****

**INTERNAL MEETING MINUTES**

|  |  |
| --- | --- |
| Meeting Purpose | TDWG Sensitive Species Extension Task Group Meeting |
| Time and Date | 19th December 2024 from 1700 AEDT |
| Attendees | Tania Laity  Cam Slatyer  Piers Higgs  Arthur Chapman  Helen Owens  William K Morris  Julia Percy Bower  James Wilson  Peggy Newman |

**Minutes**

Previous actions

* Cam and Tania to look at restructuring this into a logical table that separates these two concepts out for the group to review in the future
* Cam to find out if Nagoya is being reviewed elsewhere

Agenda

Welcome from Cam, and acknowledgement of country.

Cam recapping the last meeting; an email was sent out with proposed Sensitivity Treatments or Reasons for discussion. (due to CSIRO having issues with email this was not circulated to external attendees prior to the meeting) ***– See Attachment A***

Tania ran through the document and noted that you could potentially have multiple treatments to a dataset or record or multiple sensitivity reasons for a particular single record or across a full dataset. Instead of having every possible combination of treatment and / or reason we would like to propose that these could be provided in a 1-to-many approach either at record or dataset level – similar to how MeasurementofFact is structured and applied to DWC datasets and records.

We have included placeholders for indigenous data sensitivity reasons and treatments so that these can be used straight away regardless of not being defined more specifically at this point.

Arthur questioned the need to have another category of treatment - “withhold species”? This could possibly be dealt with under sensitivity reasons and then treatment of withhold records e.g. treatment 1.2 with a sensitivity reason that is “species of significance”.

**Sensitivity Reasons**

The Sensitivity reasons were predominantly drawn from the RASD framework, nothing else added apart from indigenous reason categories. We’ve also tried to be flexible so that if there is things that are missed we’ve included “other” categories to catch these.

Each category was described:

* Personal Identifiable Information - can use the top level (1) if they want, or can say specifically what they’re obscuring otherwise
  + Names and addresses (pretty simple)
  + CollectorIDS - can be used across all the collections, so that can be found from the other collections that use that ID (e.g. some places use the same ID across multiple collections)
  + Other personal information (catch anything else)
* Species-Location
  + This is largely broken down to have regard of the relative degrees of harm e.g. extreme to moderate risk of harm
  + Other categories include where landholders
* Species - identification
  + Only one - biosecurity species as an example
* Species - attributes
* Usage-restrictions
* Indigenous Data

James suggested rewording “not permitting release” in the usage restrictions sub categories to “restricting release” because you don’t want to conflate it with the treatment (i.e. for sensitivity reasons 5.1, 5.2, 5.3).

**Action:** Reword “not permitting release” to “restricting release” in sensitivity reasons 5.1, 5.2, 5.3

Arthur also suggested an edit to the usage restrictions categories - some of these might be temporal ie will have a time stamp on it or may be “subject to publication”. Category 5.3 should be split to include categories that include temporal periods. James agreed to split categories as it makes it easier when managing temporal restrictions to data.

**Action:** Split sensitivity reason category 5.3 into temporal (e.g. embargo) or non-temporal (informal)

Cam mentioned that 2.5 is included to pick up survey data where you can pick up the location of sensitive species through association with the survey points themselves. For example, you can flag all species within the same survey point and obfuscate the records or exclude sensitive species record entirely. The NSW DCCEEW (Aust) are handling this quite nicely and may be worth further discussion or a presentation to describe their process.

**Action:** Invite Ron Avery (NSW DCCEEW) to a future meeting of the WG to describe how they are handling survey data with sensitive species.

Arthur mentioned that another consideration in 2.5 is that if collector’s names where they are in a sequence and a sensitive species has been recorded / collected – it’s a process of elimination to work out where the missing point would be. This is difficult to handle and unsure how to do this without leaving the whole collector’s name out across the whole dataset.

**Action:** Ensure 2.5 also includes the reason of sequences of collector’s names / ids to enable identification of location of sensitive species.

**Treatments**

James questioned if there should always be a reason and where there is a treatment and vice versa. He also questioned if there could be multiple reasons and one treatment or multiple treatments and one reason. This is both at a dataset and / or record level.

Tania responded to say yes – if you have a treatment there should be a reason associated with it and vice versa. The whole point of standardising this is to make it clearer to users as to what’s been done to a record / dataset and why - so would need both. And you could have a reason and multiple treatments – for example, you could apply multiple treatments for one reason. Arthur agreed and mentioned the use case where you would obfuscate coordinates and also do something with locality field such as withhold that attribute as well. He also stressed that it’s important to document what has been obfuscated - telling people that there is more information but you can’t have it and then you can request access to it.

