, which printed a 100-page tourist guide titled "The Atlas of Crete" using software for map creation, image processing and page layout. The tourist guide was circulated in 20,000 copies and six languages.

---

**TEACHING EXPERIENCE**

---

**Instructor – University at Buffalo, SUNY**

- Data Integration (CSE 636) – Fall 08, Fall 06, Spring 06
- Database Systems (CSE562) – Spring 10, Spring 09, Fall 07
- Web Data Engineering (CSE 510) – Fall 09
- Database Seminar (CSE718) – Spring 10, Fall 08, Spring 08, Spring 07

---

**PATENTS**

---

**1. Flexible Format Query Processing System for Web Browsers**

Michalis Petropoulos, Vasilis Vassalos, Yannis Papakonstantinou  
United States Patent Number 7,203,678 (April 10, 2007)  
Assigned to BEA Systems, Inc.

---

**HONORS & AWARDS**

---

- **Best Paper Award Honorable Mention**, ACM International Conference on Management of Data (SIGMOD) for "Interactive Query Formulation over Web Service-Accessed Sources", 2006
- **Best Interdisciplinary Paper Award**, ACM International Conference on Information and Knowledge Management (CIKM) for "FACeTOR: Cost-Driven Exploration of Faceted Query Results", 2010

---

**GRANT SUPPORT**

---

- **ISTARE: Intelligent Spatio-Temporal Activity Reasoning Engine**  
Jason J. Corso (PI), Raymond Fu, Michalis Petropoulos, Venkat Krovie, Werner Ceusters (co-PIs)  
Defense Advanced Research Projects Agency (DARPA)  
6/22/2010-6/21/2013: **\$2,200,000**
- **Extending a Preservation Archive through a Social Network**  
Michalis Petropoulos (PI), Jeff Good (co-PI)

---

UB 2020 Interdisciplinary Research Development Fund (IRDF)  
12/1/2008-12/1/2009: **\$24,000**

---

SOFTWARE

---

- **CLIDE:** <http://www.clide.info/>  
Interactively formulating feasible queries over data services using query rewriting-based systems
  - **BioNav:** <http://db.cse.buffalo.edu/bionav/>  
Effective navigation on query results of biomedical databases
  - **FACeTOR:** <http://facetor.com/>  
Cost-driven exploration of faceted query results
  - **QURSED:** <http://db.cse.buffalo.edu/qursed/> (under development)  
Dynamic and declarative query forms and reports for XML data
- 

