Clodoveu A. Davis, Jr.
Departamento de Ciência da Computação – Universidade Federal de Minas Gerais, Brazil
Av. Presidente Antônio Carlos, 6627 – 31270-010 – Belo Horizonte – MG – Brazil

Frederico T. Fonseca
College of Information Sciences and Technology
The Pennsylvania State University
307E IST Building
University Park, PA - U.S.A.

Gilberto Camara
INPE - Brazilian National Institute for Space Research
Av dos Astronautas, 1.758
São José dos Campos – SP - Brazil

Environmental Sustainability: The Role of Geographic Information Science and SDI in the Integration of People and Nature

Abstract. The main purposes of Sustainability Science include understanding, integrating, and modeling nature and society. Geographic Information Science (GIScience) is important to achieve such goals, since it reviews fundamental subjects in spatially-oriented fields, such as geography and cartography, while incorporating more recent developments in cognitive and information science. This chapter identifies the necessary key research questions in GIScience to support environmental sustainability. GIScience is able to contribute to sustainability research in many of its requirements. First, there is the need to create spatially-oriented environmental models that include humans. Such models, along with large amounts of distributed spatial data, can then be used as a source for the creation of sound and enforceable environmental policies. The impact of policies needs to be measured, compared to established goals, and communicated to the society. We argue that the combination of technologies, people, and policies that defines Spatial Data Infrastructures (SDI) is probably the best approximation we have to solve these problems. The kind of sharing that motivates the creation of SDIs can be extended towards the establishment of communities of practice, in which the central theme or subject is approached in various levels of detail and complexity.

1. Introduction

The process of global change is altering the Earth's environment and climate. The implications of these changes for sustainability call for an approach that integrates the natural sciences and the human sciences. Scientists need to develop an understanding of the complexity of physical-ecological-anthropogenic systems. In this new paradigm, the Earth's environment is seen as being influenced by the dynamic interaction of natural and social systems.

One of the most important research questions today is then "How is the Earth's environment changing and what are the consequences for human civilization?" The science areas necessary to address this question are so many that only a solid interdisciplinary approach can succeed. One of the attempts to understand Global Change in an interdisciplinary way is what is called today Sustainability Science. This new undertaking has recently gained space in the National Academy of Sciences, which has approved in 2006 a new section dedicated to Sustainability Science (Clark and Dickson 2003; Clark 2007).

Sustainability science purports to understand, integrate, and model nature and society. Since most of the interventions on the environment are human choices, we need modeling tools that represent the world as seen and modified by human beings. Geographic Information Science (GIScience) is crucial for this purpose (Goodchild 2003). In order to create environmental models that include humans, we need GIScience. The key question for GIScience is whether it has the methods and techniques to support sustainability research. We think GIScience is able to contribute to sustainability research, since it reviews some fundamental themes in traditional spatially-oriented fields, such as geography, cartography, and geodesy, while incorporating more recent developments in cognitive and information science (Mark 2000).

This article identifies the necessary key research questions in GIScience to support environmental sustainability. We discuss the main topics necessary for extending GIScience for it to be capable of understanding, representing and modeling sustainability-related activities, and to support public policies of adaptation and mitigation. GIS for sustainable development has been subject of research for many years (Wheeler 1993). Now, the constant development in information technology is always bringing new opportunities in the area (Wilhelmi and Brunskill 2003). We argue that the combination of technologies, people, and policies that defines a spatial data infrastructure (SDI) is probably the best approximation we have to solve these problems.

2. GIScience research and sustainability science

The challenges for GIScience regarding the support of sustainability actions can be understood as being part of a cycle (for a full annotated bibliography of the early years, see (Shortridge 1995) and for more recent work see (Campagna 2006)). There is the need to improve our modeling skills, in order to face more complex systems and the interaction between human actions and natural systems. We also need to refine our data collection and data management tools, so that we can work in a globally distributed way, and manage increasingly large amounts of online data. Then, our knowledge discovery assets need to be revised towards working with such amounts of distributed data, in order to generate relevant and timely information. This information can then be used as a basis for policy making, and for simulations and other kinds of advanced studies. Whatever knowledge is gained in the process will probably indicate the need to improve our models and collect data again, thus forming a cycle of continuous improvement.

The four main proposed topics for the new GIScience research agenda that will help the understanding of the process of Global Change involve *modeling*, *data collection*, *knowledge discovery*, and *support for policy-making*. A similar framework, although with a focus on Spatial Decision Support Systems, was introduced earlier by Densham (1991). These topics are briefly discussed next.

2.1. Modeling

A model is a construct that is developed to help us focus on what is important and relevant in our purpose to understand a system. Modeling tries to reduce the complexity of a real-world element or phenomenon to combinations of elements, such as a set of mathematical equations (*mathematical modeling*), a number of descriptive characteristics (*database modeling*), or a set of rules and behaviors (*dynamic* or *predictive modeling*).

Scientists must use simplifications and approximations to model aspects of the reality. The inaccuracies that result from such simplifications need to be assessed, in order to check the validity of the model. One way to do so is to create *simulations*, in which the scientist uses the modeled elements and past data to verify how accurately the present conditions can be predicted. The insight on reality that can be obtained from such a process enables the formulation of *forecasting models*, by means of which trends and the effects of new policies can be anticipated.

Modeling usually reflects a particular view on reality. Modelers must select and use elements from reality, as required to solve a specific range of problems, within that particular worldview. For a geographic information scientist, however, there is the additional challenge of creating representations of geographically located real-world elements that can be used by modeling efforts from scientists in other fields of expertise (Frank 2003). Therefore, incorporating semantics to the models is an important requisite. Furthermore, semantic differences that result from modeling some real world elements in different representation scales are a challenge to GIScience (Bruegger 1995; Myers, Pancerella *et al.* 2003), along with the creation of realistic and practical spatiotemporal modeling tools.

