Enabling decisions that make a difference: guidance for improving access to and analysis of invasive species information

Rebekah D. Wallace · Charles T. Bargeron IV · Jamie K. Reaser

Received: 3 November 2019 / Accepted: 3 November 2019 / Published online: 22 November 2019 © The Author(s) 2019

Abstract The issue of how to detect and rapidly respond to invasive species before it is economically infeasible to control them is one of urgency and importance at international, national, and subnational scales. Barriers to sharing invasive species data—whether in the form of policy, culture, technology, or operational logistics—need to be addressed and overcome at all levels. We propose guiding principles for following standards, formats, and protocols to improve information sharing among US invasive species information systems and conclude that existing invasive species information standards are adequate for the facilitation of data sharing among all sectors. Rather than creating a single information-sharing system, there is a need to promote interfaces among existing information systems that will enable them to become inter-operable, to foster simultaneous access, and to deliver any and all relevant information to a particular user or application in a seamless fashion. The actions we propose include implementing a national campaign to mobilize invasive species occurrence data into publicly available information systems; maintaining a current list of invasive species data integrators/clearinghouses; establishing an agreement for sharing data among the primary US invasive species information systems; enhancing the Integrated Taxonomic Information System to fully cover taxonomic groups not yet complete; further developing and hosting data standards for critical aspects of invasive species biology; supporting and maintaining the North American Invasive Species Management Association’s mapping standards; identifying standard metrics for capturing the environmental and socio-economic impact of invasive species, including impacts and management options; continuing to support US engagement in international invasive species data sharing platforms; and continuing US membership in the Global Biodiversity Information Facility.

Keywords Invasive species · Data management · Guidelines · Standards · Formats · Protocols

Introduction

The capacity of governments to prevent and respond to biological invasions depends on ready access to the best available scientific and socio-economic information (Convention on Biological Diversity 2006). Recognizing this, Presidential Executive Order (EO) 13112 (Executive Office of the President 1999) called
for “the establishment of a coordinated, up-to-date information sharing system that utilizes, to the greatest extent practicable, the Internet.”

At this time, numerous online species information systems exist within the United States that provide data and other information resources relevant to addressing the invasive species issue at state, regional and national levels (Marsico et al. 2010; Richardson and Rejmánek 2011). Another manuscript within this special issue (Reaser et al. 2019, this issue) has an associated catalog of more than fifty online tools and databases that receive federal funding to deal with some aspect of invasive species early detection and rapid response (Simpson et al. 2019). Each of these information systems was developed to meet different goals, objectives, and standards. Rather than creating a single information-sharing system, there is a need to promote interfaces among existing information systems that will enable them to become inter-operable, to foster simultaneous access, and to deliver any and all relevant information to a particular user or application in a seamless fashion (Reaser et al. 2007; Ricciardi et al. 2000).

A number of data providers and users, including the National Invasive Species Council (NISC) and its partners and contractors, the Invasive Species Advisory Committee (ISAC), and the Western Governors Association (WGA), have called for the development of the standards, formats, and protocols needed to facilitate the inter-operability of information systems (Davis Declaration 2001; Fell 2001; Invasive Species Advisory Committee 2017; Island Conservation 2018; Western Governors Association 2018). For example, the 2016–2018 National Invasive Species Management Plan states that,

“In order to facilitate inter-operability of data and other information resources relevant to addressing the invasive species issue, establish guidance for data management standards, formats, and protocols. The guidance should target the most relevant (high priority) information systems, capitalize on existing standards, and take into consideration the work that the Global Invasive Alien Species Information Partnership already initiated to explore options for information system inter-operability” (NISC 2016).

Additionally, Presidential Executive Order 13751 (2016), which updated EO 13112, directed that, “to the extent practicable, Federal agencies shall…develop, share, and utilize similar metrics and standards, methodologies, and databases and…facilitate the interoperability of information systems, open data, data analytics, predictive modeling, and data reporting necessary to inform timely, science-based decision making.”

Finally, the Western Governors’ Association supports a number of initiatives to advance coordinated invasive species management, including development of data management standards, formats, and protocols (WGA 2016).

