## DWC Sensitive Species Extension

It is proposed that these vocabularies will be applied in a one-to-many relationship with a DWC record or dataset similar to the measurement of fact DWC extension

NB a treatment should have a sensitivity reason associated with it. A sensitivity reason may result in one or more treatments being applied. A record / dataset may have more than one sensitivity reason and therefore treatment applied.

Firstly, the intention was to be able to address either individual records or datasets so the vocabulary can be used for either. However, the purpose of having sub-categories are:

1. so that data custodians can comply with basic FAIR principles as they apply to sensitive data. The general idea is that if a data record or a dataset is being withheld, then the reason should be as specific as possible in order to enable an interrogator to understand the reasons why and plan next steps in terms of either attempting to access the data, accepting the data at the scale provided, or giving up.
2. the standard needs to be able to handle the breadth of sensitive species data management actions across data custodians. For example in Australia, one state takes a simple approach of generalising all threatened species to a 10km radius because the state government perceives a generalised threat to all such species, whereas other states “shrink wrap” their sensitive species as much as possible, with the philosophy that releasing data does more to promote species conservation than withholding it and the risk of habitat being destroyed accidentally is far, far higher than the risk of poaching.

## Sensitivity Reason Vocabulary

| SensitivityReasonID | Sensitivity Category | Sensitivity Sub-category | Use Cases Refined | Invalid usage examples |
| --- | --- | --- | --- | --- |
| 1 | Personal Identifiable Information |  | The presence of personal data – names or addresses compromise the privacy of individuals | N/A |
| 1.1 |  | Names and addresses of individuals included in observation records | * A dataset has observation records made by a third party which have the name and address of the owner of a property where the observation was made. There is no record of whether the owners have given their permission for this information to be included. The name and address need to be withheld to protect the privacy of individuals. | A survey is undertaken on a property with the owner’s permission and permission to publish the data |
| 1.2 |  | Names of individuals in association with records enable personal details to be re-identified | * A dataset has the full names of collectors, observers or identifiers and there is no record of whether the individuals have given their permission for this to be included. Note, this is different to withholding collector name to hide a trip itinerary so that restricted access records cannot be located using unrestricted specimens. See 2.5/2.6 | A dataset has been gathered publicly. In signing up, participants have to agree to releasing their names. There is therefore no reason to withhold data. |
| 1.3 |  | Species information is sensitive because its release results in trespass or property damage | Species records where the owner or manager has requested the generalisation or withholding of records to avoid trespass or property damage. Examples include:   * Restrictions on hamster records because public records result in damage to crops because of members of the public trying to obtain photos or records of hamster. * A landholder is concerned that records of pest and weed species on their land will attract the criticism of neighbours | A feral species that is spreading rapidly through a country. Withholding or generalising records will actively impede national efforts to eradicate the pest. |
| 1.4 |  | Other personal information | * Other personal data reasons for withholding or generalising data that are not covered in SensitivityReasonIDs 1.1,1.2 or 1.3 | Only use this sub category when there are other more appropriate ones in this category |
| 2 | Species-Location |  | Information about the location or habitat of a species places that species at risk if data is not either withheld or generalised | N/A |
| 2.1 |  | Species under extreme risk of exploitation / harm | Species where knowledge of a locality or occurrence potentially poses a risk of disturbance to an entire species for assessed, documented reasons such as poaching, disease or habitat fragility. Records are withheld. Should be considered in association with SensitivityReasonID 4.1. Examples are:   * A newly described, highly endemic gecko species, restricted to a single cave system that belongs to a genus that is frequently targeted for poaching. Knowledge of latitude or longitude or the cave name makes the species easy to target and disturbance represents a threat to the entire species. * A fossil locality where specimens are only available from a single site and the site needs to be protected from black market fossil collectors or simply vandalism. | A small skink that is highly endemic, but all known locations were published over forty years ago and are freely available online. Although the species is at risk, there is no way to redact existing information and no reason to withholding records |
| 2.2 |  | Release of precise locations would result in high risk of exploitation / harm / disturbance | Species are not highly restricted but have life history attributes that make coarse generalisation of all records important to reduce risks such as habitat destruction, interruption of breeding cycles or poaching. Should be considered in association with SensitivityReasonID 4.1, particularly where 4.1 can be used as an alternative to generalising or withholding all records. Examples are:   * A species of orchid has a discontinuous distribution on scattered sphagnum bogs. The species is rare but not endemic, but its habitat is fragile. Generalisation of records reduces the risk of poaching and habitat disturbance * A species of crayfish is threatened by illegal fishing. The species was formerly widespread but remaining locations are difficult to access and find, so generalisation and withholding of coordinates will be effective. | A ground-nesting bird which is threatened but nomadic. Even with records, there is very little chance that records will reveal current populations so no point in generalising records |