In order to adequately support the needs of sustainability science, we must be able to do all of the above, and also to evolve our modeling tools and skills to the point where modeling the connections between society and nature becomes feasible. There must be ways for scientists to develop a better understanding of human actions and motivations, especially is situations affect the environment. This can only be done by making the various worldviews explicit, and making sure these conceptions can be adequately represented in computational tools such as geographic information systems. The path to achieving that involves using ontologies (Fonseca, Egenhofer *et al.* 2002)as a modeling step that precedes conceptual modeling (Fonseca, Davis Jr. *et al.* 2003).

2.2. Data Collection

Data collection has certainly improved in the last decade, to the point where concerns have shifted from availability to accessibility and discovery of data sources. The Internet has certainly helped, but a relative lack of universally accepted data transfer standards makes it hard to integrate data from several sources in a meaningful and practical way. Much work on interoperability originates in such data transfer and translation difficulties (Rajabifard and Williamson 2001). A partial solution has been found by establishing neutral standards (such as the Geography Markup Language, GML) as a common ground between different data sources, but practical ways to automatically deal with semantics for integration are still the subject of much research (Fonseca 2008).

Many research initiatives currently have the need to (1) collect and organize large amounts of data using various methods, (2) integrate data from several different and distributed sources, and (3) adapt data collected within different semantic frameworks to fulfill their needs. It is usually possible, although time-consuming and error-prone, to perform such tasks manually. Research and development on fields such as data warehousing and records linkage have managed to supply scientists with a few tools and techniques, but there is still much to do.

Furthermore, when someone assembles a dataset from several different sources, chances are the data will become outdated soon. Therefore, some applications would rather count on ways to access data sources directly, instead of being caught in the extraction-transformation-load cycle. Current service-based architectures and content management technology can be combined and adapted to fulfill dynamic requests for data, thus enabling the creation of *loosely-coupled information systems*. Such systems require, fundamentally, that adequate sources of metadata are created and maintained (Kashyap and Sheth 1996). This is not a simple task, considering semantic concerns and the need to synchronize metadata and actual data, although some international metadata creation standards are available (International Standards Organization (ISO) 2003).

Collecting data for models that integrate nature and society (implying distributed global data management) requires understanding the collaborative monitoring of the Earth. There is a definite need for technologies and services that allow combining data from various (dynamic, distributed) sources to improve our capacity of measuring the state of the planet and acting upon the results.

2.3. Knowledge discovery

Dealing with large amounts of distributed data, as explained in the previous section, is already very difficult. Trying to make sense of all that data to generate useful and meaningful information is an even more complex task. There are currently data warehousing (DW), data mining (DM), and knowledge discovery from databases (KDD) techniques that are able to do so from centralized repositories (Shekhar and Chawla 2003; Han and Kamber 2005), and even some initiatives that allow for decentralized data sources, thus creating distributed data warehousing (Lau and Madden 2006). A range of DM techniques, geared towards mining data streams, can also be useful (Gaber, Zaslavsky *et al.* 2005).

GIScience takes on the challenge of putting together all kinds of knowledge discovery tools and techniques, adapting them wherever necessary to use the full potential of spatial and temporal information, in order to generate knowledge from observations, measurements, and other types of data available on the Internet. In the process, it is necessary to consider semantic frameworks to achieve integration (Fonseca, Camara *et al.* 2006), and to allow for ways to integrate without having to create centralized repositories or transferring large volumes of data across the globe. Ideally, data mining and KDD should be performed in a decentralized fashion, combining results at some location.

In the case of environmental sustainability, the challenges for knowledge discovery are even larger. It is necessary to combine and extract knowledge from spatial and temporal data (Fonseca and Martin 2004). It is also important to understand that data representing human actions and data representing nature may behave differently and generate incompatible trends. In summary, the problem of knowledge discovery, which is already complicated enough, becomes more complex when it is applied to environmental issues, understood as the result of society and nature interactions (Miranda and Saunders 2003).

There is also the technological challenge of mining data from streams of environmental measurements, then applying these data to models to be used to monitor environmental changes and also to support mitigation work. Therefore, we need to improve our capacity to discover new facts and trends in order to meet the demands of sustainable development. It is also necessary to find ways to share the new knowledge broadly and quickly (Goodchild, Fu *et al.* 2007).

2.4. Support for policy-making

In order to support policy-making, we need to use the knowledge that we acquired with the previous processes to develop policies to act upon the dynamic interactions of nature and society (Blackman and Köhlin 2008; Burtraw, Sweeney *et al.* 2008). It is necessary to communicate the results from knowledge discovery to policy makers. They also need access to the data and to well-explained versions of the models. In case of global policies, we need also to explicit any cultural assumptions behind the data and the models.

GIScience can be used in the support of the creation of new environmental policies (Walsh and Crews-Meyer 2002). How can we take actions to preserve the environment now and keep growing economically in the long run? We need to create different ways of modeling, implement and study these models (for instance, using simulation techniques), and use them to create and support policies that address sustainable development. During the current state of affairs, in which people are becoming aware of sustainability issues and starting to take immediate and long term actions, we need to monitor if our models, data and policies are correct. GIScience can help in many ways, including the creation of sustainability indices to support our decision making and to measure its effectiveness (Kates, Clark *et al.* 2001), and by facilitating the dissemination of information.

3. Connections between GIScience research and sustainability science

In this section we list the core questions for sustainability science as mentioned in *Science* by Kates et al. (2001), and discuss their repercussions. From this discussion, we propose new questions, this time specific to GIScience. Each new question is then related to one or more of the topics presented in section 2, namely modeling, data collection, knowledge discovery, and support for policy-making.

Question 1. *How can the dynamic interactions between nature and society – including lags and inertia – be better incorporated in emerging models and conceptualizations that integrate the Earth system, human development, and sustainability?*

This question poses, for GIScience and for other areas with an interest in modeling human behavior and its interactions with nature, a very big challenge. In short, it is about understanding how human societies shape and are shaped by nature, including cultural, political, social, and economic aspects. The broad scope of the question requires the capacity to generalize in a global scale, while considering local aspects and peculiarities. It also implies the need to cope with development policies and their impact on societies and on nature.