In drafting this guidance, the authors recognize that:

(a) Decision making on invasive species necessarily requires access to and analysis of information on non-native species that have not been quantitatively evaluated for evidence of harm or likelihood of harm in a particular ecosystem (references to invasive species data herein are meant to encompass the full suite of non-native species data; Reaser et al. 2007).

(b) A considerable amount of work has already been undertaken at national and international scales to identify, promote, and agree to formats, standards, and protocols for the exchange of biological information (WGA 2018; North American Invasive Species Management Association 2014; Wieczorek et al. 2012).

(c) There are substantial benefits, including cost-efficiency and the scale of analytical capacity, to aligning with the existing agreements made by standard-setting bodies, both domestic and international, that guide the exchange of biological information (Hartley 1945, OMB 2005).

(d) Data relevant to addressing the invasive species issue are contained in a wide-range of governmental and non-governmental information systems that vary in purpose, structure, operation, and public accessibility (Reaser et al. 2019, this issue).

(e) While this guidance has been drafted to improve access to and analysis of invasive species information to meet US policy and management needs, invasive species frequently originate in other countries and information held in other countries is critical to meeting US goals (Perrings et al. 2010).
Likewise, information held in US information systems is vital to addressing invasive species by other countries, as well as cooperatively among regions and along invasive pathways (Perrings et al. 2010).

Inter-operability is urgently needed to foster scientific and technical cooperation and information dissemination and exchange, within the constraints of the infrastructure currently available (Simpson et al. 2009).

The capacity to make effective policy and management decisions on invasive species issues reflects the willingness and ability of federal, state, territorial, tribal, and local governments, as well as academic institutions, non-government organizations, and the private sector, to access and utilize each other’s data to the fullest extent warranted.

Guidance on standards, formats, and protocols in US invasive species information systems

In order to maximize the accessibility, cost-efficiency, and applicability of invasive species information systems, we encourage information system managers to adhere to the following guiding principles and best practices (Convention on Biological Diversity 2002), ensuring

(a) open access;
(b) open standards in common and future usage;
(c) future extensibility and backward compatibility;
(d) phased, incremental development;
(e) use of existing services and capabilities;
(f) scalability;
(g) inclusion (e.g., facilitation of local-language queries) in design applications;
(h) language neutrality in the design of applications;
(i) inter-operability which fosters cost-efficiencies and institutional cooperation;
(j) incorporation of scientific and technical cooperation and capacity development;
(k) respect for intellectual property rights and cross boundary issues;
(l) respect for applicable rules and regulations; and
(m) cooperation across sectors and governments (domestically and internationally).

Adoption of the standards set forth in Box 1 will maximize opportunities for information system interoperability. Below we provide a description of the standards that warrant further emphasis and clarification. They are critical to ensuring the timely accessibility and reliability of invasive species occurrence data.

The Integrated Taxonomic Information System (ITIS) Taxonomic Serial Number is used to identify the species or taxon (https://www.itis.gov, accessed 29 March 2019) under consideration. In many data collection programs, shorthand or codes are used to key in species names. Many taxonomy databases have their own code (e.g., USDA Plants Database has its Plants Symbol (USDA 2018); and Mycobank has their unique number as MB# International Mycological Association 2018), but are also narrow in focus. ITIS serves all taxa occurring in the United States and has several global taxonomic treatments.

A universally unique identifier (UUID) should be assigned to each species record and then registered/maintained with a Digital Object Identifier (DOI) (or equivalent) by the resource originator. Multiple concerns arise with regards to data sharing, including issues relevant to data authenticity, duplication, correction, and updating. Version 4 UUID are a series of letters and numbers separated by hyphens in an 8-4-4-4-12 character format that are not housed or regulated by any organization but have only a 1 in $2^{122}$ chance of duplication (Chen 2016). Use of a UUID allows for duplicate record checks and error correction as data are shared. A UUID can be automatically generated by many commonly used databases (Esri 2016; Integrated Digitized Biocollections 2014) or through websites (e.g., https://www.uuidgenerator.net/version4, accessed 29 March 2019) and added to records.

