8-27-2015

Practical Guidance for Integrating Data Management into Long-Term Ecological Monitoring Projects

Robert D. Sutter
Enduring Conservation Outcomes, rsutter@enduringconservation.com

Susan Wainscott
University of Nevada, Las Vegas, sue.wainscott@unlv.edu

John R. Boetsch
National Park Service

Craig Palmer
University of Nevada, Las Vegas, palmerc@unlv.nevada.edu

David J. Rugg
United States Department of Agriculture Forest Service

Wildlife Society Bulletin; DOI: 10.1002/wsb.548

Data Management

Practical Guidance for Integrating Data Management into Long-Term Ecological Monitoring Projects

ROBERT D. SUTTER,[1] *Enduring Conservation Outcomes, 652 Maupas Avenue, Savannah, GA 31401, USA*

SUSAN B. WAINSCOTT, *University Libraries, University of Nevada, Las Vegas, Las Vegas, NV 89154 7014, USA*

JOHN R. BOETSCH, *National Park Service, North Coast and Cascades Network Inventory and Monitoring Program, 600 E Park Avenue, Port Angeles, WA 98362 6757, USA*

CRAIG J. PALMER, *School of Environmental and Public Affairs, University of Nevada, Las Vegas, Las Vegas, NV 89154 4009, USA; and Computer Sciences Corporation, Alexandria, VA 22310 3260, USA*

DAVID J. RUGG, *United States Department of Agriculture Forest Service, 1 Gifford Pinchot Drive, Madison, WI 53726 2366, USA*

ABSTRACT Long-term monitoring and research projects are essential to understand ecological change and the effectiveness of management activities. An inherent characteristic of long-term projects is the need for consistent data collection over time, requiring rigorous attention to data management and quality assurance. Recent papers have provided broad recommendations for data management; however, practitioners need more detailed guidance and examples. We present general yet detailed guidance for the development of comprehensive, concise, and effective data management for monitoring projects. The guidance is presented as a graded approach, matching the scale of data management to the needs of the organization and the complexity of the project. We address the following topics: roles and responsibilities; consistent and precise data collection; calibration of field crews and instrumentation; management of tabular, photographic, video, and sound data; data completeness and quality; development of metadata; archiving data; and evaluation of existing data from other sources. This guidance will help practitioners execute effective data management, thereby, improving the quality and usability of data for meeting project objectives as well as broader meta-analysis and macrosystem ecology research. © 2015 The Wildlife Society.

KEY WORDS data management, graded approach, iterative design, long-term ecological monitoring, metadata, quality assurance.

Long-term studies are essential for understanding the dynamics of ecological systems and the effectiveness of a wide variety of ecosystem or population management activities (Spotila et al. 1996, Nussear and Tracy 2007, Davies et al. 2012, Dodds et al. 2012, Williams and Brown 2012, Stein et al. 2013). A characteristic of successful long-term ecological monitoring (LTEM) studies is that the data are of sufficient completeness, quality, and availability to allow data interpretation by different users (Michener et al. 1997, Palmer 2003, Lindenmayer and Likens 2010). The availability of this information should be robust to changes in the status of individuals involved in the project, including gradual changes, such as remembering details of the project, or discrete events, such as staff reassignment, retirement, or death (Michener et al. 1997, Lindenmayer and Likens 2010, Vines et al. 2014). Yet, data for many LTEM studies are not synthesized (Dodds et al. 2012), consistently collected, or sufficiently documented that end users can understand data quality and usability (Lovett et al. 2007, McEachern and Sutter 2010).

Ensuring that this information is available for future data collection and analysis requires detailed and comprehensive procedures to manage the data (Fancy and Bennetts 2012). A data management section in a monitoring protocol or a separate data-management plan provides guidance to ensure that the data collected are complete, of the quality desired, available for analysis and sharing, and archived for future use. Without this guidance, the potential for data problems increases: data loss, data of inconsistent or unknown quality, inability to relocate sampling locations, and changes in sampling methods that lead to data discontinuities and/or an unintentional reduction in overall information value over the life of a study (McEachern and Sutter 2010).

Data management, description, and sharing are receiving increased attention in the literature (Reichman et al. 2011, Fancy and Bennetts 2012, Sergeant et al. 2012, Rüegg et al. 2014). The formal literature has focused more on theoretical reasons for data management, and has not provided guidance for the collection and management of data that can be easily implemented by teams designing LTEM projects. Even recent books on long-term monitoring barely mention data management or documentation (Gitzen et al. 2012). The

Received: 19 September 2014; Accepted: 27 March 2015

[1]*E-mail: rsutter@enduringconservation.com*

gray literature has provided more detailed guidance (Martin and Ballard 2010), but publications are not easily accessible. However, written data-management plans are being required by regulatory agencies and research funding organizations in the United States and European Union (Jones 2009, Dietrich et al. 2012, OSTP 2013). The implication of these developments is that data documentation should be clear to audiences external to the monitoring program team, and sufficiently detailed to reduce the potential for unintentional misuse of the data.

The objective of this paper is to provide practical guidance for the integration of effective data management into LTEM projects. We address data management as it affects collection, project stewardship, and archiving of data. The guidance is detailed enough that it can be easily implemented by those who design and implement LTEM projects, including researchers, land managers, and graduate students. We assume that documentation and description of data management activities and procedures are integrated into a larger document that clearly states the project objectives, and provides background information, the sampling/experimental design, sampling methods, and analysis methods. Use of these recommendations will help practitioners improve data management and thus improve the persistence, quality, and availability of data for decision-making and reuse by others.

A Graded and Iterative Approach to Planning: How Much Detail and When?

The level of detail required in a successful data-management plan depends on many factors, including the regional or national significance of the project, geographic scope, complexity, duration, number of participating organizations and individuals, funding source, project costs, intended use of the data, and the legal context of the project. As a result, we recommend a graded approach (USEPA 2002) to data management planning that is adaptable to the context of a project instead of a "one size fits all" approach.

Long-term ecological monitoring projects with significant implications for a federally listed species, regional economies, or human health require more detailed data planning than a local monitoring project assessing the restoration of a small preserve without any rare species. A regional monitoring program (e.g., Great Lakes Coastal Wetlands Monitoring Consortium 2008) requires detailed data-management planning to gather, store, process, and share data from multiple organizations and field crews across the region. Individual projects related to endangered species such as the Karner blue butterfly (*Lycaeides melissa samuelis*); Grundel et al. 1998, Forrester et al. 2005) also require additional data-management planning to ensure scientific and legal defensibility for regulatory decision-making. In contrast, a LTEM program to evaluate the effectiveness of invasive species removal at a site with no listed species could succeed with a simpler data-management planning effort.

The guidance provided in this paper is focused on a midlevel approach applicable to a broad range of projects. Although a more rigorous approach may be required for high-profile projects, all major components of a data management plan outlined here should be addressed at some level in most monitoring protocols. Thus, only the level of detail will differ among plans.

Data management planning is best done with an adaptive and iterative development model instead of one where all levels of detail are provided at once—similar to flexible development processes seen in other disciplines (Edmonds 1974, Smith 2007). Certain data management details must be developed prior to data collection—such as identifying data accuracy and precision requirements prior to acquiring instruments used to collect data, or creating specific steps for data quality assurance during data acquisition. However, not all details can be understood or finalized until later on—for example, specific data storage requirements, and specific destinations, mechanisms, and formats for distributing and sharing data. Long-term ecological monitoring projects necessarily involve a recurring cycle of data management activities; this cycle should include regular evaluation to maintain and improve practices (Fig. 1). Data management can then respond to the evolving needs of the project.

Components of Data Management for Long-Term Monitoring Projects

For planning and implementing a LTEM project, there are 2 key points about data management. First, successful data management is the responsibility of the lead ecologists, even if day-to-day execution is delegated to others. Second, data management considerations are woven into all phases of the project. Data management activities for a LTEM project can be idealized in a project implementation lifecycle (Fig. 1) that provides a basis for organizing and integrating data management principles into the planning and implementation of LTEM projects (Palmer 2003, National Park Service 2008, Fradkin and Boetsch 2012, Holdren 2013, Rüegg et al. 2014). We step through this sequence of activities over the course of this paper. Details about, and recommended practices for, each stage in the project implementation lifecycle can be found in Supporting Information Appendix 1.

