

ELECTRONIC SEISMOLOGIST

July/August 2003

Thomas J. Owens
Department of Geological Sciences
University of South Carolina
Columbia, SC 29208
Ph: 803-777-4530
Fx: 803-777-0906
Em: owens@sc.edu

CYBERINFRASTRUCTURE?

For a while, “software infrastructure” satisfied the ES’s “term dropping” needs. That gave way to “Information Technology”. But, everybody says that. The ES felt like a commoner. This month, Randy Keller, the ES’s embedded reporter with the GEON project, plops a real zinger on us. Cyberinfrastructure. The ES tried out dropping it at lunch today. It’s a showstopper. Read Randy’s article before you try to use it on your own, of course. He and the GEON team have clearly embraced the term and have begun, through a new NSF Information Technology Research grant, a multi-year adventure to build for the geosciences the cyberinfrastructure that we have been lacking for some time. Their approach is to address two major regional tectonic problems as the motivation for developing the necessary IT infrastructure applicable to a large segment of the geosciences community. Note the “big picture” questions that they seek to address. Then note their experimental approach: Don’t collect any more data! Shocking, isn’t it? Some would cry: Sacrilege! But, Fox Mulder would be pleased. The Truth is Out There. In our field books, on our disks, everywhere. And, therein lies the challenge. The GEON group isn’t really saying don’t ever collect more data; they are just advocating the optimal use of the data that has already been collected. GEON seeks to develop the cyberinfrastructure necessary to allow all of us to access, synthesize, and model geoscience data from a wide variety of sources and seek the answers to major geoscientific questions more efficiently. Thanks to the GEON team for taking the time to make ES readers aware of their plans. The ES looks forward to further reports as the project evolves.

GEON (GEOscience Network)—A FIRST STEP IN CREATING CYBERINFRASTRUCTURE FOR THE GEOSCIENCES

G. Randy Keller, Reporter for the GEON Team (www.geongrid.org)
Department of Geological Sciences,
University of Texas at El Paso,
El Paso, TX 79968
Ph: (915) 747-5850
Em: keller@utep.edu

A new buzz-word that we will be hearing in the future is cyberinfrastructure. This term refers to the information technology infrastructure that is needed to: 1) manage, preserve, and efficiently access the vast amounts of earth science data that exist now and the vast data flows that will be coming online as projects such as EarthScope get going; 2) foster integrated scientific studies that are required to address the increasingly complex scientific problems that face our scientific community; 3) accelerate the pace of scientific discovery and facilitate innovation; 4) create an environment in which data and software developed with public funds are preserved and made available in a timely fashion; and 5) provide easy access to high-end computational power, visualization and open source software to researchers and students. The task of creating a cyberinfrastructure for the geosciences is daunting due to the large volume and diversity of our data, as well as, the extreme differences in data formats, storage and computing systems, as well as differing conventions, terminologies, and ontological frameworks across disciplines. One way to think about this is that the ultimate goal is provide you with the tools and data that you need to do better, more creative science by minimizing the effort needed to look for data, research the background of a topic, and make software run properly. Another consideration is that when data and information are entered into an organized system, they can be easily found and unexpected relationships can be discovered via queries in a Google-like fashion. Think about discovering many of the relationships between phenomena that led to plate tectonics in days instead of years. Within the Division of Earth Sciences of the National Science Foundation, the effort to create the cyberinfrastructure that we need is referred to as *Geoinformatics* (www.geoinformatics.org) and arose out of a series of meeting and workshops, as well as, our community's energetic response to Information Technology Research (ITR) opportunities at NSF.