	Modeling	Data collection	Knowledge discovery	Support for policy-making
How do conceptualizations of sustainability vary across different cultures?	X			
How do we represent human actions in computer systems?	X			X
What is the impact of human actions in different geographical scales?	X		X	
How to deal with the variations in the perception of natural phenomena at various levels of detail?	X			
How can we merge geographic and georeferenced data from heterogeneous sources?	X	X		
How to establish the trustworthiness of data sources?		X		
How can we generate knowledge without experimenting with nature? Can we integrate and use alternative sources of knowledge, such as data from the past?		X	X	
How to assess and demonstrate the effects of development policies over natural systems?			X	X

Question 2. *How are long-term trends in environment and development, including consumption and population, reshaping nature-society interactions in ways relevant to sustainability?*

This question is intrinsically related to the previous one, since knowing more about long-term trends requires more advanced modeling and conceptualization skills. However, it presents GIScience with the need for improvements on monitoring methods and tools, in order to assess the correctness of models and the effectiveness of policies for sustainability. Kates et al. (2001) suggest the creation of sustainability indicators. We observe that such indicators should be built upon adequate spatial and temporal reference granularities, and should probably be assembled from local data, in a bottom-up fashion.

	Modeling	Data collection	Knowledge discovery	Support for policy-making
How can we agree on a set of societal and environmental variables from which indicator components can be chosen?	X	X		
How can we establish firm goals and quantifiable objectives for the sustainability effort?	X			X
How can we collect relevant data for sustainability indicators at different spatial and temporal granularities?		X		
How can we generate time series of indicators that reflect the situation in the past, so we can detect tendencies for the immediate future?		X	X	
How can we present indicators in a way that the general public can understand the evolution towards sustainability?			X	X
What kind of policies will accelerate and what kind will slow down the processes we want to control?				X

Question 3. *What determines the vulnerability or resilience of the nature-society system in particular kinds of places and for particular types of ecosystems and human livelihoods?*

Different societies and cultures may interpret and value differently vulnerability and environmental threats. Nevertheless, the actions of each society on the environment are shared by all. This is another example on how local activities affect the global environment, and on how a society (or all societies) should be held accountable for the consequences of its actions on the environment.

There are examples of fragile ecosystems that are affected by human actions that take place not only directly over them, but elsewhere, as in the case of the Great Barrier Reef, in Australia. Recently, that ecosystem has been affected both by farming, which causes pesticide- and fertilizer-based pollution in nearby basins (Devlin and Brodie 2004), and by warmer sea waters, that result from global warming (Great Barrier Reef Marine Park Authority 2008). Such an observation shows how, in several environmental issues, national borders become meaningless and the need to face problems becomes a global undertaking.

Therefore, within a GIScience understanding, the challenge presented by this question is about reaching comprehensive agreements, first on data and models, then on policies and monitoring.

	Modeling	Data collection	Knowledge discovery	Support for policy-making
How do conceptualizations of vulnerability and resilience vary across different cultures?	X			
How to express vulnerability and resilience spatially?	X			
How can we overcome national boundaries when dealing with data collection?		X		
How can we propose and implement global standards for data collection, documentation, and distributed access?		X	X	
How can we develop policies that are effective and, at the same time, fair to different cultures and lifestyles?			X	X

Question 4. *Can scientifically meaningful "limits" or "boundaries" be defined that would provide effective warning of conditions beyond which the nature-society systems incur a significantly increased risk of serious degradation?*

In a way, this question touches again on the issue of indicators, and asks whether is there a "point of no return" in relation to human actions causing degradation. If this is the case, the question implies the existence of a monitoring system, from which early warnings could be issued and action could be taken before a threshold is reached. For GIScience, this constitutes the main challenge related to this question, even though we can imagine geographic information scientists being involved in the determination of the thresholds themselves.

As a result, many demands to GIScience arise from the need to collect and analyze large amounts of data on nature-society systems, and to present results in a meaningful way.

	Modeling	Data collection	Knowledge discovery	Support for policy-making
How can we identify natural systems at risk, communities at risk, and cases of dependency between communities and natural systems? Which are the populations at immediate and long-term risk?	X	X	X	X
What are the human inputs to global climate models (land use change, carbon cycle, water cycle, and atmospheric chemistry, for instance), and where are their sources?	X	X		
How can we demonstrate and present tendencies and predict degradation risk?			X	X
How to tap into and learn from the globally distributed efforts to monitor the environmental systems?			X	
How can we isolate facts that can be used as examples and arguments to demonstrate degradation risk?			X	X
How can we isolate causes of degradation so that more efficient action can be taken against them?			X	X
How can we support the creation of a global schedule or timetable for acting against sources of degradation?			X	X

Question 5. *What systems of incentive structures – including markets, rules, norms and scientific information – can most effectively improve social capacity to guide interactions between nature and society toward more sustainable trajectories?*

Incentive systems are among the most interesting and cost-effective ways for public authorities to deal with environmental issues. In short, authorities must develop policies that make aggressions to the environment more costly than their prevention or compensation. There can be rewards for reducing impact, and/or penalties for causing degradation. In a best-case scenario, such rewards and penalties should be applied so that it becomes economically interesting, for the source of degradation, to invest in strategies and technologies to reduce impact it causes on the environment and on populations (National Center for Environmental Economics (NCEE) 2001).

However, incentive systems alone cannot ensure that society learns about threats to itself or to the environment. There are numerous cases of litigation, either involving governments and corporations, or groups of citizens and corporations, in which reparation is sought in court for health or environmental damages. Awareness of such situations should be foremost in the agenda for sustainability. For that purpose, regulations and norms that require information transparency are becoming commonplace, but communicating complex data to the general public is still a big challenge.