PUBLICATIONS

---

- |  |  |
|--|--|
| 1. <b>FACeTOR: Cost-Driven Exploration of Faceted Query Results</b><br>Abhijith Kashyap, Vagelis Hristidis, Michalis Petropoulos<br>ACM Conference on Information and Knowledge Management (CIKM), 2010  | Conference<br>Article<br>127:945 (13%)   |
| 2. <b>Effective Navigation of Query Results Based on Concept Hierarchies</b><br>Abhijith Kashyap, Vagelis Hristidis, Michalis Petropoulos, Sotiria Tavoulari<br>IEEE Transactions on Knowledge and Data Engineering (TKDE), to appear            | Journal<br>Article                       |
| 3. <b>A Data-Centric Approach to Insider Attack Detection in Database Systems</b><br>Sunu Mathew, Michalis Petropoulos, Hung Q. Ngo, Shambhu Upadhyaya<br>International Symposium on Recent Advances in Intrusion Detection (RAID), 2010         | Symposium<br>Article                     |
| 4. <b>BioNav: Effective Navigation on Query Results of Biomedical Databases</b><br>Abhijith Kashyap, Vagelis Hristidis, Michalis Petropoulos, Sotiria Tavoulari<br>IEEE International Conference on Data Engineering (ICDE), 2009                | Conference<br>Article<br>150:554 (27%)   |
| 5. <b>Exploring Biomedical Databases with BioNav</b><br>Abhijith Kashyap, Vagelis Hristidis, Michalis Petropoulos, Sotiria Tavoulari<br>ACM International Conference on Management of Data (SIGMOD), 2009  | Demonstration<br>Article<br>32:86 (37%)  |
| 6. <b>Do-It-Yourself Custom Forms-Driven Workflow Applications</b><br>Keith Kowalczykowski, Kian Win Ong, Alin Deutsch, Yannis Papakonstantinou, Kevin Zhao, Michalis Petropoulos<br>Conference on Innovative Data Systems Research (CIDR), 2009 | Conference<br>Article                    |
| 7. <b>Exporting and Interactively Querying Web Service-Accessed Sources: The CLIDE System</b><br>Michalis Petropoulos, Alin Deutsch, Yannis Papakonstantinou, Yannis Katsis<br>ACM Transactions on Database Systems (TODS), 32(4), 2007          | Journal<br>Article                       |
| 8. <b>CLIDE: Interactive Query Formulation for Service Oriented Architectures</b><br>Michalis Petropoulos, Alin Deutsch, Yannis Papakonstantinou<br>ACM International Conference on Management of Data (SIGMOD), 2007                            | Demonstration<br>Article<br>36:107 (34%) |
| 9. <b>Interactive Query Formulation over Web Service-Accessed Sources</b><br>Michalis Petropoulos, Alin Deutsch, Yannis Papakonstantinou<br>ACM International Conference on Management of Data (SIGMOD), 2006                                    | Conference<br>Article<br>58:446 (13%)    |
| 10. <b>Graphical Query Interfaces for Semistructured Data: The QURSED System</b><br>Michalis Petropoulos, Yannis Papakonstantinou, Vasilis Vassalos<br>ACM Transactions on Internet Technology (TOIT), 5(2), 2005                                | Journal<br>Article                       |
| 11. <b>Industrial-Strength Schema Matching</b><br>Philip A. Bernstein, Sergey Melnik, Michalis Petropoulos, Christoph Quix<br>ACM SIGMOD Record, 33(4), 2004   | Journal<br>Article                       |

- |  |  |
|--|--|
| <p>12. <b>Query Set Specification Language (QSSL)</b><br/>         Michalis Petropoulos, Alin Deutsch, Yannis Papakonstantinou<br/>         Sixth International Workshop on the Web and Databases (WebDB), 2003</p>                                  | <p>Workshop<br/>         Article<br/>         17:67 (25%)</p>    |
| <p>13. <b>Online Querying of Heterogeneous Distributed Spatial Data on a Grid</b><br/>         Ilya Zaslavsky, Ashraf Memon, Michalis Petropoulos, Chaitan Baru<br/>         Third International Symposium on Digital Earth, 2003</p>                | <p>Symposium<br/>         Article</p>                            |
| <p>14. <b>Generating Query Forms and Reports for Semistructured Data: The QURSED Editor</b><br/>         Yannis Papakonstantinou, Michalis Petropoulos, Vasilis Vassalos<br/>         Ninth Panhellenic Conference on Informatics (PCI), 2003</p>    | <p>Conference<br/>         Article</p>                           |
| <p>15. <b>QURSED: Querying and Reporting Semistructured Data</b><br/>         Yannis Papakonstantinou, Michalis Petropoulos, Vasilis Vassalos<br/>         ACM International Conference on Management of Data (SIGMOD), 2002</p>                     | <p>Conference<br/>         Article<br/>         42:239 (17%)</p> |
| <p>16. <b>Semantic Caching of XML Databases</b><br/>         Vagelis Hristidis, Michalis Petropoulos<br/>         Fifth International Workshop on the Web and Databases (WebDB), 2002</p>  | <p>Workshop<br/>         Article<br/>         17:68 (25%)</p>    |
| <p>17. <b>Building XML Query Forms and Reports with XQForms</b><br/>         Michalis Petropoulos, Yannis Papakonstantinou, Vasilis Vassalos<br/>         Computer Networks Journal, Special Issue on XML, Elsevier Science, 39(5), 2002</p>         | <p>Journal<br/>         Article</p>                              |
| <p>18. <b>XML Query Forms (XQForms): Declarative Specification of XML Query Interfaces</b><br/>         Michalis Petropoulos, Vasilis Vassalos, Yannis Papakonstantinou<br/>         Tenth International World Wide Web Conference (WWW10), 2001</p> | <p>Conference<br/>         Article<br/>         78:391 (20%)</p> |

