

available at www.sciencedirect.comwww.elsevier.com/locate/ecolinf

Methods for facilitating web-based participatory research informatics

Brent Pedersen^{a,1}, Faith Kearns^{a,*}, Maggi Kelly^b

^aCenter for Fire Research and Outreach, University of California, Berkeley, 137 Mulford Hall #3114, Berkeley, CA 94720

^bGeospatial Imaging and Informatics Facility, University of California, Berkeley, 137 Mulford Hall #3114, Berkeley, CA 94720

ARTICLE INFO

Article history:

Received 15 September 2006

Received in revised form

26 January 2007

Accepted 18 February 2007

Keywords:

Collaboratory

GIS

Dynamic data

Research facilitation

ABSTRACT

Scientific datasets are often generated and used by “communities”, or groups of users and creators. These communities, today often virtual, can include researchers, decision-makers, and the general public; participants in these communities may contribute different components during various iterations in the data generation process, and have different needs for resulting data products. As this “participatory research informatics” model, which includes elements of collaborative research, public participation, and digital data libraries, becomes the norm, new approaches to dynamic data management will be important. Researchers commonly use the web as a medium for research collaboration among colleagues, as well as to disseminate results to decision-makers and the public. In addition, the web can be used to facilitate dynamic collaboration by allowing research-in-progress to be accessed and edited by multiple participants. Indeed, one key benefit of using the web to facilitate research projects is that data can undergo revisions and corrections while being displayed in some form over the web. However, it can be a challenge to keep track of multiple edits to files while maintaining the integrity of a web-interface. Here we describe a system for promoting collaboration in the creation, maintenance, and use of dynamic data over the web. Using simple Microsoft Excel spreadsheets combined with traditional relational databases to facilitate dynamic data generation and updates and maintain a data schema, we are able to facilitate efficient collaboration within and among participant communities. This approach allows for the separation of web display and content, which in turn allows participants to be responsible for their own content. This separation simplifies the interaction between those responsible for the content (researchers and other participants), those responsible for the data display (web designers), and those responsible for data management, storage, and retrieval (web programmers). The approach described here is particularly well suited to large, spatially-specific ecological datasets because it can be used to facilitate the “real-time” editing and display of visual data such as web-maps, as well as to encourage multiple participants to both contribute to and interact with “beta” data and content.

© 2007 Elsevier B.V. All rights reserved.

1. Introduction

Successful research often incorporates expertise from a diverse group of individuals with a variety of skills and

backgrounds, as well as from multiple organizations (Kouzes et al., 1996). This type of interaction has been coined “collaboratory”—a combination of “collaboration” and “laboratory” (Wulf, 1993). Research within the ecological sciences is

* Corresponding author. Tel.: +1 510 643 0409; fax: +1 510 643 3490.

E-mail address: fkearns@nature.berkeley.edu (F. Kearns).

¹ Denotes equal contribution.

particularly dependent on successful collaboratories because of the multiple scales and complexity of the challenges involved, the inclusion of large datasets and rapidly changing technology, and often diverse personnel with highly specialized expertise (Pike et al., 2005). In addition, ecologists are increasingly interested in incorporating community-generated data, as well as early feedback and input from decision-makers, into the research process (Walters and Holling, 1990; National Academy Press, 1997; Haklay, 2003). This type of “participatory research informatics”—where multiple participants both generate and use information—is becoming increasingly commonplace in ecology.

Developing a research, team-based collaboratory project that adequately and appropriately includes these additional participant communities can be a challenge to technical infrastructure. For example, members in a research group and the larger community have diverse technical skills, and technical infrastructure should be able designed to accommodate diverse users. In addition, a dynamic data management system must allow streamlined content management that facilitates rapid updates from multiple groups and/or individuals in order to facilitate the participatory process. Technical constraints can be obviated through use of the Internet, which is an ideal environment to display data that are in “beta” form and encourage others to collaborate in data production, analysis, and/or use. Though most academic research projects have a web presence, to truly take advantage of the dynamic nature of the web, new approaches that enhance and facilitate the research process are key (Schweik et al., 2005). Group collaboration in managing and displaying data via the web can ensure that the most up-to-date information is accessible for viewing and editing, removing a potential barrier to collaborative research (Cummings and Kiesler, 2005), as well as an impedance to interacting with various user groups during the research process.

It is common for large scale projects with multiple users to rely on relational databases for the transactional features that prevent data loss and ease rapid data retrieval. However,

relational databases such as MySQL, PostgreSQL, and Oracle are not easily accessible by those without the ability to use a programmatic interface. Here we describe a data management system that we have developed that addresses the technical infrastructure challenges related to “participatory research informatics”. This system combines dynamic data in a widely used Microsoft Excel format, with the more static content in a MySQL relational database to ease updates, speed access, and increase reliability as data is output via the web (Fig. 1). In our implementation, this data output is often in the form of dynamic web-based maps which are an intuitive means of displaying spatial data (Kearns et al., 2003), but could be anything from tabular data to charts. The novelty of the system is the separation of research collaboratory activities through use of a spreadsheet maintained data schema, which enhances group interaction by, for example, encouraging shared responsibility for tasks such as multi-lingual content and updates. In addition to facilitating the interaction between researchers, this approach can also be used encourage participation by decision-makers, the general public, and other stakeholders.

