ITR/AP: The SCEC Community Modeling Environment—An Information Infrastructure for System-Level Earthquake Research

C.1. Problem Statement: The Need for Information Technology Research in Earthquake Science

Information technology (IT) offers new and unexploited capabilities for using the science of earthquakes to improve the assessment and mitigation of earthquake risks. In the last several years, notable advances have been made in two distinct areas of earthquake science: (1) the *dynamics of fault rupture*—what happens on a time scale of seconds to minutes when a single fault breaks during a given earthquake—and (2) the *dynamics of fault systems*—what happens within a network of many faults on a time scale of years to centuries to produce the sequencing of earthquakes in a given region. Both of these problems are highly nonlinear and coupled to one another through the complex processes of brittle deformation.

These physics-based simulations, which involve terascale computational problems, are crucial not only to gaining a fundamental understanding of earthquake phenomena; they also offer enormous practical benefits for assessing and mitigating earthquake risks based on seismic hazard analysis. Seismic hazard analysis seeks to describe, as precisely as possible, the type of shaking that can be expected at a given point of the Earth's surface due to earthquakes that are likely to occur over a specified time period. It requires the coupling of models that represent a series of complex processes:

1. *Hazard Characterization Model* (HCM): the identification and structural characterization of the faults capable of producing large earthquakes.
2. *Earthquake Forecasting Model* (EFM): the probability a particular fault will produce an earthquake of a given magnitude during the specified time period.
3. *Rupture Dynamics Model* (RDM): the dynamical description of motion across the fault during an earthquake as a function of space and time.
4. *Elastic Wave Model* (EWM): the propagation and interference of the compressional and shear waves along complex paths in the crust from an earthquake source to a target site.
5. *Site Response Model* (SRM): the excitation of ground motions in the near-surface layers at a target site, which, for strong ground motions, often involves significant nonlinearities.

The high-priority objectives of seismic hazard analysis are twofold: (a) to use physics-based numerical simulations to improve all five model types, and (b) to combine these models to produce products that include comprehensive catalogs of ground-motion simulations for use in risk assessment and earthquake-engineering analysis. Such catalogs are needed, for example, as input to research done at NSF's earthquake engineering research centers and its Network for Earthquake Engineering Simulation (NEES).

With this background, the basic problem addressed in this proposal can be stated as follows: the earthquake-science community is making great strides toward objective (a), but it does not yet have the IT infrastructure and capabilities for attaining objective (b). This task is complicated by a number of issues:

- <u>Heterogeneity and multiplicity of the models</u>. Many object types must be manipulated and many algorithms must be employed, but the pathway to a computational result is highly contingent on factors that must be evaluated at each analysis step along the way.
- <u>Distributed development of the models</u>. Models are being developed by different organizations with differing expertise, requiring data management and execution environments that span autonomous administration domains.
- <u>Variety of interaction modes</u>. Multiple access mechanisms are needed to support incremental development, interactive simulations, and coupling among models.
- <u>Requirements for model maintenance</u>. Evolution of the digital representation is driven by physical events, simulation output results, and remote-sensing input, requiring management of multiple versions.

These issues become particularly acute when "all hell breaks loose" during a big earthquake. Modern seismic information systems, such as TriNet in Southern California, can, within a few minutes, locate regional earthquakes and produce preliminary maps of ground shaking to guide emergency response [1]. The challenge is to replace the empirical maps, which are currently based on smooth interpolations of sparsely recorded motions, with more detailed maps generated using a RDM→EWM→SRM computational pathway, whose elements are described in bullets (3)-(5) above. The objective is to configure, automatically and rapidly, an appropriate set of modeling resources to assimilate seismic, geodetic, and geologic data as they are acquired in real time, create new products such as ground-motion and damage predictions, and distribute the output to multidisciplinary teams scattered across the region. These teams will then jointly visualize, manipulate, and modify the products and communicate the results to non-specialists—engineers, emergency managers, government officials, and the media. All of these operations will have to be done under potentially stressful conditions using distributed, multiply-connected computational systems that are robust to major regional disruptions in power, communications, and transportation.

