USGS
science for a changing world

NatureServe.
A Network Connecting Science With Conservation

Promoting Synergy in the Innovative Use of Environmental Data—Workshop Summary

Open-File Report 2018–1104

U.S. Department of the Interior
U.S. Geological Survey

Promoting Synergy in the Innovative Use of Environmental Data—Workshop Summary

By Healy Hamilton, Gerald F. Guala, and Annie Simpson

Open-File Report 2018–1104

U.S. Department of the Interior
U.S. Geological Survey

U.S. Department of the Interior
RYAN K. ZINKE, Secretary

U.S. Geological Survey
James F. Reilly II, Director

U.S. Geological Survey, Reston, Virginia: 2018

For more information on the USGS—the Federal source for science about the Earth, its natural and living resources, natural hazards, and the environment—visit https://www.usgs.gov or call 1–888–ASK–USGS (1–888–275–8747).

For an overview of USGS information products, including maps, imagery, and publications, visit https://store.usgs.gov.

Acknowledgments

Contents

Figures

Tables

Abbreviations

Bd	*Batrachochytrium dendrobatidis*
BISON	Biodiversity Information Serving Our Nation
Bsal	*Batrachochytrium salamandrivorans*
DEM	digital elevation model
EcoINFORMA	ecoinformatics-based open resources and machine accessibility
EDRR	early detection and rapid response
EID	emerging infectious disease
HUC	hydrologic unit code
LANDFIRE	Landscape Fire and Resource Management Planning Tools Project
NGO	nongovernmental organization

Promoting Synergy in the Innovative Use of Environmental Data—Workshop Summary

By Healy Hamilton,[1] Gerald F. Guala,[2] and Annie Simpson[2]

Abstract

From December 2 to 4, 2015, NatureServe and the U.S. Geological Survey organized and hosted a biodiversity and ecological informatics workshop at the U.S. Department of the Interior in Washington, D.C. The workshop objective was to identify user-driven future directions and areas of collaboration in advanced applications of environmental data applied to forecasting and decision making for the sustainability of biodiversity and ecosystem services. Substantial effort to recruit attendees from diverse Federal, State, and private sector organizations successfully attracted participants from 20 Federal agencies and 48 different institutions in the academic, nonprofit, State government, and commercial sectors; the total number of attendees ranged from 100 to 144 during the 3-day workshop. The first one-half of the workshop was divided into 7 plenary sessions and 3 sets of lightening talk sessions organized by sector, providing 48 oral and visual plenary presentations that shared diverse perspectives on biodiversity and ecological informatics, including original biospatial analyses from 6 graduate student map contest winners. The second one-half of the workshop focused on 10 breakout sessions with participant-driven themes from the environmental data sphere and concluded with an address by the Director of the U.S. Fish and Wildlife Service. The workshop was structured to encourage interactivity. About 80–90 percent of attendees provided direct feedback using clicker devices for specific questions related to biodiversity and ecological data uses and needs, and 10 breakout session leaders shared the highlights of their group discussions during the final workshop plenary sessions. Participants were encouraged to use the Twitter hashtag #ShareUrData. Over lunch on day 2 there were 20 simultaneous presentations of tools and apps during a special "tools café" session.

The 10 participant-defined breakout session topics are listed below:

- Ecosystem services and ecological indicators
- Inventory and monitoring
- Biogeographic map of the Nation
- Pollinators
- Invasive species
- Remote sensing
- Drivers of agricultural change
- Citizen science
- Climate
- Hydrology and watersheds

Numerous common themes that emerged from the workshop include the following:

- The vital importance of completing foundational environmental datasets that are nationally consistent and are essential to multiple sectors, such as the Soil Survey Geographic database high-resolution soils data, a minimum 5-meter resolution digital elevation model, national hydrographic data, high-resolution land cover data, time series high-resolution spatial climate data from historical to future time steps, and a national wetland inventory.

- Improved, nationally consistent environmental datasets (integrated with targeted observations) will dramatically advance forecasting capacity and support early warning systems (that is, drought, forest disease); however, multiagency coordination should focus on decision support tools that convey appropriate actions and responses to adapt to, and mitigate, potential negative consequences.

[1]NatureServe.

[2]U.S. Geological Survey.

- Digitizing and providing access to the vast stores of underused historical data that can be leveraged for this purpose is of national importance. Modern computational techniques and the ever-increasing flow of environmental data from ground and remote observations can support improved understanding of environmental change. Success of understanding patterns of change for decision making requires establishing baselines from which change can be measured. The value of digitized historical data is greater than ever before.

- There is a need to recognize the multifaceted potential of citizen science to engage the public in resource stewardship, to create the next generation of science, technology, engineering, math, and environmental leaders, and to have sufficient field personnel to monitor environmental trends, including early detection of alien invasive species, phenological shifts, shifting distribution and abundance of indicator species, and species inventories. The Federal government has an essential role in creating the infrastructure to dramatically improve mobilization of citizen science (and other) data by fostering the following: creation of data standards, creation of nationally consistent framework datasets, vertical integration of observation data, visualization and dissemination of aggregated datasets, and calculation and communication of derived trends.

- Current and near future trends in the availability of remotely sensed data (rapid expansion of satellite fleets and drones) is revolutionizing access to near-real-time ecological data. Targeted integration with ground-based observations and instrumentation has an extremely valuable role in validating remotely sensed data, filling data gaps, improving data quality, and fully realizing the potential of the near-real-time monitoring of environmental indicator trends.

- Integrated management of environmental data at the landscape scale is required even as specific actions on the ground are largely local in nature. The workshop highlighted numerous success stories; however, almost every breakout group pointed out the still-too-fragmented nature of the current data landscape.

- Management and delivery of the necessary data, tools, and analyses to sustain our Nation's environmental capital must be a collaborative effort between Federal, State, and local governments, academia, nonprofits, and the commercial sector, even though the responsibilities of each sector are different.

Introduction

A July 2011 report by the President's Council of Scientific Advisors on Science and Technology, titled "Sustaining Environmental Capital: Protecting Society and the Economy," called for the Federal government to "make better use of the knowledge that exists to inform and guide national, regional, and sectoral policies and management and to contribute to global understanding of ecosystem change" (President's Council of Scientific Advisors on Science and Technology, 2011, p. 3). The suggested process to achieve that global understanding was to "identify the most important data gaps within existing Federal and regional monitoring systems for biodiversity and other ecosystem attributes and clarify priorities and agency roles and funding for filling these," (President's Council of Scientific Advisors on Science and Technology, 2011, p. 6) and to "seek out and encourage partnerships with the private and academic sectors to develop innovative tools for data integration, analysis, visualization and decision making" (President's Council of Scientific Advisors on Science and Technology, 2011, p. iv). A biodiversity and ecological informatics workshop was convened on December 2–4, 2015, to continue the ongoing conversation among environmental data managers from all sectors, to improve communication about environmental data, to generate ideas for future collaboration, and to prioritize achievable goals.

More than 100 attendees from diverse Federal, State, and private sector organizations (table 1; fig. 1) made oral presentations and participated in structured breakout sessions during the 3-day workshop (table 2). This report provides summaries of the plenary themes and breakout session topics, results from participatory electronic questions answered during plenary sessions, and abstracts of the original biospatial analyses from six graduate student map contest winners. Through interactive electronic questioning during a workshop plenary session, almost 60 percent of all participants (90 percent of respondents) said they disseminate data to the public in their work (appendix 4). In addition to the interactive questionnaires administered during plenary sessions, workshop participants' feedback also was expressed using the Twitter hashtag #ShareUrData (appendix 5) throughout the workshop. During one luncheon session, participants demonstrated tools and projects driven by themes from the environmental data sphere. On the final day, breakout leaders reported back in plenary sessions, and the workshop concluded with an address by the director of the U.S. Fish and Wildlife Service.

Table 1. Workshop attendee demographics by sector and agency.

Number of participants from the sectors	Agencies in attendance
65 representatives of important projects and organizations from academia, industry, State and local government, and nonprofits, including several student competition winners.	Agricultural Research Service
	Bureau of Land Management
	Council on Environmental Quality
	Library of Congress
	National Aeronautics and Space Administration
75 representatives from the Federal sector, including participants from 19 Federal agencies.	National Agricultural Statistics Service
	National Oceanic and Atmospheric Administration
	National Park Service
	National Science Foundation
	Natural Resources Conservation Service
	Office of Science and Technology Policy
	Smithsonian Institution
	U.S. Agency for International Development
	U.S. Department of Defense, Natural Resources Program
	U.S. Department of the Interior
	U.S. Environmental Protection Agency
	U.S. Fish and Wildlife Service
	U.S. Forest Service
	U.S. Geological Survey

Figure 1. Chart showing work area of participants at the biodiversity and ecological informatics workshop, December 2–4, 2015, from interactive questionnaire feedback (appendix 4).

Table 2.   Agenda for workshop, "Promoting Synergy in the Innovative Use of Environmental Data," December 2–4, 2015, Department of Interior South Building, 1951 Constitution Avenue NW, Washington D.C.

[LANDFIRE, Landscape Fire and Resource Management Planning Tools Project, EcoINFORMA, ecoinformatics-based open resources and machine accessibility; GIS, geographic information system]

Time	Activity or presentation topic	Session
Wednesday, December 2, 2015		
8:00 a.m.	Individual badges, Wi-Fi and clicker registrations	Registration.
8:30 a.m.	Conference welcome and review of logistics (Healy Hamilton, chief scientist, NatureServe)	Opening plenary.
8:45 a.m.	Welcome (Bruce Rodan, assistant director for environmental health, Office of Science and Technology Policy)	
9:00 a.m.	Overview—Expectations and goals for the data synergy workshop (Stinger Guala, EcoScience Synthesis Branch lead, U.S. Geological Survey)	
9:15 a.m.	Plenary—Federal resource management agencies—Current monitoring efforts and immediate needs (moderator: Elizabeth Sellers, U.S. Geological Survey)	Plenary session 1—Monitoring in Federal resource-management agencies—Current efforts and immediate needs.
9:15 a.m.	U.S. Forest Service, Forest Inventory and Analysis Program (Greg Reams, national lead, Forest Inventory and Analysis Program, U.S. Forest Service)	
9:30 a.m.	Natural Resources Conservation Service, Ecological Site Description System and Natural Resources Inventory (Joel Brown, lead, National Ecological Site Team, Natural Resources Conservation Service)	
9:45 a.m.	LANDFIRE (Jim Smith, LANDFIRE lead, The Nature Conservancy)	
10:00 a.m.	National Park Service inventory and monitoring (Joe DeVivo, deputy chief for science, Inventory and Monitoring Division, National Park Service)	
10:15 a.m.	U.S. Fish and Wildlife Service inventory and monitoring (Todd Sutherland, data manager, National Wildlife Refuge System Inventory and Monitoring Program)	
10:30 a.m.	Bureau of Land Management (lightning talk), monitoring for adaptive management (Gordon Toevs, lead, Assessment, Inventory and Monitoring Program)	
10:40 a.m.	Federal agencies plenary clicker questionnaire and discussion	Interactive plenary discussion.
10:55 a.m.	20-minute break	Break.
11:15 a.m.	Plenary—EcoINFORMA initiatives–Current assets and future directions (moderator: Elizabeth Martín, U.S. Geological Survey)	Plenary session 2—EcoINFORMA initiatives—Current assets and future directions.
11:15 a.m.	EcoINFORMA overview (Elizabeth Martin, technical information specialist, U.S. Geological Survey)	
11:30 a.m.	Biodiversity resource hub—Biodiversity Information Serving Our Nation (Stinger Guala, EcoScience Synthesis Branch lead, U.S. Geological Survey)	
11:45 a.m.	Ecosystem services resource hub—EnviroAtlas (Annie Neale, lead, EnviroAtlas Project, Environmental Protection Agency)	
12:00 p.m.	Land cover hub—Multi-Resolution Land Characteristics Consortium (Jonathan Smith, program coordinator, Land Change Science Program, U.S. Geological Survey)	
12:15 p.m.	Proposed EcoINFORMA hub—Ag Data Commons (Cynthia Parr, technical information specialist, National Agricultural Library)	
12:30 p.m.	EcoINFORMA initiatives plenary clicker questionnaire and discussion	Interactive plenary discussion.

Table 2.   Agenda for workshop, "Promoting Synergy in the Innovative Use of Environmental Data," December 2–4, 2015, Department of Interior South Building, 1951 Constitution Avenue NW, Washington D.C.—Continued

[LANDFIRE, Landscape Fire and Resource Management Planning Tools Project, EcoINFORMA, ecoinformatics-based open resources and machine accessibility; GIS, geographic information system]

Time	Activity or presentation topic	Session
Wednesday, December 2, 2015—Continued		
12:45 p.m.	75-minute lunch on own	Lunch.
2:00 p.m.	Map and app mashup I—Student speed presentations (moderator: Annie Simpson, U.S. Geological Survey) Tiffany Yap, University of California, Los Angeles Matt Kling, University of California, Berkeley Hsun-Yi Hsieh, University of Michigan Nathan Wikle, Truman University, Ashish Gauli, Fisk University, and Ryan Yan, College of William & Mary Serenity Purcell & Enrique Montano, Smithsonian Institution Hannah Dempsey, College of William & Mary	Plenary session 2—EcoINFORMA initiatives—Current assets and future directions—Continued.
2:30 p.m.	Plenary—Citizen science—Contributions to biodiversity informatics needs (Moderator: Annie Simpson, U.S. Geological Survey)	Plenary session 3—Citizen science contributions to biodiversity informatics needs.
2:30 p.m.	eBird (Rob Stevenson, professor, University of Massachusetts, Boston)	
2:45 p.m.	The National Phenology Network (Jake Weltzin, executive director, National Phenology Network)	
3:00 p.m.	iNaturalist (Scott Loarie, co-director, iNaturalist)	
3:15 p.m.	Citizen science plenary clicker questionnaire and discussion	Interactive plenary discussion.
3:30 p.m.	30-minute break	Break.
4:00 p.m.	Plenary—Nongovernmental organizations—Biospatial data efforts in the community (Moderator: Ben Wheeler, U.S. Geological Survey)	Plenary session 4—Biospatial data efforts in the nongovernmental organization community.
4:00 p.m.	The Wilderness Society—Mapping wildland values and climate vulnerability to support conservation strategy across the United States (Greg Aplet, senior science director, and Travis Belote, ecologist)	
4:15 p.m.	Defenders of Wildlife—Integrating regulatory, biological, and geospatial data to improve Endangered Species Act implementation (Jacob Malcolm, endangered species science and policy analyst, and Ryan Covington, spatial analyst)	
4:30 p.m.	NatureServe—Biodiversity and ecosystem spatial data of NatureServe and the Natural Heritage Network (Healy Hamilton, chief scientist)	
4:45 p.m.	The Audubon Society—Audubon enterprise GIS—Bridging gaps and increasing conservation impact (Doreen Whitley, geospatial information officer)	
5:00 p.m.	Biospatial data efforts in the nongovernmental organization community—Plenary clicker questionnaire and discussion	Interactive plenary discussion.
5:15 p.m.	Formation of breakout groups	
5:30 p.m.	Adjourn for the day	Adjourn.