Here we describe the system we use and describe two case studies where our goal was to facilitate dynamic data creation, display, and editing by multiple groups of users. In both cases, we utilize this system of dynamic data display and editing via the web. We developed, tested, and refined these methods as part of research and outreach projects undertaken by a group of researchers with the recently established Center for Fire Research and Outreach (CFRO) at the University of California, Berkeley. A key goal of CFRO is to encourage and facilitate collaboration on wildfire-related research questions among academics, practitioners, decision-makers, and government agencies, as well as to provide the general public with wildfire-related information (Kearns et al., in press). The group is composed of faculty, post-docs, researchers, field technicians, web designers, and programmers, all with varying levels of computing and web skills. In addition to communicating within our research group, we have a diverse set of

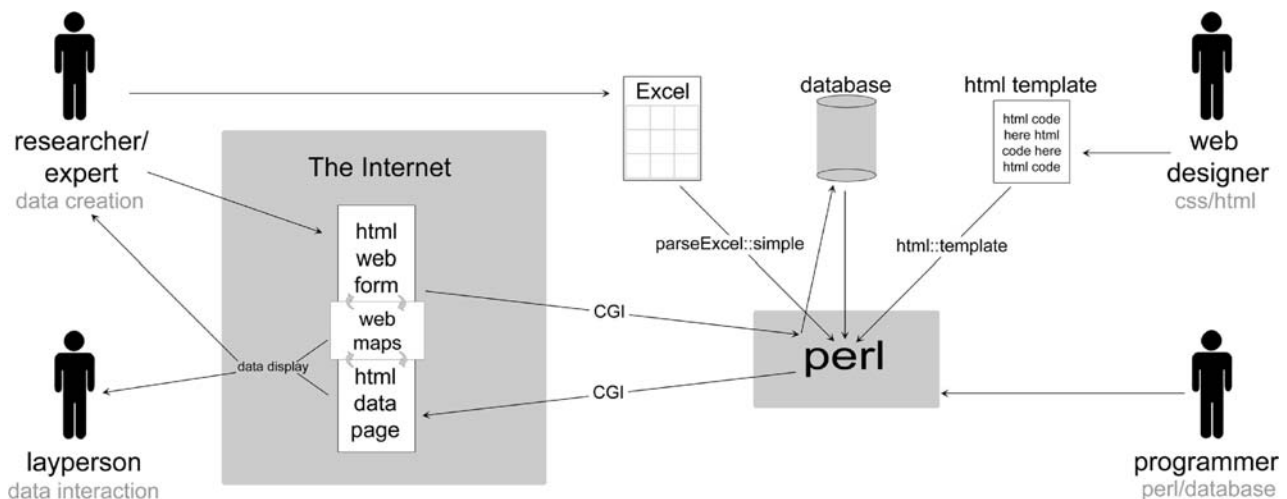


Fig. 1—A sketch of the flow of information in the system. Once the system is set up with the Perl and database interaction, the researcher can add and modify data in the Excel workbook. All the data is assembled from the database, the HTML template, and the spreadsheet by Perl and sent to the web as a map or table. Data entry also occurs through web-forms which are parsed by Perl and entered into the database.

collaborators and end-users, as well as multiple projects that involve community participation in data gathering; therefore, the goals of the group require the inter-disciplinary collaboration method that we describe here. The primary aim of this paper is to describe a simple schema for creating and managing dynamic web content and facilitating the participatory research process across a broad range of users.

2. General system structure

The primary purpose of the system we use is to ease the separation of content from presentation (<http://alistapart.com/articles/separationdilemma>). The separation of content and presentation allows each participant to be independently responsible for specific tasks. The method we have utilized to provide this separation of content, display, and data management relies on Perl programming language scripts to collect information from the dynamic Excel spreadsheets and the more static database, and then create the dynamic HTML/web content using a templating system. Data entered interactively via HTML pages are handled by the Perl scripts and returned to the database. That newly entered data is then immediately available as content for the web page. Often, it is data stored solely in a relational database such as MySQL or Oracle. Although relational databases can improve organization and the speed by which data can be stored, queried, and accessed, they require programming knowledge for even basic data interaction. This introduces a barrier between the researcher and the data or content for which she is responsible.

Due to the relative ease and frequency of use by researchers, Microsoft Excel is a good option to store and edit dynamic data in addition to the static, or web-entered data which is stored in a relational database. The use of Microsoft Office is ubiquitous in science and education. Excel provides a nearly universal editing environment that is familiar to most and fully editable by open source and free software (e.g., OpenOffice) for those who do not use a Windows platform or do not own Microsoft Office. While there are alternatives to Excel such as Google spreadsheets, the Google spreadsheets API is too slow for dynamic generation of web pages, and is not currently as widely used as Excel. In our implementation, the Excel spreadsheets do not contain raw data, only questions and explanations, and thus they remain small (for example, an Excel workbook might have less than 10 worksheets, each worksheet with an average of three columns, and a maximum of 25 questions (or rows)). The responses to these questions, which may number in the thousands (or more) are stored in a MySQL database. Using this familiar program facilitates group participation, as even those completely unfamiliar with web development can make changes in a comfortable environment.

The dynamic content from the Excel spreadsheets is used to populate HTML templates, which are a common means of presenting dynamic data in a consistent manner. The templates we employ add an additional markup syntax to an HTML file with an ID (identification) that links it to a worksheet and cell. Within the Microsoft Excel worksheets are columns for the question ID, question text, question help (which is displayed via a “tool-tip”), the possible responses to

the question, images to be displayed with the question, links further explaining the question, and the text explaining the solution to that question. The question ID is used to link the HTML, the code in the web page, and the Perl code that handles the interaction of database, Excel, and HTML. The content of the Excel spreadsheet is parsed using a Perl Module, `Parse::SpreadSheet::Excel::Simple` which reads the sheet and returns a Perl data structure which can then be manipulated in Perl. Another Perl module is used to separate the content in the Excel worksheet from its display in the web browser: `HTML::Template`. This module provides named “hooks” that correspond to the question id in an HTML file. For example, the token `<TMPL_VAR NAME=“ROOF_ID_4”>` can be replaced by the Perl program with the question text from the Excel sheet that corresponds to question ID of “ROOF_ID_4”. Many such “template parameters” are included in the HTML template file to link question ID’s to actual content (text, images, formatting) in the corresponding HTML file that is included dynamically each time the web page is loaded. Because the token (or hook) already exists in the template, it allows a web designer to create the layout of the page independent of the exact and/or final content. As long as the programmer knows the name of the template parameter, it can be replaced with dynamically generated content by the Perl script. Since the markup is only a place holder for the data stored in the worksheet, the worksheet can be modified by the researcher without changing the template or the Perl code that generates the HTML from the template and the Excel sheet. In fact, the templates support looping so that questions cannot only be altered, but they can be added and removed without change to the Perl code, or the HTML template.