	Modeling	Data collection	Knowledge discovery	Support for policy-making
What are the relations between markets and sustainability at various spatial scales?	X	X	X	
How can economic factors for sustainability be expressed and viewed spatially?	X		X	
How can we integrate structured and unstructured data for information transparency purposes?	X	X		
How does the spatial expression of markets contribute to public policies that promote sustainability?				X

Question 6. *How can today's operational systems for monitoring and reporting on environmental and social conditions be integrated or extended to provide more useful guidance for efforts to navigate a transition toward sustainability?*

In our data-intensive era, numerous data collection efforts take place simultaneously, generating large volumes of measurement data. Considering the historical accumulation of such data, these volumes compound even more, to the point where the problem of data availability has become a problem of finding and getting access to relevant data.

Naturally, if every environmental and social data source found a way to publish their data on the Internet, much of the access problem would be solved, but the data discovery problem would still remain. Furthermore, there are semantic aspects related to the paradigms that guided the data collection effort, that have to be considered when scientists need to decide whether the existing data fit their needs or not.

Metadata, in this case, become fundamentally important. Standards for geographic metadata are in place, in the form of ISO 19115 (2003) (International Standards Organization (ISO) 2003), and projects such as INSPIRE (Smits 2002) have already assembled searchable sources of geographic metadata. The current efforts can be extended to include alternatives to keyword-based searching, so that language becomes less of a hindrance and semantic aspects can be included.

There is also the matter of integration of data sources. Metadata should be sufficient to allow a scientist to decide whether two datasets could be reasonably used together, but adequate (and possibly automatic) treatment of uncertainty, level of detail, and – once again – semantics is still pending.

	Modeling	Data collection	Knowledge discovery	Support for policy-making
How can we achieve interoperability between models that are created under different scientific paradigms?	X			
How can we achieve interoperability between environmental monitoring systems?	X	X		
Can we build intelligent systems that work on the border between the environmental and the social worlds, joining data sources from both?	X		X	
Can we quickly put together new systems based on multiple and distributed data sources?	X		X	X
How can we create systems that help the design of policies and the evaluation of policies' results?				X

Question 7. *How can today's relatively independent activities of research planning, monitoring, assessment, and decision support be better integrated into systems for adaptive management and societal learning?*

The integration of scientific disciplines to promote research on interdisciplinary themes is often hard to achieve. Different worldviews, along with divergent research agendas and pigeonholed funding opportunities, constitute hurdles to groups of scientists that work on similar subjects and wish to develop integrated work.

GIScience is known to be essentially interdisciplinary, and geography can many times provide a good basis for the integration of scientific work and data from several disciplines. Therefore, the answer to this question, from a GIScience point of view, implies continuing the search for more and better ways to integrate models, people and data, and for more and better ways to communicate results and act upon them.

	Modeling	Data collection	Knowledge discovery	Support for policy-making
How can we build interdisciplinary models that reach across different and sometimes incompatible fields of knowledge?	X			
How can we integrate data coming from different sciences?		X		
How can we incorporate unstructured data coming from informal sources?	X	X		
How can we build geographical visualization systems that help public policy makers and societal stakeholders?			X	X
How do good GIS user interfaces help planners and decision makers?				X
How can a planner build scenarios using spatial decision support systems?	X	X	X	X

When applied to sustainability, implying various social and environmental conditions and a multiplicity of actors, with points of view ranging from the political to the scientific to the common citizen, Clark's questions pose enormous requirements for information science and technology, and for GIScience in particular. SDIs have the potential to address many of these problems (Keßler, Wilde *et al.* 2005; Czerwinski, Sandmann *et al.* 2007; de Man 2007), although it has to evolve in many ways to face the enormous challenges posed by sustainability.

4. Spatial data infrastructures and sustainability

Scientists involved in sustainability must be able to combine spatial data from different sources to produce new information for a study area. This activity can be very complex, for many reasons. First, phenomena occur and are modeled in various geographic scales, ranging from microbiology in specific locations to planetary climate impacts. Even though there is a general understanding about the semantic variation of phenomena across multiple geographic scales, our current tools and techniques are still primitive in comparison with the breadth of this challenge. Second, there are multiple views on the reality of the environment, including many scientific disciplines and the view of the local populations. These views are sometimes complementary, and sometimes conflicting, each one based on a particular set of concepts. Third, the complexity and level of detail of activities such as data collection and analysis range from large volumes of scientific data down to news and descriptions suited to the cultural level of local populations. From a GIScience standpoint, such requirements involve, at least, (1) efficient access to data, (2) widely accepted interoperability mechanisms (Wilhelmi and Betancourt 2005), and (3) semantic integration of data sources (Claramunt and Theriault 1996; Laurini 1998; Hakimpour and Timpf 2001).

SDI is a new approach to creation, distribution and use of geographic information that tries to address the shortcomings listed above. SDI tries to avoid the old view of GIS as an automated map distribution system, which focuses on map production and distribution of existing sources on an "as-is" basis. SDI is an enabler for understanding space. SDI does not simply deliver maps. It disseminates spatial data with associated quality control, metadata information, and semantic descriptions. In this view SDI can play an important role in the management of the environment and in the sustainable growth of our society.

The expression "spatial data infrastructure" was initially used to describe a standardized way to access to geographic information (Maguire and Longley 2005). A SDI implies the existence of some sort of coordination for policy formulation and implementation, along with more complete and standardized metadata, possibly including means to provide online access to data sources.

The first generation of SDI focused on granting a broad thematic scope, which is consistent with the current analogy between SDI and other types of infrastructure: fostering economic development by granting access to publicly-available and multiple-use goods or services. Evolution from the first generation of SDI was made possible by the recent expansion of Web-based information systems. In the USA, the Geospatial One-Stop (GOS) Web portal was created to provide widespread access to geographic information, inaugurating the concept of *geoportals* (Maguire and Longley 2005; Tait 2005), currently viewed as SDI components. While an SDI is the overarching environment formed by the confluence of several geographic data providers, each of which granting data access through specific Web services, a geoportal provides means to give humans some level of interactive access to these data resources, including Web-based viewers and metadata-based discovery tools (Figure 1).