Table 2.   Agenda for workshop, "Promoting Synergy in the Innovative Use of Environmental Data," December 2–4, 2015, Department of Interior South Building, 1951 Constitution Avenue NW, Washington D.C.—Continued

[LANDFIRE, Landscape Fire and Resource Management Planning Tools Project, EcoINFORMA, ecoinformatics-based open resources and machine accessibility; GIS, geographic information system]

Time	Activity or presentation topic	Session
Thursday, December 3, 2015		
8:30 a.m.	Day 1 summary, logistics, day 2 objectives (Stinger Guala and Healy Hamilton)	Plenary.
8:45 a.m.	Plenary—Academia—Biospatial and geospatial data efforts (moderator: Tosha Comendant, Conservation Biology Institute)	Plenary session 5—Bio- and geospatial data efforts in the academic sector.
8:45 a.m.	University of Florida—iDigBio—Enabling access to and reuse of millions of specimen and media records (Alex Thompson, information technology infrastructure lead, iDigBio)	
9:00 a.m.	University of Florida/Yale University—Map of life (Rob Guralnick, curator, Biodiversity Informatics)	
9:15 a.m.	Utah State University—Western Center for Monitoring & Assessment of Freshwater Ecosystems (Chuck Hawkins, director, Western Center for Monitoring & Assessment of Freshwater Ecosystems)	
9:30 a.m.	University of California, Berkeley—Initiative in global change biology (Michelle Koo, staff curator, biodiversity informatics and GIS)	
9:45 a.m.	College of William and Mary—Data integration and decision-making—Using GIS for climate screening with the U.S. Army Corps of Engineers (Dan Runfola, senior geospatial scientist)	
10:00 a.m.	Biospatial and geospatial data efforts in the Academic sector—Plenary clicker questionnaire and discussion	Interactive plenary discussion.
10:15 a.m.	30-minute break	Break.
10:45 a.m.	Map and app mashup II—Speed presentations (moderator: Tosha Comendant, Conservation Biology Institute) Andy Holdsworth, Minnesota Department of Natural Resources Tim Howard, New York Natural Heritage Program Jimmy Kagan, Oregon Biodiversity Information Center Chris Burkett, Virginia Department of Game and Inland Fisheries, and Jason Bulluck, Virginia Natural Heritage Program Tom Akre, Smithsonian Conservation Biology Institute William McShea, Smithsonian Conservation Biology Institute	Plenary session 5—Bio- and geospatial data efforts in the academic sector—Continued.
11:15 a.m.	Plenary—Private sector—Data integration initiatives (moderator: Stinger Guala, U.S. Geological Survey)	Plenary session 6—Private sector data integration initiatives.
11:15 a.m.	Microsoft, Inc. (Kristen Tolle, director, Microsoft Data Science Initiative)	
11:30 a.m.	The Earth Knowledge Platform (Frank D'Agnese, chief technology officer, Earth Knowledge Network)	
11:45 a.m.	Esri (Lori Pelech, manager of conservation nonprofits)	
12:00 p.m.	Private sector data integration initiatives plenary clicker questionnaire and discussion	Interactive plenary discussion.

Table 2. Agenda for workshop, "Promoting Synergy in the Innovative Use of Environmental Data," December 2–4, 2015, Department of Interior South Building, 1951 Constitution Avenue NW, Washington D.C.—Continued

[LANDFIRE, Landscape Fire and Resource Management Planning Tools Project, EcoINFORMA, ecoinformatics-based open resources and machine accessibility; GIS, geographic information system]

Time	Activity or presentation topic	Session
Thursday, December 3, 2015—Continued		
12:15 p.m.	75-minute lunch on own	Lunch
	Tools café—Rachel Carson Room	Tools café
	Informal meeting of breakout groups in Main Interior Cafeteria	Breakout pre-meet.
1:30 p.m.	Plenary—Data visualization and dissemination (moderator: Healy Hamilton, NatureServe)	Plenary session 7—Data visualization and dissemination.
1:30 p.m.	Vizzuality (Craig Mills, chief executive officer, Vizzuality)	
1:45 p.m.	Blue Raster (Mike Lippman, cofounder and principal, Blue Raster)	
2:00 p.m.	DataBasin (Tosha Comendant, senior scientist, Conservation Biology Institute)	
2:15 p.m.	National Oceanic and Atmospheric Administration Climate Resilience Toolkit (Ned Gardiner, communications and education producer)	
2:30 p.m.	Data visualization and communication plenary clicker questionnaire and discussion	Interactive plenary discussion.
2:45 p.m.	Map and app mashup III—Speed presentations (moderator: Healy Hamilton, NatureServe)	Plenary session 7—Data visualization and dissemination—Continued.
	Woody Turner, National Aeronautics and Space Administration	
	Mike Gill, Group on Earth Observations	
	Ryan Powers, Yale University	
	Michael Fitzgibbon, Point Blue	
	Regan Smyth, NatureServe	
	Mary Ellen Hannibal, Journalist	
3:15 p.m.	15-minute break	Break.
3:30 p.m.	Plenary instructions for breakout group sessions	Plenary discussion.
3:45 p.m.	Breakout sessions	Breakout sessions.
5:30 p.m.	Adjourn for the day	Adjourn.
Friday, December 4, 2015		
9:00 a.m.	Plenary keynote—Environmental data synergy—Where do we go from here? (Dan Ashe, director, U.S. Fish and Wildlife Service)	Plenary.
9:30 a.m.	Breakout session reports/discussion I	Interactive plenary reports.
10:30 a.m.	15-minute break	Break.
10:45 a.m.	Breakout session reports/discussion II	Interactive plenary reports.
12:00 p.m.	Workshop overall summary (Healy Hamilton and Stinger Guala)	Plenary summary.
1:00 p.m.	Adjourn	Adjourn.

Summary of Plenary Sessions

Plenary presentations on December 3 and 4, 2015, were grouped in categories: ecoinformatics-based open resources and machine accessibility (EcoINFORMA); Federal inventory and monitoring; academic, nongovernmental organization (NGO), and private sector initiatives; citizen science; map and app mashups; and data visualization and dissemination.

Plenary Session 1—Monitoring in Federal Resource-Management Agencies—Current Efforts and Immediate Needs

Six leaders in Federal government inventory and monitoring programs described the goals of their efforts, the data types being collected, and the decisions their programs are designed to inform (see table 2). Most major resource-management agencies were represented, including the U.S. Forest Service, the National Park Service, the U.S. Fish and Wildlife Service, the Bureau of Land Management, the Natural Resource Conservation Service, and the interagency LANDFIRE Program. In total, 80 percent of participants indicated that they use Federal agency data in their work (see appendix 4), demonstrating the reliance on Federal inventory and monitoring data (fig. 2). The importance of nationally consistent spatial datasets on biodiversity and environment was highlighted, and the need for time series data that can support trend analyses was emphasized.

Plenary Session 2—EcoINFORMA Initiatives—Current Assets and Future Directions

After an overview of EcoINFORMA and its roots in President's Council of Scientific Advisors on Science and Technology report (2011), four leaders of this initiative provided descriptions of their informatics efforts (table 2). Each speaker described their specific data contents and accessibility and drew the connections between their programs and the overall goals of EcoINFORMA. The presentations demonstrated the various stages of development of EcoINFORMA efforts, with the Biodiversity Information Serving Our Nation (BISON) mapping project of the U.S. Geological Survey and the Multi-Resolution Land Characteristics Consortium representing mature, data rich, and rapidly growing environments currently serving large volumes of decision-relevant data, EnviroAtlas offering an innovative approach to quantifying and mapping ecosystem services, and the newly available and rapidly expanding Ag Data Commons showing great promise as an aggregator of many important and largely inaccessible datasets. Fifty percent of the workshop participants had used EcoINFORMA data, and 47 percent had been providers of data to EcoINFORMA programs (see appendix 4). The session discussion noted the value of EcoINFORMA's programs in supporting community standards for various types of data aggregation, analysis, and dissemination.

Plenary Session 3—Citizen Science Contributions to Biodiversity Informatics Needs

Three pioneering and nationally significant citizen science programs (eBird, the National Phenology Network, and iNaturalist; table 2) were presented, one each from the academic, Federal, and NGO sectors. All three speakers clearly demonstrated the tremendous potential of citizen science data to contribute to biodiversity discovery and monitoring of population and phenological trends, and simultaneously to expand public engagement. In total, 48 percent of the workshop participants had contributed citizen science data efforts, and 77 percent had used citizen science data in their work (see appendix 4). The success of protocols to support contribution of eBird and iNaturalist observations to data hubs such as USGS BISON and the Global Biodiversity Information Facility conveyed the sheer volume of citizen science data and potential, emphasizing the era of big data in biodiversity observations has arrived. Substantial further discussion in multiple workshop sessions centered around the most important environmental datasets and data infrastructure needs for fully leveraging this new and growing source of environmental information. Four major issues for accelerating applicability of citizen science data were identified: improving data sharing standards, establishing appropriate quality-assurance and quality-control processes for crowd-sourced observations, improving data visualization and delivery mechanisms, and improving methods for identifying data provenance.

Plenary Session 4—Biospatial Data Efforts in the Nongovernmental Organization Community

Four science-based NGOs (the Wilderness Society, Defenders of Wildlife, NatureServe, and the Audubon Society; table 2) offered diverse, informative conservation applications of nationally consistent spatial data layers and demonstrated the value of curated digital data on biodiversity, environment, and climate. Two themes were leveraging the use of spatial data and tools for analysis and visualization and using the current and potential contribution of geographic information science to support tools, analyses, and visualizations that engage and empower. This session emphasized the benefits of investing in data access—especially spatial data, documentation such as metadata, and the bioinformatics infrastructure to facilitate the full data cycle, much of which is being addressed by EcoINFORMA initiatives. Much greater potential exists to use these data and tools to improve a stakeholder-engaged process of science-based conservation, such as creating landscape-scale conservation blueprints, identifying metrics for conservation effectiveness, identifying information gaps (fig. 3), and leveraging technology to grow public engagement.
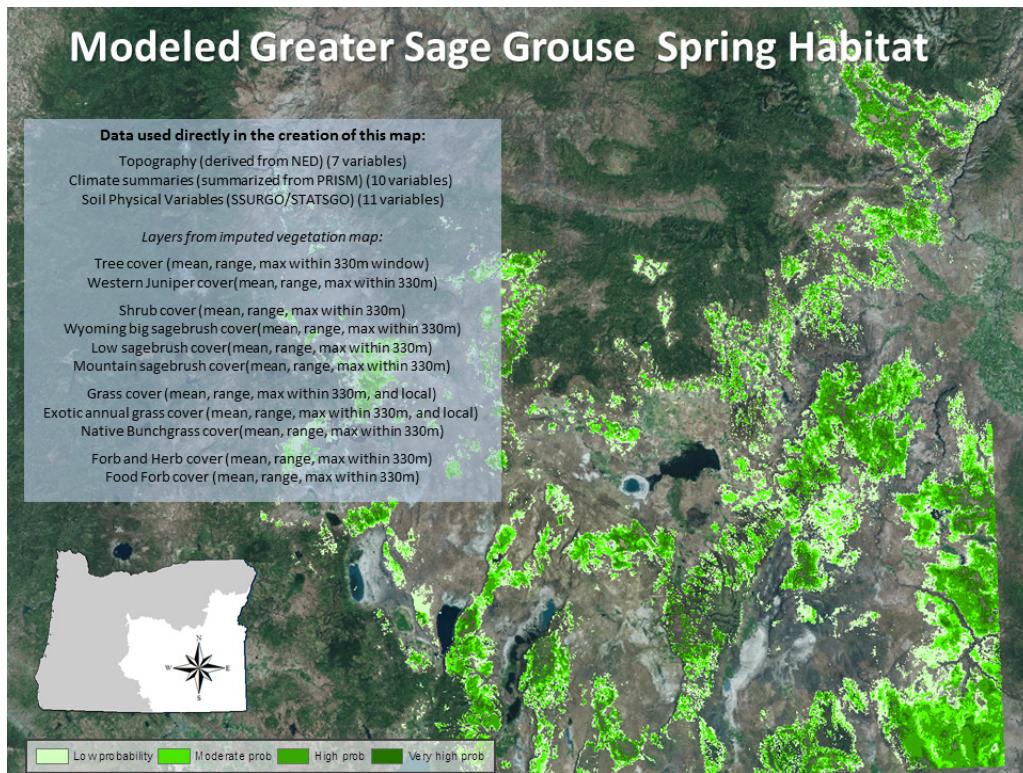
Figure 2. Map of modeled greater sage grouse habitat use in the spring. This data visualization demonstrates the importance multiple disparate environmental data layers that rely in part on Federal agency products. Image courtesy of Jimmy Kagan, Oregon Biodiversity Information Center, Portland State University; used with permission.
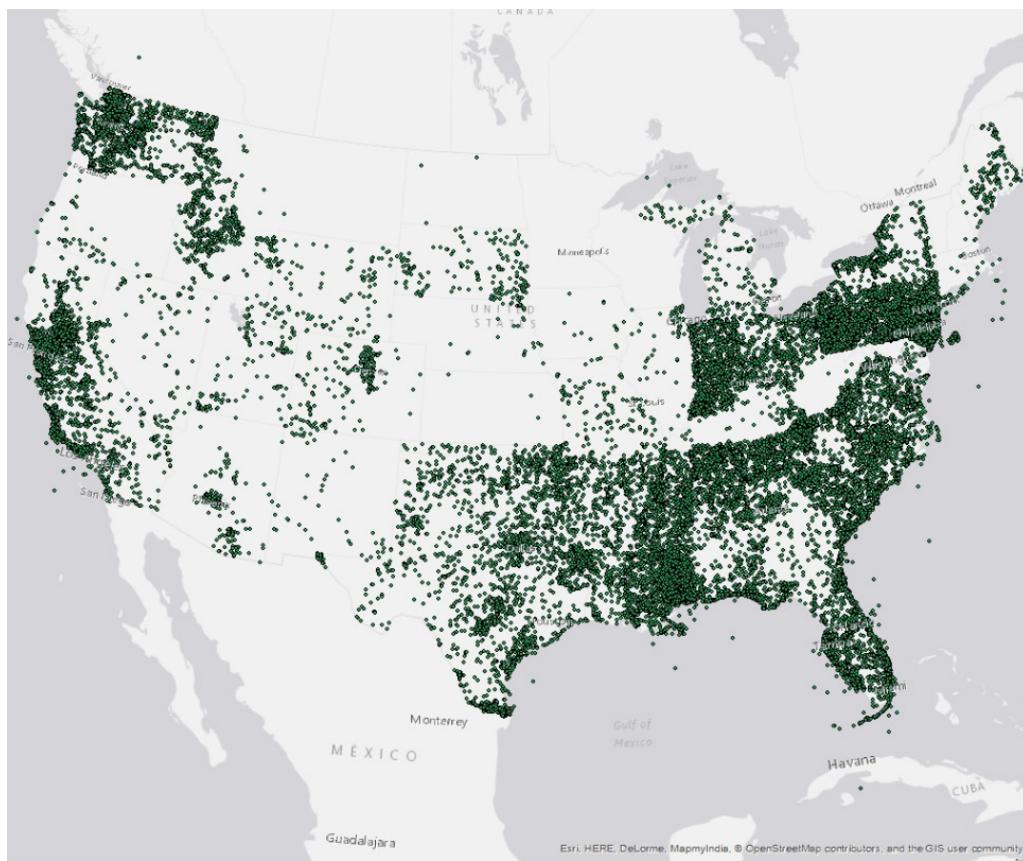
Figure 3. A map of the locations of all Endangered Species Act consultations (2008–15) for which the U.S. Fish and Wildlife Service recorded coordinates. The visible gaps in the map—most obvious at State level—are created by variation in which offices record coordinates, not the number of consultations. Map by Defenders of Wildlife, https://doi.org/10.6084/m9.figshare.5868159; used with permission.

