BIOGRAPHICAL SKETCH

Provide the following information for the key personnel and other significant contributors in the order listed on Form Page 2.
Follow this format for each person. **DO NOT EXCEED FOUR PAGES.**

NAME	POSITION TITLE
Sheth, Amit P.	LexisNexis Ohio Eminent Scholar & Professor
eRA COMMONS USER NAME APSHETH	

EDUCATION/TRAINING *(Begin with baccalaureate or other initial professional education, such as nursing, and include postdoctoral training.)*

INSTITUTION AND LOCATION	DEGREE *(if applicable)*	YEAR(s)	FIELD OF STUDY
Birla Institute of Technology & Science, Pilani, India	B.E.(Hons)	1976-1981	Electrical & Electronics Eng
Ohio State University, Columbus, OH	M.S.	1981-1983	Computer & Info Science
Ohio State University, Columbus, OH	Ph.D.	1983-1985	Computer & Info Science

Professional Positions

2010 –	Director, Ohio Center of Excellence in Knowledge-enabled Computing (Kno.e.sis). (2009-) Professor (Bioinformatics & Biomedical Track), Biomedical Sciences PhD program.
2007 –	LexisNexis Ohio Eminent Scholar; Professor of Computer Sc & Engg, Wright State University, Dayton, OH. Director, Kno.e.sis Center: http://knoesis.org
1994 – 2006	Director, Large Scale Distributed Information Systems (LSDIS) Lab, Athens, GA.
1998 – 2006	Professor, Department of Computer Science, University of Georgia, Athens, GA
1999 – 2006	Founder, Chairman and CEO, Taalee Inc (till 5/2001, then Co-founder/CTO, Semagix, Inc. thru 10/2006). (on part time sabbatical when managing Taalee)
1997 – 1999	Founder and President, Infocosm Inc., S-Corporation, Georgia.
1994 – 1998	Assoc. Professor, Dept. of Computer Science, University of Georgia, Athens, GA.
1995 – present	Adjunct Associate Professor, College of Computing, Georgia Institute of Technology, Atlanta, GA.
4/1993 – 6/1993	Visiting Faculty, Institute of Information Systems, ETH, Zurich, Switzerland.
1989 – 1994	Member of Technical Staff, Bell Communications Research (Bellcore) Inc., NJ.
1987 - 1989	Staff Research Scientist, UNISYS West Coast Research Center, CA.
1985 - 1987	Principal Research Scientist, Computer Sc. Center, Honeywell, Minneapolis, MN.

Personal Statement

Amit Sheth directs the Ohio Center of Excellence on Knowledge-enabled Computing (Kno.e.sis). With 15 faculty from Computer Science, Biomedicine, Healthcare and Cognitive Science, and approx. 50 funded PhD students, it likely the largest US academic research group of Semantic Web (SW) researchers/students. Due to Kno.e.sis' activities, Wright State University shared 4th among Universities (accessed Oct 29,2012) in the world based on 5-uear publication impact in the list of top organizations in World Wide Web by Microsoft Academic Search (http://bit.ly/www-org). Sheth and his group's experience is significant to this proposal in two ways: (a) Sheth has had extensive collaborations with life sciences and health Informatics researchers since 1994 (http://knoesis.org/amit/hcls). This has included a NIST funded Healthcare Informatics project (as the PI for about $2M+ component), a NHLBI funded Collaborative R01 with NCBO (as the PI, with U of Georgia and NCBO/Stanford as partners), as a joint PI of an R21 to evaluate abuse of pain killers through semantic analysis of social media), and a NCRR (as the coordinator of informatics component). Around 50 of his 300+ publications are co-authored with researchers in healthcare and life sciences (HCLS). Several open source tools and repositories have also resulted from these collaborations (http://knoesis.org/opensource) as well as a deployed/operational applications and commercial products (http://knoesis.org/amit/commercialization). (b) He has been extensively involved in developing SW standards and activities serving SW and biomedical research communities. He has been advisory committee member for the World Wide Web consortium- W3C since 2002, and he and his group members have been active participants in SW HCLS initiative starting with providing an early use case in 2005. His leadership also resulted in development of GLYDE II, a standard for data

exchanged adopted by world-wide complex carbohydrate research community. Sheth also has extensive experience in leading multidisciplinary and multi-institution projects he has been PI/PD/coordinator of research with over $15 million, majority of these have been multi-disciplinary and multi-institutional, and most have involved successful software development by virtual teams.

