Ecological Informatics 6 (2011) 217–227

Contents lists available at ScienceDirect

Ecological Informatics

journal homepage: www.elsevier.com/locate/ecolinf

The art and science of multi-scale citizen science support

Greg Newman [a,*], Jim Graham [a], Alycia Crall [c], Melinda Laituri [b]

[a] Natural Resource Ecology Laboratory, Colorado State University, Fort Collins, CO 80523-1499, USA
[b] Department of Forest, Rangeland, Watershed Stewardship, Warner College of Natural Resources, Colorado State University, Fort Collins, CO 80523-1785, USA
[c] The Nelson Institute for Environmental Studies, University of Wisconsin-Madison, Madison, WI 53706, USA

ARTICLE INFO

Article history:
Received 27 July 2010
Received in revised form 23 February 2011
Accepted 14 March 2011
Available online 22 March 2011

Keywords:
Citizen science
Community-based monitoring
Cyber-infrastructure
Ecological data management
Web skins

ABSTRACT

Citizen science and community-based monitoring programs are increasing in number and breadth, generating volumes of scientific data. Many programs are ill-equipped to effectively manage these data. We examined the art and science of multi-scale citizen science support, focusing on issues of integration and flexibility that arise for data management when programs span multiple spatial, temporal, and social scales across many domains. Our objectives were to: (1) briefly review existing citizen science approaches and data management needs; (2) propose a framework for multi-scale citizen science support; (3) develop a cyber-infrastructure to support citizen science program needs; and (4) describe lessons learned. We find that approaches differ in scope, scale, and activities and that the proposed framework situates programs while guiding cyber-infrastructure system development. We built a cyber-infrastructure support system for citizen science programs (www.citsci.org) and show that carefully designed systems can be adept enough to support programs at multiple spatial and temporal scales across many domains when built with a flexible architecture. The advantage of a flexible, yet controlled, cyber-infrastructure system lies in the ability of users with different levels of permission to easily customize the features themselves, while adhering to controlled vocabularies necessary for cross-discipline comparisons and meta-analyses. Program evaluation tied to this framework and integrated into cyber-infrastructure support systems will improve our ability to track effectiveness. We compare existing systems and discuss the importance of standards for interoperability and the challenges associated with system maintenance and long-term support. We conclude by offering a vision of the future of citizen science data management and cyber-infrastructure support.

Contents

* Corresponding author at: NESB/NREL, Colorado State University, Fort Collins, CO 80523-1499 USA. Tel.: +1 970 491 0410; fax: +1 70 491 1965.
E-mail address: newmang@nrel.colostate.edu (G. Newman).

1574-9541/$ – see front matter
doi:10.1016/j.ecoinf.2011.03.002

1. Introduction

Citizen science and community-based monitoring programs are emerging as significant providers of ecological data. These programs measure and monitor streams, lakes, birds, fish, invasive species, biodiversity, climate change, air quality, water quality, macroinvertebrates, astronomy, and even earthquakes (Bonney et al. 2009b, Cochran et al. 2009, Newman et al. 2010, Silvertown 2009a, b). As the number and breath of these programs increase, so does the volume of ecological data they generate (Bonney et al. 2009b). Creating and maintaining online data management systems capable of supporting the varied nature of these data is difficult for most programs. Programs fortunate enough to have their own data management systems still face user interface challenges (Newman et al. 2010) and struggle when their needs grow beyond the specificity of their current data management system.

Program-specific systems are limited to a particular domain (e.g., streams) and may not incorporate data standards or controlled vocabularies necessary for efficient data sharing or system interoperability. The benefits of integrating data from one program with another are often overlooked. For example, meta-analyses to determine climate change effects or species distributions cannot easily be conducted if data standards are not used between all programs measuring similar species and/or attributes. Additionally, given the importance of social interaction for volunteers (Bell et al. 2008a,b), systems focused solely on data entry and storage may overlook important features that facilitate communication, marketing, and social interaction among citizens, volunteer coordinators, and stakeholders (Newman et al. 2010) or that support data analysis and visualization.

