

Naveen Ashish

UC-Irvine

4308 Calit2 Building

UC Irvine, Irvine, CA 92697

Tel: 949-824-2991 Email: naveen.ashish@gmail.com Web: <http://www.ics.uci.edu/~ashish>

CURRENT POSITIONS

Assistant Research Scientist (IV) (Professional Researcher)

University of California, Irvine

California Institute for Telecommunications and Information Technology (Calit2)

April 2005 - Present

Primary Activities: Key Investigator on "SAFIRE" project, Key Investigator on "BIRN-CC" project.

Other UCI Affiliations

- Member, Center for Emergency Response Technologies (CERT)
- Member, Center for Machine Learning and Intelligent Systems (CMLIS)
- Founding Member, Center for Bio-medical Informatics (CBMI)

RESEARCH INTERESTS and EXPERTISE

Broadly in the **data management**, **semantics**, and **applied artificial intelligence** areas. Specific areas include information integration, information agents, applications of Semantic-Web technologies, semi-structured data management, information extraction from text, uncertain databases, applied machine-learning, and information semantics. I have worked in the context of different application domains for the above technologies, specifically information systems for disaster response and emergencies, and data integration in biomedical informatics.

EDUCATION

UNIVERSITY OF SOUTHERN CALIFORNIA, LOS ANGELES

PhD, Computer Science, May 2000

Dissertation Title: Optimizing Information Mediators by Selectively Materializing Data

Thesis Advisor: Prof. Craig A. Knoblock

INDIAN INSTITUTE OF TECHNOLOGY, KANPUR, INDIA

BTech, Computer Science, May 1993

PRIOR EMPLOYMENT

Research Scientist, USRA RIACS, NASA Ames Research Center

March 2002- August 2005

- R&D activities spanning research, software design and development, applications development, research and technical publications, and proposal development for new funding opportunities.

Computer Scientist, IBM Life Sciences Solutions, San Jose, CA

Dec 2000 – March 2002

- Research and development of data integration software for addressing data extraction and integration problems in the bio-informatics domain.

Visiting Assistant Professor, Department of Computer Science, University of Georgia, Athens, GA

Jan 2000 – Nov 2000

- Co-Investigator role in research activities in information integration as a member of the Large Scale Distributed Information Systems Lab.
- Teaching graduate and undergraduate level computer science courses in the areas of databases and internet information systems.

Graduate Research Assistant, University of Southern California/Information Sciences Institute
Jun 1995 – Dec 1999

- Research activities for the SIMS and Ariadne projects on information agents and information integration. Lead role on research activity on optimization aspects of information mediators as well as semi-automatic approaches to wrapper generation for internet sources. This research group also spun off Fetch Technologies (www.fetch.com) to commercialize the technology developed.

Member of Technical Staff (Summer Intern), Bell Laboratories, Lucent Technologies, Murray Hill, NJ
June 1997-Aug 1997

CONSULTING ENGAGEMENTS

Expert Consultant

Kaboodle Inc. (<http://www.kaboodle.com>) Sep 2005 – Oct 2006

Expert consultant for R&D in information extraction from internet data, applications to online shopping

Expert Consultant

Public Display Inc. Sep 2006 – May 2007

Technical consultant for automated information extraction R&D from Web data.

Technical Expert Consultant

To law firm of Scott Jones and Co. March 2007

IBM vs Amazon on litigation related to web wrapper technology.

PROGRAM COMMITTEES and ORGANIZATIONAL ACTIVITY

Chair, ISCRAM 2008 Special Session on Process Aware Disaster Management

Chair, AAAI 2006 Workshop on Event Extraction

Co-Chair, First Workshop on Semantic Web Technologies for Searching and Retrieving Scientific Data, Oct 2003, Sanibel Island, Florida

Program Committees: GeOS 2009, 9th International Semantic Web Conference (ISWC), Barcelona, 2008; SIGAW Semantic Infrastructure for Grid Computing Applications, Cardiff UK, 2005; 10th International Conference on Co-operative Information Systems, Cyprus, 2005; 1st Conference on Geo-Spatial Semantics, GeOS 2005; 9th International Conference on Co-operative Information Systems, Cyprus, 2004; 8th International Conference on Co-operative Information Systems, Trento Italy, 2003

Editor, Journal of Digital Content Technologies and Applications (JDCTA).

Guest Editor, Journal of Digital Libraries: Special Issue on Semantic Web and Science Data Interoperation, 2004

Reviewer for several journals such as IEEE Transactions on Knowledge and Data Engineering, International Journal of Co-operative Information Systems, Journal of Parallel and Distributed Computing, IEEE Transactions on Systems Man and Cybernetics, Journal of the Geological Society of America, and ACM Transactions on Internet Technology.

