Spatial Data Infrastructures for the Amazon: a First Step towards a Global Forest Information System

Frederico Fonseca
College of Information Sciences and Technology
Pennsylvania State University
University Park, PA 16802, USA
Email: fredfonseca@ist.psu.edu

Clodoveu A. Davis, Jr.
Departamento de Ciência da Computação
Universidade Federal de Minas Gerais
Av. Presidente Antônio Carlos, 6627 – 31270-010 – Belo Horizonte – MG – Brazil
Email: **clodoveu@dcc.ufmg.br**

Gilberto Câmara
INPE - Brazilian National Institute for Space Research
Av dos Astronautas, 1.758
São José dos Campos – SP - Brazil
Email: **gilberto.camara@inpe.br**

The preservation of the world's rain forests is one of the major environmental challenges of our generation. Rain forests are home to a large portion of the world's biodiversity, and play a major role in climate regulation and in the hydrological cycle of the Earth. Despite their biological richness and ecological importance, most of the world's rain forests are under increasing pressure of deforestation caused by human actions. Deforestation is particularly intense in countries with large-scale agricultural production, such as Brazil and Indonesia. In the Brazilian Amazonia, which covers an area of 4 million km^2, 17% of the original forest has been cleared, according to the Brazilian deforestation monitoring system, developed by the Brazilian National Institute for Space Research (INPE). Science Magazine called the Brazilian system "the envy of the world", but warned that, in spite of recent improvements in forest monitoring, uncertainties are still large (Kintisch 2007). Because of data and capacity gaps, current information on global deforestation is extremely patchy. There is much disagreement in the scientific community about the magnitude and extent of deforestation worldwide. As an example, the FAO 2005 Forest Resource Assessment (FAO 2006) points out that deforestation in Brazilian Amazonia from 2001-2005 equals 31,000 km^2/yr. Using data imagery from a different satellite, Hansen et al. (2008) estimated that forest clearings for Brazil in the same period were 26,000 km^2/yr. INPE´s detailed estimates are 22,000 km^2/yr for the same period (Valeriano, Monteiro et al. 2008). If we compare them to INPE's data, the FAO 2005 report and Hansen et al. overestimated deforestation in Amazonia by 63% and 18% respectively. Furthermore, making such estimates is a large effort, and comprehensive results take long to be obtained and validated. Information science and technology has a key role in the improvement of the monitoring efforts, but the efficient collection, organization, and distribution of large volumes of data still represent a great challenge for information science and technology.

In the case of environmental monitoring, which involves both a wide variety of observations and measurements, and a multiplicity of interested parties, this challenge

becomes even greater. Considering the diversity of views and interests, ranging from the scientific to the political to the common citizen, there is the need to employ the latest and the best of the available methods and techniques if we are to succeed. Spatial data infrastructures (SDI), conceived as a combination of technologies, people and politics, are probably the best approximation to the kind of multidisciplinary, multiple-scope, and multiple-scale integration that is required for global environmental monitoring. However, some elements are missing from the general SDI approach.

SDI is founded on interoperability, due to the need to join several different data-providing organizations, without interfering with their technological choices, production processes, or internal culture. There are also initiatives towards encapsulating data sources and creating online information services. However, several limitations come to mind as we think about the next step, i.e., on how the wide availability of spatial data and information can actually make a difference for complex problem-solving situations, involving multiple actors, with different (and often conflicting) world views. Therefore, even though the current framework for SDI technology, based mostly on OGC standardized components and Web services, represents a departure from the days of offline data exchange, there is much more to be done. There is a definite call to support integration and cooperation, going beyond SDI as an information provider, and using what is being learned from online social networks and from e-science efforts, in an attempt to achieve community building. People must be motivated to contribute and to participate, and better tools for data discovery must be developed, especially considering semantic aspects, since interdisciplinarity is a necessity. We need to take into consideration three actors (citizens, scientists and policy makers) and learn how they are expected to interact towards the creation of environmental policies that will lead to a sustainable planet (Davis Jr., Fonseca et al. 2009).

What is needed is an information system to enhance and augment communication and shared meanings among interested parties, thereby raising public awareness of environmental issues and showing the difficulties inherent to the process of finding solutions. Information should be widely available, easily discoverable, and usable for many different purposes.

Applying this idea to environmental monitoring, we envision a *Global Forest Information System* (Câmara, Vinhas et al. 2009) along the lines of an enhanced "Digital Earth" concept. We think that the Digital Earth metaphor can become much more than what today's virtual globes provide, and work as a large-scale repository of data, services, and models that can be discovered and used by the multiple actors involved in rainforest monitoring. A Global Forest Information System is a Web application that can increase the capacity of rainforest nations to estimate and monitor deforestation by making essential data and models easily available. The key objective is to enable cooperation and participation from the various actors involved, notably scientists, policy makers, and common citizens. Content as distinct as satellite images, spatial data infrastructures, geobrowsers, research data, laws and policies, and citizen-provided information can be indexed, searched, discovered and used by all interested parties. Although nations should be the ultimate responsible for forest monitoring and reporting, improving communication in a setting such as the one proposed here would be helpful in

many different ways, particularly for involving all interested parties and helping to eliminate speculation on the validity or reproducibility of scientific results.

