4

Revolutionizing Earthquake Engineering Research Through Information Technology

The committee was asked to assess and comment on the possible roles of information and communication technologies for collaborative onsite and remote research, the sharing of data (including the need for standardization in data reporting), metadata (data about data), and simulation codes, and to identify additional research resources that are not currently available. This chapter discusses these roles, along with opportunities, challenges, and issues, from two temporal perspectives. The first is the short term, which coincides with the initial equipment and system integration phase of NEES (i.e., plans and work that are already under way). The second is the longer-term view, looking 10 years out (2004-2014) and thinking about what needs to happen in the future. The long-term goal of NEES is to revolutionize earthquake engineering research, not just to improve it incrementally. Success here will mean making significant inroads in advancing basic and applied research in support of the overarching grand challenge: ultimately preventing earthquake disasters. A fundamental objective of NEES, and the purpose of NEESgrid, is to change the paradigm so that earthquake engineering research within the NEES Consortium becomes a collaborative effort rather than a collection of loosely coordinated research projects by individuals.

Around the globe, governments and organizations are aggressively pursuing a vision of revolutionizing research and education in science and engineering by harnessing advanced emerging information technologies. Qualitative changes in the way that scientific research is conducted are well underway. Already, meetings, workshops, training, symposia, and reviews are routinely held electronically, using AccessGrid (ANL, 2002), a large-display collaboration environment that runs over the Internet and has participants around the world. The National Science Foundation's (NSF's) TeraGrid (NSF, 2002a) and its Middleware Initiative (NSF, 2002b), the Department of Energy's Scientific Discovery Through Advanced Computing (SciDAC) (DOE, 2002), the National Aeronautics and Space Administration Information Power Grid (NASA, 2002), the United Kingdom's e-Science program (DTI, 2002), and the European Union's DataGrid project (EU, 2002) all represent significant long-term investments in building up the systems, core technologies, and domain applications that will serve as foundations for the collaboratories of the future and a new era of science and engineering.

These developments build on a decade of activity that saw the emergence of the World Wide Web, the development of the first electronic collaboratories (Olson et al., 2001; SPARC, 2002), and the appearance of grid computing (Foster et al., 1999). The Grid embodies a pair of concepts: (1) the harnessing of distributed computing, data, and information resources in a seamless manner analogous to the electric power grid in support of the activities of (2) virtual organizations composed of geographically distributed people from multiple organizations and representing multiple disciplines. The name that the United Kingdom has chosen for its program, "e-Science," is particularly descriptive of what the committee believes NEES can accomplish with Grid computing: revolutionary science and engineering enabled by distributed computing resources for distributed groups of people. The substantial efforts mentioned above are made in the belief that the Grid can do for science and engineering what the Web has done for commerce, business, and information delivery to the general public.

NEES is a member of a rapidly growing list of long-term, focused disciplinary projects that spans high-energy physics, biology, astronomy, climate, social sciences, engineering, and many others. The NEES initiative was founded on the fundamental vision of establishing a collaboratory that would realize a new paradigm for earthquake engineering research. The idea is to foster a paradigm shift in the field, moving toward integrated physical testing, model-based simulation, integration and fusion of distributed data, and collaborative participation and interaction among geographically distributed researchers.

The NEES collaboratory vision is consistent with the ideas expressed in the 1999 report of the President's Information Technology Advisory Committee (PITAC, 1999) and with the many activities around the world described above. The systems integration component of NEES, NEESgrid (NEESgrid, 2002; Kesselman et al, 2002; Prudhomme, 2002), parallels many of the ideas expressed in the recent NSF report Revolutionizing Science and Engineering Through Cyber Infrastructure (Atkins et al., 2002), and has established the following IT-related goals for enabling the research, consulting, and educational communities:

- Perform teleobservation and teleoperation of experiments,
- Maintain a repository of curated data using standardized language and format,
- Access computational resources and open-source analytical tools, and
- Access collaborative tools for experiment planning, execution, analysis, and publication.

Substantive progress in preventing earthquake disasters will require multi-disciplinary research studies of unprecedented scope and scale. In particular, major advances will be required in the computational simulation of seismic events, wave propagation, and the performance of buildings and infrastructure—all of which will rely on extensive physical testing or observation for validation of the computational models. Results from these simulations will need to couple with building inventories, historical earthquake damage, and alternative build-out scenarios and will drive performance-based system designs, pre-event mitigation planning, emergency response, and post-event assessment and recovery. Ultimately, knowledge-based systems will be developed to support decision-making by policy makers and planners. Progress on these long-term objectives will rely on major advances in information technology:

- Accuracy and computational performance of large-scale simulations,
- Hybrid physical and model-based simulation,
- Coupling between multiple analytical models,
- Analysis and visualization capabilities for both experimentation and simulation,
- Data sharing and interoperability,
- Effective collaboration across disciplines and subdisciplines, and
- Knowledge-based and geographical information systems (GIS).

