eBird Basic Dataset Metadata (v1.16)

revised 19 Feb 2025

1.16 updates: The EBD file has undergone a MAJOR revision that will be relevant to anyone who has scripts or codes using the EBD column headers. Please review the below carefully.

*First, **OBSERVER ORCID ID** is added as a new field following OBSERVER ID; OBSERVER ID section is edited slightly for clarity and the PROJECT CODE section is substantially revised.*

Second, there are significant changes to fields related to projects and protocols as follows:

1) *PROTOCOL TYPE is renamed **OBSERVATION TYPE**. It is meant to be the overarching description of the sampling event and how birds were surveyed. It has 7 possible values for modern data: Stationary, Traveling, Incidental, Historical, Banding, eBird Pelagic Protocol, and Nocturnal flight Call Count. In addition, Area, which applies to some data before Mar 2025 is also included here; after Feb 2025, this protocol will be project-specific, but still supported in those instances.*

2) ***PROTOCOL NAME** is added as a new field. A significantly reduced list of supported protocols is available in a new metadata document (protocols.txt) that describes these in full. The seven core Observation Types (see above) are also protocols, but if additional metadata is collected by an eBird Project, a distinct protocol name and code will be used.*

3) ***PROTOCOL CODE** is unchanged, but the list of protocols supported in eBird is greatly reduced (many of the rarer protocols are deprecated or moved to a project)*

4) *The **PROJECT CODE** column is removed, since the term "project" has a new and specific meaning in eBird.*

5) ***PROJECT NAMES** is added to describe the name(s) for eBird Projects. See more about eBird Projects below.*

6) ***PROJECT IDENTIFIERS** is the internal unique identifier(s) for each project; it is also used in the URLs describing projects.*

*Third, TRIP COMMENTS is renamed **CHECKLIST COMMENTS**. Since eBird now has [Trip Reports](), this change to the column name helps clarify this and is more consistent with usage elsewhere in eBird.*

In addition to the major changes above, this document has undergone a fairly substantial update. A new KEY TERMS section is added near the top of the document.

The section on handling Sampling Event Data is updated, and expanded for clarity. The BEHAVIOR CODE and REASON sections are substantially revised to improve accuracy and clarity, and five key taxonomic fields (TAXON CONCEPT ID to SUBSPECIES SCIENTIFIC NAME) are clarified with a final sentence that explains how infraspecific options are treated. The RECORDS "IN REVIEW" section is renamed UNREVIEWED RECORDS and is explained in a bit more depth. The OBSERVATION COUNT section clarifies how to handle cases with multiple infraspecific (i.e., subspecies group) entries or a checklist with a parent species and one or more entries from a subspecies group. Appendix 1 had a couple errors in Breeding Code definitions, including an erroneous definition for code "T", now fixed. Appendix 2 is substantially revised and Appendix 3 is deleted, again, because of the new support for Projects and the consequent changes in how protocols and observation types are defined. Additional minor changes include updates to the protocols section, deletion of mention of Casual or Casual Observation (now an "OBSERVATION TYPE") which has long been referred to as Incidental, inclusion of KBAs in the IBA CODE section, a caution added within the SPECIES COMMENTS and CHECKLIST COMMENTS sections, and miscellaneous minor edits for clarity.

CHARACTER SET – eBird is a global database, with a need to display characters in a variety of languages. We use the Unicode UTF-8 character set to accomplish this. When opening this file in Excel or any text editor, you must first specify the proper character set to ensure that all characters display correctly. In Excel, this process means using the 'Text import wizard', and then setting the file origin field to Unicode UTF-8 before opening the file.

KEY TERMS – The following terms are used frequently and are defined here:
- **Taxon** (plural Taxa): unique taxon of bird as defined by the eBird/Clements taxonomy; each has a CATEGORY and a unique TAXONOMIC ORDER, TAXON CONCEPT ID, COMMON NAME, SCIENTIFIC NAME and (internal) species_code.
- **Observation**: Defined by a unique OBSERVATION ID, an internal code that appears as the last element in the GLOBAL UNIQUE IDENTIFIER. Within eBird, an Observation represents a unique submission of a taxon on a Submission (checklist). Each observation has an OBSERVATION COUNT, and sometimes additional associated metadata (e.g., BREEDING CODE, EXOTIC CODE, media (see HAS MEDIA), etc.). All Observations are vetted by eBird's data quality system (see APPROVED and REVIEWED below).
- **Submission** (or Checklist): Defined by a unique SAMPLING EVENT IDENTIFIER, and sometimes referred to as a sampling event, this represents a single data collection event (checklist) in eBird. Each Submission has an OBSERVATION DATE, LOCATION, and OBSERVATION TYPE and most have Time (TIME OBSERVATIONS STARTED) and Effort (e.g., DURATION MINUTES, EFFORT

DISTANCE KM). Many submissions also record GPS tracks, which are not currently available publicly but are used to ensure the data integrity of other metadata fields (e.g., DISTANCE, DURATION, LOCATION). Submissions are also marked as Public and Not Public by eBird's data quality system; Not Public submissions are not included in the EBD.