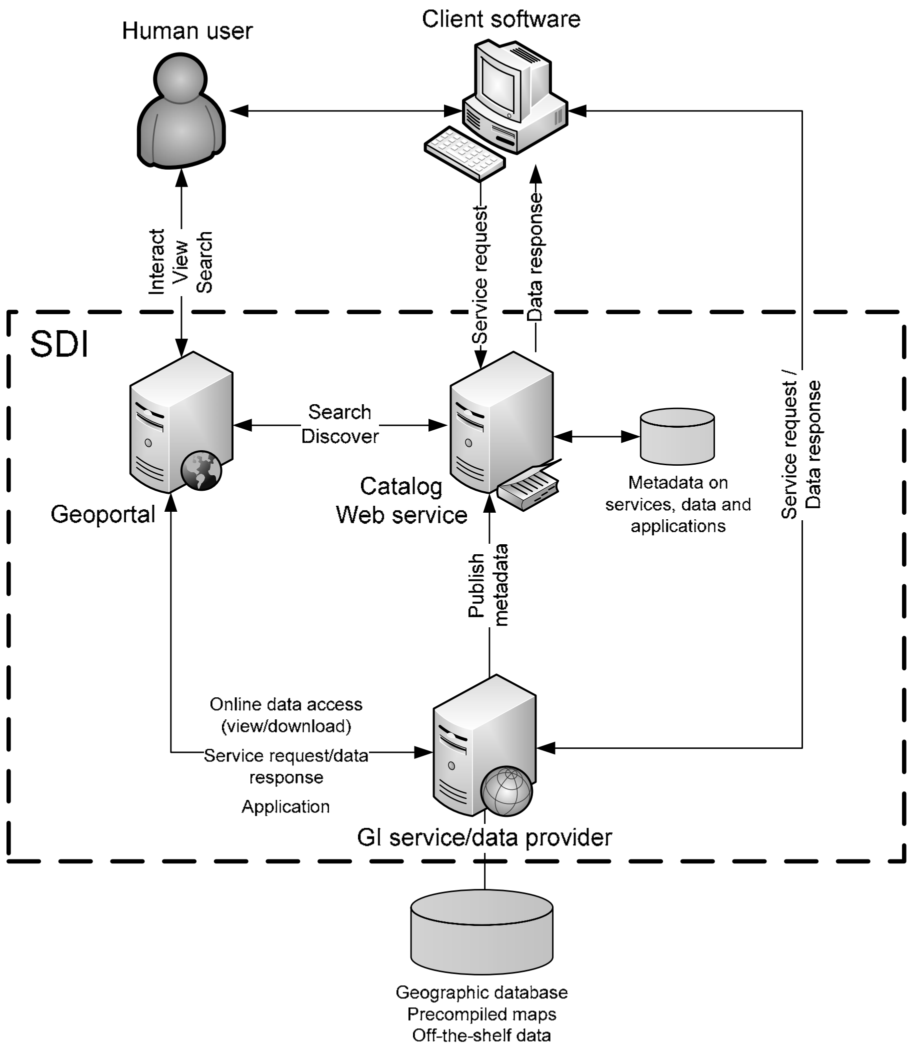

Figure 1 - Geoportals and SDI

The use of Web services to grant direct access to data is the most important distinction between first- and second-generation SDIs. In fact, the numerous possibilities that arise from using such services to encapsulate data from multiple sources, and thereby achieve interoperability, have led Bernard and Craglia (2005) to propose a new translation for the SDI acronym: *Service-Driven Infrastructures*. In fact, current SDIs include Web services as one of the possible data access channels, while maintaining links to downloadable data and existing Web applications.

The most current view on spatial information infrastructures considers their evolution into the perspective of *service-based distributed system architectures,* which have been proposed as part of a strategy for developing complex information systems based on reusable components. One of the most interesting approaches in this field is the one of *service-oriented architectures* (SOA) (Papazoglou and Georgakopoulos 2003). Services, their descriptions and fundamental operations, such as discovery, selection, and binding, form the basis of SOA. SOA supports large applications with sharing of data and processing capacity, through network-based distributed allocation of applications and use of computational resources. In this architecture, services are self-contained, which means that information on the service's description, including its capabilities, interface, behavior, and quality, can be obtained from the service itself, through a standardized set of functions. The Open Geospatial Consortium (OGC) has proposed many standards for Web service-based data access, such as the Web Feature Service (WFS), the Web Map Service (WMS), and several more, including some which are under evaluation at the time of this writing (Klopfer 2005). Of these, one of the most important is the Web Processing Service (WPS), which allows services to encapsulate analysis operations and algorithms, and provides means for service chaining, i.e., establishing sequences of steps using services, in a sort of workflow.

5. Research challenges

We consider that SDIs can be a valuable asset to develop practical solutions for the huge challenges posed by sustainability science. By organizing existing data in an unobtrusive setting of multiple and distributed sources, scientists can discover and gain direct access to relevant data, avoiding the need for time-consuming data transfer and translation. By "unobtrusive" we mean that data providers can keep their data collection and maintenance routines intact, based on the information technology tools of their choice, while being able to provide direct access to data in a timely and technologically-neutral way.

The resulting framework points towards the idea of *loosely-coupled GIS* (Alves and Davis Jr 2007), especially if the possibilities for developing and deploying more sophisticated processing and analysis services are taken into account. For instance, consider the existence of various separate sources of data on rainfall, temperature, soil types, and vegetation. From these data, a climate scientist needs to perform an analysis to determine evaporation averages. Algorithms to perform such an analysis can be documented with metadata and implemented as services, such as the WPS. As a result, chaining selected data-provision services for the four sources to a selected analysis or processing service, information can be generated without the need to transfer and install sophisticated tools at the scientist's site, and can even dismiss the need for locally-available computing power. The scientist could, in principle, execute such an analysis in the field, equipped only with some sort of mobile computing device connected to the Internet. This is a form of *cloud computing*, a concept related to Web 2.0 in the direction of providing "software as a service" (SaaS) (Buyya, Yeo *et al.* 2008; Pierce 2008). Notice that selecting among various data and processing sources is an integral part of the task, and the scientist needs to have means to discern between such alternatives. This indicates the need for *semantic discovery of services*, meaning that simple metadata schemes with keyword-based searches may not suffice.

