

AI News Bulletin

Emerging Trends in AI for the Media

FALL 2003 ■ AMERICAN ASSOCIATION FOR ARTIFICIAL INTELLIGENCE



A Note from AAAI's Media Liaison

Welcome to the first edition of the *AI News Bulletin* from the American Association for Artificial Intelligence. It is our intent in this publication to highlight some of the latest developments in the field of AI and present story ideas.

AI is a broad field with a diverse range of subfields — knowledge representation, language understanding, data mining, reasoning and machine learning, to name but a few. What unifies AI is the quest to understand and imbue machines with intelligence.

With the *AAAI News Bulletin*, we hope to assist you in keeping current on what's emerging from the various threads of AI research and application work. There are many fascinating stories in this field — many augur coming widespread trends in intelligent computing.

As the foremost AI society in North America, our membership includes many of the leading researchers and practitioners in the field. As such, we are uniquely positioned as a clearinghouse for information about the latest developments in AI — the headwaters of increasingly intelligent computing. Should you ever need assistance finding AI sources, please do not hesitate to call upon us. We will be happy to connect you with highly-respected authorities in the field.

Visit our new e-Press Room!

AAAI has also recently put up an e-press room on its web site as a resource for the media. It offers a variety of news sources to assist you in working on stories about the science and applications of AI.

Again, welcome!

Sara Hedberg

AAAI Media Liaison (hedberg@aaai.org)

What's Inside

- ✓ Current AI Applications: Where AI Is Having a Huge Impact
- ✓ AI at Google: AI Application Used Several Million Times a Day
- ✓ Paul Allen Funds AI Research
- ✓ AI & Molecular Biology: A Growing Success Story
- ✓ Other News from this Year's AI Conference

2004 AI Conferences Scheduled

The 2004 AI conferences are a good opportunity to learn of many advances in the research and application of AI, and to meet noted researchers and practitioners in the field. The AAAI-sponsored conferences will be held jointly July 25–29, 2004 in San Jose, CA. Visit www.aaai.org/Conferences/National/2004/ or contact hedberg@aaai.org for more information.

The *AI News Bulletin* is published periodically by The American Association for Artificial Intelligence, 445 Burgess Drive, Menlo Park, California 94025. Copies are available from the association's web site at www.aaai.org/Pressroom/.

For more information about the contents of this newsletter, contact Sara Hedberg, AAAI Media Liaison at hedberg@aaai.org.

Current AI Applications where AI Is having a Huge Impact

The question continues to be asked — “So where is AI being used these days?” Two companies using recent high-impact applications include Google and the NASDAQ.

Google News automatically generates and refreshes the news from 4,500 news sources every fifteen minutes using several AI techniques. Google News is used several million times a day.

NASDAQ surveillance for insider trading and fraud is done with AI-based data mining software that digs through between 8,500 and 18,000 news wire stories, approximately 1,000 quarterly and annual SEC filings from corporations, and evaluates price-volume models for 25,000 securities each day, identifying and linking items of interest that may warrant further investigation. The system has raised the red flag on a number of cases — including some very high profile, instantly recognizable cases — to the SEC and Justice Department for further investigation and/or prosecution. (For further information about this application, please visit our e-press room at www.aaai.org/Pressroom/Releases/releases.html.)

Other recent high-impact applications include the 2003 winners of AAAI’s Innovative Applications of Artificial Intelligence award, including the following:

Educational Testing Service, a leading supplier of fair and valid assessment instruments, received an award for a commercially deployed, web-based system that provides automated scoring and evaluation of student essays. It provides specific feedback for students to improve their writing skills. The application is intended to relieve the load on teachers faced with reading and providing detailed feedback on perhaps thirty essays or more every time a topic is assigned.

National Association of Securities Dealers — creators of the surveillance software for insider trading and fraud (See the description of NASDAQ application above).

Tokyo University of Science, Information Media Center (Japan) created a sales support system for door-to-door sales using a cellular phone with a camera, e-mail software, and a web browser to customize and automate product recommendations to the individual prospective customer.

Norwegian State Railways devised an automated scheduling system that schedules and manages the work of 1,800 personnel, 1,000 engine drives and 800 ticket inspectors allocated to 39 bases across Norway.

For more information about these award-winning examples of fully deployed innovative applications of AI, see AAAI’s e-press room (www.aaai.org/Pressroom/Backgrounders/backgrounders.html). This site also includes summaries of winning applications for the previous two years.

Contact hedberg@aaai.org for further information.

Paul Allen Funds AI Research

What do Paul Allen (cofounder of Microsoft with Bill Gates) and AI have in common? Allen’s investment company, Vulcan Ventures, has been funding research for a project called Halo. “Mr. Allen has a personal passion for AI,” according to Halo project manager Noah Friedland.

Halo is a staged research effort inching towards a “digital Aristotle,” named after the Greek philosopher said to have known the answer to any question about science. As envisioned, Halo will be capable of producing user and domain-appropriate answers and justifications to novel (previously unseen) questions in an ever-growing number of domains.