One manifestation of the ITR program is the GEON (GEOscience Network) project that was recently funded by a large ITR grant. The focus of GEON is the pressing need in the geosciences to: 1) craft the many relatively raw data sets in the Earth Science community into mature databases that can grow and evolve; 2) interlink and share these multidisciplinary databases; 3) create a robust toolbox of open source software for analysis, modeling, and visualization; and 4) provide the information technology (IT) infrastructure to manage and explore a highly distributed and diverse network. GEON is a true partnership between Computer Science and Earth Science researchers, and the scientific goal of this project is to facilitate efforts to understand complex problems focusing on the 4-D structure and evolution of continents. To rise to this challenge, we formed a coalition of researchers with key Computer Science expertise and researchers representing a broad cross-section of Earth Science sub-disciplines. The creation of GEON is a first step in developing the critical cyberinfrastructure necessary to achieve the vision of *Geoinformatics* and facilitate other research initiatives, in particular *EarthScope*. GEON is working closely with organizations such as IRIS, the U. S. Geological Survey, SCEC, and UNAVCO as well as other IT efforts within the Earth Science community. In particular, the U. S. Geological Survey has joined as a major partner and has made creation of key GEON-related databases a priority effort over the next several years.

Creating the GEON cyberinfrastructure to integrate, analyze, and model 4D data poses fundamental IT research challenges due to the extreme heterogeneity of geoscience data formats, storage and computing systems and, most importantly, the ubiquity of "*hidden semantics*" and differing conventions, terminologies, and ontological frameworks across disciplines. As the prototype for a national cyberinfrastructure in the geosciences, our guiding principle is embracing heterogeneity at all levels—hardware, software, networking, as well as information

structures. Creating this infrastructure involves basic as well as applied research in Information Technology and use of state-of-the-art information technologies. GEON IT research focuses on modeling, indexing, semantic mediation, and visualization of multi-scale 4D data, and creation of a prototype *GEON Grid*. An important contribution will be embarking on the definition of a *Unified Geosciences Language System (UGLS)*, to enable semantic interoperability. For example, the P_n phase means different things to controlled source and earthquake seismologists, and think of all the possible answers to a Google-like query about the magnitude of an earthquake and of all the places where one might find this information. We will create a portal to provide access to the GEON environment, which will include advanced query interfaces to distributed, semantically-integrated databases, Web-enabled access to shared tools, and seamless access to distributed computational, storage, and visualization resources and data archives. Various GEON-like grid efforts, such as GriPhyN, NEESGrid, and BIRN, all indicate the readiness of the Computer Science community to provide the necessary interoperable infrastructure, and testify to the value of integration of IT with major science and education initiatives.

In order to insure that the scope of GEON was manageable, linkage and refinement of existing and emerging databases is being emphasized, and two testbeds (the mid-Atlantic and Rocky Mountain region) were identified to focus the GEON geoscience research effort geographically. These regions were selected due to the variety of geological issues embodied within them that require integration of multi-disciplinary database, and because they are areas of expertise for the GEON geoscience research team. The ultimate goal of GEON research is to significantly impact large multi-scale geoscience research programs such as *Earthscope*, as well as individuals and smaller groups of researchers with the goal of facilitating the development of a culture in which data are shared, archived and rapidly disseminated across our discipline, much like IRIS has done within seismology. Many disciplinary geoscience database projects are already underway, indicating the readiness of the community to participate in such a national-scale effort. By facilitating the use of large and diverse data sets, we believe that the scientific community will make major scientific discoveries and create new and exciting scientific paradigms that lead us into the post-plate tectonics era.

Test Bed 1: Mid-Atlantic Region

The mid-Atlantic region records a complete Wilson cycle that starts with the breakup of eastern continental North America with the development of a passive margin and formation of an ocean basin (570-480 Ma—million years ago), includes development of convergence related island arcs (formed on multiple terranes?), and ends with their collision with the passive margin (460 Ma) followed by continent–continent collision as the final phase of the Appalachian orogeny (300-260 Ma). This cycle started over by the breakup of eastern North America again at approximately 210 Ma with the development of a passive-margin and ocean-basin formation that continues today. Evidence of the end of a preceding cycle is also preserved in the form of the Grenville orogeny. The complex geological events that accompanied this Wilson cycle as a function of time are well preserved in the Mid-Atlantic test bed that extends from Virginia to eastern Pennsylvania and New Jersey. Thus the area provides an ideal opportunity to attempt to reconstruct the geological processes involved through multiple time slices leading to a 4D assessment of continental growth. Some key research questions for this area that we will address within the context of integrated analysis of diverse geological and geophysical data sets are:

(1) Significant continental growth and modification arises by addition of ‘foreign’ crustal fragments, currently called terranes. Can these be unambiguously identified? What attributes can be recovered from geological record to assign terranes? Can terrane boundaries be identified in 3D space and through time and what is the nature of these boundaries? Are old terrane boundaries the locus of later reactivation? How do such terranes become a ‘permanent’ part of a continent? Are terranes accreted piecemeal or are they assembled and then accreted as larger superterrane units? Are there long-term differences in crustal character between juvenile terranes and those comprised of older continental crust after they have been incorporated into the continent?