| 2.3 |  | Release of precise locations would result in risk of exploitation / harm / disturbance | Species are broadly distributed across the landscape but there are specific life history traits that threaten individuals in such a way as to increase risk to the species, particularly feeding or breeding aggregations or permanent nest sites. Fine scale generalisation is sufficient to protect these traits. Should be considered in association with SensitivityReasonID 4.1, particularly where 4.1 can be used as an alternative to generalising or withholding all records. Examples are:   * Colonies of a widespread penguin species, where burrows are not obvious and locations increase the risk of trampling. * Records of a widespread dwarf cactus species that is sensitive to poaching but difficult to find, making targeting of records difficult | A threatened tree species with discontinuous populations over a broad area but which is at serious threat from woodland clearance. The minimal value of restricting data is outweighed by the significant value of publicising occurrences to encourage people not to clear populations. |
| 2.4 |  | Survey information can be used to infer location of at-risk species | Species that have been assessed as restricted access species, where their records were collected as part of a systematic survey or collecting trip expose their localities. Examples include:   * A regional flora survey has been undertaken in an arid area. Species records are associated with plot IDs. Restricted access species records have been generalised but they risk being exposed by the SiteID information. A decision needs to be made whether to withhold the restricted access species records or withhold the SiteID information for plots including restricted access species. * A plant collector undertook a field trip in 2015 collecting species belonging to a particular plant family. Some of the species are restricted access species. Because the specimens have a continuous collector’s number sequence, it is possible to identify the location of the restricted access species. Records of all specimens from a particular locality need to be withheld or generalised. * A museum is doing multi-year study reusing the same pit traps. The pit trap locations are therefore sensitive because they can be used by pet collectors to easily gather specimens in the latter years of the study |  |
| 2.5 |  | Species at risk of exposure from contextual information such as habitat being used to locate specimens | Species that have associations with other species that enable individuals of a species to be identified and place them at risk. Examples include:   * An endemic butterfly species has a commensal relationship with a food plant. The food plant is widespread, but the butterfly is highly restricted. Records of the non-sensitive food plant need to be withheld or obfuscated within the range of the butterfly. | Where the association is publicly understood, there is no value in applying this SensitivtyReasonID. For example, a wader whose migratory feeding areas are included in publications and websites. Withholding records to avoid disturbance is pointless because the data is freely available |
| 2.6 |  | Release of precise location information would result in potential harm to people | Release of records where the precise location represents a risk to observers. Examples include:   * There is a dataset covering records from a Defence Training Area including species that are popular for photographs. A decision to withhold records reduces the risk to the public from the risks of trespass including unexploded ordinance or live firing * There is a dataset of feral animals that is being used by hunters to illegal enter private property and shoot posing a threat to landholders. * Deathcap mushrooms (*Amantia spp.*) are poisonous to humans and are sometimes collected deliberately. Generalising records reduces the public risk of records of this type. |  |
| 2.7 |  | General data restriction for species | Instances where resources do not permit the species by species assessment of the risk posed by making species records publicly available. It is strongly recommended that this is not to be used as a justification for withholding data generally. Examples include:   * An agency has a large number of orchid species occurring within its lands. The species are known to be attractive to plant collectors and there are historical records of poaching. The agency takes the decision to generalise all orchid records. | Only use this sub category when there are other more appropriate ones in this category |
| 3 | Biosecurity species- |  |  | N/A |
| 3.1 |  | Restriction for biosecurity reasons | Specifically for high biosecurity risk species. Data contains declared biosecurity species which may be of concern to biosecurity or trade which have been seized and eradicated. Public display of records might be incorrectly interpreted and have trade implications. Examples include:   * A species of pinhole borer is detected at a port. The species is widespread overseas and many countries erect timber import restrictions based on the presence of the species. The borer infestation is eradicated and specimens are sent to the local museum, which serves specimen data publicly. The records should be withheld as they might be used to incorrectly suggest that the borer is present in this country. * A weed species is at an early stage of incursion. The incursion is featuring in media and has become politicised. In an attempt to encourage landholders to report the weed, confidentiality of records has been guaranteed until the weed is eradicated. | Don’t use this category for non-biosecurity species |