This past August, Friedland and members of the three contract groups involved in the first phase of Halo, participated in a lively panel presentation and discussion about the project at AAAI’s Innovative Applications of AI conference (IAAI-03.) In phase one, a six month effort, Vulcan funded three competing teams to use their distinctive knowledge representation, reasoning and query technologies to first encode 71 pages of advanced placement inorganic chemistry into laws and rules the computer could understand. Then each team developed a query language and formulated 100 free-form questions. Phase two of the project is now underway in this multistage project being managed in a DARPA RFP-like fashion.

The three competing teams for phase one included Cycorp Inc. (Austin, Texas), ontoprise GmbH (Karlsruhe, Germany), and a team composed of SRI International (Menlo Park, Calif.), the University of Texas at Austin, and Boeing Phantom Works (Seattle, WA). Information about the project and the results of the phase one bake-off can be found at www.projecthalo.com.

AI and Molecular Biology: A Growing Success Story

AI's role in assisting research molecular biologists in using an exploding number of data sources was the topic of Lawrence Hunter's talk during AAAI's Innovative Applications of AI conference (AAAI-03) this year. Hunter, Director of the Center for Computational Pharmacology at the University of Colorado Health Sciences Center, is the first recipient of the newly-established AAAI Robert S. Englemore Memorial Lecture award. This annual keynote lecture has been established by AAAI in honor of Dr. Englemore's extraordinary service to AAAI, *AI Magazine*, and the applied AI community.

In his talk, Hunter surveyed how successful AI has been in molecular biology. "Tens of thousands of molecular scientists use AI tools all the time," he asserted. "Lots of people are using the technology, although they may not think of it as AI."

"There are huge sources of information about life," he says. "Molecular biologists have access to dozens of general and hundreds of specialized databases with little if any standardization.... There's an exponential explosion of data. We're drinking from the fire hose. More than 12,000 articles are published each week."

Hunter cited an AI-based data mining tool developed by Georgia Tech, GOMiner (Gene Ontology Miner) as an example of a literature-based information extraction and presentation system that performs meta analysis making connections between genes. (<http://discover.nci.nih.gov/gominer/>).

Other News from this Year's AI Conference

Information about other aspects of this year's conference, such as the Mobile Robot competition, Trading Agents competition, AAAI Fellows awards, and more can be found in the AAAI e-Press Room (www.aaai.org/Pressroom/Releases/releases.html).

AI at Google: AI Application Used Several Million Times a Day

Google Senior Research Scientist Mehran Sahami was an invited speaker at the International Joint Conference on AI (IJCAI-03). He spoke of the role of AI today and down the road at the popular web search engine company. Sahami is also a lecturer at Stanford University in the computer science department, where he received his Ph.D. His research interests include machine learning, data mining, intelligent agents and information retrieval on the Web. (<http://robotics.stanford.edu/users/sahami/bio.html>)

Sahami portrayed some of the issues Google faces in the Web information retrieval business, such as the following:

Data collection in an environment of varying page quality and trustworthiness, misspellings, bad formatting, spam and misinformation: Adversarial interests, for instance, try to fool the rankings, compromising the trustworthiness of the documents. The freewheeling nature of the web, with no editorial filters or mediation to ensure well-authored documents adds to the chaos. In some cases, misinformation gets served up by sites that "cloak" — serving the Google search engine one page, and the users another, or replacing words in otherwise legitimate text.

User queries that are short, have limited if any specificity, and can have many different misspellings. Google has documented, for instance, over 800 different spelling of Britney Spear's name and has received general queries such as "map" "davis" "homework help" and "people in Phoenix."

Speed and reliability of response with millions of users in a rapidly escalating growth of use. Sahami reported that the number of web users has increased from 140 million in 1999 to 320 million in 2002. In the same time period, the number of web pages has swelled from 500 million to 3–8 billion, and the number of searches has quintupled from 100 to 500 million in the same three years. (Source: Warburg Dillion Read LLC). Users want fast response time.

Google, one of the most popular search engines, manages colossal amounts of data and computing resources, indexing 4 billion documents, 3 billion URLs, 390 million images, 35 million file types, 700 million individual messages, and 30 terabytes of data in 35-plus languages. It services 200 million queries per day — or 1,000 per second. Google services this workload with more than 20,000 machines acting as servers and 1,000 terabytes of disk storage.

"AI applications are using the infrastructure to get people useful information in interesting ways," according to Sahami. For instance, Google News is automatically generated from 4,500 news sources each 15 minutes — using several AI techniques, says Sahami. "There is no human intervention," notes Sahami. "Google News is an example of where AI is making a huge difference. It's used several million times a day."

Sahami also gave a glimpse of AI-based research in progress at Google that has yet to be deployed, such as voice-driven search and query results clustering to help users navigate. "We want to combine information retrieval, large systems, and AI to work together towards the next generation of search engines," concluded Sahami.