Mapping Standards For Program Managers

Approved by:
North American Invasive Species Management Association

October 17, 2018

Table of Contents

Contents

Introduction

"Standards are something established for the use of a rule or basis of comparison in measuring or judging capacity, quantity, quality, content, extent, value, etc." - Webster's New World Dictionary

The definition above served as a guiding reminder of purpose to the original Mapping Standards Committee of the North American Weed Management Association (NAWMA). The standards were officially adopted by NAWMA in early 2001, after years of development by the committee. Their reason for creating the minimum mapping standards for invasive weeds was to increase the ability to share weed-mapping information - to create minimum standards so that the most basic information on infestations would be compatible between organizations and jurisdictions. The standards addressed the minimum base information necessary to compare and combine invasive weed maps across tribal, county, state/provincial, national, and even international borders:

The Five Basic Elements of Invasive Species Inventories:

1. What species was documented?
2. Where on the landscape was this species documented?
3. How large was the area infested by the species documented?
4. When was the information on this species infestation documented?
5. Who collected the documentation of this species infestation?

These standards were enthusiastically adopted by many North American weed management agencies and have been modified since their 2001 origins after feedback from those using them. Invasive species mapping and data sharing technology has improved and expanded considerably since the standards were first adopted, in ways that the original NAWMA Mapping Standards Committee could scarcely have imagined. NAWMA itself has evolved and expanded in scope to become the North American Invasive Species Management Association (NAISMA). **In 2014, the NAISMA Mapping Standards Committee was charged with amending and expanding their existing weed mapping standards to encompass all invasive species. In this 2018 version, the standards now include considerations for mapping aquatic invasive species, based on recommendations provided by the USGS – Nonindigenous Aquatic Species database, and for mapping biocontrol.**

Why Use These Standards?

These standards are intended to be compatible with most existing invasive species inventories. They are not intended to discourage other organizations from collecting additional information on invasive infestations. By using these minimum standards, information collected can be incorporated within inventories serving other purposes, thereby widening the usefulness of the collected information. These standards are intended to be as user friendly as possible, while still providing information essential at every level of invasive species management, from the site of the infestation to regional, national, and international levels. By adopting these standards you will be joining the invasive species community in making data more shareable across boundaries.

Inventory and Monitoring Standards

This chapter describes the basic information necessary to inventory and monitor populations of invasive species. These data and mapping standards represent the minimum or core information necessary to characterize a population of invasive species.

There are basic elements to invasive species inventories that enable them to be easily shared. This chapter contains those data fields required to satisfy these basic inventory elements, as well as optional fields that may be included. A sample field form can be obtained at the NAISMA web site, www.NAISMA.org.

Each data field/subject area is divided into the following subheadings:

Field Name(s): This is the name that will appear on the inventory form and on requests for information between agencies, states, and management areas. It will be the name used to share equivalent information between users. It will provide common vocabulary for sharing information. Words separated by commas are recommended to be in separate columns, form fields, or database fields. Words separated by "OR" allow for one field name or the other alternative name to be used.

Definition: Provides a description and explanation of the data field.

Why it is Useful: Describes why this information may be important and how it will be useful in describing infestations.

Core Element: This tells you whether this is a 1) required (core), 2) recommended, or 3) optional data field. Some data elements are very common and useful for invasive species inventories but will not be required for information sharing; these will be called optional fields. Recommended fields are optional fields of high importance.

Coding: Describes the proper way information should be entered.

Data Value: Describes the types of characters that will be accepted by a field. Options include numeric, text, or alphanumeric (consisting of both text and numbers).

Example: Provides a sample of the proper coding.

Record Level Identifiers

Universally Unique Identifier

Field Name: UUID

Definition: A universally unique identifier that can be assigned to each record to ensure that as data is collected, each record is distinct and can be referenced and queried. There is not a single database of assigned UUIDs (many UUID generator services are freely available on the web), but the chance of a generated UUID being duplicated is extremely small.

Why it is useful: This will allow for records to be shared and ensure that duplicates are not entered into any dataset aggregating system. Users and database managers alike will be able to reference a specific record across databases as the UUID travels with the record.

Core Element: This is a required field as of 2020.

Coding: The Version 4 UUID is a 128-bit number used to identify information that is made of alphanumeric characters in an 8-4-4-4-12 format. These can often be generated within a database (e.g. SQL, ArcGIS, etc.) or through a variety of websites. A listing of resources on how to generate UUID in databases and websites for UUID generation can be found in the Technical Resources section at the end of the document.

Data Value: Alphanumeric

Example:

UUID: defeaa69-703d-403a-a3cd-69eee650fd8c

Persistent Identifier

Field Name: PID

Definition: A unique identifier that can be assigned to each record to ensure that as data is entered into a database and published each record is distinct and can be referenced and queried back to a resolvable location or reference. This can include a website URL for a specific record, such as those created in iNaturalist or EDDMapS, a DOI to a journal article documenting the first occurrence of a newly introduced species, or other systems of publishing a record in a persistent manner. There is not a sole generator of PIDs, but most generators of these identifiers utilize technologies to ensure they

remain unique.

Why it is useful: This will allow records to be traced back to the original record or source and, when shared, it can help to ensure that duplicates are not entered into an aggregate system.