Even though there is a potentially large number of data sources, there is a definite challenge in finding and combining them adequately. One view of SDI (Craglia, Goodchild *et al.* 2008) associates it with official mapping agencies. We think it is unreasonable to expect that systematic cartography can keep up with spatial data needs, especially in environmental frontiers, such as the Amazon. However, this point has been the subject of much debate recently. While cartography advocates stand for a more traditional approach, with an SDI based on a broad agreement on its data contents, others envision a multitude of apparently disconnected data sets, each of which with its own intended uses, but with possible applications in other areas as well. The first approach implies a stronger presence of official data providers in the definition of what should be available, while the second expects users to proactively discover the data they need, in a broader definition of the SDI concept. This redefinition of SDIs is based on the strong need for data availability, requiring simple and practical Web-based resources. We understand that, in such an SDI, a user should be able to assess data provenance (Myers, Pancerella et al. 2003), and to make an informed choice between official and other data sources.

Modeling in such a complex setting is also a challenge. We need to create different ways to implement and study these models (possibly using simulation techniques), and use them to create and support policies that address sustainable development. During the current state of affairs, in which people are becoming aware of sustainability issues and starting to take immediate and long term actions, we need to monitor if our models, data and policies are correct. One of the solutions points toward the creation of sustainability indices to support our decision making and to measure its effectiveness, as suggested by Kates et al (2001). This is a definite requirement if we intend our models to succeed in situations that are much more complex than the usual geographic application, such as in the broad modeling of the interaction between Society and Nature.

6. Future trends and prospects

The potential volume of data sources and the complexity of geospatial analysis algorithms pose interesting and important challenges for loosely-coupled GIS and cloud computing. Application requirements for large volumes of data transfer can be costly and time consuming, indicating that users might prefer to keep copies at more convenient (although also Web-based) locations, and therefore some kind of synchronization should take place. There is also the need for more research and development on services integration, chaining, and orchestration, with better and easier to use tools, along with the need for specialized services, designed to assist the use of geospatial cloud computing resources with temporary data storage and synchronization methods (Alves and Davis Jr 2007). Furthermore, more and better tools for mobile SDI-based geospatial computing need to be developed, including geospatial viewers specifically designed for small screens, and location-aware services, which can count on the growing availability of GPS receivers in cellular phones and other devices. There are also several concerns about computational performance, protection of sensitive data, and the security of partial results, although these concerns are shared by the general SOA and cloud computing development efforts.

The kind of sharing that motivates the creation of SDIs can be extended towards the establishment of communities of practice (Lesser, Fontaine *et al.* 2000; de Man 2007), in which the central theme or subject is approached in various levels of detail and complexity. In order to support sustainability efforts, there is clearly the need to improve information dissemination not only among members of the same group (scientists, policy makers, citizens), but among groups. The recent phenomenon of online communities of social interaction on the Web demonstrates that this kind of integration is possible, and highly desirable as a means to motivate people to participate and contribute to solve real problems.

In our point of view, the center of a community of practice for a subject as wide as environmental sustainability should be built along the lines of the Digital Earth paradigm (Craglia, Goodchild et al. 2008), which consists on a very wide array of data and information sources, ranging from the governmental mapping agency cartographic data (SDI), to data collected in research projects, associated to academic publications. It should also range from structured data sources, such as remote sensing products, to simplified and less structured sources such as volunteered geographic information about themes of interest for the community. Geobrowsers can play the role of integrators for popular contributions, and should be enhanced by scientific data and elements from official policies. Furthermore, we expect that a new generation of semantic integration tools and techniques will be able to enrich community building, by providing actors with some level of semantic support, such as translation of concepts, automatic links to educational resources, and discovery of services or applications.

7. Acknowledgments

Clodoveu Davis' work was partially supported by the Brazilian National Institute of Science and Technology for the Web (CNPq grant 573871/2008-6), CNPq (grants 306370/2006-9 and 551037/2007-5), and Fapemig (CEX APQ 0551-5.01/07), Brazilian agencies in charge of fostering research and development.

Frederico Fonseca's work was partially supported by Fapemig (CEX 00038/07), and by the generous support of Penn State's College of Information Sciences and Technology.

Gilberto Camara's work is partially funded by CNPq (grant PQ 550250/2005-0) and FAPESP (grant 04/11012-0).

8. Authors

Clodoveu A. Davis, Jr.

Departamento de Ciência da Computação – Universidade Federal de Minas Gerais, Brazil

Av. Presidente Antônio Carlos, 6627 – 31270-010 – Belo Horizonte – MG – Brazil

clodoveu@dcc.ufmg.br

Short bio: Clodoveu Augusto Davis Junior received his B.S. degree in Civil Engineering in 1985 from the Federal University of Minas Gerais (UFMG), Brazil. He obtained M.Sc. and Ph.D. degrees in Computer Science, also from UFMG, in 1992 and 2000, respectively. He led the team that conducted the implementation of GIS technology in the city of Belo Horizonte, Brazil, and coordinated several geographic application development efforts. Currently, he is a professor and researcher at the Federal University of Minas Gerais. His main research interests include geographic databases, urban GIS, spatial data infrastructures, and multiple representations in GIS.

Frederico T. Fonseca

College of Information Sciences and Technology

The Pennsylvania State University

307E IST Building

University Park, PA - U.S.A.