Long-term partnerships among researchers, practicing engineers, computer and computational scientists, and social and policy scientists will be key to success in these endeavors. Of even greater importance will be the education and training of the next generation of earthquake engineering talent.

FOUNDATIONS FOR NEES

Crafting a collaboratory like NEES involves the integration of a variety of enabling technologies, including general Web capabilities, mobile software (e.g., Java), Grid computing and, more recently, Web services (W3C, 2002). The Globus Toolkit (Globus, 2003) serves as one of the premiere examples of middleware for building Grid environments. The architecture of NEESgrid (Kesselman et al., 2002; Prudhomme et al., 2002) incorporates all of these current and emerging technologies as well as a top layer of new, generalized collaboratory software (Knoop et al., 2001) that has emerged from the building, operating, and assessment of scientific collaboratory environments over the past several years.

Obviously, IT is evolving and changing at a furious rate. It is important to note here that system integration is perhaps not the best descriptor for the NEESgrid activities. It is, of course, that—but it is also an *initial* development effort that will need to grow in capability, leverage new information and communication technologies as they emerge (which is often), and expand dramatically in applications and user interfaces over the next decade. Building a collaboratory is not yet quite the same thing as building a machine or an instrument. As an example, in the 2003-2004 time frame, the four general technology areas mentioned above will be coalesced into a new generation of Grid technology (Foster et al., 2002); provisions need to be in place to enable NEES to leverage such new developments and expand its functionality and usability. It is reasonable to expect the technology change and advancement ramp to be steep.

COLLABORATIVE ENVIRONMENTS AND DIRECTIONS

Information technology can enable collaboration in a variety of ways. E-mail, Web sites, and mailing lists provide asynchronous collaboration; these technologies are already widely and effectively used in this manner. Web-based collaboration environments expand on this level of capability and provide chat rooms, online data, and document sharing and have the potential to enable broad participation across the university, industrial, and education communities. Videoconferencing between individuals and groups is another important mode of collaboration. Dramatic reductions in the cost of entry have come about here, making these technologies available and practical for researchers and project participants with varying levels of infrastructure support. The past 2 years have also brought exciting and rapid advances in high-end, real-time, group-to-group collaboration, with the AccessGrid (Childers et al., 2000) serving as a prominent example. The AccessGrid environment leverages advances in network bandwidth and connectivity, commodity personal computer technology, inexpensive projection systems, and open-source software (PITAC, 2000). The AccessGrid has recently achieved remarkable penetration, especially in the academic community but also in the corporate and agency realms. There are now more than 100 sites around the world and perhaps double that number of actual systems (more than one system at each site). With its combination of capability, sense of presence, and growing availability, the AccessGrid is ushering in the beginnings of a revolution in how we meet, work together, and advance research goals.

NEESgrid is developing the collaboration environment by building Web-based portals that are based on collaboration frameworks developed at the University of Michigan, on electronic notebooks, and on Grid middleware. For videoconferencing (i.e., real-time human collaboration), NEES has adopted commercial solutions for the first stage, emphasizing simplicity and cost-effectiveness. For the additional tasks of telepresence, data manipulation, and analysis and visualization, the adoption of a primarily Web-based approach democratizes access to NEES. Cost of entry and requirements for local support and infrastructure are minimal with this approach and are key to gaining use and buy-in from the larger community.

FIGURE 4.1 An AccessGrid Session on NEESgrid.

Advances in IT will bring many opportunities for improved capability to NEES. AccessGrid has excellent possibilities for use in the NEES context and is already used by the Consortium Development group (See Figure 4.1). One challenge heard from the earthquake engineering research community was that of maintaining close communications and mentoring ties between students using the NEES facilities and their distant advisers. Live group collaboration environments such as the AccessGrid could be valuable in this regard. In addition, work is under way within the community to integrate additional collaboration capabilities, high-definition video, and analysis and visualization tools into the AccessGrid environment. All of these could one day be valuable to NEES, especially the high-definition video, which could be useful for teleobservation and telepresence activities.

Providing capabilities for teleoperation (i.e., remote control) of NEES resources has long been a stated goal of the effort, but it is a contentious issue. While collaboration capabilities in general appear to generate broad enthusiasm on the part of NEES sites and potential NEES users alike, it was the committee's observation that NEES sites were generally skeptical of the idea of remote control of the experimental equipment, citing deep concerns about security, varying levels of investigator proficiency, human safety, and the integrity of expensive equipment. The NEESgrid user requirements team delved into this matter in its survey (Finholt et al., 2002) of potential NEES users and found that a significant number of Ph.D.-level respondents asserted that teleoperation would be valuable to them. So the idea of enabling the remote operation of experimental equipment has some attraction, and from a technical standpoint, it is generally feasible.