Core Element: Recommended

Coding: There are many different formats and sources that can serve as PIDs: URL, URN, DOI, URI, etc. are examples of resolvable identifiers that can link back to an original source for an individual record.

Data Value: Alphanumeric

Example:

PID: https://doi.org/10.1109/5.771073

Catalog Number

Field Name: Catalog Number

Definition: A unique identification number within a dataset that records the museum/herbarium catalog number or other collection number

Why it is Useful: This number can help to locate data or place it within an existing dataset. Many contributors to herbaria and museums will include numbers or identifiers for records within a collection. The Catalog Number can also come from herbaria and museum collections.

Core Element: Optional

Coding: Enter the information as provided; generally a unique value within a dataset.

Data Value: Alphanumeric

Example: Record comes from a digitized account of a specimen submitted to Harvard University Herbaria with an ID of HVD-154356

Catalog Number: HVD-154356

Taxonomy/Subject of Report

Species Name

Field Name(s): Kingdom (required), Family (required), Genus (required), Species (recommended), Subspecies (optional), variety (optional) Authorship (optional)

Definition: The scientific or taxonomic name of the species of the occurrence. The scientific name follows the standard of binomial nomenclature and consists of the genus name followed by the species name, expressed together and written in italics. Some species are further classified into subspecies, variety, or hybrids. Lastly, the individual who first classified the species and assigned the scientific name (formatted according to the conventions of the applicable nomenclature code) is called the scientific name authorship.

Why it is Useful: Scientific names may seem intimidating and cumbersome for some to learn, but they have a decided advantage over common names. They provide a universal code or language for naming species, so people all over the world will use the same name. Even when the name changes due to new discoveries or new information, a trail of synonyms or conserved scientific names is retained, so the species can still be identified. Scientific names also show how groups of species are related.

Core Element: Kingdom, Family, and Genus are required elements, species is highly recommended (and required if known), subspecies or variety as appropriate or known. Authorship is optional.

Coding: Enter as in field guide, ITIS (https://www.itis.gov), USDA PLANTS Database http://plants.usda.gov, or other reputable or published source.

Data Value: Text

Example: Scientific name for yellow star thistle: *Centaurea solstitialis* L.

Kingdom: Plantae **Family:** Asteraceae **Genus:** Centaurea **Species:** solstitialis **Authorship:** L.

Common Name

Field Name: Common name

Definition: The species name expressed in the common language(s) of the country and/or region, which is generally English, but may also be another language in North America. Common names are not capitalized unless they are based on a place or a person's name.

Why it is Useful: These are the names most commonly used in conversation. They are often descriptive, like Asian carp, and are always in the spoken language(s) of the country. The common names are easy to pronounce and remember. The common name cannot be used alone because there may be several common names for the same species, and/or, the same name may refer to several different species.

Core Element: Optional

Coding: Enter common name from field guide or another reputable source.

Data Value: Text

Example: For *Centaurea stoebe* L. ssp. *micranthos* (Gugler) Hayek, you would record:

Common Name: spotted knapweed

Taxonomic Serial Number

Field Name: TSN

Definition: The Taxonomic Serial Number from the Integrated Taxonomic Information System.

Why it is Useful: This is a universal identifier for species' names; it is helpful for database matching of species. The ITIS database also contains synonyms and older names. This is helpful especially when sharing historic data or between databases, so that names are easily referenced, checked for validity, and are made current while not eliminating the provided scientific name.

Core Element: Optional, but highly recommended

Coding: The codes used in the Integrated Taxonomic Information System can be found at: https://www.itis.gov.

Data Value: Numeric

Example: The TSN for zebra mussel, *Dreissena polymorpha*, would be as follows:

TSN: 81339

Location Data

Country

Field Name: Country Code

Definition: The abbreviation for the nation or country in which the infestation is located. Separate records or mapping polygons will be created for infestations that cross international boundaries.

Why it is Useful: This information facilitates free exchange of information across international boundaries. Information can be separated or summed based on national affiliations. Statistics on infested area of an individual species within and/or across country borders can easily be obtained.

Core Element: Required

Coding: ISO 3166-1 alpha-2 is a two-character code for the country. These are the same as country-level internet domain names or postal codes.

Data Value: Text

Example: An African rue (*Penganum harmala* L.) infestation was found on the Sonoran Desert in northern Mexico. The information would be entered as follows:

Country: MX

State OR Province

Field Name: State Code OR Province Code

Definition: The two-letter code for the next largest political division below country (state, province, territory, etc.) where the infestation is located.

Why it is Useful: This allows the infestation to be located within a geographic area. It also allows the easy and quick summation of information on invasive species at a political boundary below the country level.

Core Element: This is a required field.

Coding: ISO 3166-2 is coding within the 3166-1 system that denotes the principal subdivision within

a country. This field uses the two-character code after the country code, which is also the standard postal code.

Data Value: Required

Example: Dalmatian Toadflax (*Linaria dalmatica* L.) was found in Vancouver, British Columbia.

Province: BC

County OR Municipality

Field Name: County OR Municipality

Definition: The county, Parish, Borough (US, Mexico, and Canada) or municipality (Canada) where infestation is located.

