

## ENABLING DECISIONS THAT MAKE A DIFFERENCE

### GUIDANCE FOR IMPROVING ACCESS TO AND ANALYSIS OF INVASIVE SPECIES INFORMATION

#### A. INTRODUCTION

1. The capacity of governments to prevent and respond to biological invasions depends on ready access to the best available scientific and socio-economic information. Recognizing this, Presidential [Executive Order 13112](#) called for “the establishment of a coordinated, up-to-date information sharing system that utilizes, to the greatest extent practicable, the Internet.”
2. At this time, numerous information systems exist within the United States that provide data and other information resources relevant to addressing the invasive species issue. 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.
3. Numerous data providers and data users, including the [National Invasive Species Council](#), [Invasive Species Advisory Committee](#), and [Western Governors Association](#), have called for the development of the standards, formats, and protocols needed to facilitate the inter-operability of information systems. For example:
  - a. 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;”
  - b. [Executive Order 13751](#) 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”; and

This document advances [2016–2018 NISC Management Plan](#) action 2.4. It was originally produced under a NISC Secretariat contract to the University of Georgia enacted by Rebekah D. Wallace, Charles T Barger IV, David J. Moorhead, Joseph H. LaForest. Subsequent advances were made with input from federal agency staff, including Pam Fuller, Stinger Guala, Andrew Mendoza, Jeff Morisette, Matt Neilson, Jamie K. Reaser, and Annie Simpson. Input was also invited during the Western Governors Association Invasive Species Data Management Workshop (March 14-15, 2018). This is a living document; it will be periodically updated as needed to maintain consistency with advances in information technologies, information system management, and standard setting.

- c. *Western Governors' Association Policy Resolution 2016-05* supports a number of initiatives to advance coordinated invasive species management, including development of data management standards, formats, and protocols.
4. In drafting this guidance, the authors recognized 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);
  - 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;
  - 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;
  - 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;
  - e. While this guidance has been drafted to improve access to and analysis of invasive species information to meet U.S. policy and management needs, invasive species frequently originate in other countries and information held in other countries is critical to meeting U.S. goals;
  - f. Likewise, information held in U.S. information systems is vital to addressing invasive species by other countries, as well as cooperatively among regions and along invasive pathways;
  - g. Inter-operability is urgently needed to foster scientific and technical cooperation and information dissemination and exchange, within the constraints of the infrastructure currently available; and
  - h. 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.

## **B. GUIDANCE ON STANDARDS, FORMATS, AND PROTOCOLS IN U.S. INVASIVE SPECIES INFORMATION SYSTEMS**

1. Guiding principles for the application of standards, formats, and protocols for achieving inter-operability include:
  - a. **Open access**;
  - b. **Open standards** in common and future usage;
  - c. Future extensibility and backward compatibility;
  - d. Phased, incremental development;
  - e. Building on existing services and capabilities;
  - f. Scalability;
  - g. Inclusion (e.g. facilitate local-language queries) in design of applications;
  - h. Language neutrality in the design of applications;
  - i. Use of inter-operability as a tool for fostering 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 among governments (domestically and internationally).
2. The authors encourage U.S. data managers to adopt the following formats, standards, and protocols in

order to enable policy and management decisions that lead to the prevention, eradication, and control of invasive species in a timely and cost-effective manner.

- a. *Recommended standards:*
    - i. [Darwin Core II](#)
    - ii. North American Invasive Species Management Association ([NAISMA](#)) mapping standards
    - iii. Global Invasive Species Information Network ([GISIN](#))
    - iv. [International Digital Object Identifier \(DOI\) Foundation](#);
    - v. [International Organization for Standardization \(ISO\) 19115 spatial metadata](#);
    - vi. [ISO 23950 interoperability](#);
    - vii. [ISO 25964-1:2011 thesauri](#);
    - viii. [ISO 3166 country codes](#);
    - ix. [ISO 639-2 language codes](#);
    - x. [ISO 8601 time and date representations](#);
    - xi. Internet Engineering Task Force ([IETF](#)) Requests for Comment ([RFCs](#)) – *various*;
    - xii. [Open Geospatial Consortium](#);
    - xiii. [Version 4 Universally Unique Identifier \(UUID\)](#);
    - xiv. [Worldwide Web Consortium \(W3C\)](#);
  - b. *Recommended formats:*
    - i. [JSON](#)
    - ii. [HTML](#)
    - iii. [PDF](#)
  - c. *Recommended protocols*
    - i. [https](#)
3. In the context of invasive species information management, the following standards warrant further emphasis and clarification as they are critical to ensuring the timely accessibility and reliability of invasive species occurrence data:
- a. **Use the Integrated Taxonomic Information System ([ITIS](#)) Taxonomic Serial Number to identify the species or taxon.** 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 their 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.
  - b. **Assign a universally unique identifier (UUID) to species records and register/maintain information 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<sup>122</sup> chance of duplication (Chen 2016). Use of a UUID allows for duplicate record checks and error correction as data are shared. UUID can be automatically generated by many commonly used databases (ESRI 2016; iDigBio 2014) or through [websites](#) and added to records. Use of a DOI enables reference to the exact information source and, per membership in a DOI assigning organization, 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). DOI are available through services which have a membership with the [International DOI Foundation](#), including data set repositories, journal publishers, and more (International DOI Foundation 2017).
  - c. **Ensure that invasive species occurrence data are exportable and fully compatible with the NAISMA mapping standard format.** Invasive species occurrence data are not consistent in such parameters as