Prof. Sheth will serve as the contact PI for this project, and provide leadership for overall ISE development that will integrate his team's expertise in ontology alignment, semantic association and ranking, semantic query formulation and processing, and visualization with Kavuluru's expertise in access control and semantic annotation of literature, Sahoo's expertise in provenance and semantic query processing, and Miller's expertise in data access, integration and mapping (esp. EupathDB). He will host regular conference calls with both development and user teams, ensure tool's testing and evaluation with bench researchers (working with Minning and Tarleton, then Kissinger, and later Louis), will ensure coordination with the usability advisory team of Tarleton/Kissinger/Louis as well as coordinate broader evaluation and dissemination with EuPathDB, VectorBase and NCBO comunities. He has worked for 4 to 14 years (and extensively published) with each of the investigators in this project including all work identified as preliminary work for this application.

Awards and other Professional Activities

- **Trustee's award** – Wright State University, 2010**:** "Highest award given by the University … intended to honor those who serve as the most outstanding of role models for all faculty."
- **IEEE Fellow** for contributions to information integration and workflow management, class of 2006.
- Among the top 100 authors in Computer Science (h-index of 76: see http://www.cs.ucla.edu/~palsberg/h-number.html), in WWW (listed among top 3; see http://j.mp/www-0113), in database (among top 20 based on all time) -- all based on a Microsoft Academic Search ranking, accessed January 08, 2013. 24,000+ citations on Google Scholar, 57 papers with 100+ citations each. Two best paper awards at conferences, four additional papers nominated for best paper awards.
- Presented 45 **keynote addresses**, 40+ invited talks at International Conferences and Workshops, and 200+ invited talks/colloquia/tutorials/panels.
- Commercial impact of research, from (a) five significant commercial products (InfoHarness from Bellcore, METEOR EAppS workflow management system from Infocosm, MediaAnywhere from Taalee, SCORE/Freedom from Semagix, **Active Semantic Electronic Medical Record** system from Athens Heart Center – in active deployment since 01/2006) (b) technology licensing resulting in two companies (managed as COB/CEO/CTO), (c) two patents including the first patent on Semantic Web, and (d) many deployed/operational commercial and scientific applications. Contributions to standards activities and community efforts (initiator and/or member of working group) in Semantic Web Services (WSDL-S and SAWSDL-now a W3C recommendation); co-chairing Semantic Sensor Networking Incubator. Also initiator and co-author of GLYDE, **GLYDE-CT** -- now adopted by the Glycomics scientific community as a standard XML protocol for exchange of glycan structural data. Guided/co-guided development of 10+ open-source data and tools. Details: http://knoesis.org/amit/commercialization , http://knoesis.org/opensource
- **Editor in Chief**, Intl Journal on Semantic Web & Information Systems (ranked among top 5 in WWW: http://bit.ly/www-j). Joint **Editor-in-Chief**, Distributed & Parallel Databases – an International Journal, Springer. Co-editor of a book series on Semantic Web & Beyond, Springer. Member of 5 Journal Editorial Boards, Served on 200+ Program committees and organized (as chair/co-chair) 75+ international conferences and workshops.
- **Advisory Committee Member,** World Wide Web Consortium, since 2002. **Board of Directors**, **Healthcare Open Systems & Trials (HOST) consortium**, 1995—1999.
- **Five innovation/challenge awards** from IBM (Eclipse and UIMA), HP Labs and Microsoft Research. Five Recognition of Service Awards/Certificates of Appreciation from the IEEE TC on Data Engineering and the IEEE Computer Society, 1989, '90, '91, '92, from the ACM, 1993. **ACM Lecturer**, 91-'92/92-'93.

