Anastasios Arvanitis

E-mail: tasosarvanitis@gmail.com   **URL:** http://www.dblab.ntua.gr/~tasosarvanitis/

LinkedIn: http://gr.linkedin.com/in/tasosarvanitis

Education

Jan 2013

PhD in Computer Science

National Technical University of Athens (NTUA), Electrical and Computer Engineering, Greece

Dissertation Title: "*Personalized Data Management Systems*",

Advisors: Prof. Yannis Vassiliou and Prof. Timos Sellis

DiplEng (MEng) in Electrical and Computer Engineering – Major: Computer Science   **Jul 2005**

National Technical University of Athens (NTUA), Electrical and Computer Engineering, Greece

Diploma Thesis: "*System and Application Integration using Web Services*"

Thesis Supervisor: Prof. Timos Sellis

Work Experience

Oracle Corporation, Database Manageability Group, Redwood Shores, CA, USA   **Oct 2014 – present**

Senior Member of Technical Staff (C, PL/SQL, Oracle RDBMS)

- ✓ Responsible for the development of new features and enhancements of SQL Tuning Advisor and SQL Tuning Sets components in Oracle RDBMS 12.2.
- ✓ Worked on the integration of SQL Tuning with the Optimizer Statistics Advisor and the Reporting Framework as well as on enabling import and tuning of SQL statements from remote and pluggable databases (PDBs).

Vanderbilt University Medical Center, Vanderbilt Institute for Clinical and Translational Research (VICTR), Nashville, TN, USA

Health Systems Software Engineer II (Java, Perl, Python, IBM Netezza)   **Sep 2013 – Sep 2014**

- ✓ Worked as a member of the team that maintains the data infrastructure of the Research Derivative (RD) and Synthetic Derivative (SD) data warehouses of the Vanderbilt University and Medical Center. RD and SD data warehouses maintain over 200M medical records for about 2.2M patients, with more than 100,000 new records being added every day.
- ✓ Responsible for enhancing, migrating and maintaining the unstructured data pipeline (ETL) for RD.
- ✓ Reengineered MEDEX Natural Language Processing (NLP) tool as a distributed application that can now handle millions of medical records of VUMC. Integrated MEDEX with the daily data loading pipeline.
- ✓ Designed and developed a data cleaning and normalization workflow for medication annotations found in medical text. Implemented a suite of tools that discover relationships between medication-related terms (e.g. brand names, generics, ingredients) and retrieve associated drug names and indications.
- ✓ Built a tool to parse problem list medical records, extract medications, allergies, procedures and other medical terms and perform data analysis tasks on the output.
- ✓ Implemented various NLP tools including a section tagger to extract allergic reactions and family history from medical records, as well as specialized parsers for labs, vital signs and echocardiograms.

University of California, Riverside, Department of Computer Science and Engineering, Riverside, CA, USA

Research Scholar/Postdoctoral Researcher - UCR Database Lab   **Jun 2012 – Aug 2013**

- ✓ Worked on keyword search and document similarity queries on large biomedical databases (e.g. electronic medical records) using rich domain ontologies **(implemented in Java, MySQL) [EDBT'14]**
- ✓ Developed efficient and effective algorithms for diversified retrieval of microblogging posts **[EDBT'14]**
- ✓ Studied the freshness factor on a class of web search queries with implicit temporal characteristics. Proposed ranking methods for improving the relevance for such queries. **[CIKM'13]**

National Technical University of Athens (NTUA), School of Electrical and Computer Engineering, Athens, Greece

Graduate Research Assistant/PhD Student - Knowledge & Database Systems Lab   **May 2007 – May 2012**

- ✓ Designed and implemented the PrefDB prototype system. PrefDB acts as a personalization layer on top of a RDBMS and facilitates engineering and execution of preference-aware applications (e.g. applications that adapt or filter the query output based on collected user preferences). PrefDB is built on top of PostgreSQL and includes a profile manager, a query parser, optimizer and execution engine as

well as an administration/monitoring tool. **(implemented in Java, PL/pgSQL) [ICDE'12, TKDE'13]**
- ✓ Worked on developing PrefDBAdmin, a graphical tool that helps application designers in the process of building, testing and fine-tuning preferential queries on top of a RDBMS. Designed a profile manager backend and an API for connecting PrefDBAdmin with PrefDB core. **(technologies used: Java, jQuery, CSS, DWR, WebSocket, JSON, PostgreSQL, Tomcat) [demonstrated at SIGMOD'12]**
- ✓ Designed novel algorithms with applications in personalized advertising and product positioning. The proposed algorithms outperform the state-of-the art approaches in terms of efficiency, scalability (scale to millions of products and customers and multiple dimensions) and progressiveness of result calculation. **(implemented in C++ STL, Boost) [CIKM'12]**
- ✓ Developed a probabilistic adaptation of skyline query algorithms for both non-indexed and indexed (using R-tree indexes) data. **(implemented in C++ STL, Boost) [ICDE'10]**.

Institute for the Management of Information Systems, "Athena" Research and Innovation Center, Athens, Greece

Graduate Research Assistant   **Dec 2008 – Jun 2010**
- ✓ Worked on the EU-funded research project TALOS **(Task-Aware Location-Based Services)**. Contributed on (a) designing a novel task description model and language, (b) building a web-based task authoring tool and (c) developing the backend application server, web services and database for supporting task modeling, content annotation and task-based content retrieval. **(used Java EE Technologies (JPA/Hibernate, JSP), jQuery, JSON, SQL Server, Tomcat)**

Satways LTD, Athens, Greece

Software Engineer (Java (JPA/Hibernate, JSP, Swing, Apache AXIS), SQL   Jun 2007 - Nov 2008 Server, JBoss, Tomcat)
- ✓ Worked as a subcontractor of Siemens SBT/SES/CCS for Athens C4I (Command, Control, Communications, Computers and Intelligence) project. Responsible for the design, development and deployment of AIS (Automatic Identification System) which is operational at the Greek Coast Guard. The system allows real-time monitoring, identification and visualization of commercial and coast guard vessels and interoperates with ELS/GEOFIS (Crisis Management) and Asset Management systems.
- ✓ Involved in the development of the AutoTrack AVL (Automatic Vehicle Location) system, which allows users to monitor mobile resources in real time, perform GIS operations, get customized reports, etc.