- **Location**: Defined by a unique name (LOCALITY) and LOCALITY ID, a location in eBird is meant to represent where a Submission took place. Each Location has a LATITUDE and LONGITUDE, and we append additional metadata by adding the COUNTRY, STATE/PROVINCE, and COUNTY based on the latitude and longitude. We also indicate if it falls in certain types of protected areas, such as IBAs, KBAs, and U.S. National wildlife Refuges; see those sections below for more information. eBird has two key types of locations: 1) Hotspots, which are shared locations used across the eBird community and are managed for accuracy in naming and latitude/longitude by a team of hotspot reviewers; 2) Personal locations, which are created, names, plotted, and maintained by individual users. Some personal locations can be shared between users (e.g., from a shared checklist) but the location always owned and managed by a single user.

TAXONOMY – We use the eBird/Clements taxonomy for the species-level taxonomy and nomenclature for eBird. Please see our eBird taxonomy page for more information on the species lists used in eBird; downloadable versions of the full taxonomy are also provided there. The eBird/Clements taxonomy is the reference source for information contained within the following fields: Taxonomic Order, Category, Common Name, Scientific Name, Subspecies Common Name, and Subspecies Scientific Name. This taxonomy is updated once annually in October to conform with the latest science; therefore, the version of the EBD including records through October (extracted in mid-November) will be the first to be revised to the latest taxonomy. Note however that the transition to the new taxonomy requires updating of large numbers of records and sometimes this process will not be fully completed until the January EBD version.

DATA INCLUDED – This dataset includes all records treated as "Accepted" in eBird, including those that have been vetted by eBird reviewers as well as those that are not flagged as unusual. See APPROVED, REVIEWED, and REASON fields below to better understand this.

UNREVIEWED RECORDS – The "Unvetted" dataset is also available. These are records flagged as unusual but not evaluated by the eBird data quality process; the records may be in the process of review or simply not yet assessed by a reviewer. These records should be used with caution and should be assessed on their merits by the analyst. Unvetted data, when requested, will be provided in a separate .txt file from the EBD and have a file name ending with "_provisional.txt".

SENSITIVE SPECIES – Certain taxa that are at risk from targeted hunting or other pressures are not included in this dataset. Observations are considered Sensitive either

globally, by country (e.g., Indonesia) or by state (e.g., Wisconsin, US). They can also be Sensitive year-round or only within a certain date range (e.g., 5 March-25 August). See the [Sensitive Species policy and list of species](#).

Sensitive Species data can be requested by emailing us at ebird@cornell.edu with the subject line written exactly "Sensitive Species Data Request". In the email please indicate the species you are interested in and a brief abstract of the proposed data use.

SAMPLING EVENT DATA

eBird's data are a richer source of information when both the presence and the absence of a taxon on a given checklist (i.e., sampling event) are taken into account (e.g., see Guillera-Arroita et al. 2015). The EBD by itself is a form of "presence-only" data, with one record for each taxon (typically species) reported. Information about non-detections ("absences") can be added by additionally downloading the Sampling Event Data (SED) that contains one row for each eBird checklist included in your query of observations, including the date, time, location, and the amount of search effort expended, but no information about bird species. For complete checklists–checklists for which observers indicated that all detected species were reported–if there is a record in the SED but no record of a species in the EBD or unvetted data, then a count of zero individuals of that species can be inferred, creating the non-detection/absence information. The R package auk (described below) is designed to create these non-detection/absence records. Downloading both EBD and SED is highly recommended; note that downloading the SED will automatically download the "Provisional" file as well (see UNREVIEWED RECORDS above, since the SED may include sampling events restricted to that file. Importantly, checklists that report zero species will not have any records in the EBD but will appear in the SED. Any checklist treated as "Not Public" (see APPROVED below) is excluded from both the EBD and SED.

Guillera-Arroita, G., J. J. Lahoz-Monfort, J. Elith, A. Gordon, H. Kujala, P. E. Lentini, M. A. McCarthy, R. Tingley, and B. A. Wintle (2015). Is my species distribution model fit for purpose? Matching data and models to applications. Global Ecology and Biogeography 24:276-292.