It is proposed here to address the challenges of system-level earthquake research by creating an integrated framework for creating, executing, and managing computational pathways for earthquake simulations, such as the ones required to predict strong ground motions. Besides multidisciplinary earthquake science, this will require bringing together three distinct computer science disciplines: (a) knowledge representation and reasoning, (b) digital libraries and information management, and (c) Grid resource-sharing environments.

C.1.a. The Need for a Large, Integrated Project

The U.S. organization most active in coupling the advanced aspects of earthquake physics to seismic hazard analysis is the Southern California Earthquake Center (SCEC), a consortium of 40 universities and research organizations. The NSF, USGS, and other agencies support SCEC to gather new information about earthquakes in Southern California, to integrate knowledge into a comprehensive and predictive understanding of earthquake phenomena, and to communicate this understanding to engineers, emergency managers, government officials, and the general public. Recent FEMA estimates ascribe nearly half of the national earthquake risk to Southern California, with one-quarter concentrated in Los Angeles county alone [2]. SCEC thus serves a high-risk population of more than 20 million people as its regional center for earthquake information and coordinated earthquake studies. This coordination is essential to the development of the comprehensive data sets, consensus models, and consistent scientific judgements needed for public policy in earthquake risk management and mitigation.

SCEC has embarked on an ambitious program to develop the capabilities for numerical simulation in each of the five areas described above, and it has proposed a scientific framework for integrating these capabilities and applying them to the problems of earthquake risk assessment and mitigation (Fig. 1). Central to this program will be the construction of a *Community Modeling Environment*, in which the appropriate computational resources will be developed, documented, and maintained on-line for use by SCEC and other earthquake researchers. This environment will function as a virtual collaboratory for the purposes of knowledge quantification and synthesis, hypothesis formulation and testing, data conciliation and assimilation, and prediction. The SCEC Community Modeling Environment will greatly facilitate the system-level understanding of earthquake phenomena, and it has the potential to improve substantially the utilization of seismic hazard analysis in reducing earthquake losses.

We have formed a SCEC/IT Partnership to develop an advanced information infrastructure for system-level earthquake science in Southern California. Our partnership comprises SCEC, USC's Information Sciences Institute (ISI), the San Diego Supercomputer Center (SDSC), the Incorporated Institutions for Research in Seismology (IRIS, a 100-institution consortium), and the U.S. Geological Survey. The funding requested in this proposal will support four project elements:

- Fundamental IT research by ISI and SDSC on how to integrate knowledge representation and reasoning, digital libraries and information management, and Grid resource sharing into a methodology that will support the SCEC Community Modeling Environment.
- Application of this new methodology by SCEC and USGS scientists to seismic hazard analysis for the purpose of reducing earthquake losses in Southern California.
- Transfer of the methodology to other regions and extension to other Earth science problems by IRIS and (at no cost to NSF) the USGS.
- Use of products from the SCEC Community Modeling Environment to educate students at all levels and inform the general public about earthquake hazards.

The scope of this effort clearly requires a large, integrated project involving an extended collaboration among many disciplines and research organizations that spans both earthquake and computer sciences.

C.2. Geophysics Approach and Research

SCEC's basic research program will be supported primarily through its funding from NSF/GEO and the USGS. SCEC activities under the proposed ITR project will thus be focused in adapting its data bases and simulation engines specifically to the Community Modeling Environment, diagrammed in Fig. 1. The primary data bases include a Unified Structural Representation (USR) that combines a Community Fault Model (CFM) and a Community Velocity Model (CVM) to portray the surface and subsurface structure of Southern California. The USR will provide the structural information needed by the various simulation engines. These engines will include versions of the five community models (HCM, EFM, RDM, EWM, SRM) described in §C.1, appropriately wrapped for syntactical interoperability.