These presentations by the NGO community suggested several derived national data layers that are in high demand, such as ecological representation of current conservation estate, percent loss of natural vegetation, habitat connectivity, climate exposure, species richness, and various ecological metrics of resistance and resilience. Building awareness of data access and applicability is essential because there are few organizations and insufficient resources to properly manage the full biodiversity data life cycle (from collection, curation, and analysis to visualization and dissemination) for major biodiversity and environmental datasets.

Plenary Session 5—Bio- and Geospatial Data Efforts in the Academic Sector

Five presentations from the academic community demonstrated the creative strength of that sector in modeling biotic response to global change (iDigBio by the University of Florida, Map of life by the University of Florida/Yale University, Western Center for Monitoring & Assessment of Freshwater Ecosystems by Utah State University, Initiative in global change biology by the University of California, Berkeley, and GIS for Climate Screening by the College of William and Mary; table 2). Each participant offered a powerful application based on digitizing, aggregating, and visualizing biodiversity and ecoinformatics datasets. In spite of the high quality of the academic applications shared during the academic sector plenary and that 72 percent of 82 workshop questionnaire respondents said they use data and applications from academia in their work, 25 percent of them found out about data or tools from academic sources completely opportunistically and 23 percent by surfing the web (appendix 4). A common theme among the bio- and geospatial applications from academia was the essential nature of geolocated data, with latitude and longitude as a central organizing data attribute, or analyzing the dynamics of global change. To extract the most value out of event-based data, such as species observations, its counterpart—spatial environmental data or base layers, such as downscaled climate models—needs to be invested in. Major university-led efforts are underway to digitally capture historical biodiversity datasets and apply them to improve the understanding of baseline ecological conditions and to establish historical ranges of variation. From user-driven dynamic modeling tools for refining biodiversity knowledge to predictive species lists for a watershed, the fundamental building blocks enabling these and similar efforts are nationally consistent biospatial and geospatial data layers. More focus is warranted on the rapidly growing field of geovisualization tools, especially given their importance in connecting audiences, including decision makers, to biodiversity analysis products.

Plenary Session 6—Private Sector Data Integration Initiatives

Presentations from the private sector (Microsoft, Inc., the Earth Knowledge Platform, and Esri; table 2) strongly emphasized data interoperability as a fundamental challenge and opportunity for increasing the relevance of and improving outcomes from applications of biodiversity and environmental data. Investment in connectivity, open data access, and data interoperability can transform resource management decision making from reactive to proactive, which can save money and improve outcomes. The rapidly emerging sensor revolution, from dense networks of environmental data loggers to unmanned aerial vehicles and new generations of satellites, highlights the power of increasing the scale of data interoperability; however, interoperability alone cannot transform information into decision-relevant insight. The demand for real-time dashboards that synthesize, analyze, and visualize disparate forms of data for decision support is rapidly increasing. In total, 84 percent of 89 questionnaire respondents felt that public-private partnerships are either essential or very important to creating the big data processing infrastructure required to support information needs of managers and practitioners that must sustain resources in the face of global change (appendix 4). Almost 50 percent of the 90 respondents said they were currently part of such public-private partnerships. Cloud-enabled data aggregation, management, and analysis will provide improved characterization of uncertainties, will accelerate interoperability, and will advance the application of machine learning algorithms that can fill data gaps. Part of this big data challenge is integrating data across space and time, such as establishing and implementing spatially consistent domains for modeling. In total, 50 percent of the 94 questionnaire respondents (appendix 4) said that lack of opportunities to develop collaborations were the limiting factor in broadening public-private collaboration, and 26 percent felt that lack of common goals (knowledge versus profit) was the limiting factor.

Plenary Session 7—Data Visualization and Dissemination

The final plenary set of presentations (on Vizzuality, Blue Raster, DataBasin, and the National Oceanic and Atmospheric Administration Climate Resilience Toolkit; table 2) offered perspectives from private sector, government, and NGO efforts focused on visualizing and disseminating biodiversity and environmental data. This session was widely regarded as being one of the most successful, and 84 percent of 98 participants surveyed said they use web-based data visualization tools. It exposed a community largely focused on data collection, aggregation, and analysis to modern, creative approaches for visualizing and communicating data to target audiences. Through advances in technology, the art and design communities have developed high capacity for dynamic, visually
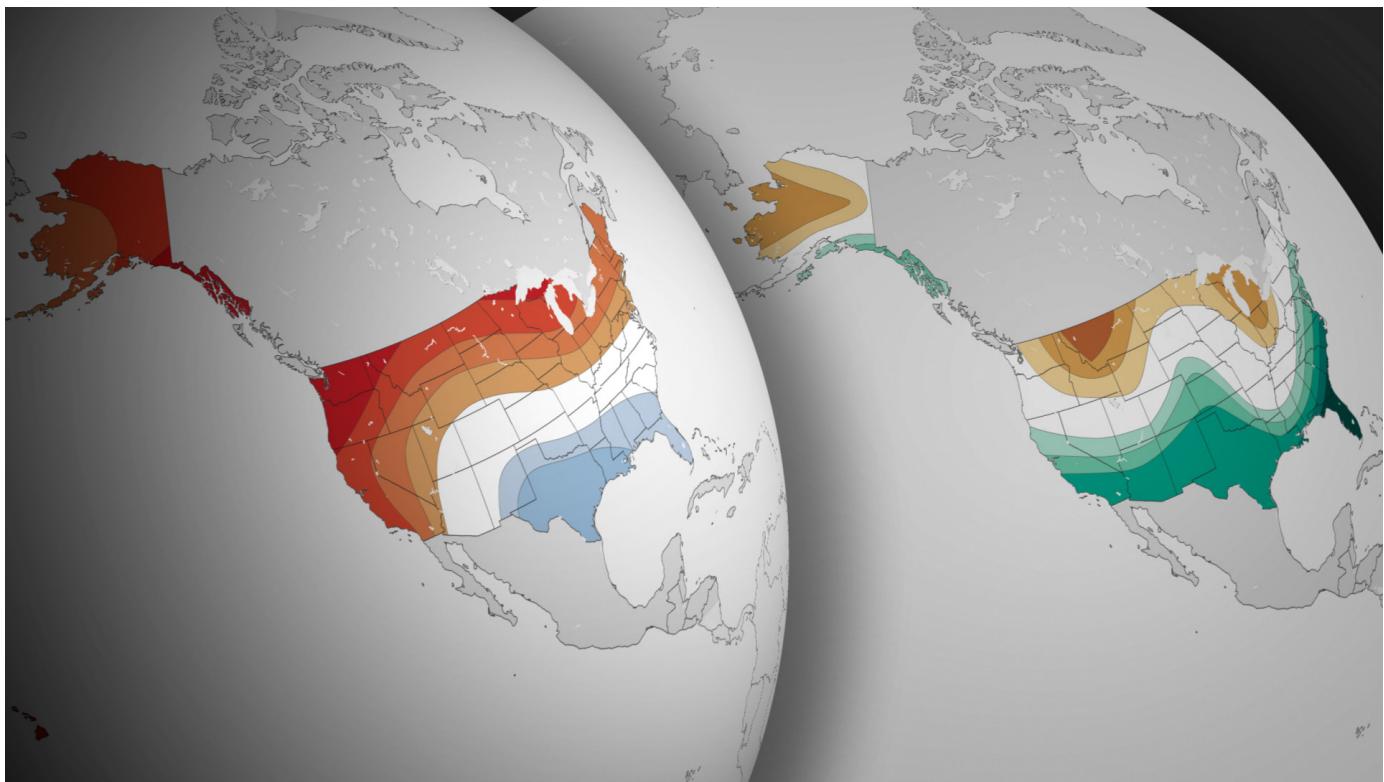
Figure 4. Geovisualization (or geographic visualization) refers to a set of tools and techniques supporting the analysis of geospatial data using interactive functionality. Image courtesy of Edward Gardiner and Climate.gov; used with permission.

intuitive representations of a wide range of environmental data (fig. 4). Geospatial mapping, animation, and web technologies offer unlimited opportunities to store, create, and tell data-based stories in new visual ways, reaching more diverse audiences. These same technologies mobilize powerful peer-to-peer learning networks around tools, data, and expertise, creating communities of practice that can be directed toward issues in ecological informatics. There is demonstrated value of data-sharing platforms that offer aggregations of disparate environmental data and syntheses of lessons learned. Speakers advocated that more opportunities should be created for the art, design, and biodiversity science communities to generate public-private partnerships for ecological data analysis and visualization.

"Take Homes" from the Breakout Sessions

Based on the issues raised in the presentations and participant-declared interests, breakout sessions with 10 groups were held in the afternoon and into the evening of day 2 and presented results consisting of answers to questions posed to them (see appendix 1) and suggestions for future directions.

Ecosystem Services and Ecological Indicators

Although a Whitehouse Memorandum (Executive Office of the President of the United States, 2015) directed Federal agencies to use ecosystem services in planning and decision making, some driver is needed for agencies to finish and maintain foundational datasets that will allow them to accurately assess and define the extent of ecosystem services and their effects on the environment. The breakout group reported that there needs to be a provision for directing agencies to complete essential and foundational datasets like the Soil Survey Geographic Database, the high-resolution National Hydrography Dataset, the 5-meter digital elevation model (DEM), and the National Wetland Inventory, and make them EcoINFORMA hubs. Group members also asserted we do not have social data sets that are comparable to the biophysical data that we have. While providing a call for more social science to be directed at the problem, the lack of comparable social/biophysical data sets also highlights the difficulty of working with disparate datasets collected for other purposes—a theme that arose repeatedly among the breakouts. Scale dependencies in the data are often very limiting. The ecoinformatics community also has not sufficiently mined the data that we do have and has not adequately cast it in terms of benefits and beneficiaries (for example, the National Aquatic

Resource Survey and the Forest Inventory and Analysis were mentioned). Highly integrative and ongoing analysis of ecosystem services is needed, but the many pieces are spread across many disconnected areas in many agencies and across the private sector. A much higher level of coordination will be needed to achieve the stated goals on this issue.

Inventory and Monitoring

The group encapsulated their goal as follows: shift from making conservation decisions and designs based on best available information to making decisions based on the most complete and appropriate information that exists. To further clarify this, they pointed out that "best available" does not necessarily mean "good" or "appropriate," and this discrepancy is what needs to change. Although informatics can help enormously (in the form of data integration and modeling) to focus efforts and improve efficiency, there is no substitute for comprehensive, long-term data on the species and ecosystems in question. Further, there is a compelling and obvious need for the consistent and comprehensive use and, in some cases development of, national standards integrated with global standards for data gathering and dissemination. Optimizing resources through national strategies for monitoring, rather than ad hoc local- or single organization-oriented ones, also is needed. Optimization is challenging, especially regarding compartmentalizing data, funding, and priorities. Leadership at the highest levels is needed to ameliorate this, and some successes because of cooperation would greatly help

the process. Although marine systems were not specifically addressed in this workshop, the group also pointed out the need for integration with, and within, that arena.

Biogeographic Map of the Nation

In contrast to other breakout sessions that highlighted the need for a central resource to integrate and visualize data on species and their distributions, to identify gaps in our knowledge, to assess threats, and to efficiently use and manage resources, this breakout session was charged with visioning such a central resource (fig. 5). Although BISON has moved aggressively toward meeting this need during the past 3 years and served as the EcoINFORMA hub for biodiversity, leading efforts in the Federal realm, it is only part of the story. The group pointed to the need for a full information infrastructure to record, manage, and synthesize all data relevant to documenting and understanding the spatial distributions of organisms, the factors that determine distributions, and how distributions have changed and will change over time. They emphasized that the resource must be dynamic and function across scales. It must support simple interfaces for decision making and more complex interfaces for research and re-use such as Linked Open Data. Finally, they pointed out that although further community building is key, most of the critical components of this much-needed central, integrating resource were highlighted in this workshop. More workshops like this, coupled with sharing of students across projects, were discussed as high-value ways to move forward.
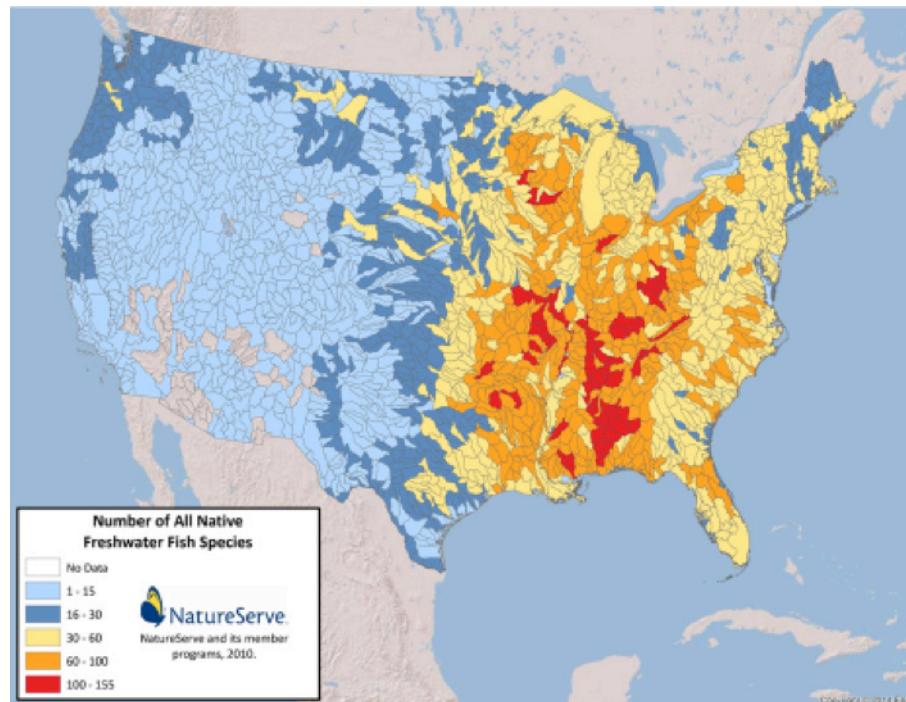

Figure 5. Many of the critical components of a national biogeographic map of the Nation were showcased during the workshop. Image courtesy of NatureServe; used with permission.