| 4 | Species - Attributes |  |  | N/A |
| 4.1 |  | Specific habitat information about species regarding their habitat or physical attributes can be withheld or generalised while all other records of a species are not sensitive | Instances where records of a species are not sensitive but a particular life history attribute or descriptor is. Can be used as an alternative to generalising or withholding all records. Examples include:   * There is a reference to breeding behaviour in the notes attribute. The notes field should be removed to protect breeding habitat. * Withholding stratigraphic unit or geological formation names to protect fossil sites * Withholding the name of a unique land feature such as a bog or a cave because it is the only place that a restricted access species occurs | The need to protect the species against the need to protect it’s important habitats needs to be weighed up here (like the tree example in 2.3) |
| 4.2 |  | Species regarded as sensitive because physical, habitat or management information but taxa not individually assessed | * Were specifics of the reason is unknown only that exposing additional information about the taxon may be detrimental to it or people. | Only use this sub category when there are other more appropriate ones in this category |
| 5 | Usage-restrictions |  |  | N/A |
| 5.1 |  | Legal contract restricting the release of raw data | This includes a negotiated legal contract or a standard form data licence agreement that in effect is a legal contract limiting the transfer of the data to a third-party. Examples include:   * Data gathered under a signed contract with a third-party contractor * Data from a research project provided under legal or licence agreement that relates to control of third-party data use. This limits who else the data can be shared with. It does not concern data where there are informal agreements in place. |  |
| 5.2 |  | Non--legal reasons restricting the release of raw data (embargo period) | This covers non-legal concerns that require a temporary embargo of records. Examples include:   * A company has access to data about endangered species in relation to a development they are involved with where third-party knowledge of that data (by competitors) might impact the economic viability of the project. This dataset is embargoed until a certain date after which time it can be publicly available. * Researchers involved in the development of new food crops need to withhold phenotype and monitoring data until after results are published to retain competitive commercial advantage * A researcher working on a PhD project has a significant number of observations about a new species. The researcher is happy to share the data but does not want the data to be made public until after the publication of their results. An alternative example might be a researcher’s data where the location can be revealed but other fields such as measurements must be withheld. This dataset might be embargoed until a certain date. * A project has a condition as part of participation in the project that all data gathered will be withheld for a specified time (eg two years). After that embargo period the data may be shared. | Species that need to be restricted because they have a commercial value should be dealt with under SensitivityReasonID 2 categories not here |
| 5.3 |  | Non-Legal restrictions on release of raw data (no time limit) | This principally applies to legacy datasets where it is no longer possible or too onerous to contact participants and there is therefore uncertain over sharing permissions. Examples include:   * A dataset includes a mixture of citizen and specialist provided records gathered between 1990 and 2000. There were no terms and conditions over the project at the time and too many people were involved and there are no contact details, so as a consequence there is uncertainty over the ability to share the data. |  |
| 6 | First Nations data |  |  | N/A |
| 6.1 |  | Cultural sensitivities relating to the species | Instances where a species has importance for First Nations communities that make sharing a record inappropriate. Examples include:   * A First Nations community regards a species as a totem animal. While the species is widespread, the community have agreed that they do not wish records to be made public has it conflicts with community beliefs * First Nationals rangers have gathered data on species in their reserve. Because the data belongs to community under CARE principles, the data can only be shared with community agreement |  |
| 6.2 |  | Cultural sensitivities relating to First Nations information about the species | Relates to First Nations traditional knowledge about a species that is inappropriate to share without permission. Examples include:   * An herbarium specimen has field notes that include information gathered by the collector about how a First Nations community uses the plant for medicinal purposes. The information belongs to the community and should not be shared without community agreement * An endangered tree record mentions that the tree has an important ceremonial carving on it. That information should not be shared without first nations community agreement |  |
| 6.3 |  | Cultural sensitivities relating to location | Relates to species records that reveal information about First Nation sites. Examples include   * Species records that include information about an important ceremonial location that should not be shared outside the community * Species records falling within the boundary of a Sacred Site that the First Nations community do not wish to share because (for example) it might lead to unwanted / inappropriate visitation |  |