16802-6823

fredfonseca@ist.psu.edu

Short bio: Frederico Fonseca graduated from the Federal University of Minas Gerais with a degree in Data Processing (1977) and the Catholic University of Minas Gerais with a B.S. in Mechanical Engineering (1978). Subsequently he obtained a Master's in Public Administration from the João Pinheiro Foundation, Brazil (1995). Dr. Fonseca was an Assistant Professor in Computer Science at the Catholic University of Minas Gerais, and held professional appointments as senior systems analyst, GIS analyst and programmer. He completed his Ph.D. with Dr. Egenhofer at the University of Maine in 2001. His thesis covered the area of GIS interoperability. His research provides a theoretical basis for semantic interoperability, which is a necessary foundation for a working, interoperating environment. With his focus on system design, Dr. Fonseca demonstrates how complicated processes can be integrated to the benefit of users. His newly developed concept of ontology-based GIS is highly interdisciplinary as it brings together various research methods from artificial intelligence, software engineering, and GIS. Currently, he is an Associate Professor at the College of Information Sciences and Technology at the Pennsylvania State University.

Gilberto Camara

INPE - Brazilian National Institute for Space Research

Av. dos Astronautas, 1.758

São José dos Campos – SP

Brazil

gilberto.camara@inpe.br

Short bio: Dr. Gilberto Camara is currently the general manager of the Brazilian National Institute for Space Research – INPE, appointed by the Brazilian president for the period 2006 to 2010. He is responsible for the administration of INPE's research and development groups in Space Science, Space Engineering, Earth Observation and Weather and Climate Studies. Previously, he was head of INPE's Image Processing Division from 1991 to 1996 and Director for Earth Observation from 2001 to 2005. His research interests include: geographical information science and engineering, spatial databases, spatial analysis and environmental modeling. He has published more than 120 full papers on refereed journals and scientific conferences.

9. References

Alves, L.L. and Davis Jr, C.A. (2007). Evaluation of OGC Web Services for Local Spatial Data Infrastructures and for the Development of Clients for Geographic Information Systems. Advances in Geoinformatics. Davis Jr, C.A. and Monteiro, A.M.V. Berlin, Springer**:** 217-234.

Bernard, L. and Craglia, M. (2005). SDI: From Spatial Data Infrastructure to Service Driven Infrastructure. Research Workshop on Cross-Learning between Spatial Data Infrastructures (SDI) and Information Infrastructures (II), Enschede, Holanda.

Blackman, A. and Köhlin, G. (2008). Environment for Development: RFF Joins Initiative to Strengthen Green Policies in Developing Countries. Resources. **Fall 2008:** 4-5.

Bruegger, B. (1995). Theory for the Integration of Scale and Representation Formats: Major Concepts and Practical Implications. Spatial Information Theory—A Theoretical Basis for GIS, International Conference COSIT '95, Semmering, Austria. Frank, A. and Kuhn, W. Berlin, Springer-Verlag. **988:** 297-310.

Burtraw, D., Sweeney, R. and Walls, M. (2008). Crafting a Fair and Equitable Climate Policy: A Closer Look at the Options. Resources" (Fall 2008, Number 170). **Fall:** 20-23.

Buyya, R., Yeo, C.S. and Venugopa, S. (2008). Market-oriented cloud computing: vision, hype, and reality for delivering IT services as computing utilities. 10th IEEE International Conference on High Performance Computing and Communications (HPCC-08), Dalian, China.

Campagna, M. (2006). GIS For Sustainable Development, CRC Press.

Claramunt, C. and Theriault, M. (1996). Toward Semantics for Modelling Spatio-temporal Processes within GIS. International Symposium on Spatial Data Handling, Delft, NL:47-63.

Clark, W.C. (2007). "Sustainability Science: A room of its own." Proceedings of the National Academy of Sciences **104**(6): 1737.

Clark, W.C. and Dickson, N.M. (2003). Sustainability science: The emerging research program. Proceedings of the National Academy of Sciences, National Acad Sciences. **100:** 8059-8061.

Craglia, M., Goodchild, M.F., Annoni, A., Câmara, G., Gould, M., Kuhn, W., Mark, D., Masser, I., Maguire, D., Liang, S. and Parsons, E. (2008). "Next-Generation Digital Earth." International Journal of Spatial Data Infrastructures Research **3**: 146-167.

Czerwinski, A., Sandmann, S., Stšcker-Meier, E. and Pluemer, L. (2007). "Sustainable SDI for EU noise mapping in NRW – best practice for INSPIRE." International Journal of Spatial Data Infrastructures Research **2**: 90-111.

de Man, E. (2007). "Beyond Spatial Data Infrastructures there are no SDIs - so what?" International Journal of Spatial Data Infrastructures Research **2**: 1-23.

Densham, P.J. (1991). Spatial decision support systems. Geographical Information Systems: principles and applications Maguire, D.J., Goodchild, M.F. and Rhind, D. London, Longman Scientific & Technical 403-412.

Devlin, M.J. and Brodie, J. (2004). "Terrestrial discharge into the Great Barrier Reef Lagoon: nutrient behavior in coastal waters " Marine Pollution Bulletin **51**(1-4): 9-22.

Fonseca, F., Camara, G. and Monteiro, A. (2006). "A Framework for Measuring the Interoperability of Geo-Ontologies." Spatial Cognition and Computation **6**(4): 307-329.

Fonseca, F., Egenhofer, M., Agouris, P. and Câmara, G. (2002). "Using Ontologies for Integrated Geographic Information Systems." Transactions in GIS **6**(3): 231-257.

Fonseca, F. and Martin, J. (2004). "Space and Time in Eco-Ontologies." AI Communications - The European Journal on Artificial Intelligence **17**(4): 259-269.

Fonseca, F.T. (2008). Ontology-Based Geospatial Data Integration. Encyclopedia of GIS. Shekhar, S.X., Hui London ; New York, Springer.

Fonseca, F.T., Davis Jr., C.A. and Câmara, G. (2003). "Bridging Ontologies and Conceptual Schemas in Geographic Applications Development." Geoinformatica **7**(4): 355-378.

Frank, A. (2003). Ontology for Spatio-temporal Databases. Spatio-Temporal Databases: The Chorochronos Approach. Koubarakis, M. and Sellis, T. Berlin Heidelberg New York, Springer-Verlag**:** 9-78.