Why it is Useful: Allows the infestation to be located in a sub-state/province political boundary area. It also allows the easy and quick summation of information on invasive species at the county, parish, borough, or municipality level.

Core Element: Required

Coding: This is a text-based field to note the county or municipality of the infestation. There is no standardization across North America for coding of these legal designations, but it is suggested that the words County, Parish, Borough, or Municipality not be included.

Data Value: Text

Example: There is an infestation in Humboldt County, Nevada.

County: Humboldt

Location

Field Name: Location

Definition: A text based description of the place of the occurrence.

Why it is Useful: Location information is essential for invasive species mapping. It allows invasive species sites to be located when exact location coordinates are not obtainable or recorded.

Core Element: Location is an optional field, but highly recommended when coordinates or other spatial information are not recorded.

Coding: Enter description as in data record unless field size will not allow full description.

Data Value: Text

Example: Zebra mussels were observed in Lake Powell on a boat dock on the east side of the lake.

Location: Dock on east side of Lake Powell.

Coordinates

Field Name(s): Latitude, Longitude, and Datum

Definition: The exact place of the location of an infestation (in decimal degrees), which will refer to the center of the infestation, or the center of the polygon that defines it. The field Centroid Type will be used to denote the centroid status of the coordinates. The Datum, the geodetic data system that the coordinates are based on, is required and can be found on the GPS unit, smartphone, or other technology used for recording the coordinates. If the geodetic Datum is not known, enter Unknown.

Why it is Useful: Location information is essential for invasive species mapping. It allows invasive species sites to be located on a map, be plotted across landscapes, and allows users to relocate a site.

Core Element: Coordinates are a required, if available, field. If polylines or polygons are mapped with the report, the user must also use the Data Type field, note the spatial type there, and enter the center location information for the center of the polygon or polyline as Latitude, Longitude, and Datum.

Coding: Enter the Latitude and Longitude fields as numbers in decimal degrees, with negative signs used where appropriate. Enter the Datum field as text.

Data Value: Latitude and Longitude: Numeric, with negative symbol allowed; Datum: Text.

Example:

Location: Latitude: 42.608897 **Longitude:** -114.332635 **Datum:** WGS84

Geographic Well-Known Text

Field Name: WKT

Definition: A text representation of the exact geographic shape of the infestation. This can be provided by the spatial reference system or software used to document the infested area (e.g. ArcGIS, SQL Spatial, etc.). WKT offers a compact machine- and human-readable representation of geometric objects.

Why it is useful: The Well-Known Text can be shared easily and is interchangeably usable in a variety of spatial software to draw the shape of the infestation.

Core Element: This is a recommended field for sharing records with polygons and linestrings. Datum must be included to be able to interpret the WKT.

Coding: Software should generate the WKT in the ISO 19162:2015 (WKT 2) format.

Data Value: Alphanumeric

Example:

WKT: POINT (-84.306466 32.565250)

Centroid Type

Field Name(s): Centroid type

Definition: A descriptor for records with non-point spatial information. To clarify that any occurrence coordinates representing the center of a non-point spatial shape is a centroid, Centroid Type must be included in the data set.

Why it is Useful: Subsequent users will be notified that the coordinates represent the center of a larger shape as opposed to an exact point.

Core Element: This field is required if coordinates represent a centroid. If there is no value provided, Latitude and Longitude values are assumed to be point locations.

Coding: If coordinates represent an actual point, this field can be ignored or the field left empty. If coordinates are the centroid of a non-point shape, enter a descriptor for the type of shape, such as 0.1 Degree, City, County, HUC8, Zipcode, Unknown, etc. in this field for those records.

Data Value: Alphanumeric

Example: Location is a polygon of a pond infested with hydrilla. Latitude, Longitude, and Datum were noted in the appropriate fields representing the centroid of the polygon

Centroid Type: HUC12

Data Type

Field Name: Data Type

Definition: The nature of the geographic mapped shape of the infestation (not the same as Area Surveyed). Shapes can be described as Point, Polyline, or Polygons by the mapping program used. Point records will have one definite and specific Latitude and Longitude in the decimal degrees format. Polyline records have a series of latitude and longitudes with a defined beginning and end that do NOT connect (e.g. rivers, trails, fencerows, etc.). Polygon records can be any shape with a definite interior area bound by "sides" (e.g. rectangle, circle, etc.).

Why it is Useful: Allows for documentation of the kind of shape of an infestation, which provides more context for other fields, such as Infested Area.

Core Element: This is a required field if the provided location for the record is a polyline or polygon. If not provided, locational information is assumed to be a point.

Coding: Point, Polyline, or Polygon are the available options.

Data Value: Text

Example: Eastern red cedar growing in the fencerows of pastures

Data Type: Polyline

Coordinate Uncertainty in Meters

Field Name: Coordinate Uncertainty

Definition: The variability of a pair of latitude and longitude values, in meters. Due to availability of satellites and surrounding geography, there may be some variability in the accuracy of coordinates. Often, smartphones, GPS-enabled tablets, and GPS units record this information automatically. The term provides an estimate of the number of units (coordinate uncertainty) of distance in meters within which the actual occurrence may be found.

Why it is Useful: Allows for the documentation of the accuracy of a location.

Core Element: Optional

Coding: Do not include 'meters' because it is understood. Enter number value.

