****

**INTERNAL MEETING MINUTES**

| Meeting Purpose | TDWG Sensitive Species Extension Task Group Inaugural Meeting |
| --- | --- |
| Time and Date | 27th September, 2024 from 16:30 AEST (UTC+10) |
| Attendees | Tanya Laity Cam Slatyer  Andrew Rodrigues Piers Higgs Mieke Strong  William Morris Kiah Grogan  Kate Ingenloff  Julia Percy-Bower  Alice C Hughes |

**Minutes**

Introductions

Cam welcoming the group as the Chair, and will be also recruiting some others around the globe.

Today’s purpose is a scene setting exercise, to do some introductions.

Cam recapped the TDWG Hobart 2023 workshop and the work that Gaia Resources and GBIF

Introductions around the room.

* Cam Slatyer - Managing the taxonomic backbone of the Atlas of Living Australia, and was the PM of the Restricted Access Species Data Framework
* Tania Laity - work with Cam, worked on RASD, and working on some of the other projects at the ALA
* Piers Higgs - CEO of Gaia Resources
* Andrew Rodrigues - Data Partnerships officer at GBIF - looking at the use of data for decision making, handling a lot of sensitive data
* Julia Percy-Bower - Database manager at WA Herbarium
* Kiah Grogan - Working at DBCA - working in the Biodiversity Information office on Dandjoo
* William (Will) Morris - Finnish museum of national history as IT specialist in bioinformatics unit, and node manager of GBIF Finland; run a service for authorities to access sensitive data separately to the public domain data (and a sensitive data sceptic)
* Kate Ingenloff - from Copenhagen, science team joining for curiosity and working on the side of mobilising survey monitoring data to GBIF, and sensitive data and ILK is going through that onboarding approach
* Alice Hughes - in an airport in China - working from primary species distribution data to describing species, some of which are sensitive (e.g. range restricted and vulnerable to the pet trade) so at the pointy end of data collection, from the University of Hong Kong
* Mieke Strong - doing a lot of work on data standards - also being involved in the Biodiversity Data Repository project, so really important to know what is coming from this space

Cam talking about conflation of sensitivity and restricted data - not necessarily withholding data, but restricting the flow of data, create a standard that enables data custodians to take a sensible approach to their data, to try to get as much data into the public estate as possible, not an excuse for withholding data.

TDWG Processes for Standards

TL: Slide deck being shown, the main tasks:

* Set up a task group - done in TDWG Hobart 2023
* Develop a charter for the group - that goes to the Executive for approval (done)
* Goals and output and outcomes to be done:
  + How the sensitive species extension relates to other TDWG stuff
  + How this can be included
  + Developing a specification
* When we have consensus on that extension, deliver to the TDWG Executive
* Follow the documentation standard to get the data for approval of all the different documentation provided
* Also a vocabulary maintenance spec that is also included including a public feedback period

KI: One of the things in the Humboldt group was to develop use cases that explained the usability, and they could get feedback from people along the way as well, and also submitted an implementation report that summarised the feedback from those use cases - showing that it has been tried out on various types of data - so potentially consider each type of sensitive data and write a use case for each.

Background to the Task Group

CS: Talking to the background slide

* Arthur Chapman’s 2020 “Current best practices for generalising species occurrence data” and RASD Framework were two of the latest things.
* Arthur proposed an extension to Darwin Core
* GBIF recently looked at how the current sensitive fields are being used and found gaps
* With the expansion of biodiversity data there is a lot of need to look at restricted access species data, and there are various other initiatives under way as well

ACH (via chat): This missed a key point. Newly described small range species in genera very vulnerable to pet trade are not captured in CITES and are amoung the most sensitive

for example cave geckos, a number of which are thought to have been collected to extinction in the wild due to heavy collection for pet trade, sometimes using species descriptions to locate sites

AR: The report did not find that global trigger lists based on IUCN Global Red Lists and CITES Appendices were necessarily the best approach for identifying sensitive species

CS: Australian jurisdictions all have vastly different ways of dealing with sensitive data, but moving from broad brush to specific targeted species, the NSW process where it’s an add on to the threatened species process but discrete have been very good at picking up species as they have been described and then contacting aggregators to remove data before it’s listed (but there is only a list of 200 - but they are genuinely at high risk)

ACH: Cross-cutting lessons - where data is not more readily available - e.g. will not give a type locality when describing species - and bringing information such as Andrew discussed and then the NSW stuff we should be able to help with that.

AR: The majority of the world do not have the lists that we would need - so defining some traits might be useful across regional areas.