Pollinators

The importance of pollinators (fig. 6) for food security and many other ecosystem services cannot be understated. Substantial work within the past 5 years on the strategic needs for pollinator assessment and management, such as the Pollinator Research Action Plan (Pollinator Health Task Force, 2015), has moved the conversation forward greatly. However, many of the most basic information needs remain to be met; for example, a national list of pollinators, or of the plants that require animal mediated pollination services, does not exist. Distribution data for all but the most common species are very poor, and even co-occurrence data (with pollinated plant species) are extremely rare compared to the actual number of pollinator associations. Although progress has been made, a comprehensive standard to document pollinator plant interactions and the factors that affect them has not be created. A major effort is needed to acquire and interpret data represented in specimen collections that are not digitized, data in the literature, and new syntheses to determine baselines and the temporality of pollinator interactions and how they may change.

Invasive Species

The early detection of and rapid response (EDRR) to non-native invasive species is critical to managing and maintaining our natural ecosystems, ecosystem services, and habitat resilience. The importance of EDRR is emphasized in a newly released interdepartmental report that was described during the workshop (U.S. Department of the Interior, 2016; fig. 7). U.S. Department of the Interior (2016) proposed a national framework consisting of a coordinated set of actions to find and eradicate potential invasive species before they spread and cause harm. A nationally comprehensive database of non-native species does not exist and is clearly a central need. More comprehensive expert-vetted lists are necessary for determining what is (and is not) native in any given area, and the taxonomy needs to keep up. Existing data networks collaborate, but much more reporting is needed for effective EDRR. Citizen science should play a major role, especially for species that are easy to recognize and tag as invasive, and we should concentrate on those species at a national scale. For more cryptic species, some citizen science networks are developing substantial identification expertise and should be leveraged. The potential to massively scale up observations of invasive species to support EDRR decisions highlights the need to ensure connection with major repositories and to support a data-to-decisions pipeline, moving beyond the current state of taxonomically or geographically restricted monitoring and databases.

Figure 6.   Native pollinators such as the swallowtail are essential to the balance of natural ecosystems. Photo courtesy of Gretchen Westhoff, https://creativecommons.org/licenses/by/2.0/; used with permission.


Figure 7.   The national framework for invasive species early detection and rapid response (U.S. Department of the Interior, 2016) describes parameters for addressing the invasive species issue. Image courtesy of the U.S. Department of the Interior; used with permission.

Remote Sensing

The Federal government invests heavily in the remote sensing of many critical variables for assessing and managing biodiversity and ecosystem change. Landsat-derived products have been indispensable, but training data for classification are still limited. These products must be combined with directly measured data on the ground and higher-resolution data to be most useful in this context. Standards for image processing are well defined, but large gaps still exist in the availability and use of standards and comparative data for interpretation (for example, what is a water body or wetland). Land use interpretation is very challenging because local and municipal data are largely lacking and poorly defined. Also, multiple levels of privacy and access concerns must be addressed in the data, especially for higher-resolution data. An integrated system of Landsat, Sentinel, light detection and ranging (lidar), hyperspectral, and aerial imagery would be ideal (fig. 8). Remote sensing of individual plant species extents and ecosystem health and richness are within reach. This group also suggested that EcoINFORMA should have a role in helping gauge usage and solicit feedback.

Drivers of Agricultural Change

Agriculture, from intensively managed microfarms to passively managed forests, accounts for most of our Nation's food and fiber production and the largest commercial interface with ecosystems. Substantial investments have gone into accounting and assessment across the landscape with respect to crops and factors that affect their production (fig. 9). However, a much more integrated picture is needed to assess and manage biodiversity and ecosystems in concert with agriculture and to make ecosystems sustainable and resilient. Forecasting is key and numerous temporally fixed and predictive models exist, but they are rarely well validated. The combination of agricultural with nonagricultural lands and environmental factors beyond the immediate interface is rare. Largely due to economic and privacy issues, but also in many cases for less compelling reasons, much of the data are proprietary and difficult to obtain in a timely fashion. Further, highly significant spatial data on the microbiome, pollinators, pathogens, and other important variables either do not exist or are spotty at best. There is enormous potential for more collaboration across the agricultural interface.
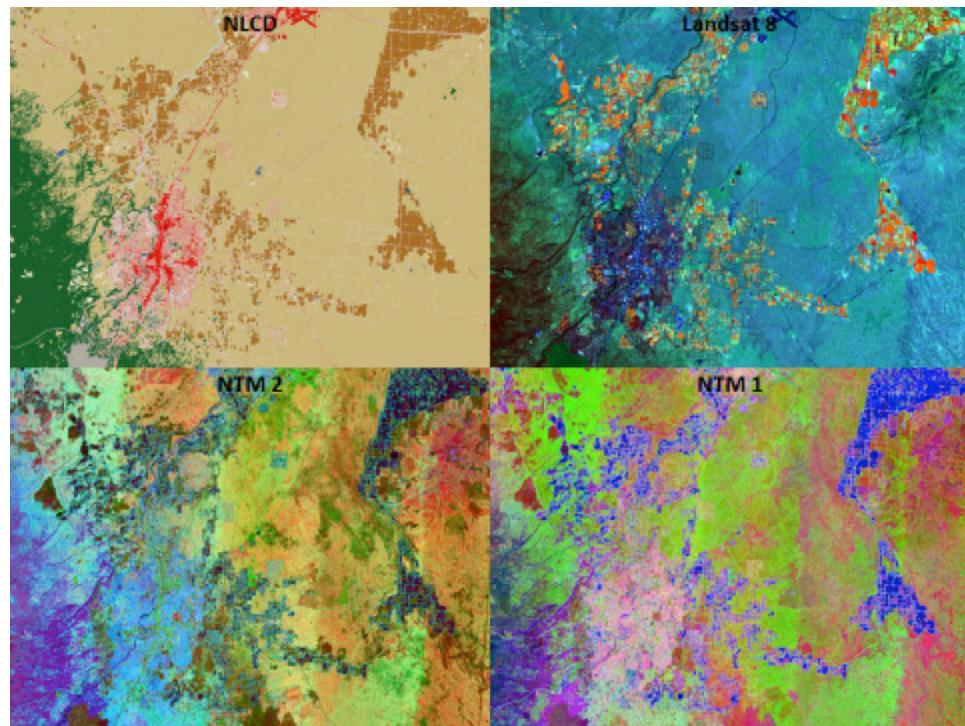

Figure 8.   Modern imagery provided by satellites, such as Landsat 8 and Sentinel 2, and by publicly available air photography, such as the National Agriculture Imagery Program, provide analysis tools for land use and land cover interpretation. These are four images of the same location, with the National Land Cover Database shown on the top left, Landsat 8 on the top right, and two sets of different texture metrics derived from air photography shown in the bottom two images. Image courtesy of Jimmy Kagan, Oregon Biodiversity Information Center, Portland State University; used with permission.
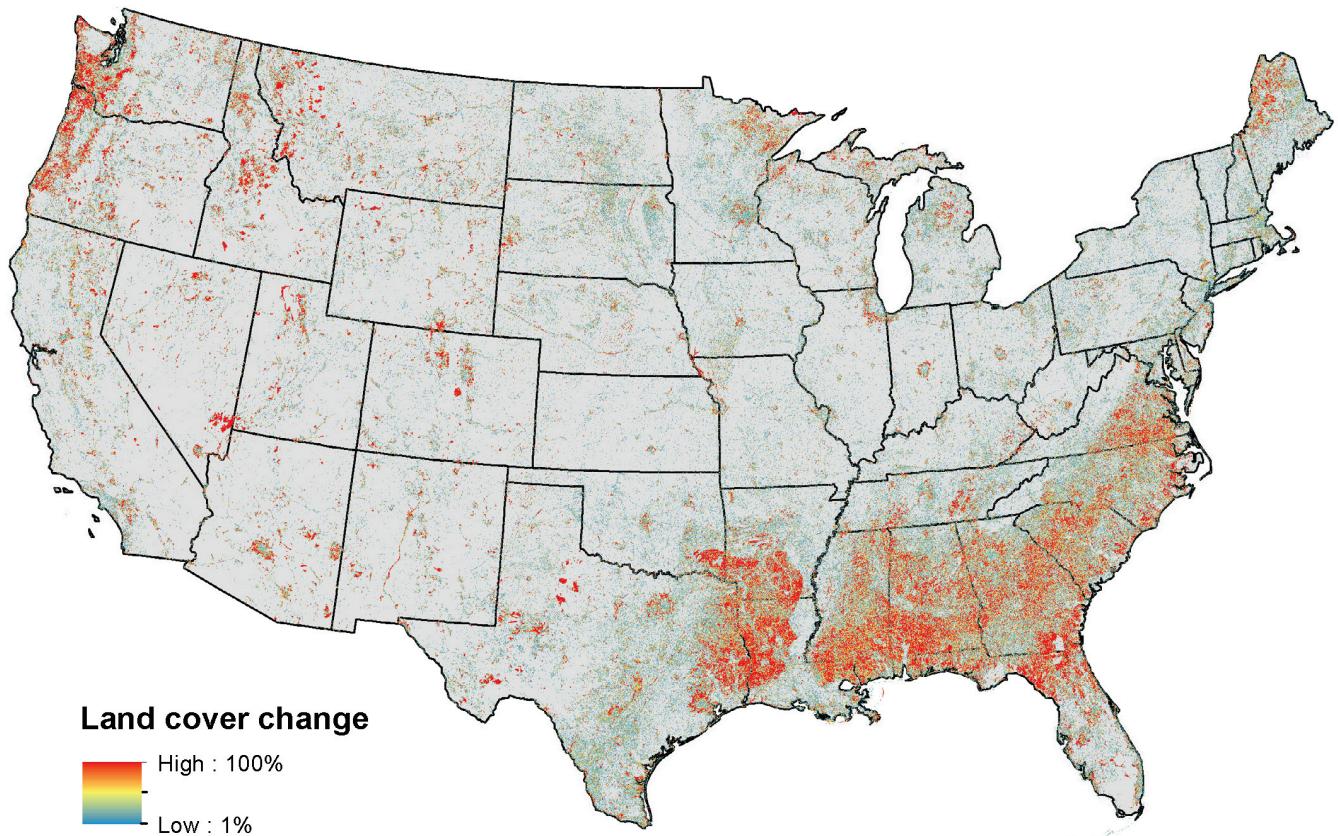
Figure 9.   Maps such as this one, which displays the magnitude of spatial change over time, are part of the National Land Cover Database, a product of the Multi-Resolution Land Characteristics Consortium. Image from Homer and others (2015); used with permission.

Citizen Science

Citizen science has become increasingly important in understanding large-scale phenomena that require large amounts of distributed data and that are easily observed and communicated with little scientific training (fig. 10). The data needed to answer many of the major questions in biodiversity management fit squarely within this realm, and numerous breakout groups highlighted the importance of citizen science in the future of their disciplines. The power of citizen science comes from the large volume and spatial scale of the data and the broad appreciation from the public that the activity promotes. Examples of novel and time-tested methods to ensure data quality and reasonable sampling are readily available. However, what is currently needed is more national-scale leadership to enhance and manage collaboration and to promote more complementary and comprehensive coverage and support of citizen science. Small and spatially or taxonomically restricted citizen science projects often generate high enthusiasm but fail to collaborate effectively up the progressively broader hierarchy of appropriate national efforts. This often reduces the level of data quality as well as the relevance, and hence sustainability, of the project. Ultimately the data need to feed into national-level resources with synthetic analyses clearly visualized, with appropriate credit, expertise, and resources flowing back to the source.

Climate

The climate breakout session participants discussed large-scale climatic data and its uses related to projecting climate futures and biodiversity vulnerability. The rate, magnitude and nature of climate change must be considered in applied research and decision making across the realm of ecosystems and biodiversity. These forecasted changes are very complex, and uncertainties need to be accounted for and communicated. Validation through hindcasting should be actively supported, and comparison among climate models, emissions pathways, and quantification of differences in downscaling methods should be encouraged. Improved methods of determining climate change exposure for species and ecosystems are needed. Establishing an overall baseline as well as appropriately downscaled climate data layers for specific regions and analyses is critical. The fundamental inclusion of specific ecologies into analyses of the past or future effects of climate change also is necessary for credible outcomes. The ecological forecasting community needs to be supported to identify, develop, and deploy the most important spatial data layers for predicting biodiversity and ecosystem response to climate change.
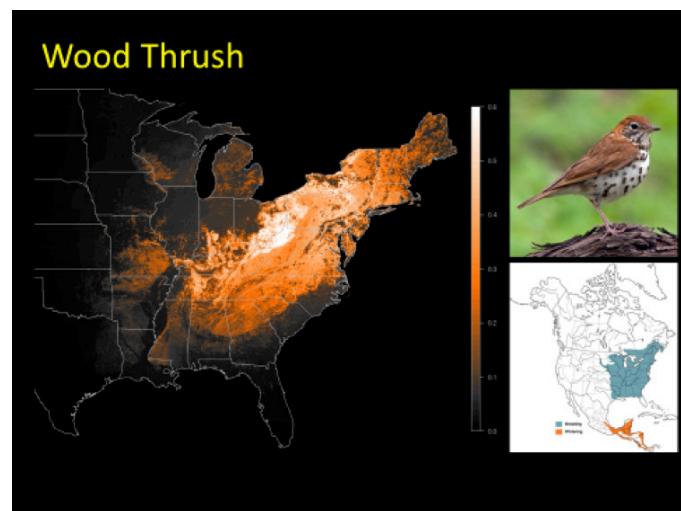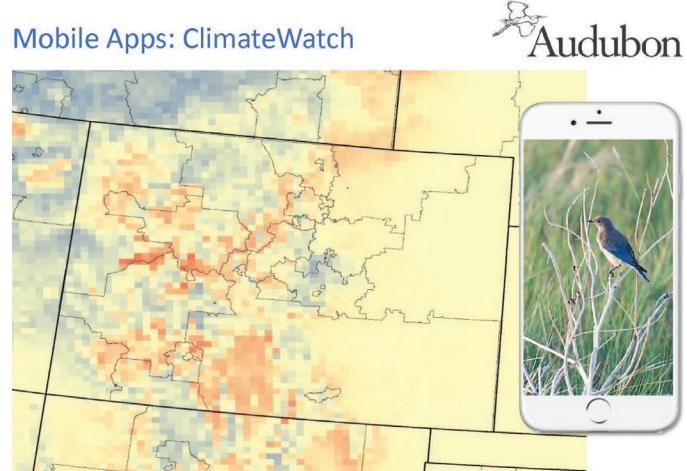

Figure 10.   Top image, Spatio-Temporal Exploratory Mapping model for Wood Thrush (*Hylocichla mustelina*) that breed throughout eastern forests and are present east of the Great Plains from central Florida and east Texas north to southern Canada. The lighter color indicates their spring arrival in April and May. Image courtesy of Steve Kelling, Cornell Lab of Ornithology; used with permission. Bottom image, Climate Watch App, an open online platform designed to empower users with climate data visualizations. Image courtesy of the National Audubon Society; used with permission.

