

**Agenda/Meeting notes (November working session):**

Participating:

- Rod Page
- Raïssa Meyer
- Camila Plata
- David Fichtmueller
- Ian Engelbrecht
- Jutta Buschbom
- Soulaine Theocharides
- Ben Norton
- Kit Lewers
- Sarah Vincent
- Rukaya Johaadien
- Carlos Martínez
- Holly Little
- James Macklin
- JJ Dearborn
- Jonathan Rees
- David Shorthouse
- Wouter Addink
- Brenda Daly
- John Wieczorek

Comments and notes taken during the meeting are in red.

NOTE: please make comments as text directly in the document rather than in marginal comments.

As with previous meetings, Steve will record the meeting for note taking and later viewing (this time publicly since it's a working session).

- I. Set time and date for next meeting. Suggest 2023-01-09 13:00 UTC. **No objection**
- II. For new attendees: about the TAG and review of recent activities (Steve). Slides posted in the TAG GitHub repository at <https://github.com/tdwg/tag/blob/master/meetings/2022-11-07-working-session-slides.pdf>
- III. Follow-up on items from the previous meeting
  - A. 3D Task Group will probably move ahead regardless of resolution of the dc:type issue. To be discussed at the Audubon Core Maintenance Group working session Thursday, 10 November 17:00 UTC. Close <https://github.com/tdwg/tag/issues/27> **No objection.**

- B. Propose recommendation to Executive to abandon effort to make GGBN a TDWG standard until a champion emerges. Close <https://github.com/tdwg/tag/issues/17>  
Ocean Teacher is having a course for data publishing using GGBN for several projects in Colombia in November, Camila can try to bring the topic with the members of GGBN one last time. Raissa also mentioned that GBWG was meeting this week and she would bring it up with them one more time. But she agreed that for now we could close this issue. No other objections.
- C. Brief report back on Memorandum of Understanding between TDWG and the Genomic Standards Consortium (GSC) [https://github.com/tdwg/gbwg/blob/main/dwc-mixs/MoU/MemorandumOfUnderstanding\\_TDWG-GSC.md](https://github.com/tdwg/gbwg/blob/main/dwc-mixs/MoU/MemorandumOfUnderstanding_TDWG-GSC.md) (output of the [Sustainable DwC-MIxS Interoperability TG](#)) Raissa reported that we now have an officially signed agreement. She noted that this is an important step for TDWG with respect to exploring how to work with other standards bodies. Steve noted that this agreement has provided us with a model for the technical details of how such mappings can be done.

#### IV. Policy on boolean values.

- A. Draft controlled vocabulary: <https://github.com/tdwg/tag/blob/master/build/boolean/index.md>
- B. Ancillary document Steve will work in creating a sort of "recipes" document showing how to serialize the values in the different systems that Ben outlined in his detailed background document.
- C. Does this need to go through a formal standards process? Or do we just recommend it in a manner similar to the tdwgutility: namespace terms? See <https://github.com/tdwg/tag/issues/13> Options: go through official standards process with the TAG as the maintaining group or simply maintain it as another utility vocabulary that is not an official TDWG vocabulary, but that is recommended for use as a best practice. Ben noted that this documentation can provide a defense for choices about how to handle booleans, and not require the implementing group to have to document and explain. One can simply refer to these documents. David noted that in the future, if people don't follow these recommendations, they could be asked why and if they have a good reason, that's OK. But otherwise, they should follow the recommendations. There was a consensus that it isn't necessary to advance this through the standards process.
- D. How do people find this? Where do we put TAG-developed best practices? Consensus: on the website, precise place not yet decided. On a subpage of the TAG, linked to from several other reasonable places as well, e.g. the process page: <https://www.tdwg.org/about/process/> . Other suggestions were: in a new subcategory on the standards page: <https://www.tdwg.org/standards/> DS in chat: TDWG website pulled from GitHub. Can make pull requests here <https://github.com/tdwg/website> & Stan Blum will evaluate. Jutta suggested that there should be a "Recommendations" section on the website – the website is the first place she goes for guidance. Steve noted that if it's on the website, there needs to be some high-level navigation mechanism for people to find the recommendations. If they are buried somewhere deep in the website (e.g. under the TAG page) nobody will find them.

- E. The process established here will set the stage for the effort to address arrays (item VI) and is directly connected to the Latimer Core review.