Recommendations from TDWG 2023 Workshop

Three recommendations from the workshop:

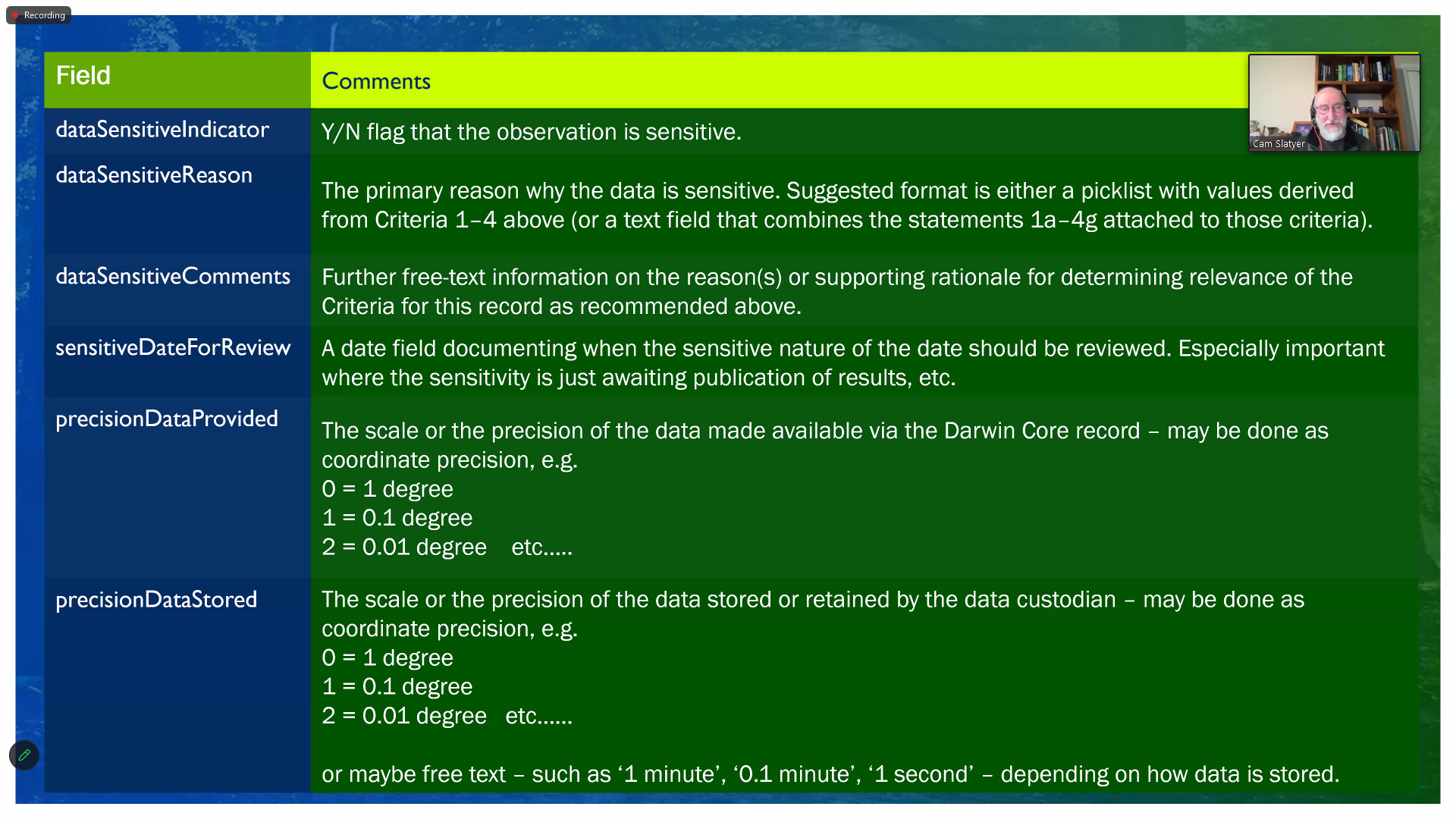
* Amend Arthur’s proposed extension
* Define a standard set of extension fields / vocabularies / ontologies
* Consider the creation and adoption of national checklists

(Alice has left for a flight)

Recording has started in Zoom

(quick recap by Cam for the benefit of those watching it later)

Popped up the slide number 10 for visuals



Additional amendments in the workshop that were proposed also discussed in this section. *Minutes from here no longer are in temporal order, but were captured and documented according to the fields they relate to.*

* *While the inclusion of Indigenous and cultural sensitivity might be covered under* ***dataSensitiveReason****, fields other that coordinates may be affected by sensitivity and this would need to be picked up*

CS: Cultural sensitivity might be out of the remit of this group but is included for context

* *Where data is embargoed for a specified period, an* ***embargoDate*** *field is also required. This was considered a good idea due primarily to withhold records prior to publication – delisting might be another reason. Addition consideration under this suggestion needs to be given to consideration of obfuscated pre-embargo release product prior to embargo end date instead of embargoing entire dataset / and all other records in the dataset that isn’t sensitive; and clarity around the publication which the data is going to be used in so that others can make a call if the embargo period has ended once the paper is published*

CS: Example of the data from a PhD for example being able to provide data to an aggregator but not have it published until after the PhD

KG: i certainly support an embargo field - it's something that has come up a lot recently

MS: Do we have guidelines for each ‘type’ of embargo - so recommendations on different categories?