Planning and Preparation

The initial planning stage of a project develops a well-defined scope and associated set of objectives, and begins the process of providing sufficient detail about the subsequent steps of project implementation (Fig. 1). Project objectives that are specific and measurable provide information needed to plan a monitoring effort beginning with a study design and a monitoring protocol. Other data-management planning and preparing activities include establishing data management infrastructure and processes, delineating roles and responsibilities, and providing appropriate training and workspace.

Study design.—Planning information is often encapsulated in a set of documents that comprise a monitoring protocol, which is an essential part of the legacy of a project and its resulting data. A well-developed protocol should address the following basic questions, many of which we elaborate upon further later in this paper:

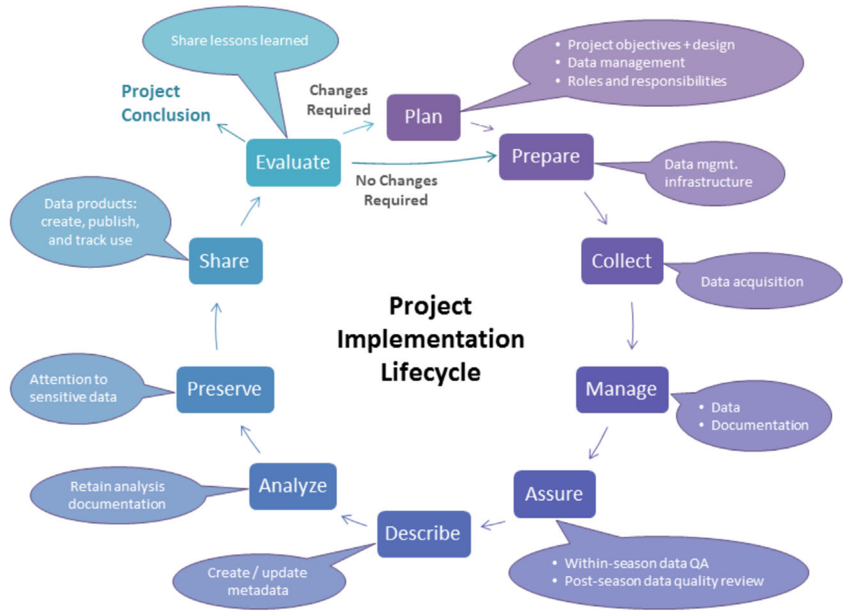

Figure 1. Stages of an ecological monitoring project. Stages that involve data management planning are shown in boxes, with the corresponding data management components in the bubbles. The stages are presented as a cycle to reflect the recurring nature of monitoring at intervals. Arrows represent decision points to implement or conclude the project, or to revise project plans based on feedback or internal evaluation.

1) Why?—Documentation should clearly lay out the relationship between monitoring objectives and the data being collected.
2) How?—Data collection procedures should be described in sufficient detail so that data can be collected in a consistent manner within and across seasons. Details should include the order in which data are collected, specific procedures for data collection, equipment use, and quality assurance routines built into the data collection.
3) Where?—The reference frame, sampling design, plot placement, coordinates, etc.
4) When?—The seasonality, time of day, years, and sampling conditions relevant to deciding when to sample should be clearly articulated.
5) What?—Detailed descriptions of the data being collected should include sampling event information (e.g., location information, date, observer names, parameters, data types, and domain values [e.g., pick list options]).
6) Who?—The roles and responsibilities for data collection should be clearly indicated, along with descriptions of any training or certification required by participants.

A key data-management consideration when creating a monitoring protocol is to promote consistency even if project staff members change. Documentation should be detailed and clear enough that a person not associated with the project could understand how well the data represent the population or phenomenon of interest. The reader should be able to properly interpret results with full knowledge of assumptions and limitations.

Although the measurements, units of measurement, and samples collected by a LTEM project are usually well-documented in project reports, all too often important aspects of data collection are not sufficiently described, such as specific and unique collection rules (e.g., gap rules for line intercept vegetation cover measurements), or how unidentified species and species groupings are treated and recorded. Documentation should include definitions of any terms that are not globally understood, and clearly define any acronyms, codes, and shorthand used in field forms or in project documents. Data quality requirements for measurements—also vitally important but only rarely documented sufficiently—should describe the precision or accuracy of individual measurements, why the described level of data quality is important to the project's success, and how well the methods or instruments used during data collection meet these requirements. This information can be effectively documented in a table with a separate column for each metric (e.g., Appendix 2 in the Supporting Information, from a project measuring occupancy of the Mojave desert tortoise [*Gopherus agassizii*; Sutter et al. 2012]). Another approach is to develop a project plan document based on a template that includes prompts to ensure that essential information about the project is being developed from the very start; we provide such a template in Appendix 3 in the Supporting Information. In the planning stage, design information describes the data that will be collected: What is their relevance? How will they be collected? What are feasible values? And so on. Once the project has actually collected some data, the design information becomes one part of what is referred to as "metadata" or "data about data." Metadata are sufficiently important to the usability of data from LTEM projects that we provide an in-depth discussion later in the paper.

Data management infrastructure and procedures.—Effective data management planning for monitoring projects requires putting sufficient thought and planning into the details about data acquisition (collecting, quality assurance, and associated documentation), data stewardship (storage structures, integrity protection, and backup), data preservation, and data sharing. Ideally, most details will be decided before any data collection begins; however, as noted above, it may not be possible to develop specific workflows and procedures until the data actually begin to flow in. Nevertheless, at least developing a comprehensive sketch of basic needs and workflows in advance of data acquisition is vital to ensuring adequate staffing, funding, and infrastructure to carry out the project and meet data management objectives (i.e., quality, completeness, availability, and usability for the long term). A properly designed file management system provides a balance between security and accessibility in support of project objectives. For established LTEM programs, there may already be guidelines, standards, and policies in place covering important topics such as data ownership and data quality. Existing outlets, systems, and procedures may also be in place for ensuring data security, for sharing data with others, and for developing and archiving data products.

Roles and responsibilities.—There are many roles and responsibilities associated with any LTEM project. Clearly identifying and communicating these is a crucial aspect of ensuring that key data-management tasks can be coordinated and carried out effectively and consistently. Typical roles include project lead, data collection lead, field technicians, data manager, Geographic Information System (GIS) specialist, data analyst, archivist, and records manager. The responsibilities distributed among the roles include leading or managing the project, collecting or compiling different types of data (field data, GIS, literature, etc.), managing and processing the data, analyzing the data, archiving the data, managing the data network, and communicating results to decision-makers. Assignment of individuals to each role differs among projects and organizations. For more extensive projects and in larger organizations, there may be specific teams assigned to each responsibility, but often just a few individuals manage these responsibilities.

To ensure project efficiency and data quality, each person involved in the project should understand their role in ensuring good data collection and management. The objectives of delineating these roles and responsibilities (National Park Service 2008) are as follows:

1) To establish data ownership throughout all phases of the project;
2) to instill data accountability; and
3) to ensure that adequate data quality and metadata are maintained on a continuous basis.

Regardless of where these roles and responsibilities are defined within project documentation, describing them provides a quick reference table for all project members. Such a summary may also assist future users of the data by identifying the individuals involved with each aspect of the project. Roles and responsibilities can be summarized with a simple list (Table 1), whereas more detailed information on qualification requirements and team reporting structure can be provided elsewhere in the plan or monitoring protocol. For larger projects, it may be useful to include an organizational chart or a table with responsibilities and a time schedule (for an example, see Supporting Information Appendix 4).

Training.—The level of training field crew members receive is an important predictor of data quality (Geissbuhler and Kuchler 2002, Ahrends et al. 2011). Classroom and field training are effective in explaining monitoring protocols and ensuring consistent methods for recording and managing data. Training should include the use of the intended equipment, data collection procedures, quality control and/or assurance methods, and crew safety. This approach will ensure that the flow of data acquisition and data entry is indeed efficient under field conditions. By establishing a standard set of training objectives and applying a variety of training modalities (e.g., hands-on, plus discussion and/or scenario-based, plus formal), the consistency of observations across years can be maintained and nonsampling error minimized. Testing at the end of training can be used as a readiness review to ensure that field crew members can conduct all procedures according to established protocols, are collecting the data within the quality requirements for precision and accuracy, can enter these data accurately and completely, and can upload the data to the project database. Training framed in the context of project objectives for data quality fosters awareness that data quality is largely determined during data acquisition (Michener et al. 1997).