Will flagged that one treatment is missing from the list. In Finland they obfuscate (not withhold) time as well as location in some instances e.g they provide a year or decade but not a precise time. James asked if there was a set of rules around when a year, when a decade is chosen. Will mentioned that they have an independent body/panel of experts from various stakeholder groups (mainly government) that go through this and recommend restrictions / treatments to taxa. Arthur mentioned the use case of obscuring month for particular species due to nesting times etc.

**Action:** Add a time obfuscation category to Treatment Vocabulary

James asked for clarification that if there were treatments and reasons attached to a record / dataset that there would be a reference list of taxa and recommended treatments for uses to refer to. Cam clarified the need to link to a live sensitive species list / lists. Data custodians / aggregators and users want to know what was being obfuscated otherwise it’s impossible to know if their records were being treated properly. Arthur suggested that one way of making it more obvious to users where data has been withheld is to replace values in attributes fields e.g. if location is withheld – include text saying location information is restricted due to sensitivity reasons or include a free text field where this is not appropriate. Helen mentioned that this needs to be articulated in the metadata level where treatments have been applied to records within a dataset or a whole dataset.

Helen suggested that the treatment table should include where notifications to users regarding treatments applied e.g. either at dataset / record / attribute level (e.g. which ones need a dataset alert and which ones need a record alert - use asterisks or hashes to put that into the table)

**Action:** Modify treatment table to indicate what level user alerts should appear.

**Wrap Up and Next Steps**

We will work on editing this table and provide a before and after version to the working group for further comment prior to the next meeting. We will request use cases to illustrate where and how reasons and treatments could be applied and why. This will help us test the reasons and treatments. James is interested in testing the use cases against Arthur’s document and RASD. Aurthur suggested that we look at a survey he did with GBIF years ago that includes a number of use cases.

**Action:** Tania to make Arthur’s earlier documents available and compile some initial use cases from these.

**Action:** Task Group Members to send through use cases to Piers to collate by end of January 2025

**Action:** James to test the use cases against the risk assessment in Arthur’s document and the treatment process in RASD

**Next Meeting: February the 6th**

## Attachment A: Proposed Sensitivity Treatments and Reasons

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

## Sensitivity Reason Vocabulary

| SensitivityReasonID | Sensitivity Category | Sensitivity Sub-category |
| --- | --- | --- |
| 1 | Personal Identifiable Information |  |
| 1.1 |  | Names and addresses of individuals |
| 1.2 |  | Collector IDs of individuals (where this can be reconciled with collector) |
| 1.3 |  | Other personal information |
| 2 | Species-Location |  |
| 2.1 |  | Species under extreme risk of exploitation / harm |
| 2.2 |  | Release of precise locations would result in high risk of exploitation / harm / disturbance |
| 2.3 |  | Release of precise locations would result in moderate risk of exploitation / harm / disturbance |
| 2.4 |  | species information that the landholder or land manager considers sensitive from a privacy perspective |
| 2.5 |  | Species at a high to moderate risk of exploitation that have been recorded as part of a survey where a survey identifier can be resolved to a precise locality |
| 3 | Species - Identification |  |
| 3.1 |  | Data contains declared biosecurity species which may be of concern to biosecurity or trade e.g. a species of concern recorded as a quarantine intercept |
| 4 | Species - Attributes |  |
| 4.1 |  | information about species regarding their habitat or physical attributes which could result in exploitation or increased risk to populations if released e.g. nesting sites, reproductive status etc. |
| 4.2 |  | information regarding the management of species that the land manager considers sensitive e.g. pest control activities on private property |
| 5 | Usage-restrictions |  |
| 5.1 |  | Legal contract not permitting release of raw data |
| 5.2 |  | Commercial in confidence (legal-financial) reasons not permitting release of raw data |
| 5.3 |  | Non-Legal – other restrictions not permitting the release of raw data e.g. embargoed data, informal agreement between parties |
| 6 | Indigenous data |  |
| 6.1 |  | Cultural sensitivities relating to entire dataset |
| 6.2 |  | Cultural sensitivities relating to species |
| 6.3 |  | Cultural sensitivities relating to location |
| 6.4 |  | Cultural sensitivities relating to other attributes |

## Treatment Vocabulary

|  |  |  |
| --- | --- | --- |
| TreatmentID | Treatment Category | Treatment sub-category |
| 1 | Withhold data |  |
| 1.1 |  | Withhold entire dataset |
| 1.2 |  | Withhold records |
| 1.3 |  | Withhold attributes |
| 1.4 |  | Withhold entire dataset for Embargo period |
| 1.5 |  | Withhold specific records for Embargo period |
| 1.6 |  | Withhold specific attributes for Embargo period |
| 2 | Obfuscate Coordinates |  |
| 2.1 |  | Location coordinates generalised to 1 decimal place |
| 2.2 |  | Location coordinates generalised to 2 decimal places |
| 2.3 |  | Other method of generalisation / obfuscation of coordinates |
| 3 | Other treatments |  |
| 3.1 |  | Treatments specified by indigenous owners(?) custodians(?) |