Progress on HPCC and NII

Marianne Winslett
Computer Science Department
University of Illinois
Urbana, IL 61801
winslett@cs.uiuc.edu

Abstract

In this issue we briefly touch on the continuing turmoil over NSF, ARPA, and HPCC, and the brighter news regarding the US National Information Infrastructure plan. We then describe funding opportunities from NSF, ARPA, the National Security Agency, the National Center for Automated Information Research, the Air Force, and NASA.

1 Funding Information Available On Line

All the funding announcements that are used in this column and are publicly distributable can be perused on-line at the SIGMOD Information Server. Many of the news stories covered here are also accessible on line, as is a wide variety of information related to obtaining funding. To access the server, use anonymous ftp to bunny.cs.uiuc.edu or, preferably, use Mosaic to browse the document http://bunny.cs.uiuc.edu:80/README.html. For more information on using the SIGMOD Information Server, see the September 1993 issue of the *Record*, or read the file pub/howToBrowseAndWhy.txt on bunny.cs.uiuc.edu.

The SIGMOD Information Server now runs on the more powerful machine bunny.cs.uiuc.edu, rather than swine.cs.uiuc.edu. Requests submitted to swine will be transparently forwarded to bunny for the next six months or so, but you should go ahead and update your Mosaic "hot list" to point directly to bunny. In the near future, we will introduce a more mnemonic moniker for the server.

We welcome offers of assistance in maintaining and/or augmenting the capabilities of the SIGMOD Information Server. For more information, send email to winslett@cs.uiuc.edu.

2 HPCC and the National Information Infrastructure

The troubles at NSF and ARPA, discussed in recent issues of this column, have continued. The biggest news is that the US Senate did not pass the follow-on version of the High Performance Computing Act before it ended its 1993 session, and it is not entirely clear that it *will* pass. The January 1994 issue of *Computing Research News* covers the history and sad current state of the legislation in detail in several articles. To summarize, given the number of special interest groups who have a stake in the High-Performance Computing and Communication Act, "it is surprising that any bill has been able to move through Congress. Because it has moved forward in the House and may well move, in some form, through the Senate next session attests to the deeper consensus on general technological priorities for the nstion. Everyone differs in the details. [...] Unfortunately, as is often heard on the Hill, 'The devil is in the details.' "

The January issue of *Computing Research News* also tracks the much more encouraging progress of the National Information Infrastructure Pilot Projects Program, appropriated US$26 million for fiscal 1994.

For more detail on these recent funding events, I recommend the January issue of *Computing Research News*, available on-line through Mosaic at URL "file://cra.org/pub/cra_b". (cra.org has a very slow connection to the internet, so browsing may be quite slow at peak times.)

3 US Department of Defense News

In November, Dr. Anita Jones ("director Defense Research & Engineering, equivalent to an Assistant Secretary of Defense") attended a meeting of the Defense Science Study Group (a small invitation-only group of professors in science and engineering, with typically 5–10 years experience in academia, who are given an overview of US defense needs and organization over the course of a year or so) at the Institute for Defense Analyses, and gave her views on the relationships between the US Department of Defense and university research. We excerpt from the report of that meeting below.

"Anita Jones has been the DDR&E (director Defense Research & Engineering, equivalent to an Assistant Secretary of Defense) since early July. Her background, which she described to us, includes training and a faculty post at CMU, with research in compilers for parallel computing. She described her association with the AFSAB as reinvigorating, because she found herself 'bored in her career' by the time she had tenure. Her point was to encourage the DSSG members to accept any future calls to service in DoD, but I also view her statements as evidence of a disenchantment with academics and research in general. Although she views CMU as the ideal research environment, it is clear that she views advisory boards such as the AFSAB and science policy as a higher level of service. Since then, she has served in industry at a company that her husband, Bill Wulf (well known himself in CS research) [founded], and at the University of Virginia, where she chaired the CS Department.

"She stated that her emphasis would be on the 'information technologies,' and she stated that she believes that there is 'gold to be mined in the interfaces between the sciences, particularly between information science and other fields.' My impression is that she will be quite sympathetic to the military need for ATR (Automatic Target Recognition, the focus of the study group) and brilliant missile design, and that the interdisciplinary nature of image understanding research, particularly neurophysiological guidance to computer vision development, fits well with her views. She sees information technologies as an enabler.