Hydrology and Watersheds

Hydrologic data collection needs cooperators, but we must realize that freshwater and marine systems are very different. The workshop discussion emphasized freshwater systems, and the breakout session participants determined the following needs and wants within this area:

1. Nationwide (or at least in the lower 48 States) completion of the National Hydrography Dataset high-resolution stream coverage. If this coverage were able to include the attributes of the National Hydrography Dataset plus data, it would be exceptionally useful for aquatic species modeling. Fifth order and larger streams should have both stream boundary lines estimated.

2. Modeled riparian area and flood plain boundaries need to be completed, using the 5- or 1-meter DEM. Work underway by the Federal Emergency Management Agency and State partners needs to be standardized and integrated with high-resolution land cover data to allow flooding, shading, and aquatic habitat conditions to be modeled.

3. Nationally consistent, 14-digit hydrologic unit code (HUC) boundaries updated with the 1- or 5-meter DEM and distributed by the U.S. Geological Survey or U.S. Environmental Protection Agency as Federal Geographic Data Committee standard units. Alternatively, 12-digit HUC boundaries would be acceptable. If possible, 14-digit HUCs could include valley type and have flow volume summarized and attributed.

4. National standards for aquatic stream habitats (not conditions), so mapping pools, riffles, cascades, and other major stream types is possible.

5. Methodologies for linking fish or aquatic invertebrate observations and collections to stream covers in a consistent way.

6. Attributed nation-wide valley type maps (possibly using the Rosgen Stream Classification Technique).

7. An aggregated wetlands layer across the lower 48 States, Hawaii, Puerto Rico, and other parts of the United States (aside from Alaska), based on the 1- or 5-meter DEM, that only includes attributes needed to model ecosystem service provision. This layer would include all adjacent NWI wetland polygons to be analyzed as to water volume, based on system averages described in Cowardin and others (1979), and then merged into a single wetland (or wetland complex) map.

Student Projects

To facilitate student involvement in the workshop, a competition was held inviting student submissions of self-contained or openly deployed mobile or web applications, map mashups, or compelling and dynamic visualizations that demonstrate innovative approaches to the integration, analysis, and visualization of environmental data. Students from the five winning submissions were awarded travel grants to attend the workshop, and this section describes their products.

ComFlo Invasion Model

Models to create projections of invasive species spread are critical in identifying high-priority shipping centers and pathways of bioinvasion, essential information to prevent environmental and economic harm caused by invasive species. In this work, Ashish Gauli (Fisk University), Nathan Wikle (Truman State University), and Ryan Yan (College of William and Mary) developed an interactive web-based simulation system called ComFlo to project the movement of invasive species across the United States. Freight Analysis Framework commodity shipping data from 1997 to 2012, provided by the Federal Highway Administration, were used to produce the model. Specifically, the data were used to generate the transition matrix at the core of a stochastic matrix model that projects terrestrial invasive species spread in the United States through the domestic commodity flow network. In addition, environmental predictors, including precipitation and temperature, were used to predict the movement of the invasive population to regions that have a suitable habitat for the species. The resulting model is unique in its reliance on FAF shipping data in conjunction with the environmental data and can be used for a variety of species that may be transported within commodity shipments. The hope is to continue to improve the model to identify specific domestic shipping routes and trade hubs that are primary conduits of invasive species dispersal. Such information should prove useful to managers tasked with limiting the spread and damage caused by invasive species within the United States.

The invasive model can be accessed here: http://www.nimbios.org/sre/comflo.

Mutualism Network in Space

Azteca sericesur is a keystone ant species in Neotropical coffee agroecosystems in Chiapas, Mexico. Its spatial distribution significantly affects the density and incidence of *Coccus viridis* (the green coffee scale, which also is a

mutualism partner of the ant) and the natural enemies of the ant and the green coffee scale. The ant nest clustering app by Hsun-Yi Hsieh (University of Michigan, Ann Arbor) applies linear algebra and network theory to generate an unlimited number of mutualism networks in space as a basis for further explorations of spatial relations between the mutualism and other organisms.

Using the principals of linear algebra and network theory, "critical distances" (0–70 meters) were used to generate an unlimited number of networks. More specifically, these networks were created based on distance between two ant nests that form mutualism with the green coffee scale; if two nests fall within a critical distance (such as within 5 meters), they belong to the same cluster. The components in each network signify the clusters of mutualism. They are represented in different colors. Advanced uses of these networks include mapping them with the densities and incidence of organisms associated with the mutualism networks, such as the biocontrol agents of the green coffee scale.

The app can be accessed here: https://hhsieh.shinyapps.io/Shiny

Local Climate Effects: Ephemera

The Ephemera app by Matthew Kling (University of California, Berkeley) visualizes the ever-changing interactions between climate, geography, and biotic communities in an era of anthropogenic climate change. The app is a Shiny R application and is a place-based tool aimed at helping users understand dimensions of local climate change by placing that change in the context of climate variation across geography, seasons, and vegetation types. Temperature and precipitation data from three periods (1948–80, 1981–2012, 2041–70) highlight future projected change and trends already observed during the last 60 years. Data sources include academic (climate data from Climate North America) as well as government and NGO (vegetation data from LANDFIRE.gov).

This is a Shiny R application that runs out of RStudio. It runs smoothly on the latest version of R and Studio on the machines tested. Note that it autoinstalls several public packages from the Comprehensive R Archive Network. Download and package installation could take several minutes. During installation, a popup dialogue or a message in the console may ask about installing from local binaries; choose "yes."

Predicted Threat to North American Salamander Biodiversity from the Lethal Pathogen, *Batrachochytrium salamandrivorans*

Emerging infectious diseases (EIDs) are increasingly affecting wildlife and are a major threat to global biodiversity. Chytridiomycosis, an EID caused by two fungal pathogens, *Batrachochytrium dendrobatidis* (Bd) and *Batrachochytrium salamandrivorans* (Bsal), has been implicated in an

amphibian pandemic. Whereas Bd has infected more than 500 species and caused declines and extinctions in more than 200 species globally, little is known about the newly described Bsal. Lethal to salamanders and likely from Asia, Bsal has been documented in the pet trade and poses a substantial threat to salamander populations worldwide. A Bsal invasion in North America could lead to rapid declines and extinctions in the world's salamander biodiversity hotspot. The potential biological effect and geographic scope of Bsal in North America were estimated by producing a vulnerability model using open source software and publicly available environmental data. A Bsal habitat suitability model was created by Tiffany Yap (University of California, Los Angeles) using maximum entropy software, Maxent (http://biodiversityinformatics.amnh.org/open_source/maxent/), potential carrier occurrence data from Vertnet (http://www.vertnet.org), and bioclimatic data from WorldClim (www.worldclim.org). Host availability (salamander richness) was estimated from ranges obtained from International Union for Conservation of Nature and Natural Resources (http://www.iucnredlist.org) and AmphibiaWeb (https://amphibiaweb.org) using R (https://www.r-project.org). The habitat suitability model was weighted with salamander richness and indicated the most likely ports of entry from U.S. salamander trade data (U.S. Fish and Wildlife Service Freedom of Information Act) to produce a salamander vulnerability model. The resulting model provides targeted areas for Bsal surveillance and monitoring, and multiple agencies are using it to aid in the response to this biodiversity crisis. Studies are ongoing, using similar methodologies to continue addressing amphibian biodiversity conservation.

In Search of the "Common" Trees

This study by Enrique Montaño (University of Maryland and Goddard Space Flight Center) and Serenity Montaño (Smithsonian Institution) examined tree species distributions and their contributions to the composition of State, regional, and national tree species abundance to determine "common" tree species.

A self-contained R script was developed that will, in a single step, query the U.S. Geological Survey's Biodiversity Information Serving Our Nation (USGS BISON) data (all providers) and output a meaningful visual analysis of all non-hybrid tree species composition and distribution in the United States, including tables and maps. As configured, the script can run in a single step, accessing data publicly available on BISON.

The package as written runs the initial query, and then the results are analyzed discreetly for each standardized region of the Natural Resource Conservation Service (NRCS) boundaries (West Region, Central Region, Southeast Region, Northeast Region).

Species occurrence data were examined for 1,389 species of vascular plants determined by U.S. Department of

Agriculture's Plant List of Accepted Nomenclature Taxonomy and Symbols to qualify as trees; query results detailed species present in each region, species distribution across each region, cumulative contribution of species occurrence as percentages of the regional composition, and national and regional occurrence counts per species. The results include occurrences from greater than 50 to 500,000.

Findings demonstrated that 33 species account for the top 50 percent of cumulative composition across all recorded occurrences across all regions.

Summary and Conclusions

This diverse group of workshop participants, offering perspectives from 19 Federal agencies and many more nongovernmental organization, academic, and private sector representatives, was highly aligned in identifying and prioritizing national biodiversity informatics needs. For example, participants recognized the importance of building and supporting the infrastructure to leverage the volumes of biodiversity data being observed and shared by our increasingly networked world. The rapidly growing demand for ecological modeling applications highlights the need for investment in nationally consistent data layers on patterns of climate, biodiversity, and environment. The advent of virtually unlimited cloud-based storage and analysis capacity can support intensive modeling efforts that have not previously been possible, such as a distribution model for every U.S. species. These activities are ready to come together in support of decision-driven ecological informatics.

For example, substantial expansion in our understanding of the distributions of species and ecosystems is within our grasp. Leveraging existing data and investing in modeling capacity, we could dramatically improve knowledge of the probable location of species and ecosystems, information needed for regulatory, management, and conservation purposes. Inventory and monitoring programs of Federal and State agencies, academia, and nongovernmental organizations already collect a vast amount of biodiversity data. Citizen science has the capacity to generate huge volumes of biodiversity observation data. Investment should be directed toward developing the optimal set of environmental monitoring metrics that can be fed by existing inventory and monitoring and citizen science efforts, can detect biodiversity response to global change, and can be aggregated across ecological and political boundaries. Monitoring targets can be strategically designed to maximize trend detection capacity. There is high demand for nationally consistent predictor layers to support ecological modeling made possible by increasing volumes of and easy access to biodiversity observation data. We have available all the components of a dynamic biodiversity assessment, inventory and monitoring regime that involves citizen science to check the pulse of key metrics over time. Participants' opinions were almost equally divided in what they see

as the biggest challenge in data dissemination: ensuring data are not misinterpreted, dealing with the issue of data overload, tracking how data are used, standards in web service metadata, and "other" (appendix 4).

The next step is to articulate in greater detail one or more large-scale initiatives that integrate various components of biodiversity and environmental data, both observed and modeled, to measure and communicate biotic response to one or more drivers of change. The goal is to build on existing investments in data aggregation and access, to leverage and expand the community of data providers and consumers, and to do analyses that inform management and conservation actions.

References Cited

Executive Office of the President of the United States, 2015, Incorporating ecosystem services into Federal decision making: Washington, D.C., The White House archives, memorandum M–16–01, accessed February 16, 2017, at https://obamawhitehouse.archives.gov/sites/default/files/omb/memoranda/2016/m-16-01.pdf.

Homer, C.G., Dewitz, J.A., Yang, L., Jin, S., Danielson, P., Xian, G., Coulston, J., Herold, N.D., Wickham, J.D., and Megown, K., 2015, Completion of the 2011 National Land Cover Database for the conterminous United States—Representing a decade of land cover change information: Photogrammetric Engineering and Remote Sensing, v. 81, no. 5, p. 345–354.

Pollinator Health Task Force, 2015, Pollinator research action plan: Washington, D.C., The White House archives, 85 p., accessed February 16, 2017, at https://obamawhitehouse.archives.gov/sites/default/files/microsites/ostp/Pollinator%20Research%20Action%20Plan%202015.pdf.

President's Council of Scientific Advisors on Science and Technology, 2011, Sustaining environmental capital—Protecting society and the economy: Washington, D.C., White House archives, 126 p., accessed February 16, 2017, at https://obamawhitehouse.archives.gov/sites/default/files/microsites/ostp/pcast_sustaining_environmental_capital_report.pdf.

U.S. Department of the Interior, 2016, Safeguarding America's lands and waters from invasive species—A national framework for early detection and rapid response: Washington D.C., 55 p., accessed February 16, 2017, at https://www.doi.gov/sites/doi.gov/files/National%20EDRR%20Framework.pdf.

Appendixes 1–5

Appendix 1.   Seven Questions for Every Breakout Session

Members of each breakout session were asked to address the following questions during their discussion. Results of their discussion and members' additional input were presented in plenary oral and subsequent written reports and are summarized elsewhere within this report. The answers shown below are fictitious examples only.

1. What is the goal of the breakout—clearly and concisely?

    A national assessment of the distribution, availability, and future prospects for four leaf clovers.

2. Why is this a national priority?

    Four leaf clovers represent a highly underused resource. With the way that things are these days, all that we have to rely on is luck. A national assessment would quantify our Nation's access to this important source of luck and elucidate the various factors that will affect its availability in the future.