Teleoperation in the NEES context is an idea worth exploring in a cautious and intelligent manner, bearing in mind security and safety considerations and that such a capability might be appropriate at some sites and for some purposes but not others. For example, there appear to be good reasons for enabling remote control for useful but safe operations, such as triggering measurements and orienting imaging devices (as opposed to large-mass or destructive operations). This type of remote control is sometimes referred to as telepresence. Teleoperation is already being explored as an educational tool (UCIST, 2002) and may also have interesting applications in training, which could be used to gain experience and familiarity with the feasibility of this mode of operation and the related security, safety, and technical issues.

MANAGING, CURATING, AND SHARING DATA

The sharing of data is an important underpinning for scientific collaboration. Success in this endeavor necessitates a number of capabilities: basic access to electronic data, common formats allowing data to be easily shared and reused, metadata standards that ultimately play a critical role in finding and using scientific data effectively, and Web-based data portals that provide sophisticated management and access functions for anyone with a commodity desktop system.

Many scientific communities have undertaken large-scale coordinated efforts aimed at developing broadly useful standards for data formats, software interfaces, and metadata conventions. In some cases, these activities have been going on for a decade or more. Metadata are essentially "data about data"—information about scientific data that is fundamental to discovering data, establishing their context, understanding their progeny, sharing them easily, using them correctly, and interoperating fluidly among diverse software systems. The Extensible Markup Language (XML) is emerging as a basic enabling technology and lingua franca that provides the glue for associating data and metadata, along with related information such as scientific papers, documentation, and the electronic services that provide access to data and metadata. Numerous examples of XML-based scientific markup languages have recently begun to emerge, spanning the gamut of science, engineering, and commerce. Bioinformatic Sequence Markup Language (BSML) (BSML, 2002), Chemical Markup Language (CML) (Murray-Rust et al., 2002), and Astronomical Markup Language (AML) (Oasis, 2001) are just a few examples drawn from hundreds. Comparable efforts in the earthquake engineering community are only in their infancy. Model Testing Markup Language (MTML) (Kutter et al., 2002) is an initial metadata effort for geotechnical physical model testing and the resulting data. COSMOS/PEER also has a project under way aimed at developing an environment for classifying, archiving, and disseminating geotechnical data over the Web (COSMOS/PEER, 2002).

The earthquake engineering community is just beginning to gain traction in the area of data standards and metadata, and effective strategies here are crucial to the data-sharing goals. With this operation in mind, NEES has established a task force devoted to factoring out common requirements for metadata across the NEES community. Through this task force it is developing metadata schemas, catalog services, and harvesting capabilities. This work is in its early stages but may serve as a catalyst for bringing the community together on important data issues. NEESgrid must track and leverage other efforts in metadata development and various efforts in developing databases and data services. For example, the concept of a National Spatial Data Infrastructure (NSDI) was first advanced by the Mapping Science Committee of the National Research Council in 1993. The Federal Geographic Data Committee, supported by the efforts of the NRC Mapping Science Committee, has been instrumental in fostering partnerships to encourage the documentation of data according to national standards to facilitate their sharing; and to encourage the use of geospatial data in new applications (NRC, 1993, 1995, and 2001).

Security and the protection of intellectual property are important concerns. In addressing them, NEES must have a data infrastructure that provides for flexible specification of access and management permissions. Over the course of its lifetime, a piece of NEES data would generally be assessed by populations ranging in size from individuals or small groups to large communities. Confidence and participation by investigators will hinge on the trustworthiness of these capabilities.

The NEES Consortium will be tasked with developing and maintaining a curated data repository. Curation will need to be undertaken on a level that goes beyond data integrity and persistence to data (and metadata) correctness. Careful consideration will need to be given to accomplishing this task, bearing in mind that discipline experts will probably need to be heavily involved. A voluntary program in this area may not be sufficient; instead dedicated resources may be required to achieve success. Similarly, much work will be required to advance data and metadata methods and standards; discipline expertise will be critical to long-term success. Financial support and reward structure require careful consideration and, possibly, proactive steps. That said, early indications are that NEES researchers have begun to contribute advances in these areas in a commendable fashion.

Looking toward the future, it can be seen that researchers who are engaged in grand challenge work will need to draw on data that span large-scale numerical simulation (e.g., basin-scale earthquake models), NEES resources (e.g., structural performance of individual buildings and collections of buildings), GIS (e.g., infrastructure, lifelines, buildings), and other data sources required for assessing event damage and appropriate responses. NEES, as a resource, is only one component in all of this. The NEES Consortium faces a challenge and enjoys the opportunity to foster partnerships in which other technology efforts and NEESgrid efforts collaboratively define the evolution of NEES data strategies in support of frontier research problems.

In addition, emerging work in translational strategies among metadata standards will ultimately need to be explored, possibly building on emerging work in the development of domain ontologies—a necessary step for the fusion of multiple disparate data holdings for integrative experimentation, simulation, and impacts studies (i.e. the grand challenges set out in this report). In this context, ontology is essentially a formal definition of the terms and relationships associated with a given domain. This concept forms the basis for a new area that has been receiving much attention as of late, the Semantic Web (Berners-Lee, 1998). Computer science researchers could carry out forward-looking work on the Semantic Web in the areas of information technology, digital libraries, and others. Examples of potentially synergistic efforts include the Alexandria Digital Earth Prototype Project (UCSB, 2002), the Digital Library for Earth System Education (DLESE, 2002), and the Geosciences Network (GEON, 2002).