"We (the members of the DSSG) questioned her closely on the role of universities in sponsored research and defense science, seeking an appreciation of a university as a center of learning and breeding ground for research talent and a research environment that serves defense needs broadly. The answer was that 'We are defense,' and that sponsored research is a different side to the university endeavor, separate from education, while acknowledging the AASERT program for encouraging US citizens to pursue graduate degrees. (Ira Skulnick from Arpa noted that the URI [University Research Initiative] program has also provided education support.) She reiterated a recurring theme in the current administration that researchers in general must increasingly address the commercial and civilian potential in order to gain government funding. As such, we did not view her as an academic representative, in the way that might be expected. However, the news for defense labs was somewhat worse. She stated, simply, that the labs would be downsizing and restructuring. There was nothing about quality or roles or dogma in the justification of this prediction, but simply "budget constraints." The role of the labs is viewed as making the military into smart buyers, placing money, and providing for maintenance and upgrading of material, especially on a very short schedule, as when crises occur that call for a concerted dedicated effort. Within the domain of ATR research, it seems that the role of the labs in facilitating research, making it more efficient, and the resulting cost-savings, together with partnerships with industry and academia, is the best strategy in light of her views."

4 US Security Community Gears Up For Larger Databases

The advent of the internet age opens up many new questions in computer security and individual privacy. (The January 1994 issue of *Computing Research News* contains no fewer than four separate articles on recent government activity relating to privacy and security; space constraints prevent us from discussing the latest government actions here, but you may browse them on-line with Mosaic at URL "file://cra.org/pub/cra_b".) As part of the recent activity in this area, the Advanced Technology Office of the Community Management Staff in Washington DC issued a call for abstracts last November, for a workshop to be held at the beginning of February. Unfortunately, we received the call for submissions after the deadline.

The call for abstracts states that "Future intelligence systems must effectively manageme massive amounts of digital data (i.e., multi-terabytes or greater). Issues such as scalability, design, and integration need to be addressed to realize a wide spectrum of intelligence systems ranging from centralized terabyte and petabyte systems comprised of many large objects (e.g., images) to distributed heterogeneous databases that contain many small and large objects (e.g., text). The Community Management Staff's Massive Digital Data Systems (MDDS) Working Group[,] on behalf of the intelligence community, is sponsoring a two day invitation-only unclassified workshop on the data management of massive digital data systems with government, industry, and academia. [...] The objective of the workshop is to make industry and academia aware of intelligence community needs, stimulate discussion of the technical issues and possible solutions, and identify potential research efforts that warrant further investigation for possible government funding. The amount of funding estimated for investments is three to five million dollars per year over the next 2–3 years. [... Abstracts should be] related to the issues of the data management of massive digital systems including (but not limited to) scalability, architecture and data models, and database maagement functions. The focus of the abstract should be on potential solutions for the longer term research challenges (i.e., 5–10 years out) that must be addressed today in order to effectively manage data of massive proportions in the future." The call for abstracts also included a nine-page overview of the issues in massive digital data systems, including scalability of different database architectures (including distributed, parallel, extensible, and federated architectures), data models (multimedia, federated databases), and functions (querying in a federated environment, query processing in an extensible architecture, inference, data distribution, browsing, and filtering; long transactions, recovery techniques, replication, federated environments; access methods and index strategies (selective indexing, multimedia, storage structures, compression; metadata management; integrity constraints; real-time processing)).

Clearly the call for abstracts touches almost all areas of database research. I assume the workshop will be followed by some sort of request for proposals, perhaps from the National Security Agency; quite possibly the RFP will have a short deadline, and it might not be distributed widely. If I receive such an RFP, I will post it on-line immediately, so keep an eye on the SIGMOD Information Server. I do not currently have any additional information available, on or off line, about this program; you might wish to contact Dr. David Charnovia, the Director of the Advanced Technology Office, if you would like more information.

5 US National Security Agency

Another possible source to keep in mind if you work on database security is the US National Security Agency's INFOSEC program. Every year the NSA issues a request for proposals for its University Research Program (URP). The 1993 URP had a focus on multi-level secure distributed systems, with multi-level secure databases a tertiary concern. The URP announcement is issued every year in October, with proposals due in December. URP contracts range up to $60,000 or $100,000 (depending on area of focus) per year for two years, and include a number of strict reporting and eligibility requirements. For more information, contact Ms. Dorothy Darnauer, dld@tycho.ncsc.mil, phone (301) 688-0851, fax (301) 688-0255, at the National Security Agency.