A web designer is able to make changes to the way the content is displayed: layout, font, and images. A programmer is able to make changes to the handling and storage of the data in the MySQL database. The key is that each of these occurs independently with a series of Perl programs linking them all together and enabling immediate display of modified data.

In addition to the Excel spreadsheet schema, we have made as many accommodations as possible for a wide variety of users on the “client side” of this system. For example, because many of our users may be in rural areas with slow connections, web pages were consistently optimized for both speed and style. Using cascading style sheets, automatically generating print versions of many pages were created, all text (except that embedded in images) is fully resizable, all images include alternative text that can be read by text readers for the visually impaired, and the color scheme is distinguishable by the color blind.

3. Case studies

3.1. Case study 1: Wildfire hazard assessment and ranking

One of the first large CFRO projects has been the development of a parcel-level, science-based wildfire hazard assessment and ranking approach for communities and homeowners (sai[Frontiera et al., submitted for publication]), including a web-based Geographic Information System (webGIS) interface for data exploration. This approach

web version | printer-friendly version

Fire Information Engine

Homeowner Wildfire Assessment

BETA



For explanatory information on specific questions, roll your mouse pointer over the question. You can also click here to see the glossary, which contains definitions for specific terms. Click here for Frequently Asked Questions.

ROOFING

1. Is your roof combustible? yes no n/a
2. Does your roof have any unstopped or unscreened open roof tiles? yes no n/a
3. Is your roof, or a portion of your roof, in poor condition, potentially exposing the underside of your roof? yes no n/a
4. Is your roof valley design something other than woven that may expose the underside of your roof? yes no n/a
5. Do you have a flat roof or a complex roof that may create vulnerable corners and/or surfaces? yes no n/a
6. Do you have any unscreened chimney openings? yes no n/a
7. Are your roof gutters unscreened? yes no n/a
8. Is there any combustible vegetation, vegetative debris, or other combustible material on or touching the roof and/or are three tree limbs within six feet of the roof? yes no n/a

VENTS

1. Do you have any vents on the outside of your home that are unscreened and/or screened with a mesh size that is less than 1/4-inch? yes no na
2. Do you have vents located in areas where you also have combustible or poor condition siding? yes no n/a
3. Do you have any gaps around your exterior doors? yes no n/a
4. Do you have any gaps around your garage doors? yes no n/a
5. Do you have any other gaps located on your home? yes no n/a
6. Do you have any man-made fuels located alongside vents and/or gaps in your home? yes no n/a
7. Do you have any combustible vegetation, vegetative debris, or other combustible materials located alongside vents and/or gaps in your home? yes no n/a

Fig. 2 – Web-based form allows homeowners to answer a series of questions about their homes; once the answers are submitted, they get back a “report card” outlining suggestions for improving their vulnerability to wildfire.

is available via the web as an HTML form (Fig. 2—<http://firecenter.berkeley.edu/homeassessment/>). Once the form is filled out by a homeowner or community official, wildfire hazards are categorized as low, medium, or high, and a “report card” that explains each hazard and outlines suggestions for mitigating those hazards is generated. The results can also be quickly mapped to a simple Google Maps interface (Fig. 3).

3.2. Dynamic content

Both the wildfire hazard assessment web-form and its output display were revised numerous times during the research process based on user and researcher input. At the same time, the hazard assessment web-form and report card have been available in whatever their current forms on the web. The methods that we have developed facilitated the incorporation of changes by participants involved at any level of the project. For example, a researcher is able to make changes to the content by adding, removing, and revising questions and the rating (high/medium/low) of a given response to a question. In addition, anyone in the group could add the url of an image in

the Excel spreadsheet, along with some explanatory text, which would then be included in the web page. A community member is then able to enter the results of a self-assessment of their property via the web.

It should be noted that this project was developed by a relatively small group—one programmer, one web developer, and approximately five researchers. The group meets weekly over the course of a year to develop the project and come to agreement on terminology and definitions. While in this regard, this process was relatively easy because of the size of the group, a larger group and a project with “legacy” data faces a larger challenge, as described in the second case study below.

3.3. The system

Knowing that the content of the hazard assessment and rating system was under constant development, we stored the questions, and the suggestion for mitigation matching each fire hazard as simple text in a Microsoft Excel spreadsheet, which is easily editable by anyone. More static content is stored in a traditional relational database (MySQL) to ensure that the application is responsive. A simple versioning system