From the foregoing discussion, it is clear that the NEES Consortium will develop massive quantities of data from both experimental and analytical research programs. Only if they are carefully managed and curated will these data be of use to the research community, policy makers, educators, and the general public. NEES will need to invest considerable effort in developing both the technology and the policies for storing, managing, and sharing these data. The key elements of a NEES data management and curation program are described below:

- *Raw project data.* Participants in a particular research project will need to share raw data as they are gathered from experiments. These data are not suitable for public distribution until they have been reviewed and processed into a form that is understandable by the research community at large. Thus, there must be a secure system for individual project researchers to share unprocessed data, work with the data, and convert them into a form that can be released for use by others.
- *Data for other NEES participants*. Researchers within the NEES consortium will want to access data from other NEES participants, particularly those working on projects with similar themes. A NEES data clearinghouse could create opportunities for such collaboration. The clearinghouse should be set up so as to encourage spontaneous collaboration among NEES researchers, allowing NEES data to be easily shared, compared, and combined.
- *Data for use by non-NEES researchers.* Once NEES data have been evaluated and reviewed, they should be published for use by researchers anywhere. A data repository would be created and data placed in the repository would be provided in standard formats. Unless formats can be defined in advance for researchers, the effort to convert project data into this standard format will be considerable and not within the budget of individual research projects. Consideration should be given to designating funding specifically for the development of standard NEES data formats, conversion of research data into the NEES formats, and curation of the data sets stored in the data repository.
- *Data for standards writers, practitioners, and educators.* In addition to being included in the repository of detailed data described above, research results should also be summarized and stored in a format that highlights the significant technical findings of the research. These summary data will be the most useful format for standards writers, practitioners, and educators. Again, consideration should be given to funding the considerable effort required to create a data synopsis for each NEES research project.
- *Data for policy makers, the press, and the general public.* The technical data repositories and summaries described in the categories above are neither intended nor suitable for use by policy makers, the press, and the general public. To maximize the impact of NEES research results, the NEES Consortium should generate public policy briefs, press releases, and educational resources for the general public. Development of these resources would be administered by a committee of NEES researchers, consortium managers, and public relations specialists.

In summary, ease of access to quality data developed by NEES and ease of collaboration among researchers over NEESgrid are among the most important aspects of the collaboratory and will strongly influence the success of NEES in ultimately preventing earthquake disasters. Significant efforts must be made to ensure that NEES data are of good quality and are released in a timely fashion. Major advances have been made over the past 30 years in experimentation in earthquake engineering, but generally speaking the data generated have only been available for use by the investigators who conducted the research.

BEYOND EXPERIMENTATION: SIMULATION, DATA ANALYSIS, VISUALIZATION, AND KNOWLEDGE SYSTEMS

One of the primary goals of NEES is to foster a movement toward integrated computer simulation and physical testing. Initial NEES efforts encompass identifying simulation codes that are of interest to the community and providing repositories and integrated execution sites for them so that participants may use them readily. NEES activities currently under way in data analysis and visualization focus on delivering analysis, simple visualization tools, and capabilities for accessing video streams and imaging from within the collaborative Web-based portal environment. NEES should also seek inputs from other disciplines in these technical areas to assist in the development of more appropriate tools and techniques. Initial dialogue has begun with the OpenSEES (PEER, 2002) effort, a framework for constructing simulation models, as an initial candidate for the simulation repository. Ongoing work in the development of community models—such as the Southern California Earthquake Center's (SCEC's) Southern California Velocity Model (Magistrale et al., 2000) and ground motion simulation tools (Bao et al., 1998), and, more generally, it's Community Modeling Environment—should be considered as well. SCEC has embarked on an ambitious program to develop physics-based models of earthquake processes and to integrate these models into a new scientific framework for earthquake hazard analysis and risk management. The Community Modeling Environment is under development at SCEC with an NSF Information Technology Research grant in support of the seismic-analysis and risk-management efforts. It will function as a virtual collaboratory for the purposes of knowledge quantification and synthesis, hypothesis formulation and testing, data conciliation and assimilation, and prediction. Given that the purpose of this modeling environment is entirely consonant and complementary with that of NEES, significant potential exists for collaboration between the two activities. Early dialogue between the NEES and SCEC communities should be strongly encouraged.

