

# Keynotes

Contributed by Sebastian Salvucci  
Friday, 26 October 2007  
Last Updated Friday, 23 November 2007

Hector Garcia Molina, Martin L. Kersten and Luis von Ahn confirmed as keynote speakers.

PhotoSpread: A Spreadsheet for Managing Photos  
Hector Garcia-Molina, Stanford University

PhotoSpread is a spreadsheet system for organizing and analyzing photo collections. It extends the current spreadsheet paradigm in two ways: (a) PhotoSpread accommodates sets of objects (e.g., photos) annotated with tags (attribute-value pairs). Formulas can manipulate object sets and refer to tags. (b) Photos can be reorganized (tags and location changed) by drag-and-drop operations on the spreadsheet.

The PhotoSpread design was driven by the needs of field biologists who have large collections of annotated photos. In the talk I will describe the PhotoSpread functionality and the design choices made. I will also describe some of the other data management tools we have developed with field biologists.

Hector Garcia-Molina is the Leonard Bosack and Sandra Lerner Professor in the Departments of Computer Science and Electrical Engineering at Stanford University, Stanford, California. He was the chairman of the Computer Science Department from January 2001 to December 2004. From 1997 to 2001 he was a member the President's Information Technology Advisory Committee (PITAC). From August 1994 to December 1997 he was the Director of the Computer Systems Laboratory at Stanford. From 1979 to 1991 he was on the faculty of the Computer Science Department at Princeton University, Princeton, New Jersey.

His research interests include distributed computing systems, digital libraries and database systems. He received a BS in electrical engineering from the Instituto Tecnológico de Monterrey, Mexico, in 1974. From Stanford University, Stanford, California, he received in 1975 a MS in electrical engineering and a PhD in computer science in 1979. Garcia-Molina is a Fellow of the Association for Computing Machinery and of the American Academy of Arts and Sciences; is a member of the National Academy of Engineering; received the 1999 ACM SIGMOD Innovations Award; is on the Technical Advisory Board of DoCoMo Labs USA, Yahoo Search & Marketplace; is a Venture Advisor for

Diamondhead Ventures, and is a member of the Board of Directors of Oracle and Kintera.

The Database Architecture Jigsaw Puzzle  
Martin L. Kersten, CWI, Amsterdam, The Netherlands

Each DBMS represents a solution in a design space covering hundreds of parameters. The sheer size of this space leaves large parts unexplored, but also requires courage. The open-source MonetDB system is used to exemplify the pitfalls and opportunities of such an exploration into the realm of column-stores. We illustrate the vistas of high-risk projects based on radical changes in the design parameters, e.g., database cracking for self-organization, informative query summaries and database storage rings where the database is on the move. The missing jigsaw pieces identified are important for real innovations and provide an inspiration for changing the legacy architecture embodied in (relational) database products.

Dr. Martin L. Kersten devoted most of his scientific career on the development of database systems. The latest incarnation is the open-source system MonetDB, which demonstrates viability of the column-storage approach as an sufficient basis for both an efficient SQL and XQuery database engine. The system is developed by the Database Architectures and Information Access group of CWI, which he established in 1985, and which hosts a strong group of experimental scientists.

Kersten is head of the Information Systems department of CWI and a full professor of the University of Amsterdam. He is an author of more than 100 papers and recipient of multiple large international research grants to steer multi-media and database research. He is a member emeritus of the VLDB Endowment.

Human Computation  
Luis von Ahn, Carnegie Mellon University

Construction of the Empire State Building: 7 million human-hours. The Panama Canal: 20 million human-hours. Estimated number of human-hours spent playing computer solitaire around the world in one year: billions. A problem with today's computer society? No, an opportunity.

What if this time and energy could be channeled into useful work? What if people could play computer games and accomplish work without even realizing it? What if billions of people collaborated to solve important problems for humanity or generate training data for computers? My work aims at a general paradigm for doing exactly that: utilizing human processing power to solve computational problems in a distributed manner.

In particular, I focus on harnessing human time and energy for addressing problems that computers cannot yet solve. Although computers have advanced dramatically in many respects over the last 50 years, they still do not possess the basic conceptual intelligence or perceptual capabilities that most humans take for granted. By leveraging human skills and abilities in a novel way, I want to solve large-scale computational problems and/or collect training data to teach computers many of these human talents. To this end, I treat human brains as processors in a distributed system, each performing a small part of a massive computation. Unlike computer processors, however, humans require an incentive in order to become part of a collective computation. Among other things, I use online games as a means to encourage participation in the process.

In this talk, I will describe my work in the area of Human Computation.

Luis von Ahn is an assistant professor in the Computer Science Department at Carnegie Mellon University. He is the recipient of a MacArthur Fellowship, and was named one of Popular Science Magazine's "Brilliant 10" scientists of 2006. His research interests include encouraging people to do work for free, as well as catching and thwarting cheaters in online environments.

