CONTRACTOR'S REPORT

NATIONAL INVASIVE SPECIES COUNCIL SECRETARIAT
U.S. DEPARTMENT OF THE INTERIOR · OFFICE OF THE SECRETARY
1849 C STREET NW · WASHINGTON, DC 20240

Information management relevant to invasive species early detection and rapid response programs

Rebekah D. Wallace,* Charles T. Bargeron IV,* David J. Moorhead,* Joseph H. LaForest*

* Center for Invasive Species and Ecosystem Health, University of Georgia, Tifton, GA

JUNE 2018

ABSTRACT

The *National Invasive Species Council Management Plan: 2016–2018* calls for an evaluation of current information systems relevant to early detection and rapid response programs. Information systems are important resources in invasive species programs for identification, management, modeling, education, and more. These systems both facilitate and are reliant upon information sharing and aggregation, which the federal government and many professional journals are requiring as part of funding and publication. There are many challenges associated with aggregating, validating, storing, querying, analyzing, and sharing information, as the amount of data and information increases nearly as fast as the advancements in computer hardware and processing power. In order to determine the types of data that are collected and shared, where and how those data are shared, which information systems are being accessed, and what information is being used from the information systems, a survey was sent out to representatives from federal agencies that are associated with invasive species programs. The respondents' results and recommendations are summarized. Most agencies collected data that are relevant to their programs, research, and management needs, and they share data either through their own portals or directly with those who request it. The respondents all recommended that support needs to increase for invasive species programs so that data collection can spread to more areas of interest and towards prevention and proactive programs. The authors have also proposed additional recommendations for information standardization, data collection minimums, metadata and provenance preservation and standardization, and features necessary for functional information systems.

Keywords: Early detection and rapid response (EDRR); invasive species; technology; information systems; data management; data standards; big data

INTRODUCTION TO INFORMATION SYSTEMS AND THEIR ROLES IN EDRR PROGRAMS

Information management systems are the end product of efforts to understand our world. These systems are places where documentation and measurements can be standardized, housed, and retrieved for later analysis and publication, as well as places where the publications, press releases, summaries, and anything relating to information can be stored and retrieved from or made freely available. As it relates specifically to invasive species, they are central storage locations for observational reports, management plans and actions, risk assessments, taxonomic information, species descriptions, and more. There are innumerable information systems available on the internet or housed internally and are run by individual programs or collaborative efforts that span from federal, state, private, and nonprofit. Identification guides and tools, taxonomic information, research on species and other programs, species of threat, etc., are sourced from information that is built up by many people and programs, which can be accessed from databases. Existing risk assessments and models to conduct new assessments are available from publications and through databases. Risk assessments help create the Watch Lists to identify potential threats of the U.S. Department of Agriculture (USDA) Forest Service plan. Response options can be evaluated from other programs successes and failures.

Information that is discovered and made available by one management or eradication program (e.g. through reports, research, popular press, etc.) can be used by other programs. One of the first large-scale, federal agency-led eradication efforts was the Witchweed Eradication Program (Westbrooks and Eplee 2011). Witchweed (*Striga asiatica*) was discovered in North Carolina in 1956, and as it is a parasitic plant that affects grasses, including several crops important to the U.S., the eradication program was established and funded under the Federal

Organic Act (1944) and the Federal Plant Pest Act (1957) in 1958 (Westbrooks and Eplee 2011). Witchweed has yet to be eradicated in the U.S., but the infested area has been reduced to 1,271.2 acres in 2015 from a high of 432,000 acres in 1970 across North and South Carolina (Lassiter 2015; Westbrooks and Eplee 2011). Although this program has been successful, it is also expensive, and therefore smaller scale eradication efforts and early detection and rapid response (EDRR) programs have become preferable to cover local infestations. A lot of experience and information was gained from this program, which has influenced subsequent EDRR programs. According to Westbrooks et al. (2014) one of the main lessons learned was that lack of including landowners and farmers, two of the primary stakeholders for this species, in active roles in the eradication program almost certainly extended the length of time the program has taken to achieve results. This same issue was noted in the goatsrue (*Galega officinalis*) infestation in Utah, the relevant stakeholders (landowners to the Bureau of Land Management, USDA Forest Service, and U.S. Fish and Wildlife) were not significantly involved in the eradication program and the program ran only from 1981 to 1996 and was ended due to lack of continued funding (Westbrooks et al. 2014).

Later, when *Caulerpa taxifolia* was discovered in southern California in 2000, a meeting of stakeholders occurred within seven days of identification, and control steps began within seventeen days (Anderson 2005). As that species was added to the Federal Noxious Weed List by the Aquatic Nuisance Species Task Force in 1999, one year prior to the detection, the proper agencies were aware of the potential impact of the species (Anderson 2005). The Federal Noxious Weed List contains terrestrial, parasitic, and aquatic plant species that meet the definition of a "quarantine pest." The International Plant Protection Convention (IPPC) and the North American Plant Protection Organization (NAPPO) define a quarantine pest as "a pest of potential economic importance to the area endangered thereby and not yet present there, or present but not widely distributed and being officially controlled." Once a plant has been determined to fit that definition, a risk assessment must be prepared according to the process set out by the United States Department of Agriculture's Animal and Plant Health Inspection Service (n.d.). The risk assessment includes: 1) Identification of the species, 2) Consequences associated with the introduction/spread, 3) Likelihood of introduction/spread, and 4) References (United States Department of Agriculture – Animal and Plant Health Inspection Service n.d.). The next steps include publishing the proposed name addition in the Federal Register, analyzing and responding to public comments, and, if deciding to continue with species addition, publishing the final rule in the Federal Register (United States Department of Agriculture – Animal and Plant Health Inspection Service n.d.). This process allows for a thorough review of the species and its potential impact if introduced to the United States.

In addition to being on the Federal Noxious Weed List, *Caulerpa taxifolia* and eight others within the *Caulerpa* genus are banned from California (Anderson 2005). By evaluating the issues that other regions were having with this invasive species, adding the species to the Federal Noxious Weed List, and rapid identification when the species was introduced, the program was able to respond quickly, involve the appropriate stakeholders, secure funding, and establish goals for local eradication of *Caulerpa taxifolia* (Anderson 2005). In 2006, *Caulerpa taxifolia* was declared eradicated from California (United States Department of the Interior and National Invasive Species Council Secretariat 2016).

While noxious weeds, introduced species, and "pests" have been legislated to varying degrees for decades, the first true models and frameworks for EDRR systems were proposed starting after the 1999 Executive Order on Invasive Species (Executive Order 13112 1999). In 2002, the Federal Interagency Committee for the Management of Noxious and Exotic Weeds (FICMNEW) published their "National Early Warning and Rapid Response System for Invasive Plants in the United States," which outlines the objectives for implementing a national early warning system for invasive plants:

- Organize committees and leadership for the system
- Develop a system to report suspected invasive species by the public and professionals
- Establish standards for protocols of early detection, reporting, specimen submission, identification, verification, and information databasing
- Foster interagency partnerships and collaborations
- Conduct species risk assessments
- Establish protocols for reporting and managing introduced/suspected regulatory species
- Build up regional capabilities and management protocols for rapid response
- Facilitate the sharing of information and reporting via the internet
- Monitor plants brought in via e-commerce
- Develop outreach and awareness strategies
- Work towards a global early warning system
- Secure funding

The early warning system was structured linearly, with the relevant agencies, planned programs, etc. involved with each step:

1. Early Detection
2. Identification/Vouchering
3. Rapid Assessment/Record Archival
4. Stakeholder Consultation and Program Planning
5. Rapid Response

This basic structure is replicated in most other EDRR programs of the time with variable details and steps for each section between the proposals. In 2003, the National Invasive Species Council (NISC) proposed the "General Guidelines for the Establishment and Evaluation of Invasive Species Early Detection and Rapid Response Systems," which follows:

1. Early Detection (Identification and Vouchering included)
2. Rapid Assessment
3. Rapid Response (Stakeholder and Program Planning included)

In 2003, FICMNEW published a larger document, "A National Early Detection and Rapid Response System for Invasive Plants in the United States: Conceptual Design" (Federal Interagency Committee for the Management of Noxious and Exotic Weeds 2003). This document maintained the same EDRR plan structure as the 2002 document but added more details to their objectives and more action items within each objective. In 2004, the U.S. Forest Service published "The Early Warning System for Forest Health Threats to the United States" that laid out their plan for an EDRR program for invasive species that could affect forests and states in the document that it was modeled on the FICMNEW plan. Their plan was organized around the following actions:

1. Identify Potential Threats
2. Detect Actual Threats
3. Assess Impacts
4. Respond

One of the most recent proposed frameworks for EDRR is in *Safeguarding America's Lands and Waters from Invasive Species: A National Framework for Early Detection and Rapid Response* (United States Department of the Interior 2016). This report highlights the role of preparedness, which it defines as "Having the knowledge, financial resources, tools, trained personnel, and coordination structures in place to streamline activities at each of stage in the EDRR process (United States Department of Interior 2016)." Preparedness goes beyond identifying potential threats, as is included in the 2003 FICMNEW plan, and is defined as a process that "establishes the plan, coordination networks, tools, training, and necessary resources for deployment of detection, rapid assessment, and rapid response actions" (United States Department of Interior 2016). This step is missing or simply incomplete in previous plans but is important in ensuring success of the overall program and goals.

In all of these programs and subsequent proposals, the information needed for each step must come from somewhere.

Deciding which species should be or are part of EDRR programs and how to monitor or survey for them comes through evaluation of potential harm and pathways of introduction (Table 1). Who should be involved in EDRR programs at different steps or in specific capacities is dependent on what the species is, where it is found, and the potential stakeholders affected (Table 1). Identifying the potential effects of the species is ascertained by determining the accuracy of the identification and the information that is known about the species (Table 1). Response protocols may exist for a given species if another program has encountered it previously (Table 1). However, this information is typically scattered across many databases. Some of the reasons for this are understandable, such as having databases dedicated solely to taxonomy, but others are due to the lack of a cohesive database for a topic or the inability to easily share information. Many data are sourced from work done under federal and state agencies or with funding from agencies, but until recently, there was not a primary directive or policy that mandated that information must be shared. This has fortunately been changing in the last several years to allow for more high quality data to be shared.