Success in addressing the grand challenge of ultimately preventing earthquake disasters will be intertwined with related grand challenges in information technology. Large-scale integrative simulation activities will push the envelope of what is possible both scientifically and technically. With terascale computational platforms already available and petascale systems on the horizon in 10 years, it will be technically possible to perform integrations of tremendous resolution for tsunamis and regional seismic events. Uncertainties about the source of seismic events and the soil material properties at the scale needed to model ground motion and system performance for frequencies of engineering interest make it necessary to introduce stochastic modeling. The requirements posed by analysis, visualization, and storage management will be formidable, perhaps even comparable to those posed by high-energy physics, cosmology, meteorology, and turbulence. Researchers will need new tools that scale to the complexity and size of the problem. These tools are in turn dependent on addressing the myriad data challenges. Management of the massive amount of information that will be generated by NEES experiments, field observations, and simulations was discussed in detail above. In addition, the visualization of this information will need to be a key component of the NEES effort. Visualization is essential to researchers, helping them to guide the design and execution of experiments and computer simulations. Moreover, with the huge amounts of data expected from NEES, the availability of tools for visualizing complex data sets will be crucial, allowing the researcher to interpret the results of experiments, observations, and simulations, which will in turn lead to the discovery of new results. Most important, sophisticated visualization tools will be essential for communicating the results and implications of the investigations to stakeholders, such as public officials and other policy makers, practicing engineers, students and teachers, and the public at large.

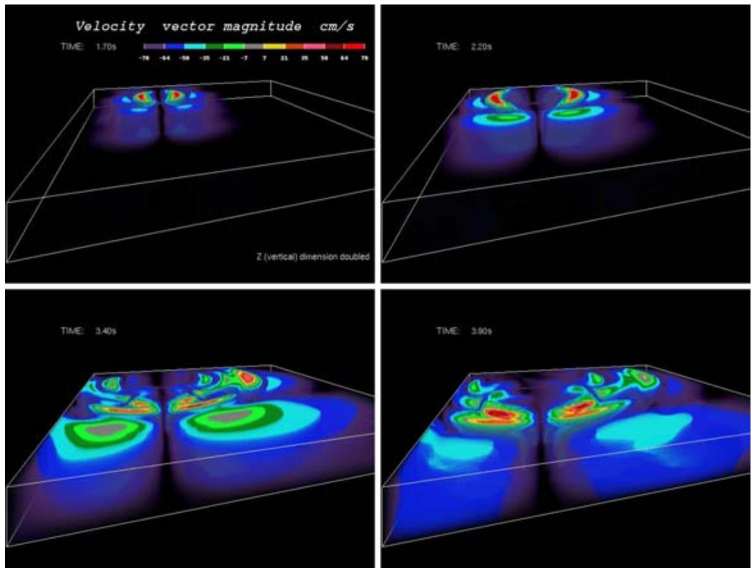

FIGURE 4.2 Visualization of the wave propagation in a layer over a half space due to an earthquake generated over an extended strike-slip fault. The distribution of the fault-parallel component of the horizontal velocity at different times following the onset of the excitation is shown. Source: simulation by Antonio Fernández and Jacobo Bielak, Carnegie Mellon University; visualization by Greg Foss, Pittsburgh Supercomputer Center.

The current NEES integration effort is addressing the need for storing and displaying visualizations, but except for demonstration projects such as shown in Figure 4.2, it is not addressing the need for visualization tools. Even though some generic tools are available commercially or are in the public domain, there has been little effort to date to develop a set of tools that will serve specifically the needs of NEES users in particular and the earthquake engineering community in general. For NEES to fully realize its promise, the development of suitable visualization tools needs to be explicitly encouraged and supported.

NEES has the potential to play a pivotal role in enabling frontier research activities. Through teleparticipation, the ready access to research results provided by NEESgrid, and the collaboratory nature of NEES in general, the social and policy sciences will be able to influence the course of research and application of the results. Earlier and better participation of the social and policy sciences in engineering research will aid in the development of the new loss estimation models and decision-making systems (King, et al., 1997) that will be needed for government and businesses to engage in effective pre-event mitigation and post-event emergency response and recovery. Ultimately, some of the problems posed will not fully submit to traditional analytical approaches, and a movement toward knowledge-based systems will be required if sustained progress is to be achieved. Such research endeavors would span the numerical simulation of seismic events and ground motion over large geographic regions; the simulation of tsunamis and the modeling of flood inundation; the physical and numerical simulation of infrastructure and building performance, event-impact prediction and assessment, collateral hazard analysis, damage evaluation, and emergency response. Rapidly evolving GIS capabilities will be invaluable for managing and analyzing the massive amounts of data that will be available from numerical simulations and other distributed databases. See, for example, Longley, et. al., 2001; Goodchild, et. al.,1999; and Greene, 2002.

BUILDING COMMUNITY

NEES is a national facility targeted at fundamentally redefining traditional modes of earthquake engineering research. Building on the concept of a collaboratory, NEES has an explicit charter to enable and broadly serve researchers at universities across the United States as well as practitioners and researchers at private corporations and government facilities. NEES also has a responsibility to contribute to the education of students, the continuing education of faculty, and the elevation of public awareness of earthquake engineering and earthquake hazards in general. Also, while NEES is a national effort, earthquake research is a global concern. NEES should be expected to play a long-term role in advocacy, partnership, and joint research with other national and international projects. NEES should demonstrate leadership that will not only advance U.S. research interests but will also serve as an example for other nations and programs and as a catalyst for enhanced international cooperation in pursuing mutual research interests. Furthermore, NEES potentially has the opportunity to transfer new technology to developing countries. Over the past 3 years, the United States has spent in excess of $50 million on direct humanitarian aid and disaster relief (USAID, 2000, 2001, 2002). Proactive investments in technology transfer to developing countries could generate goodwill as they mitigate these expenses, and they also could lead to export revenues for U.S. companies.