formatting, field definitions, and data type per field. This limits capacities for data sharing and quality control. In order to overcome this challenge, NAISMA is in the process of revising their standards for mapping invasive species data. NAISMA (formerly North America Weed Management Association) has had mapping standards in place for plants since 2002 and revised these standards in 2014 to address all taxa of invasive species.

- d. **Leading by example, encourage all invasive species data holders to make their data public and digitally available to data aggregators using recognized standards.** While multiple federal regulations have been signed that direct federally-funded data and information to be made open, transparent, machine readable, free, and rapidly accessible, 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](#) 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.
- e. **Compel and support data aggregators to ensure data attribution, accuracy, authority, and timeliness, as well as enable inter-operability 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. 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.

## C. PRIORITY ACTIONS

4. To make effective use of this guidance, additional priority actions will need to be accomplished:
  - 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 herein.** 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 National Invasive Species Council ([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 Biodiversity Information Serving Our Nation ([BISON](#)) information system, Early Detection and Distribution Mapping System ([EDDmaps](#)), [iMap Invasives](#) data management system, and Nonindigenous Aquatic Species ([NAS](#)) information resources. It should also be open to participation from organizations facilitating the collection of invasive occurrence data by citizen scientists (e.g., [iNaturalist](#)).

- d. **Mobilize invasive species occurrence data into central data aggregators that are coupled with appropriate analytical and decision support tools.** As a minimum, BISON should be utilized in this capacity. A broad, openly-accessible, authoritative 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.
- e. **Enhance the Integrated Taxonomic Information System (ITIS) to fully cover taxonomic groups not yet complete.** 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; especially seeking 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. **Development and hosting of 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 GISIN (Global Invasive Species Information Network 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 (i.e. TDWG).
- g. **Support and maintain the NAISMA mapping standards.** NAISMA standards are currently not complete, as they are missing aquatic standards and there are fields that are currently unresolved for data type (North American Invasive Species Management Association 2014). NAISMA will be 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) and Socio-economic Impact Classification of Alien Taxa (SEICAT) 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 experts 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. 2018). 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 U.S. engagement the international information frameworks and platforms that advance invasive species data sharing in keeping with the guidance herein.*** For example:
  - i. 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.
  - ii. Continue U.S. membership in the Global Biodiversity Information Facility (GBIF) to enable invasive species data sharing and analyses at multi-national scales, including those data relevant to understanding invasion risks and pathways. BISON serves as the U.S. hub for GBIF.

## D. 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 impacts 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 are adequate for the facilitation of data sharing among all sectors. What we need now is the will to enable the greatest possible benefits to all.

## REFERENCES

- Bacher, S, Blackburn, TM, Essl, F, Genovesi, P, Heikkilä, J, Jeschke, JM, Jones, G, Keller, R, Kenis, M, Kueffer, C and Martinou, AF (2018) Socio-economic impact classification of alien taxa (SEICAT). *Methods in Ecology and Evolution*, 9: 159-168. <http://onlinelibrary.wiley.com/doi/10.1111/2041-210X.12844/epdf>. Accessed 7 March 2018
- Burwell SM, VanRoekel S, Park T, Mancini DJ (2013) Open data policy: managing information as an asset. The White House Office of Management and Budget. <https://obamawhitehouse.archives.gov/sites/default/files/omb/memoranda/2013/m-13-13.pdf>. Accessed 30 October 2017
- Chen R (2016) When you start talking about numbers as small as  $2^{-122}$ , you have to start looking more closely at the things you thought were zero. *The Old New Thing*. <https://blogs.msdn.microsoft.com/oldnewthing/20160114-00/?p=92851>. Accessed 1 February 2018
- ESRI (2016) How To: Calculate unique identifier values similar to global IDs. ESRI. <https://support.esri.com/en/technical-article/000011677>. Accessed 2 February 2018
- Executive Order No. 13642 (2013) Making open and machine readable the new default for government information. 78 FR 28111. <https://obamawhitehouse.archives.gov/the-press-office/2013/05/09/executive-order-making-open-and-machine-readable-new-default-government>. Accessed 30 October 2017
- Executive Order No. 13751 (2016) Safeguarding the Nation from the impacts of invasive species. 81 FR 88609. <https://obamawhitehouse.archives.gov/the-press-office/2016/12/05/executive-order-safeguarding-nation-impacts-invasive-species>. Accessed 30 January 2018