EDRR Action	Information Needed
Detection	What should be reported?
	How is it reported?
Responders Alerted	Who has jurisdiction?
	Is it new to the county, state, U.S.?
Rapid Assessment	Is the identification correct?
	Has a risk assessment been done?
Response	What management options are available?
	What was done and was it effective?

Table 1 Information Management for EDRR.

However, not all information systems are equal and there are many issues associated with the existing systems that are in use today. To discover how invasive species related information systems are viewed, structured, and accessed by various federal agencies, a call for data was distributed to federal agencies with a series of questions relating to how they collect, store, and use information relating to invasive species. While some responses were missing or incomplete, the replies received indicated that information systems are important but that the current state of many agencies' contribution to and implementation of information available could be improved. NISC has undertaken, as part of their *2016-2018 Management Plan*, an evaluation of the availability of information systems, how they are used in early detection and rapid response (EDRR) programs, what information and capacity gaps exist, and what needs to be done to build a national EDRR framework (National Invasive Species Council 2016).

Data within Information Systems

Information systems usually limit their contents to serve a purpose or goal. The Integrated Taxonomic Information System

(ITIS) is focused on providing accurate, current, and complete information related to the taxonomic history and classification of species worldwide. These data are made available to the public through their website and are downloadable so as to be easily integrated into other databases and information systems. Accurate taxonomic information is one vital piece of invasive species data management and it can be incorporated with data from other systems to provide a full picture of the invasive species profile. Species distribution is another crucial data point that can be accessed from multiple public, academic, or government databases. Global Biodiversity Information Facility (GBIF), United States Department of Agriculture - Plants Database, the Early Detection and Distribution Mapping System (EDDMaps), USGS Biodiversity Information Serving Our Nation (BISON), USGS - Nonindigenous Aquatic Species (USGS-NAS), Global Invasive Species Information Network (GISIN), and others are all resources that provide venues to view, query, and access species distribution data and often other information as well. Some programs have data primarily focused on a topic, region, or other parameter. Data within existing information systems can be used to create more data or knowledge.

Data and information are built upon and added to over time. As more information becomes available, it is published and distributed in a variety of ways, but it will eventually be included into an information database. One of the first tasks when addressing invasive species data is correct identification. Usually, the first option that individuals turn to are local guides and lists of invasive species. These can be very useful, as it will limit the options to the species that are known, and often common, to an area. If available as handouts or small publications, these are especially useful for volunteers and newly trained citizen scientists. Traditional dichotomous keys are also still widely used by experts and well-trained or educated identifiers and are published in many books and websites for broad, region-wide identification, such as the *Flora of the Southern and Mid-Atlantic States* (Weakley 2015), or for keying out a small set of species in a particular area, such as the "Florida Preying Mantids" (University of Florida Entomology and Nematology Department 2003). Another option is the multi-access key system, which uses a dynamic function that allows for more options on paths of identification for certain species groups. Lucid Keys are one system used by experts, which are developed and published for general use. Almost 300 multi-access keys are available through LucidCentral.org and the top five Lucid Keys have been downloaded over 1000 times each (LucidCentral 2017). However, individuals, citizen scientists, and even experts may not be able to identify a species on their own using keys. In those situations, specialist or additional expert help is available from herbaria, museums, and diagnostic labs.

Instructions for pressing plants, preserving insects, and photography for identification are available widely on the internet from many institutions (Burkett 2009; Lacey et al. 2001; Rawlins 2009; University of Florida Herbarium 2017). The internet has also made it easier than ever to find experts and specialists for various species and for finding herbaria and museums, all of which are options for aid in identification. For species that are not plant or animals, there are also regional diagnostic labs to help identify fungi and diseases, which are an area often lacking in publicity, but can be devastating to an ecosystem. Diagnostic labs are available to the public to help with identification of causal agents for damage to plants, which can include invasive diseases (Miller et al. 2009). Diseases like laurel wilt, which has killed almost half a billion red bay trees over 250,000 square kilometers (Hughes et al 2016), and chestnut blight, killed more than 4 billion trees over almost 81 million hectares (Burhans and Hebard 2012), have dramatically changed the landscape. However, herbaria and museums are, unfortunately, being consolidated and removed as a presence at many academic institutions and agencies, usually due to funding cuts and the advancements in DNA analysis for taxonomic classification (Deng 2015; Kemp 2015).

Until relatively recently, classification and identification was reliant upon visual and behavioral characteristics, however, new tools are now available to assist and fast identification is key to any early detection and rapid response program. With invasive species, some are so unique to a system that visual identification is sufficient, but there are other tools available for identification when there is not a high confidence in visual cues due to similarities or lack of a complete specimen. In these cases, DNA testing may be the best option, and this field has been developing in the last several years to serve as a quick identification option. DNA barcoding involves taking an unknown DNA and comparing it to known and unique sequences that have been categorized in a DNA barcode database (Cristescu 2014). This is limited to one sample, whereas DNA metabarcoding and environmental DNA (eDNA) metabarcoding will accept mixed samples (Cristescu 2014, Thomsen and Willerslev 2014). Metabarcoding involves a sample of mixed DNA, such as from an insect trap or stomach contents, and evaluating the samples using the high-throughput method for the species present (Cristescu 2014, Thomsen and Willerslev 2014). eDNA identification uses an extraction of the the DNA contained within a sample of environmental material (e.g. water sample, soil sample, etc.), amplifies it, and then uses metabarcoding to quickly identify the species represented in the sample (Thomsen and Willerslev 2014). Metabarcoding can be very useful for identifying new introductions and in monitoring, but it does not detect if the resulting species are alive and present in a propagating capacity, as it may contain DNA from dead samples, body fragments, etc. (Thomsen and Willerslev 2014). In addition, barcoding and metabarcoding are still reliant upon a genetic barcode having been described and catalogued adequately (Nilsson et al. 2006; Trebitz et al. 2015), and upon the inquirer possessing knowledge of the process, how it works, who to seek out for such a service, and the funding to pay for testing. As DNA barcoding databases are still being populated, image databases, physical specimens, species description databases, and expert assistance are still key to accurate identification.

As mentioned in the survey results, most of the data collected on invasive species are relevant to presence, size, location, etc. Risk assessments involve taking all of the known factors that influence invasion and comparing species characteristics to the resulting rubric (Diez et al. 2012). Native distribution, as well as invasive introduction pathways, can be used to fill in variables in risk assessments, horizon scanning, and other predictive models (Jiménez-Valverde et al. 2011). Ability to cross-reference the locations where an individual species is established against other data known about those locations (e.g. habitat, disturbance history, migratory patterns, soil characteristics, precipitation, etc.) can inform species invasiveness potential (Diez et al. 2012). Models such as Australia's Weed Risk Assessment (WRA), USDA-APHIS – Plant Protection and Quarantine Plant Risk Evaluation Tool, and the U.S. Fish and Wildlife Service's (USFWS) Ecological Risk Screening Summary will evaluate known data on assessments done on closely related species and the life histories, climate, physiology, reproductive and dispersal strategies, pathways of introduction, impact, etc., of the species of concern (Australia Department of Agriculture and Water Resources 2015; Koop et al. 2012, United States Fish & Wildlife Service 2016). Risk assessment models themselves, such as the WRA, are often useful or adaptable across regions, which means that not every region needs to conduct their own assessment (Gordon et al. 2008). In developing the WRA, the model was evaluated against weeds both known to be invasive and non-invasive in Australia and was shown to not accept (i.e. reject or re-evaluate) 84% of known serious and minor weed species. Only 7% of "reject" species were non-weeds (Pheloung et al. 1999). Evaluation across plant species have determined that while there are also no "universal traits" that can be applied to determine invasiveness, there are some correlations with performance related traits (e.g. physiology, growth rate, size, leaf area, etc.) (Diez et al. 2012; van Kleunen 2010). This underscores the importance of broad data evaluation and sharing to investigate potential invasive species. Risk assessments, as mentioned above, look at many variables, some of which fluctuate in importance and standing from model to model. There are efforts to aggregate models and assessments in databases, such as by the Invasive Species Centre (2017), to make them broadly available and accessible to any EDRR program. Risk assessments of potential or known invasive species are sought by organizations to gain perspective on species that should be prioritized for biosecurity, regulatory, and EDRR programs. The European and Mediterranean Plant Protection Organization (EPPO) maintains a list of pest species scores as graded by the Computer Assisted Pest Risk Analysis (CAPRA) program and uses the assessments to work with other regional plant protection organizations to recommend International Standards for Phytosanitary Measures (European and Mediterranean Plant Protection Organization 2017).

Why Federal Data are Being Shared

Sharing data in publications and though presentations and demonstrations has broadly advanced knowledge across all fields and as communication methods have drastically advanced, information can spread further and faster than ever before. Not only has this made sharing completed project results easier, real-time, long distance collaborations and raw data sharing are now possible. Researchers have readily made use of communication technologies which have facilitated collaboration, especially internationally, such as e-mail, shared electronic whiteboards, scheduling/calendar tools, videoconferencing, shared applications, electronic lab notebooks, and more (Sonnenwald 2007). Wagner et al. (2015) evaluated the growth of international collaboration and found that internationally authored papers on Web of Science went from 10% in 1990 to nearly 25% in 2011. International communication is important in EDRR programs, because as firsthand experience, models, and risk assessments are published in other countries, they can be adapted to other countries' programs and laws. Countries around the world are coming to the conclusion that sharing data openly aids all countries in quality, quantity, and reuse of data, and as such, the Open Data Charter was signed by G8 leaders promoting the concept and adoption of similar policies (United Kingdom Government Cabinet Office 2013).