Gaber, M.M., Zaslavsky, A. and Khrishnaswami, S. (2005). "Mining Data Streams: a Review." ACM SIGMOD Record **34**(2): 18-26.

Goodchild, M.F. (2003). "Geographic Information Science and Systems for Environmental management." Annual Review of Environment and Resources **28**: 493-519.

Goodchild, M.F., Fu, P. and Rich, P. (2007). "Sharing Geographic Information: an assessment of Geospatial One-Stop." Annals of the Association of American Geographers **97**(2): 250-266.

Great Barrier Reef Marine Park Authority. (2008, August 25, 2008). "Coral bleaching on the Great Barrier Reef." from http://www.gbrmpa.gov.au/corp_site/key_issues/climate_change/climate_change_and_the_great_barrier_reef/coral_bleaching_on_the_great_barrier_reef.

Hakimpour, F. and Timpf, S. (2001). Using Ontologies for Resolution of Semantic Heterogeneity in GIS. 4th. AGILE Conference on Geographic Information Science, Brno, Czech Republic:385-395.

Han, J. and Kamber, M. (2005). Data Mining: Concepts and Techniques, 2nd. Edition, Morgan-Kaufmann.

International Standards Organization (ISO). (2003). "Geographic information - Metadata."   Retrieved August 25, 2008, from http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=26020.

Kashyap, V. and Sheth, A. (1996). Semantic Heterogeneity in Global Information System: The Role of Metadata, Context and Ontologies. Cooperative Information Systems: Current Trends and Directions. Papazoglou, M. and Schlageter, G. London, Academic Press**:** 139-178.

Kates, R.W., Clark, W.C., Corell, R., Hall, J.M., Jaeger, C.C., Lowe, I., McCarthy, J.J., Schellnhuber, H.J., Bolin, B. and Dickson, N.M. (2001). Environment and Development: Sustainability Science. Science. **292:** 641-642.

Keßler, C., Wilde, M. and Raubal, M. (2005). Using SDI-based Public Participation for Conflict Resolution. 11th EC-GI & GIS Workshop, Alghero, Sardinia.

Klopfer, M. (2005). Interoperability & Open Architectures: an analysis of existing standardisation processes & procedures. OGC White Paper. Consortium, O.G., Open Geospatial Consortium**:** 26p.

Lau, E. and Madden, S. (2006). An integrated approach to recovery and high availability in an updatable, distributed data warehouse. Proceedings of the 32nd International Conference on Very Large Databases, Seoul, Korea:703-714.

Laurini, R. (1998). "Spatial Multi-database Topological Continuity and Indexing: a Step Towards Seamless GIS Data Interoperability." International Journal of Geographical information Science **12**(4): 373-402.

Lesser, E.L., Fontaine, M.A. and Slusher, J.A. (2000). Knowledge and Communities, Butterwoth-Heinemann.

Maguire, D.J. and Longley, P.A. (2005). "The emergence of geoportals and their role in spatial data infrastructures." Computers, Environment and Urban Systems **29**(1): 3-14.

Mark, D. (2000). Geographic Information Science: defining the field. Foundations of Geographic Information Science. Duckham, M., Goodchild, M.F. and Worboys, M.F. New York, Taylor & Francis**:** 3-18.

Miranda, S.M. and Saunders, C.S. (2003). "The Social Construction of Meaning: An Alternative Perspective on Information Sharing." Information Systems Research **14**(1): 87-106.

Myers, J.D., Pancerella, C., Lansing, C., Schuchardt, K.L. and Didier, B. (2003). Multi-Scale Science: Supporting Emerging Practice with Semantically Derived Provenance. Semantic Web Technologies for Searching and Retrieving Scientific Data Workshop at the 2nd International Semantic Web Conference, Sanibel Island, FL.

National Center for Environmental Economics (NCEE) (2001). Chapter 3: The cost effectiveness and environmental effects of incentive systems. The U.S. Experience with economic incentives for protecting the environment. Washington, DC, Environmental Protection Agency (EPA).

Papazoglou, M.P. and Georgakopoulos, D. (2003). "Service-Oriented Computing." Communications of the ACM **46**(10): 25-28.

Pierce, M.E. (2008). "Grids and Geospatial Information Systems." Concurrency Computation: Practice and Experience **20**: 1611-1615.

Rajabifard, A. and Williamson, I.P. (2001). Spatial Data Infrastructures: Concept, Hierarchy, and Future Directions. GEOMATICS'80, Tehran, Iran.

Shekhar, S. and Chawla, S. (2003). Spatial Databases: a tour, Prentice-Hall.

Shortridge, A. (1995). Multiple Roles for GIS in US Global Change Research: Annotated Bibliography. Santa Barbara, CA, University of California, Santa Barbara.

Smits, P. (2002). "INSPIRE Architecture and Standards Position Paper." Retrieved 04/2005, from http://inspire.jrc.it/reports/position_papers/inspire_ast_pp_v4_3_en.pdf.

Tait, M.G. (2005). "Implementing geoportals: applications of distributed GIS." Computers, Environment and Urban Systems **29**(1): 33-47.

Walsh, S.J. and Crews-Meyer, K.A., Eds. (2002). Linking People, Place, and Policy: A GIScience Approach. Berlin, Springer.

Wheeler, D.J. (1993). "Commentary: Linking Environmental Models with Geographic Information Systems for Global Change Research." Photogrammetric Engineering & Remote Sensing **59**(10): 1497-1501.

Wilhelmi, O.V. and Betancourt, T.L. (2005). "Evolution of NCAR GIS Initiative: Demonstration of GIS Interoperability." Bulletin of American Meteorological Society **86**(2): 176-178.

Wilhelmi, O.V. and Brunskill, J.C. (2003). "Geographic Information Systems in Weather, Climate, and Impacts." Bulletin of American Meteorological Society **84**(10): 1409–1414.