Publications (from a total of over 325)

1. Amit Sheth and James Larson, “Federated Databases: Architectures and Issues,” ACM Computing Surveys, 22 (3), September 1990, pp. 183-236.
2. K. Kochut, J. Arnold, A. Sheth, J. Miller, I.B. Arpinar, J. Cardoso, IntelliGEN: A Distributed Workflow System for Discovering Protein-Protein Interactions, Distributed and Parallel Databases, An International Journal, Special Issue on Bioinformatics, 13 (1), 01/ 2003, pp. 43-72.
3. S S. Sahoo, C. Thomas, A. Sheth, C. Henson, W. S. York, GLYDE - An expressive XML standard for the representation of glycan structure. Carbohydrate Research, 340 (18), Dec 30, 2005: pp. 2802-2807. Epub 2005 Oct 20. PMID:16242678.
4. S. S. Sahoo, C. Thomas, A. Sheth, W. York, S. Tartir, Knowledge Modeling and its application in Life Sciences: A Tale of two Ontologies, Proc. 15th Int. World Wide Web Conf (WWW2006), May 2006.
5. Amit P. Sheth, S. Agrawal, Jonathan Lathem, Nicole Oldham, H. Wingate, P. Yadav, and K. Gallagher, 'Active Semantic Electronic Medical Record,' Proc. of the 5th Intl Semantic Web Conference, Athens, GA, November 6-9, 2006, pp. 913-926.
6. S. S. Sahoo, K. Zeng, O. Bodenreider, and A. Sheth, “From ‘Glycosyltransferase’ to ‘Congenital Muscular Dystrophy’: Integrating Knowledge from NCBI Entrez Gene and the Gene Ontology,” in MEDINFO 2007: Proceedings of the 12th World Congress on Health (Medical) Informatics, Studies in Health Technology and Informatics, Vol. 129, IOS, August 2007, pp. 1260–04. PMID: 17911917.
7. Satya S. Sahoo, Olivier Bodenreider, Joni L. Rutter, Karen J. Skinner, and Amit P. Sheth, “An Ontology-Driven Semantic Mash-up of Gene and Biological Pathway Information: Application to the Domain of Nicotine Dependency,” special issue: Semantic Biomedical Mashups, Journal of Biomedical Informatics, 41 (5), October 2008, pp.752-765. PMID: 18395495.
8. S. S. Sahoo, O. Bodenreider, P. Hitzler, A. Sheth and K. Thirunarayan, 'Provenance Context Entity (PaCE): Scalable provenance tracking for scientific RDF data.' In the 22nd International Conference on Scientific and Statistical Database Management (SSDBM), Heidelberg, Germany 2010.
9. Ashwin Manjunatha, Paul Anderson, Ajith Ranabahu, and Amit Sheth, 'Identifying and Implementing the Underlying Operators for Nuclear Magnetic Resonance based Metabolomics Data Analysis', Third Intl conf on Bioinformatics and Computational Biology, New Orleans, LA, USA, March 23 - 25, 2011.
10. Delroy Cameron, Ramakanth Kavuluru, Olivier Bodenreider, Pablo Mendes, Amit Sheth, Krishnaprasad Thirunarayan, “Semantic Predications for Complex Information Needs in Biomedical Literature,” 5th International Conference on Bioinformatics and Biomedicine BIBM11, Atlanta GA, Nov 12-15, 2011.
11. David Wild, Ying Ding, Amit Sheth, Lee Harland, Michael Lajiness, ‘Systems chemical biology and the Semantic Web: what they mean for the future of drug discovery research, Drug Discov Today. 2011 Dec 29. PMID: 22222943 .
12. Satya Sahoo, Vinh Nguyen, Olivier Bodenreider, Priti Parikh, Todd Minning and Amit Sheth. "A Unified Framework for Managing Provenance Information in Translational Research." BMC Bioinformatics 2011, 12:461 doi:10.1186/1471-2105-12-461. PMID: 22126369 [Highly Accessed]
13. Ramakanth Kavuluru, Christopher Thomas, Amit Sheth, Victor Chan, Wenbo Wang, Alan Smith, “An Up-to-date Knowledge-Based Literature Search and Exploration Framework for Focused Bioscience Domains,” IHI 2012 - 2nd ACM SIGHIT Intl Health Informatics Symposium, January 28-30, 2012.
14. Priti Parikh, Todd A. Minning, Vinh Nguyen, Sarasi Lalithsena, Amir H. Asiaee, Satya S. Sahoo, Prashant Doshi, Rick Tarleton, and Amit Sheth. “A Semantic Problem Solving Environment for Integrative Parasite Research: Identification of Intervention Targets for Trypanosoma cruzi.” PLoS Negl rop Dis 6(1): e1458. doi:10.1371/journal.pntd.0001458, 2012. PMID: 22272365.