EXUS, Business Process Management Department, Athens, Greece

Software Engineer (Java EE (JSF, Facelets, Struts, Hibernate, JSP), JavaScript,   Jun 2006 - May 2007 ASP, SQL Server, JBoss)
- ✓ Worked on the development of iPerform workflow engine v2.0
- ✓ Involved in migrating iPerform deployments on Piraeus Bank (largest financial/banking institution in Greece) and Wind Hellas Telecommunications S.A. (3rd largest telco in Greece)
- ✓ Responsible for the deployment and testing of iPerform v2.0 at Wind Hellas Telecommunications S.A.

Greek Army General Staff, Athens, Greece

Software Engineer, Network/System Administrator (Jave EE Technologies (EJB   Oct 2005 – Jul 2006 2.1), Oracle RDBMS, BEA Weblogic)–During Military Service and as a contractor
- ✓ Worked on the development and support for the middleware components of the Greek Army C4I

Selected Publications

- ✓ **Anastasios Arvanitis**, Matthew Wiley, Vagelis Hristidis, *Efficient Concept-based Document Ranking,* International Conference on Extending Database Technology (**EDBT'14**), Athens, Greece, March 24-28, 2014
- ✓ Shiwen Cheng, **Anastasios Arvanitis**, Marek Chrobak, Vagelis Hristidis, *Multi-query Diversity in Microblogging Posts,* International Conference on Extending Database Technology (**EDBT'14**), Athens, Greece, March 24-28, 2014 (3 citations)
- ✓ Shiwen Cheng, **Anastasios Arvanitis**, Vagelis Hristidis, *How Fresh Do You Want Your Search Results?,* ACM International Conference on Information and Knowledge Management (**CIKM'13**), San Francisco, California, USA, October 27 - November 1, 2013 (7 citations)
- ✓ **Anastasios Arvanitis**, Georgia Koutrika, *PrefDB: Supporting Preferences as First-Class Citizens in Relational Databases*, IEEE Transactions on Knowledge and Data Engineering (**TKDE**), v.26,no.6 (3 citations)
- ✓ **Anastasios Arvanitis**, Antonios Deligiannakis, Yannis Vassiliou, *Efficient Influence-based Processing of Market Research Queries*, ACM International Conference on Information and Knowledge Management (**CIKM'12**), Maui, Hawaii, USA, October 29 - November 2, 2012 (9 citations)
- ✓ **Anastasios Arvanitis**, Georgia Koutrika, *PrefDB: Bringing Preferences closer to the DBMS*, **ACM SIGMOD'12**, Scottsdale, Arizona, USA, May 20-24, 2012 (demo paper) (3 citations)
- ✓ **Anastasios Arvanitis**, Georgia Koutrika, *Towards Preference-aware Relational Databases*, IEEE International Conference on Data Engineering (**ICDE'12**), Washington DC, USA, April 1-5, 2012 (10 citations)
- ✓ Dimitris Sacharidis, **Anastasios Arvanitis**, Timos Sellis, *Probabilistic Contextual Skylines*, IEEE International Conference on Data Engineering (**ICDE'10**), Long Beach, California, USA, March 1-6, 2010 (14 citations)

Technical Skills

Extensive experience and keen interest in:

- ✓ Database design, engineering and optimization. Worked with all major relational DBMSs including Oracle, Microsoft SQL Server, IBM Netezza, PostgreSQL, MySQL in both industry and academic positions. Experience with various data access technologies including SQL, PL/SQL, PL/pgSQL, JDBC, JPA/Hibernate, Spring JDBC, EJB 2.1 & 3.0
- ✓ Data structures and data-driven algorithmic design with applications in information retrieval, personalization, recommender systems, spatial databases and healthcare informatics.

Proficient or familiar with a wide array of programming languages, concepts and technologies, including:

- ✓ Java Frameworks and Tools (JSP, Servlets, JSF, JBoss Seam, Facelets, Struts, log4j, JAXB, JAXP, Swing, JNDI, JBoss Drools, JMF)
- ✓ Programming Languages: C/C++ (STL, Boost), JavaScript/jQuery, Perl, Python
- ✓ Web Technologies: AJAX, CSS, HTML, XML/XML Schema, JSON, WebSocket
- ✓ Middleware (JBoss, Oracle Weblogic, Apache Tomcat), Service-Oriented Architectures (Apache AXIS framework, SOAP and REST protocols)
- ✓ Unix/Linux network administration basics, shell scripting, regular expressions
- ✓ Software Development Processes and Tools: Test-driven development, Continuous integration, jUnit, Maven, Ant, Git, SVN, CVS, gdb, Eclipse, NetBeans, Oracle Developer, SQL*Plus, Emacs, issue tracking systems (JIRA, Trac, MantisBT), UML and process modeling with BPMN
- ✓ NLP tools (e.g. SVMlib, Apache cTakes, Apache UIMA, MEDEX, Negex**)**
- ✓ ESRI ArcGIS software stack - Certificate on *"Geographic Information Systems (GIS)"* from the Geography and Spatial Analysis Laboratory, School of Rural and Surveying Engineering, NTUA (2005)