BETA Fire Information Engine
know your risk / lower your risk

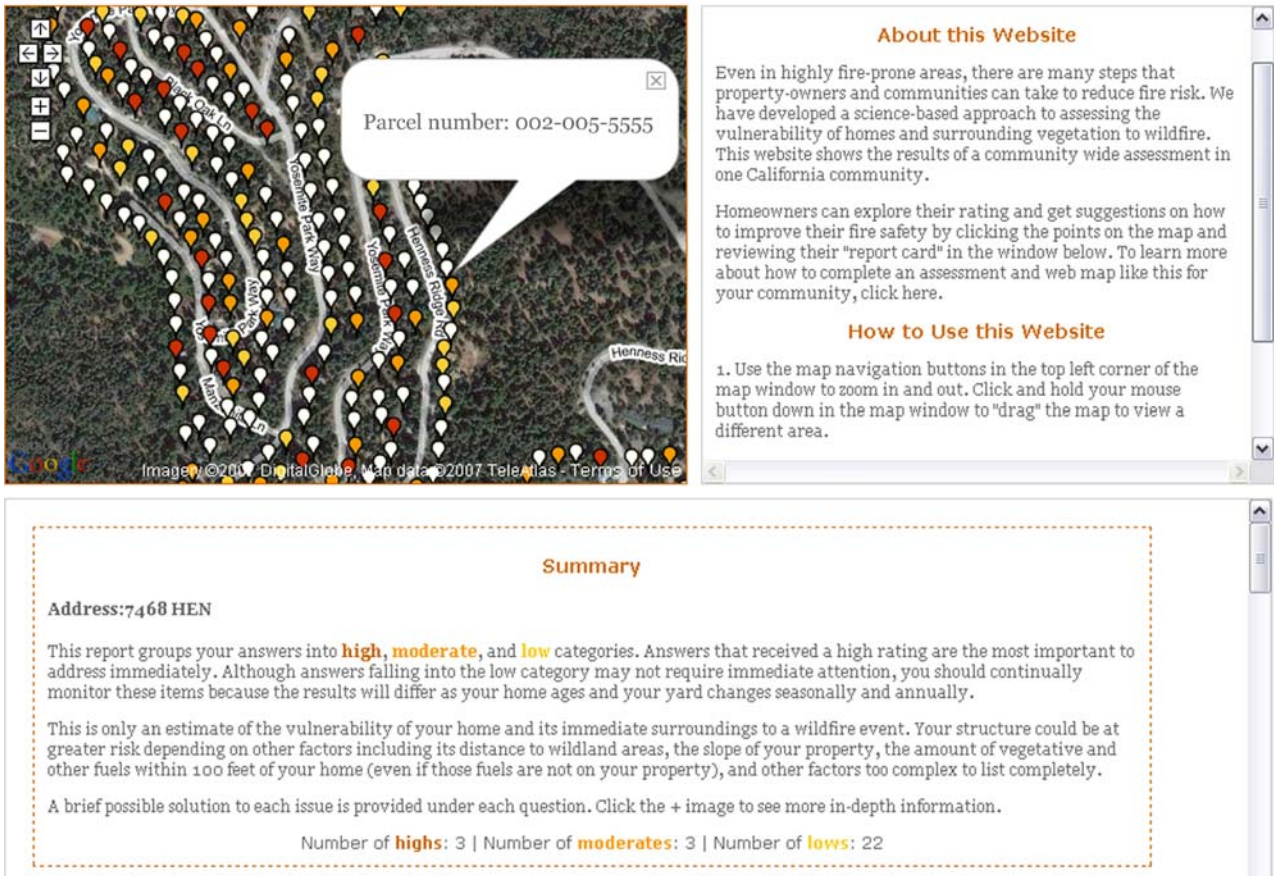


Fig. 3 – Google Maps interface displays the results of either community- or expert-generated parcel-level fire hazard assessment results.

was introduced for the dynamic Excel sheets where, for example, fireform_v11.xls would be version 11 of the form and report card. If changes were made, the new version would be named, for example, fireformv12.xls, though the web page would default to version 11 until it was deemed that version 12 was stable enough for public display. At any point, any recent version of the form can be seen at the URL above by adding "index.pl?VERSION=10". This allows us to "remember" the state of the form for communities that adopted the rating system at a particular step in the process.

Our first pilot community for the parcel assessment rating system using the above methodology was in a remote area where the accuracy of the handheld GPS was poor. For this reason, the spatial data had to be checked and updated multiple times throughout the process. Adding new data to a webGIS can require multiple steps and multiple people, delaying the instant visual feedback of the interactive web-map. To alleviate this delay, and the need for multiple people to be involved in updating spatial data, we developed a system where those responsible for the data in the shapefile (a common spatial data format), can upload the shapefile to the server and immediately see and query the uploaded data in a public webGIS. As an added benefit, this same system actually allows us to provide a plug-in for participating communities

that wish to adopt our rating system. An interested community can download a version of our form that, when used on a handheld GPS will create a shapefile with answers (once filled) to the questions for our parcel-level rating system. This same shapefile can then be uploaded to the website and the simple queryable map mentioned above is displayed and a link provided so that information can be disseminated by the community without the need for creating their own webGIS. Again, the usefulness of this type of data management system is the instant feedback that arises when each participant is responsible for her own data, and that data can be viewed and checked immediately through the website.

3.4. Case study 2: Global fire assessment

The Global Fire Partnership (GFP) is an international research collaboration between The Nature Conservancy (TNC), the World Conservation Union (IUCN), the World Wildlife Fund (WWF), and our Center for Fire Research and Outreach (CFRO). The overall objective of the research project was to complete a global assessment of native fire regimes that can be used in decision-making related to biodiversity conservation. A 2004 meeting of global fire experts and policy makers resulted in a coarse scale assessment of the extent to which fire is

beneficial or harmful to biodiversity; the results of this workshop are now called the Global Fire Assessment ([The Nature Conservancy, 2004](#)). Since that first workshop, the GFP has been focused on improving the assessment results by hosting a series of three international workshops aimed at gathering participatory knowledge for specific regions of world. The first workshop was held in February 2006 in Berkeley, California, US; the second in April 2006 in Bogor, Indonesia; and the third in July 2006 in Santiago, Chile.