Data Value: Numerical.

Example: A GPS unit with a coordinate uncertainty of 10 meters was used to plot the latitude and longitude in decimal degrees of an infestation.

Coordinate Uncertainty: 10

Georeferenced Protocol

Field Name: Source of Location

Definition: How the Latitude and Longitude coordinates of the observation were determined. Describes process by which survey was conducted (aerial survey, ground survey, helicopter, etc.).

Why it is useful: Knowledge of how the location information was collected can help data reviewers evaluate the technology used, and track changes in data collecting.

Core Element: Required

Coding: Enter how the location was recorded.

Data Value: Text

Example: Observation was recorded on a ground survey with a smartphone application.

Source of Location: Ground Survey, Smartphone GPS

Ecosystem

Field Name: Ecosystem

Definition: A descriptive term for the environment where the subject was observed. Determination of ecosystem type can include vegetation, soil-type, climate/weather patterns, management, etc.

Why it is useful: Knowledge of ecosystem where species are currently found can be utilized in future

surveying and modeling of potential species spread. Documentation of ecosystem can also aid in identification of species, as many species will only infest areas that are similar to their native habitat.

Core Element: Optional

Coding: Enter brief ecosystem description, preferably from an established system with a controlled vocabulary.

Data Value: Text

Example: Observation occurred in an area surrounded by pine trees.

Ecosystem: Conifer Forest

Record Status

Occurrence Status

Field Name: Occurrence Status

Definition: Whether or not a species is found at a location during a survey. Invasive species can be mapped as a Detected or Undetected. Detected indicates that the species was found during the survey. Undetected indicates that the species was surveyed for and not found; this covers instances where the organism was previously found at the location but not detected at time of this survey (disappeared/not detected), instances where the organism was previously treated and has not reoccurred, and instances where it has not yet been found.

Why it is useful: Detected will allow for surveys with positive results to be documented. Undetected will allow for surveys with negative results to be documented, which allows for more precise identification of when a range expansion occurs, and when monitoring areas where the species was historically found, but that could be reinvaded due to a history of known invasion.

Core Element: Required

Coding: A text field with options: Detected or Undetected.

Data Value: Text

Example: Silver carp (*Hypophthalmichthys molitrix*) was surveyed for in the Suwannee River within Woods Ferry Conservation Area and not found at time of the survey, but has been known to occur for the last several years and has been determined to be self-sustaining.

Occurrence Status: Undetected

Population Status

Field Name: Population Status

Definition: Whether or not a group of a species is observed in an area. Populations can be mapped as: Recently Introduced, Established, Extirpated, Failed, or Unknown. Recently Introduced indicates an occurrence that is not (yet) self-sustaining. Established indicates that the population is self-sustaining or is constant or continuous due to repeated introduction. Extirpated is defined as a local extinction of a population; in this document it is expected to be due to human intervention (e.g., pesticide/mechanical treatments, biological control agent releases, trapping, etc.). Failed indicates that the population is no longer present and did not successfully establish due to inability to reproduce or influence by other species (e.g., predation, competition, parasitism, etc.). Unknown indicates that the current status of the population is not known or undetermined at time of survey.

Why it is useful: This will allow for the overall population status to be documented independent of the physical presence at time of survey.

Core Element: Recommended

Coding: A text field with these options: Extirpated, Established, Failed, Unknown.

Data Value: Text

Example: Silver carp (*Hypophthalmichthys molitrix*) was surveyed for in the Suwannee River within Woods Ferry Conservation Area and not found at time of the survey, but has been known to occur for the last several years and has been determined to be self-sustaining.

Population Status: Established

Management Status

Field Name: Management Status

Definition: Allows the collector to denote if the organism or population, at time of observation, is

being managed or if control efforts are being conducted. Management Status can be mapped as Newly Discovered and Treated, Newly Discovered and Untreated, Untreated, Previously Treated, Treated, Possibly Eradicated, Not Found, or Unknown. Newly Discovered indicates a first occurrence at the site, which is either treated or untreated. Untreated indicates that the organism or population have not been and are not treated. Previously Treated means some form of treatment has occurred in the past. Treated means treatment is applied at time of survey (e.g., chemical, mechanical, etc.). Possibly Eradicated indicates prior treatment may have eliminated the population because it was not found. Not Found means the organism was searched for but was not present at time of survey, and no previous treatment is known. Unknown means status of prior treatments is unknown at time of survey.

Why it is useful: This will allow for subsequent users of the data to know if the population is being monitored or treated. Users could filter for the exact data that they are interested in for their needs.

Core Element: Required

Coding: A text field with these options: Newly Discovered and Treated, Newly Discovered and Untreated, Untreated, Previously Treated, Treated, Possibly Eradicated, Not Found, or Unknown.

Data Value: Text

Example: Cogongrass (*Imperata cylindrica*) was sprayed with glyphosate on April 5, 2013.