Workspace and file management.—We recommend establishing a well-defined project workspace and a well-designed file management system. The workspace plus file-management system provides appropriate privileges for contributors to edit and delete files in a documented structure for organizing content. At present, there are many options available to accommodate project collaborators who may be able to share access to a network drive, including cloud-based solutions (e.g., Dropbox, Google Drive). The important considerations are the ability to retain an administrative history documenting the decisions made during the life of the project, in addition to the need to clearly mark and segregate files containing sensitive information from those with no distribution limitations.

In general, project files should be maintained in a manner that ensures the long-term integrity and usability of project information. As discussed earlier, a common theme for project file management is the need to clearly separate working materials (e.g., draft documents associated with the planning stage, or files accumulated during the current implementation year) from finalized products and files. Standardized file-naming conventions and file-folder naming structures can save time and confusion. File names should uniquely and concisely reflect file content (e.g., protocol documents, data exports, draft analytical products) and clearly distinguish file versions. Procedures and file structures should be put in place to store raw data in its original format prior to any manipulation or processing, preferably in a read-only state.

Regular backup and restore operations are a critical practice for minimizing the risk of information loss over time. Additionally, many agencies or corporations have record

Table 1. Roles and responsibilities for Intertidal Monitoring in the North Coast and Cascades Inventory and Monitoring Network (adopted from Fradkin and Boetsch 2012).

Role	Responsibilities
Project lead	Project oversight and administration.
	Track project budget, information requirements, and progress toward meeting project objectives.
	Facilitate communications between National Park Service and cooperators.
	Primary point of contact for data requests.
	Coordinate and ratify changes to protocol.
	Provide training to Field Lead, assist in conducting field work.
	Maintain and archive project records.
	Review and certify each season's data for quality and completeness.
	Complete and deliver reports, certified data, metadata, and other products as scheduled and according to Inventory and Monitoring (I&M) Program specifications.
Data analyst	Perform data summaries and analyses, interpretation, and report preparation.
Field lead	Train and ensure safety of field crew.
	Plan and execute field visits.
	Acquire and maintain field equipment.
	Oversee data collection and data entry.
	Conduct database audits to verify complete and accurate data transcription.
	Review and archive field forms.
	Complete a field season report.
Technicians	Conduct field work to collect field data, data photos, and sediment samples.
	Enter field data into the project database, and verify database records for accuracy.
	Download and import temperature data from data loggers to the project database.
	Download and process photographic images.
	Score rocky shoreline photo-plot images and enter scoring data into the project database.
	Process sediment samples for composition.
Data manager	Consult on data management activities.
	Facilitate check-in, review, and posting of data, metadata, reports, and other products to national databases and clearinghouses according to schedule.
	Maintain and update database application.
	Provide database training as needed.
GIS specialist	Consult on spatial data collection, GPS use, and spatial analysis techniques.
	Facilitate spatial data development, GPS data processing, and map output generation.
	Work with Project Lead and Data Analyst to analyze spatial data and develop metadata for spatial data products.
	Primary steward of GIS data and products.
Network program manager	Review annual reports for completeness and compliance with I&M standards and expectations.
Park curator	Receive and catalogue voucher specimens.
	Receive and archive copies of annual reports, analysis reports, and other publications.
	Facilitate archiving for other project records (e.g., original field forms, etc.).

retention policies that may put project records (including data) at risk of deletion; we recommend gaining a firm understanding of relevant retention policies and how they affect project records. As appropriate, ensure that the organization's policies clearly identify raw data, protocol documentation, analytical inputs and documentation, metadata, and derived data products as permanent records.

Data Collection and Management

There are several essential data-management practices that should be planned in advance and built into the data collection stage, including 1) developing data capture systems (e.g., digital or paper field forms) that include design elements such as data dictionaries, option pick lists, and other features that help to ensure complete and accurate data acquisition; 2) incorporating periodic evaluations into data collection activities to help ensure that data are being collected consistently and according to established procedures; 3) regular instrument calibration to minimize measurement errors; and 4) fostering a culture among field crews that encourages careful and complete data collection. The first and the last practices are discussed below, with the middle 2 discussed in later sections.

A key step in planning data collection is determining whether to use hand-written paper datasheets, digital data-collection field forms, or both. The decision involves balancing a number of considerations, remembering that human error is possible with any medium. Under certain circumstances, such as when the cost of having to repeat data acquisition is too high (e.g., high travel costs for sites that are difficult to access), where the phenomenon of interest is ephemeral (e.g., wildlife observations or desert annuals), or where the sites are altered by measurement sampling activities, field data acquisition using both digital field forms and paper datasheets may be the optimal approach.

Digital field forms and paper datasheets should be designed to mimic the logical flow of discovery and data collection. The flow of data entry can also be designed to minimize time handling organisms or minimize foot traffic within a sampling unit (Sutherland 2006). To help promote broader usability during data analysis, we recommend recording likely correlates that may affect the validity or applicability of the observed data.

In addition, to prompt field staff to be mindful of these factors, we recommend providing a separate comment box for each category of relevant correlates or other sampling events of interest. Both digital field forms and paper datasheets benefit from a pilot test and early peer review by a data manager.

As discussed below, paper datasheets are part of the permanent record for project data and should be handled in a way that preserves their future interpretability and information content. If changes to data recorded on paper datasheets need to be made, the original values should not be erased or rendered illegible. Instead, protocols should establish acceptable annotations such as: drawing a horizontal line through the original value, writing the new value adjacent to the original value with the date and initials of the person making the change, and making changes in a format such that it is easy for subsequent viewers to retrace the edit history.

Tabular, relational, and geospatial data.—Data collected in any format are likely to be transcribed, stored, and analyzed as digital tabular files. Best practices for data table design exist in the literature (Borer et al. 2009, Hook et al. 2010), and should be consulted early in project design. A fundamental recommendation from statisticians and data archivists is to organize data measurements to have only 1 measurement in each database field or spreadsheet cell. Both of the above papers provide other practical advice, such as ensuring that each table includes a header row of descriptive parameter names, and that each subsequent row contains data from one observation or measurement event. These practices greatly enhance data usability, and are most easily put in place before data collection.

We recommend that prior to data collection, a concerted effort be made to evaluate available technology and make informed choices about which software system to use, and on how the expected needs for data processing, summarization, and exporting will be met. Depending on the complexity of project data, relational databases, or NoSQL databases might be appropriate because they provide several advantages over spreadsheets. Kolb et al. (2013) provide a good overview of relational databases for ecologists. It may be necessary to recruit or hire help to ensure a robust solution that is truly able to support project requirements for data storage, processing, and summarization.

In many cases, it will be desirable or necessary to create an application interface to allow users to interact with different presentations of the data depending on the specific workflow tasks they are performing (e.g., data entry, quality review, data reduction, and summarization). Such interfaces are usually implemented as a separate "layer" that interacts with data in a more fundamental "data layer" (i.e., the data record themselves, which are usually contained in a separate database file or system). For example, a data entry screen might contain several forms and subforms that present specific data records from multiple related tables to the user (as opposed to entering records directly into tables in the absence of such an application interface). In such cases, we recommend that the flow and layout of such forms be developed to reflect the layout of any physical paper forms used to record data. Development of such data entry interfaces permits the incorporation of important data checks and constraints to facilitate data-quality assurance during data entry (e.g., built-in option pick lists, required fields, and user notifications in case of logical inconsistencies or missing or out-of-range values). The caveat for applications is that they usually require specialized programming skills to develop, maintain, and update; however, for many projects, the benefits of being able to efficiently access and manipulate the data throughout the project life cycle will more than justify the costs and risks associated with this decision.

An effective database solution will segregate working data that are actively being edited and processed (e.g., from the current year) from data collected in previous years that have already been reviewed and finalized. Finalized data should be protected by building safeguards into the database system that restrict edit privileges. In addition, depending on the staff size associated with a project, there may be a need to provide different levels of privileges to data (e.g., by allowing seasonal staff to add new records but not edit or delete existing records).