Use of a DOI enables reference to the exact information source and, per membership in a DOI assigning organization such as DataCite (https://www.datacite.org, accessed 29 March 2019), any changes in location/URL to the information must be reflected in the metadata of the DOI database to avert broken links or inaccessibility (International DOI Foundation 2017). DOIs are available through services which have a membership with the International DOI Foundation (https://www.doi.org, accessed 29 March 2019), including data set repositories, journal publishers, and more (International DOI Foundation 2017).
It is important that invasive species occurrence data be exportable and fully compatible with the North American Invasive Species Management Association (NAISMA 2018) mapping standard format. Invasive species occurrence data are often inconsistent in formatting, field definitions, and data type per field. This limits data sharing and quality control capacities. In order to overcome this challenge, NAISMA (formerly North America Weed Management Association) is in the process of revising their standards for mapping invasive species data. NAISMA has had mapping standards in place for plants since 2002 and revised these standards in 2014 and 2018 to address all taxa of invasive species.

Invasive species data holders are encouraged to make data public and digitally available to data aggregators using recognized standards. While multiple regulations have been signed that direct federally funded data and information to be made open, transparent, machine readable, free, and rapidly accessible (Office of Management and Budget 2016; Office of Science and Technology Policy 2014; Interagency Working Group on Open Data Sharing Policy 2016; Burwell et al. 2013; Executive Office of the President 2013; Holdren 2013), the compliance and promotion of these policies have been lacking. To ensure that data (not just information summaries) are available, any research proposal, grant funding, or contractual agreement should include a plan for data management, preservation, and accessibility. Promotion and adoption of the NAISMA standards (NAISMA 2018) and other standards listed in this document will aid data incorporation into aggregate databases, making the data more broadly available and applicable to timely and reliable decision making.
Finally, data aggregators (who might also be regarded as data integrators) need to be compelled and supported to ensure data attribution, accuracy, authority, and timeliness, as well as to enable interoperability with emerging technology platforms for data acquisition and analysis. Data aggregators have the responsibility to ensure that the information publicly available through their information platforms is sufficiently reliable for policy and management decision-making, as well as ensuring adequate and appropriate attribution to their data sources (Reichman et al. 2011). They also have a role in establishing a seamless relationship between the information systems they manage and the best available analytical and decision support tools.

**Priority actions**

To make effective use of this guidance, additional priority actions will need to be accomplished. We encourage relevant federal and state agencies and their partners to undertake the following actions, which are echoed and expanded upon in Reaser et al. (2019), this issue.

(a) Create and implement a national campaign to mobilize invasive species occurrence data into publicly available information systems according to the principles, standards, formats, and protocols described in this paper. Effective policy and management decisions on invasive species issues necessitate that all levels of government, as well as academic institutions, non-government organizations, and the private sector, are willing to make invasive species occurrence data publicly accessible. Data need to be actively mobilized from a wide range of sources (e.g., databases, technical reports, peer-reviewed and gray literature, social media) to information systems that are managed according to the guidance herein.

(b) Create and routinely update a list of data aggregators/clearinghouses through which relevant data can be openly shared. A considerable amount of invasive species data is not currently available in widely accessible information systems (e.g., data generated from individual research projects, biological surveys not intentionally focused on invasive species, and environmental impact assessments). Lack of accessibility limits our capacity to apply this information for policy and management decision-making. A listing of repositories or clearinghouses is needed to help mobilize federal and non-federal data sets, with the ultimate goal of encouraging data contribution for data application. The public availability of information also enables greater expert review and data quality assurance. Ideally, this list would be accessible on the NISC website, but also include reference to non-federal information systems.

(c) Establish an agreement for sharing data among the primary information systems for non-native/invasive species occurrence data in the United States. As a minimum, this should include the USGS Biodiversity Information Serving Our Nation (BISON) all-species information system (https://bison.usgs.gov, accessed 29 March 2019), Early Detection and Distribution Mapping System (EDDMapS; https://www.eddmaps.org, accessed 29 March 2019), iMap Invasives (https://www.imapinvasives.org, accessed 29 March 2019) data management system, and the USGS Nonindigenous Aquatic Species (NAS; https://nas.er.usgs.gov, accessed 29 March 2019) information resources. It should also be open to participation from organizations facilitating the collection of invasive occurrence data by citizen scientists (e.g., iNaturalist; https://www.inaturalist.org, accessed 29 March 2019) and others.