(2) Links between foreland and hinterland evolution should help us assess the relationship of orogeny to the preserved stratigraphic record as well as use the stratigraphic record to help decipher orogenesis. Is the sedimentary record in foreland basins related to the geomorphology of the hinterland and to climate? Can details of the environmental history of depositional basins obtained through careful faunal analysis as well as sedimentological interpretations be related to orogenic processes? Do major sediment transport corridors and accumulation sites reflect underlying structural or other features related to deformation caused by the tectonic processes in the area? Do thick sedimentary accumulations in basins influence later deformation?

(3) In this test bed area, the early Paleozoic is characterized by thick-skinned deformation; this evolved to thin-skinned deformation in the late Paleozoic. What caused this transition or change? Was it linked with one or several identifiable tectonic event(s)? What in the record identifies this change?

(4) Rates of collisional and extensional tectonic processes appear to vary over time. How do these rates vary with time? Are variations in rate related to identifiable tectonic events such as a major change in convergence vector? Addressing this issue requires linking information in the geological record, which commonly is punctuated, to models that assume tectonic processes are continuous rather than sporadic.

Investigating continental growth by terrane accretion as part of the Wilson Cycle provides an example of how we will utilize the data integration and visualization capabilities of GEON. Disciplinary databases need to be integrated on spatio-temporal attributes as well as by their “process context” (i.e., how they relate to the Wilson Cycle). In the mid-Atlantic region, continental growth through accretion of approximately 12 terranes has been proposed, but current models do not agree on criteria for identifying the terranes and present various accretionary histories. To assess this mechanism of continental growth, the models must be constrained by multidisciplinary observational data sets within the region that will be assembled by the GEON geoscience team, including members of the U. S. Geological Survey.

This test bed area occupies a key position in the geographic center of the Appalachian mountain belt. As the project evolves it will provide a template for extending the information and interpretations derived both to the north into New England and maritime Canada, and to the south into the southern Appalachians. In addition to understanding the processes that control the growth of continents, this focus area will also provide insights into the interpretation of modern geologic features e.g. relationship of modern earthquakes and drainage systems to paleostructures and the influence of intraplate deformation and erosion on development of present day landforms.

Test Bed 2: The Rocky Mountain Region

The Rocky Mountain region is the apex of a broad dynamic orogenic plateau that lies between the stable interior of North America and the active plate margin along the west coast. For the past 1.8 billion years, the Rocky Mountain region has been the focus of repeated tectonic activity, and has experienced complex intraplate deformation for the past 300 million years. This activity has included the formation of numerous, often superimposed, basins that contain a rich stratigraphic record and some of the prime fossil localities in the world. Thus, it is an ideal area to investigate the interaction between tectonics and basin evolution in an Intraplate setting. The Continental Dynamics program at NSF is presently funding two large projects, the CD-ROM and Yellowstone experiments that are producing large volumes of data focusing on this region. In concert with these projects, GEON is addressing questions such as :

- (1). What is the nature of the processes that formed the continent during the Proterozoic?
- (2). What is the Influence of old structures on the location and evolution of younger ones?
- (3). What were the processes at work during the numerous phases of intraplate deformation, including the Yellowstone region?
- (4). What were the processes that caused uplift of the mountains and high plateaus that are seen in this region today?
- (5). What are the effects of mountain building on the distribution of mineral, energy, and water resources?
- (6). What is the nature of interactions among Paleozoic, Laramide, and late Cenozoic basins?