Management Status: Treated

Basis of Record

Field Name: Record Basis

Definition: Many reports come from in-field observation, but new technology is allowing for other methods of observation and data collection. This field, which comes from the Darwin Core Metadata Standard, allows the collector to document how the occurrence was determined. Basis of Record can be mapped as Human Observation, Living Specimen, Machine Observation, Moving Image, Physical Object, Preserved Specimen, Sound, Still Image, or Unknown. Human Observation means the record was collected by a person viewing the organism. Living Specimen means the organism is within a collection such as a zoo, bacterial culture, or botanical garden. Machine Observation means the data was collected by an aerial photo, animal/motion detection camera, remote sensing, image recognition, DNA sensor, or similar. Moving Image means the record is a video or similar. Physical Object means the record consists of a material sample (or resample) which may be either preserved or destructively processed (e.g. for DNA analysis). Preserved Specimen is a part or whole of the organism that is preserved and included as part of a formal collection. Sound is a record of an audio recording (such as a bird or insect call). Still Image is a visual record of all or part of the organism. Unknown is only included for purposes of historic datasets, because this is a required field.

Why it is useful: This will allow for subsequent users of the data to know how the occurrence was determined. Users could filter for the exact kind of data that they are interested in for their needs.

Core Element: Required

Coding: A text field with these options: Human Observation, Living Specimen, Machine Observation, Moving Image, Physical Object, Preserved Specimen, Sound, Still Image, or Unknown.

Data Value: Text

Example: Spotted wing drosophila (*Drosophila suzukii*) was surveyed by setting traps and counting captured individuals which were not collected/preserved.

Record Basis: Human Observation

Record Type

Field Name: Record Type

Definition: Conveys the purpose of the record. This field helps to describes the broad scope of data that could be collected using these standards, especially since more than invasive species occurrences may be mapped with these standards.

Why it is useful: This will allow for subsequent users to know the purpose of the record and the reason the data is being collected. Users could filter for the exact data that they are interested in for their needs.

Core Element: Required

Coding: Select from the following options: Biological Control Agent Survey, Biological Control Agent Release, Invasive Species Survey, Invasive Species Treatment, Monitoring, Other, Unknown.

Data Value: Text

Example: Examiners released black dot spurge flea beetle (*Aphthona nigriscutis*) as a biological control agent on leaf spurge (*Euphorbia esula*).

Record Type: Biological Control Agent Release

Data Collection Method

Field Name: Method

Definition: A descriptor for the methodology by which the data was collected. This covers data collected by field surveys, including the survey method, but also data collected by other methods, such as aerial surveys, sketch mapping, animal camera, LiDar analysis, etc.

Why it is Useful: Noting how data is collected lets the subsequent users include desirable or eliminate undesirable data methods and provides the verifier with additional information.

Core Element: Optional

Coding: Enter the methodology used to take the data, preferably choosing terms from a list.

Data Value: Text.

Example:

Method: Digital Aerial Sketch Mapping

Verification Method

Field Name: Verification Method

Definition: How the species was identified or verified. This can include everything from expertise based on education and training to photographs, samples, or laboratory procedures such as DNA testing.

Why it is useful: Due to the inability for everyone to inspect every report of invasive species, documenting how the species was identified lends additional credibility to the report. Noting the method used provides more information especially for outliers, first in region, and difficult to identify species.

Core Element: Recommended

Coding: Text based field due to the number of ways to identify invasive species, but suggested values are: Expertise, Photographs, DNA Testing, Diagnostic Laboratory Sampling, and Specimen

Data Value: Text

Example: Identification of kudzu bug (*Megacopta cribraria* (Fabricius)) based on the reporter's knowledge of species' unique characteristics.

Verification Method: Expertise

Information Source

Reference

Field Name: Reference

Definition: A citation or descriptive method of finding the source of literature that data comes from.

Why it is useful: This will allow for subsequent users of the data to know the source of the data and examine it, if needed. Users could filter for the exact data that they are interested in for their needs.

Core Element: This is a recommended field for literature based data.

Coding: An open text field that follows any widely-accepted and consistent format for citations.

Data Value: Text

Example: Canada thistle (*Cirsium arvense*) in Philips County, Montana was documented in *Flora of Montana*.

Reference: Booth, W.E., and J.C. Wright. 1966. Flora of Montana. Montana State University, Bozeman

Examiner

Field Name: Examiner

Definition: The full name of the individual who collected the information in the field or is the contact person for the data. Also known as Collector, Observer, or Recorded By.

Why it is Useful: Within a management area, many different public and private sector individuals may all have contributed to the survey. A name allows the person compiling the inventory to serve as a contact person and to verify and correct any questions on the information.

Core Element: Optional, though it may be useful at the field office level or to estimate how many people have contributed to the dataset over time.

Coding: Enter the full name of the individual who collected the data.

Data Value: Text

Example:

Examiner: Ronald J. Weed

Data Source

Field Name: Data Source

Definition: The owner or manager of the data (either a person of an organization). This may be a different person or entity from the owner, or the person who collected the data. It may be an office manager or a database specialist. This is the entity responsible for answering questions about the data, or for responding to data requests.

Why it is Useful: Provides contact point for questions about data. Allows consolidation/coordination of information requests. Bridges gap between those collecting information, and those managing the data.

Core Element: Required

Coding: Names are written out with no acronyms or abbreviation, because those may vary.

Data Value: Text

Example: Banff National Park in Alberta, Canada has been mapping invasive plants.