Through several acts, laws, executive orders, memorandums, and other forms of official communication, the U.S. government has driven policy towards an open and transparent stance with regards to information and data (Burwell et al. 2013; Holdren 2013; Office of Science and Technology Policy 2014; Orszag 2009). As the internet has made sharing of information faster than ever, the policies and guidelines published have included accessibility through various websites and data portals, as well as ensuring that the data are machine-readable, available for free when possible and in a timely manner, and to continue to evaluate ways to encourage collaborative efforts (Burwell et al. 2013; Executive Order 13642; Interagency Working Group on Open Data Sharing Policy 2016; Johnson 2005; Office of Management and Budget 1998; Office of Management and Budget 2016; Orszag 2009). In addition to data accessibility, information quality has been addressed through multiple government publications, such as the Information Quality Act (IQA), which required that each agency create and maintain guidelines toward ensuring that published data meets certain standards of quality, objectivity, utility, and integrity (Information Quality Act 2001). The Office of Management and Budget (OMB) was tasked with establishing guidelines from which other agencies would build their guidelines. The OMB published their guidelines in 2001 and, in addition to defining terms used in the IQA, including "information," "quality," "objectivity," "utility," and "integrity," the guidelines instructed agencies to: 1) establish their own data standards as appropriate to the information they would disseminate, 2) develop a review process within the agency and for affected citizens, 3) establish an officer in charge of compliance to guidelines, 4) respond to complaints in a timely fashion, 5) and also report on the complaints and how they were resolved in a yearly report (Office of Management and Budget 2001). However, there was some criticism of the implications of the published guidelines and how they could be used to stymie new health, safety, and environmental standards due to the vagueness of the wording in the IQA (Copeland and Simpson, 2004; OMB Watch 2004). Analysis of the OMB's first and subsequent reports on the complaints are conflicted on the true number of correction requests that are directly attributable to the IQA guidelines (many other complaints received under the IQA were similarly received prior to the Act) depending on who is evaluating the reports by individual agencies (Office of Management and Budget, 2004; Copeland and Simpson, 2004; OMB Watch 2004; Office of Management and Budget 2005). These different acts, laws, regulations, etc., all mandate the sharing, accessibility, and quality of data, which directly impacts the most current research being made available to inform decisions in EDRR programs.

Challenges in Data Sharing

While it is required by government funded projects (Holdren 2013), researchers may be particularly recalcitrant to contribute data due to protection of the data collector's publishing rights, effort spent collecting data, privacy policies, data misuse potential, and more (Savage and Vickers 2009; Tenopir et al. 2011). Requiring or encouraging data to be broadly published is an increasingly common requirement of not just the government, but also from journals (PLOS Journals 2014; Nature Research 2016; Elsevier 2017) and organizations which encourage open access (Simmons College - School of Library and Information Sciences, 2017). Data sharing and contribution to data repositories or aggregators is an important part of overall data management. While there are many impediments to combining data into an aggregate database or repository, such as quality, standardization, documentation, ethical concerns, funding, poor planning, and technological barriers (Soranno et al. 2015, Van der Eynden et al. 2011), there are also many benefits to data sharing. Data sharing addresses the most important concern: preventing ultimate data loss/destruction. However, it can also encourage scientific debate, promote new data uses, discourages re-collection of data, encourage collaborations among disparate programs, stimulate inquiry into new observations, provide a wider audience for the data, and provide resources for training and education (Duke and Porter 2013, Van der Eynden et al. 2011; Tenopir et al. 2011). Vines et al. (2014) found that of 516 articles published between 1991 and 2001, only 121 data sets were extant and accessible, and found that as more time passed, data accessibility decreased by 17% per year. As authors move positions, they may become increasingly difficult to contact. Wren et al. (2006) found that corresponding authors of articles over 10 years old had an 84% email delivery error rate, which means that even if the data are accessible, the author may not be to the inquirer.

An issue with static materials is not only their "snapshot" nature, but that there is not consistent conformation to necessary information. For example, "Guidelines for Coordinated Management of Noxious Weeds: Development of Weed Management Areas" (United States Department of the Interior et al. n.d.) does not have a date of publication on the document nor on any websites that the document was included in. This can be an issue as the document contains references to programs, funding sources, links to websites, references to laws, contact information for many individuals, and more. Readers do not have a sense of if the programs described are still active, if the relationships described are current or historical, if the people are still employed in the same position, etc. but these materials will be perpetuated on websites and in informational databases for years. Regular audits and quality checks could aid in making sure that data presented is current. One of the easiest things to do is to make sure all resources contain a date of publication so that, even if the material is perpetuated, the public will be informed that it was published years or decades before, and the reader can decide if more current information is needed. Authorship is another important piece of information for published materials. The document "Florida Preying Mantids" does not list a proper author on the document, but rather states "this key was originally generated from an Insect Classification exercise and has since been tested and modified by several groups of students" under the title (University of Florida Entomology and Nematology Department 2003). As the document is housed in the University of Florida's Entomology and Nematology Department website, this article has had to cite that institution as the "author," which may not be accurate, but authorship cannot be determined otherwise. This can be an issue when reconciling the document in aggregate databases, who to contact regarding information within documents, and proper interpretation of the information. A similar issue is seen with resources that have no clear title, which are difficult to reference and include in databases. Many formally produced or peer reviewed documents do have these basic requirements, but many entirely digital resources may be lacking them.

Examples of static electronic resources include presentations, recorded webinars, videos, digital images, electronic documents, etc. Digital resources can be assigned a digital object identifier (DOI) which is intended to be a static location dedicated to that resource (International DOI Foundation 2017). However, these are not "free" resources, usually issued by a DOI Registration Agency, and the database is reliant upon the DOI issuer to maintain the links to prevent dead-ends (International

DOI Foundation 2017). These resources can be highly produced and edited or extremely casual, and the resources can often be shared, embedded on other websites or in other media, and/or downloaded which means that it is important to include enough information in the product itself to be easily assessed for quality and timeliness. A good example of including important information in static electronic media can be found in the "Understanding Forest Ecology: Fire, Water, and Bark Beetles" video on the U.S. Department of Agriculture's Forest Service YouTube channel. The video provides locations and dates for the information that they are presenting, which is important information, as knowledge and practices can vary in time and location (United States Department of Agriculture - Forest Service 2018). The presenters' information is also included in the video, which allows the viewer to look up the presenters to assess their qualification on the topic (United States Department of Agriculture - Forest Service 2018). Metadata for electronic resources are mainly associated with ensuring that the resources will be discovered during searches. There are many metadata schemas (e.g. Dublin Core, Schema.org, MARC, Open Graph, etc.) for information and their applications are primarily divided by the needs of the system/field or how the information is designed to be accessed or used. Data preservation, accessibility, visibility is the most important part of an overall information system.

Challenges with Big Databases

The first challenge with large databases is long-term support and maintenance. Any tool or resource that is created must be maintained and that will take dedicated long-term staff and funding. "What's Invasive" was a project that had a five-year funded plan, it was well marketed, and users uploaded thousands of invasive species observations (University of California Los Angeles 2010). However, the project did not have dedicated funding after the initial project timeline, and the University of Georgia's Center for Invasive Species and Ecosystem Health acquiesced to taking over the smartphone app and incorporating the data into the EDDMaps database. Without this transferal, users would have continued to upload data on a system that was not being maintained or monitored and any EDRR species reports would have been missed by relevant state and federal agency contacts. As of now, due to lack of funding, the websites and applications have not been updated or maintained beyond ensuring their continued existence, but any data that is contributed becomes part of the larger EDDMaps database and submitted to the verification network, which includes federal contacts (Wallace et al. 2017). Without continued funding, abandoned resources will not remain current or relevant for very long and may become a liability. Even with adequate funding, the quantity of data in these databases could cause challenges with information storage and retrieval.

There are many issues with regards to data and information aggregation/integration, storage, and recall. Databases increasingly have to handle more and more information, especially as research can now collect data with tools that will record data automatically and on a regular basis (National Research Council 2013a). This is seen not just in natural sciences experiments, with variables like temperature, air quality, etc. but also in many other fields, especially marketing, social sciences, and forensics (National Research Council 2013b). As computers and processing have advanced, data collection has increased in volume and complexity (National Research Council 2013c). Consider that initially, computers could only accommodate text and numbers, a relatively small amount of data especially by today's capacity. As technology advanced, and computer processing and memory increased, descriptions of species could then include photographs. The first photograph scan occurred in 1957, and the image was only 176 X 176 pixels (Newman 2007). Now, smartphones can take images at 12 megapixels (Apple 2017; Samsung 2017) and applications can contain text, images, videos, and can access a wireless or mobile network to look at a database's or website's real-time maps of invasive species locations, so that anyone can identify a species, see if it has been reported in their area, and report the occurrence while out in the field (Bugwood Apps 2017). In planetary science, the amount of data collected by robotic exploration in 2002-2012 was 100 times the amount collected in the previous forty years (National Research Council 2013c). While most researchers still maintain their own databases, the U.S. government issued the Cloud First Policy in 2011 to ensure that data and information was housed in an environment which was safe, efficient, and technologically innovative (Kundra 2011). While cloud servers are much safer from data loss than traditional methods, a downside to cloud servers, and many other databases that house very large amounts of data, is that transfer of the data is becoming increasingly difficult as the sizes move into the terabytes and petabytes (Curry 2011; National Research Council 2013c). As the hardware has advanced, there are also a number of software tools being developed to query and analyze data.

In many of these fields, the data are intended to be collected and analyzed quickly and in large amounts at a time. However, even in programs where smaller quantities of data are collected, they may not have the support to manage and share data (Crall et al. 2010). There is also the issue of databases aggregating data from disparate datasets, as data from different sources comes with their own standards, methodologies, etc. that makes analysis and sharing difficult (National Research Council 2013c). Currently, there is no singular set procedure or protocol with maintaining metadata or provenance with datasets, which can create issues as data are shared and reused (National Research Council 2013d). There is also the added issue that as fields evolve and new technology is presented and employed, data standards can change that would necessitate a change in databases and data management (National Research Council 2013d). However, even with advancements in data analysis, there is still a lack of effective data quality management tools, and so this task is still reliant upon a data manager (National Research Council 2013d; National Academies of Sciences, Engineering, and Medicine 2017b). Data managers will evaluate data for quality, formatting, and errors that could affect use, quality, and validity of the data (National Academies of Sciences, Engineering, and Medicine 2017b). Even with data managers and multiple people evaluating data, there are still errors that occur, which may take a long time to track down and may only become known with broad exposure when included in aggregate databases.

For example, there are currently no known valid reports of yellow starthistle (*Centaurea solstitialis*) in Minnesota. In 2010, data from the Biota of North America Program (BONAP) was incorporated into EDDMaps, which showed reports of yellow starthistle in 22 counties in Minnesota via two sources. Minnesota Department of Agriculture employees tried to track down these populations and, after research, no populations were discovered. While BONAP does not publish the source of their reports on their website, this information was provided to EDDMaps when it was shared and the source for most of this data was the journal *Weed Science*.