In some cases, there is a consensus methodology for the simulations (e.g., finite-difference computation of elastic-

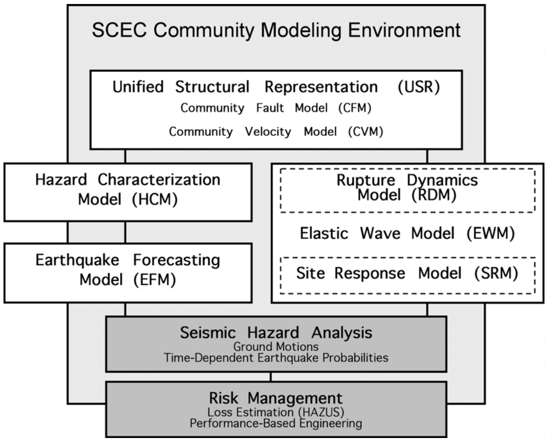

Figure 1. *Schematic of the SCEC Community Modeling Environment, showing the main areas for the development of community models (white boxes) and their areas of application (shaded boxes). The computational pathways flow from the top to the bottom of the diagram. The right side represents calculations for individual earthquakes; the left sided represents calculations for sequences of earthquakes. The Community Fault Model (CFM) and Community Velocity Model (CVM) are data bases embedded in a Unified Structural Representation. The Rupture Dynamics Model (RDM) and Site Response Model (SRM) are plug-and-play modules within a standardized Elastic Wave Model (EWM).*

wave propagation for the EWM), while in others the methodology remains the subject of active research (e.g., finite-difference, finite-element, and boundary-integral-element codes for the RDM), or there is choice of algorithms, depending on the geological situation (e.g., soil type for the SRM or the geometry of faulting for RDM). Working groups within SCEC will sponsor benchmarking and code intercomparisons, and they will wrap the competitive candidates for plug-and-play within the Community Modeling Environment. The availability of many modules with many configurations for each step in the simulation will provide a multiplicity of computational pathways to arrive at an "answer." But which answers are most suitable for the purposes at hand? Automating such decisions will be required for rapid earthquake response, as well as for generating the full ensembles (many thousands) of scenario simulations needed for seismic hazard analysis. Task-oriented decision automation of this type is a challenging ITR problem.

For the purposes of this project, the SCEC/IT partnership will focus on the two important computational problems diagrammed in Fig. 2. The first is the "earthquake scenario" pathway for calculating ground motions from a specified seismic source. The second is the "seismicity simulator" pathway for generating earthquake sequences within the Southern California fault system. Both of these pathways feed their output into seismic hazard analysis. Earthquake scenarios provide the *ground-motion catalogs*, which are needed to drive earthquake engineering simulations of structural response (e.g., for use in the NEES program), while seismicity simulations furnish *earthquake likelihoods*, which are needed for probabilistic seismic hazard analysis. The long-term goal of seismic hazard analysis is to employ the time-dependent probabilities from the latter analysis to structure the scenarios for the former.

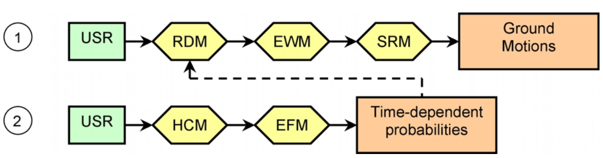

Figure 2. *Two computational pathways that will be the foci of the proposed project. (1) The earthquake scenario pathway to estimate ground motions. (2) The seismicity simulation pathway to estimate time-dependent earthquake probabilities. The dashed line indicates that (2) will eventually be used to drive (1). At each computational step, decisions must be made about the choice and configuration of the algorithm; hence, the multiplicity of pathways is very large.*

C.3. Computer Science Approach and Research

The proposed integration framework for composing community models to perform analysis tasks poses many difficult challenges from the perspective of computer science. First, we need to be able to characterize all the important aspects of available models so that models relevant for a specific pathway can be identified, and then the most appropriate model set selected, based on consideration of all pertinent constraints. This is complicated not only by the fact that the characteristics of the models are very complex and that their inputs and outputs might require syntactic and semantic translation and adaptation before they can be connected, but also, that there may be additional constraints imposed by the availability of resources needed to execute the model, or the pathway as a whole. For example, while there may be several models well suited for propagating high-frequency waves from a rupture on thrust fault, we may select one over another based on the computational requirements and the availability of a certain type of supercomputer with enough compute cycles available within the next 24 hours.