A Global Forest Information System would have to address some requirements. First, data would be more readily available and in larger volumes, shortcutting issues such as image acquisition and preparation. A Global Forest Information System would also allow the integration of images of different resolutions to be used for different purposes. The communication aspect would enable not only the implementation of open data policies but also their discussion and enhancement. Second, there is the possibility of the different actors sharing and understanding the meaning of the scientific models explaining deforestation processes. This way information from the different sources can be used as a communication tool, in order to motivate common citizens, scientists, and the society at large, to contribute with the monitoring effort and to influence policy making and enforcement. Third, this Web application and the databases associated with it will serve as a research ground for scientific innovation, by allowing the discovery of patterns, the identification of practices, the development, calibration and testing of models, and the correlation of the observed phenomena with other social and economic information. Finally, scientists and technicians from nations with a lower level of readiness can have access to best practices, models, methods and solutions developed elsewhere, and can count on the support from more experienced peers in the adaptation of such tools to their nation's reality.

Considering this preliminary list of activities, it is clear that creating such a system presents numerous challenges. The papers in this special issue take an initial step in addressing some of the challenges related to creating a spatial data infrastructure for the Amazon, covering both theoretical and practical aspects. The papers that address the practical aspects deal mainly with the Amazon region, while the theoretical papers have a broader scope.

The first paper in the special issue is "Spatial Data Availability and its Implications for Sustainable Development of the Brazilian Amazon" by Amaral and d'Alge. They take a look at some data dissemination initiatives and present an inventory of the spatial data available for the Amazon. They also address issues in spatial information sharing. Interestingly enough, they found that the problem is more on the people side of the equation than on the technological side. They say that "the limitation on the promotion of spatial information and data sharing in Amazonia is neither computational nor technological. Rather, it is a matter of informing those who generate the data about the benefits of knowledge as a public good and promoting collaborative work". They also ask for some Web-based solutions that might promote a wider distribution of information about the Amazon. This is one of the themes of the next paper.

Souza et al. participate in the special issue with "Web-Oriented GIS Portal for Monitoring, Conservation and Law Enforcement of the Brazilian Amazon". They present what we could call an engaged Web interface for the Brazilian Amazon. They go beyond mere visualization of spatial data. By providing the results of previously processed spatial analyses and targeting protected areas, they aim at participating directly in the protection

of the Brazilian rainforest. Much in the spirit of our proposed Global Forest Information System, they use only public domain data and open source software in order to make their results clear and analyzable by other actors in the rain forest monitoring process.

Enabling broader participation in a Global Forest Information System setting is the theme of the next paper. With "Getting Across Information Communities – Embedding Semantics in the SDI for the Amazon", Maué and Ortmann are looking for a way to enable the participation of common citizens in the discovery and evaluation of geographical information. They argue that to have a sustainable Amazon we need the participation of more than just scientists and institutions. They point that local populations are often underrepresented in traditional SDI. Their solution, based on rule-based semantic annotations of geographic information sources, is their way to support this kind of broader participation.

The next paper, "Searching for metadata using knowledge base and topic maps in Spatial Data Infrastructures" by Correa Silva et al., seeks new approaches for the recovery and sharing of geospatial information. They try to overcome some of the limitations of SDI, mostly the fact that search is limited to keywords, spatial coordinates, temporal or thematic classification. Their focus is on use of thesauri and topic maps to support information retrieval in SDI. They describe the creation of a SDI-oriented thesaurus and how an information system supported by thesauri and topic maps could be used to retrieve metadata on SDI.

In the same way that our proposal for a Global Forest Information System is pushing SDI forward, we can see the next two papers as looking into the future of Cyberinfrastructure (CI), defined by Pinheiro Silva et al. as "the set of organizational practices, technical infrastructure and social norms that collectively provide for the smooth operation of scientific work at a distance". While Pinheiro Silva et al. look at the problem from a more practical point of view, Georgiadou et al. use a theoretical approach. Pinheiro Silva et al. in "CI-Miner: Semantically Enhancing Scientific Processes" consider that alternative ways of collaborating in science will need a decentralized approach to multiple-source data integration. They think that this leads to a "new generation of research activities that are more collaborative and multidisciplinary in nature" with scientists depending on innovative ways to collect, share and use data. This also takes us to one of the requisites in a Global Forest Information System, which is that all knowledge should be represented and available for all of the actors. The article gives the first step in this direction. While Maué and Ortmann applied semantic annotations to geographic information sources, Pinheiro Silva et al. used them to aggregate knowledge about scientific processes and their products. Annotations, as used in the CI-Miner project, benefit scientists only. In the future the annotations might be translated and adapted to other actors involved in the process of understanding and monitoring rain forests.