The authors of the *Weed Science* article sent out a nation-wide questionnaire and herbaria survey asking for known populations of yellow starthistle, and it was published stating that 22 counties in Minnesota had infestation sites (Maddox et al. 1985). Questions have arisen about this report, as no herbaria samples were reported and the three Minnesota respondents, Dr. Gerald B Ownbey, Dr. Oliver Strand, and Dr. Paul Monson, were highly unlikely to misidentify this plant due to their high level of qualifications; two were curators of herbaria and one a professor of weed science (Bell Museum of Natural History 2017; Holmstrand 2004; Maddox et al. 1985; Pioneer Press 2010; University of Minnesota 2016). Further, in a subsequent book by Ownbey, there is no mention of yellow starthistle, whereas other introduced species, including Russian thistle, were noted (Ownbey and Morley 1991). As misidentification is unlikely, it is likely that there was an error either in filling out the questionnaire or in transcribing the questionnaire into the database. With no other support for this data, no voucher specimens, no photographs, no current observations, lack of ability to access the original data, and no way to contact the original survey respondents (all now deceased) (Chandler 2011; Holmstrand 2004; Pioneer Press 2010; University of Minnesota 2016) these reports were removed from the EDDMaps database. The remaining report was deemed a likely historical waif, which never established, but the author was also unable to access his original records for a definitive answer and no subsequent reports have occurred (Chandler 2011). When databases share their data, it helps fill in gaps in other databases, but it also brings increased exposure to the data, often to people who have an investment in the topic. In this case, it has helped to bring attention to a potential data issue before it traveled to other databases via EDDMaps. For aggregate databases, maintaining portions of the data as original and retaining provenance information can reduce the number of questions as the data are shared and allow for tracking duplicates. Rapidly evaluating data for errors and ensuring maximal sharing lies in standardization.

Data Standardization

While having data standardized across disparate disciplines can be impractical and time/cost prohibitive, standardization even within a specific scientific field is still an issue, especially when comparing data collected for different purposes. A few of the more public data standards relative to documentation of invasive species occurrence data are the North American Invasive Species Management Association (NAISMA), Darwin Core, and GISIN, though every information system and database are customized to a degree for the user or target audience's needs. GBIF is primarily based on Darwin Core. EDDMaps was originally based on the North American Weed Management Association (NAWMA became NAISMA in 2012) standards. Data collected by researchers is usually formatted to standards set by a professional society or the leading journal. Data collected by people in "applied" settings is usually in a format that is easiest for field technicians, land managers, volunteers, etc. Across the nation, even something as basic as location data are not standardized. Previous mapping standards recommendations accommodated regional preferences for how location data are collected in the field: state and county, township/city/borough, Latitude and Longitude, UTM, Public Land Survey System (PLSS), Metes and Bounds, etc., are all common ways that location is recorded, and some may not be easily converted to or useable in other systems (North American Weed Management Association 2002). Cities may cross multiple counties, PLSS cannot be converted to a county-based or GPS point-based coordinate system, and, for example, even the vocabulary used can be imprecise; "townships" can be civil, survey (PLSS), or charter (Michigan only) (Charter Township Act 1947; United States Census Bureau 2017; United States Geological Survey 2017). This can lead to confusion if data are housed in a single system and it can be time consuming to convert data to conform to another standard.

Additional steps to contribute information to a database makes it less likely for individuals to participate without motivation, either internal (e.g., assisting in other research, broadening the knowledge base, etc.) or external (e.g., it is inherent to the position to contribute data, there is a benefit or tangible reward, etc.). Incorporating existing data into a different database can have similar issues, formatting can be troublesome especially if metadata is nonexistent. The best practice for data management is for the different agencies and initiatives to enact data management standards in budgets and contractual agreements and to have a data manager named (National Academies of Sciences, Engineering, and Medicine 2017a). There are many challenges associated with data aggregation, most of which cannot be solved solely by technology. Occurrence data can come from many different disciplines, which have varying goals and may not collect pertinent information associated with invasive species observations. This issue has been observed in ecological restoration data synthesis, monitoring data may come from projects focused on wetland creation, barrier islands, river diversions, fisheries management, and more and all of these projects will collect data that suits their needs, but may be difficult to put together for an overall picture of ecosystem restoration (National Academies of Sciences, Engineering, and Medicine 2017b). Data synthesis will often require a manager to oversee error checking, data cleaning, taxonomy checking, data formatting (bringing data to common units, data types, etc.), and more (National Academies of Sciences, Engineering, and Medicine 2017b).

METHODS: THE CALL FOR DATA

With a vast amount of information available in a variety of places, questions arose about how data were collected, accessed, and used by others, specifically within the federal government. In 2016, the National Invasive Species Council (NISC) Secretariat sent out a request for response to a series of questions to multiple federal agencies asking the following:

- Describe the information systems that your agency uses for early detection and rapid response EDRR activities (e.g. assessment of risks and horizon scanning; pathway analysis; specimen identification and species biology; mapping, monitoring, and inventory; rapid response planning and actions; information sharing internally and with other agencies and non-federal partners).
- Describe the types of data used in those information systems or otherwise collected by your agency.
- Detail the budgetary expenditures within your agency that support each of those systems from fiscal year 2014 through to the President's proposed budget for fiscal year 2017.
- Could these information systems and data contribute to a national alert system to report the introduction of new potentially invasive species? If so, how?
- What types of information are missing? What types of information are not currently available?
- Recommend measures and key priorities necessary to build institutional capacities to collect, store, and disseminate information and data relevant to EDRR activities, including for a national alert system.

While not all agencies responded in full, their feedback and additional research into publicly available data from these and other agencies provided a picture of how data are collected and used by various federal agencies. The responses have been summarized below to give an overview on how information is accessed, collected, and used.

RESULTS

"Describe the information systems that your agency uses"

Most of the responses listed the information systems used but did not specifically break out how the data were used and for what category of activity. Agencies such as U.S. Geological Survey (USGS) rely upon a wide variety of information systems, some of their own design and development and some from outside of their agency. The individual agencies may maintain some information systems that are very specific, such as USDA, which has several internal programs that cover pesticide data, crop safety, and animal health drugs, the USGS Asian Carp Telemetry database, the USGS-NAS database, and the National Park Service (NPS) uses the National Invasive Species Information Management System (NISIMS) that was based on the Bureau of Land Management's system of the same name. While others are more broad in their data, such as USGS BISON, Integrated Taxonomic Information System (ITIS), and Earth Resources Observation Systems program. Some agencies have very narrow focuses for their own databases, but may contribute in other ways, as USDA does by hosting the Invasive Species Information Center at their National Agricultural Library. Many agencies also use information systems and databases developed by other agencies and non-governmental or highly collaborative sources, such as the Center for Invasive Species and Ecosystem Health/EDDMaps, iNaturalist, GBIF, World Register of Marine Species, and Encyclopedia of Life.

"Describe the types of data used or collected"

Data collected is reliant upon the funding and purpose of each particular project. NPS has several programs that collect data on plant species within their lands and even each park may collect data specific to their needs. NPS does not have a central database that all collected data are submitted to, and so standardization of data and collection protocols are not uniform agency-wide. The most common data collected related to invasive species is species and location, though, as listed above, location is not standardized among different organizations. Many programs will also collect information about the specimen (e.g., size, treatment/control/management efforts, percentage land cover, phenology/morphology, observational counts, land usage, etc.)

"Budgetary expenditures FY2014 through FY2017"

The following numbers are based on information reported in the survey. USDA Agricultural Research Service allocates $170,000 towards the Invasive Species Information Center. NPS has allocated $14-34,000 towards the NISIMS database over the last four years and has also allocated $1.9 million towards the Heartland Inventory & Monitoring Network to support Inventory and Monitoring and the Exotic Plant Management Teams programs. USFWS does not track how much funding goes towards information systems specifically, but approximately 23% of the Fish and Aquatic Conservation Aquatic Invasive Species program's budget of $3.151 million over four years was allocated towards EDRR work. USGS is the only agency that reported the actual amounts spent on information systems. For 2015, which was the last year final expenditures were known at the time of the survey, USGS spent $671,000 on information systems, which were reported on the NISC crosscut budget and $533,000 on invasive species information systems, which were not reported in the NISC crosscut budget based on the information provided by USGS.

"Potential for Information systems and data to contribute to a national alert system"

NPS noted that their NISIMS database is currently not set up to function as an EDRR system, as data are not updated frequently enough and it does not have any alert system included, though certain parks are using EDDMaps and/or iNaturalist to serve this function within their properties. USGS and USFWS mention that the USGS NAS database has an alert system in their program that users can sign up for, though the NAS database is currently concerned primarily with aquatic species (Fuller et al. 2013). USDA Animal and Plant Health Inspection Service (APHIS) has an alert system for agricultural pests, though it is limited specifically to that scope currently.

"Information not currently available or collected"

The common theme among responding agencies was focused on acquiring resources needed to be able to collect additional data. NPS would find benefit in more reporting systems allowing for the recording of absence data (surveys of where species have been searched for and not found) for prediction modeling, risk assessments, and assessing early detection program effectiveness. NPS also noted that spatial data are limited on invasive animals, whereas this is not as pervasive an issue with plants and aquatic species, whose advancements are slower or restricted by habitat. The comments from USFWS focused on habitat-specific monitoring and impact of invasive species introduction, and mentioned that economic impact data are especially desirable.

"Recommendations by agencies"

Almost every agency responded with recommendations including increased funding to grow programs, collect additional data, add functionality to existing programs, and connect with other programs and their databases/information systems. The USFWS recommended growing the USGS NAS's Alert System, expanding USDA-APHIS identification capabilities, introducing species identification and other needs at ports through the Law Enforcement Management Information System, as well as providing overall support for data sharing and standardization. USGS has proposed increased funding and technical support, adding more functionality and resources to the NAS program, and the possibility of expanding the NAS database to be inclusive of more taxa (i.e. terrestrial plants, terrestrial animals, non-insect invertebrates, and non-bird vertebrates).