Sampling Event Data include
There are two ways to download Sampling Event Data:
1) Custom Download, accessed by checking the box in the Options section titled "Include sampling event data" under the Custom Download. This will provide the Sampling Events that match your other query parameters defined above (i.e, you must define a region, species, or date range). These have a file name ending in "sampling" like: ebd_OM_delpri1_smp_relNov-2024_sampling.

2) Global Sampling Event Data, delivered as a .tar file, includes information for all Public eBird checklists in the database and will match to the Observations in the World file under "Prepackaged Options"

RECOMMENDED R PACKAGE -- AUK
The auk R package has been developed by Cornell Lab staff and may prove useful for working with the EBD downloaded data. For more information, see: https://cornelllabofornithology.github.io/auk/

----

eBIRD BASIC DATASET FIELDS
The below list names and describes the contents of all the data fields for the eBird Basic Dataset download file.

GLOBAL UNIQUE IDENTIFIER – A unique alphanumeric code assigned to each observation (of a single taxon within a single checklist) that stays with it through database revisions, updates, and edits.

LAST EDITED DATE – The date and time of the most recent edit to any observation on the checklist (see SAMPLING EVENT IDENTIFIER below); this is useful for determining whether an observation should be updated if a copy of these data is being stored locally. Edits include changes to any of the following:
- Location: e.g., latitude, longitude, location name, county, state, or country
- Submission: e.g., date, effort, complete checklist
- Observation: e.g., species, count, review status, exotic code, or other edits made by user or editor.

TAXONOMIC ORDER – The numeric value assigned to this taxon in the eBird/Clements taxonomy to arrange the species in taxonomic sequence. These values are updated with each annual taxonomic update (e.g., for the November EBD; see TAXONOMY above) along with other revisions to the taxonomy.

CATEGORY – The category (e.g., species, hybrid, slash) assigned to this taxon in the eBird/Clements taxonomy. A full description of these categories is here: https://support.ebird.org/en/support/solutions/articles/48000837816.

TAXON CONCEPT ID – Unique taxonomic identifier meant to identify a specific taxonomic concept. The same English or scientific name may be applied to different populations over time, even across versions within a single taxonomy. A Taxon Concept ID refers to a population with shared characteristics and a specific range circumscription and not to the treatment of these populations within a specific taxonomy. For example, when well defined subspecies or subspecies groups are elevated to species level (i.e., a

split), the Taxon Concept ID will not change even though the common and scientific names will change. A change to the genus or the spelling of a scientific name will not result in a change to the Taxon Concept ID. Taxon Concept IDs are the best way to track taxonomic changes through time or across datasets; note that the TAXON CONCEPT ID applies to the most specific identification offered, so if the SUBSPECIES COMMON NAME and SUBSPECIES SCIENTIFIC NAME fields are populated, the TAXON CONCEPT ID applies to those taxa; if not, TAXON CONCEPT ID applies to the COMMON NAME and SCIENTIFIC NAME fields. We use the Taxon Concept IDs developed and maintained by Avibase (https://avibase.bsc-eoc.org/) which also maintains linkages between taxonomic authorities and versions through time. **Example:** In 2021, eBird lumped Northwestern Crow (*Corvus caurinus*) with American Crow (*Corvus brachyrhynchos*), resulting in no name change for American Crow but a substantial change in the range of American Crow, since populations in coastal British Columbia and Alaska are now included in American Crow (*Corvus brachrhynchos*). Consequently, this lump results in 1) deletion of the Northwestern Crow (*Corvus caurinus*) as a species from the dataset; 2) conversion of all those former Northwestern Crow records to American Crow (*Corvus brachyrhynchos*); 3) change in the TAXON CONCEPT ID for American Crow (*Corvus brachyrhyncho*s) from avibase-9E9F2476 to avibase-69544B59. Thus, the TAXON CONCEPT ID helps to track taxonomic changes, including those where the name might not change, such as American Crow (*Corvus brachyrhynchos*). NOTE: the TAXON CONCEPT ID refers to the exact taxon reported, not the parent species in cases where an infraspecific option is shown.

COMMON NAME – The primary English common name of the taxon in the eBird/Clements taxonomy. When a user reports an infraspecific option (categories of ISSF, DOMESTIC, or INTERGRADE; sometimes a FORM), the value for the parent species is shown here.

SCIENTIFIC NAME – The scientific name of the taxon in the eBird/Clements taxonomy. When a user reports an infraspecific option (categories of ISSF, DOMESTIC, or INTERGRADE; sometimes a FORM), the value for the parent species is shown here.

