GIS Ireland 2004, Dublin, Irish Organisation for Geographic Information (IRLOGI)

The New "Explornography" in the Age of Digital Earth, Digital Government, and Cyberinfrastructures

Dawn J. Wright
104 Wilkinson Hall, Department of Geosciences, Oregon State University, Corvallis, OR, USA 97331-5506, dawn@dusk.geo.orst.edu, http://dusk.geo.orst.edu

Abstract. What might the concept of exploration and the notion of discovery mean to geographers and GIS practitioners today? Exploration of our planet through fieldwork, and hence discovery of new places, is still ongoing, but so also is the exploration of environmental databases, even of information spaces that do not necessarily include spatial data. Therefore, "discovery" of a place does not necessarily mean having to "be there" in the field. Presented in this context are the themes of data sharing and the benefits thereof in the United States, and the emergence of cyberinfrastructures (i.e., the use of high-end information technology in day-to-day activities, not just for the occasional supercomputer job), which are taking hold in basic and applied research, but also within the realm of "digital government." Under the umbrella of cyberinfrastructures, exciting new research topics are being developed in the areas of web GIS (e.g., modeling, algorithms, data structures, stability, performance, and other computing issues), ontological libraries and semantic interoperability within web GIS, and networks of data and metadata clearinghouses that are being built with open specification web mapping services and web feature services.

Introductory Terminology

"Explornography" is a term first coined by John Tierney in a 1998 *New York Times* article, in which he defined it as "the vicarious thrill of exploring when there is nothing left to explore." His discussion of the term was actually meant to be a critique of the Peary expedition to the North Pole in particular and in general of some forms of extreme tourism to exotic or dangerous places (clearly beyond the intended scope of this paper). But if one extends this beyond the notion of just exploring physical places on the Earth's surface, one can think of exploration and discovery in a new way. We are now in what many call a "second age of discovery," where virtual worlds of real and imagined phenemona may be explored through computers on a desktop, in large visualization theaters in small handheld devices or soon even through small devices on our clothing or eyewear. But thankfully there is still much left to explore physically. For example, in terms of the surveying and mapping of the Earth's surface, very little is still known about the fine-scale topography and structure of the global seabed. There now exists satellite altimetry covering all the world's oceans, from which low-resolution bathymetry can be derived. But slower, more spatially-restricted shipboard measurements must still be made at sea in order to gather the higher resolution data required for tectonic studies or the baseline, framework data sets needed for a host of applications from marine cable laying to conservation of marine protected areas. Still only 35-40% of the entire Earth's surface (including the seabed) has been mapped at a similar resolution of a common hiking map, or of topographic maps of other planets such as Mars and Venus.

So in our quest to build a "digital *Earth*," the global access to all possible geographic data about a place on the surface and the subsurface, researchers and practitioners face may enticing challenges, including the development of visualization systems with user-friendly interfaces that enable the analysis, modeling and simulation of data, as well as just the simple viewing of it.
For several years the U.S. National Aeronautics and Space Administration led the Digital Earth Initiative that included the development of a prototype visualization system, a large globe that a user could manipulate with special gloves and glasses, "a very visual Earth explorer that lets scientists, both young and old, examine information about the Earth to learn how the forces of biology and geology interact to shape our home planet." The initiative has evolved more into a data sharing, data standards enterprise, under the purview of the Geospatial Applications and Interoperability Working Group within the U.S. government's Federal Geographic Data Committee as part of the National Spatial Data Infrastructure. It is thus an example of helping to build the second age of discovery through geographic information science, recognizing that technologies give rise to questions about their appropriate and more efficient use, questions that need theoretical frameworks in order to be solved. For example, interoperability is one of many research topics that geographic information science, computer science and other communities still grapple with. At times we pay the price for building technology in the absence of good theory.

In the U.S. the term "cyberinfrastructure" is being used with greater frequency to refer to how the traditional modes of science scientific research (e.g., experimentation in the lab, observation in the field, processing/analyzing on a single calculator or computer, calculating on the back of an envelope) or being extended or replaced by information networks. Indeed, as physical "infrastructure" has represented roads, bridges, railroad lines, power grids, etc. as fundamental components of modern communities, cyberinfrastructure now refers to the fundamental components of modern scientific and engineering methodologies (i.e., information technology, digital communications, and distributed computing). As stated by a recent blue ribbon advisory panel of the U.S. National Science Foundation (NSF) one of the primary funders of basic and applied research in the U.S. (Atkins et al., 2003):

> Cyberinfrastructure will become as fundamental and important as an enabler for the enterprise as laboratories and instrumentation, as fundamental as classroom instruction, and as fundamental as the system of conferences and journals for dissemination of research outcomes.