3. What resources exist within the Federal government for achieving it?

    The U.S. Department of Agriculture's Plant List of Accepted Nomenclature Taxonomy and Symbols (USDA PLANTS) database and the U.S. Geological Survey's Biodiversity Information Serving Our Nation mapping project (USGS BISON) provide significant access to data on the distribution of different clover species. The National Science Foundation would consider proposals for new scientifically defensible and innovative techniques to identify and map four leaf clovers in particular.

4. What resources exist within the non-Federal purview for achieving it?

    Discoveries during this millennium in the private sector (for example, https://doi.org/10.2135/cropsci2009.08.0457) have elucidated the genetic basis of the four leaf phenotype. Significant citizen science knowledge on the current state of the resource is available, especially in kindergartens. Substantial citizen collections are available pressed among the pages of old books and between pieces of wax paper across the country.

5. How can these resources be more complimentary and synergistic?

    Now that we know the genetic basis, we can move beyond the highly reliable, but minimally scalable, method of leaf counting for identification. A National Science Foundation proposal will be developed to identify four leaf clovers on a broad scale through environmental deoxyribonucleic acid (eDNA) methods facilitated by lawnmower mounted sensors in yards across the country. The Natural Resources Conservation Service will examine mandatory reporting of four leaf clovers on Farm Plans.

6. What are the major gaps that need to be filled?

    Ancillary research is needed on the effects of five- to eight- leaved clovers to determine the luck levels that they provide. Luck as an ecosystem service must be clarified and parameterized.

7. What are the most significant challenges that stand in the way of this goal becoming a reality?

    Kindergarteners generally do not do well with forms and are difficult to organize. Metadata from pressed citizen science collections is generally lacking. Every herbarium sheet that led to the clover records in USDA PLANTS and USGS BISON must be re-examined to determine the number of leaves. Large-scale modeling efforts will be needed to prioritize efforts for survey resources.

Appendix 2.   Tools Café Program

All interested participants were invited to demonstrate their digital tools and applications with other members during an extended lunch hour on day 2 of the workshop. A summary of those demonstrations is shown in table 2.1.

Table 2.1.   Demonstrations during day 2 of the biodiversity and ecological informatics workshop, December 3, 2015.

Title	Description	Presenter (and affiliation)
In Search of the "Common" Trees (student competition winner)	We developed a self-contained R script that will, in a single step, query the U.S. Geological Survey's (USGS) Biodiversity Information Serving Our Nation mapping project (BISON)database and output a meaningful visual analysis of all nonhybrid tree species composition and distribution in the United States, including tables and maps.	Enrique Montaño (University of Maryland and Goddard Space Flight Center) and Serenity Montaño (Smithsonian Institution)
ComFlo invasion model (student competition winner)	Interactive web-based simulation system to project the movement of invasive species across the United States. Freight Analysis Framework (FAF) commodity shipping data from 1997 to 2012, provided by the Federal Highway Administration, produce our model. Environmental predictors, including precipitation and temperature, from the National Oceanic and Atmospheric Administration website, were used to predict the movement of the invasive population to regions that have a suitable habitat for the species.	Ashish Gauli (Fisk University), Nathan Wikle (Truman State University), and Ryan Yan (College of William and Mary)
Predicted threat to North American salamander biodiversity from the lethal pathogen, *Batrachochytrium salamandrivorans* (student competition winner)	We estimated the potential biological effect and geographic scope of *Batrachochytrium salamandrivorans* (Bsal), an emerging amphibian pathogen, by producing a vulnerability model using open source software and publicly available environmental data. The model uses maximum entropy software, Maxent, potential carrier occurrence data from Vertnet, and bioclimatic data from WorldClim to predict the most likely ports of entry from U.S. salamander trade data (U.S. Fish and Wildlife Service Freedom of Information Act) and produce a salamander vulnerability model for North America. Read more: https://www.sciencemag.org/content/349/6247/481	Tiffany Yapp (University of California, Los Angeles)
Ephemera (student competition winner)	This is a Shiny R application that runs out of Rstudio and visualizes the ever-changing relations between climate, geography, and biotic communities in an era of anthropogenic climate change. It is a place-based tool aimed at helping users understand dimensions of local climate change by placing that change in the context of climate variation across geography, seasons, and vegetation types.	Matthew Kling (University of California, Berkeley)
Mutualism network in space (student competition winner)	This app applies linear algebra and network theory to generate an unlimited number of mutualism networks in space, as a basis for the exploration of the spatial relationships between the Mexican keystone ant *Azteca sericesur* and *Coccus viridis* (the green coffee scale), which is a mutualism partner of the ant. The components in each network signify the clusters of mutualism. Advanced uses of these networks include mapping them with the densities and incidence of organisms associated with the mutualism networks, such as the biocontrol agents of the green coffee scale. Access here: https://hhsieh.shinyapps.io/Shiny	Hsun-Yi Hsieh (University of Michigan, Ann Arbor)

Table 2.1.   Demonstrations during day 2 of the biodiversity and ecological informatics workshop, December 3, 2015.—Continued

Title	Description	Presenter (and affiliation)
Scalable, transparent, and efficient scientific workflows— USGS–R geoknife and dataRetrieval packages	Gridded time series data (that is, climate projections, historical precipitation, land use) and water cycle observations (that is, flow, water quality, water availability) are important for many ecosystem inquiries, but the volume and complexity of these data can make access, quality control, and data manipulation cumbersome. Two packages developed by the USGS–R assist scientists in getting, cleaning, and analyzing data in the R programming environment and use USGS data portal web services.	Emily Read (U.S. Geological Survey)
Facilitating species distribution model using the Software for Assisted Habitat Modeling	The Software for Assisted Habitat Modeling (SAHM) is a set of free modules within the VisTrails workflow management and provenance software that facilitates species distribution modeling including preprocessing data, exploring data, fitting models, and visualizing results. SAHM can harvest species location data directly from BISON as input to models and combine it with standard predictor variables such as climate data to produce a habitat suitability model.	Catherine Jarnevich (U.S. Geological Survey)
MAPIT (Mapping Application for Freshwater Invertebrate Taxa)	The Western Center for Monitoring and Assessment of Freshwater Ecosystems and the Bureau of Land Management's National Aquatic Monitoring Center provide access to distribution data for about 5,500 freshwater invertebrate taxa contained in a large biodiversity database (about 1.5 million records from about 50,000 samples) that the 2 centers jointly manage, compiled from a variety of sources. The Mapping Application for Freshwater Invertebrate Taxa (MAPIT) allows users to search for freshwater invertebrate taxa of interest to display in Google Maps format. Information on individual records can be viewed by clicking on mapped locations, and records of interest can be downloaded as a .csv file. Data are used to build individual and multitaxon species distribution models and to inform the biodiversity status of freshwater ecosystems in the United States.	Chuck Hawkins (Utah State University)
GeoPlatform Web Map Viewer	The GeoPlatform web map viewer was developed to increase users' ability to search, integrate, and visualize open geospatial information. The viewer is an effective tool to see and interact with spatial data from the primary hubs of the ecoInforma community. Read more and access here: https://www.geoplatform.gov/	Derek Masaki (U.S. Geological Survey)
EnviroAtlas	An online decision support tool that government (local to Federal) agencies, industry, communities, and public interest groups can use to help inform policy and planning decisions impacting the places we live, learn, work and play. EnviroAtlas also provides a valuable resource for educators, researchers, and the public combining maps, analysis tools, fact sheets, use stories, and accessible data into an easy-to-use, web-based resource. Access here: https://epa.gov/enviroatlas	Annie Neale (Environmental Protection Agency)
National Conservation Easement Database	The National Conservation Easement Database is a national spatial database of conservation easement information that helps agencies, land trusts, and other organizations plan strategically, advance public accountability and that helps make the case for conservation. Access here: https://www.conservationeasement.us/	Jes Skillman (Ducks Unlimited)
Deepwater Horizon Project Tracker	The Deepwater Horizon Project Tracker is a centralized directory of projects funded because of the *Deepwater Horizon* oil spill and provides a comprehensive picture of the location, type, cost, funding sources, and scope of recovery, restoration, and research projects. Access here: http://dwhprojecttracker.org/	Jes Skillman (Ducks Unlimited)

Table 2.1.   Demonstrations during day 2 of the biodiversity and ecological informatics workshop, December 3, 2015.—Continued

Title	Description	Presenter (and affiliation)
Biodiversity tools for Virginia	The mission of the Virginia Natural Heritage Program is to conserve biodiversity. To that end, the program has developed tools that facilitate scientifically informed conservation actions, including ConservationVision, the Virginia Wetlands Catalog, and species distribution models and conservation sites on the Natural Heritage Data Explorer.	Joe Weber and Jason Bulluck (Virginia Department of Conservation and Recreation)
Reporting and aggregating invasive species mapping data with EDDMapS and regional smartphone apps	EDDMapS is a database-driven website for aggregating and mapping invasive species data across North America. Invasive species can be reported through EDDMapS or one of the regional Bugwood apps to contribute to the distribution maps, and data also can be downloaded for use in research projects, modeling programs, and more.	Rebekah Wallace (University of Georgia)
BioData—An authoritative data source for over 17,000 USGS aquatic community samples	BioData is an authoritative source of USGS community-level data for fish, macroinvertebrate, and periphyton data that currently stores over 17,000 fish, macroinvertebrate, and algae samples collected since 1993 from more than 2,000 sites across the Nation. In aggregate, the data include more than 1 million species occurrence and abundance records. These data are available to the public and can be retrieved using a simple web interface and are delivered using a standard, up-to-date taxonomic system that resolves synonymy and ambiguity across the entire period of record of the dataset. Access here: https://aquatic.biodata.usgs.gov/landing.action	Peter Ruhl (U.S. Geological Survey)
The Water Quality Portal (WQP)—a single, integrated source for EPA STORET, USGS Water Mission Area, and Department of Agriculture ARSS water quality and biology data	The Water Quality Portal (WQP) is a cooperative web service sponsored by the USGS, the U.S. Environmental Protection Agency (EPA), and the National Water Quality Monitoring Council (NWQMC) that serves more than 250 million water-quality records from more than 2.2 million sites that were sampled by more than 400 State, Federal, Tribal, and local agencies. During 2015 the EPA and USGS collaborated to add USGS and Storage and Retrieval database (STORET) fish community using the Water Quality Exchange (WQX) data exchange standard. Data are available online through a simple drop-down menu format or through web services. Access here: http://www.waterqualitydata.us/	Peter Ruhl (U.S. Geological Survey)
Data Basin	DataBasin.org is an open access, science-based mapping and analysis platform facilitating learning, research, and sustainable environmental stewardship. The platform allows you to download and upload spatial data, to create custom visualizations, drawings, and analyses, and to use collaborative group tools.	Tosha Comendant (Conservation Biology Institute)
Ecoengine API & Explore Tools	The Berkeley Initiative for Global Change Biology's efforts to mobilize and access biological and ecological data resulted in the Ecoinformatics Engine ("EcoEngine") application programming interface (API), which brings together data from natural history museums, field stations, sensor networks, archives, and climate models. The Ecoengine API allows development of rich web tools to explore and compare data, enabling researchers to address questions of global change biology. The scope is global with an initial focus on western North America where the University of California, Berkeley's holdings are richest. Access here: https://holos.berkeley.edu	Michelle Koo (University of California, Berkeley)

Appendix 3.   List of Participants of the Biodiversity and Ecological Informatics Workshop, December 2–4, 2015

A list of participants' names based on badge pickup and agenda and subgroup participation is provided in table 3.1 and is current as of the workshop in December 2015.

Table 3.1.   Participants of the biodiversity and ecological informatics workshop, December 2–4, 2015.

Last name	First name	Job title	Affiliation
Abrahms	Briana	Ph.D. candidate	University of California, Berkeley
Akre	Thomas	Ecologist/program director	Smithsonian Institution
Aplet	Greg	Senior science director	The Wilderness Society
Ashe	Dan	Director	U.S. Fish and Wildlife Service
Aulenbach	Steve	Biologist	U.S. Geological Survey
Balaji Bhaskar	Maruthi Sridhar	Assistant professor	Texas State University
Barragan	Claudia	Founder at The City Ecologist	American University
Beach	Jim	Assistant director for informatics	Kansas Biodiversity Institute, University of Kansas
Belote	Travis	Ecologist	The Wilderness Society
Benson	Abby	Biologist	U.S. Geological Survey
Bernal	Blanca	Postdoctoral fellow	Smithsonian Institution
Blum	Stanley	Biodiversity information architect	Taxonomic Databases Working Group
Brown	Joel	Team leader	National Ecological Site Team, Natural Resources Conservation Service
Bulluck	Jason	Natural heritage information manager	Virginia Department of Conservation and Recreation
Burkett	Chris	Wildlife action plan coordinator	Virginia Department of Game and Inland Fisheries
Cannister	Matt	Biologist	U.S. Geological Survey
Cannon	John	Analyst	Marstel-Day
Casey	Frank	Economist	U.S. Geological Survey
Cole	Daniel	Geographic information system (GIS) coordinator	Smithsonian Institution
Coleman	Kaycee	Statistician (biology)	U.S. Fish and Wildlife Service
Comendant	Tosha	Senior scientist	Conservation Biology Institute
Covington	Ryan	Conservation GIS analyst	Defenders of Wildlife
Cragin	Melissa	Staff associate	National Science Foundation
D'Agnese	Frank	President and chief technology officer (CTO)	Earth Knowledge Network
Daniel	Jessica	Oak Ridge Institute for Science and Education (ORISE) fellow	Environmental Protection Agency
Dempsey	Hannah	Student	College of William and Mary
DeVivo	Joe	Deputy chief for science	Inventory and Monitoring Division, National Park Service
Eccles	Jordan	Analyst	Marstel-Day
Fay	Frank	Fire ecologist	Landfire, U.S. Department of Agriculture (USDA), Forest Service
Fitzgibbon	Michael	CTO	Point Blue

Table 3.1.   Participants of the biodiversity and ecological informatics workshop, December 2–4, 2015.—Continued