PUBLICATIONS

BOOK CHAPTERS

Naveen Ashish and Sharad Mehrotra

XAR: An Integrated Framework for Semantic Extraction and Annotation

Cases on Semantic Interoperability for Information Systems Integration: Practices and Applications

Yannis Kalfoglou (Editor), IGI Global, June 2009

Naveen Ashish, Ronald Eguchi, Rajesh Hegde, Charles Huyck, Dmitri Kalashnikov, Sharad Mehrotra, Padhraic Smyth and Nalini Venkatasubramanian

Situational Awareness Technologies for Disaster Response

Terrorism Informatics: Knowledge Management and Data Mining for Homeland Security (Chen et al Eds)

Springer, 2008

JOURNAL ARTICLES

Naveen Ashish and David A. Maluf

Intelligent Information Integration: Reclaiming the Intelligence

International Journal of Intelligent Information Technologies (IJIT), 5(3), 2009

Naveen Ashish, Craig A. Knoblock and Cyrus Shahabi

Selectively Materializing Data in Mediators by Analyzing User Queries

International Journal of Cooperative Information Systems (IJCIS), 11(1/2), pp 119-144, 2002

Craig Knoblock, Steve Minton, Jose Luis Ambite, Naveen Ashish, Ion Muslea, Andrew Philpot and Sheila Tejada

The ARIADNE Approach to Modeling Web Sources for Information Integration

International Journal of Cooperative Information Systems (IJCIS) Special Issue on Intelligent Information Agents: Theory and Applications, 10(1/2), pp 145-169, 2001

Naveen Ashish, Sharad Mehrotra and Pouria Pirzadeh

Incorporating Integrity Constraints in Uncertain Databases

CoRR:arXiv:0907.1632

REFEREED CONFERENCE AND WORKSHOP PROCEEDINGS

Naveen Ashish, Jay Lickfett, Sharad Mehrotra and Nalini Venkatasubramanian

The Software EBox: Integrated Information for Situational Awareness

IEEE Intelligence and Security Informatics (ISI), June 2009, Dallas

Naveen Ashish, Sharad Mehrotra and Pouria Pirzadeh

XAR: An Integrated Framework for Free Text Information Extraction

IEEE CSIE Conference, April 2009, Los Angeles

Jay Lickfett, Naveen Ashish, Sharad Mehrotra, Nalini Venkatasubramanian and Jacob Green

The RESCUE Disaster Portal for Disasters and Emergency Response

The 5th International ISCRAM Conference, May 2008, Washington DC

Naveen Ashish, Tom Chidester, Deepak Kulkarni, Rich Keller, Bob Lawrence, Bob Lynch, Hemil Patel, Irving Statler, Yao Wang and May Windrem

Secure Integration of Aviation Data

NASA Tech Briefs, Mar 2007 (patent pending on this work)

Yiming Ma, Dmitri Kalashnikov, Sharad Mehrotra, Nalini Venkatasubramanian, Ram Hariharan, Naveen Ashish and Jay Lickfett "On-Demand Information Portals for Disaster Situations"
IEEE International Conference on Intelligence and Security Informatics (ISI) 2007, May 23-24, New Brunswick, NJ

Entity Search: A New Paradigm for People Search on the Web
Dmitri Kalashnikov, Rabia Nuray, Stella Chen, Naveen Ashish, Sharad Mehrotra
ICDE International Conference on Data Engineering, April 2007

David A. Maluf, David G. Bell, Naveen Ashish, Peter Putz, Yuri Gawdiak: Business Intelligence in Large Organizations: Integrating Which Data?
ISMIS 2006: 248-257

Dmitri V. Kalashnikov, Yiming Ma, Sharad Mehrotra, Ramaswamy Hariharan, Nalini Venkatasubramanian, Naveen Ashish
SAT: Spatial Awareness from Textual Input
Extending Database Technology (EDBT) 2006

Naveen Ashish, Dmitri Kalashnikov, Sharad Mehrotra and Nalini Venkatasubramanian
Event Based Approach to Situational Awareness
CoRR:arXiv:0906:4096

David Maluf, David Bell, Naveen Ashish, Peter Tran and Chris Knight
Semi-structured Data Management in the Enterprise: A Nimble, High-Throughput, and Scalable Approach
International Database Engineering and Applications Symposium (IDEAS) 2005, Montreal Canada