SUBSPECIES COMMON NAME – The common name of the subspecies or subspecies group in the eBird/Clements taxonomy. Certain other taxa (such as intergrades, and certain forms and domestics) also included here. When a user reports an infraspecific option (categories of ISSF, DOMESTIC, or INTERGRADE; sometimes a FORM), the exact taxon reported is shown here.

SUBSPECIES SCIENTIFIC NAME – The scientific name of the subspecies or subspecies group in the eBird/Clements taxonomy. Certain other taxa (such as intergrades, and certain forms and domestics) also included here. When a user reports an infraspecific option (categories of ISSF, DOMESTIC, or INTERGRADE; sometimes a FORM), the exact taxon reported is shown here.

EXOTIC CODE – Exotic Codes are applied to eBird Observations (i.e., any unique GLOBAL UNIQUE IDENTIFIER) when the species is believed to be exotic (i.e., non-native); observations of taxa native to the region will have a null value in this field. The three Exotic Codes are: N (Naturalized), P (Provisional), and X (Escapee). Exotic Codes are defined by eBird and applied as part of the eBird Review Process by eBird reviewers, using both an automated and manual process. See Exotic Species in eBird https://docs.google.com/document/d/1VHA1bpLI5zOk89WtbytUMfJME567SGRJUrHxm Sr7Or8/edit?usp=sharing for more information. The full definitions of these Exotic Codes are as follows:

- **N (Naturalized):** *Exotic population is self-sustaining, breeding in the wild, persisting for many years, and not maintained through ongoing releases (including vagrants from naturalized populations). These count in official eBird totals and, where applicable, have been accepted by regional bird records committee(s).*
- **P (Provisional):** *Either: 1) member of exotic population that is breeding in the wild, self-propagating, and has persisted for multiple years, but not yet* Naturalized*; 2) rarity of uncertain provenance, with natural vagrancy or captive provenance both considered plausible. When applicable, eBird generally defers to bird records committees for records formally considered to be of "uncertain provenance". Provisional species count in official eBird totals.*
- **X (Escapee):** *Exotic species known or suspected to be escaped or released, including those that have bred but don't yet fulfill the criteria for* Provisional*. Escapee exotics do not count in official eBird totals.*

OBSERVATION COUNT – The count of individual birds of a given taxon (each with unique values for TAXON CONCEPT ID, COMMON NAME, and SCIENTIFIC NAME) made at the time of observation. If no count was made, an 'X' is used to indicate presence. Note that the separate OBSERVATION COUNTs needed to be added when the same SAMPLING EVENT IDENTIFIER contains a multiple records for a species, such as when one record is of the CATEGORY of species, and one or more records are for the CATEGORY of subspecies.

BREEDING CODE – The highest-level breeding information reported for the species on a given checklist. The Breeding Code will reflect the value entered by a user except in cases where reviewers of atlas data have deemed it a likely typo or other error; in those cases, the original Breeding Code will appear on the public eBird checklist, but the corrected code is provided in the data here. For a complete list of codes and their definitions see Appendix 1.

BREEDING CATEGORY – Four categories used to describe a species' breeding status based on the 'BREEDING CODE' reported on the eBird checklist: C1 – Observed; C2 – Possible; C3 – Probable; C4 – Confirmed. In most cases, these are the default values corresponding with the breeding code reported by the observer. But in some cases reviewers of atlas data may reinterpret a breeding category, and that reinterpretation is

reported here. For instance, a tern species might be seen carrying food (typically C4 – Confirmed), but since terns feed young away from the nesting area it would be reinterpreted as a lower breeding category. For more information on reinterpretation of breeding codes, see this document: https://support.ebird.org/en/support/solutions/articles/48000837520

BEHAVIOR CODE — The highest-level behavior observed for the species on this checklist. In most cases, this will match the BREEDING CODE, differing only for official eBird Breeding Bird Atlases (BBAs) where BBA reviewers have acted on the record in a way that resulted in a difference between the observed behavior and the breeding meaning for the species on this checklist (e.g., a migrant songbird singing on migration in places where it doesn't breed might have S for BEHAVIOR CODE, but no BREEDING CODE). When BREEDING CODE and BEHAVIOR CODE differ on a given eBird checklist, the public display shows the originally reported code with a strikethrough and the new, corrected code shows as the official code on the list. For a complete list of codes and their definitions see Appendix 1.

AGE/SEX – The reported number of each age and sex combination for a species on a given checklist. Age categories are: adult, immature, and juvenile. Sex categories are: male, female, and unknown.

