

Statement of Interest Submitted to the AAAI-98 Workshop on AI and Information Integration

Yigal Arens
USC/ISI
arens@isi.edu

Research Interests and Background

I have been working in the field of information integration since 1990, when I established the SIMS (Single Interface to Multiple Sources) research project at the University of Southern California's Information Sciences Institute. Since that time, the SIMS project has expanded into a group of several coordinated research efforts involving approximately 15 research staff and students, including some of the top researchers in the field today. The group has conducted and is currently conducting research on all aspects of the problem of information integration mentioned in the call for papers, including:

- High-level query facilities
- Modeling the contents of information systems
- Flexible approaches to selecting relevant sources
- Approaches for efficient and flexible query planning
- Learning or constructing wrappers for extracting data
- Resolving inconsistencies in identifying objects

In addition, I am currently heading (together with Prof. Salvatore Stolfo of Columbia University) an effort to establish CARDGIS, the Center for Applied Research in Digital Government Information Systems. The center's purpose is the application of modern information technology to U.S. Government information systems. AI in general and information technology of the kind developed under the SIMS project in particular are central to the aims of this new center.

Relevant Publications

1. Yigal Arens. *SIMS: Integrating Heterogeneous, Distributed Information Sources*. Proceedings of

the 1997 ONDCP/CTAC International Technology Symposium. Office of National Drug Control Policy, Executive Office of the President. Chicago, IL. August 18-21, 1997.

2. Yigal Arens. *SIMS: Addressing the Problem of Heterogeneity in Databases*. Proceedings of SPIE-The International Society of Optical Engineering Symposium: *Command, Control, Communication, and Intelligence Systems for Law Enforcement*. SPIE Proceeding Series Vol. 2938. Edward M. Carapezza and Donald Spector, Chairs/Editors. Boston, MA. November 19-21, 1996.
3. Yigal Arens, Craig A. Knoblock and Chunnan Hsu. Query Processing in the SIMS Information Mediator. In *Advanced Planning Technology*, edited by Austin Tate. The AAAI Press, Menlo Park, CA. May 1996, ISBN 0-929280-98-9.
4. Yigal Arens, Craig A. Knoblock, and Wei-Min Shen. Query Reformulation for Dynamic Information Integration. *Journal of Intelligent Information Systems*, Special Issue on Intelligent Information Integration. Vol. 6, No. 2/3, 1996, ISBN 0925-9902, pp. 99-130.
5. José-Luis Ambite, Yigal Arens, Naveen Ashish, Chin Y. Chee, Chun-Nan Hsu, Craig A. Knoblock, Wei-Min Shen, and Sheila Tejada. *The SIMS Manual, Version 1.0*. USC/Information Sciences Institute Technical Manual No. ISI/TM-95-428. July 1995.
6. Yigal Arens and Craig A. Knoblock. Intelligent Caching: Selecting, Representing, and Reusing Data in an Information Server. In *Proceedings of the International Conference on Information and Knowledge Management (CIKM-94)*. NIST, Gaithersburg, MD, November 29-December 1, 1994.

7. Craig A. Knoblock, Yigal Arens and Chun-Nan Hsu. Cooperating Agents for Information Retrieval. In *Proceedings of the Second International Conference on Cooperative Information Systems (CoopIS-94)*. Toronto, Canada, May 17–20, 1994.
8. Contributed to *R&D for the NII: Technical Challenges*, report of a symposium held in cooperation with the Technology Policy Working Group of the US Government's Information Infrastructure Task Force Committee on Applications and Technology, and with the National Institute of Standards and Technology. Edited by M.K. Vernon, E.D. Lazowska, and S.D. Personick. February 28–March 1, 1994.
9. Yigal Arens, Chin Y. Chee, Chun-Nan Hsu, Hoh In and Craig A. Knoblock. Query Processing in an Information Mediator. In *Proceedings of the ARPA/RL Knowledge-Based Planning and Scheduling Initiative Workshop*. Tucson, AZ, February 21–24, 1994.
10. Craig A. Knoblock and Yigal Arens. Cooperating Agents for Information Retrieval. In *Proceedings of the AAAI Spring Symposium on Software Agents*. Stanford, CA, 1994.
11. Yigal Arens Chin Y. Chee, Chun-Nan Hsu and Craig A. Knoblock. Retrieving and Integrating Data from Multiple Information Sources. *International Journal of Intelligent and Cooperative Information Systems*. Vol. 2, No. 2, 1993. Pp. 127–158.
12. Yigal Arens and Craig A. Knoblock. SIMS: Retrieving and Integrating Information From Multiple Sources. Demonstration system description in *Proceedings of SIGMOD-93*. Washington, DC, May 26–28, 1993.
13. Yigal Arens and Craig A. Knoblock. Planning and Reformulating Queries for Semantically-Modeled Multidatabase Systems. In *Proceedings of the First International Conference on Information and Knowledge Management (CIKM-92)*. Baltimore, MD, November 8–11, 1992. Pp. 92–101.
14. Yigal Arens and Craig A. Knoblock. SIMS: Services and Information Management for decision Support (slides). In *Proceedings of Joint RL/AFOSR Workshop on Intelligent Information Systems*. Eds. Raymond A. Liuzzi and Abraham Waxman. In-House Report RL-TR-92-57, Rome Laboratory, AFSC, Griffiss AFB, NY, pp. G-1–G-21, April 1992.
15. Yigal Arens. Services and Information Management for Decision Support. In *AISIG-90: Proceedings of the Annual AI Systems in Government Conference*. George Washington University, Washington, DC, May 7–11, 1990.