The increased role of spatial analysis in ecological research has led to a concomitant increase in the importance of placing ecological studies into a spatial context. Geospatial data require additional management considerations. First, these data can be of 2 basic types: vector (point, line, and polygon) and raster. The data type should be noted in the metadata, along with other vital information such as data projection, coordinate system, and error tolerances. Selecting appropriate Global Positioning System (GPS) units and postprocessing methods will help to ensure that the accuracy of the resulting data is documented and matches project objectives (e.g., sub-meter, 3–5-m, or >10-m accuracy). Prior to data collection, it will be important to establish standards for elevation masks, signal-to-noise ratios, minimum positional accuracy thresholds, and number of satellites. In addition, there may be associated attributes about each geospatial feature stored in related tables, including positional accuracy information. Depending on software format, these features and attributes may be stored in separate files that may be dependent upon stable folder substructure and file names for full functionality within the GIS software. Geographic Information System specialists should be consulted when file naming conventions and file storage structures are designed for these data formats. Additionally, when considering GIS or other software upgrades, special attention should be paid to keeping the format of older data up-to-date, for use by the project and future users.

Multimedia content.—Audio and image content often has substantial value to a LTEM project, whether as training material during the project or as documentation material after project conclusion. Storage for this content should take advantage of nonproprietary, commonly used formats (example, JPEG or TIFF). Images should be captured and stored at no less than a project-defined minimum file resolution (e.g., 300 dpi or 6 megapixel: National Archives and Records Administration [NARA], 2003) so that future use will not be constrained by insufficient resolution (e.g., publication-quality images to identify features or taxa within the image). The key consideration is that file resolution and format should both be conscious choices made in advance of acquisition, and that the decision should weigh storage space against quality and future use considerations.

If digital audio or visual files are processed, the steps should be well-documented to ensure proper use of the data in the future. Workflows should include naming conventions for folders where image data files will be stored. Adopting a file naming convention that concisely embeds metadata in file names can go a long way to preserve basic information about each image (Borer et al. 2009). For digital images, metadata regarding who, when, where, and how the image was created; what the image represents; and keywords to locate the image can also be included in the headers of the digital file or in an associated database. A useful reference for image data management is The Nature Conservancy's guide to photography (http://www.lllevin.com/users/LarryLevin3660/docs//photo_manual.pdf).

Voucher specimens.—Tissue samples, full specimens, water, soil, or other materials may be collected for off-site analysis or measurement, or as vouchers for proof of occurrence and confirmation of taxonomic identity (Wheeler 2003, Kageyama et al. 2006). Voucher specimens should be collected and preserved in a manner that supports long-term storage and identification. Each voucher should be uniquely identified with an accession number that can be linked to spatial and nonspatial data about the voucher. Location data associated with vouchers should also include information that will allow secondary users to properly gauge the precision of location data. Procedures should also make sure that specimen curation and delivery to an appropriate repository are part of a regular workflow; in some cases, agreements with such repositories may need to be worked out in advance.

Calibration and quality control for instruments and equipment.—A key component of data quality control is ensuring that instruments and equipment are operating as intended. Measurement accuracy is determined by the combination of equipment design and operator practices. Equipment should be calibrated before each new field season or as recommended by the manufacturer, or before each use for more sensitive equipment or sensitive data. Clocks and timestamps should also be synchronized on cameras, GPS receivers, and data loggers to assist with matching data records from each device to the same data collection event. Calibration events should be recorded for each instrument along with the unique identity of calibrated instruments used to collect each subset of data, to allow identification of data collected with a questionable instrument.

Assuring Data Quality

During data collection.—Quality assurance measures taken as data are being acquired have a large influence on overall data quality. Data should be reviewed as they are accumulated during a field season to minimize error propagation and data loss. Whenever errors or deficiencies are identified, steps should be taken to identify and mitigate their sources—whether through additional training, calibration, or by making changes in the equipment or procedures used to collect data.

Estimation of the quality of biological sampling efforts is often overlooked in ecological monitoring efforts (Blocksom and Flotemersch 2008). However, reliability and precision are important attributes of ecological monitoring methods (Bauer and Ralph 2001). Although not often included in monitoring protocols, remeasurement (duplicate or triplicate sampling) of plots (McCune et al. 1997, Barker et al. 2002, Brandon et al. 2003), of field samples (Haase et al. 2010), or of recorded sounds (Genet and Sargent 2003) could be used to improve data collection methods and training procedures, to identify potential data limitations, and to better analyze and interpret data. Remeasurement provides an opportunity to quantify and trace the source of measurement errors to causes such as instrument error, observer variability, or observer bias. Remeasurements should be conducted independently, preferably near in time to the original measurements so that temporal variation is minimized.

Debriefings should be conducted periodically (e.g., after each tour, weekly) and at the end of each field season to promptly identify issues that may have affected data quality. Debriefings may include the use of preplanned questions and allow for field crew comments. It is valuable to have field crew members rate their confidence in the quality of the measurements they conducted and to solicit ideas for improving the project, including logistics. As with training and audits, debriefing results should also be captured in project documentation.

Data quality review: verification, validation, and certification.—An important aspect of project quality assurance is to complete a thorough assessment of data quality at regular intervals. In addition to review as data are collected, a natural assessment time is after each season's data have been collected and entered, and before they are used for analysis or other purposes. The purposes of a data quality review are 1) to ensure that any problems and errors are known prior to using the data; 2) to fix as many errors as is feasible; and 3) to document the overall condition of the data set and any gaps or use limitations so that data can be confidently used now and into the future.

It seems intuitive to suggest that the best time to find, correct, and document data problems is soon after they are collected, and that this responsibility rests with the same individuals who generated the data. However, without incorporating this step into project workflows, with adequate resources, a thorough quality review is unlikely to happen—resulting in gaps and errors that remain unknown, unfixed, or undocumented. The risk is that subsequent users of the data either will not realize the limitations of the data and possibly draw incorrect conclusions, or that they will discover undocumented errors and lose confidence in the data so that the data will not be used to their full potential.

We recommend a data-quality review process that includes 3 essential steps:

1) Verification—Are all of the data there? Are they accurately represented?

2) Validation—Do the data make sense? What are the errors?
3) Certification—Is the data set ready for use and sharing? What is the overall quality?

Data verification evaluates newly acquired data for completeness, correctness, and conformance with acquisition specifications (e.g., min. accuracy thresholds). This may include tasks such as 1) visually reviewing data after they have been entered or uploaded to ensure correct correspondence between stored data and the original source; 2) ensuring that all data from the full set of sampling events are completely represented in the data set; and 3) ensuring that there are no duplicate records in the data set. Data verification may take place in the field before leaving a monitoring site, and more complex verification tasks can take advantage of access to the full data in the office; for example, comparing recent measurements against several years of data in the project database. It is particularly important to ensure that key fields such as date and location are recorded and transcribed correctly to maintain the structural integrity of the resulting data. Another aspect of data verification is to evaluate data against predetermined specifications for data quality (USEPA 2002). Finally, data verification may also be used to confirm that data processing steps such as transformations or other calculations have been implemented correctly.

Data validation evaluates the data for their content and reasonableness from a scientific standpoint. This can be accomplished through comparisons such as range checks to identify outlier values beyond a reasonable range, and checks for structural integrity and logical consistency. We recommend developing automated data validation procedures (e.g., database queries or scripts) to identify and flag outlier values, missing values, illogical data combinations, and potential inconsistencies in the data set in a repeatable fashion. Such automated quality checks are essential when working with large volumes of environmental sensor data (e.g., from temperature data loggers) that are frequently collected as correlates for ecological observations (Campbell et al. 2013). Data validation is also a good time to flag and annotate data that are not fully compliant with the data collection protocol. This supports the repeatability of analyses and helps users determine what subset of the data is relevant for their uses.