Data Source: Parks Canada, Banff National Park

Date and Time

Collection Date

Field Name: Collection Date

Definition: The date the observation was made in the field. It does not refer to the date information

was entered into the computer (that can be an optional additional field, often 'Last Update' or similar. If the date is estimated, choose a date at the mid-point of the estimated date range and include the accuracy of the selected date in the Date Accuracy in Days field.

Why it is useful: This field tells you when the occurrence was observed. Phenology and morphology can change over the course of the year, so notation of observation/collection date is important. This field also tells you how old the information is. These cues will help you decide how reliable the information is and whether a follow-up visit to the site may be warranted.

Core Element: Required

Coding: A date field; because data is liable to be shared internationally, ISO 8601 is recommended. Increasingly, data is being collected directly in GPS units, smartphone applications, online forms, and other data logging technology. Date formatting can be established within the logging units and forms without having to train collectors in a new format. The ISO 8601 format for reporting dates with time is yyyy-mm-ddThh:mm:ss.ffffff, where yyyy=4-digit year, mm=2-digit month, dd=2-digit day, T=time value follows, hh=2-digit hour, mm=2-digit minute, ss=2-digit second, and ffffff=fraction of a second with up to six digits of accuracy. Time should be standardized to local time zone of collection. For data without time known coding is yyyy-mm-dd.

Data Value: Alphanumeric

Example: A knapweed site was visited on October 3, 2002. You would record:

Collection Date: 2002-10-03

Date Accuracy in Days

Field Name: Date Accuracy in Days

Definition: The range of days within which the data was collected. The exact date may not always be obtainable, so collectors can mark the accuracy of the observation date with the number of days around the noted observation date the observation could have been made.

Why it is useful: This allows subsequent users to know that the observation data is not exact, but an estimation. As date is sometimes used in verification, an out of place date could cause concern without the additional information that the observation date is an estimation.

Core Element: Optional

Coding: A number that represents the estimated accuracy to the nearest day (within the month = 15;

within the year = 182).

Data Value: Numeric

Example: Survey was conducted during an undefined period in June 2015, Collected Date is recorded as June 15, 2015 (2015-06-15) and accuracy is ± 15 days:

Date Accuracy in Days: 15

Quantifying the Infestation

Infested Area

Field Name(s): Infested Area, Infested Area Units

Definition: The calculated area containing one invasive species type. An infested area is defined by drawing a line around the actual perimeter of the infestation. It is highly recommended that only a single species be entered for each infested area. Different areas and programs will use different units for area, so it is important to know which unit of measure (Infested Area Units) was used to measure the infestation.

Why it is Useful: Infested area can be defined in many ways and there is little consistency between individuals, counties, states, and countries. Is an acre of weeds: one weed plant in an acre, an acre covered with weeds, or all the lands threatened with invasion from an existing infestation? This definition provides a consistent and common method of describing infestations; the calculated perimeter of one species' infestation. This is the data field that will be used to sum and report total infested area across all ownerships.

Core Element: Both Infested Area and Infested Area Units are required fields for invasive plant species that were detected.

Coding: Infested Area: Enter the number that correctly reflects the area infested. Infested Area Units is selected from the following options: Square Feet, Acres, Square Meters, Square Kilometers, Hectares, Other.

Data Value: Infested Area: Numeric; Infested Area Units: Text.

Example: 1.6 hectares of oxeye daisy (*Leucanthemum vulgare Lam.*) found outside Vancouver, BC.

Infested Area: 1.6 **Infested Area Units:** Hectares

Area Surveyed

Field Name: Area Surveyed, Area Surveyed Units

Definition: The entire land area surveyed for an invasive species, whether or not invasive species were found. Information will be recorded in two fields, Area Surveyed and Area Surveyed Units.

Why it is Useful: These fields record information on the extent, or the total area that was surveyed. It allows landowners and land managers to maintain records on the areas that have been surveyed for invasive species, and those areas where no surveys have occurred. This field can be used in tandem with Report Status to denote a specific area surveyed for an invasive species that was (detected) or was not (undetected) found.

Core Element: Optional

Coding: Area Surveyed: Enter number that correctly reflects the area surveyed. Area Surveyed Units: Enter as appropriate: Square Feet, Acres, Square Meters, Square Kilometers, Hectares, or Other.

Data Value: Area Surveyed: Numeric; Area Surveyed Units: Text.

Example: In the summer of 2000, Jasper National Park completed a 17-hectare survey in and around park headquarters facilities for Canada thistle (*Cirsium arvense*), which was found sporadically throughout the survey area.

Area Surveyed: 17 **Area Surveyed Units:** Hectares

Incidence

Field Name: Incidence

Definition: The proportion (as a percentage) of the host community that is affected by an insect or pathogen, or the proportion of an area that is infested by independent living organisms (e.g., plants, wildlife, etc.). The Area Surveyed would provide context for the area considered and the Incidence notes the percentage within is infested/infected.

Why it is Useful: These fields record information on the extent of the spread of an insect or pathogen within a population or an area, or describes the extent of an infestation within an Area Surveyed.

Core Element: This is an optional field.

Coding: Incidence is considered a percentage of the whole Area Surveyed and is coded as a number not to exceed a maximum value of 100.