(d) Mobilize invasive species occurrence data into central data aggregators that are coupled with appropriate analytical and decision support tools. At a minimum, BISON should be utilized in this capacity (WGA 2018). A broad, openly accessible, authoritative all-species database that meets the standards, formats, and protocols described herein is needed to serve as the national focal point for invasive species occurrence data. Multiple federal agencies and non-federal partners have already made substantial investments in BISON to achieve this aim (Jackson et al. 2016; Sutton and Armsworth 2014). BISON contributes to the USGS National Biogeographic Map, which is a prototype application designed to provide a platform
for data analysis packages about what species are where, how well protected they are, how vulnerable they are to stressors, and what management options will have the greatest benefit for conservation of biodiversity (USGS 2019).

(e) Enhance the Integrated Taxonomic Information System (ITIS) to fully cover taxonomic groups not yet complete, with particular emphasis on those from taxonomic groups prone to invasiveness. Currently, ITIS has virtually complete taxonomy for plants, bacteria, vertebrates, most insects, and other important groups but is lacking in some other categories of increasing importance for invasions, such as many fungi and viruses. As there are many invasive diseases caused by fungi and viruses, ITIS should ensure all invasive pathogens and parasites are included in its system and seek resources to comprehensively address fungi and working with the community to develop and adopt a single consistent classification for viruses. It also needs to be quickly informed of any new non-native species that arrive in the United States so that its treatment of invasive species is comprehensive. The current effort to fully deploy the ITIS global taxonomic workbench to dramatically streamline the name addition and vetting process should be fully supported. ITIS providing taxonomic serial numbers across all taxa will facilitate data sharing and reduce errors in taxonomy due to inconsistent, shorthand, or custom species coding, as this number never changes, even when the accepted names evolve (Integrated Taxonomic Information System 2018).

(f) Develop and host data standards for critical aspects of invasive species biology and population parameters (e.g., resource use, pathways of movement, types and degree of impacts). Work on these standards has been initiated by Global Invasive Species Information Network (GISIN 2018), but priority attention is warranted. These metrics are needed to help distinguish which non-native species are invasive (i.e., harmful), as well as to prioritize and plan response measures. The appropriate global platform for invasive species data standards development is the Biodiversity Information Standards working group (also known as the Taxonomic Database Working Group (TDWG; https://www.tdwg.org/community/tnc/, accessed 3 July 2019).

(g) Support and maintain the NAISMA mapping standards. NAISMA standards are being modified to include aquatic standards and in the 2014 version there are fields that are unresolved for data type (NAISMA 2014). NAISMA is updating their mapping standards with information gathered from multiple recent workshops. NAISMA will also seek endorsement from multiple agencies and organizations to promote the adoption of the standards as broadly as possible. Fields in the NAISMA mapping standard, as appropriate, should be mapped to or harmonized with their Darwin Core II equivalent.

(h) Identify the standard metrics for capturing the environmental and socio-economic impacts of invasive species. Risk analyses necessitate both ecological and socio-economic impact metrics. The Environmental Impact Classification for Alien Taxa (EICAT; https://www.iucn.org/theme/species/our-work/invasive-species/eicat, accessed 3 July 2019) and Socio-economic Impact Classification of Alien Taxa (SEICAT; Bacher et al. 2017) standards should be assessed for relevant applicability. While there are an increasing number of ecological impact assessments, socio-economic impact studies lag far behind. In general, impact research tends to be very narrowly focused and ill-defined. An expert’s consultation to define “socio-economic impact” parameters and then identify the metrics by which to evaluate species for actual or predicted impact would aid in clarifying communication between stakeholders and scientists (Jeschke et al. 2014; Bacher et al. 2017). Once these metrics are identified and agreed to, various analytical tools can be developed, tested, and utilized in tandem with existing predictive models, such as weed risk assessments.