COUNTRY – The country where the observation was made. Follows ISO 3166-2. Country names and codes used in eBird can be downloaded from this page: https://support.ebird.org/en/support/solutions/articles/48000838205

COUNTRY CODE – Abbreviation for country name. Follows ISO 3166-2. Country names and codes used in eBird can be downloaded from this page: https://support.ebird.org/en/support/solutions/articles/48000838205

STATE – The state/province where the observation was made. Follows ISO 3166-2. State/province (or subnational1) names and codes used in eBird can be downloaded from this page: https://support.ebird.org/en/support/solutions/articles/48000838205.

STATE CODE – Abbreviation for state/province name. Follows ISO 3166-2. State/province (or subnational1) names and codes used in eBird can be downloaded from this page: https://support.ebird.org/en/support/solutions/articles/48000838205.

COUNTY – The county where the observation was made. No international county standard is available, so the best available county layer is used when available. County (or subnational2) names and codes used in eBird can be downloaded from this page: https://support.ebird.org/en/support/solutions/articles/48000838205.

COUNTY CODE – Alphanumeric code representing county name in this format COUNTRY-STATE-COUNTY (e.g., US-AK-016). County (or subnational2) names and codes

used in eBird can be downloaded from this page:
https://support.ebird.org/en/support/solutions/articles/48000838205.

IBA CODE – The alphanumeric code for an Important Bird Area and/or Key Biodiversity Area. If an observation falls within an IBA or KBA, it is given this code. Some locations may fall in more than one IBA or both a KBA and an IBA, and when that happens, these IBA codes are delimited in the field with a pipe ("|"). A list of codes and their corresponding site names is included in the metadata bundle (IBACodes.txt and BirdLifeKBACodes.txt). For more information on IBAs and KBAs see
https://www.birdlife.org/projects/ibas-mapping-most-important-places/.

BCR CODE – The alphanumeric code for a Bird Conservation Region. If an observation falls within a particular BCR, it is given this code. A list of codes and their corresponding site names is included in the metadata bundle (BCRCodes.txt). More on BCRs here:
http://www.nabci-us.org/bcrs.htm.

USFWS CODE – The alphanumeric code for a United States Fish and Wildlife Service land holding. If an observation falls within a particular USFWS polygon, it is given this code. A list of codes and their corresponding site names is included in the metadata bundle (USFWSCodes.txt).

ATLAS BLOCK – Sampling units called blocks have been established for specific atlas projects run within eBird. Blocks are established using a grid system (for example, in the United States these are based on 7.5-minute topographic quadrangle maps (quads) prepared by the U.S. Geological Survey). Each quad has a unique identifier. For atlasing purposes, each quad is divided into 6 blocks, each roughly 3 × 3 miles and encompassing about 23 sq km (9 sq mi). Each block has been coded with a 2-letter code: either northwest (NW), northeast (NE), center-west (CW), center-east (CE), southwest (SW), or southeast (SE). ATLAS BLOCK is only assigned to a record when an eBird Atlas PROJECT CODE (see below) is selected by the observer. See the list of eBird Atlases here:
https://ebird.org/about/portals/

LOCALITY – The reported location name for the observation. Observers can create and name their own locations, or choose from existing locations, including those curated by the eBird Hotspot team (these curated hotspots are seen here:
https://ebird.org/ebird/hotspots).

LOCALITY ID – Unique alphanumeric code for a location.

LOCALITY TYPE – In some cases location names can be confusing. This code is meant to help define the type of location used, as participants in eBird can plot specific locations on a map (personal locations, P), choose existing hotspot locations from a map (H), or choose to submit data for a town (T), postal code (PC), county (C), or state (S). Submissions from an entire county or state are not included in the EBD and not shown

publicly in eBird, since the location information is exceedingly imprecise. Abbreviations: State (S), County (C), Postal/Zip Code (PC), Town (T), Hotspot (H), Personal (P).

LATITUDE – Latitude of the observation in decimal degrees.

LONGITUDE – Longitude of the observation in decimal degrees.

OBSERVATION DATE – Date of the observation expressed as year-month-day (YYYY-MM-DD).

TIME OBSERVATIONS STARTED – The time observations were initiated based on the 24-hour clock (sometimes known as 'military time').

OBSERVER ID – Unique alphanumeric identifier associated with each Cornell Lab of Ornithology account. Every observation in eBird is associated with an account and unique Observer ID.

OBSERVER ORCID ID – Unique ORCID (Open Research and Contributor ID; https://orcid.org/) number (ORCID iD) associated with each eBird observer. This field will be null unless the observer explicitly provides and permits their ORCID iD to be associated with their eBird account. For more information on ORCID, see: https://info.orcid.org/what-is-orcid/.