Data certification represents a benchmark indicating that the data are in a finalized state and can be used for analysis and shared with others. The role of certification is to establish a clear process, timeframe, and set of responsibilities for finalizing a data set. In the absence of such a benchmark, data sets are susceptible to being in left in a perpetual draft state, with no clear distinction between data that have been reviewed and data that still need additional quality assurance —thereby, complicating the creation of data set versions for analysis, publication, and sharing. Certification does not imply that the data are completely free of errors, because not all errors can be fixed. Rather, the certification step documents that the data are complete for the period of record, they have undergone and passed quality assurance checks, and any gaps or data of questionable quality have been identified. Certification is also an appropriate time to ensure that a knowledgeable project staff member prepares a description of the results of the data quality review. Upon completion of data certification, project metadata can be updated with the results of the quality review, and a read-only copy of the certified data set can be prepared for archiving and distribution. Any subsequent changes to certified data should be logged and recertified, and a new version of the data set issued.

Projects that carry high political risk may also be required to provide an audit trail of all changes to the data, including the nature of the edit as well as the time and identity of the user who edited the data. Monitoring projects may also require a more rigorous process for documenting the review and approval of the quality assurance steps. In such cases, consultation with a records management professional or archivist may be advisable.

Describing Data

Data documentation is critical for ensuring that data sets are usable for their intended purposes well into the future. Despite this importance, Hernandez et al. (2012) found that of 434 environmental or ecological science graduate students surveyed in California, USA, only 26% had ever created metadata to document their research data, and 28% reported they did not know what it meant to create metadata. Clearly written and comprehensive metadata are a key component for successfully managing active data over time and for reusing ecological data for a variety of purposes. Metadata are intended to describe what a data set user needs to understand to successfully use a data set—what the data include and represent, why they were collected, how they were collected, data-quality checks performed and results, idiosyncrasies of the data set, and who to contact if there are questions about the data. Metadata are particularly useful for long-term data series to help the original team remember what was done and why over time, and to provide crucial orientation information to new team members. In a general sense, metadata may include descriptive text, diagrams, flow charts, images, video, sound files, or other descriptive files—such items constitute informal metadata, and we have discussed their creation and maintenance above. A formal metadata record provides a structured, high-level overview of the data set, and often includes references to detailed information in other sources. A template for collecting project information for metadata is illustrated in Appendix 3 in the Supporting Information.

Approved standards for formal metadata currently rely on expressing metadata content in eXtensible Markup Language (XML) to facilitate computer searching and conversion of the document into multiple formats (e.g., HyperText Markup Language [HTML] or plain text). Because of the complexity of the standards, most metadata today are written using a specialized tool. The tool might be general purpose (e.g., an XML editor such as Altova's XMLSpy [Altova 2013]), or specific to a given standard (e.g., Metavist [Rugg 2004] for the Biological Data Profile or Morpho [Knowledge Network for Biocomplexity 2011] for Ecological Metadata Language [Fegraus et al. 2005]). There are a number of metadata standards relevant to ecology and environmental sciences (Table 2; see also http://www.dcc.ac.uk/resources/metadata-standards). The International Organization for Standardization family of standards is the most complex and flexible (Holdren 2013). Internationally, it is replacing the Consultative Committee for Space Data Systems standard, and it includes the Federal Geographic Data Committee (FGDC) Biological Data Profile (FGDC 1998, 1999).

Regardless of the standard used by a project, well-written metadata conform to the 4 Cs: 1) correct, 2) complete, 3) comprehensive, and 4) comprehensible. "Correct" means that the metadata content has no errors; "complete" means that all relevant metadata elements are present, even the optional ones; "comprehensive" means that the metadata content fully describes the data set, and "comprehensible" means that the metadata content can be understood by someone who is not an expert in the field that motivated collection of the data set. Each "C" builds on the previous C's to establish a higher quality bar. Comprehensibility is an aid not just to people outside the field, but also minimizes the chance that experts in the field will lose their ability to understand the metadata as jargon changes over time.

If the LTEM project team writes any formal metadata, it is common to write it just before sharing the data with other users. We recommend writing a version of the formal metadata when developing the study protocols, followed by refining and augmenting that version over the course of the project. This allows the team to capture idiosyncrasies of the data set as they happen, rather than having to recall these details later. It also allows the project to use questions asked by new team members as insight into what additional metadata content future team members and external data users will find useful.

Formal metadata also serves as an important "marketing tool" that enhances discovery, value, and usability. Metadata are often the source of data set information when performing searches for relevant data sets (Rüegg et al. 2014). This emphasizes the importance of complete and well-written metadata, because this makes it significantly easier to determine the suitability of a data set for a given use.

Analysis and Product Development

With the completion of quality assurance, data certification and updates to documentation, analysis and product development can proceed (Fig. 1). From a data management perspective, the primary objectives are to ensure the repeatability of analysis routines over time, and ensure that people not associated with the project are able to understand how analysis data sets and other derived products (e.g., maps, models) were developed. The team should establish procedures to retain copies of the data scripts and input data sets used for analysis, and to document the steps taken during analysis. This can be accomplished in several ways. Code scripts can be enhanced with descriptions of the steps taken by each block of code. An overarching document describing the overall flow of steps taken during analysis could be written, and would assist with production of metadata for any data products that are developed. As with data edits and annotations made during the data quality review, it is important to maintain a detailed log of any data modifications or manipulations made during analysis, and use a file naming and versioning system. This provides a record for the team, transparency to users of analysis products, and helps to meet potential legal or political requirements.

Preserving and Sharing Data and Data Products

Special attention should be given to ensuring that the essential products of monitoring projects—certified data, metadata, reports, and analysis products—are available and usable for the long term (i.e., on the scale of multiple decades or more) by archiving and/or publishing products to appropriate repositories. Projects and programs may develop a set of products for targeted distribution in order to reach specific audiences and stakeholders. Data management planning should also take into consideration how data requests will be handled and tracked.

Data archiving is distinct from backups of project files in that archiving generally occurs as data sets and the products derived from them are finalized—either on an annual or recurring basis for monitoring projects, or at the end of more short-term projects—whereas automated backups occur in the background throughout project implementation and cover all project files, including temporary and intermediate products. The archived form of the data should provide sufficient context and structure so others can use it after the project is over.

For archival purposes, data and metadata should be stored in lossless and nonproprietary formats to promote long-term preservation and usability (for example, TIFF not JPEG; XLSX or ODF not XLS; XML or CSV not MDB; see also Borer et al. 2009). It may also be necessary to establish a plan for migration to new software and hardware formats. If a nonproprietary format is not available, using a format that is in broad use is recommended. Archived data also should be periodically verified, to ensure that the data are retrievable and bit integrity is maintained. Archived products should be classified as permanent records in the organization's records management policy and records retention schedule. Select

Table 2. Metadata standards relevant to ecology and environmental sciences.

Standard name	Target data	References
Content Standard for Digital Geospatial Metadata (CSDGM)	Spatial data	Federal Geographic Data Committee (FGDC 1998)
CSDGM—Biological Data Profile (BDP)	Biological data with or without a spatial data component	FGDC Biological Data Working Group and U.S. Geological Survey Biological Resources Division (1999)
Ecological Metadata Language (EML)	Ecological data	Fegraus et al. (2005), Knowledge Network for Biocomplexity (2011)
ISO 19115-1	Revision to 19115	International Organization for Standardization (ISO 2014)
Darwin Core	Biodiversity data	Darwin Core Task Group (2009)
ISO 19115-2	ISO 19115 with extensions for imagery and gridded data	ISO (2009)

appropriate file formats, storage media, records-retention schedule updates, and an ultimate archive location; as appropriate, seek the advice of an archive specialist when making these decisions. There are a number of web sites with reliable information on archiving scientific data (see the California Digital Library's Data Management Plan Tool, https://dmp.cdlib.org/; Digital Curation Center http://www.dcc.ac.uk/resources/data-management-plans; and Whitlock 2011). Similarly, there may be organizations willing to host the data, and project managers can investigate services such as Databib (http://databib.org/index.php) to search for appropriate subject archives. Monitoring projects with a high level of political risk may also be required to provide proof of the authenticity (provenance) of digital or other data files in a court of law, in which case a records management professional or archivist should be consulted.

There are 2 broad classes of archives: "user maintained" and "actively curated." For user-maintained archives, the responsibility for updating all content (metadata, data, data formats, etc.) rests on an identified member of the project team, often the project leader. This class of archive frequently holds project content in the formats originally used by the project team, and there is minimal review of deposited content. This class has the benefit of being easy for the project team to use. However, over time there is an associated cost of declining usability as the project loses the team's attention and as members of the team leave the project. Websites maintained by project teams are an example of internal, user-maintained archives. Two examples of external, user-maintained archives are Knowledge Network for Biocomplexity (http://knb.ecoinformatics.org/index.jsp) and figshare (http://figshare.com/). Both examples also allow data self-publishing and support the use of digital object identifiers (DOIs) that are tied to a stable internet location for the data files for the self-published data sets.