The survey results and additional topics were discussed at the EDRR Assessments Discussion to Support Project Contractors workshop in Washington, D.C. in 2017. This meeting was attended by and included presentations contributed by stakeholders, those involved in invasive species, and developers of models for incident response initiatives to help discuss plans for the best EDRR models. Based presentations and discussion during the workshop, some of the challenges of information systems include (Merony 2017; Smith 2017; Harris 2017; Keen 2017; Wheatley 2017; Martin 2017; Atsedu 2017; Burgiel 2017):

- Lack of a central location for data storage, even within agencies
- Consistency issues in self-reporting
- Lack of common data standards and methods
- Data storage and accessibility
- Data heterogeneity
- Subjects and species covered are uneven

Central Location for Data Storage

There is no singular supported information system, even within agencies. Most agencies maintain their own database or multiple databases for their research or other data collection and analysis needs, and will also use outside information systems for additional data, current taxonomy and known synonymy, risk assessments, management strategies, media (e.g. images, video, graphs, etc.), and other information. Additionally, there are also agencies that do not have their own centralized databases and data are fragmented by region or even by state/site. The NPS for their Exotic Plant Management Teams' data is contained in the NISIMS database, however the Inventory and Monitoring data is in its own database and most individual parks' data is housed on site, and the Forest Service's Forest Inventory and Analysis program data held in a database separate from other USFS occurrence data. This could cause issues with standardization, as it is unlikely that funding exists on a local basis for data management and dedicated data sharing. Creation of a centralized database or information system is difficult to conceive, as there are many political, technical, economic, legal, and social issues that can and do prevent freely shared information (National Research Council 2013c).

Issues in Sharing (Self-Reporting)

Without pressure, either intrinsic or extrinsic, to share data with others, it is difficult for data collectors to justify allocating the time and cost it can take to respond to data requests. They may feel that there is too much effort to find a place to share data or that it will take too much effort to do so (Kim and Zhang 2015; Soranno et al. 2015). Some intrinsic reasons that researchers will name as to why they share data include public benefit, transparency and re-use, accessibility, and preservation. Extrinsic reasons include that it is the general practice in their field; they want more visibility for their research; it is required by funding, institution, or journal; and a perceived career benefit (Ferguson 2014; Kim and Zhang 2015). Some reasons that data are not shared include intellectual property or confidentiality issues, fear of research being "scooped" or lost publication opportunities, that they will not be credited, misinterpretation or misuse of the data, or simply that they are not required to do so (Ferguson 2014; Kim and Zhang 2015; Soranno et al. 2015).

Standardization

The issues of data standards, methods, and storage and accessibility could be resolved by having a data manager within each agency or program who could set standards for data reporting, maintain an internal database and/or contribute to aggregate databases or data repositories, and maintain metadata associated with projects (National Academies of Sciences, Engineering, and Medicine 2017b). Data managers are not the sole solution though, as they may not have the ability to control or influence research, such a subjects researched or how the program sources their data. In the workshop and in the survey, it was discussed that most species are studied after their introduction has occurred or before a risk assessment has been conducted, so programs are often reactive rather than proactive. Data are also becoming more difficult to query and analyze as databases grow in volume and complexity. As larger quantities of information pass through a multi-parameter query or analysis, the more it costs in computer processing and it can be difficult if the data are sourced from many different projects.

Data Heterogeneity

Heterogeneous data, variability in data as a result of data collected from a variety of source causes issues with analysis, quality control, and overall confidence. Data collection practices depend on the program and the data collection purpose. Many programs rely on temporary/seasonal employees, interns, or volunteers for invasive species reporting and documentation, while others have robust research being conducted and data could be collected via scientists or professionals/employees that makes aggregation and analysis difficult (National Research Council 2013c). However, researchers often found that data collected by trained volunteers (i.e. citizen scientists) could be considered valuable (Brandon et al. 2003; Crall et al. 2015; Danielsen et al. 2005). However, even with professionals, data errors occur through transcription, identification, or technical issues (Goodwin et al. 2015; Olson et al. 2014; Whelan et al. 2017). Having quality assurance checks, proper training, and quality control data audits will reduce data errors (Crall et al. 2010; Johnson et al. 2011). However, aggregating, and further synthesizing or analyzing the data grow increasingly difficult as the volume of data grows (National Research Council 2013c; National Academies of Sciences, Engineering, and Medicine 2017b).

Lack of critical data

In the survey and the workshop, respondents mentioned a lack of data and information in a variety of areas. Comments indicated that is often due to a lack of support and funding for proactive analysis and research. According to a report by the Committee on the Scientific Basis for Predicting the Invasive Potential of Nonindigenous Plants and Plant Pests in the United States, while the APHIS port inspection program has stopped a number of introductions (53,000 in 1999 [USDA - Animal and Plant Health Inspection Service 2002a]), they do not record negative results in their Port Information Network database due to lack of resources (USDA - Animal and Plant Health Inspection Service 2002b). In fact, at the time, they had the capacity for inspecting less than 2% of the transportation that brings products and people into the U.S. (Committee on the Scientific Basis for Predicting the Invasive Potential of Nonindigenous Plants and Plant Pests in the United States 2002a). Compounding this, while federal laws have likely prevented plant pathogen introductions due to trade (Committee on the Scientific Basis for Predicting the Invasive Potential of Nonindigenous Plants and Plant Pests in the United States 2002a), they tend to lag behind research for covering the prevention of new species introductions, especially with concern to the pet trade and green industry. Due to the incompleteness of available information and slow reaction to novel invasive species by the federal government, and with concern for their own natural spaces, some individual states or municipalities have passed their own laws or regulations. One federal law, the Nonindigenous Aquatic Nuisance Prevention and Control Act of 1990, has directly addressed prevention, as well as control, of aquatic species and rallied together several agencies to participate in the effort. Aquatic species being the focus of a federal law is not surprising as they are increasingly prominent in the invasive species community, since they affect water quality, recreation, food supply, transport and shipping, etc. Most laws are similarly limiting, restricted to covering certain species or categories of invasive species, such as the Asian Carp Prevention and Control Act (2010), Noxious Weed Control and Eradication Act (2004), and the Public Lands Corps Healthy Forests Restoration Act (2005), which covers forest pests. However, there is a lack of legislation and research support for critical, but less used, habitat such as wetlands. Without research into potential problem species and critical resource and conservation areas, it is difficult to encourage legislation or policy that would prevent new introductions. However, as more information is shared across borders, examples can be taken from outside of our country.

DISCUSSION

Respondents to the survey indicated that they contribute to invasive species information sharing in different ways. USDA houses the Invasive Species Information Center at their National Agricultural Library. Several agencies contribute data to external databases and mention their use of other information systems, essentially, virtual collaboration across agencies that leverage off of shared resources and existing systems. Agencies, organizations, and working groups use various reporting and crowdsourced observational databases to be notified of new invasive species reports, range expansions, and published maps. In many of these databases a verification system is in place where experts are employed, volunteer their time, or are tasked as part of their own work responsibilities to ensure data quality (Dean, 2005; iNaturalist 2017; EDDMaps 2017). However, data are still disparate, and as data sharing and aggregation becomes more common, a lack of metadata, lack of provenance, and data duplication could become an increasing issue. This is potentially a larger issue if incorrect data are duplicated and, if provenance was not maintained, the source of the error cannot be tracked down. There is currently no standard in place to maintain data lineage or provenance as it is aggregated and shared. The first step to ensuring that data maintains its integrity is to encourage data standardization at the time of collection.

Data Standards and Recommendations

Follow the recommendations for Standards, Formats, and Protocols published in "Enabling Decisions that Make a Difference: Guidance for Improving Access to and Analysis of Species Information."

The guidance document lays forth recommendations for not only data standardization in established formats, but also file formats and protocols and provides steps for ensuring that data is broadly available (National Invasive Species Council Secretariat 2018). Specifically, for mapping, the document recommends that shared data is compatible with the NAISMA standards (National Invasive Species Council Secretariat 2018). Data standards for mapping invasive species is a topic that has been discussed for over a decade. Several standards have developed but most systems have required at the very least the who, what, when, and where:

- Reporter/Collector
- Species
- Date of observation or sampling
- Location

Even these few "required" fields may not always be documented, and when they are, they may not follow any recognized standard. In addition, this may not be enough information for certain species, and it gives very little information for EDRR needs. When documenting insects or diseases, the host may be important in positive identification of the suspected invasive species (North American Invasive Species Management Association 2014). As positive identification is among the first steps in EDRR programs, and a delay due to incomplete information could be costly. Location is often recorded in a format as is needed by the program. However, exact coordinates are vital in EDRR programs to target efforts and have been heavily recommended by NAISMA for documentation (NAISMA 2014). This requirement has become increasingly reasonable, as even the average person likely owns a standalone GPS unit or smartphone/tablet with GPS. While minimum data field collection may be all that can be expected of casual reporters, professionals, especially federal and state agency employees, could be tasked with collecting additional data to increase the overall usability. Many modelers and managers would find estimates or actual infestation sizes or counts of observed species, treatment plans, habitat, percent cover, etc., useful for assessing the scope of an infestation, and these could be factored into the initial EDRR plans. Requiring Universally Unique Identifiers (as a Record ID) for data or a DOI for other information would aid aggregate databases, such as USGS NAS, BISON, and EDDMaps, in tracking errors and preventing duplication (National Invasive Species Council Secretariat 2018). Many metadata standards have been discussed and adapted for individual programs, such as BISON being based on Darwin Core and EDDMaps being based on NAISMA, but none have been universally adopted as needs are diverse and compromise is usually necessary. Scientists will publish using international units, and applied professionals may use acres and square feet, or "forest" instead of a more specific habitat type. With the variety of sources to consider, maintaining the provenance of the data will aid in ensuring data quality and standards are adhered to as the data is shared, harvested, and reused.

Recommendations for Avoiding Abandoned Resources

Recommendations for retaining the following information from the original collected data are proposed:

1. Original reported species name
2. Reporter/Collector
3. Original record ID or collection number
4. Contact for the data (if the data was submitted by someone other than the owner)

Maintaining this information will allow data tracking across databases for corrections, updates, querying for duplicates, which has been an issue in the past, as mentioned in the *C. solstitialis* anecdote.

Error duplication in aggregate databases should be rare, but it can be hard to correct once it has been shared. With the potential for the replication of errors, not only should there be a way for data to be shared, there should be a venue for data corrections and retractions to be aggregated. In the same system that data or files could be uploaded for sharing and for aggregators to source information, data downloading can be tracked and emails can be sent out by either the uploader or other aggregate databases if some piece of the data are determined to be erroneous. This will aid in maintaining a better data quality for the data collectors and for subsequent users. The next issue comes from the databases themselves.

Recommendations for Avoiding Abandoned Resources

Dedicated funding to not only start a project, but continued funding and support for further development and maintenance.