(i) Encourage and accommodate information on invasive species impacts and management options. Information on invasive species impacts and management approaches could provide valuable insight to the wider invasive species policy and management communities. This could include not only background
information on management options, but specifics on when and where such management was applied, the resources required for the actions, and the effectiveness of the action. Ideally, this information would be linked to occurrence data to enable context-specific interpretation of the impact and management parameters and options.

(j) Continue to support US engagement in international information frameworks and platforms that advance invasive species data sharing in keeping with the guidance herein. For example, support TDWG’s ongoing effort to develop and publish a formal Darwin Core II extension for invasive species data. Given that the bulk of invasive species occurrence data globally is held (or exportable) in Darwin Core II format, a well-designed and documented enterprise extension to accommodate the salient business rules and required augmentations of the NAISMA standards is needed. This would allow for the seamless accommodation of a much larger group of relevant data in the software systems and analysis libraries that already exist for Darwin Core II.

(k) Continue US membership in the Global Biodiversity Information Facility (GBIF; https://www.gbif.org, accessed 29 March 2019) to enable invasive species data sharing and analyses at multi-national scales, including those data relevant to understanding invasion risks and pathways. The National Science Foundation, USGS, and many other US organizations contribute and provide leadership to GBIF and play a critical role in mobilizing data, promoting standards, and providing access to data. BISON serves as the US hub for GBIF and includes within it all GBIF species occurrence data for the United States and Canada (Hanken 2013).

Conclusion

The invasive species issue is one of urgency and importance at international, national, and subnational scales. Broad collaboration among government agencies, non-governmental organizations, academia, and the private sector is needed to ensure that “We can do this!”—we can minimize the impact of invasive species on the environment and economy, as well as human, animal, and plant health. Substantial public will, financial resources, and institutional collaboration have been invested to this end; it is thus imperative that we achieve effectiveness and cost-efficiency by maximizing the return on these investments. Barriers to sharing invasive species data—whether in the form of policy, culture, technology, or operational logistics—need to be addressed and overcome. Existing standards for biodiversity information (such as those listed listed in Box 1) are adequate for the facilitation of invasive species data sharing among all sectors. What we need now is the will to enable the greatest possible benefits to all.

Acknowledgments This document advances action 2.4 of the 2016–2018 NISC Management Plan. A previous version of the manuscript was produced under a NISC Secretariat contract to the University of Georgia and was made available on the web. The authors are grateful to everyone who contributed to the federal agency surveys and supplemental inquiries. The authors thank Mike Frame, Pam Fuller, Gerald Guala, Joseph H. LaForest, Marcia McNiff, Andrew Mendoza, Enrique Montaño, David J. Moorhead, Jeff Morisette, Matt Neilson, and Annie Simpson for their review and constructive comments. Input was also invited during the Western Governors Association Invasive Species Data Management Workshop (March 14–15, 2018). The findings and conclusions in this article are those of the author(s) and do not necessarily represent the views of the US Government. The majority of Jamie K. Reaser’s contributions occurred while Executive Director of the National Invasive Species Council.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

Bacher S, Blackburn TM, Essl F, Genovesi P, Heikkilä J, Jeschke JM, Jones G, Keller R, Kenis M, Kueffer C, Martinou AF (2017) Socio-economic impact classification of alien taxa (SEICAT). Methods Ecol Evol 9:159–168
Basques K (2018) Why HTTPS matters. Web fundamentals, Google developers. https://developers.google.com/web/fundamentals/security/encrypt-in-transit/why-https. Accessed 17 April 2018
Burwell SM, VanRoekel S, Park T, Mancini DJ (2013) Open data policy-managing information as an asset. Office of
Holdren JP (2013) Increasing access to the results of federally government-supported scientific data and research findings through international scientific cooperation. Interagency Working Group on Open Data Sharing Policy, the White House Executive Office of the President, Washington, DC. https://obamawhitehouse.archives.gov/sites/default/files/microsites/ostp/NSTC/iwgodsp_principles_0.pdf. Accessed 17 Jan 2018

International DOI Foundation (2017) Frequently asked questions about the DOI® System. DOI.org. http://www.doi.org/faq.html. Accessed 1 Feb 2018