SAMPLING EVENT IDENTIFIER – The unique alphanumeric identifier associated with the sampling event (eBird checklist). Each sampling event has a unique combination of location, date, observer, and start time. For a sampling event to exist in the EBD it must can contain observations of one or more taxa, all of which share this unique identifier. See section SAMPLING EVENT DATA near the top of this document; sampling events that record no taxa of birds will exist in that dataset and are important for establishing absences. The alphanumeric SAMPLING EVENT IDENTIFIER code begins with an "S" and is the final element of an eBird checklist link (e.g., https://ebird.org/checklist/S40515915).

OBSERVATION TYPE – The type of survey associated with this sampling event. These basic observation types define the core data entry fields for eBird and also the fundamental methodology being used to search for birds. The three primary observation types are:

- Traveling
- Stationary
- Incidental

Four secondary observation types are also provided to support less common scenarios:

- Historical
- Banding
- eBird Pelagic Protocol
- Nocturnal Flight Call Count

See also PROTOCOL NAME and PROTOCOL CODE below.

PROTOCOL NAME – A protocol is a specific survey with specific requirements to document effort, bird counting methodology, or other metadata related to the observation event (i.e., Submission) or bird counts (i.e., Observation). The seven Observation Types above are standalone Protocols in eBird when no additional metadata is collected for the Submission or Observation. However, in some cases a Project (see below) may require additional metadata fields from its volunteers. Additional Project-specific metadata is available to Project admins through their Project, not the EBD. The combination of an observation type with additional metadata fields may define a new Protocol, with a unique PROTOCOL NAME and PROTOCOL CODE. See the protocols.txt file provided with the download for the complete list of protocols, with their names, codes, and descriptions.

eBird previously supported more than 45 different protocols. The release of eBird Projects in March 2025 now allows researchers and conservationists to use the eBird platform to collect, download, and analyze specific data collection initiatives. This has allowed for eBird protocols to be simplified.

PROTOCOL CODE – This short alphanumeric code is used to identify the specific protocol. See the protocols.txt file provided with the download for the complete list of protocols, with their names, codes, and descriptions.

PROJECT NAMES – eBird Projects provide a framework to conduct targeted data collection efforts using the eBird platform, and access, download, and explore those Project-specific data. eBird Projects may support customized data entry for: 1) Project-specific locations, curated and maintained by that project; 2) additional metadata at the observation level (e.g., counts collected inside or outside of up to three specific distance bands); 3) additional metadata at the submission level to describe additional data relevant to that survey, location, or location during that survey (e.g., tide, habitat conditions etc.); 4) data aggregation within a system of grids, used to incentivize participation and promote more uniform spatial coverage (i.e., eBird Atlases, often Breeding Bird Atlases). When an existing OBSERVATION TYPE incorporates additional metadata at the level of the Observation of Submission, we treat that as a distinct PROTOCOL in eBird. Learn more about eBird Projects here: https://support.ebird.org/en/support/solutions/articles/48001268392. Since a given checklist may be part of more than one Project, these values are given as a list delimited by a pipe "|".

PROJECT IDENTIFIERS – Unique internal code (used in e.g., URLs) to define each eBird Project. Since a given checklist may be part of more than one Project, these values are given as a list delimited by a pipe "|".

DURATION MINUTES – The duration of the sampling event reported in minutes.

EFFORT DISTANCE KM – The distance traveled during the sampling event reported in kilometers.

EFFORT AREA HA – The area covered during the sampling event reported in hectares.

NUMBER OBSERVERS – The total number of observers participating the sampling event.

ALL SPECIES REPORTED – A critical field that separates eBird checklist data from most other observational datasets. Observers answer 'yes' to this question when they are reporting all species detected by sight and by ear to the best of their ability on a given checklist (sampling event). Observers answer 'no' to this question when they are only reporting a subset of species from an outing, usually the highlights or unusual birds. When observers report all species, it allows one also to infer which species were not detected. Given sufficiently large samples of records with ALL SPECIES REPORTED in a region, it is possible to estimate the probability that a nondetection represents the true absence of a species. (1 = yes; 0 = no).

GROUP IDENTIFIER – When multiple observers participate in the same sampling event, they can share checklists. If a checklist is shared between multiple observers (i.e., multiple copies are created from the original checklist, with one copy for each observer with whom the original checklist is shared), this group of duplicate checklists is given a GROUP ID number. These checklists can be edited by each observer so may or may not be exact copies in terms of the taxa involved, counts, comments, or even effort. Use GROUP ID to eliminate duplicate data when multiple observers are sharing data.