- Global Invasive Species Information Network (2018) GISIN Protocol. GISIN. [http://www.gisin.org/cwis438/websites/GISINDirectory/Tech/Protocol\\_Home.php?WebSiteID=4](http://www.gisin.org/cwis438/websites/GISINDirectory/Tech/Protocol_Home.php?WebSiteID=4). Accessed 1 February 2018
- Holdren JP (2013) Memorandum for the heads of executive departments and agencies: increasing access to the results of federally funded scientific research. The White House Office of Science and Technology Policy. [https://obamawhitehouse.archives.gov/sites/default/files/microsites/ostp/ostp\\_public\\_access\\_memo\\_2013.pdf](https://obamawhitehouse.archives.gov/sites/default/files/microsites/ostp/ostp_public_access_memo_2013.pdf). Accessed 30 October 2017
- iDigBio (2014) UUID excel generator. Integrated Digitized Biocollections. <https://www.idigbio.org/wiki/images/o/o3/GUIDgeneration.pdf>. Accessed 1 February 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. The White House Executive Office of the President. [https://obamawhitehouse.archives.gov/sites/default/files/microsites/ostp/NSTC/iwgodsp\\_principles\\_o.pdf](https://obamawhitehouse.archives.gov/sites/default/files/microsites/ostp/NSTC/iwgodsp_principles_o.pdf). Accessed 15 January 2018
- International DOI Foundation (2017) Frequently asked questions about the DOI® System. DOI.org. <http://www.doi.org/faq.html>. Accessed 1 February 2018
- International Mycological Association (2018) MycoBank database. MycoBank. <http://www.mycobank.org>. Accessed 1 February 2018
- Integrated Taxonomic Information System (2018) What is an Integrated Taxonomic Information System “TSN?”. ITIS. [https://www.itis.gov/pdf/faq\\_itis\\_tsn.pdf](https://www.itis.gov/pdf/faq_itis_tsn.pdf). Accessed 1 February 2018
- Jeschke JM, Bacher S, Blackburn TM, Dick JTA, Essi F, Evans T, Gaertner M, Hulme PE, Kühn I, Mrugała A, Pergl J, Pyšek P, Rabitsch W, Ricciardi A, Richardson DM, Sendek A, Vila M, Winter M, Kumschick S (2014) Defining the impact of non-native species. *Conserv Biol* 28:1188-1194. <http://dx.doi.org/10.1111/cobi.12299>
- Johnson III C (2005) Improving public access to and dissemination of government information and using the federal enterprise architecture data reference model. The White House Office of Management and Budget. <https://georgewbush-whitehouse.archives.gov/omb/memoranda/fy2006/mo6-02.pdf>. Accessed 30 October 2017
- Mervis J (2017) Data check: U.S. government share of basic research funding falls below 50%. *Sci Mag*. <https://doi.org/10.1126/science.aal0890>
- National Invasive Species Council (2016) Management plan: 2016–2018. <https://www.doi.gov/sites/doi.gov/files/uploads/2016-2018-nisc-management-plan.pdf>. Accessed 6 March 2018
- North American Invasive Species Management Association (2014) North American invasive species mapping standards. NAISMA. [http://www.naisma.org/images/NAISMA\\_Mapping\\_Standards2014\\_1.pdf](http://www.naisma.org/images/NAISMA_Mapping_Standards2014_1.pdf). Accessed 4 November 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. The White House Office of Management and Budget. [https://georgewbush-whitehouse.archives.gov/omb/inforeg/2005\\_cb/final\\_2005\\_cb\\_report.pdf](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. The White House Office of Management and Budget. <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. The White House Office of Science and Technology Policy. [https://obamawhitehouse.archives.gov/sites/default/files/microsites/ostp/us\\_open\\_data\\_action\\_plan.pdf](https://obamawhitehouse.archives.gov/sites/default/files/microsites/ostp/us_open_data_action_plan.pdf). Accessed 15 January 2018
- Orszag P (2009) Open government directive M-10-06. The White House Office of Management and Budget. <https://obamawhitehouse.archives.gov/open/documents/open-government-directive>. Accessed 30 October 2017
- United States Department of Agriculture (2018) The PLANTS database. USDA Natural Resources Conservation Service. <https://plants.usda.gov/faq.html>. Accessed 1 February 2018