A dedication by funders of projects and by staff working on the projects to seek natural evolution of the products is needed; both are key to staying relevant and finding and maintaining funding.

Abandoned databases and websites can cause issues in and of themselves, as they are still operational, searchable, etc., but contain information that is not current. There are many invasive species related materials, especially outreach material, which can quickly become outdated. Almost anything with data that are continuously collected is outdated as soon as it is printed, especially publications delineating the spread of a fast moving invasive species, which can become outdated within months or years. In 2006, *Agrilus planipennis* was present in five states and Ontario, as of 2017, it is present in 31 states and Ontario and Quebec (Emerald Ash Borer Information Network 2017). Information from 2006 would also be missing that *A. planipennis* has been discovered on white fringetree (*Chionanthus virginicus*) in addition to *Fraxinus spp.*, which was only discovered in 2015 (Cipollini 2015). Maintenance of information and resources within a website or database takes constant and regular funding to support a staff to add new and important information as it becomes available.

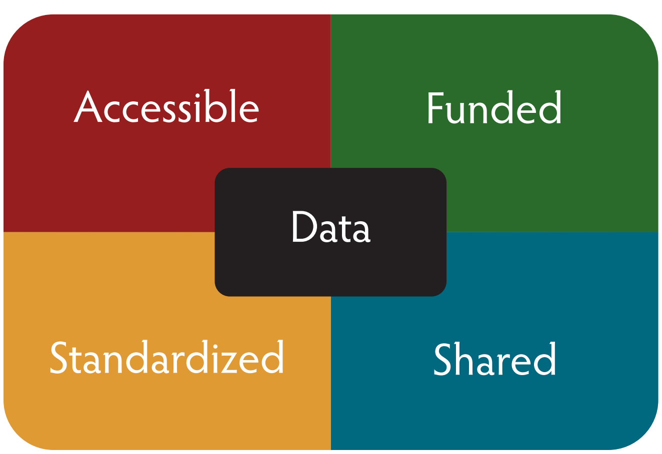

Figure 1 Information Management Recommendations.

Recommendations for All Printed and Static Electronic Resource Informational Materials

To house information in a standard format for ease of search and retrieval, the following minimum standards are proposed for new publications:

1. Date of Publication
2. Author (Person, Agency, or Organization that produced the materials)
3. Title

These recommendations are fairly minimal and common, as dates, authorship, and titles are often a requirement in data collection and publication, so it should be in all informational materials as well.

CONCLUSION

Information systems rely on data and the interpretation and summarization of data into information. The data/information recommendations by the authors and the agencies can be assigned under four general categories: Funded, Accessible, Shared, and Standardized (Fig 1).

Increased funding and critical projects, high priority species, imperiled habitats, and monitoring programs that report absence data were broadly recommended. All of the data and information will need to be housed in a database that is broadly accessible and able to be shared. The system will also need continuous funding and support to ensure continued growth, availability of tools, and novel tool development. The system must promote the use of standards for printed materials, data contributions, and provenance. The system must have a way to publicize data corrections and for those who requested or accessed the data to be notified of the corrections for their own records. To ensure that an information system is successful, feedback and endorsement from major data collectors and data users would aid in its publicity and use. All of these recommendations are pertinent to continued success of EDRR programs and other programs that rely on the information housed in these systems.

REFERENCES

Anderson L (2005) California's reaction to *Caulerpa taxifolia*: A model for invasive species Rapid response. Biol Invasions 7:1003–1016. doi:10.1007/s10530-004-3123-z

Apple (2017) iPhone 8 technical specifications. https://www.apple.com/iphone-8/specs. Accessed 31 October 2017

Asian Carp Prevention and Control Act (2010) Public Law 111-307. https://www.gpo.gov/fdsys/pkg/PLAW-111publ307/html/PLAW-111publ307.htm. Accessed 2 November 2017

Atsedu M (2017) Invasive Species Risk Analysis and Horizon Scanning Tools - Customs and Border Protection. In: Presentation at EDRR Assessments Discussion to Support Project Contractors, Washington D.C., March 2017

Australian Government - Department of Agriculture and Water Resource (2015) Weed risk assessment system. http://www.agriculture.gov.au/biosecurity/risk-analysis/weeds/system/weed_risk_assessment. Accessed 2 November 2017

Bell Museum of Natural History (2017) Minnesota Biodiversity Atlas. http://bellatlas.umn.edu. Accessed 4 November 2017

Brandon A, Spyreas G, Molano Flores B, Carroll C, Ellis J (2003) Can volunteers provide reliable data for forest vegetation surveys? Natural Areas Journal 23:254-262.

Bugwood.org (2012) Preservation and curation of insects. BugwoodWiki. https://wiki.bugwood.org/Preservation. Accessed 2 November 2017

Bugwood Apps (2017) Center for Invasive Species and Ecosystem Health Mobile Applications. https://apps.bugwood.org/apps. Accessed 2 November 2017

Burhans B and Hebard FV (2012) Restoring the American chestnut tree. USDA Forest Service Proceedings, RMRS-P-68. https://www.fs.fed.us/rm/pubs/rmrs_p068/rmrs_p068_024_025.pdf. Accessed 2 November 2017

Burgiel S (2017) Information Systems: Overview of Input from Federal Agencies. In: Presentation at EDRR Assessments Discussion to Support Project Contractors, Washington D.C., March 2017

Burkett N (2009) Macrophotography. BugwoodWiki. https://wiki.bugwood.org/Macrophotography. Accessed 2 November 2017

Burwell SM, VanRoekel S, Park T, Mancini DJ (2013) Office of Management and Budget - Open data policy - Managing information as an asset. M13-13. https://obamawhitehouse.archives.gov/sites/default/files/omb/memoranda/2013/m-13-13.pdf. Accessed 30 October 2017

Campbell EG, Bendavid E (2003) Data-sharing and data-withholding in genetics and the life sciences: Results of a national survey of technology transfer officers. J of Health Care Law Policy 6:241–255

Chandler M (2011) Subject: RE: where is yellow starthistle? E-mail correspondence

Cipollini D (2015) White fringetree as a novel larval host for emerald ash borer. J. Econ. Entomol. 108:370–375

Committee on the Scientific Basis for Predicting the Invasive Potential of Nonindigenous Plants and Plant Pests in the United States (2002a) Predicting invasions of nonindigenous plants and plant pests - Chapter 1: Introduction. http://www.nap.edu/catalog/10259.html. Accessed 2 November 2017

Committee on the Scientific Basis for Predicting the Invasive Potential of Nonindigenous Plants and Plant Pests in the United States (2002b) Predicting invasions of nonindigenous plants and plant pests - Chapter 2: Immigration: Predicting the Mode and Pathways of Introduction. http://www.nap.edu/catalog/10259.html. Accessed 2 November, 2017

Convention on Biological Diversity (2012) Operational plan for global invasive alien species information partnership. Conference of the parties to the Convention on Biological Diversity Item 13.9 of the Agenda. https://www.cbd.int/doc/meetings/cop/cop-11/information/cop-11-inf-34-en.pdf. Accessed 3 November 2017

Copeland CW, Simpson M (2004) The information quality act: OMB's Guidance and initial implementation. RL32532. Congressional Research Service - Library of Congress. https://fas.org/sgp/crs/RL32532.pdf. Accessed 22 May 2018

Crall AW, Jarnevich CS, Young NE, Panke BJ, Renz M, Stohlgren TJ (2015) Citizen science contributes to our knowledge of invasive plant species distributions. Biol Invasions 17:2415-2427. doi:10.1007/s10530-015-0885-4

Crall AW, Newman GJ, Jarnevich CS, Stohlgren TJ, Waller DM, Graham J (2010) Improving and integrating data on invasive species collected by citizen scientists. Biol Invasions 12:3419-3428. doi:10.1007/s10530-010-9740-9

Cristescu ME (2014) From barcoding single individuals to metabarcoding biological communities: towards an integrative approach to the study of global biodiversity. Trends in Ecology & Evolution 29:566-571. doi: 10.1016/j.tree.2014.08.001

Curry A (2011) Rescue of old data offers lesson for particle physicists. Science. 331:694-695.

Danielsen F, Burgess ND, Balmford A (2005) Monitoring matters: examining the potential of locally-based approaches. Biodiversity and Conservation 14: 2507-2542. doi:10.1007/s10531-005-8375-0

Dean J (2005) Mapping invasive species in NY http://adkinvasives.com/wp-content/uploads/2015/07/iMapInvasives-ADK-training-Backcountry-Aquatics.pdf. Accessed 2 November 2017

Deng B (2105) Plant collections get pruned back. Nature 523:16. doi:10.1038/523016a

Diez JM, Hulme PE, Duncan RP (2012) Using prior information to build probabilistic invasive species risk assessments. Biol Invasions 14:681–691. doi:10.1007/s10530-011-0109-5

Duke CS, Porter JH (2013) The ethics of data sharing and reuse in biology. BioScience 63:483-489

EDDMaps (2017) How to verify a record http://bugwoodcloud.org/CDN/eddmaps/tools/HowtoVerifyarecordinEDDMapsNew.pdf. Accessed 2 November 2017

Elsevier (2017) Research Data. https://www.elsevier.com/about/our-business/policies/research-data. Accessed 2 November 2017

Emerald Ash Borer Information Network (2017) Timelines. http://www.emeraldashborer.info/timeline.php. Accessed 4 November 2017

European and Mediterranean Plant Protection Organization (2017) EPPO and Pest Risk Analysis. https://www.eppo.int/QUARANTINE/Pest_Risk_Analysis/PRA_intro.htm. Accessed 1 November 2017

Executive Order 13112 (1999) Invasive species. FR64:6183-6186. https://www.gpo.gov/fdsys/pkg/FR-1999-02-08/pdf/99-3184.pdf. Accessed 3 November 2017

Executive Order 13642 (2013) Making open and machine readable the new default for government information. 3 C.F.R. https://obamawhitehouse.archives.gov/the-press-office/2013/05/09/executive-order-making-open-and-machine-readable-new-default-government-. Accessed 30 October 2017

Federal Interagency Committee for the Management of Noxious and Exotic Weeds (2002) National Early Warning and Rapid

Response System for Invasive Plants in the United States. https://www.invasive.org/edrr/EarlyWarningActionPlan031502-1.pdf. Accessed 3 November 2017