CFRO researchers were invited to help to refine the assessment methodology and aid in related outreach efforts by helping to organize and participate in these workshops, as well as develop a multi-lingual web-interface to display a series of questions on the fire characteristics of a particular ecoregion and a database to store the responses. After logging in to the web-interface in the desired language, participants choose an ecoregion from a simple custom mapping interface developed specifically for this project. They are then directed to a page where they can enter data about the dominant fire regime, the status and trend of the dominant fire regime, major threats to the fire regime, and potential actions to alleviate those threats. All data from the entry form is then stored in a database and is immediately available and retrievable as a large table via the web-interface.

3.5. Dynamic content

During the first workshop in Berkeley, CA, which focused on refining fire regime data for the US, Canada, and Mexico that had been collected as a first version of the assessment, participants were divided into groups by the geographic region with which they were most familiar. They were then given a series of paper

maps and worksheets, as well as a workstation with ESRI ArcMap, and asked to update the shapefile by ecoregion. The workshop generated a wealth of information. Organizing and updating that new data proved to be a challenge, largely because participants were editing ecoregions that crossed national boundaries—leaving multiple entries for a single region—and because fields were updated inconsistently by various participants. Following that meeting, it became apparent that: (1) it was laborious to compile data from multiple shapefiles with inconsistent entries, and (2) the simple table structure (rows, columns) imposed by the shapefile was not an ideal way to store the complex data. In fact, the shapefile had over 80 columns of data, many of which were empty. In addition, because participants interpreted questions differently, the answers had to be “normalized”, introducing a level of separation between the actual reply, and the data saved in the shapefile.

For this project, we were faced with a much larger group of researchers/experts, and with the issue of “legacy” data—data that had been a part of the first global assessment, as well as data collected during the workshop. In addition, the definitions for various terms were also debated during the workshop and in continuing emails and subsequent meetings. In the end, we decided that although participants might not agree on all definitions, as long as we could find a clear way for users to reference definitions while thinking about a given issue, we could minimize confusion.

3.6. The system

Following this first workshop and in anticipation of two more workshops, we began by developing a web-interface and a relational database schema that would simplify data entry,

The screenshot shows a web-based data entry form for the Global Fire Assessment project. The interface is divided into several sections:

- Navigation:** A menu at the top with five items: 1. About, 2. User Information, 3. Choose an Ecoregion, 4. Enter Data, and 5. See Previous Entries.
- Language:** Options for ENGLISH and SPANISH.
- Fire Regime Characterization:**
 - Question 1: "Please choose the dominant fire regime for this ecoregion:" with radio buttons for Fire dependent, Fire sensitive, Fire independent, and no data.
 - Question 1a: "If the dominant fire regime is fire dependent/influenced or fire sensitive, are the severity and frequency (Choose one):" with radio buttons for Low to mixed severity, 0 to 35 year frequency; Replacement severity, 0 to 35 year frequency; Mixed severity, > 35 year frequency; Replacement severity, 35 to 200 year frequency; Replacement severity, > 200 year frequency; and No data.
 - Question 1b: "Are there embedded ecosystems within the ecoregion that have a different fire regime?" with radio buttons for Yes and No (selected).
- Fire Regime Status and Trends:**
 - Question 2: "Please choose the dominant fire regime status and trend for this ecoregion:" with radio buttons for Intact/stable, Intact/declining, Very degraded/improving, Degraded/improving, Degraded/declining, Degraded/stable, Very degraded/stable, Very degraded/declining, and No data.
 - Question 2a: "If the dominant fire regime status is degraded or very degraded, please choose the primary way it is degraded (choose one):" with radio buttons for Fires are too frequent, Fires are too infrequent, Intensity is too low, Intensity is too high, Size is too big, Size is too small, Fire occurs outside of natural seasons, and No data.

The background of the form features a world map with orange and yellow highlights indicating fire-prone regions.

Fig. 4—Web-based data entry form for the Global Fire Assessment project. Experts choose an ecoregion via a simple map interface and then answer questions about the fire regime based on expert knowledge and ancillary data.

reduce confusion, and normalize potential answers while still providing participants the flexibility to respond thoroughly. The basis of the web-based questionnaire is a series of fire threat categories and potential actions to mitigate those threats (Fig. 4). The questions and sections of that questionnaire were entirely decided by the researcher. Experts on a particular ecoregion can go to the web-form and choose the top three threats and actions for that ecoregion. In addition, information such as the dominant fire regime, the frequency of fire, and characteristics of sub-regimes within the larger ecoregion are saved in the database. This was all framed in a webGIS to provide an intuitive spatial context for the data being entered.

As with the wildfire hazard assessment project, this has been a dynamic process. Multiple TNC and CFRO participants have changed the number and wording of questions, the categories of answers, the threat and action categories, and incorporated feedback from participants. Rather than having a chain of editing wherein requested edits and a programmer edits the database—a process found to be error prone and time consuming—we were able to give them full control over the wording of the questions, as well as the list of potential responses (for those questions that have a number of choices) via the use of an Excel workbook. In fact, the workbook is used as a sort of mini-database, where each sheet is a table, and there are columns for a question identifier, question text, and question help. This is depicted for a single sheet in Table 1 for English (a) and Spanish (b). For example (and see Table 1a):

FS	Fire Sensitive	Ecosystem structure and composition tend to inhibit...
----	----------------	--

Where “Fire Sensitive” is one of four options describing the dominant fire regime. If the expert chooses this option, then all that is saved in the MySQL database for their response is

Table 1a – Example content from a Microsoft Excel spreadsheet used to generate content for the website

a. English translation		
FD	Fire dependent	Fire is essential in maintaining predominant ecosystem composition, structure, function and extent. If fire is removed, or if a fire regime is altered beyond its historical range of variability, the ecosystem changes to something else; dependent species and their habitats decline or disappear. Fire is not a disturbance but rather an essential process.
FS	Fire sensitive	Ecosystem structure and composition tend to inhibit ignition and fire spread. The majority of species generally lack adaptations to respond positively to fire. Fire can influence ecosystem structure, relative abundance of species, and/or limit ecosystem extent. Fire may create habitats for key species by creating gaps, regeneration niches, or by initiating or affecting succession.
FI	Fire independent	Fires characteristically would not occur because of a lack of fuel and/or ignition sources. Fire regimes can be altered by a change in fuels (e.g., invasive species) or ecologically-inappropriate human-caused ignitions.
ND	no data	No data about the dominant fire regime.