---

PROFESSIONAL ACTIVITIES

---

*NSF Panel Review Member*

- 2007, 2008

*Workshop Co-Chair (with Alexandros Labrinidis):*

- WebDB 2009 - International Workshop on the Web and Databases (SIGMOD Workshop)  
<http://webdb09.cse.buffalo.edu/>

*Tutorials Chair:*

- ICWE 2007 - International Conference on Web Engineering

*Local Arrangements Chair:*

- SIGMOD 2003 - ACM International Conference on Management of Data

*Program Committee Member:*

- ICDE 2011 - International Conference on Data Engineering
- WWW 2010 - International World Wide Web Conference
- EDBT 2010 - International Conference on Extending Database Technology
- NTII 2010 - International Workshop on New Trends in Information Integration (ICDE Workshop)
- WebDB 2008, 2006 - International Workshop on the Web and Databases (SIGMOD Workshop)
- EDBT 2008 Demonstrations - International Conference on Extending Database Technology
- ICWE 2007 - International Conference on Web Engineering
- ISDIKS 2007 - International Symposium on Data, Information & Knowledge Spectrum
- CIKM 2006 - Conference on Information and Knowledge Management
- WWW 2002 *Posters* - International World Wide Web Conference

*Reviewer:*

- ACM Computing Surveys
- IEEE Transactions on Knowledge and Data Engineering (TKDE)
- Elsevier Journal of Data & Knowledge Engineering (DKE)

*Member:*

- Association of Computing Machinery (ACM) – since 2002

- ACM Special Interest Group on Management of Data (SIGMOD) – since 2002
- W3C XForms Working Group (2000)

---

## RESEARCH SUPERVISION

---

### *Academic Advisees:*

- Abhijith Kashyap – PhD (since 09/2006)
- Daniel Schlegel – PhD (since 09/2009)

### *PhD Committees:*

- Denis Mindolin (advisor: Jan Chomicki)
- Xi Zhang (advisor: Jan Chomicki)
- Slawomir Staworko (advisor: Jan Chomicki)

---

## PRESENTATIONS

---

- **Interactive Query Formulation over Web Service-Accessed Sources**  
University of Pennsylvania, Computer Science Dept., Philadelphia PA, 2008  
ACM SIGMOD Conference, Chicago IL, 2006  
Athens University of Economics and Business, Computer Science Dept., Athens, Greece, 2006  
Foundation for Research and Technology – Hellas (FORTH), Crete Greece, 2006
- **Exporting and Utilizing Database Interfaces on the Web**  
University at Buffalo, SUNY, Department of Computer Science and Engineering, 2005  
University of Illinois at Urbana-Champaign, Department of Computer Science, 2005  
Purdue University, Department of Computer Science, 2005  
Texas A&M University, Department of Computer Science and Engineering, 2005  
University of Houston, Department of Computer Science, 2005
- **Web Services Interfaces**  
Microsoft Research, Database Group, Redmond WA, 2003  
Gemo Group (INRIA) - UCSD Database Group Workshop (GEMSAD), San Diego CA, 2003
- **Query Set Specification Language (QSSL)**  
International Workshop on the Web and Databases (WebDB), San Diego CA, 2003
- **QURSED: QUerying and Reporting SEMistructured Data**  
ACM SIGMOD Conference, Madison WI, 2002  
University of Pennsylvania, Department of Computer and Information Science, 2002
- **XML Query Forms (XQForms): Declarative Specification of XML Query Interfaces**  
Tenth International World Wide Web Conference (WWW10), Hong Kong S.A.R., 2001
- **Relational Wrapper for Navigation-Driven Lazy Mediator**  
San Diego Supercomputer Center, La Jolla CA, 1999