Second, once the appropriate choices are made, we must actually execute the computational pathway. This is complicated by the distributed nature of the model components, the necessarily distributed nature of the computing resources that are being used to execute the components, the potentially high levels of performance required by some modeling elements and the number, size, and variety of the data products being produced. Usability requirements demand that the complexities of interacting with both high-performance and distributed resources be minimized, while addressing the issues of executing within the presence of failure (exacerbated if this system is used for disaster management immediately after an earthquake).

Finally, we must be able to manage the results of pathway execution. Each pathway has the potential to generate very large amounts of data. These data will be of many diverse types, with complex relationships between the data products. To be useful, they must be organized and stored in such a way as to make it possible to support discovery and understanding of specific data elements, to understand the relationships between the data products from the current and previous analysis tasks, and to share the data with a broader community of interested researchers. Finally, we observe that repeatability requirements demand that we record as much information as possible about the process and starting conditions for the simulations, to be able to understand the relationships between the data elements produced by the analysis, and to be able to either regenerate the data exactly or formulate a new analysis task derived from the knowledge gained from a previously performed analysis.

We address these challenges with an integrated approach that brings together three distinct computer science disciplines: knowledge representation and reasoning (KR&R), Grid environments, and digital libraries and information management. Specifically, we propose to:

- Use knowledge representation techniques to capture the complex relationships between the physical processes and the model algorithms, between the algorithms and the simulation codes, and between the simulation codes and the data products. Knowledge reasoning techniques (e.g. inference) will be used to apply knowledge representations to the problems of task composition, execution, and data retrieval.
- Use Grid technologies to support the execution of analysis tasks in a distributed, heterogeneous resource environment.
- Use advanced digital library technology to manage both the information produced by the computational pathway, as well as information about the pathway and its execution.

The integration requirements addressed in this proposal are not unique to the area of earthquake research, but are becoming an increasingly important issue in all areas of large-scale information technology applications. While this framework will be developed within the context of earthquake research, many of the results will be of direct interest and applicability to other domains, such as climate modeling and earthquake engineering, as well as to the computer science community in general. We now touch on each of the major technologies in slightly more detail.

Work in the area of knowledge representation will follow two major thrusts. The first is to investigate *knowledge sharing and reuse of task-oriented (procedural) knowledge*. Previous research on ontology sharing and reuse has shown that this is possible in modest settings that integrate software components, specifically in the areas of image processing software and configuration design [3]. The scale and heterogeneity of the earthquake-science application will drive new research on knowledge-based techniques that support knowledge sharing and reuse of procedural knowledge. The second major challenge is to define an adequate representation vocabulary to characterize the pertinent aspects of the physical processes. The knowledge bases must also define how different simulation algorithms are related through spatial accuracy, time resolution, degree of convergence of the results, etc. We plan to build on recent work on unifying problem solving, planning, and process representations and reuse [4], and ontology-based mediation [5].

Important research issues to be addressed include: (a) How can the relevant physics and modeling assumptions implicit in simulation components be made explicit in a precise yet concise form to allow effective semantic interoperation between heterogeneous components? (b) Is it possible to develop an overarching common ontology (as the medical community is doing with the Unified Medical Language System (UMLS) and as some industrial and manufacturing groups are investigating), building on terms from the Community Modeling Environment? (c) How to enable users who are not experts in knowledge engineering to understand, build and extend the semantic annotations describing components developed by them and others? (d) How can we manage knowledge in the distributed and collaborative environment that will create the community models?

Grids technology is becoming an increasingly used infrastructure for sharing heterogeneous resources in distributed, multi-institutional environments [6]. As such, it is a natural target for the execution of analysis tasks produced by our integrated framework. Grids not only provide a means for remote access to resources such as computers, storage, data-sets, and software, but they allow resources to be aggregated, facilitating the execution of the extremely large problems often found in geophysics [7,8].

There are several aspects of the proposed integration that are challenging to current Grid environments, including (a) creating services for managing multi-version component models, (b) coordination of the end-to-end analysis task in a failure-prone environment, (c) development of Grid-enabled software architectures and development methodologies that enable the composition of independently developed model components into a end-to-end data analysis task, (d) integration of Grid data-management services into digital library collection management procedures, and (e) development of end-user tools that deliver Grid resources to non-expert users.