Data Value: Numeric

Example: An area within Paynes Prairie State Preserve was evaluated for laurel wilt (*Raffaelea lauricola*) on red bay trees (*Persea borbonia*) and the area had 20 percent of the trees affected.

Incidence: 20

Severity

Field Name: Severity, Severity Units

Definition: The proportion (percentage) of an individual sample (i.e. one plant, one leaf, one animal, etc.) that is affected by a pathogen. Incidence and Severity are independent from each other. A pathogen may be wide spread within a population (a high Incidence), but may not be very impactful to the individual plants/sample (a low Severity). Or, the converse may be true, only a few plants affected (low Incidence) but the plants are killed (high Severity). This data can be used to describe "hot spots" for areas of targeted management (low Incidence/high severity).

Why it is Useful: These fields (taken together) provide context for how injurious a pathogen is to the individual sample.

Core Element: Optional

Coding: Severity can be marked as a percentage or as a number relative to a sampling amount. Severity Units it describes the methodology of either the percent or the number of the Severity.

Data Value: Severity: Numeric; Severity Units: Text.

Example: Twenty leaves of one red bay tree (*Persea borbonia*) with laurel wilt (*Raffaelea lauricola*) were assessed for the amount of disease signs and 10 leaves of 20 were showing signs.

Severity: 50 **Severity Units:** percent of sampled leaves from one plant

---

Organism Quantity

Field Name(s): Organism Quantity, Organism Quantity Units

Definition: A count of subjects observed, captured, treated, or released (for biological control agents) within the surveyed/infested area (Organism Quantity); a descriptor of what kind of quantity was measured (Organism Quantity Units). Enter an exact or estimated number of observed subjects. The Quantity Units can be a description of the life stage of the subject (e.g., adults, eggs, nymphs, etc.), or if more practical, description of the organism or object that the subjects were observed on (e.g., trees, dock posts, bridge piers, etc.).

Why it is useful: A count of the species observed is more appropriate for invasive species populations that may migrate or experience population changes quickly.

Core Element: This is a required field for insect, animal, plant pathogen, and biocontrol surveys.

Coding: Enter an exact or estimated number of observed subjects and description of the organism life stage or description of what the subject was observed on.

Data Value: Organism Quantity: Numeric; Quantity Units: Text

Example: Observed 50 hemlock trees infested with hemlock wooly adelgid

Organism Quantity: 50 **Organism Quantity Units:** Trees

Subject Characteristics

Life Stage

Field Name: Life Stage

Definition: A brief description of the phenological or life stage of development of the plants, animals, insects, biocontrol agents, or pathogens observed. Individuals or entire infestations can express multiple life stages at once, so multiple options can be used.

Why it is useful: The life stage at time of observation can be important in identification and management of the infestation. Notation of the phenological aspects or life stage of subjects observed, e.g. nymph, mature, adult, eggs, in flower, in fruit, seedling, etc., can be utilized, in combination with other data (e.g. the date of observation, host species, etc.) by future surveyors, can help determine appropriate treatment, or can be incorporated into modeling programs and aid in determining the timing of the life cycle for that species in a particular environment.

Core Element: Optional

Coding: Using the singular, enter a description of the observed subject from a list of options with multiple stages separated by a comma; seed, seedling/rosette, sapling/immature, mature, in flower, in fruit, dormant/dead, egg, nymph, larva, pupa, juvenile, adult, or other.

Data Value: Text

Example: Several life stages of emerald ash borers were found infesting ash trees.

Life Stage: Adult, Larva, Egg

Sex

Field Name: Sex

Definition: A brief description of the sex of the organism observed. For one organism observed, one option may be chosen. Individual species or entire populations can express multiple sexes at once, so multiple options can be used.

Why it is useful: The sexes observed can be utilized by data managers to note trends or frequency of the sex of observed subjects. Some species have different coloration or patterns between males and females; identification of some species is therefore easier with one gender over the other. This information can be utilized in combination with other fields to lend certainty to identification.

Core Element: Optional

Coding: Enter a description of the sex of the observed organism(s): male, sterile male, female, sterile female, hermaphrodite, asexual, other, or unknown.

Data Value: Text

Example: The observed wild boar was male.

Sex: Male

Host Species

Field Name: Host Genus, Host Species, Host Subspecies OR Host Variety, Host Authorship

Definition: Scientific name of the host of the observed subject (infestations or diseases or biological

control agent). Include genus, species, subspecies/variety (if applicable), and the authorship (if applicable).

Why it is useful: The use of the scientific name provides a definite label to the host species. Knowledge of the host for biocontrol agents, insects, and pathogens can aid in positive identification for the observed species. It can also influence areas searched in future surveys, and help to predict a potential range of spread for pathogens and biocontrol agents.

Core Element: This is a required field for biocontrol observations, invasive insects, and pathogens, unless collected in traps or otherwise unassociated with host. Host genus and species (if known) are required. Subspecies/variety and authorship are optional.

Coding: Enter name of species as it appears in a reliable taxonomic authority, such as ITIS.

Data Value: Text

Example: For a biocontrol agent found on Dalmatian Toadflax: *Linaria dalmatica* ssp. *dalmatica* (L.) Mill. you would record:

Host Genus: Linaria **HostSpecies**: dalmatica
Host Subspecies: dalmatica **Host Authorship: (**L.) Mill.