Citizen science programs are created for many purposes. Examples include: long term monitoring; scientific research; community networking; social empowerment; science literacy improvement; environmental education; youth career development in science, technology, engineering, and mathematics; community service; and the preservation of traditional ecological knowledge. Citizen science program objectives are equally varied. Examples include: contributing quality data, helping scientists answer questions, informing local decisions, engaging in social networks, and/or offering opportunities to enjoy nature. Meeting these objectives requires data management systems with many capabilities. For example, effective systems must announce training events, offer educational materials, perform automated data quality checks, provide tools for metadata support, automate summary statistics, create reports, enable data uploads and downloads, offer tools for analysis and modeling, exchange data with other databases, and provide decision support capabilities. End users demand flexible systems capable of integrating data across domains and scales while also accommodating diverse needs. Bonney et al. (2009b) articulate these challenges clearly: "… as citizen science [programs] grow in scope, …innovative tools in database management, scientific analysis, and educational research [will be needed], … networking technologies and… database solutions [will be] imperative, [and] computationally efficient geospatial analysis and imaging techniques [will be needed] … to handle … massive amounts of monitoring data … collected across vast geographic scales." Thus, we sought to: (1) briefly review existing citizen science approaches and data management needs; (2) propose a framework for multi-scale citizen science support; (3) develop a cyber-infrastructure designed to support citizen science program needs; and (4) describe the lessons we learned. We compare existing systems and discuss the importance of standards for interoperability and the challenges associated with system maintenance and long-term support. We conclude by offering a vision for the future of citizen science data management, informatics, and cyber-infrastructure support.

2. Existing approaches and data management needs

At the forefront, it is important to review various citizen-based approaches and summarize their respective data management needs. Unfortunately, terminology remains confusing (Table 1) and includes phrases such as Community-Based Monitoring or Citizen-Based Monitoring, Citizen Science, Decision Support Systems, Environmental Decision Support Systems, Environmental Collaborative Monitoring Networks, Volunteered Geographic Information, Participatory Geographic Information Systems, Participatory Monitoring Networks, Public Participation Geographic Information Systems, Indigenous Mapping, Community Networking, Participatory Action Research, and, more recently, Public Participation in Scientific Research. These approaches can be categorized as contributory, collaborative, or co-created (Table 1; Bonney et al. 2009a). For the purposes of this paper, we use the term citizen science broadly to encompass all of these approaches.

Citizen science represents scenarios in which citizens participate in the scientific process along with professionals (Bonney et al. 2009b). Citizen science programs require significant oversight, coordination, protocol development, protocol refinement, training, data management infrastructure, and financial support (Bonney et al. 2009b, Cohn 2008a, b). Some programs focus on public engagement, with goals and objectives less data collection oriented and more policy oriented (Powell and Colin 2008), while others enlist citizens to "volunteer" their personal computers for causes such as monitoring seismic activity (Cochran et al. 2009), celestial bodies (e.g., Galaxy Zoo), or posting disaster information online (Laituri and Kodrich 2008). Preeminent North American examples include the programs coordinated by the Cornell Lab of Ornithology, including Project FeederWatch, PigeonWatch, NestWatch, NestCams, Great Backyard Bird Count, eBird, Celebrate Urban Birds, CamClickr, BirdSleuth, and Birds in Forested Landscapes (Bonney et al. 2009b, Cornell Lab of Ornithology 2008) while exemplar European programs include iSpot for citizen-based nature sharing (McAndrew et al. 2010) and OpenStreetMap for community-based street mapping (Haklay and Weber 2008).

In addition to these notable large-scale programs, however, are countless smaller efforts, including 115+ programs listed in the Citizen Science Central registry (Cornell Lab of Ornithology 2008) and 272+ listed at scienceforcitizens.net (scienceforcitizens.net 2011). Small programs (e.g., programs with less than 100 active volunteer members and a support staff of 10 or fewer) often lack the internal capacity to develop their own online data management system and may benefit most from cyber-infrastructure support. Regardless of situation or size, the data management needs of citizen science programs encompass data beyond mere species observations, such as auxiliary environmental data, participant information, volunteer hours, land manager contact information, training event schedules, species attributes, site characteristics, and user preferences for alerts related to new observations. Citizen science programs require features that support communication; teach field skills online; store field data collected by citizens; offer analysis and reporting capabilities; and collect, store, and analyze standardized program evaluation data in a single comprehensive system.