Federal Interagency Committee for the Management of Noxious and Exotic Weeds (2003) A National Early Detection and Rapid Response System for Invasive Plants in the United States: Conceptual Design. https://www.invasive.org/edrr/FICMNEW.pdf. Accessed 3 November 2017

Ferguson L (3 November 2014) How and why researchers share data (and why they don't). Wiley Exchanges https://hub.wiley.com/community/exchanges/discover/blog/2014/11/03/how-and-why-researchers-share-data-and-why-they-dont?referrer=-exchanges. Accessed 30 October 2017

Fuller P, Neilson M, Huge DH (2013) The NAS alert system: a look at the first eight years. Fisheries 38:128-138 doi:10.1080/03632415.2013.767241

Goodwin ZA, Harris DJ, Filer D, Wood JRI, Scotland RW (2015) Widespread mistaken identity in tropical plant collections. Current Biology 25:R1066-R1067

Gordon DR, Onderdonk DA, Fox AM, Stocker RK (2008) Consistent accuracy of the Australian weed risk assessment system across varied geographies. Diversity Distrib. 14:234–242 doi:10.1111/j.1472-4642.2007.00460.x

Harris C (2017) Wildlife Disease Surveillance and Response. In: Presentation at EDRR Assessments Discussion to Support Project Contractors, Washington D.C., March 2017

Holdren JP (2013) Office of Science and Technology Policy - Increasing Access to the Results of Federally Funded Scientific Research https://obamawhitehouse.archives.gov/sites/default/files/microsites/ostp/ostp_public_access_memo_2013.pdf. Accessed 30 October 2017

Holmstrand L (2004) In Memorium – Dr. Paul Monson. The Life Scientist - UMD Biology Department News. https://www.d.umn.edu/biology/news/documents/bio2004s.pdf. Accessed 4 November 2017

Hughes MA, Riggins JJ, Koch F, Cognato A, Anderson C, Dreaden TJ, Fomby JP, Ploetz RC, Smith JA (2016) The laurel wilt story: Introduction and impact of an exotic vector (Xyleborus glabratus) and pathogen (Raffaelea lauricola). Phytopathology 106S4:124

iNaturalist (2017) Help - Frequently Asked Questions https://www.inaturalist.org/pages/help#quality. Accessed 2 November 2017

Information Quality Act, Pub. L. No. 106-554, §515 (2001)

Interagency Working Group on Open Data Sharing Policy (2016) Principles for promoting access to federal government-supported scientific data and research findings through international scientific cooperation. https://obamawhitehouse.archives.gov/sites/default/files/microsites/ostp/NSTC/iwgodsp_principles_0.pdf. Accessed 15 January 2018

International DOI Foundation (2017) The DOI® System. https://www.doi.org. Accessed 30 March 2018

Invasive Species Centre (2017) Risk Assessments Database. http://www.invasivespeciescentre.ca/LEARN-ABOUT-INVASIVE-SPECIES/Risk-Assessments. Accessed 2 November, 2017

Jiménez-Valverde A, Peterson AT, Soberón J, Overton JM, Aragón, Lobo JM (2011) Use of niche models in invasive species risk assessments. Biol Invasions 13:2785–2797. doi:10.1007/s10530-011-9963-4

Johnson III C (2005) Office of Management and Budget - Improving public access to and dissemination of government information and using the federal enterprise architecture data reference model. M-06-02. https://georgewbush-whitehouse.archives.gov/omb/memoranda/fy2006/m06-02.pdf. Accessed 30 October 2017

Johnson CL, Temple GM, Pearsons TN, Webster TD (2011) An evaluation of data entry error and proofing methods for fisheries data. Transactions of the american fisheries society 138:593-601. doi:10.1577/T08-075.1

Keen A (2017) State Department Bureau of Intelligence and Research approach to disease incident response. In: Presentation at EDRR Assessments Discussion to Support Project Contractors, Washington D.C., March 2017

Kemp C (2015) Museums: The endangered dead. Nature 518:292–294. doi:10.1038/518292a

Kim Y, Zhang P (2015) Understanding data sharing behaviors of STEM researchers: The roles of attitudes, norms, and data repositories. Library & Information Science Research 37:189-200. doi:10.1016/j.lisr.2015.04.006

Koop AL, Fowler L, Newton LP, Caton BP (2012) Development and validation of a weed screening tool for the United States. Biol Invasions 14:273-294. doi:10.1371/journal.pone.0121053

Kundra V (2011) Federal cloud computing strategy. https://obamawhitehouse.archives.gov/sites/default/files/omb/assets/egov_docs/federal-cloud-computing-strategy.pdf. Accessed 15 January 2018

Lacey J, Short S, Mosley J (2001) How to collect, press and mount plants. Montguide MT 198359 AG. http://store.msuextension.org/publications/agandnaturalresources/mt198359ag.pdf. Accessed 2 November 2017

Lassiter B (2015) Plant Industry – Plant Protection Section – Witchweed Eradication Project http://www.ncagr.gov/PLANTINDUSTRY/plant/weed/witchnc.htm. Accessed 7 December 2017

LucidCentral (2017) Welcome to the Lucidcentral Key Search. http://www.lucidcentral.com/en-us/keys173;/searchforakey.aspx. Accessed 30 October 2017

Maddox DM, Mayfield A, Poritz NH (1985) Distribution of yellow starthistle (*Centaurea solstitialis*) and Russian knapweed (*Centaurea repens*). Weed Sci 33:315-327

Martin C (2017) Invasive Species Risk Analysis and Horizon Scanning Tools - United States Fish and Wildlife Service. In: Presentation at EDRR Assessments Discussion to Support Project Contractors, Washington D.C., March 2017

Merony K (2017) Wildfire Surveillance and Response. In: Presentation at EDRR Assessments Discussion to Support Project Contractors, Washington D.C., March 2017

Miller SA, Beed FD, Harmon CL (2009) Plant disease diagnostic capabilities and networks. Annu Rev Phytopathol 47:15-38. doi:10.1146/annurev-phyto-080508-081743

National Academies of Sciences, Engineering, and Medicine (2017a) Data stewardship. In: Effective Monitoring to Evaluate Ecological Restoration in the Gulf of Mexico. National Academies Press: Washington, D.C. doi:10.17226/23476

National Academies of Sciences, Engineering, and Medicine (2017b) Synthesis and integration In: Effective Monitoring to Evaluate Ecological Restoration in the Gulf of Mexico. National Academies Press: Washington, D.C. doi: 10.17226/23476

National Invasive Species Council (2003) General Guidelines for the Establishment and Evaluation of Invasive Species Early Detection and Rapid Response Systems. http://digitalcommons.unl.edu/cgi/viewcontent.cgi?article=1017&context=natlinvasive. Accessed 3 November 2017

National Invasive Species Council (2016) Management Plan: 2016-2018. Washington, DC. https://www.doi.gov/sites/doi.gov/files/uploads/2016-2018-nisc-management-plan.pdf. Accessed 15 January 2018

National Invasive Species Council Secretariat (2016) Data Call: Assessing risks: In support of 2016-2018 NISC Management Plan Action Item 5.1.3. Washington, DC.

National Invasive Species Council Secretariat (2018) Enabling decisions that make a difference: guidance for improving access to analysis of invasive species information. Version March 26, 2018. Washington, DC. https://www.doi.gov/sites/doi.gov/files/uploads/isim_guidance.pdf. Accessed 30 March 2018

National Research Council (2013a) Summary. In: Frontiers in massive data analysis. National Academies Press: Washington, D.C. https://www.nap.edu/read/18374/chapter/2. Accessed 30 October 2017

National Research Council (2013b) Introduction. In: Frontiers in massive data Analysis. National Academies Press: Washington, D.C. https://www.nap.edu/read/18374/chapter/3. Accessed 30 October 2017

National Research Council (2013c) Massive data in science, technology, commerce, national defense, telecommunications, and other endeavors. In: Frontiers in Massive Data Analysis. National Academies Press: Washington, D.C. https://www.nap.edu/read/18374/chapter/4. Accessed 30 October 2017

National Research Council (2013d) Scaling the infrastructure for data management. In: Frontiers in Massive Data Analysis. National Academies Press: Washington, D.C. https://www.nap.edu/read/18374/chapter/5. Accessed 30 October 2017

Nature Research (2016) Data availability statements and data citations policy: guidance for authors. https://www.nature.com/authors/policies/data/data-availability-statements-data-citations.pdf. Accessed 30 October 2017

Newman ME (2007) Fiftieth anniversary of first digital image marked. National Institute of Standards and Technology. https://www.nist.gov/news-events/news/2007/05/fiftieth-anniversary-first-digital-image-marked. Accessed 31 October 2017

Nilsson RH, Ryberg M, Kristiansson E, Abarenkov K, Larsson KH, Kõljalg U (2006) Taxonomic reliability of DNA sequences in public sequence databases: a fungal perspective. PLOS ONE 1:e59

Nonindigenous Aquatic Nuisance Prevention and Control Act of 1990, Pub. L. 101-646 (1990). https://www.anstaskforce.gov/Documents/nanpca90.pdf. Accessed 2 November 2017

North American Invasive Species Management Association (2014) North American Invasive Species Mapping Standards. http://www.naisma.org/images/NAISMA_Mapping_Standards2014_1.pdf. Accessed 4 November 2017

North American Weed Management Association (2002) North American invasive plant mapping standards. Endorsed by Federal Interagency Committee for the Management of Noxious and Exotic Weeds. http://www.naisma.org/images/Mapping_Standards.pdf. Accessed 2 November, 2017

Noxious Weed Control and Eradication Act, Pub. L. No. 108-488 (2004). https://www.gpo.gov/fdsys/pkg/PLAW-108publ412/pdf/PLAW-108publ412.pdf. Accessed 2 November, 2017

Office of Management and Budget (1998) Uniform administrative requirements for grants and agreements with institutions of higher education, hospitals, and other non-profit organizations. Circular No. A-110. https://www.whitehouse.gov/omb/fedreg_a-110-finalnotice. Accessed 30 October 2017

Office of Management and Budget (2001) Guidelines for ensuring and maximizing the quality, objectivity, utility, and integrity of information disseminated by federal agencies. https://obamawhitehouse.archives.gov/omb/fedreg_final_information_quality_guidelines. Accessed 30 October 2017