6 NSF Research Equipment Grants

Every year, NSF solicits proposals for grants for research equipment. This year's deadline was February 1; the deadline for next year is October 1, 1994, so it's not too early to make plans.

"The objective of a grant for research equipment is to improve the quality or broaden the scope of the research and education that will be conducted at the proposing institution. Awards are for the purchase of new research equipment or for upgrading of existing equipment. Reasonable costs for the assembly and fabrication will be considered in cases where it is not possible to purchase "off the shelf" items that are needed for the research. Local computing equipment (including workstations, specialized processors, superminis, and local area networks) may be supported under this program. General-purpose office equipment is not eligible for support. Particular emphasis is given to those unique or new research capabilities that will ensue from the acquisition of the equipment. Although there are no specified minimum or maximium amounts for Engineering Research Equipment Grants, awards typically range between US$20,000 and US$200,000.

"Each proposal must include a paragraph indicating the institutional support proposed. Institutions must contribute at least one-third of the total cost of the equipment."

You can read more about this program through the SIGMOD Information Server. Follow the hyperlinks to NSF's full text search facility, and then type in "nsf93155". (You can also use the same approach to find out information about any NSF program.)

7 NSF Academic-Industrial Liaison Grants

Recent issues of the SIGMOD *Record* funding column have explained how Congress is pressuring the government funding agencies to focus on "strategic" research. One problem to be overcome in making this move is getting proposers interested in, and in fact simply aware of, strategic research issues: "There is a gap between the creation of new knowledge and technology in universities and its implementation by U.S. industry. There is also an inability of some university researchers to address industrially relevant issues without having access to industrial-scale equipment and facilities. These problems are some of the factors which are hampering the ability of the U.S. economy to produce competitive products and services in a timely fashion. It is becoming increasingly important that the research supported by the [NSF] Directorate of Engineering foster strategic links between academe and industry that will enhance the competitiveness of U.S. industry." Thus NSF has introduced programs to help with this problem by allowing faculty to spend time working with industry.

In the first program, Engineering Faculty Internships, a faculty member spends three to eight months in industry as a visitor. The visit is intended to provide a basis for future joint efforts between the faculty member and the visited lab. Up to US$25,000 may be requested, with indirect costs limited to 10%. The industrial sponsor must provide matching funds and a co-investigator. The participation of students is encouraged.

The second program, the Combined Research-Industrial Scholarship Projects, provides support for shorter faculty visits to industry that are part of a longer-term research project. The long-term project is three years in length, with 2–6 month industrial visits occurring at the beginning of the project, and the rest of the work occurring back at the university. "There is an opportunity to experiment and transfer the research results to industry by the end of the project."

Finally, the Industry-University Cooperative Research Projects program is intended to support longer-term projects, but with full participation of both faculty and industrial scientists.

All three of these programs look to support "long-term, generic, pre-competitive research. The investigators are encouraged to combine the research objectives with educational and human resources goals."

The target date for submissions was February 15, 1994. However, I'm not sure what a target date is, as opposed to a deadline; so you might want to contact NSF for more information if you are interested in submitting a proposal. If the deadline is truly past, you might think about plans for a future submission, as I predict that programs of this sort will become more widespread in the future.

You can read more about this program through the SIGMOD Information Server. Follow the hyperlinks to NSF's full text search facility, and then type in "iucrp". (You can also use the same approach to find out information about any NSF program.)

8 Other Recurring Programs at NSF

NSF has a number of programs that run every year and may be of interest to the database community. These include fellowships for beginning or second-year graduate students, research initiation awards, research opportunities for women faculty who have had a career interruption or are changing their area of specialty, and others that are best discovered by browsing, such as NSF programs for curricular development or retention of female students. I recommend that all strong beginning graduate students apply for NSF fellowships. In addition, all new faculty members should automatically apply for research initiation awards.

You can obtain more information about any of these NSF programs through the SIGMOD Information Server. Follow the hyperlinks to NSF's full text search facility, and then type in an evocative keyword or two. (You can also use the same approach to find out information about any NSF program.)

9 Recurring Programs at Rome Lab

Don't forget that the US Air Force has a number of recurring programs at Rome Lab that are relevant to database researchers, many of which will be accepting proposals for the next several years. Space doesn't permit us to describe all of these here; take a look at them on-line in the SIGMOD Information Server.