## Treatment Vocabulary

The user notification level shows at what level the alert or notification to users that the dataset / record have been changed should appear.

| TreatmentID | Treatment Category | Treatment sub-category | User Notification Level | Refined Use Cases |
| --- | --- | --- | --- | --- |
| A | Withhold data |  |  |  |
| A.1 |  | Withhold entire dataset | Metadata / dataset level | Examples where an entire dataset might be withheld include:   * A dataset of species from a defence training area where people trespassing in order to see a rare species run the risk of stepping on unexploded ordinance * A dataset from First Nations rangers that requires community permission to use * A PhD dataset on a rare tree only known from one site that is threatened by a fungal disease spread on soil carried by shores or car tyres |
| A.2 |  | Withhold records | Dataset Level | Examples where records might be withheld include:   * Records of an endemic species that only occurs in a fragile habitat that is easily identifiable if coordinates are provided such as caves or bogs * Records of a severe international crop pest that were found at an airport and were eradicated, but public release of the records might confuse overseas observers into thinking the species is present in a country |
| A.3 |  | Withhold attributes | Record level and Dataset Level | * Some attributes of the record are withheld – these may be where the locality of the species is sensitive, other information is sensitive or allows the user to extrapolate the locality of sensitive species e.g. habitat information, collectors name, personal information, site descriptions, survey information, breeding / life stage information |
| A.4 |  | Withhold entire dataset for Embargo period | Dataset Level | * A researcher may want to withhold a dataset prior to publication of the results of their research. |
| A.5 |  | Withhold specific records for Embargo period | Record level and Dataset Level | * Where data contributors impose an embargo period on selected records under agreement. |
| A.6 |  | Withhold specific attributes for Embargo period | Record level and Dataset Level | * Where a time period is sensitive where it allows the tracking of movement of taxa in near-real time e.g. pest species or other sensitive species. |
| B | Obfuscate Coordinates |  |  |  |
| B.1 |  | Location coordinates generalised by reducing number of decimal places | Record level and Dataset Level | Coordinates have the decimal places reduced – commonly this may include either zero, 1 or 2 decimal places. Treatment descriptions for the record should state the degree to which the coordinates have been truncated. Examples include:   * Highly sensitive species such as a recently rediscovered species which has no specific habitat requirements and is nomadic so coarse generalisation is sufficient to protect the species * Species for which the release of precise locations would subject it to a risk of exploitation / disturbance |
| B.2 |  | Location coordinates replaced by a standardised grid rather than points | Record level and Dataset Level | Treatment descriptions for the record should contain information about the standardised grid including its dimensions and origin.   * Record contains information regarding the management of species that the land manager considers mildly sensitive e.g. pest control on private property |
| B.3 |  | Location coordinates are randomised within a defined polygon of specified proportions | Record level and Dataset Level | Records latitude and longitude are randomised within a defined polygon. The defined polygon and its proportions should be included in the treatment descriptions for the records Case study examples similar to the rest of the category. |
| B.4 |  | Location coordinates replaced by centroid of a geographic region | Record level and Dataset Level | This can include the centroid of a defined geographic region e.g. county, jurisdiction, bioregion, map sheet etc. Treatment descriptions for the record should include which geographic region, type and definition of that region which was used to define the centroid.  Examples include:   * Records that a landholder considers sensitive from a privacy perspective |
| C | Generalise attributes |  |  |  |
| C.1 |  | Replace attribute with a generalised value | Record level and Dataset Level | * generalise scientific name to higher taxonomic level, nearest named place for locality, region name for locality etc where these attributes could expose information about a sensitive species |
| D | Obfuscate Time / Date |  |  |  |
| D.1 |  | Date generalised to Month | Record level and Dataset Level | * To avoid inferences being made about the locations of restricted plant species by using collections of common plants made around the same time by the same collector |
| D.2 |  | Date generalised to Year | Record level and Dataset Level | * To avoid inferences being made around breeding / nesting times in particular areas for some taxa |
| D.3 |  | Date generalised to Decade | Record level and Dataset Level | * for species like seals which do not have an annual breeding cycle where inferences can be made from the locality (eg. Species that come to shore to breed) and the period between breeding events. |
| E | First Nations |  |  |  |
| E.1 |  | Treatments specified by First Nations communities | Record level and Dataset Level | * where the species, it’s locality or other information within the record has significance or sensitivities according to the traditional owners. |
| F | Other Treatments |  |  |  |
| F.1 |  | Obfuscate or withhold records of non-sensitive taxa due to associations with sensitive taxa / site / or collector etc | Record level and Dataset Level | * Where a non-sensitive taxon is associated with a sensitive taxon so that its presence in an area can be extrapolated to divulge the presence of a sensitive species. |