Wisam Dakka

Columbia University   (office) 212-939-7116
Computer Science Department   (mobile) 646-643-1306
1214 Amsterdam Avenue   wisam@cs.columbia.edu
New York, New York, 10027   www.cs.columbia.edu/~wisam

Education

Ph.D. (candidate) Computer Science, May 2008, Columbia University, New York
> **Dissertation Title:** *Faceted Searching and Browsing Over Large Collections of Textual and Text-Annotated Objects*
> **Advisor:** *Luis Gravano*
> **Co-Advisor:** *Panagiotis G. Ipeirotis*

M.Phil. Computer Science, November 2007, Columbia University, New York

M.Sc. Computer Science, May 2004, Columbia University, New York

B.S. Mathematics and Computer Science, July 1998, Ben Gurion University, Israel
> *With distinction*
> *Certificate of recognition for thesis project*

Research Interests

Web and faceted search, Web mining, and information retrieval
Large-scale data analysis
Parallel computing

Professional Employment

September 2002 – Present: *Graduate Research Assistant,* Columbia University, New York, New York, U.S.
> *Time-Faceted Searching and Browsing (Advisor: Luis Gravano)*
> *Extracting Useful Facet Terms (Advisor: Panagiotis G. Ipeirotis)*
> *Multifaceted Browsing Interfaces (Advisor: Panagiotis G. Ipeirotis)*
> *Efficient Summarization-Aware Search (Advisor: Luis Gravano)*

June 2007 – September 2007: *Research Intern,* Machine Learning and Perception Group, Microsoft Research, Cambridge, U.K.
> *Making Sense of Queries to Rank Document Spans (Advisor: Michael J.Taylor)*

June 2006 – September 2006: *Research Intern,* Text Mining Search and Navigation Research, Microsoft Research, Redmond, Washington, U.S.
Augmenting Wikipedia with Named Entity Tags (Advisor: Silviu Cucerzan)

May 2005 – August 2005: *Research Intern,* Dynamic Ranking Group, MSN Search, Redmond, Washington, U.S.
Analysis of Query Logs and Index Reachability (Advisor: Andy Laucius)

August 2001 – September 2002: *Customer Development Manager,* Dalet Digital Media Systems, New York, New York, U.S.
Managed software development for media clients XM Satellite, NPR and others

August 1998 – July 2001: *Senior Software Engineer,* Dalet Digital Media Systems, Beer Sheva, Israel, and Paris, France
Lead developer on a virtual distributed network of user-defined agents

January 1997 – July 1998: *Software Engineer,* Gefen Dekel Technology, Beer Sheva, Israel
Key member of Israeli start-up that built large-scale media management software

Teaching Assistantships

Columbia University: *Database Systems* (Spring 2005)
Ben Gurion University: *Operating Systems* (Spring semesters 1999-2001); *System Programming* (Fall semesters 1999-2001); *Operating Systems* (Fall 1998)

Programming Skills

Fluent in C++ (eight years), C# (two years), Java (five years), Perl, Tcl, JavaScript, PHP, VB, and SQL
Experience in functional programming languages such as Lisp and F#
Experience with large-scale databases and parallel computing systems such as Hadoop, client-server systems, and Web-based applications

Patents Submitted (Microsoft)

On-Site Search Engine for the World Wide We*b*, Microsoft
M. Richardson and W. Dakka , MS DOCKET NO. 317712.01

Document clustering based on entity association rules, Microsoft
S. Cucerzan and W. Dakka, MS DOCKET NO. 318289.01

Working Papers

Processing General Time Sensitive Queries Using Language Models
W. Dakka, L. Gravano, and P. Ipeirotis

Query by Document
D. Y. Yang, N. Bansal, W. Dakka, P. Ipeirotis, N. Koudas, and D. Papadias

Making Sense Of Queries to Rank Document Spans
M. Taylor, W. Dakka, P. Cowans, N. Craswell, and D. Meyerzon

NE-NET: Named Entity Network
S. Maskey and W. Dakka

Papers in Refereed Conferences

Automatic Extraction of Useful Facet Terms from Text Documents (To appear)
W. Dakka and P. Ipeirotis
Proceedings of the 24th IEEE International Conference on Data Engineering
(ICDE 2008), 2008

Augmenting Wikipedia with Named Entity Tags
W. Dakka and S. Cucerzan
Proceedings of the Third International Joint Conference on Natural Language
Processing (IJCNLP 2008), Hyderabad, 2008

Multifaceted Browsing over Large Databases of Text-Annotated Objects
W. Dakka, P. Ipeirotis, and K. Wood
Proceedings of the 23rd IEEE International Conference on Data Engineering,
Demonstrations (ICDE 2007), 2007

Efficient Summarization-Aware Search for Online News Articles
W. Dakka and L. Gravano
Proceedings of the 2007 ACM+IEEE Joint Conference on Digital Libraries (JCDL
2007), 2007

**Comparing American and Palestinian Perceptions of Charisma Using
Acoustic-Prosodic and Lexical Analysis**
F.Biadsy, J. Hirschberg, A. Rosenberg, and W. Dakka
Proceedings of the eighth conference in the annual series of INTERSPEECH, 2007

Automatic Discovery of Useful Facet Terms
W. Dakka, R. Dayal, and P. Ipeirotis

ACM SIGIR 2006 Workshop on Faceted Search, 2006

Automatic Construction of Multifaceted Browsing Interfaces
W. Dakka, P. Ipeirotis, and K. Wood
Proceedings of the 2005 ACM CIKM International Conference on Information
and Knowledge Management , 2005

References

Luis Gravano, *Associate Professor*
Computer Science Department, Columbia University
(office) 212-939-7064
gravano@cs.columbia.edu

Panagiotis G. Ipeirotis, *Assistant Professor*
Leonard N. Stern School of Business, New York University
(office) 212-998-0803
panos@stern.nyu.edu

Silviu-Petru Cucerzan, *Researcher*
Microsoft Research (MSR)
(office) 425-722-7547
silviu@microsoft.com

Julia Hirschberg, *Professor*
Computer Science Department, Columbia University
(office) 212-939-7114
julia@cs.columbia.edu

Michael J. Taylor, *Researcher*
Microsoft Research (MSRC)
(office) 44-1223-479-802
mitaylor@microsoft.com

Michael Codish, *Professor*
Computer Science Department, Ben Gurion University
(office) 972-8-646-1654
mcodish@cs.bgu.ac.il

Michael Elhadad, *Senior Lecturer*
Computer Science Department, Ben Gurion University
(office) 972-8-647-7874
melhadad@dalet.com