Last name	First name	Job title	Affiliation
Gardiner	Ned	Communications and education producer	National Oceanic and Atmospheric Administration
Gauli	Ashish	Student	Fisk University
Gergely	Kevin	Biologist	U.S. Geological Survey
Gill	Mike	Vice-chair	Group On Earth Observations Biodiversity Observation Network
Guala	Stinger	Workshop co-organizer, director of the Integrated Taxonomic Information System (ITIS) and the Biodiversity Information Serving Our Nation mapping project (BISON)	U.S. Geological Survey
Guralnick	Rob	Curator	Biodiversity Informatics, Florida Museum of Natural History
Hamilton	Healy	Workshop co-organizer, chief scientist	NatureServe
Hannibal	Mary Ellen	Journalist	Journalist
Hasson	Tomer	Science coordinator	Department of the Interior
Hawkins	Charles	Director and professor	Aquatic Ecology & Watershed Science, Utah State University
Hitchcock	Ann	Senior advisor	Science Collections & Environmental Safeguards, National Park Service
Holdsworth	Andy	Supervisor	Data and Performance Unit, Minnesota Department of Natural Resources
Honey	Leslie	Vice president, conservation services	NatureServe
Howard	Tim	Director of science	New York Department of Environmental Conservation
Hsieh	Hsun-Yi	Ph.D. student	University of Michigan
Jarenvich	Catherine	Research ecologist	U.S. Geological Survey
Kagan	Jimmy	Director	Oregon Biodiversity Information Center
Keller	Kaylene	Data manager, ecologist	Inventory and Monitoring Program, U.S. Fish and Wildlife Service
Kingston	Simon	Data ranger	National Park Service
Klein	Mary	President and chief executive officer (CEO)	NatureServe
Kling	Matthew	Graduate student	University of California, Berkeley
Kodad	Yassine	Computer scientist	U.S. Geological Survey
Koo	Michelle	Staff curator of biodiversity informatics and GIS	University of California, Berkeley
Kumar	Anjali	Economist	U.S. Agency for International Development
Lacher	Iara	Landscape ecologist	Smithsonian Conservation Biology Institute
Laurent	Ed	Executive director	Connecting Conservation
Laurent	Kevin	Computer scientist	U.S. Geological Survey
Lippmann	Michael	Principal and co-founder	Blue Raster
Loarie	Scott	Co-director, iNaturalist.org	California Academy of Sciences
Long	John	Computer scientist	U.S. Geological Survey
Malcom	Jacob	Endangered species science & policy analyst	Defenders of Wildlife
Martin	Elizabeth	Biologist	U.S. Geological Survey

Table 3.1.   Participants of the biodiversity and ecological informatics workshop, December 2–4, 2015.—Continued

Last name	First name	Job title	Affiliation
Masaki	Derek	BISON product owner	U.S. Geological Survey
McClees-Funinan	Ricardo	Physical scientist	U.S. Geological Survey
McDermott	Mike	Deputy directory	Core Science Analytics Synthesis, and Libraries, U.S. Geological Survey
McKerrow	Alexa	Biologist	U.S. Geological Survey
McNiff	Marcia	Technical information specialist	U.S. Geological Survey
McShea	William	Research ecologist	Smithsonian Institution
Milam	Amy	Director and founder	EcoLogic LLC
Mills	Craig	CEO	Vizzuality
Mitchell	Richard	Assessment lead	National Rivers & Streams Assessment, U.S. Environmental Protection Agency
Mize	Aaron	Branch chief	Conservation Planning and Policy, U.S. Fish and Wildlife Service
Montano	Enrique	Geotaxonomist	Smithsonian Institution
Montano	Serenity	Geographer	Smithsonian Institution
Moore	Gerry	Team leader	National Plant Data Team, USDA's Plant List of Accepted Nomenclature Taxonomy and Symbols database
Morris	Cheryl	Director	Core Science Analytics Synthesis, and Libraries Program, U.S. Geological Survey
Murray	Reina	GIS analyst	National Trust for Historic Preservation
Neale	Anne	Project lead	EnviroAtlas, Environmental Protection Agency
Newman	Jana	Branch chief	Inventory & Monitoring, Fish and Wildlife Service
Nightingale	Emily	Science policy fellow	Institute for Defense Analyses (IDA), Science and Technology Policy Institute
OConnell	Allan	Fish ecologist	U.S. Geological Survey
Orrell	Thomas	Research biologist	Smithsonian Institution
Parker	Alison	ORISE fellow	Environmental Protection Agency
Parr	Cynthia	Technical information specialist	USDA National Agricultural Library
Peterjohn	Bruce	Chief	U.S. Bird Banding Lab, U.S. Geological Survey
Petropoulos	Zoe	Science policy fellow	IDA Science and Technology Policy Institute
Phang	Christina	Analyst	Blue Raster
Posner	Stephen	Policy engagement associate	COMPASS
Powers	Ryan	Postdoctoral fellow	Yale University
Rajbhandary	Jasmine	Biologist	U.S. Geological Survey
Rau	Becky	Wildlife biologist, data administrator	U.S. Fish and Wildlife Service
Read	Emily	Scientist	U.S. Geological Survey
Reams	Greg	National leader	Forest Inventory and Analysis, USDA Forest Service
Ries	Leslie	Assistant professor	Georgetown University
Rose	Robert	Director	Center for Geospatial Analysis, College of William and Mary
Ruhl	Pete	Biological data manager	U.S. Geological Survey
Runfola	Dan	Senior geospatial scientist	College of William and Mary

Table 3.1.　Participants of the biodiversity and ecological informatics workshop, December 2–4, 2015.—Continued

Last name	First name	Job title	Affiliation
Sandborn	Avery	Statistician	National Agricultural Statistics Service
Schulz	Katja	Program coordinator	Smithsonian Institution
Schuster	Rudy	Branch chief, social scientist	U.S. Geological Survey
Sellers	Elizabeth	BISON data team	U.S. Geological Survey
Seltzer	Carrie	Program manager	Citizen Science Program, National Geographic Society
Sieracki	Jennifer	Data manager	National Park Service
Silverman	Emily	Statistician	U.S. Fish and Wildlife Service
Simpson	Annie	Biologist & information scientist	U.S. Geological Survey
Skillman	Jane (Jes)	GIS analyst	Ducks Unlimited
Smith	Jim	Project lead	LANDFIRE Project, The Nature Conservancy
Smith	Jonathan	Program coordinator	Land Change Science Program, U.S. Geological Survey
Smyth	Regan	Spatial ecology project manager	NatureServe
Spurrier	Carol	Rangeland ecologist	Bureau of Land Management
Stevenson	Robert	Professor	University of Massachusetts
Strader	Ginger	Director	Smithsonian Institution Scholarly Press
Sutherland	Todd	Data manager	U.S. Fish and Wildlife Service
Taylor	Charlotte	Consultant	Department of Defense Natural Resources Program
Thatcher	Ben	Assistant national coordinator	Landscape Conservation Cooperatives
Thomas	Alexis	Director	University of Florida GeoPlan Center
Thompson	Alexander	IT infrastructure lead	Advanced Computing Lab, University of Florida
Toevs	Gordon	Program lead	Assessment, Inventory, & Monitoring Program, Bureau of Land Management
Tolle	Kristin	Director	Data Science Initiative, Microsoft
Turner	Woody	Program scientist	National Aeronautics and Space Administration
Von Holle	Betsy	Program director	National Science Foundation
Wallace	Rebekah	Data coordinator	University of Georgia Center for Invasive Species & Ecosystem Health
Weber	Joe	Natural heritage conservation biologist	Virginia Department of Conservation & Recreation
Wee	Brian	Chief of strategic alliances	National Ecological Observatory Network
Weltzin	Jake	Ecologist	National Phenology Network, U.S. Geological Survey
Wheeler	Ben	U.S. Global Biodiversity Information Facility (GBIF) node manager	U.S. Geological Survey
Whitley Rogers	Doreen	Geospatial information officer	The Audubon Society
Wikle	Nathan	Student participant	Truman University
Winthrop	Robert	Senior social scientist	Bureau of Land Management
Yan	Ryan	Student	College of William and Mary
Yap	Tiffany	Student	University of California, Los Angeles

Appendix 4.   Questionnaire Results

Feedback from workshop participants was solicited during plenary sessions of the workshop through the Option-Power® interactive audience response system from Option Technologies (https://www.optiontechnologies.com/). To obtain immediate feedback about workshop plenary sessions from the audience, questions with numbered answer options were projected on a screen after each session and polling hardware (commonly called question clickers) distributed to all participants for their vote. Between 22 and 101 participants responded to questions at different times during the workshop.

Interactive Questions

Workshop:
Promoting Synergy in the Innovative
Use of Environmental Data

December 02-04, 2015   Department of Interior South Bldg, 1951 Constitution
Avenue NW, Washington DC

How far did you travel to attend this workshop?

1. 50 miles or less
2. More than 50 miles
3. More than 100 miles
4. More than 500 miles
5. More than 1000 miles
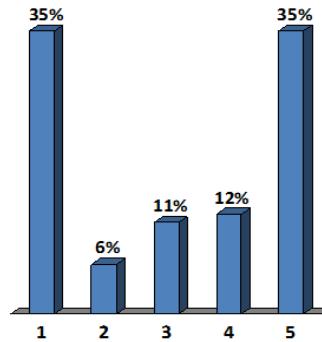
Responses:

#actual/#possible

98 / 150
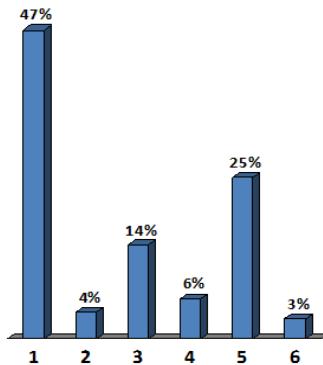
Which of the following best describes your primary role(s) at this workshop? Select the best answer.

1. Data generator / provider
2. Data user
3. Data aggregator
4. Data manager
5. Applications / software developer
6. Other
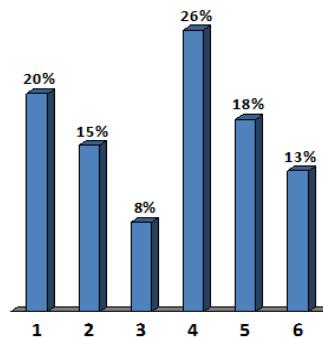
Responses:

#actual/#possible

85 / 150

Which of the following best describes the kind of work you do? Select only one.

1. Inventory and monitoring
2. Modeling
3. Other biological / ecological research
4. Taxonomic collections
5. Technical / scientific data & information dissemination
6. Information technology development
7. Other
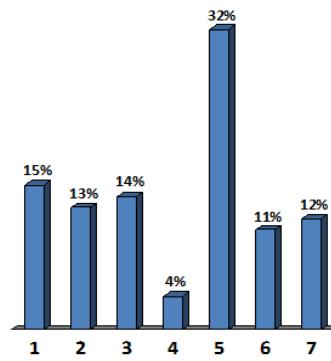
85 / 150

Have you used inventory and monitoring data from federal agencies in your line of work?

1. Yes
2. No
3. I'm not a data user

80%

11%   9%

1   2   3
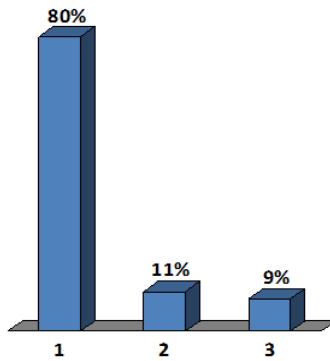
Responses:

#actual/#possible

101 / 150

Have you used inventory and monitoring data from federal agencies in your line of work?

1. Yes
2. No
3. I'm not a data user

79%

12%   9%

1   2   3
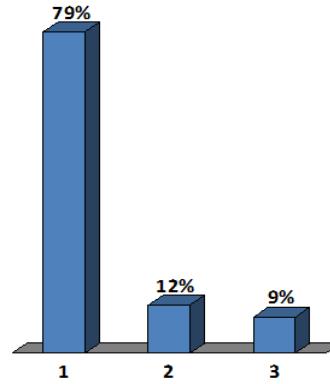
Responses:

#actual/#possible

101 / 150

Have you used analyses or applications from inventory and monitoring program in federal agencies for your line of work?

1. Yes
2. No

71%

29%
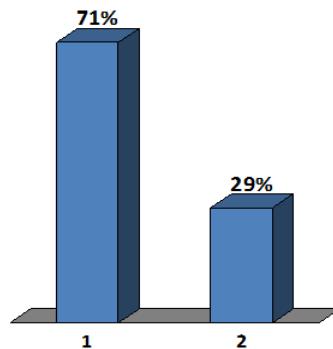
Responses:

#actual/#possible

103 / 150

1   2

Which of the following should be of highest priority for improving accessibility of inventory / monitoring data across spatial scales and jurisdictions? Select only one.

1. Improved data sharing standards
2. Availability of appropriate data aggregators / repositories
3. Improved data agreements and access permission processes
4. Improved methods for presenting data provenance and acknowledgements
5. Improved data delivery mechanisms including visualization
6. Other not listed above

35%

20%

30%

3%   2%

9%

1   2   3   4   5   6

93 / 150

My initial questions about EcoINFORMA were addressed today.

1. Strongly agree
2. Agree
3. Neither agree or disagree
4. Disagree
5. Strongly disagree

48%

34%

14%

2%

1%

1   2   3   4   5
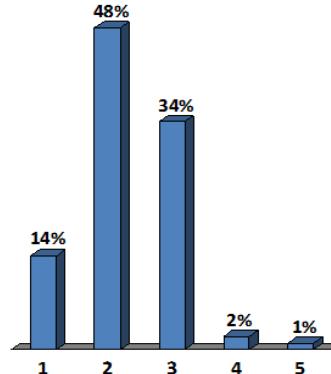
Responses:

#actual/#possible

99 / 150

Have you used data from an EcoINFORMA resource hub (BISON, EnviroAtlas, MRLC)?

1. Yes
2. No
3. I'm not a data user

50%

39%

12%

1   2   3
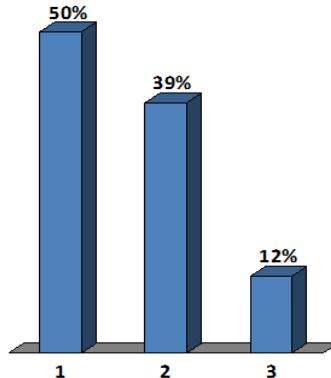
Responses:

#actual/#possible

101 / 150

Have you contributed data to an EcoINFORMA resource hub (BISON, EnviroAtlas, MRLC)?

1. Yes
2. No
3. I'm not a data provider

47%

42%

11%
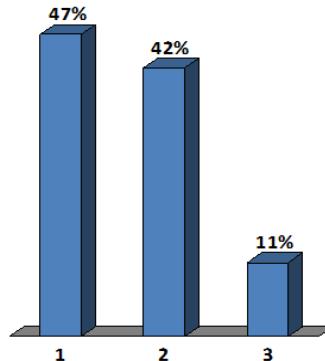
Responses:

#actual/#possible

96 / 150

1   2   3

Which do you think should be EcoINFORMA's highest priority in the coming year? Select only one.