10 ARPA Research in High-Performance Computing

ARPA issued a very nice RFP in November that has doubtless received many submissions from the database community. The technical contact for the RFP was Randy Katz, and the topic area was high-performance computing software and software development environments. The announcement was well publicized (i.e., we got a copy early on and posted it on the SIGMOD Information Server), so all interested parties had a reasonable chance to submit before the January 7 deadline. We quote at length from the announcement, describing the four target areas of research:

"HPC Environments and Languages - to enable development of applications that can take advantage of the full range of HPC resources–workstations to large-scale HPC systems. Thrusts of current interest include: development of comprehensive HPC environments, compilers/optimizers,advanced languages, computational tools, program visualization/tuning aids, and program debugging aids. Projects are expected to focus on development of common architectures and frameworks to enable other research to build upon initial prototypes."

"Systems Software and Services - to provide a stable, effective operating system and system services upon which applications and new experimental services can be developed. Current research has shown that the boundary between application compiler services, application libraries, operating systems and system services is flexible, and that moving specific functionality between these components can have significant impact on performance and correctness. Efforts under this topic will develop new or extend existing operating systems and support services to include new functionality as well as to explore the design space of performance and correctness. Functionality of interest includes but is not limited to: [...] high integrity/privacy/resistance to software attack enabling development of highly trustworthy applications and private operations, high performance input/output, location independent naming and access to other services (i.e. processing and storage), real-time support for embedded processing needs, and providing services under dynamically changing resource availability or extremely limited resources (to include both fault tolerance and mobile computing needs). [...]

"Prototype Information Infrastructure Services - to provide a base for implementation of National Information Infrastructure applications. Conceptually these services are extensions of those developed in 'Systems Software and Services' above to include greater support for ubiquity, mobility, and embedded intelligence. [...] These may include (but are not limited to): services in support of information repositories and digital libraries based on techniques for scalable storage management, persistent object bases, multimedia objects, distributed and dispersed indexes, and inter-object linking, services in support of electronic commerce including transfer of money, authentication, audit trail, digital signatures, electronic solicitations and contracting, electronic certification of bidders, and intellectual property management, services that enable cooperation of applications on computer systems while managing the trust which is placed in remote systems, rich trust structures consisting of thousands of domains which may or may not have consistent views of trust, services that support location-sensitive and situation-aware processing by applications, such as using caching and location information to provide the illusion of connectivity and application tailoring to the users specific context.

"Application/Integrated Demonstrations - While it is expected that all projects will include a demonstration of at least minimal application support using the services built under this BAA, a limited number of focused demonstrations on specific applications areas may be cofunded in coordination with other agencies. [...] Examples of possible integrated demonstrations might be: semiautomatic categorization/cataloguing/retrieval of medical imaging, seismic, RADAR, or sonar data, or application of computational finite element or computational fluid analysis of a significant military or commercial vehicle/structure/chemical which could not be accomplished without HPC levels of computing."

11 National Center for Automated Information Research

As we mentioned last year, the US National Center for Automated Information Research has a recurring program that seeks proposals related to the legal and accounting professions. Funding ranges up to US$150,000. This year's deadline for submissions was February 1. For more information, refer to past issues of the *Record*. The past issues can be browsed on-line through the SIGMOD Information Server.

12 NASA Needs Help With Data Management

NASA has issued a new request for proposals, with submissions due at the beginning of March, under program number NRA-94-MTPE-02. (Although you may be reading this after the deadline for submissions, I recommend that you take a close look at the request for proposals, as there will undoubtedly be similar ones in the future.) The announcement requests proposals for information system technology relevant to NASA's EOSDIS (Earth Observing Systems The announcement is extremely lengthy, and can be obtained or browsed in its entirety with Mosaic at gopher://gopher.mtpe.hq.nasa.gov/pub/nra/94-mtpe-02. (I was also pleased to find NASA's new directory of on-line information servers at URL http://www.sti.nasa.gov/www.html.) We quote from the announcement below.

"The United States Government has initiated the U.S. Global Change Research Program (USGCRP) to develop a predictive understanding of the global environment. In the past decade or so, science and technology have reached levels that permit assessment of global environmental change. This assessment is an enormous and continuing task. No one scientist, laboratory, university, agency, or nation acting alone has the resources to develop the comprehensive understanding required. There is a collective responsibility to integrate the existing, diverse intellectual and technological capabilities, and to bring them to bear on the issues of global change.