HAS MEDIA – Indicates whether a particular observation of a taxon is supported by rich media stored at the Cornell Lab of Ornithology's Macaulay Library (http://macaulaylibrary.org/). If an observation has media, the media are displayed in the eBird checklist view (e.g., https://ebird.org/checklist/S40515915, see "SAMPLING EVENT IDENTIFIER above) and on the Macaulay Library specimen page for the observation.

APPROVED – The status of the record within the eBird data quality process. If "Accepted", the record is deemed acceptable. If "Not Confirmed" the record has been deemed unacceptable by our review processes. Note that the eBird data quality process reviews the accuracy of both the checklist and the observation; when an entire checklist is marked "Not Public" in eBird (e.g., because the location is inaccurate or the checklist applies to multiple dates), all records are treated as "Not Confirmed" and are not

included in this dataset. "Not Confirmed" records that have not yet been reviewed (see REVIEWED below) are included in the unvetted dataset only (see UNREVIEWED RECORDS above). Unvetted data can be requested by checking the box "include unvetted data" on the download page with your data request. Unvetted data come in a separate file, and have not yet been reviewed by our regional editor network. It is not advisable to use unvetted data in any kind of analysis (1 = yes; 0 = no).

REVIEWED – "Not Reviewed" means that the record passed through our automated filters without problems, that the species, date, and count were within expected levels, and that the record has otherwise not been reviewed by a reviewer. "Reviewed" means that the record triggered a higher-level review process, either through an automated or manual process, and that it was vetted by one of our regional editors. (1 = yes; 0 = no).

REASON – The reason the record was "Not Confirmed". In this dataset, this field will probably be empty, since the EBD does not include "Not Confirmed" species that have been reviewed by reviewers. (Prior to the release of Exotic Species in eBird this field may have included "Species-Introduced/Exotic", but the revised Exotic Species practices (as of 2023) mean this is no longer used in this way.). See APPROVED above for more on eBird's review process.

CHECKLIST COMMENTS – General comments about the sampling event (checklist) provided by the observer as of free-form text. Because the text in this field is not under any constraints, it may cause problems with importing the data into a computer program, for example if the field separator used by the EBD is used by the observer as they enter their TRIP COMMENTS.

SPECIES COMMENTS – Comments about this particular species observation provided by the observer as free-form text. Because the text in this field is not under any constraints, it may cause problems with importing the data into a computer program, for example if the field separator used by the EBD is used by the observer as they enter their TRIP COMMENTS.

Appendix 1 – eBird Breeding and Behavior Code Definitions
The sequence of these breeding and behavior codes matches that shown during data entry on eBird and eBird atlases; the sequence of the codes represents the breeding code prioritization, with the better evidence for breeding listed higher in the list.

NY--Nest with Young – Nest with young seen or heard. Typically considered Confirmed.

NE--Nest with Eggs – Nest with eggs. Typically considered Confirmed.

FS--Carrying Fecal Sac – Adult carrying fecal sac. Typically considered Confirmed.

FY--Feeding Young – Adult feeding young that have left the nest, but are not yet flying and independent (should not be used with raptors, terns, and other species that may move many miles from the nest site). Typically considered Confirmed.

CF--Carrying Food – Adult carrying food for young (should not be used for corvids, raptors, terns, and certain other species that regularly carry food for courtship or other purposes). Typically considered Confirmed.

FL--Recently Fledged young – Recently fledged or downy young observed while still dependent upon adults. Typically considered Confirmed.

ON--Occupied Nest – Occupied nest presumed by parent entering and remaining, exchanging incubation duties, etc. Typically considered Confirmed.

UN--Used nest – Unoccupied nest, typically with young already fledged and no longer active, observed and conclusively identified as belonging to the entered species; note that this breeding code may accompany a count of "0" if no live birds were seen/heard on the checklist. Typically considered Confirmed.

DD--Distraction Display – Distraction display, including feigning injury. Typically considered Confirmed.

NB--Nest Building – Nest building at apparent nest site (should not be used for certain wrens, and other species that build dummy nests). Typically considered Confirmed, sometimes Probable.

CN--Carrying Nesting Material – Adult carrying nesting material; nest site not seen. Typically considered Confirmed, sometimes Probable.

PE--Brood Patch and Physiological Evidence – Physiological evidence of nesting, usually a brood patch. This will be used only very rarely. Typically considered Confirmed.

B--Woodpecker/Wren nest building – Nest building at apparent nest site observed in Woodpeckers (Family: Picidae) or Wrens (Family: Troglodytidae)—species in both families known to build dummy nests or roost cavities. Typically considered Probable.

A--Agitated behavior – Agitated behavior or anxiety calls from an adult (ex. "pishing" and strong tape responses). Typically considered Probable.