Actively curated archives have a dedicated staff responsible for content updates, rather than the project team. This class of archive usually holds project content in the formats submitted by the project team, but also holds the content in archival formats that are maintained over time. These archives review metadata documents before final acceptance of the data package. Working with these archives takes more effort from the project team, but the final product is generally of higher quality and will be usable for a substantially longer time. In many respects, actively curated archives are analogous to scientific journals for data. Examples of actively curated archives include the Oak Ridge National Laboratory Distributed Active Archive Center (http://daac.ornl.gov/) and the Forest Service Research Data Archive (http://www.fs.usda.gov/rds/archive/). Both examples provide publishing capability and use DOIs. Actively curated archives usually design their systems to conform to the Open Archival Information System Reference Model (CCSDS 2012, ISO 2012).

Relational databases pose special archival challenges. They are often stored in proprietary software formats, even when the software is open source, and migration to newer versions of the software is an additional maintenance requirement (Christensen et al. 2011). More complex relational database structures may also warrant a more rigorous approach to test the fidelity and stability of those structures within archival storage systems. The data architecture (e.g., table structure, relationships among tables, and field names) appropriate for short-term analysis purposes may be different than that for long-term archiving and/or sharing. In addition, relationships among tables and indexes may be lost in file format migrations. Thus, it is important to include documentation of these aspects of data architecture with the data archive (Horsburgh et al. 2008). We recommend exporting archive copies of the data in a format that promotes long-term storage and usability (e.g., Unicode text format or XML), along with an image depicting relationships among tables and fields. For short-term purposes, data may also be stored and shared in a widely used format (e.g., SQL Server or MySQL), which may simplify reuse by other teams; although they are proprietary, these formats can often be read by other software, and free versions of those applications may be available for use with smaller data sets.

If digital images of paper datasheets or other physical materials are to be archived, in addition to following guidelines for image data management, we recommend careful consideration of minimum quality settings for any handwritten materials. Best practices for archiving video and sound data in physical media formats are well-established, and are nicely summarized in a reference guide provided by the Image Permanence Institute (Adelstein 2009). This guide contains references to a number of International Organization for Standardization standards for a variety of physical media types. Recommendations for long-term storage of digital video and sound recordings, however, are not yet well-developed within the archives community. Helpful guidelines for federal digital video records storage may be found on the NARA website: http://www.archives.gov/records-mgmt/policy/transfer-guidance.html.

When creating other derived data products such as reports, maps, and graphic representations, it may be necessary to establish a development and internal review process that takes into consideration target audiences and their needs, and to identify and/or develop standards for graphic identity, document format and content, and other usability requirements (e.g., compliance with accessibility requirements for the visually impaired).

Policies regarding legal confidentiality (e.g., for rare species, or personally identifying information), as well as any prepublication period during which data may not be shared, may constrain the manner in which data can be shared and published. Statements regarding data ownership and use limitations should be clearly recorded in the metadata associated with each data set. It is also important to understand the interests and the legal and/or regulatory framework of cooperating agencies, landowners, and other stakeholders. Sensitive information (e.g., the locations of protected resources such as caves, archaeological sites, or threatened or endangered species in U.S. national parks) should be clearly identified prior to product development and publication, and removed from distribution copies of data sets and derived products when disclosure would be in violation of law and/or agency requirements. We recommend creating workflows and procedures to review data sets and derived products for sensitive information prior to distribution; to flag and remove or reduce resolution of such information as needed to prevent disclosure; and to establish confidentiality agreements as needed to permit disclosure to cooperators and contractors.

Evaluation

Conducting regular postseason review meetings provides an opportunity for LTEM project staff to review, learn, and improve data quality, metadata quality, and other documentation—as well as other aspects of project operations (e.g., safety and logistics). Evaluations can be an important part of instilling a "learning organization" perspective into the implementation of a LTEM project. Postseason discussions should cover 1) review of data collection protocols for usability and repeatability; 2) feedback and suggested improvements to training materials; 3) effectiveness of quality assurance routines; 4) status and disposition of data and other files (e.g., images) accumulated during the most recent season; and 5) suggested changes to equipment, workflows, data processing routines, and infrastructure that may improve efficiency. The concept of evaluation can also be applied to other aspects of project implementation and may include, as assessments of the usability and availability of project information 1) the completeness of metadata records, and 2) the impact and value of project information for stakeholders and other intended audiences. The intensity of such evaluation approaches should depend on project objectives.

Key project documents such as protocols and standard operating procedures should be periodically updated to reflect changes in practice and to incorporate lessons learned. The change history of these documents should clearly track, through a document history log, information including who made changes, when changes were made, and the rationale behind the changes. It may also make sense to maintain separate documentation of key administrative decisions during a project.

DISCUSSION

Managers and researchers who implement long-term ecological research formally or informally implement many aspects of data management. In this paper, we present a comprehensive framework of data management components as they relate to the planning and implementation stages of ecological monitoring projects (Fig. 1). Basic recommendations for each step of the framework are found in the Supporting Information Appendix 1.

Although the management of data is an important component of all monitoring and research projects, well-designed data management is essential for long-term monitoring and research. Effective data management ensures data quality, completeness, repeatability of methods, and long-term availability and usability. Maintaining the managed data will ensure the integrity and usability of data if funding for the project is reduced or eliminated. It is valuable to proactively consider options if funding is reduced. Having complete and updated data may also help in maintaining or increasing funding for projects through its presentation and preliminary analyses. If a project is defunded, then the managed data will be available for future start-up and meta-analysis.

What is called for is a culture shift away from viewing data as a single-purpose, "consumable" item, toward that of developing a valuable, irreplaceable resource that may even increase in value over time. Such a shift will enhance the ability of scientists and managers to pursue integrated analyses, and thereby, advance our shared scientific understanding. Of course, the primary purpose of data management planning is to ensure good data-management practices in support of project objectives. However, effective data management can also help to build a larger body of reliable information about ecological systems for the benefit of all.

The essential and complex nature of data management requires a proactive, planned approach. Data management is as important to the success of a long-term ecological study as any other aspect (project objectives, sampling design, or data analysis). Data management should be integrated into the protocol of long-term monitoring projects, rather than being relegated to an appendix or a separate document. As much as possible, documentation should be completed prior to the onset of data collection (Van den Eynden et al. 2011). Additionally, adequate funding should be available for data management throughout the project life cycle, from planning, data collection, analysis, archiving, and publishing.

There is no "one-size-fits-all" template for incorporating data management into the diversity of LTEM projects. The detail and depth of data documentation is dependent on factors such as the regional or national significance of the project, geographic scope, complexity, duration, number of participating organizations and individuals, funding source, project costs, intended use of the data, and the legal context of the project. A graded approach ensures a sufficient level of data quality without expensive overkill, and an integrated approach increases the likelihood that data management protocols will be followed consistently.

We agree with Rüegg et al. (2014: page 25) that most environmental scientists lack expertise in data management tools and that "As a result, data management practices frequently become an (unfunded) afterthought rather than a carefully planned process that can improve complex science." Although it is true that large and complex monitoring projects may require a "skilled data professional" as part of the team, for many projects data management is the responsibility of individuals who have other roles. We support greater training of ecologists in data management to help meet the need to better integrate best practices into ecological monitoring.

Good data-management sections in protocols or separate data-management plans take time to produce and implement. Although developing guidance for data management requires resources, we firmly believe that the value of a plan and its implementation outweighs these costs, and will save time later on at the reporting and archiving stages. Ultimately, good data management will lead to better long-term data collection, quality, and persistence, and from there a greater ability to answer important ecological questions on how to conserve and manage natural resources.

