

Challenges and Opportunities in Scientific and Statistical Databases

Yannis E. Ioannidis

Dept. of Informatics, University of Athens, Hellas (yannis@di.uoa.gr)

Dept. of Computer Sciences, University of Wisconsin-Madison (yannis@cs.wisc.edu)

It is almost twenty years after the 1st SSDBM was held in Berkeley (then called Workshop on Statistical Database Management). It is almost ten years after the first of a few workshops organized by the National Science Foundation on scientific databases. How many of the original challenges have been accomplished? How many are still here with us and need serious attention? What are the new challenges that have emerged that were not put on the table until recently?

There are several recent technological advances that present new challenges but also offer major opportunities. The Web is clearly the first such development that comes to mind, and so are recent trends in hardware. On the other hand, the nature of experimental science has been changing as well: there are increasingly larger, distributed and multidisciplinary teams of scientists working together and putting continuously higher demands for supporting more complex models and instruments, and managing larger amounts of data. Where and how does the new technology meet the

new demands?

In the recent “Asilomar Report on Database Research” proposing important research directions for the future, a “grand challenge” has been proposed for the database research community to try to achieve in a decade:

The Information Utility: Make it easy for everyone to store, organize, access, and analyze the majority of human information online.

Does this subsume the current challenges in the scientific and statistical database arena? If we are successful on “The Information Utility”, will the entire SSDBM community be satisfied? If the answer to the two questions above is negative, what would be a corresponding “grand challenge” for scientific and statistical databases?

Going one step beyond the above, in general, how different is scientific and statistical data management from “regular” data management these days? The most re-

cent calls for papers for SSDBM have much in common with those of SIGMOD and VLDB; that was not the case earlier. Is it that the demands for “mainstream” data management finally reached the complexity that existed in scientific and statistical databases all along? Are there other reasons?

The panel will initiate discussion and attempt to provide answers to these and other questions related to the future of scientific and statistical data management. The panelists include researchers with extensive experience in the field, coming from both the database area and domain disciplines.