Comments

Field Name: Comments

Definition: Additional notes to be included with the record to document important information which cannot be assigned to another field.

Why it is useful: It is not possible for all of the important data to be accounted for in a standard, so comments allow users to include additional data that does not have to conform to the parameters described in an existing field.

Core Element: Recommended

Coding: This field should be able to accommodate open ended, longer entries of data. It is recommended that unrelated comments for one record be separated by periods. Do not use tab characters within a comment (or any other) field in the dataset.

Data Value: Alphanumeric.

Example:

Comments: Population appears to have some unknown insect herbivore feeding on it; revisit with supplies to sample insect population.

Glossary of Terms

Area Surveyed: Intended to show general location and population information. It is an area that was surveyed for specific species. Unlike *Infested Area,* the area is defined by drawing a line around the general perimeter of the survey, not the area covered by individual or groups of invasive species. Area surveyed may contain significant parcels of land that are not occupied by the target organism.

Collection Date: The date the record's information was collected in the field. It does **not** refer to the date the record's information was entered into the computer.

Datum: A model of the earth's shape. Geodetic datums define the size and shape of the earth and the origin and orientation of the coordinate system used to map the earth. The most common Datums used will be NAD83 and WGS84, but there are older and more locally specific systems which may have been or be used.

Georeference: Defining the location of an object using coordinates and assigning

GIS (Geographic Information System): A computerized system for the collection, storage, management, retrieval, changing, modeling, analysis and display of spatial data, used to create a representation of the real world.

GPS (Global Positioning System): A global navigation system based on a system of high orbiting satellites. The GPS receiver uses at least 4 satellites to compute position.

Infestation: The presence of an unusually large number of invasive species in a place, typically enough to cause damage or disease.

Infested Area: Calculated or estimated area containing one invasive species. Defined by drawing a line around the actual perimeter of the infestation as defined by the area covered by the target organism, excluding areas not infested. An area containing only occasional target organisms per acre does **not** equal one acre infested. Generally, the smallest area of infestation mapped will be 1/10th (.10) of an acre (or 0.04 hectares).

Latitude: The angular distance (distance measured in degrees) north or south of the equator. Latitude is 0 degrees at the equator, 90 degrees at the north pole, and –90 degrees at the south pole. Latitude is also described by N or S (direction north or south of the equator) instead of + or -.

Longitude: The angular distance (distance measured in degrees) east or west of the prime meridian. Longitude is 00 at the prime meridian, and is measured + 180 going east, and –180 going west. Longitude is also described by E or W (direction east or west) of the prime meridian, instead of + or -.

Persistent Identifier: A globally unique identifier that is a reference to a digital object. They are often actionable, allowing a user to accesses the original digital resource using a long-term, persistent link through an internet browser. There are many formats for Persistent Identifiers (PIDs), the most

common of which is Digital Object Identifiers (DOI), but also include Uniform Resource Identifier (URI), of which Uniform Resource Locator (URL) and Uniform Resource Name (URN) are two specializations of URI, and Uniform Resource Characteristic (URC).

Universally Unique Identifier: Following the version 4 UUID format of 32 alphanumeric characters in five groups of characters separated by hyphens in 8-4-4-4-12 form (a total of 36 characters), the probability of a duplicated UUID is negligible. A UUID is used to ensure that an individual record can be positively identified within a system or across systems and to help prevent duplications when aggregating data across many datasets and between aggregate databases. UUID (called Global Unique Identifier [GUID] in Microsoft products) can be generated in many software and database programs, as well as through UUID generator websites.

Well-Known Text: Well-Known Text (WKT) is a text based representation of the geometry of objects on a map and should use the ISO 19162:2015 (WKT 2) format. WKT can be used to represent points, polygons, linestrings, and more. Most common mapping programs can accommodate WKT and draw a geometric shape from the text, which can be a way for groups using different or proprietary mapping software to share data.

Technical Resources

Persistent Identifier – PID can be any number of ways that a record is resolvable to a definite source. This can include Digital Object Identifiers (DOI) links to an individual resource (e.g., journal article, press release, data set, etc.) that describes the record, a website URL that links to the original record on a state, county, or other website (such as EDDMapS, iNaturalist, GISIN, etc.), and more. DOI are under the domain and maintenance responsibility of specific registering agencies, which can be accessed via https://www.doi.org/registration_agencies.html. URL may be more inconsistent, as those are reliant upon the domain holder for upkeep and so broken links and dead-ends may be more of an issue.

Universally Unique Identifier – UUID can be generated in a variety of programs and databases, but this process can differ from version to version as software is updated. Below are some websites with instructions on how to generate UUID for each common program used to collect and/or store invasive species data.

Excel: https://www.idigbio.org/wiki/images/0/03/GUIDgeneration.pdf
ArcGIS: http://moravec.net/2015/08/5-ways-to-make-a-guid-or-uuid-in-arcgis/
SQL: https://docs.microsoft.com/en-us/sql/t-sql/functions/newid-transact-sql?view=sql-server-2017

As there may be some instances where a program is using another software or may be less technologically inclined, there are many websites that generate UUID and those can been added to files and databases as data is entered.

https://www.guidgenerator.com/
https://www.uuidgenerator.net/version4