Establishing a new paradigm for earthquake engineering is a sociotechnical problem. Choosing IT foundations that promote ready participation across all of the interested communities will be key to fostering the participation, buy-in, and feedback processes that are critical to long-term success. Technology and opportunity alone will not necessarily galvanize the community in new modes of work. Engaging in research in a highly collaborative mode is rather new to the earthquake engineering community. It will require sustained community building and the development of trust—trust among people and trust in the technology that manages precious data and protects intellectual investments. This should be considered a role of paramount importance for the NEES Consortium.

EDUCATION AND OUTREACH

Earthquakes, tsunamis, and natural disasters in general are enormously relevant and interesting to society at large and to students in the classroom in particular. IT has an important role to play in enabling educational and outreach programs that would leverage NEES investments and enhance awareness and visibility. Web-based environments for posing questions, running simple idealized simulations, and even engaging in simulated disaster response management will offer exciting possibilities for projects that leverage NEES offerings.

Sidebar 4.1 Collaboratories, the Grid, Cyberinfrastructure, and the Future of Science and Engineering

The term "collaboratory"–a contraction of "collaboration" and "laboratory" – first came to light in a National Research Council report in 1989. The basic idea was (and still is) simple and compelling: to establish "laboratories without walls"–virtual electronic spaces where geographically distributed researchers can explore, learn, share knowledge, and partner and collaborate in order to solve hard science and engineering problems. This idea caught on nicely, and during the 1990s a number of collaboratory projects began in various agencies and problem domains. Perhaps one of the best known was the Upper Atmosphere Research Collaboratory (UARC) and its predecessor, the Space Physics and Aeronomy Research Collaboratory (SPARC). UARC is now a member of the Smithsonian Institution's permanent collection of information technology research.

Work continues in the development of collaboratories today, and there are now even collaboratories for studying collaboratories. To date, most of the collaboratories have been unique creations, each one carefully crafted for its given problem domain and scientific community. The earliest examples were even pre-Web. Obviously, such efforts were very challenging, not to mention expensive. It was pioneering work, and common sharable technologies and established methods generally did not exist.

The grand challenges in medicine, the environment, physics, and engineering demand the best and the brightest from all countries and continents–a perfect proving ground for the collaboratory idea. However, attacking many (probably most) of these difficult problem domains would require the use of a broad array of distributed computational resources: supercomputers, huge data archives, fast networks, sensors, and other complex instruments including spaceborne observation platforms. How could all of these systems be used in concert? Enter the Grid.

The 1998 book *The Grid: A Blueprint for a New Computing Infrastructure* focused on the concept of a new generation of computational capability, where geographically distributed resources were shared for virtual organizations. The concept of distributed computing was, of course, not at all new at that time and had its own long record of technological development and progress. What the new document offered was the broad concept of harnessing a collection of distributed computing resources for use by a virtual, distributed community. The choice of the term "grid" established a new metaphor: computing as a utility that one taps in much the same way as into a wall receptacle and draws on the national power grid without knowing how or where the power was generated or what entity was responsible for it. In its most general form, a Grid harnesses a collection of computing (data, storage, networking, instrumentation) resources that are geographically distributed and not necessarily under any form of centralized control.

Now, in 2003, we hear about all sorts of "Grids": access grids, data grids, computing grids, and a wide variety of grids for specific scientific and engineering endeavors. The avid reader will find articles about grid computing in the *Economist*, *Wired* magazine, and the *New York Times*. Businesses such as IBM and Sun Microsystems have made grid computing a prominent element of their long-term corporate strategies. So there is a lot of activity in this area and a number of different categories of Grid technologies and environments. But the long-term vision sees the Grid, a global interoperable fabric for computation and interaction that is ubiquitous and analogous to the Web. For this to happen, a global consensus must be reached on common protocols, interfaces, and services that can turn the vision into reality. Responding to this challenge, a new body called the Global Grid Forum (GGF) focuses on precisely these issues several times a year.

As mentioned above, grid computing is a metaphor for tapping the power grid, which is one important element of our societal infrastructure. Building our future collaboratories and knowledge environments is going to require an enormous amount of technological infrastructure, much more than we now have. Shared interoperating technology for federated data systems, digital libraries, visualization, collaboration, computation, security, and more will be needed. A new term has recently been coined to describe all of this: "cyberinfrastructure". One of the central ideas and, indeed, the potential key benefit of cyberinfrastructure is enabling much more rapid and cost-effective development of next-generation systems, environment, and applications. The Atkins Report (Atkins et al., 2002) put it this way:"If *infrastructure* is required for an *industrial* economy, then we could say the *cyberinfrastructure* is required for a *knowledge* economy."