15. D. Cameron, O. Bodenreider, H. Yalamanchili, T. Danh, S. Vallabhaneni, K. Thirunarayan, A. P. Sheth, T. C. Rindflesch, A Graph-Based Recovery and Decomposition of Swanson's Hypothesis using Semantic Predications, Journal of Biomedical Informatics (2012) http://dx.doi.org/10.1016/j.jbi.2012.09.004

Research Projects Ongoing or Completed During Last 3 Years

ONGOING

NIH (NHLBI) R01: A. Sheth (Role: PI) [1R01HL087795-01A1] 05/01/2008-03/31/2013

Semantics and Services Enabled Problem Solving Environment (SPSE) for Trypanosoma cruzi

Collaborative R01 led by Wright State Univ. with University of Georgia and Stanford University as partners. The scientific analysis of the parasite *T. cruzi*, the principal causative agent of human Chagas disease. The SPSE allows data analysis and knowledge discovery through the dynamic integration of lab and public data to answer biological questions at multiple levels of granularity. [Also supported by an ARRA supplement.]

AFRL through Ball Aerospace: A. Sheth (Role: PI) 10/15/2010-10/14/2013

LVC Sensors Integration for Data Fusion in Operations and Training

Perform research related to the fusion of mobile, social, and sensor data, with applications to emergency (including medical emergency) response during disasters.

NIH (NIDA): A. Sheth (Role: PI, a multi-PI project) [R21 DA030571-01A1] 07/01/2011-06/30/2013

A Study of Social Web Data on Burprenorphine Abuse Using Semantic Web Technology

Web-based study will generate new information about burprenorphine/naloxone and buprenorphine abuse practices to inform public health interventions and policy. Also, contributing to the advancement of public health and substance abuse research methods by providing automatic coding and information extraction tools to handle rapidly growing Web-based data.

ezDI: A. Sheth (Role: PI) 06/01/2011-03/31/2013

Sponsored research involving Semantic Web technology and ontology supported data mining of Cardiology EMR and transcription data to reduce errors and improve adherence to medical guidelines. Includes development of a comprehensive Cardiology ontology with clinical perspective.

NSF-EAGER: A. Sheth (Role: PI) 09/01/11 – 08/31/13

Expressive Scalable Querying over Integrated Linked Open Data (ESQUiLO): This project develops exploratory techniques to richly interlink components of Linked Open Data (LOD) and then addresses the challenge of querying the LOD cloud, i.e., of obtaining answers to questions which require accessing, retrieving and combining information from different parts of the LOD cloud.

NSF-SOCS: A. Sheth (Role: PI) 09/01/11 – 08/31/14

Collaborative Research: Social Media Enhanced Organizational Sensemaking in Emergency Response: This project seeks to leverage Twitter posts (tweets) as the primary source of citizen inputs and couple relevant content and network information along with microworld simulations involving human role players to measure effectiveness of various organized sensemaking strategies. Evaluations will involve modeling disaster situations and typical organizational structures.

COMPLETED (partial list)

Jackson Foundation (AFRL) (Role: PI) 6/1/2008-7/31/2010

Human Performance Cognition Ontology

Development of comprehensive knowledgebase using unique approach of extracting initial taxonomy from relevant Wikipedia pages and then enhancing it through entity and relationship extraction from domain specific scientific literature. The project involved development of SCONNER, a semantic browser to support knowledge exploration by biologists from biomedical literature, and is now in use by AFRL biologists.

NSF-ITR (Medium): A. Sheth (Role: PI) 10/1/2003 – 12/312009

SemDIS: Discovering Complex Relationships in the Semantic Web

This research focus on the design, prototyping and evaluation of a system, called SemDIS (Semantic Discovery) that supports semantic associations (querying of complex semantic relationships and is driven by notions of information trust and provenance and models of hypotheses. Also developed techniques for complex entity and relationship extraction from biomedical text.