N--Visiting probable Nest site – Visiting repeatedly probable nest site (primarily hole nesters). Typically considered Probable.

C--Courtship, Display or Copulation – Courtship or copulation observed, including displays and courtship feeding. Typically considered Probable.

T--Territorial defense – Permanent territory presumed through defense of breeding territory by fighting or chasing individuals of same species. Typically considered Probable.

P--Pair in suitable habitat – Pair observed in suitable breeding habitat within breeding season. Typically considered Probable.

M--Multiple (7+) singing males. Count of seven or more signing males observed in a given area. Typically considered Probable.

S7--Singing male present 7+ days – Singing male, presumably the same individual, present in suitable nesting habitat during its breeding season and holding territory in the same area on visits at least 7 days apart. Typically considered Probable.

S--Singing male – Singing male present in suitable nesting habitat during its breeding season. Typically considered Possible.

H--In appropriate habitat – Adult in suitable nesting habitat during its breeding season. Typically considered Possible.

F--Flyover – Flying over only. This is not necessarily a breeding code, but can be a useful behavioral distinction, as it indicates that a bird was not interacting with the landscape underneath it.

Appendix 2. eBird Observation Types and Definitions
The below observation types are the seven current, active protocols available in eBird (http://ebird.org) for data entry. Additional protocols may be defined by a project, but all will use one of the below observation types. Additional information on specific protocols is available in the PROTOCOL NAME field and in the protocols.txt document provided with each EBD download.

This list is sequenced as in eBird, with more common protocols at the top of the list. The corresponding Protocol Code is listed parenthetically after the name of the protocol (thus, "Traveling" uses protocol code P22).

Traveling (P22) – This is used when you move more than 30 meters (100ft) during your checklist. You must be able to estimate the distance that you traveled during your outing, which can be walking, driving, or even by boat. If you aren't sure of the distance or area you covered, please enter your observation as an Incidental or Historical checklist. Examples of Traveling checklists are: walking a trail at a local park, driving an auto loop at a National Wildlife Refuge, participating on a pelagic (boat) trip, or even birding while jogging through your neighborhood. Required Date/Effort fields: Date,

Start Time, Duration, and Distance Covered.

Stationary (P21) – This is used when you do NOT move more than 30 meters (100ft) during your checklist. This does not mean you must stand completely still as you record the birds, but you should remain in an area approximately 30 meters in diameter while you are recording birds. If you move farther than that, you should use Traveling. Examples of Stationary checklists are: a hawk watch, lake watch, or sea watch, or even sitting in your backyard for a period of time identifying birds. Required Date/Effort fields: Date, Start Time, and Duration.

Incidental (P20) – This is used when birding is not your primary purpose: you notice some birds while doing something else and want to record them quickly. Checklists that do not include or distance/area components are classified as Incidental checklists. Examples of an Incidental checklist are: an oriole that flies by while you are checking your mail, a hawk flying past the window while you work, or a gull just off the side of the road while you travel to work. Note: This formerly was referred to as "Casual Observation." Required Date/Effort fields: Date.

eBird Pelagic Protocol (P60) – Only use this when you are birding from a boat more than 2 miles from land. If you're moving, keep the duration of your checklist to ≤60 minutes. If you're stationary (e.g., anchored), keep a single checklist and return to ≤60 minute checklists once you move again. Please visit this URL for the full protocol description: https://support.ebird.org/en/support/solutions/articles/48000950859-guide-to-ebird-protocols#anchorPelagic.

Nocturnal Flight Call Count (P54) – Use at night when stationary and detecting migrant birds by their vocalizations (by human ear or microphone). Separate checklists by date if they span midnight. More specifics can be found here: https://support.ebird.org/en/support/solutions/articles/48000950859-guide-to-ebird-protocols#anchorNFC.

Historical (P62) – Birding is your primary purpose, but you do not know start time, duration, and/or distance; use Traveling or Stationary if you know all of these values. This protocol allows start time, duration, and distance to be entered, but none is required. Required Date/Effort fields: Date

Banding (P33) – At least one bird on your checklist is captured during banding/ringing and would not have been detected otherwise. Answer "no" to "are you reporting all species" if you are ONLY reporting netted birds; answer "yes" if you include captured AND observed birds. Please read full protocol at this link: https://support.ebird.org/en/support/solutions/articles/48000950859-guide-to-ebird-protocols#anchorBanding.

One additional legacy protocol exists:

Area (P23) – Complete coverage of an area, used when exhaustively searching a known area for a specific purpose (Traveling counts that collect a GPS track using eBird Mobile generally preferable). Available only as a part of eBird Projects. Required Date/Effort fields: Date, Start Time, Duration, and Area Covered.