Alon Y. Halevy, Naveen Ashish, Dina Bitton, Michael J. Carey, Denise Draper, Jeff Pollock, Arnon Rosenthal and Vishal Sikka
Enterprise information integration: successes, challenges and controversies.
ACM SIGMOD Conference 2005: 778-787, Baltimore MD

David Maluf, David Bell and Naveen Ashish
Lean Middleware
ACM SIGMOD International Conference on Management of Data, June 2005, Baltimore MD

Andre Goforth, Naveen Ashish, Gabor Karsai, Robert Neches and Pedro Szekeley
Attaining Program Affordability Through Integration of Logistical Operations and Health Maintenance
AIAA 1st Space Exploration Conference, Orlando FL, Jan 2005

Naveen Ashish and Andre Goforth
Intelligent Information Fusion in Aviation Domain: A Semantic-Web based Approach
AIAA Intelligent Systems Conference, Chicago IL, Sep 2004

Naveen Ashish, Deepak Kulkarni and Yao Wang
Source Update Capture in Information Agents
IJCAI Workshop on Information Integration on the Web, 2003, Acapulco, Mexico

Tarcisio Lima, Amit P. Sheth, Naveen Ashish, Mukesh Guntamadugu, Sriram Lakshminarayanan, Narayanan Palsena and Dilpreet Singh.
Digital Library Services Supporting Information Integration over the Web.
Workshop on Information Integration on the Web, Rio de Janeiro, Brazil, April 2001

Naveen Ashish, Craig A. Knoblock and Cyrus Shahabi
Selectively Materializing Data in Mediators by Analyzing User Queries
Fourth IFCS Conference on Cooperative Information Systems, CoopIS 1999, Edinburgh, Scotland

Naveen Ashish, Craig A. Knoblock and Cyrus Shahabi
Selectively Materializing Data in Mediators by Analyzing Query Distribution, Source Structure and Maintenance Cost
ACM CIKM 2nd Workshop on Web Information and Data Management (WIDM), 1999, Kansas City, Missouri

Naveen Ashish, Craig A. Knoblock and Cyrus Shahabi
Intelligent Caching for Information Mediators: A KR Based Approach
Knowledge Representation meets Databases (KRDB) 1998, Seattle, WA

Craig Knoblock, Steve Minton, Jose-Luis Ambite, Naveen Ashish, Pragnesh Modi, Ion Muslea, Andrew Philpot and Sheila Tejada
Modeling Web Sources for Information Integration
AAAI 1998, The Fifteenth National Conference on Artificial Intelligence, July 1998, Madison WI
[More than 200 citations. Source: Google Scholar]

Jose-Luis Ambite, Naveen Ashish, Craig A. Knoblock, Steven Minton, Pragnesh J. Modi, Ion Muslea, Andrew Philpot, and Sheila Tejada
Ariadne: A System for Constructing Mediators for Internet Sources
ACM SIGMOD International Conference on Management of Data, June 1998, Seattle WA

Naveen Ashish, Craig A. Knoblock and Cyrus Shahabi
Semantic Caching for Information Integration
AAAI Workshop on AI and Data Integration, July 1998, Madison, WI

Naveen Ashish, Craig Knoblock and Alon Levy
Information Gathering Plans with Sensing Actions.
Fourth European Conference on Planning, 1997, Toulouse, France

Naveen Ashish and Craig Knoblock.
Semi-automatic Wrapper Generation for Internet Information Sources.
Second IFCS Conference on Cooperative Information Systems (CoopIS) 1997, Charleston, South Carolina

Naveen Ashish and Craig Knoblock.
Wrapper Generation for Semi-structured Internet Sources.
Invited paper, ACM SIGMOD Record, December 1997, Special Issue on Management of Semi-structured Data.
[More than 300 citations. Source: Google Scholar. Was in **Top 10 SIGMOD Record papers** for the 10 year 1995-2005 period]

Naveen Ashish and Craig Knoblock.
Wrapper Generation for Semi-structured Internet Sources.
ACM SIGMOD Workshop on Management of Semi-structured Data, 1997, Tucson , Arizona.