ACKNOWLEDGMENTS

The authors gratefully acknowledge the data management community of the National Park Service Inventory and Monitoring Program as a source of some of the creative approaches and pragmatic, project-based perspectives reflected in this paper. We thank the staff members of the Clark County, Nevada, Desert Conservation Program for maintaining a focus on practical data-management solutions and embracing an iterative, adaptive approach to the design of ecological monitoring protocols and data management plans. C. Lampert and S. Southwick of the University Nevada Las Vegas Libraries' Digital Collections provided insightful comments on an early version of the metadata section. We thank R. Applegate, Associate Editor of the Wildlife Society Bulletin; C. Ribic, Editor in Chief; and 2 anonymous reviewers for their comments that greatly improved the paper. We especially thank the U.S. Forest Service Research Data Archive for assistance with publication expenses. Mention of trade names or commercial products does not constitute endorsement for use by the U.S. Government.

LITERATURE CITED

Adelstein, P. Z. 2009. IPI media storage quick reference. Second edition. Image Performance Institute, Rochester, New York, USA. https://www.imagepermanenceinstitute.org/webfm_send/301. Accessed 29 Nov 2013.

Ahrends, A., C. Rahbek, M. T. Bulling, N. D. Burgess, P. J. Platts, J. C. Lovett, V. W. Kindemba, N. Owen, A. N. Sallu, A. R. Marshall, B. E. Mhoro, E. Fanning, and R. Marchant. 2011. Conservation and the botanist effect. Biological Conservation 144:131–140.

Altova. 2013. XMLSpy. http://www.altova.com/xmlspy.html. Accessed 13 Mar 2013.

Barker, J. R., M. Bollman, P. L. Ringold, J. Sackinger, and S. P. Cline. 2002. Evaluation of metric precision for a riparian forest survey. Environmental Monitoring and Assessment 75:51–72.

Bauer, S. B., and S. C. Ralph. 2001. Strengthening the use of aquatic habitat indicators in clean water act programs. Fisheries 26:14–25.

Blocksom, K. A., and J. E. Flotemersch. 2008. Field and laboratory performance characteristics of a new protocol for sampling riverine macroinvertebrate assemblages. River Research and Applications 24:373–387.

Borer, E. T., E. W. Seabloom, M. B. Jones, and M. Schildhauer. 2009. Some simple guidelines for effective data management. Bulletin of the Ecological Society of America 90:205–214.

Brandon, A., G. Spyreas, B. Molano-Flores, C. Carroll, and J. Ellis. 2003. Can volunteers provide reliable data for forest vegetation surveys? Natural Areas Journal 23:254–261.

Campbell, J. L., L. E. Rustad, J. H. Porter, J. R. Taylor, E. W. Dereszynski, J. B. Shanley, C. Gries, D. L. Henshaw, M. E. Martin, W. M. Sheldon, and E. R. Boose. 2013. Quantity is nothing without quality: automated QA/QC for streamlining environmental sensor data. BioScience 63:574–585.

Christensen, S. W., C. C. Brandt, and M. K. McCracken. 2011. Importance of data management in a long-term biological monitoring program. Environmental Management 47:1112–1124.

Consultative Committee for Space Data Systems [CCSDS]. 2012. Reference model for an Open Archival Information System (OAIS), recommended practice. Issue 2. 650.0-M-2. http://public.ccsds.org/publications/archive/650 × 0b1s.pdf. Accessed 15 Apr 2015.

Darwin Core Task Group. 2009. Darwin core. http://www.tdwg.org/standards/450/. Accessed 13 Mar 2013.

Davies, G. M., J. D. Bakker, E. Dettweiler-Robinson, P. W. Dunwiddie, S. A. Hall, J. Downs, and J. Evans. 2012. Trajectories of change in sagebrush steppe vegetation communities in relation to multiple wildfires. Ecological Applications 22:1562–1577.

Dietrich, D., T. Adamus, A. Miner, and G. Steinhart. 2012. De-mystifying the data management requirements of research funders. Issues in Science and Technology Librarianship 70 (Summer). http://www.istl.org/12-summer/refereed1.html. Accessed 15 Apr 2015.

Dodds, W. K., C. T. Robinson, E. E. Gaiser, G. J. A. Hansen, H. Powell, J. M. Smith, N. B. Morse, S. L. Johnson, S. V. Gregory, T. Bell, T. K. Kratz, and W. H. McDowell. 2012. Surprises and insights from long-term aquatic data sets and experiments. Bioscience 62:709–721.

Edmonds, E. A. 1974. A process for the development of software for nontechnical users as an adaptive system. General Systems 19:215–218.

Fancy, S. G., and R. E. Bennetts. 2012. Institutionalizing an effective long-term monitoring program in the US National Park Service. Pages 481–497 *in* R. A. Gitzen, J. J. Millspaugh, A. B. Cooper, and D. S. Licht, editors. Design and analysis of long-term ecological monitoring studies. Cambridge University Press, Cambridge, Massachusetts, USA.

Federal Geographic Data Committee [FGDC]. 1998. Content standard for digital geospatial metadata. FGDC-STD- 001–1998. Federal Geographic Data Committee, Washington, D.C., USA.

Federal Geographic Data Committee Biological Data Working Group and U.S. Geological Survey Biological Resources Division [FGDC]. 1999. Content standard for digital geospatial metadata—biological data profile, FGDC-STD-001. 1–1999. Federal Geographic Data Committee, Washington, D.C., USA.

Fegraus, E. H., S. Andelman, M. B. Jones, and M. Schildhauer. 2005. Maximizing the value of ecological data with structured metadata: an introduction to Ecological Metadata Language (EML) and principles for metadata creation. Bulletin of the Ecological Society of America 86:158–168.

Forrester, J. A., D. J. Leopold, and S. D. Hafner. 2005. Maintaining critical habitat in a heavily managed landscape: effects of power line corridor management on Karner blue butterfly (*Lycaeides melissa samuelis*) habitat. Restoration Ecology 13:488–498.

Fradkin, S. C., and J. R. Boetsch. 2012. Intertidal monitoring protocol for the North Coast and Cascades Network. National Park Service, Natural Resource Report NPS/NCCN/NRR—2012/512, Fort Collins, Colorado, USA.

Geissbuhler, S., and M. Kuchler. 2002. Influence of the quality of vegetation data on the ecological description of sites. Botanica Helvetica 112:1–11.

Genet, K. S., and L. G. Sargent. 2003. Evaluation of methods and data quality from a volunteer-based amphibian call survey. Wildlife Society Bulletin 31:703–714.

Gitzen R. A., J. J. Millspaugh, A. B. Cooper, and D. S. Licht. 2012. Design and analysis of long-term ecological monitoring studies. Cambridge University Press, Cambridge, England, United Kingdom.

Great Lakes Coastal Wetlands Monitoring Consortium. 2008. Great Lakes coastal wetlands monitoring plan. http://64.9.200.103/glc/files/docs/Great-Lakes-Coastal-Wetlands-Monitoring-Plan-FINAL-March-2008.pdf. Accessed 15 Apr 2015.

Grundel, R., N. B. Pavlovic, and C. L. Sulzman. 1998. Habitat use by the endangered Karner blue butterfly in oak woodlands: the influence of canopy cover. Biological Conservation 85:47–53.

Haase, P., S. U. Pauls, K. Schindehutte, and A. Sundermann. 2010. First audit of macroinvertebrate samples from an EU water framework directive monitoring program: human error greatly lowers precision of assessment results. Journal of the North American Benthological Society 29: 1279–1291.

Hernandez, R. R., M. S. Mayernik, M. L. Murphy-Mariscal, and M. F. Allen. 2012. Advanced technologies and data management practices in environmental science: lessons from academia. BioScience 62:1067–1076.

Holdren, J. P. 2013. Increasing access to the results of Federally funded scientific research. Memorandum from the Executive Office of the President, Office of Science and Technology Policy. https://www.whitehouse.gov/sites/default/files/microsites/ostp/ostp_public_access_-memo_2013.pdf. Accessed 25 Feb 2013.

Hook, L. A., S. K. Santhana-Vannen, T. W. Beaty, R. B. Cook, and B. E. Wilson. 2010. Best practices for preparing environmental data sets to share and archive. Oak Ridge National Laboratory Distributed Active Archive Center, Oak Ridge, Tennessee, USA. http://dx.doi.org/10.3334/ORNL-DAAC/BestPractices-2010. Accessed 15 Apr 2015.

Horsburgh, J. S., D. G. Tarboton, D. R. Maidment, and I. Zaslavsky. 2008. A relational model for environmental and water resources data. Water Resources Research 44:5.