Georgiadou et al.'s approach to cyberinfrastructure opens the more theoretically-oriented part of this special issue. First, Georgiadou et al., Saab, and then de By and Lemmens discuss some of our current approaches on SDI and the theory behind it. They propose to advance towards more encompassing environmental information collection,

sharing, and use. With "Framing the use of geo-information in government: A tale of two perspectives" Georgiadou et al. propose an interesting alternative view on how people perceive the use of geoinformation, showing two different perspectives, one market-oriented and other policy-oriented. The political (polis) view is often overwhelmed by technological and practical concerns, so that people often lose perspective of the reasons that move the market. If we intend to put forward Global Forest Information System, we must really understand what motivates people to participate. This is where Georgiadou et al.'s paper makes an important addition to the scope of this special issue.

The next paper, "A conceptual investigation of the ontological commensurability of spatial data infrastructures among different cultures" by David Saab, does a conceptual study of SDI. Saab explores the nature of cultural schemas as expressions of indigenous ontologies and the ensuing challenges in communication. He tackles one of the aspects mentioned by Georgiadou et al.: "In different application domains—health, promoting investment in land, transport, et cetera—the specific values, practices and rules may differ considerably. … Values are continuously socially constructed and provide an arena for more particular contestations, as shown with the values of efficiency and equity. Human practices emanate from conflicting world-views (hierarchism, egalitarianism, individualism), shape management and decision styles and are shaped by political contexts". Saab sees SDI as shared socio-technological-informational ensemble that need to incorporate indigenous cultural schemas. He wants indigenous communities contributing to the collection of geospatial data. This should be done in a way that recognizes their contribution as a legitimate form of knowledge. In order to achieve this, Saab proposes to use ontologies as manifestations of cultural schemas rather than as expressions of logical formalisms. He uses Heidegger to show the parallels between the idea of cultural schemas and the human experience to propose SDIs more in line with Global Forest Information Systems.

"A skeleton design theory for Spatial Data Infrastructure Methodical construction of SDI nodes and SDI networks" from de By and Lemmens closes both the theoretical section and this special issue. They propose the early stages of a theory for designing spatial data infrastructures as geoservice systems. Although at a first look this could seem not exactly connected to the special issue's main thrust, the theme is important for SDI design and implementation. There is a need for such an organization in the context of complex, multi-thematic SDIs. The paper is focused on a service-oriented SDI framework, so that we can start thinking about how to design the loosely-coupled geographic information systems that are dynamically set up in their environment. The notion of a dynamically-configured GIS is problematic, since not all users are likely to be able to assemble their own SOA/SDI-based GIS on the fly. Therefore, it is necessary to develop some techniques for designing SDIs and their applications, and this is where this paper's contribution fits.

Thus, the special issue covers several aspects of a spatial data infrastructure for the Amazon and its possible extension into a Global Forest Information System. We hope the readers will find the material about what is currently being done on the subject useful in their research. We would like to thank Dr. Lubia Vinhas and Dr. Dalton Valeriano who participated with us in the creation of the Global Forest Information System concept. We

also would like to thank Earth Science Informatics and especially Dr. Hassan Babaie for their support and opportunity in hosting this special issue. We are thankful to the reviewers for their excellent and expedite work that made this special issue possible. Finally, we thank all the authors that submitted their papers for their excellent research and willingness to work together with us and the reviewers to make this special issue on Spatial Data Infrastructures for the Amazon: a First Step towards a Global Forest Information System a reality.

References

Câmara, G., L. Vinhas, et al. (2009). Geographical Information Engineering in the 21st Century. Research trends in geographic information science. G. Navratil. Dordrecht; New York, Springer**: 203-218.

Davis Jr., C. A., F. T. Fonseca, et al. (2009). "Beyond SDI: Integrating Science and Communities to Create Environmental Policies for the Sustainability of the Amazon." International Journal of Spatial Data Infrastructures Research **4**: 156-174.

FAO (2006). Global forest resources assessment 2005: progress towards sustainable forest management. Rome, Food and Agriculture Organization of the United Nations.

Hansen, M. C., S. V. Stehman, et al. (2008). "Humid tropical forest clearing from 2000 to 2005 quantified by using multitemporal and multiresolution remotely sensed data." Proceedings of the National Academy of Sciences **105**(27): 9439-9444.

Kintisch, E. (2007). "Carbon emissions: improved monitoring of rainforests helps pierce haze of deforestation." Science **316**(5824): 536.

Valeriano, D., A. Monteiro, et al. (2008). Monitoramento da Floresta Amazônica Brasileira por Satélite. INPE - Instituto Nacional de Pesquisas Espaciais