CS: Agree we should have the ability to do that. Shouldn’t have embargo by default!

KG: I also back Mieke's comment on dataSensitiveReason with the use case of WA having threat codes from the BC and EPBC Act and wanting to capture that

PH: Should we have the embargoDate as embargeoEndDate or embargoLiftDate?

KI (via chat): Yes to embargoStartDate and embargoEndDate

**DatasensitiveReason**

MS: Can you have multiple values in there that might interact with other values?

CS: Probably needs some further discussion - experience in Australia is you can have more than one reason, and they may not necessarily apply to the same field, so may need to revisit that.

KI (via chat): Both precisionDataProvided and precisionDataStored should be divided into 2 terms each to follow other DwC terms ---> precisionDataProvidedValue and precisionDataProvidedUnit, and precisionDataStoredValue and precisionDataStoredUnit

KI (via chat): dataSensitiveReason should allow multiple entries with | between each entry

**Language and openness/restriction**

WM: I like them but they are also premised on the fact that ther is a benefit to the reason to restrict data. Do we have a dataNOTsensitivereason. Supports the sensitivedatefor review approach

CS: Would be good that the flavour needs to be there as a last resort for important issues, not a first resort for absolutely everything

TL: Would like some guidance that this is open by default and stays there - Arthur goes into that at the start, has the decision tree - but then very separate from the data field recommendations

MS: A good reason for making things sensitive is necessary - and an opportunity is there to call out the justification and making it public and contact the data provider to find things out - find the balance between having some restriction, but make as much data available as possible - does it really need to be restricted?

**Generalising fields - time**

WM: Not only generalise coordinates in their data but also generalist time - precision of time might need to be generalised as well

TL: Started a discussion around that previously

CS: Related context - hacking other data to find open data of the same time stamp to find out where those restricted data were actually located

KI (via chat): Agreed to date generalisation terms

**Other fields being generalised**

MS: There are other fields that also might be restricted - people’s names or localities as well.

TL: There are precedence here ie.g. One to many relationship that can be set up in Darwin core, like the measurementofFact attributes in DwC, like that - a record can have one or many generalisations on it. Might be able to be set up with various generalisations in teh same way

*(back to temporally ordered minutes from here)*

Priorities

CS Proposing the prioritisation here.

For example, we need to workshop the proposed extension fields - some general agreement - rewording or alternate wording to make these exceptions not the rule, and the subtleties we started to discuss, merits more discussions

Break the three recommendations up into three work packages, as a series of sequential units? Landing the extension unit first is fundamental to the rest of these, proposing the packages done in this order:

1. Discuss and agree Sensitive species extension fields and definitions
2. Define a standard set of extension vocabularies / ontologies
3. Consider and propose principles for the creation and adoption of national checklists

CS: Any ideas?

MS: Work package 2 is going to be touched on a lot in work package one…

CS: Agreed - maybe we populate the work packages as we go in the Github repository

KI: Also think about the fields and definitions - populate examples as well as we go, so we can capture the examples, and things like that as we go.

CS: So we should extend the wording of the work package to also include “and with examples” at the end

KI: Yep

KI (via chat): Perhaps a WP 4 for use case development and implementation testing of the proposed extension

TL: Can be done as we develop each, but

Agreement:

* Alter the wording of the first one to include examples (see above)
* Work through it in a sequential but overlapping way
* Add in a fourth work package that is about use case development and implementation testing

Timeframes

CS: Would be really keen to try to knock off WP1 over the next 12 months. Is that realistic?

KI: Yes, with WP1 and 2 it’s going to be ongoing until we get to the open public process before ratification - so will have an ongoing timeline that this will continue right through to this

CS: Propose to use the 1st October 2025 as the working timeline and need to agree a reasonable frequency of meetings. Anyone with an objection of monthly meeting, on a different evening than a Friday night… does this time work?

General agreement this is a decent time to work with and experiment next month but for a one-hour meeting time slot.

CS: Can I propose Thursdays at this time?

General agreement, working on a fortnight that TL is available on those weeks.

AR: The COP is coming so just watch those dates in October

CS: Will aim for early in the month for November and early December and will workshop dates offline via email.

CS: Will circulate an agenda for the next meeting, pick up discussion from the previous one, and move the discussion on via email. Will send out the notes and look at a revised extension from the notes today.

CS Thanks people for coming and closes the meeting - the emphasis on this being a last resort not a first resort is a key outcome from today

Meeting Closed: 15:56 AEST (UTC+10)