Digital libraries are currently being used to manage digital objects for a wide range of application domains [9, 10, 11, 12]. A current area of active research is the incorporation of knowledge into a digital library by mapping domain concepts from the digital library metadata attributes [13]. The goal is to support concept-based queries against the collection in order to identify relevant data sets that have specified relationships described within the knowledge base. A second issue we plan to pursue is how to maintain sufficient information about the construction and execution of specific computational pathways so as to facilitate query based on the means by which the information was produced in order to replicate the production of identical or similar data analysis.

Important research issues to be addressed include: (a) How can collection-based execution environments build upon existing digital library technology for organizing simulation and remote-sensor output within the Community Modeling Environment? (a) How can concept spaces be related across physical processes, instantiations of physical processes as simulation codes, and the resulting digital representations of the physical environment within which earthquakes occur?

C.4. Integration

A key element to our approach is a plan to couple these very different technologies. KR&R technologies need to capture information about resources and data, as well as correlate physical processes to simulation output, which necessitates integration with Grid and digital library technologies. Digital libraries will use the relationships inherent in the information they store to perform knowledge-based retrieval, which provides an interface to KR&R. Furthermore, the digital libraries must interact with the Grid environment to manage output files and to manage descriptions of the storage and communication resources needed to access data.

In addition to integration across information technology efforts, there must also be a tight coupling between information technology activities and the earthquake science. To help ensure that this takes place, we have structured the effort so that each major institution has participants in both information technology and earthquake science. We have also included funds in the budget to support an SCEC/IT Workshop to bring larger groups within each field together on an annual basis.

C.5. Knowledge Transfer, Education, and Outreach

SCEC maintains a very successful, nationally recognized Education and Outreach (E&O) program to transfer the knowledge gained by SCEC researchers to end-user communities, to educate students at all levels, and to engage the general public in earthquake-related issues. SCEC also coordinates its activities with its partners in earthquake engineering and risk management. The UCSD NPACI program maintains an E&O program to apply advanced technology in support of human resource development, increase participation of underrepresented groups, and promote national programs in education. We will combine the output of the proposed project, which will include many products of general interest (e.g., 4D visualizations of earthquake and ground-motion simulations), with the capabilities in both E&O programs to improve public understanding of earthquakes. These products should find extensive use in SCEC's excellent compendium of curricular and on-line educational materials.

The proposed project will produce new methodologies and analysis products of potentially great utility to other technical groups outside of Southern California earthquake science, and the inclusion of the USGS and IRIS in the SCEC/IT partnership will facilitate this technology transfer. SCEC and the USGS are collaborating in the RELM Project to erect a webportal for providing users with access to advanced methodologies in seismic hazard analysis [14], and the Community Modeling Environment will contribute substantially to this activity. The USGS will take the lead in applying the information infrastructure developed in this project to real-time operations and post-earthquake emergency response. The USGS will also ensure that the results of this ITR effort are exported to other regions of earthquake risk, as well as to other USGS activities. The development of a USR for Southern California provides a strong basis for SCEC/IRIS cooperation in the USArray component of the EarthScope project [15]. In particular, IRIS will ensure that the USR methodology developed by SCEC for the study of Southern California earthquakes can be transported to other regions and applied to other Earth-science problems, which include requirements for the representation of cratonic structures, orogenic belts, and active volcanic provinces absent in Southern California.

C.6. Management Plan

The management of this proposal will be built on the existing management structure of the Southern California Earthquake Center (SCEC). On December 1, 2000, SCEC submitted a proposal to NSF/EAR and the USGS for a 5-year continuation of Center activities. The University of Southern California (USC) will continue as the SCEC managing institution, with T. Jordan, the Principal Investigator of this proposal, as the new Center Director.