Table 1b

b. Spanish translation		
FD	Dependiente del fuego	El fuego es esencial para el mantenimiento de la composición, la estructura, la función y la extensión del ecosistema predominante. Si se eliminan los incendios, o si el régimen del fuego se altera más allá de su rango histórico de variabilidad, el ecosistema cambia a otra cosa; las especies dependientes y sus hábitats se deterioran o desaparecen. El fuego no es un disturbio, sino más bien un proceso esencial.
FS	Sensible al fuego	La estructura y la composición del ecosistema tienden a dificultar la ignición y la propagación del fuego. Por lo general, la mayoría de las especies carece de las adaptaciones necesarias para responder al fuego de manera positiva. El fuego puede influir en la estructura del ecosistema, la abundancia relativa de las especies y/o limitar la extensión del ecosistema. El fuego puede crear hábitats para especies clave mediante la creación de vacíos, rincones de regeneración o iniciando o afectando la sucesión.
FI	Independiente del fuego	Característicamente, el fuego no ocurre debido a la falta de combustible y/o fuentes de ignición. Los regímenes del fuego pueden ser alterados por un cambio en los combustibles (por ej., especies invasoras) o igniciones ecológicamente no adecuadas causadas por seres humanos.
ND	No hay datos	No hay datos sobre el régimen de fuego predominante.

Multiple languages are supported simply by adding a workbook with the appropriate translations. Since all data is stored in the relational database only by the id in the first column (FD, FS, FI, ND) regardless of the language in which the data was entered, the data could later be displayed in any language just by referencing the ID in the first column.

the question id: ‘FS’. When looking up the response, FS can later be associated with the question text, or the explanatory text using the Excel workbook. This is in contrast to the shapefile where all information had to be stored in order to be accessed later.

In addition, this method had a further benefit which we encountered serendipitously: providing a multi-lingual site only requires that we switch from using the spreadsheet with the English translation to those of another language. For example, the entry in the Spanish worksheet corresponding to the English one above is (and in Table 1b):

FS	Sensible al fuego	La estructura y la composición del ecosistema tienden...
----	-------------------	--

Where “Sensible al fuego” is displayed as the text in the HTML page in place of “Fire Sensitive” that would have been displayed in the English version. Again, if the researcher chooses this option, all that is stored is “FS”. The implications of this are that data can be entered in Spanish, and viewed in English or vice-versa. In fact, all that is required to add

additional language is to provide a translation of the Excel workbook. Once that file, named, for example, *german.xls*, is included in the appropriate directory on our server, then a link will automatically appear at the top of each HTML page (next to the current Spanish and English) to view that page in German. This is a useful consequence of maintaining the separation between content and delivery.

4. Discussion and conclusions

Ecological research is often conducted by multiple collaborators at distant sites, working across various spatial and temporal scales, in data-rich and dynamic environments (Turner et al., 2006). As this collaborative model becomes more common in scientific research, it becomes increasingly important to also understand the factors necessary for successful scientific collaboration. While some researchers stress the social dimensions of successful collaboration—for example, personal compatibility, work connections, and incentives (Hara et al., 2003)—the importance of computer infrastructure that facilitates interaction between many diverse participants cannot be under-emphasized (vanAalst and vanderMast, 2003; Fruchter, 2004).

Indeed, we originally implemented this data management system because there was frequent transferring of data between those who were involved in the projects. Every time a change in content was made by a researcher, the programmer had to change the code, and the designer had to change the display. By separating each of these roles, we were able to create a more dynamic and less error prone system, as well as reduce the redundancy of tasks and allow less experienced members of the group make more effective contributions. The most novel and useful part, for us, has been the use of simple text content stored in an Excel workbook to define the content of a web page. This gives anyone, regardless of programming and/or web design knowledge, responsibility for the content of a site. This means that if a researcher does not like the wording of a question, or notices a misspelling, all that is required is to download the latest version of the Excel worksheet, update it, increment the version number, and re-upload to the appropriate directory (it should be noted that all members of the group were considered “expert” enough to make these types of changes). The programmer and the designer can both be completely unaware of this change, thereby creating a less redundant and more productive work environment. Likewise, the designer can make changes to the display of the web page without modifying the content. We have used this system in other projects, and it is easily adopted. The addition of the versioning system makes it usable even in systems where others are dependent on the data as any version is available at any time in the development process.

Another critical component of the participatory research model that is facilitated by this approach is the inclusion of outside review and input from multiple participants. As with many research centers, our group is interested in community involvement in the research process, whether it be to aid in the data gathering process, to get feedback and input into a particular project, or simply to disseminate research results. Indeed, our system easily allows for and encourages interac-

tion from participants outside of the immediate research “team”. At a basic level, because data entered is immediately available as web content, users are able to immediately view and provide feedback. However, this type of dynamic data facilitation also makes it possible to encourage potential users to become actual participants in the research process. For example, the wildfire hazard assessment data gathered by communities, as outlined in the first case study, can be analyzed by researchers and/or the community itself to determine what type of fire-related problems occur most frequently. With that knowledge, participants can encourage local cooperation, as well as push their decision-makers for needed changes to reduce wildfire risk.