Distributed computing is a particularly important part of the equation, as the computing power in cyberinfrastructure for serving, rendering, analyzing, simulating may be as distributed as the data sets themselves (and this distribution often implies that data producers and providers are willing and able to share their products, often in near real-time). As such, research in cyberinfrastructure deals with the interoperability of technologies, as well as their efficiency, connectivity, and usability within the realms of large systems such as university consortia, large research collaboratives, and local/state-country/federal governments.

Current Initiatives in the U.S.

Two recent programs at the NSF have been created to provide federal dollars to cyberinfrastructure researchers are the NSF Division of Shared Cyberinfrastructure (www.cise.nsf.gov), with a focus on acquisition and upgrading of supercomputing facilities, high-capacity mass storage systems, enterprise software suites and programming environments, support staffers, etc. for the academic community; and the NSF Digital Government Program (www.digitalgovernment.org) with a mission to link academic research in information technology (including cyberinfrastructure) to the mission, directives, and activities of government at the federal and state levels, and to evaluate the overall resulting impact on governance and democracy. These “e-science” programs point to the priorities placed by our government on these areas and the recognition that new subdisciplines may be created as a result. *There has also been great interest expressed in funding collaboratives between U.S. researchers and European partners, and that cyberinfrastructure developed in the U.S. be interoperable with that being developed and deployed in other countries* (see further details in Atkins et al., 2003 and upcoming program information at www.cise.nsf.gov).

There are many examples of cyberinfrastructure projects in development, far too numerous to highlight in this paper or the accompanying talk, but one currently underway in the U.S., the Oregon Coastal Atlas (www.coastalatlas.net) has many connections to the Marine Irish Digital Atlas or MIDA (O’Dea et al., 2004). The Oregon Coastal Atlas is funded primarily by the NSF Digital Government Program and is a collaboration between the State of Oregon’s Ocean-Coastal Management Program (state government), Oregon State University (academia) and Ecotrust (non-profit environmental organization). The heart of the atlas is an interactive map, data, and metadata portal for coastal zone managers and coastal planners, with additional outreach sections for scientists, secondary school educators, and the general public. The portal enables users to obtain data sets, but also to understand their original context, and to use them for solving a spatial problem via online tools. The design of the atlas draws from the reality that resource decision-making applications require much more than simple access to data. Resource managers commonly make decisions that involve modeling risk, assessing cumulative impacts, and weighing proposed alterations to ecosystem functions and values. These decisions involve pulling together data sets and thus knowledge from disparate disciplines such as biology, geology, oceanography, hydrology, chemistry and engineering. Practitioners within each one of these disciplines are often vested in the technologies that dominate the market within their particular field. This presents significant data integration difficulties for investigators involved in management decisions that are as inherently interdisciplinary as those in the coastal zone. The goal of the atlas effort is to address these problems by incorporating a variety of geospatial data and analysis tools within a common framework. Advanced GIS tools to date that are available within the atlas include the Erosion Hazard Suite, Watershed Assessment Tool, and Marine Visioning Tool (Haddad et al., in press). In this way, the collaborative seeks to improve universal participation in coastal decision-making among communities within the state of Oregon by extending infrastructure to public offices that would otherwise face difficulties accessing these services and resources.

Examples of cyberinfrastructures that have been developed on a much broader scale (regional to national) include the Biomedical Research Network, a collaboration of three U.S. west coast universities (the University of California at San Diego or UCSD, the University of California at Los Angeles or UCLA, and the California Institute of Technology) with Duke University on the east coast to distribute and integrate multiscale biomedical data for human disease studies. GEONGrid (www.geongrid.org), a large, 5-year collaborative effort spearheaded by the San Diego Supercomputer Center, San Diego State University, and the Pennsylvania State University, to foster interdisciplinary research among geologists and geophysicists. These and many other collaboratives all participate to some extent in geodata.gov, the new incarnation of the nationwide network of geospatial metadata clearinghouses at the heart of the U.S. National Spatial Data Infrastructure. Geodata.gov has launched their Geospatial One-Stop Initiative (toward “one-stop” shopping of free government and academic data), part of the ongoing technological and “e-government” trend toward collecting and maintaining data sets locally or regionally, and sharing them nationally or internationally (*in some cases as fulfillment of a grant deliverable or contract, which must be completed before being eligible to apply again for future funding*).