Looking to the future, we can envision knowledge environments–new spaces where the collective experience and understanding of a global community can be synthesized, recorded, indexed, shared, and leveraged. If we can readily develop these for many problems and communities, there exists a unique opportunity to revolutionize the conduct of science, engineering, and education. Cyberinfrastructure, with grid computing as one component, will serve as a critical enabling technology foundation for framing the efficient and sustainable development of these new environments.

NEES sits somewhat uniquely at an intersection: As so aptly put by Kim Mish at a recent workshop, NEES is "where infrastructure meets cyberinfrastructure".

REFERENCES

Argonne National Laboratory (ANL). 2002. The Access Grid Project. Available at http://www.accessgrid.org [May 28, 2003].

Atkins, D.E., K.K. Droegemeir, S.I. Feldman, H. Garcia-Molina, M.L. Klein, P. Messina, D.G. Messerschmitt, J.P. Ostriker, and M.H. Wright. 2003. Revolutionizing Science and Engineering Through Cyberinfrastructure: Report of the National Science Foundation Blue-Ribbon Advisory Panel on Cyberinfrastructure. Arlington, Va.: National Science Foundation.

Bao, H., J. Bielak, O. Ghattas, L.F. Kallivokas, D.R. O'Hallaron, J.R. Shewchuk, and J. Xu. 1998. Large-scale Simulation of Elastic Wave Propagation in Heterogeneous Media on Parallel Computers. *Computer Methods in Applied Mechanics and Engineering*. 152(1-2): 85-102.

Berners-Lee, T. 1998. Semantic Web Roadmap. Available at http://www.w3.org/DesignIssues/Semantic.html [May 28, 2003].

Bioinformatic Sequence Markup Language (BSML). 2002. The Bioinformatic Sequence Markup Language (BSML). Available at http://www.bsml.org/i3c/Docs/BSML_Background.doc [May 28, 2003].

Childers, L., T.L. Disz, R. Olson, M.E. Papka, R. Stevens, and T. Udeshi. 2000. Access Grid: Immersive Group-to-Group Collaborative Visualization. In Proceedings of *Immersive Projection Technology*. Ames, Iowa: IEEE.

Consortium of Organizations for Strong-Motion Observation Systems and Pacific Earthquake Engineering Research Center (COSMOS/PEER). Archiving and Web Dissemination of Geotechnical Data Project Website. Available at http://geoinfo.usc.edu/gvdc/generalinfo.htm [May 29, 2003].

Digital Library for Earth System Education (DLESE). 2003. A community centered resource for anyone interested in learning about the Earth. Available at http://www.dlese.org [May 28, 2003].

Department of Energy (DOE). 2002. SciDAC—Scientific Discovery through Advanced Computing. Available at http://www.osti.gov/scidac/index.html [May 28, 2003].

Department of Trade and Industry (DTI). 2002. The U.K. e-Science Program. Available at http://www.escience-grid.org.uk/index.htm [May 29, 2003].

European Union (EU). 2002. The DataGrid Project. Available at http://eu-datagrid.web.cern.ch/eu-datagrid/ [May 28, 2003].

Finholt, T., E. E. Wierba, J. P. Birnholtz, and E.Hofer. 2002. NEESgrid User Requirements Document, Version 2.0. Available at http://guru.ncsa.uiuc.edu/neesgrid/NSFreview/NEESgrid_UR_Feb15.2002.pdf [May 29, 2003].

Foster, I., C. Kesselman, J.M. Nick, and S. Tuecke. 2002. The Physiology of the Grid: An Open Grid Services Archictecture. Available at http://www.globus.org/research/papers/ogsa.pdf [May 28, 2003].

Foster, I. and C. Kesselman. 1999. *The Grid: Blueprint for a new computing infrastructure.* San Francisco: Morgan Kaufmann Publishers.

The Geosciences Network (GEON). 2002. Cyberinfrastructure for the Geosciences. Available at http://www.geongrid.org/ [May 28, 2003].

The Globus Project. 2003. The Globus Project. Available at http://www.globus.org/ [May 28, 2003].

Goodchild, M., M. Egenhofer, R. Fegeas and C. Kottman (eds.), 1999. *Interoperating Geographic Information Systems*, Boston: Kluwer.

Greene, R.W. 2002. *Confronting Catastrophe: A GIS Handbook*, Redlands: ESRI Press.

Kesselman, C., R. Butler, I. Foster, J. Futrelle, D. Marcusiu, S. Gullipalli, and L. Pearlman. 2002. NEESgrid System Architecture Version 1.0.

King, S.A., A. S. Kiremidjian, N. Basöz, K. Law, M. Vucetic, M. Doroudian, R. A. Olson, J. M. Eidinger, K A. Goettel, and G. Horner. 1997. Methodologies for Evaluating the Socio-Economic Consequences of Large Earthquakes. *Earthquake Spectra.* 13(4).