International Mycological Association (2018) MycoBank. database. MycoBank. http://www.mycobank.org. Accessed 1 Feb 2018

International Organization for Standardization for improved exchange of biodiversity-related information, convention on biological diversity, Montreal, Canada. https://www.cbd.int/doc/meetings/chm/chmifmps-01/official/chmifmps-01-02-en.doc. Accessed 07 Feb 2019

Convention on Biological Diversity (2002) Report of the joint CBD/GISP informal meeting on formats, protocols and standards for improved exchange of biodiversity-related information, convention on biological diversity, Montreal, Canada. http://www.itis.gov/pdf/faq_itis_tsn.pdf. Accessed 1 Feb 2018

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. Interagency Working Group on Open Data Sharing Policy, the White House Executive Office of the President, Washington, DC. https://obamawhitehouse.archives.gov/sites/default/files/microsites/ostp/NSTC/iwgodsp_principles_0.pdf. Accessed 15 Jan 2018

Convention on Biological Diversity (2006) Alien species that threaten ecosystems, habitats or species. Convention on biological diversity’s sixth conference of the parties decision VI-23. https://www.cbd.int/decision/cop/default.shtml?id=7197. Accessed 07 Dec 2018

Davis Declaration (2001) Workshop on development of regional invasive alien species information hubs, including requisite taxonomic services, in North America and Southern Africa, 14–15 February 2001, Davis, California. https://www.doi.gov/sites/doi.gov/files/uploads/davis_declaration_on_invasive_species_2001.pdf. Accessed 15 Nov 2018

Esri (2016) How to: calculate unique identifier values similar to global IDs. Esri. https://support.esri.com/en/technical-article/000011677. Accessed 2 Feb 2018

Executive Office of the President (1999) Executive Order 13112, 84 FR 6183-6186, 8 February 8 1999

Executive Office of the President (2013) Executive Order 13642, 78 FR 28111-28113, 9 May 2013

Executive Office of the President (2016) Executive Order 13751, 81 FR 88609-88614, 5 December 2016

Fell A (2001) Davis declaration calls for action on invasive species. University News, UC Davis. https://www.ucdavis.edu/news/davis-declaration-calls-action-invasive-species. Accessed 15 Nov 2018

Global Invasive Species Information Network (2018) GISIN ontology for invasive species information. Biodiversity Information Standards Taxonomic Databases Working Group https://github.com/tdwg/gisin Accessed 7 Dec 2018

Hanken J (2013) Biodiversity online: toward a network integrated biocollections alliance. Bioscience 63(10):789–790

Hartley P (1945) International biological standards: prospect and retrospect: walter ernst dixon memorial lecture. Proc R Soc Med 4(29):45–58

Holdren JP (2013) Increasing access to the results of federally funded scientific research. Office of Science and Technology Policy Memorandum, United States Government, Washington, DC. https://obamawhitehouse.archives.gov/sites/default/files/microsites/ostp/ostp_public_access_memo_2013.pdf. Accessed 06 Dec 2018

Integrated Digitized Biocollections (2014) UUID excel generator. Integrated Digitized Biocollections. https://www.idigbio.org/wiki/images/0/03/GUIDgeneration.pdf. Accessed 1 Feb 2018

Integrated Taxonomic Information System (2018) What is an Integrated Taxonomic Information System “TSN”? ITIS. https://www.itis.gov/pdf/faq_itis_tsn.pdf. Accessed 1 Feb 2018

Jackson ST, Duke CS, Hampton SE, Jacobs KL, Joppa LN, Kassam KAS, Mooney HA, Ogden LA, Ruckelshaus M, Shogren JF (2016) Toward a national, sustained US ecosystem assessment. Science 354(6314):838–839. https://doi.org/10.1126/science.aah5750

Jeschke JM, Bacher S, Blackburn TM, Dick JTA, Essi F et al (2014) Defining the impact of non-native species. Conserv Biol 28:1188–1194. https://doi.org/10.1111/cobi.12299

Library of Congress (2013) ISO 639-2: codes for the representation of names of languages. https://www.loc.gov/standards/iso639-2/php/code_list.php. Accessed 17 April 2018