Similar to geodata.gov is The National Map, a digital government effort by the U.S. Geological Survey to provide “a consistent framework for geographic knowledge needed by the nation. …public access to high-quality, geospatial data and information from multiple partners to help inform decision making by resource managers and the public” (nationalmap.usgs.gov). The National Map is about data access but through interactive map services (especially the Open Geospatial Consortium’s web mapping service (WMS) specification and web feature service (WFS) specification). Through a series of partnerships with other federal agencies, states, counties, municipalities, universities, and commercial companies, data may be published in such a way that a map drawn by the user on the web may come from more than one partner and from several different databases. Other features include “clip and ship,” where users may select an area on a map and receive data constrained to that area via FTP. In some cases the actual vectors may be streamed, edited online by the user, and then returned to an archive, or additional tools may be added to an archive’s interface. One of the most important features of the National Map is the *series of willing partnerships to ensure that data and services are provided with no charge or use restriction*, so that availability is in as many ways as possible, projects can move forward quickly, and data may even be updated and improved by local data maintainers who are most knowledgeable about it (“participants working on their day-to-day mapping responsibilities locally--being a star-and sharing the results with the nation--being part of the constellation”; Reed, 2004).

A Concluding Eye to the Future

The following is a small sampling of the most compelling cyberinfrastructure research topics being undertaken within the U.S. GIS/geographic information science community:

- Ontology and ontology cataloging, where ontology is briefly defined as the formalization of concepts and terms used in a practice or discipline (for background see Gruber, 1993; Mark et al., 2003). Ontologies can thus provide the semantic aspects of metadata, including lists of terms with definitions, more complex relationships between terms, rules governing those relationships, and potential values for each term.
- Closely related is the area of semantic interoperability and the semantic web (Fonseca and Sheth, 2003). Despite ontologies, words may still mean different things to different people within an interdisciplinary community and how does one, for example, search effectively through shared databases based on the words in the metadata (e.g., “coastline” vs. “shoreline”, “seabed” vs. “seafloor”, engineering vs. ecological resilience, “resilience” vs. “robustness”, scale vs. resolution (!), wetland buffering vs. GIS buffering, and so on.
- Spatialization, or the process of mapping out non-geographic information, again, in an attempt to improve distribution, search, and visualization of data and information (e.g., Skupin, 2002; Skupin et al., 2003).
- Development of domain-specific data models, with their accompanying distribution protocols and toolsets; data models for web GIS (e.g., Wright, 2003; Wright and Halpin, in press).
- Grid computing (Grid GIS, distributed agent GIS, peer-to-peer or P2P GIS), where the computing power may be as distributed as the data sets themselves (e.g., one might execute data on one machine, render it on another, send it back to another machine for GIS analysis and mapping, etc., and then deploy a prototype that ties all of these processes on all these servers together in a seamless interface).
    - Stability, performance, and connectivity issues
    - Design, architecture, algorithmic, data structure issues

- ❖ Data mining/knowledge discovery, visualization (e.g., Miller and Han, 2001; Shekar et al., 2003)
- ❖ Distributed GIS education (distance education)

So as work continues in these areas, and collaborations and funding levels remain at least at the present levels, the future appears bright for a new kind of exploration and discovery (even productive *re*-discovery), of physical places, environmental databases, information spaces, spatial data infrastructures and the like through cyberinfrastructures.

References

Atkins, D.E., Droegemeir, K.K., Feldman, S.I., Garcia-Molina, H., Klein, M.L., Messerschmitt, D.G., Messina, P., Ostriker, J.P., and Wright, M.H., 2003. *Revolutionizing Science and Engineering Through Cyberinfrastructure*, Washington, D.C., National Science Foundation, 175 pp., www.cise.nsf.gov/div/index.cfm?div=sci

Fonseca, F., and Sheth, A., 2003. *The Geospatial Semantic Web,* UCGIS Short-Term Research Priority White Paper, Leesburg, VA, USA, University Consortium for Geographic Information Science, www.ucgis.org → Research Priorities, 2 pp.

Gruber, T.R. 1993. A translation approach to portable ontologies, *Knowledge Acquisition*, 5(2):199-220.

Haddad, T., Wright, D., Dailey, M., Klarin, P., Dana, R., Marra, J., and Revell, D., in press. The tools of the Oregon Coastal Atlas, in Wright, D.J. and Scholz, A.J. (eds.), *Place Matters—Geospatial Tools for Marine Science, Conservation, and Management in the Pacific Northwest*, Corvallis, OR, USA, Oregon State University Press, http://dusk.geo.orst.edu/aaas.