Office of Management and Budget (2004) Information quality: a report to congress. Fiscal year 2003. https://obamawhitehouse.archives.gov/sites/default/files/omb/assets/omb/inforeg/fy03_info_quality_rpt.pdf. Accessed 8 October 2017

Office of Management and Budget (2005) Validating regulatory analysis: 2005 Report to congress on the costs and benefits of federal regulations and unfunded mandates on state, local, and tribal entities. https://georgewbush-whitehouse.archives.gov/omb/inforeg/2005_cb/final_2005_cb_report.pdf. Accessed 8 October 2017

Office of Management and Budget (2016) Managing information as a strategic resource. Circular No. A-130. https://obamawhitehouse.archives.gov/sites/default/files/omb/assets/OMB/circulars/a130/a130revised.pdf. Accessed 30 October 2017

Office of Science and Technology Policy (2014) U. S. open data action plan. https://obamawhitehouse.archives.gov/sites/default/files/microsites/ostp/us_open_data_action_plan.pdf. Accessed 15 January 2018

Olson DD, Bissonette JA, Cramer PC, Green AD, Davis ST, Jackson PJ, Coster DC (2014) Monitoring wildlife-vehicle collisions in the information age: How smartphones can improve data collection. PLOS ONE 9:e98613. doi:10.1371/journal.pone.0098613

OMB Watch (2004) The reality of data quality act's first year: A correction of OMB's report to congress. http://webarchive.loc.gov/all/20041030034947/http://www.ombwatch.org/info/dataqualityreport.pdf. Accessed 30 October 2017

Orszag P (2009) Office of Management and Budget - Open government directive. M-10-06. https://obamawhitehouse.archives.gov/open/documents/open-government-directive. Accessed 30 October 2017

Ownbey GB and Morley T (1991) Vascular plants of Minnesota: a checklist and atlas. Minneapolis: University of Minnesota Press

Pheloung PC, Williams PA, Halloy SR (1999) A weed risk assessment model for use as a biosecurity tool evaluating plant introductions. J of Environ Manag 57:239-251

Pioneer Press (2010) Gerald B. Ownbey. http://www.legacy.com/obituaries/twincities/obituary.aspx?n=gerald-b-ownbey&pid=143832881. Accessed 4 November 2017

PLOS Journals (2014) Data availability. http://journals.plos.org/plosone/s/data-availability. Accessed 30 October 2017

Public Land Corps Healthy Forests Restoration Act of 2005 (2005) Public Law 109-154 (Dec 30, 2005) https://www.gpo.gov/fdsys/pkg/PLAW-109publ154/pdf/PLAW-109publ154.pdf. Accessed 2 November 2017

Rawlins KA (2009) Module: herbarium specimens. The Center for Invasive Species and Ecosystem Health. BW–2009–104. http://www.eddmaps.org/training/collecting.pdf. Accessed 12 December 2017

Samsung (2017) Galaxy S8 specifications. http://www.samsung.com/global/galaxy/galaxy-s8/specs. Accessed 31 October 2017

Savage CJ, Vickers AJ (2009) Empirical study of data sharing by authors publishing in PLOS journals. PLOS ONE 4: e7078. doi: 10.1371/journal.pone.0007078

Simmons College - School of Library and Information Sciences (2017) Open Access Directory. http://oad.simmons.edu/oadwiki/Main_Page Accessed 30 October 2017

Smith A (2017) Natural Hazards Surveillance and Response. In: Presentation at EDRR Assessments Discussion to Support Project Contractors, Washington D.C., March 2017

Sonnenwald D (2007) Scientific collaboration. Ann Rev of Information Sci and Tech 41:643-681 doi:10.1002/aris.2007.1440410121

Soranno PA, Cheruvelil KS, Elliott KC, Montgomery GM (2015) It's good to share: Why environmental scientists' ethics are out of date. BioScience 65:69-73. doi: 10.1093/biosci/biu169

Tenopir C, Allard S, Douglass K, Aydinoglu AU, Wu L, Read E, Manoff M, Frame M (2011) Data sharing by scientists: practices and perceptions. PLOS ONE 6(6): e21101.doi:10.1371/journal.pone.0021101

The Charter Township Act (1947) Legislative Council, State of Michigan Act 359. http://www.legislature.mi.gov/(S(vornbpzzubxbunfzqno1lx1b))/printDocument.aspx?objName=mcl-Act-359-of-1947&version=txt. Accessed 2 November 2017

Thomsen PF, Willerslev E (2014) Environmental DNA - An emerging tool in conservation for monitoring past and present biodiversity. Biological Conservation 183:4-18. doi: 10.1016/j.biocon.2014.11.019

Trebitz AS, Hoffman JC, Grant GW, Billehus TM, Pilgrim EM (2015) Potential for DNA-based identification of Great Lakes fauna: match and mismatch between taxa inventories and DNA barcode libraries. Sci Reports 5:12162. doi:10.1038/srep12162

United Kingdom Government Cabinet Office (2013) Open data charter. https://www.gov.uk/government/publications/open-data-charter/g8-open-data-charter-and-technical-annex. Accessed 15 January 2018

United States Census Bureau (2017) Population of interest: municipalities and townships. https://www.census.gov/govs/go/municipal_township_govs.html. Accessed 2 November 2017

United States Department of Agriculture – Animal and Plant Health Inspection Service (No Date) Guide to the listing process for federal noxious weeds. https://www.aphis.usda.gov/plant_health/plant_pest_info/weeds/downloads/listingguide.pdf. Accessed 30 March 2018

United States Department of Agriculture - Forest Service (2018) Understanding forest ecology: fire, water, and bark beetles. YouTube Video File. https://www.youtube.com/watch?v=H8vkzRoxdrU. Accessed 30 March 2018

United States Department of Interior, Bureau of Land Management, National Park Service, and U.S. Forest Service (No Date) Guidelines for coordinated management of noxious weeds: Development of weed management areas. https://library.ndsu.edu/ir/bitstream/handle/10365/3267/1966gov.pdf?sequence=1&isAllowed=y. Accessed 14 January 2018

United States Department of the Interior and National Invasive Species Council Secretariat (2016) Invasive Species Early Detection and Rapid Response: Resource Guide. https://www.doi.gov/sites/doi.gov/files/uploads/EDRR%20Resource%20Guide_02_17_2016_Final.pdf. Accessed 30 March 2018

United States 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. https://www.doi.gov/sites/doi.gov/files/uploads/national_edrr_framework_2016.pdf. Accessed 30 March 2018

United States Fish and Wildlife Service (2016) Standard operating procedures for the rapid screening of species' risk of establishment and impact in the United States. https://www.fws.gov/injuriouswildlife/pdf_files/ERSS-SOP-Final-Version.pdf. Accessed 2 November 2017

United States Forest Service (2004) The early warning system for forest health threats to the United States. https://www.fs.fed.us/foresthealth/publications/EWS_final_draft.pdf. Accessed 3 November 2017

United States Geological Survey (2017) The Public Land Survey System (PLSS) https://nationalmap.gov/small_scale/a_plss.html. Accessed 2 November 2017

Van Kleunen M, Weber E, Fischer M (2010) A meta-analysis of trait differences between invasive and non-invasive plant species. Ecology Letters 13:235-245. doi:10.1111/j.1461-0248.2009.01418.x

University of California - Los Angeles (2010) What's Invasive & Project Budburst. Center for Embedded Networked Sensing: Participatory Sensing, (PART) Research Projects. http://web.archive.org/web/20100708022918/http://research.cens.ucla.edu/urban/2010/part03.pdf. Accessed 4 November 2017

University of Florida Entomology and Nematology Department (2003) Florida Preying Mantids http://entnemdept.ufl.edu/choate/mantid_key2_03.pdf. Accessed 30 October 2017

University of Florida - Herbarium (2017) Preparation of plant specimens for deposit as herbarium vouchers. https://www.floridamuseum.ufl.edu/herbarium/voucher.htm. Accessed 2 November 2017

University of Minnesota (2016) Oliver E. Strand: Strand Memorial Herbarium. https://webapps2.cfans.umn.edu/herbarium/OliverStrand.aspx. Accessed 4 November 2017

Van der Eynden V, Corti L, Woollard M, Bishop L, and Horton L. May (2011) Managing and Sharing Data – Best Practices for Researchers. Third edition. UK Data Archive. University of Essex.

Vines TH, Albert AYK, Andrew RL, Débarre F, Bock DG, Franklin MT, Gilbert KJ, Moore JS, Renaut S, Rennison DJ (2014) The availability of research data declines rapidly with article age. Current Biology 25:94-97

Wagner CS, Park HW, Leydesdorff L (2015) The continuing growth of global cooperation networks in research: A conundrum for national governments. PLOS ONE 10(7):e0131816zi: 10.1371/journal.pone.0131816

Wallace RD, Bargeron CT, LaForest JH (2017) Data and knowledge preservation. Earthzine https://earthzine.org/2017/05/22/data-and-knowledge-preservation. Accessed 3 November 2017

Weakley AS (2015) Flora of the Southern and Mid-Atlantic States. UNC Herbarium, North Carolina Botanical Garden, University of North Carolina at Chapel Hill. http://www.herbarium.unc.edu/FloraArchives/WeakleyFlora_2015-05-29.pdf. Accessed 30 March 2018

Westbrooks RG and Eplee RE (2011) Early detection and rapid response. In: Encyclopedia of Biological Invasions. University of California Press, Berkeley, pp. 169-177

Westbrooks RG, Manning ST, Waugh JD (2014) Early detection and rapid response: a cost-effective strategy for minimizing the establishment and spread of new and emerging invasive plants by global trade, travel and climate change. In: Ziska LH and Dukes JS (eds) Invasive Species and Global Climate Change. CAB International, Oxfordshire UK, pp 305-325

Wheatley P (2017) National Geo-Spatial Administration capacities applicable to building a national incident response program for invasive species. In: Presentation at EDRR Assessments Discussion to Support Project Contractors, Washington D.C., March 2017

Whelan NV, Johnson PD, Garner JT, Strong EE (2017) On the identity of Leptoxis taeniata: a misapplied name for the threatened painted rocksnail (Cerithioidea, Pleuroceridae) ZooKeys 697:21-36. doi:10.3897/zookeys.697.14060

Wren JD, Grissom JE, Conway T (2006) E-mail decay rates among corresponding authors in MEDLINE. European Molecular Biology Organization Reports 7:122-127. doi:10.1038/sj.embor.7400631