Knoop, P.A. and D. Kiskis. 2001. CHEF: A CompreHensive collaborativE Framework. in Proceedings of *Global Grid Forum 1*. Amsterdam.

Kutter, B.L., D.W. Wilson, and J.P. Bardet. 2002. Metadata Structure for Geotechnical Physical Model Tests. 1st International Conference on Physical Modeling in Geophysics, July 10-22, St. Johns, Newfoundland.

Longley, P.A., M.F. Goodchild, D.J. Maguire, and D.W. Rhind, 2001. *Geographic Information Systems and Science*, New York: Wiley.

Magistrale, H., S. Day, R.W. Clayton, and R. Graves. 2000. The SCEC Southern California Reference Three-Dimensional Seismic Velocity Model Version 2. *Bulletin of the Seismological Society of America.* 90(6B): 65-76.

Murray-Rust, P. and H.S. Rzepa. 2002. The Chemical Markup Language (CML): A Position Paper. Available at http://www.xml-cml.org/information/position.html [May 29, 2003].

National Research Council (NRC). 1993. Toward a Coordinated Spatial Data Infrastructure for the Nation. Washington, D.C. National Academies Press.

NRC. 1995. A Data Foundation For The National Spatial Data Infrastructure. Washington, D.C. National Academies Press.

NRC. 2001. National Spatial Data Infrastructure Partnership Programs: Rethinking the Focus. Washington, D.C. National Academies Press.

National Science Foundation (NSF). 2002a. The TeraGrid. Available at http://www.teragrid.org/ [May 29, 2003].

NSF. 2002b. The NSF Middleware Initiative. Available at http://www.nsf-middleware.org/ [May 29, 2003].

NEESgrid. 2002. Available at http://www.neesgrid.org [May 29, 2003].

Oasis. 2001. Astronomical Markup Language. Available at http://xml.coverpages.org/aml.html [July 31, 2003].

Olson, G.M., D. Atkins, R. Clauer, T. Weymouth, A. Prakash, T. Finholt, F. Jahanian, and C. Rasmussen. 2001. Technology to support distributed team science: The first phase of the Upper Atmospheric Research Collaboratory (UARC). In *Coordination Theory and Collaboration Technology*, edited by G.M. Olson, T. Malone, and J. Smith. Hillsdale, NJ: Lawrence Erlbaum Associates.

President's Information Technology Advisory Committee (PITAC). 1999. Information Technology Research: Investing in Our Future. Available at http://www.nitrd.gov/pitac/report/pitac_report.pdf [July 29, 2003].

PITAC. 2000. Developing Open Source Software for High End Computing. Available at http://www.nitrd.gov/pubs/pitac/pres-oss-11sep00.pdf [July 29, 2003].

Pacific Earthquake Engineering Research Center (PEER). 2002. OpenSEES—Open System for Earthquake Engineering Simulation. Available at http://opensees.berkeley.edu/index.html [May 29, 2003].

Prudhomme, T. 2002. NEESgrid System Overview. Available at http://guru.ncsa.uiuc.edu/neesgrid/html/TR_2002/NEESSystemOverview2.doc [May 29, 2003].

Space Physics and Aeronomy Research Collaboratory (SPARC). 2002. The Space Physics and Aeronomy Research Collaboratory. Available at http://intel.si.umich.edu/sparc [May 29, 2003].

USAID (United States Agency for International Development). 2000. FY 1999 Annual Report of the U.S. Agency for International Development, Office of U.S. Foreign Disaster Assistance. Washington, D.C.: United States Agency for International Development. Available at http://www.usaid.gov/hum_response/ofda/publications/annual_reports/1999/1999annual-508.pdf [May 15, 2003].

USAID. 2001. FY 2000 Annual Report of the U.S. Agency for International Development, Office of U.S. Foreign Disaster Assistance. Washington, D.C.: United States Agency for International Development. Available at http://www.usaid.gov/hum_response/ofda/publications/annual_reports/2000/2000annual-508.pdf [May 15, 2003].

USAID. 2002. FY 2001 Annual Report of the U.S. Agency for International Development, Office of U.S. Foreign Disaster Assistance. Washington, D.C.: United States Agency for International Development. Available at http://www.usaid.gov/hum_response/ofda/publications/annual_reports/2001/2001annual-508.pdf [May 15, 2003].

University of California at Santa Barbara (UCSB). 2002). The Alexandria Digital Earth Prototype Project (ADEPT). Available at http://www.alexandria.ucsb.edu/

University Consortium on Instruction Shake Tables (UCIST). 2002. Instructional Shake Tables: A Cooperative Effort in Earthquake Engineering Education. Available at http://wusceel.cive.wustl.edu/ucist/ [May 29, 2003].

World Wide Web Consortium (W3C). 2002. Web Services Activity. Available at http://www.w3.org/2002/ws/ [May 29, 2003].