Mark, D.M., Egenhofer, M., Hirtle, S., and Smith, B., 2003, *Ontological Foundations for Geographic Information Science*, UCGIS Long-Term Research Challenge White Paper, Leesburg, VA, USA, University Consortium for Geographic Information Science, www.ucgis.org → Research Priorities, 8 pp.

Miller, H.J., and Han, J. (eds.), 2001. *Geographic Data Mining and Knowledge Discovery*, London, Taylor & Francis.

O'Dea, E., Dwyer, N., Cummins, V., and Dunne, D., 2004. The design of the Marine Irish Digital Atlas, *Proceedings of GIS Ireland 2004*, IRLOGI, Dublin, Ireland.

Reed, C., 2004. The National Map: Open for business, *Geospatial Solutions*, 14(5), 1-2, www.geospatial-online.com → June 2004.

Shekhar, S., Vatsaval, R.R., Gahegan, M., Miller, H.J., Buttenfield, B., and Yuan, M., 2003. *Geographic Data Mining and Knowledge Discovery,* UCGIS Short-Term Research Priority White Paper, Leesburg, VA, USA, University Consortium for Geographic Information Science, www.ucgis.org → Research Priorities, 2 pp.

Skupin, A., 2002. A cartographic approach to visualizing conference abstracts, *IEEE Computer Graphics and Applications*. (Jan/Feb Issue).

Skupin, A., Fabrikant, S. and Couclelis, H., 2003. *Spatialization: Spatial Metaphors and Methods for Handling Non-Spatial Data,* UCGIS Short-Term Research Priority White Paper, Leesburg, VA, USA, University Consortium for Geographic Information Science, www.ucgis.org → Research Priorities, 2 pp.

Wright, D.J., 2003. *Pervasive Computing,* UCGIS Short-Term Research Priority White Paper, Leesburg, VA, USA, University Consortium for Geographic Information Science, www.ucgis.org → Research Priorities, 2 pp.

Wright, D.J., and Halpin, P.N., in press. Spatial Reasoning for *Terra Incognita*: Progress and Grand Challenges of Marine GIS in Wright, D.J. and Scholz, A.J. (eds.), *Place Matters—Geospatial Tools for Marine Science, Conservation, and Management in the Pacific Northwest*, Corvallis, OR, USA, Oregon State University Press, http://dusk.geo.orst.edu/aaas.

For Further *Exploration* on the Web

Cyberinfrastructure (U.S. National Science Foundation)
    www.cise.nsf.gov
Digital Government
    www.digitalgovernment.org
Geodata.gov / Geospatial One-Stop
    www.geodata.gov/gos
GEONGrid
    www.geongrid.org
The National Map
    nationalmap.usgs.gov
Oregon Coastal Atlas
    www.coastalatlas.net
UCGIS Research Priority White Papers (University Consortium for Geographic Information Science)
    www.ucgis.org

Dawn Wright received her Ph.D. in Physical Geography and Marine Geology from the University of California-Santa Barbara in 1994. She is currently professor of Geography and Oceanography at Oregon State University. Dr. Wright's research interests include geographic information science, marine geography, tectonics of mid ocean ridges, and the processing and interpretation of high-resolution bathymetric, video, and underwater photographic images. She has completed oceanographic fieldwork in some of the most geologically active regions of planet, including the East Pacific Rise, the Mid-Atlantic Ridge, the Juan de Fuca Ridge, the Tonga Trench, volcanoes under the Japan Sea and the Indian Ocean, and has also dived three times in the deep submergence vehicle *Alvin*. Dr. Wright serves on the editorial boards of the *International Journal of Geographical Information Science*, *Transactions in GIS*, and *Geospatial Solutions*, and on the U.S. National Academy of Sciences' National Needs for Coastal Mapping and Charting Committee. Her recent books include *Marine and Coastal Geographical Information Systems* (with Darius Bartlett, published by Taylor & Francis, 2000) *Undersea with GIS* (published by ESRI Press, 2002) and *Place Matters: Geospatial Tools for Marine Science, Conservation, and Management in the Pacific Northwest* (with Astrid Scholz, to be published by Oregon State University Press, forthcoming in April 2005).

Dr. Wright is the past recipient of a U.S. National Science Foundation CAREER award, Excellence in Mentoring awards by the Oregon State University College of Oceanic & Atmospheric Sciences, and Woman of the Year in Education by *Clarity* magazine.