The modern topography of the Rocky Mountain region dates back at least to the Laramide orogeny, which probably represents the best-documented example of basement-involved foreland deformation on Earth. This impressive and enigmatic intracontinental restructuring (reactivation?) event was mostly superimposed at high angles to the pre-existing lithospheric architecture. The mechanisms controlling basement-involved foreland deformation, including the role of basement anisotropy and lithosphere-asthenosphere interactions, remain poorly known. There is no agreement on how plate-tectonic processes, both west of and underneath the Rockies, connect with the complex pattern of shortening in the upper crust. In fact, even the largest-scale geometries of crustal deformation are controversial due to a lack of integrated analysis. Recent hypotheses have predicted everything from lithospheric buckling and upwarping the Moho, to ductile crustal intrusion. The depth of disagreement is surprising considering the quantity and quality of available geologic and geophysical data from this region.

The Laramide orogeny was followed by an extensive phase of volcanic activity in Colorado, New Mexico, and west Texas that was in turn followed by formation of the Rio Grande rift in this same region. During all of this activity, the Colorado Plateau was only modestly affected, and its resistance to deformation and its uplift history are also major topics of scientific debate that we will address, using the computational power and integrated database infrastructure of GEON. Shallow-marine rocks exposed on the Colorado Plateau indicate that this plateau was near sea level during much of the Phanerozoic, but it stands ~2 km above sea level today. Substantial crustal thickening may be involved in the uplift, but geophysical models disagree about thickness of the crust today and an integrated analysis is needed to resolve this issue. Various mechanisms have been proposed, including magmatic injection, displacement of the lower crust, ductile channel flow from the uplifted Sevier hinterland, or change in the mantle thickness or density. Most of these models assign a major role to the hypothesized flat

subduction of the Farallon plate during the Laramide orogeny, but the cause of the flat subduction and the geodynamics of the coupled Farallon-North America plates remain poorly understood. Overall, our emphasis on Laramide and post-Laramide tectonism offers the potential to study active processes in the western U.S., link these to emerging images of the structure and evolution of the Proterozoic lithosphere, and hence decipher lithospheric-asthenospheric processes through time.

While computer simulations have provided useful insights into some of these problems, the lack of both computing power and integrated observational constraints has limited previous models to various 2D simplifications. Many of these limitations will be removed by GEON activities. For this case study, we will develop sets of fully 3D geodynamic models to investigate the lithosphere-asthenosphere interactions and crustal deformation during the Laramide orogeny, and their control on uplift of the Colorado Plateau. The grid computing power of GEON—including desktop workstations and cluster computer nodes—will allow innovative numerical approaches in the proposed modeling. For example, multiple faults are difficult to include in traditional finite element (FE) modeling of continental deformation. New techniques are being developed to incorporate fault networks in FE models by decomposing the model domain along fault surface. This allows independent finite element discretization of blocks across the fault, which can then be solved using parallel and grid computing techniques.

The databases that we plan to employ include:(1) temporal and spatial variations of strain and strain rates, including GPS measurements, slip rates estimated from Quaternary fault movement, and strain history from structural analyses; (2) stress field, both present stresses derived from in situ measurements and earthquake focal mechanisms, and paleo-stresses inferred from fault development and measurement of shear strains; (3) uplift history, both with respect to sea level as inferred from proxies of paleoaltimetry, and exhumation of crustal rocks as indicated by thermochronologic and petrologic data and sedimentary records; (4) history of regional crustal deformation and three dimensional tectonic boundary conditions inferred from geological reconstructions; (5) thermal history, including present distribution of surface heat flow and paleo-heat flow and volcanism; and (6) lithospheric structure, as indicated in seismic velocity, seismic tomography, seismic anisotropy, gravity anomalies, and other geophysical data.

Summary

In the case of both test beds, our goal is to pursue an ambitious research agenda in Earth Science stressing integrated studies while working with Computer Science colleagues to create a cyberinfrastructure that pushes the envelop in their field. We hope to create an environment in which researchers will be able to see benefits for their personal efforts and will want to contribute data, software, and ideas. You can stay tuned by visiting our website.

For more information see:

NSF Geosciences Beyond 2000 - <http://www.geo.nsf.gov/adgeo/geo2000.htm>

Background on the Geoinformatics initiative - www.geoinformatics.org

New report on Cyberinfrastructure - http://www.communitytechnology.org/nsf_ci_report/