International Organization for Standardization [ISO]. 2009. Geographic information - metadata - part 2: extensions for imagery and gridded data. ISO 19115-2:2009. International Organization for Standardization, Geneva, Switzerland.

International Organization for Standardization [ISO]. 2012. Space data and information transfer systems—open archival information system (OAIS)—reference model. ISO 14721:2012. International Organization for Standardization, Geneva, Switzerland.

International Organization for Standardization [ISO]. 2014. Geographic information - metadata - part 1: fundamentals. ISO 19115-1:2014. International Organization for Standardization, Geneva, Switzerland.

Jones, S. 2009. A report on the range of policies required for and related to digital curation. Version 1.2. Mar 2009. Digital Curation Centre, Glasgow, Scotland.

Kageyama, M., R. Monk, R. Bradley, G. Edson, and R. Baker. 2006. The changing significance and definition of the biological voucher. Pages 259–266 *in* S. Williams and C. Hawks, editors. Museum studies: perspectives and innovations. Society for the Preservation of Natural History Collections, Washington, D.C., USA.

Knowledge Network for Biocomplexity. 2011. Morpho 1.10.0 user guide. Version 1.4. http://knb.ecoinformatics.org/software/morpho/MorphoUserGuide.pdf. Accessed 13 Mar 2013.

Kolb, T. L., E. A. Blukacz-Richards, A. M. Muir, R. M. Claramunt, M. A. Koops, W. W. Taylor, T. M. Sutton, M. T. Arts, and E. Bissel. 2013. How to manage data to enhance their potential for synthesis, preservation, sharing, and reuse—a Great Lakes case study. Fisheries 38:52–64.

Lindenmayer, D. B., and G. E. Likens. 2010. The science and application of ecological monitoring. Biological Conservation 143:1317–1328.

Lovett, G. M., D. A. Burns, C. T. Driscoll, J. C. Jenkins, M. J. Mitchell, L. Rustad, J. B. Shanley, G. E. Likens, and R. Haeuber 2007. Who needs environmental monitoring? Frontiers in Ecology and the Environment 5:253–260.

Martin, E., and G. Ballard. 2010. Data management best practices and standards for biodiversity data applicable to bird monitoring data. U.S. North American Bird Conservation Initiative Monitoring Subcommittee. http://www.nabci-us.org/aboutnabci/bestdatamanagementpractices.pdf. Accessed 15 Apr 2015.

McCune, B., J. P. Dey, J. E. Peck, D. Cassell, K. Heiman, S. Will Wolf, and P. N. Neitlich. 1997. Repeatability of community data: species richness versus gradient scores in large-scale lichen studies. Bryologist 100:40–46.

McEachern, K., and R. Sutter. 2010. Assessment of eleven years of rare plant monitoring from the San Diego Multiple Species Conservation Plan. U.S. Geological Survey-Western Ecological Research Center-Channel Islands Field Station Administrative Report 2010-01, Ventura, California, USA.

Michener, W. K., J. W. Brunt, J. J. Helly, T. B. Kirchner, and S. G. Stafford. 1997. Nongeospatial metadata for the ecological sciences. Ecological Applications 7:330–342.

National Archives and Records Administration [NARA]. 2003. Expanding acceptable transfer requirements: transfer instructions for permanent electronic records: digital photographic records. Issued Nov 12, 2003 updated in 2005. http://www.archives.gov/records-mgmt/initiatives/digital-photo-records.html. Accessed 29 Nov 2013.

National Park Service. 2008. Data management guidelines for inventory and monitoring networks. National Park Service Natural Resource Report NPS/NRPC/NRR-2008/035, Fort Collins, Colorado, USA.

Nussear, K. E., and C. R. Tracy. 2007. Can modeling improve estimation of desert tortoise population densities? Ecological Applications 17:579–586.

Palmer, C. J. 2003. Approaches to quality assurance and information management for regional ecological monitoring programs. Pages 211–225 *in* D. E. Busch and J. C. Trexler, editors. Monitoring ecosystems: interdisciplinary approaches for evaluating ecoregional initiatives. Island Press, Washington, D.C., USA.

Reichman, O. J., M. B. Jones, and M. P. Schildhauer. 2011. Challenges and opportunities of open data in ecology. Science 331:703–705.

Rüegg, J., C. Gries, B. Bond-Lamberty, G. J. Bowen, B. S. Felzer, N. E. McIntyre, P.A. Soranno, K. L. Vanderbilt, and K. C. Weathers. 2014. Completing the data life cycle: using information management in macrosystems ecology research. Frontiers in Ecology and the Environment 12:24–30.

Rugg, D. J. 2004. Creating FGDC and NBII metadata with Metavist 2005. U.S. Department of Agriculture, Forest Service, General Technical Report NC-255, North Central Research Station, St. Paul, Minnesota, USA.

Sergeant, C. J., B. J. Moynahan, and W. F. Johnson. 2012. Practical advice for implementing long-term ecosystem monitoring. Journal of Applied Ecology 49:969–973.

Smith, P. G. 2007. Flexible product development: building agility for changing markets. Jossey-Bass, San Francisco, California, USA.

Spotila, J. R., A. E. Dunham, A. J. Leslie, A. C. Steyermark, P. T. Plotkin, and F. V. Paladino. 1996. Worldwide population decline of *Dermochelys coriacea:* are leatherback turtles going extinct? Chelonian Conservation and Biology 2:209–222.

Stein, B. A., A. Staudt, M. S. Cross, N. S. Dubois, C. Enquist, R. Griffis, L. J. Hansen, J. J. Hllmann, J. J. Lawler, E. J. Nelson, and A. Paris. 2013. Preparing for and managing change: climate adaptation for biodiversity and ecosystems. Frontiers in Ecology and the Environment 11:502–510.

Sutherland, W. J., editor. 2006. Ecological census techniques: a handbook. Second edition. Cambridge University Press, Cambridge, Massachusetts, USA.

Sutter, R., L. Bice, L. Mata, M. Peters, and A. Barnett. 2012. Desert tortoise occupancy covariate monitoring protocol. Version 1.0. Unpublished Report to Clark County Desert Conservation Program, Project Number 2009-ECO-801A, Las Vegas, Nevada, USA.

U.S. Environmental Protection Agency [USEPA]. 2002. EPA requirements for quality management plans, EPA QA/R-2. EPA/240/B-01/002. U.S. Environmental Protection Agency, Washington, D.C., USA. http://www.epa.gov/QUALITY/qs-docs/r2-final.pdf. Accessed 15 Apr 2015.

Van den Eynden, V., L. Corti, M. Woolard, L. Bishop, and L. Horton. 2011. Managing and sharing data. Third edition. UK Data Archive, University of Essex, England, United Kingdom.

Vines, T. H., A. Y. Albert, R. L. Andrew, F. Débarre, D. G. Bock, M. T. Franklin, and D. J. Rennison. 2014. The availability of research data declines rapidly with article age. Current Biology 24:94–97.

Wheeler, T. A. 2003. The role of voucher specimens in validating faunistic and ecological research: a brief prepared by the Biological Survey of Canada (Terrestrial Arthropods). Biological Survey of Canada (Terrestrial Arthropods), Document Series no. 9, Ottawa, Ontario, Canada.

Whitlock, M. C. 2011. Data archiving in ecology and evolution: best practices. Trends in Ecology & Evolution 26:61–65.

Williams, B. K., and E. D. Brown. 2012. Adaptive management: the U.S. Department of the Interior applications guide. U.S. Department of the Interior, Adaptive Management Working Group, Washington, D.C., USA.

Associate Editor: Applegate.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of this article at the publisher's web-site.

Appendix 1. Stages of a project-implementation lifecycle applied to an ecological monitoring project.

Appendix 2. An example of using a table to list descriptors, definitions, collection procedures, and data verification and validation for a set of metrics for monitoring Mojave desert tortoise (Sutter et al. 2012).

Appendix 3. Example Resource Monitoring/Management Project Plan Outline.

Appendix 4. Yearly project tasks organized by project stage, from a National Park Service protocol for monitoring sandy and rocky intertidal communities in 3 parks in the Pacific Northwest (Fradkin and Boetsch 2012).