1. Incorporate new resource hubs for identified data gaps
2. Improve data delivery and visualizations within each resource hub
3. Identify environmental issues that could be addressed with existing resource hubs data
4. Identify collaborator needs that could be addressed with an EcoINFORMA collaboration
5. Improve cross-hub data visualizations
6. Develop cross-hub data analyses
7. Other not listed above

17% 18% 19% 20%

6% 9% 10%

1   2   3   4   5   6   7
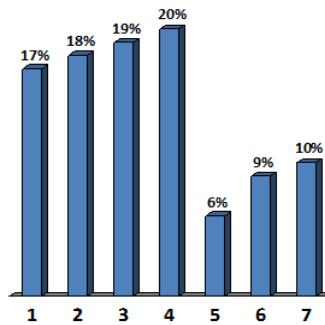
99 / 150

Please select the most preferred option for gathering your input on next steps for EcoINFORMA.

1. Have cross-cutting EcoINFORMA questions for every breakout group
2. Have a focused EcoINFORMA breakout group
3. Have thematic breakout groups of existing EcoINFORMA resource hubs (biodiversity, ecosystem services, land cover dynamics)
4. Have thematic breakout groups of potential new resource hubs (e.g. agriculture, climate, health, etc.)
5. Other not listed above

93 / 150

43%		1
14%		2
26%		3
12%		4
5%		5

NSF   NatureServe   USGS
A Network Connecting Science With Conservation   science for a changing world
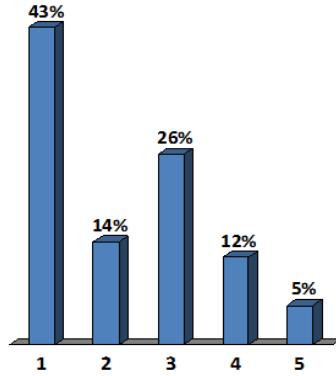
Interactive Questions
# Citizen Science Plenary

Workshop:
Promoting Synergy in the Innovative
Use of Environmental Data

December 02-04, 2015   Department of Interior South Bldg, 1951 Constitution Avenue NW, Washington DC

Have you participated or contributed data via a citizen science program like the ones described today?

1. Yes
2. No
3. I'm not a data provider
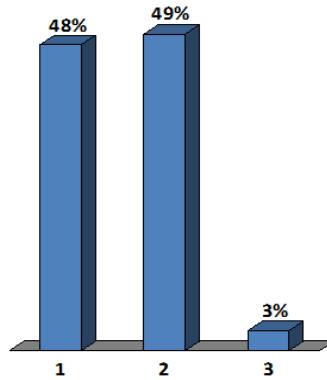
Responses:

#actual/#possible

63 / 150

48% 49% 3%

1   2   3

Have you used citizen science data in your line of work?

1. Yes
2. No
3. I'm not a data user

Responses:

#actual/#possible

22 / 150

77% 23% 0%

1   2   3
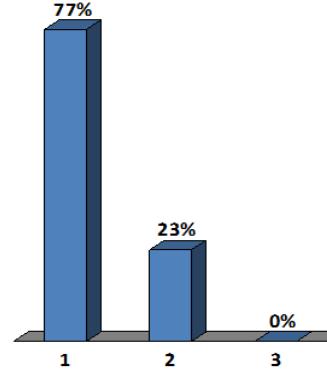
Have you used analyses or applications from citizen science programs in your line of work?

1. Yes
2. No

54%

46%

1   2
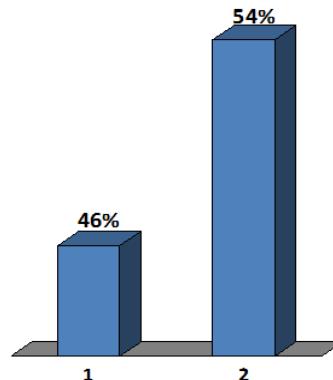
Responses:

#actual/#possible

39 / 150

Which of the following should be a priority for increasing collaboration with citizen science programs in EcoINFORMA? Select only one.

1. Improved data sharing standards
2. Availability of appropriate data aggregators / repositories
3. Improved data agreements and access permission processes
4. Improved methods for presenting data provenance and acknowledgements
5. Improved data delivery mechanisms including visualization
6. Other not listed above

30%

9%

7%

19%   19%

16%

1   2   3   4   5   6
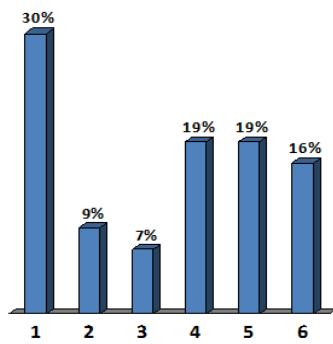
43 / 150

Which citizen science programs strength should be a priority for an <u>EcoINFORMA</u> collaboration? Select only one.

1. Sharing large volume of data
2. Data analysis and use
3. Educational and public outreach components
4. Mobilization of large numbers of volunteers
5. Other not listed above
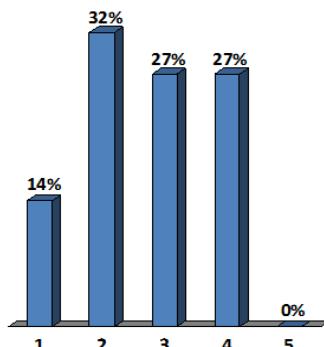

Responses:

#actual/#possible

22 / 150

Interactive Questions: Academia Plenary

Workshop: Promoting Synergy in the Innovative Use of Environmental Data

December 02-04, 2015   Department of Interior South Bldg, 1951 Constitution Avenue NW, Washington DC

How do you usually find out about data or tools from academic sources that are relevant to your work?

1. Project-specific collaborations (data or tool designed for your needs)
2. Workshops or conferences
3. Surfing the Web
4. Reading scientific literature
5. Students, interns, new hires
6. Completely opportunistically
7. None of the above
8. I am not a data user
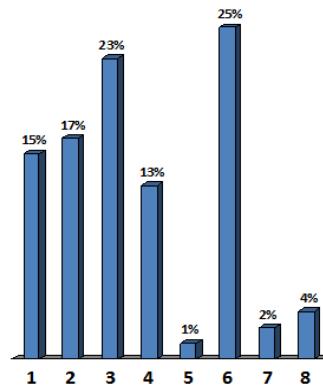
84 / 150

How can we better connect biodiversity and ecoinformatics efforts from the academic community to resource managers and decision makers?

1. Dedicated workshops
2. Dedicated webinars
3. Build funding for outreach into tool development awards
4. Internships and sabbatical opportunities
5. Case studies connecting tools to potential applications
6. Funding and training for staff dedicated to connecting tools to users
7. All of the above
8. None of the above
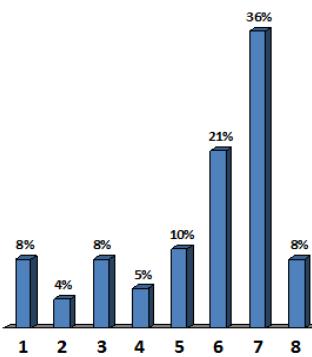
84 / 150

Have you used data from academia in your line of work?

1. Yes
2. No
3. I'm not a data user

72%

20%

9%

1   2   3
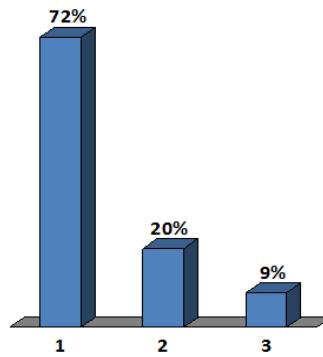
Responses:

#actual/#possible

82 / 150

Have you used analyses or applications from academia in your line of work?

1. Yes
2. No

87%

13%

1   2
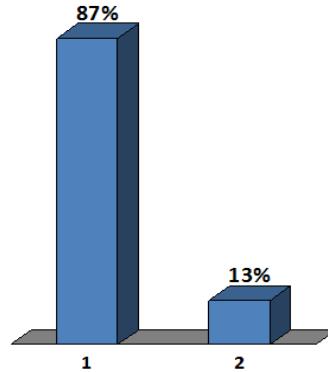
Responses:

#actual/#possible

79 / 150

Interactive Questions:
Private Sector Plenary

Workshop:
Promoting Synergy in the Innovative
Use of Environmental Data

December 02-04, 2015   Department of Interior South Bldg, 1951 Constitution
Avenue NW, Washington DC

How important do you think public/private partnerships are
for generating and disseminating priority national
biodiversity and ecoinformatics datasets?

1. Essential
2. Very important
3. Somewhat important
4. Not relevant
5. No opinion
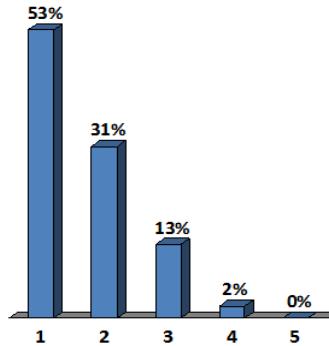
Responses:

#actual/#possible

Q 1/4

89 / 150

53% 31% 13% 2% 0%

1   2   3   4   5

Are you currently part of a public/private partnership described above?

1. Yes
2. No
3. Doesn't apply to me

47%   44%

9%

1   2   3
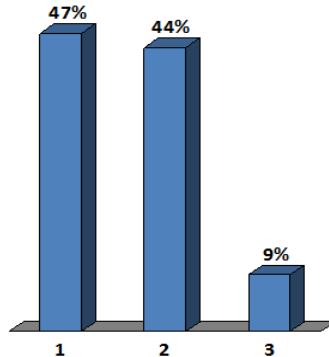
Responses:

#actual/#possible

90 / 150

Are you part of a project or organization that could benefit from greater collaboration with the private sector?

1. Yes
2. No
3. Doesn't apply to me

82%

10%   9%

1   2   3

Responses:

#actual/#possible

94 / 150

What is most important to you in using web tools to generate data visualizations?

1. Telling a compelling story
2. Source data are defendable
3. Updating visualizations with new data is an elegant process
4. All of the above
5. Other

| Responses: |
| --- |
| #actual/#possible" |

97 / 150
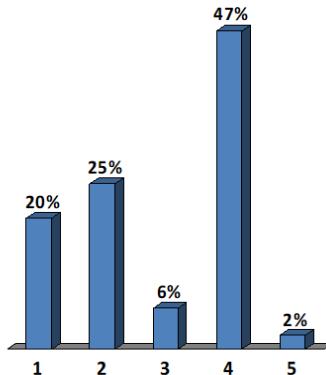
How do you disseminate data within your organization?

1. Data download
2. Direct database connections
3. Web services
4. I do not disseminate data within my organization
5. All of the above
6. Other

Responses:
#actual/#possible"

93 / 150
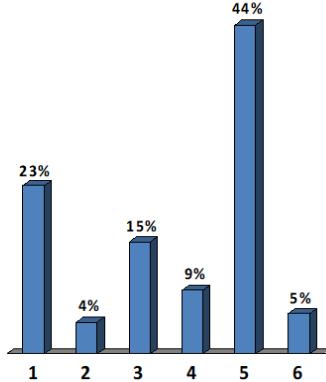
In your work, do you disseminate data to the public?

1. Yes
2. No

Responses:

#actual/#possible

98 / 150

90%

10%

1   2
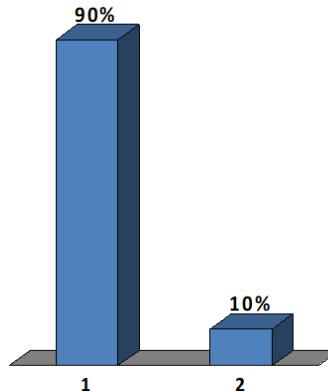
What do you see as the biggest challenge in data dissemination?

1. Standards in web service metadata
2. Tracking how data are used
3. Ensuring data are not misinterpreted
4. Dealing with the issue of data overload
5. Other

Responses:

#actual/#possible

97 / 150

18% 19% 23% 23% 19%

1   2   3   4   5
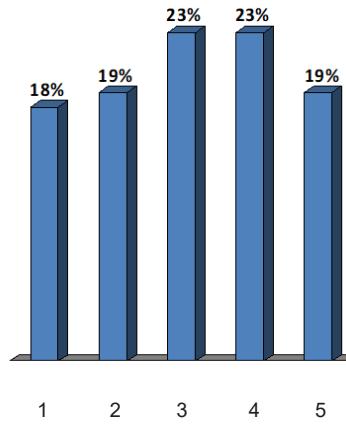
Appendix 5.   Social Media Posts

Throughout the workshop, participants were encouraged to tweet their reactions using the hashtag #ShareUrData. This appendix contains a very brief summary of some of the opinions and observations about the workshop that were expressed on Twitter. The full dialogue, with images, can be found at https://twitter.com/search?q=#ShareUrData.

Dr. Lara Lachner:

- [Create] dynamic data, go beyond observations, integrate w/ private sector, share tools and case studies. Portray data as art, engage students, curate data!

- Identify charismatic species to inspire multi-species conservation across states, organizations, and scientific and public realms.

Annie Simpson:

- Too many environmental scientists are painfully unaware of data sharing standards.

- EcoINFORMA breakout report: watershed and water quality weren't addressed at the workshop. How about creating an EcoINFORMA water hub?

- Inventory and monitoring breakout report: inventory and monitoring information is never complete, but we do need more and better data; it is needed to avoid an impeding sixth mass-extinction.

- Pollinators breakout report: there is a need for a map of chemical applications that may affect pollinators and for the creation of an authoritative list of U.S. pollinators.

- Citizen science breakout report: there is a strong push by U.S. government to have grant writers include citizen science in their proposals.

- Agriculture breakout report: we need to use simulations to fill in data gaps caused by difficulties in scale. We are using data on genetics of agricultural adaptability for predicting responses to climate change.

- Link to open data, semantically enabled, to achieve full synthesis of data within a biogeographic information infrastructure.

- U.S. Fish and Wildlife Director Dan Ashe said, "Think about species that are surrogates for ecosystem health and emphasize them…"

Mary Ellen Hannibal:

- The big thing I'm seeing is most people here don't know what is already out there in citizen science.

- Can a 4th grader provide good science? An iNaturalist anecdote hits a data quality nerve [among workshop participants].

Jes Skillman:

- U.S. Department of Agriculture has created an agricultural data hub at https://data.nal.usda.gov

- U.S. Fish and Wildlife Service Director Dan Ashe said, "If we can't do conservation at a landscape scale, we will be unsuccessful."

Amy Milam:

- Work with environmental data, especially GIS? Then you need to know about https://ecosystems.data.gov

- Jim Smith from LANDFIRE calls for 'real' data [collected by people] on the ground, not more satellites, and for collaborators not just data consumers.

USGS