In the case of the Global Fire Assessment (GFA) described in the second case study, not only is most of the data participant-generated, but current data can also be downloaded as a shapefile so that anyone interested in the project can do their own spatial analysis or make their own maps (the unit of data entry for this project is a global ecoregion). Since all entries are saved, regardless of redundancy, the data can be analyzed to see, for example, which ecoregions were assigned different fire regimes by different participants, or which fire regime threats changed through time.

The range of possible responses to both case studies is essentially infinite as each expert has different knowledge of any given study area, as well as a different means of expressing that knowledge. This unstructured freedom can lead to unmanageable and unanalyzeable data files. To remedy that, we have constrained the responses to questions contained in the templates. This means that there are common definitions of data fields, and a limited range of responses to most questions. In effect, for client-side users of the system, this exemplifies a common mantra in web development, “convention over configuration”, where we have limited the options as much as possible to reduce the confusion and problems with data storage. For example, in the case of the GFA, an expert may wish to convey that from 1900–1950, an ecoregion was dominated by fire, while from 1950 on, it has had very few fires. They are able to do so in the free text comment field. However, in the response to the specific question, we limit them to choosing the fire regime from a list of choices. This is important because if all 10K entries have any one of five choices for fire regime, we can analyze and display the data coherently. However, if all responses are complex conditions and non-assertive statements, no such analysis is possible. The free text box is there to retain any information the expert wishes to share, and is available should a more complex analysis be undertaken.

Both of these case studies utilize a dynamic approach. For example, if we were to create a fire vulnerability rating for each parcel in a given community, and leave it at that, there would be less motivation for change. However, more actively involving community participants by having the capacity to immediately reflect the real-world changes that they make to the data via the web promotes action. Likewise, with the GFA, rather than merely creating a tool for one-time data collection, we offer the means to update and display the latest content in a simple format while still retaining all records should more detailed analysis be desired. By doing this, we provide the tools for the participants to create change, rather than retaining

stale data in an obscure database. In essence, this approach gives “ownership” of data to participants. This ownership is key, particularly in natural resource management, because, ultimately, action has to be taken by landowners (including the “owners” of public lands), not researchers.

This type of information collaboration and exchange with diverse participants is situated within the solidifying global knowledge infrastructure that has become widely accessible through wireless, wiki, blog, and related communication technologies. Several specific fields within the sciences have already begun to capitalize on this. For example, the field of adaptive environmental management, which is increasingly being tried in the United States (Stokstad, 2005), faces challenges common to other natural resource management programs in areas characterized by natural, social, and historical complexity. Successful adaptive management requires the sharing and discussion of information about the human and natural components of the system being managed (Walters and Holling, 1990; Haklay, 2003). New tools such as Internet message boarding, “counter-mapping”, and Geographic Information Systems (GIS) have been used separately to gather new information, optimize the use of available information, and ensure that all parties can effectively participate in the decision-making process (Aberley, 1993; Crampton and Stewart, 2004; Sieber, 2002). These tools collectively form a distributed model of communication between diverse and distributed participants, and are critical components of a natural resource adaptive management process (Meredith et al., 2002).

A second example of sub-fields within the sciences already immersed in these issues is in the context of the geographic information sciences and specifically the challenges of participatory GIS (PGIS). Examples of PGIS cover a variety of natural resource issues, including wetland management, coastal planning, conservation and forest diseases (Jordan, 1998; Aspinall, 2002; Ball, 2002; Craig et al., 2002; Kyem, 2002; Kelly and Tuxen, 2003; Voss et al., 2004; Baker, 2005; Mathiyalagan et al., 2005). The proliferation of interactive geospatial tools, including web-based GIS, has led to the evolution of two paradigms of public interaction. First, these technologies can lead to empowerment of individuals through better and more rapid access to public data, the ability to present data in more persuasive ways, improve communication and technologies to support collaboration, and the power of the Internet for rapid publication and dissemination of ideas and data (National Academy Press, 1997; Crampton and Stewart, 2004). However, these technologies can also marginalize people and communities through increased restrictions of public access and increased privatization of data (Harris and Wiener, 1998); legal, financial and technical restrictions (Sieber, 2003); and barriers produced by differing paradigms of interpretation, the incorporation of non-standard data, and differing modes of communication (Meredith et al., 2002).

Although what we have described here goes beyond geographic information, some of the same possibilities for marginalization exist, and need to be honestly addressed by work such as ours. In the case of the GFA, we provided multi-lingual content in a web-interface which can be accessed from anywhere, with data stored on our server, to reduce the load on the users’ system. However, because both projects that we

have described are web-based, they are largely limited to participants who have access to the web. In the case of the wildfire hazard and assessment project, because we worked with a community in a rural area with limited access to the internet beyond dial-up modems, we countered this by mailing surveys to participants, encouraging them to complete the survey online if they had access, or mailing it back to us if not; we also mailed the results of a community-wide survey to the participants. While these kinds of actions may compensate for some concerns about participatory research, there remains an undeniable divide between those with quick and easy access to technology, and those without. Here we propose a new term, “participatory research informatics” (PRI), which we hope will allow the discussion about participatory research to expand into the larger research community. The PRI system we have presented was created during two substantial participatory research projects. This approach provides a simple organization and separation of data that promotes participation amongst groups with diverse, and potentially non-technical backgrounds. As PRI becomes more pervasive, dynamic, accessible systems such as this one will be increasingly useful.