Naveen Ashish, Craig Knoblock and Steve Minton
Extracting Information from Online Semi-structured Sources
Autonomous Agents, Marina del Rey CA, 1997

MAGAZINE ARTICLES

Naveen Ashish, Chris Knight and David Maluf
EII: Achieving Scalability and Cost-Effectiveness
DM Review, Jan 2008

Naveen Ashish, Doug Appelt, Dayne Freitag and Dmitry Zelenko
The AAAI 2006 Workshop on Event Extraction: Co-chair's report
AI Magazine, 2007

TECHNICAL REPORTS

Naveen Ashish, Deepak Kulkarni and Yao Wang
Source Update Capture in Information Agents
NASA Center: Ames Research Center
Publication Year: 2003 Added to NTRS: 2006-08-03
Document ID: 20060015632

Naveen Ashish, David Bell and David Maluf
Lean Middleware
NASA Center: Ames Research Center
Publication Year: 2005 Added to NTRS: 2005-08-25
Document ID: 20050184137

Naveen Ashish and Andre Goforth
Intelligent Information Fusion in the Aviation Domain: A Semantic-Web based Approach
NASA Center: Ames Research Center
Publication Year: 2005 Added to NTRS: 2005-02-01
Document ID: 20050040756

Naveen Ashish
Semantic-Web Technology Applications at NASA
NASA Center: Ames Research Center
Publication Year: 2004 Added to NTRS: 2005-01-25
Document ID: 20050010168

Naveen Ashish
NASA and The Semantic Web
NASA Center: Ames Research Center Publication Year: 2005
Added to NTRS: 2005-01-25
Document ID: 20050019521

INVITED TALKS and TUTORIALS

Invited presentation on "NASA and the Semantic Web", At the *Dagstuhl Seminar on Semantic Interoperability and Integration*, Sep 2004, Dagstuhl Germany.
Note: Available as NASA Technical Report NTRS: 2005-01-25 (20050010168)

Information Mediation: Integrating Information from Multiple Information Sources (with Amit P. Sheth)
The Tenth International Conference on Management of Data (COMAD), December 2000, Pune, India

TECHNOLOGY DEVELOPMENT

Strong interest in development and distribution of open-source technologies for community use.

- Primary architect and developer of the XAR information extraction platform. This is an extraction system developed using many off-the-shelf and open-source text analysis tools. Provides the capability of building powerful information extraction applications from extraction from free text in an easy and intuitive manner. System is available (open-source) at <http://www.ics.uci.edu/~ashish/xar> under a Creative Commons License.
- Key developer of NETMARK, an XML based data management and integration system developed at NASA Ames Research Center. Open-source version available for non-profit use.
- Primary developer of TGRAM, a taxonomy driven graph mining engine.

INVITED TALKS AND SEMINARS

Indian Statistical Institute, Calcutta, India, March 1999 ; Indian Institute of Management , Calcutta, India , March 1999 ; Infosys Technologies Limited, Bangalore, India , March 1999 ; Virginia Tech, Blacksburg CA, Jan 2000 ; Enosys Markets Inc, San Diego CA, Aug 2000 ; Nimble Technologies, Seattle WA, Aug 2000 ; San Diego Supercomputer Center, San Diego CA, Sep 2001 ; NASA Ames Research Center, Moffet Field CA, Dec 2000 ; SRI International, Menlo Park CA, Jan 2002 ; Accenture Technology Labs, Palo Alto, Oct 2003 ; Arizona State University, Tempe, Oct 2003 ; Jet Propulsion Lab (JPL), Pasadena, May 2004, University of California – Irvine, July 2004, Institute for Defense and Government Analysis (IDGA), Washington DC June 2006; Bell Labs India, Bangalore India, April 2007; Disney Research, Glendale Jan 2009, CSU – Channel Islands, April 2009.

HONORS AND AWARDS

- NASA Ames Space Act Board Award for Aviation Data Integration System (ADIS), 2004
- NASA Ames Software Release Award for Aviation Data Integration System (ADIS), 2004
- NASA Ames Patent Application Inventor Award, 2004
- NASA Ames Honor Award for contributions to Aviation Data Integration Project (ADIP), 2003
- AAI Doctoral Consortium Award, American Association of Artificial Intelligence, 1998
- Best Speaker, First IMSC student retreat, University of Southern California, 1998
- Academic Proficiency Certificate, From Senate, IIT Kanpur, India, 1990
- Founders Gold Medal, For topping grade XII examination, La Martiniere for Boys, Calcutta, India, 1989

SPECIAL INVITATIONS

- Invitation to Dagstuhl Seminar on *Semantic Interoperability and Integration*, Dagstuhl Germany, Sep 2004
- Invitation to Dagstuhl Seminar on *Machine Learning for the Semantic Web*, Dagstuhl Germany, Sep 2005

THESIS COMMITTEES

(All at the University of Georgia, Department of Computer Science)

1. Dilpreet Singh, MS, 2000, 2. Narayanan Palsena, MS, 2000, 3. Mukesh Guntamagadu, MS, 2000, 4. Sriram Lakshminarayanan, MS, 2000

PERSONAL INFORMATION

Indian Citizen; US Permanent Resident; Married

Erdos Number: 4