SCEC is an institutionally-based organization governed by a Board of Directors representing its core institutions. The fourteen core institutions currently enrolled in the Center renewal proposal are Caltech, Columbia, Harvard, MIT, San Diego State, Stanford, USGS Golden, USGS Menlo Park, USGS Pasadena, UCLA, UCSD, UCSB, University of Nevada, and USC. The Center Director acts as Chair of the Board and the representative of the managing institution (USC). The Center Director is also the Chief Executive Officer of the Center and bears ultimate responsibility for the Center's programs and budget. The Center has an External Advisory Council that serves as an experienced advisory body to the Board of Directors, which comprises a diverse membership representing all aspects of Center activities. Knowledge transfer, education, and public outreach are managed by a Vice-Director for Communication, Education and Outreach (CEO), who supervises a staff of CEO specialists.

In its new mode, SCEC will function as a matrix of activities with relationships approximated by the diagram in Fig. 3. Disciplinary committees in seismology, geodesy, geology, and rock mechanics will be responsible for planning and coordinating data-gathering and disciplinary infrastructure, including field programs, centralized data processing, and the distribution of data products. The interdisciplinary research more relevant to this proposal will be organized by four focus groups—Structural Representation, Fault Systems, Earthquake Physics, and Seismic Hazard Analysis. These project-oriented groups will be responsible for the development, verification, release, maintenance, and improvement of the SCEC Community Models.

The SCEC/IT Partnership will be managed by a SCEC/IT Steering Committee, which will oversee all aspects of this project. Members of the Steering Committee will include the P.I. and three Co-P.I.'s (C. Kesselman, R. Moore, and J.-B. Minster), representatives from IRIS (T. Ahern) and the USGS (N. Field), and the SCEC Information Technology Architect (to be named). The Steering Committee will be responsible for planning all Partnership activities and reporting to NSF on the results of the proposed project. The IT Architect will coordinate the day-to-day activities of the Partnership and will be responsible for managing the SCEC Community Modeling Environment, including the software standards for data structures and model interfaces.

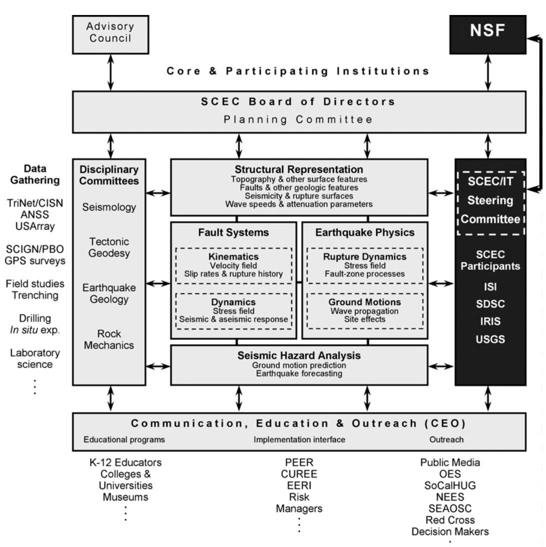

Figure 3. *Diagram showing the relationship of the SCEC/IT Partnership (black box) to other activities in the SCEC matrix (gray boxes). The Partnership will be managed by a Steering Committee that will be responsible for project planning and reporting to NSF. Within SCEC, four focus groups (central boxes) will organize the interdisciplinary research needed for the SCEC Community Modeling Environment. Knowledge transfer, education, and outreach will be managed through SCEC's CEO office by a CEO Vice-Director.*

REFERENCES:

1. D. Wald, V. Quitoriano, T. Heaton, H. Kanamori, C. W. Scrivner & C. B. Worden, *Earthquake Spectra* **15**, 537-556, 1999; http://www.trinet.org/shake.
2. *HAZUS®99 Estimated Annualized Earthquake Losses for the United States*, Federal Emergency Management Agency Report 366, Washington, D.C., September, 2000, 32 pp (http://www.fema.gov/pdf/FEMA366.pdf).
3. F. Fisher, S. Chien, E. Lo & R. Greeley, Using Artificial Intelligence Planning to Automate SAR Image Processing for Scientific Data Analysis, *Proceedings of the National Conference on Innovative Applications of Artificial Intelligence (AAAI/IAAI-98)*, 1998; J. Runkel, W. Birmingham & A. Balkany, Solving VT by reuse, *International Journal of Human-Computer Studies*, **44**, 1996.
4. W. R. Swartout and Y. Gil, EXPECT: A User-Centered Environment for the Development and Adaptation of Knowledge-Based Planning Aids, in *Advanced Planning Technology: Technological Achievements of the ARPA/Rome Laboratory Planning Initiative*, A. Tate (Ed), AAAI Press, Menlo Park, CA, 1996; Y. Gil & E. Melz, Explicit Representations of Problem-Solving Strategies to Support Knowledge Acquisition, *Proceedings of the Thirteen National Conference on Artificial Intelligence (AAAI-96)*, Portland, OR, August 4-8, 1996; J. Kim & Y Gil, Acquiring Problem-Solving Knowledge from End Users: Putting Interdependency Models to the Test, Proceedings of the Fifteenth National Conference on Artificial Intelligence (AAAI-2000), Austin, TX, July 30-August 3, 2000; D. Fensel, V. Benjamins, E. Motta, & B. Wielinga, UPML: A framework for knowledge system reuse. In *Proceedings of the International Joint Conference on AI (IJCAI-99*, 1999; F. Tissot & M. Gruninger, *NIST Process Specification Language*. Technical report, NIST, 1999.
5. H. Chalupsky. OntoMorph: A Translation System for Symbolic Knowledge, in A.G. Cohn, F. Giunchiglia, and B. Selman, editors, Principles of Knowledge Representation and Reasoning: Proceedings of the Seventh International Conference (KR2000), San Francisco, CA, 2000. Morgan Kaufmann; D. Pynadath, M. Tambe, Y. Arens, H. Chalupsky, Y. Gil, C., Knoblock, H. Lee, K. Lerman, J. Oh, S. Ramachandran, P. S. Rosenbloom & T. Russ, Electric Elves: Immersing an Agent Organization in a Human Organization, in *AAAI Fall Symposium on Socially Intelligent Agents: The Human in the Loop*, November, 2000.
6. I. Foster & C. Kesselman, *The Grid: Blueprint for a New Computing Infrastructure*, Morgan Kaufmann, San Francisco, 677 pp., 1999.
7. S. Smallen, W. Cirne, J. Frey, F. Berman, R. Wolski, M. Su, C. Kesselman, S. Young, Combining Workstations and Supercomputers to Support Grid Applications: The Parallel Tomography Experience, *Proc. 9th Heterogeneous Computing Workshop*, to appear.
8. K. Czajkowski, I. Foster, C. Kesselman, Resource Co-allocation in Computational Grids, *Proc. of the 8th IEEE Symp. on High Performance Distributed Computing,* 1999.
9. R. Moore, C. Baru, P. Bourne, M. Ellisman, S. Karin, A. Rajasekar, S. Young, Information Based Computing, *Proceedings of the Workshop on Research Directions for the Next Generation Internet*, May, 1997.
10. R. Moore et al., D-Lib Magazine, Collection-Based Persistent Digital Archives, March 2000, Volume 6 Number 3 [Part 1], D-Lib Magazine, April 2000, Volume 6 Number 4 [Part 2].
11. C. Baru, Managing Very Large Scientific Data Collections, poster session, *5th International Conference on High Performance Computing HiPC'98*, Dec. 17-20, 1998, Chennai, India.
12. C. Baru, R. Moore, A. Rajasekar & M. Wan, The SDSC Storage Resource Broker, *Proc. CASCON'98 Conference*, Nov.30-Dec.3, 1998, Toronto, Canada.
13. A. Gupta, B. Ludäscher, M. E. Martone, Knowledge-Based Integration of Neuroscience Data Sources , 1*2th Intl. Conference on Scientific and Statistical Database Management (SSDBM)*, Berlin, Germany, IEEE Computer Society, July, 2000.
14. The Regional Earthquake Likelihood Models Project is describe at http://www.scec.org/research/RELM.
15. EarthScope is an NSF initiative to employ new technologies for synoptic observation of the active tectonics and structure of the North American continent.(http://www.earthscope.org).