REFERENCES

- Aberley, D., 1993. *Boundaries of Home. Mapping for Local Empowerment*. New Society Publishers, Gabriola Island, BC.
- Aspinall, R., 2002. A land-cover data infrastructure for measurement, modeling, and analysis of land-cover change dynamics. *Photogrammetric Engineering and Remote Sensing* 68, 1101–1105.
- Baker, T.R., 2005. Internet-based GIS mapping in support of K-12 education. *Professional Geographer* 57, 44–50.
- Ball, J., 2002. Towards a methodology for mapping ‘regions of sustainability’ using PPGIS. *Progress in Planning* 58, 81–140.
- Craig, T., Harris, T.M., Weiner, D. (Eds.), 2002. *Community Participation and Geographic Information Systems*. Taylor and Francis, London, pp. 205–217.
- Crampton, J.W., Stewart, D.J., 2004. Community mapping as a solution to digital equity. In: Janelle, D.G. (Ed.), *WorldMinds: Geographical Perspectives on 100 Problems*. Kluwer Academic Publishers, pp. 523–527.
- Cummings, J.N., Kiesler, S., 2005. Collaborative research across disciplinary and organizational boundaries. *Social Studies of Science* 35, 703–722.
- Frontiera, P.L., Kearns F.R., Leonard J.L., Quarles S.L., Moritz, M.A., submitted for publication. A new approach to assessing parcel-level vulnerability to wildfire in the wildland–urban interface. Under revision, submitted to *Landscape and Urban Planning*.
- Fruchter, R., 2004. Degrees of engagement in interactive workspaces. *AI & Society* 19, 8–21.
- Hara, N., Solomon, P., Kim, S., Sonnenwald, D.H., 2003. An emerging view of scientific collaboration: scientists’ perspectives on collaboration and factors that impact collaboration. *Journal of the American Society for Information Science and Technology* 54, 952–965.
- Haklay, M.E., 2003. Public access to environmental information: past, present and future. *Computers, Environment and Urban Systems* 27, 163–180.
- Harris, T., Wiener, D., 1998. Empowerment, marginalization, and “community-integrated” GIS. *Cartography and Geographic Information Systems* 25.

- Jordan, G., 1998. A public participation GIS for community forestry user groups in Nepal: Putting people before the technology, Varenus Specialist meeting—"Empowerment, Marginalization and Public Participation GIS", Santa Barbara, CA.
- Kearns, F.R., Goldstein, N., Pedersen, B.S., Moritz, M.A., in press. The Fire Information Engine: A web-based toolkit for wildfire-related needs. In press, accepted to *Journal of Map and Geography Libraries*.
- Kearns, F.R., Kelly, N.M., Tuxen, K.A., 2003. Everything happens somewhere: using webGIS as a tool for sustainable natural resource management. *Frontiers in Ecology and the Environment* 1, 541–548.
- Kelly, N.M., Tuxen, K., 2003. WebGIS for monitoring "Sudden Oak Death" in coastal California. *Computers, Environment and Urban Systems* 27, 527–547.
- Kouztes, R.T., Meyers, J.D., Wulf, W.A., 1996. Collaboratories: doing science on the internet. *Computer* 29, 40–46.
- Kyem, P.A.K., 2002. Promoting local community participation in forest management through a PPGIS application in Southern Ghana. In: Craig, W.J., Harris, T.M., Weiner, D. (Eds.), *Community Participation and Geographic Information Systems*. Taylor and Francis, New York, pp. 218–231.
- Mathiyalagan, V., Grunwald, S., Reddy, K.R., Bloom, S.A., 2005. A webGIS and geodatabase for Florida's wetlands. *Computers and Electronics in Agriculture* 47, 69–75.
- Meredith, T.C., Yetman, G.G., Frias, G., 2002. Mexican and Canadian case studies of community-based spatial information management for biodiversity conservation. In: Craig, W.J., Harris, T.M., Weiner, D. (Eds.), *Community Participation and Geographic Information Systems*. Taylor and Francis, New York.
- National Academy Press, 1997. *The Future of Spatial Data and Society*, NRC Commission on Geosciences, Environment, and Resources, Washington DC.
- Pike, W., Yarnal, B., MacEachren, A.M., Gahegan, M., Yu, C.Q., 2005. Retooling collaboration: a vision for environmental change research. *Environment* 47, 8–21.
- Schweik, C.M., Stepanov, A., Grove, J.M., 2005. The open research system: a web-based metadata and data repository for collaborative research. *Computers and Electronics in Agriculture* 47, 221–242.
- Sieber, R.E., 2002. Geographic information systems in the environmental movement. In: Craig, W.J., Harris, T.M., Weiner, D. (Eds.), *Community Participation and Geographic Information Systems*. Taylor & Francis, London, pp. 153–172.
- Sieber, R.E., 2003. Public participation geographic information systems across borders. *Canadian Geographer* 47, 50–61.
- Stokstad, E., 2005. Learning to adapt. *Science* 309, 688–690.
- The Nature Conservancy, 2004. *Fire, Ecosystems and People: a Preliminary Assessment of Fire as Global Conservation Issue*. Tallahassee, FL. October 2004. <http://nature.org/initiatives/fire/science/>.
- Turner, W., Bowker, F., Gasser, L., Zackland, M., 2006. Information infrastructures for distributed collective practices. *Computer Supported Cooperative Work* 15, 93–110.
- vanAalst, J.-W., vanderMast, C., 2003. Performer: an instrument for multidisciplinary courseware teams to share knowledge and experiences. *Computers and Education* 41, 39–48.
- Voss, A., Denisovich, I., Gatalsky, P., Gavouchidis, K., Klotz, A., Roeder, S., Voss, H., 2004. Evolution of a participatory GIS. *Computers, Environment and Urban Systems* 28, 635–651.
- Walters, C.J., Holling, C.S., 1990. Large-scale management experiments and learning by doing. *Ecology* 71, 2060–2068.
- Wulf, W.A., 1993. The collaboratory opportunity. *Science* 261, 854–855.