"Progress in understanding global environmental change depends upon the integration and management of numerous data and information sources, extensive data holdings, and diverse data products. Yet, to achieve true success, this endeavor requires an information system that enables and stimulates cooperation among many researchers. The multiagency Global Change Data and Information System (GCDIS) must provide for the management of data, the sharing and harvesting of information, the dissemination of research results, and the establishment of a widespread community of collaborators in direct support of the USGCRP. The Earth Observing System Data and Information System (EOSDIS) is NASA's portion of this system.

"[...] EOSDIS will manage the data resulting from NASA's Earth science research satellites and field measurement programs, other data essential for the interpretation of these measurements, and information produced by these activities or from these data. It will also provide access to data held in the archives of other government agencies, organizations, and countries. EOSDIS will generate useful data products, and will facilitate the combination and manipulation of data from many sources as well as their incorporation into models of the global environment. Concomitant to fulfilling its data management functions, EOSDIS should encourage interdisciplinary research and assist in breaking down the intellectual barriers between the traditional disciplines of Earth science.

"EOSDIS will supply data archive, distribution, and information management services; for EOS satellites, EOSDIS will also provide data product generation and command and control functions. EOSDIS will provide data sets generated by assimilation of applicable observations into global climate models. To this end, EOSDIS must perform a wide variety of functions, supporting individuals located in various organizations and carrying on several distinct types of activities, including: Mission planning, scheduling, and control; Instrument planning, scheduling, and control; Resource management; Communications; Research and analysis at investigator sites; Generation of standard data products; Generation of special data products; Archiving of data, products, and research results; Data and information cataloging, searching, browsing, and ordering; Effective distribution of all data and information holdings; User support."

"Examples of areas of interest include, but are not limited to: data organization and access; management of large active archives; data characterization and content-based searching; visualization integrated into data querying; distributed computing; software development environments; software portability; and data quality assurance support. Proposals should demonstrate the repeatability and scaleability of the proposed technology development or adaptation (e.g. that it could readily be extended to additional Earth science data sets and be scaled up to accommodate the data volumes and performance requirements anticipated in EOSDIS)."

"It is estimated that the total annual research funding for this program will be \$3-5 million. NASA anticipates acceptance of approximately ten two-year proposals, with typical budgets of \$300,000 to \$800,000 per annum. Proposals may include options for a third year of activity. Projects of shorter duration, or lower budgets will also be considered. [...] Proposers should be willing to host in the second year of their investigation, as a working member of their team, someone from the EOSDIS implementation team, to facilitate technology transfer back to EOSDIS.

Participation in this program is open to all categories of organizations, both domestic and foreign: industry, educational institutions, nonprofit institutions, NASA centers, and government agencies. Proposals that demonstrate effective partnerships and cooperative arrangements, between a computer or information scientist principal investigator and a practicing Earth scientist co-investigator (active in global change research and a current or potential user of the EOSDIS), are encouraged."

The announcement goes on to give detail about the organization of NASA's internal data centers, policies, and plans. Individuals desiring additional information are referred to Dr. Dixon M. Butler at NASA, phone 202-358-1935, fax 202-358-2770, email dbutler@mtpe.hq.nasa.gov.

13 ARPA TRP Programs News

In the last issue of this column, we mentioned that ARPA's Technology Reinvestment Program had not awarded any of what one might call 'pure database' contracts. The list of all awardees has been released now, and can be browsed on-line at the SIGMOD Information Server. Perhaps the most database-like of the awardees is Enterprise Integration Technologies, a Palo Alto startup firm of former AI-ers who want to establish a National Information Infrastructure-style network in Silicon Valley.

14 Small Business Grants

The US Department of Defense has announced a new program for Small Business Technology Transfer (STTR). STTR is a pilot program which will fund small businesses, which, jointly with a research institution (a university or a federally funded research and development center) to perform cooperative research and development to bring research into commercial practice. The STTR is structured similarly to a Small Business Innovative Research grant (see previous issues of this column).

A booklet with details and forms is available from Defense Technical Information Center / Attn DTIC/STTR / Building 5, Cameron Station / Alexandria VA 22304-6415 USA; or phone (800) 225-3842 or (703) 274-6903. The next deadline is April 15, 1994. Topic include high performance computing and simulation, sensors and information processing, advanced human-computer interfaces, strategic planning tools bsed on object-oriented technology, and something intriguingly titled "Surprises and Opportunities".