For reference here is how CSVW (the corresponding W3C recommendation) handles the question of choice of representation of booleans  
<https://www.w3.org/TR/tabular-data-primer/#boolean-format>

V. GUID issues

A. Open issues related to this:

1. <https://github.com/tdwg/tag/issues/14> (Convene an identifiers task group)
2. <https://github.com/tdwg/tag/issues/2> (Deprecate LSID Applicability Statement)
3. <https://github.com/tdwg/tag/issues/9> (Revise GUID Applicability Statement)

B. Options:

1. Convene a task group to update the GUID Applicability Statement. Needs a champion to convene the group.
2. Create documents (not to be included in the standard) that explain the history, current practice within TDWG, and non-binding recommendations. Include links to these documents on the standards landing page. Work to be done by TAG members as part of normal operations.
3. Do nothing.

C. Discussion and decision. Record decision while closing ancient issues. Open a single new issue if necessary, clearly defining the task.

Comment (Ian Engelbrecht): I support Option 1, and would like to add Roger Hyam and Wouter Addink as additional names to the list of potential participants included in minutes from an earlier meeting (Wouter is agreeable to this). My sense is that there's a real need for TDWG to provide a guideline on identifiers for the community, but that there's also still quite serious work to be done to resolve what exactly we want from identifiers (pointers to database records versus disambiguating entities like taxon names). I'd like to propose another option, to set a meeting date for early 2023 for that proposed list of participants to discuss forming a task group, identifying a (respectable) convenor, setting a plan of attack, etc. I'm happy to organize that meeting and hopefully get the ball rolling. (He is happy to play the administrative role of organizing that meeting, but does not plan to be the convenor of the task group.)

Comment from Wouter during meeting: He noted that are several policies existing in Europe for PIDs (permanent identifiers) and he thinks it would be a good idea for TDWG to endorse a PID policy. Example from chat: EOSC PID policy: <https://doi.org/10.2777/926037>

Comment (Lyubomir Penev): There is a lot of work going on in the [BiCICKL project](#) on recommendations for use of identifiers in various aspects. One of the first outcomes is a large paper with recommendations of use of identifiers for article- and subarticle-level elements in biodiversity publishing (expected to be published in RIO Journal in the next 10 days). Another important outcome is the forthcoming "Best practice manual for interoperability of infrastructures" developed under the lead of Wouter Addink.

Comment (Rod Page): There are millions of LSIDs published in articles and on web sites. Failure to be able to resolve these looks bad. Some progress made at TDWG 2022  
<https://lsid.io>

Comment made after meeting (Tim R): I support Option 1 and also note that there is significant use of LSID (one of the most used formats in terms of objects identified) as Rod says, so should be addressed as part of the discussion.

## VI. Recommendations for expressing complex values

- A. In several task groups there has been discussion about what to recommend in cases where properties may require multiple values, or in cases where values of several properties need to be linked in order for them to be meaningful. Rather than having each group adopt its own ad hoc solution, it was suggested that the TAG might coordinate TDWG-wide recommendations.
- B. There is a review of current precedents within TDWG at [https://github.com/tdwg/tag/blob/master/complex\\_values/existing\\_precedents.md](https://github.com/tdwg/tag/blob/master/complex_values/existing_precedents.md)
- C. There is a list of use cases to be considered at [https://github.com/tdwg/tag/blob/master/complex\\_values/use\\_cases.md](https://github.com/tdwg/tag/blob/master/complex_values/use_cases.md)
- D. Is this something that the TAG wishes to take up? If so, timely action would be necessary as several of the groups are in the last stages of developing their standards. **Comments by Ben:** It's easy to handle nesting of multiple complex values with a structured serialization like JSON (see his example in the use cases document), but how do you handle them in CSV files? How do we handle these across the board? Latimer Core has several of these and they've come up with several types of arrays. He thinks we need to have some kind of white paper to address each type of use case.
- E.

From James in the chat: Have we looked beyond us to other communities for a best practice? It strikes me that this is a global issue and someone must have addressed this...

Steve mentioned the IIIF standard: <https://iiif.io/api/cookbook/> although there is a single serialization: JSON-LD, whereas TDWG supports multiple serializations. Steve suggests that Rob Sanderson (of the TAG) may have thoughts on this.

Steve suggested that putting bits of JSON snippets in spreadsheet cells seems like a simple solution, but he doesn't have enough experience to know if it's a good idea. It's how dynamicProperties works in Darwin Core.

John Wiczorek said that JSON was recommended for dynamicProperties because there was too much variability in what was going in before. The term does get used and it's one of the only ways to express information that doesn't have a home elsewhere in Darwin Core. If it's formatted correctly it's more useful any other solution because it allows for key:value pairs,

dictionaries, etc. Other than people not following the instructions, there haven't been any reported problems with it.

James notes that most data are being captured through collection management systems and spreadsheets. Those are only as good as what they allow in a given field and they are only as good as the documentation and best practices implemented in a particular institution. We could re-interpret them as JSON when we share them if they were put into the system in a way that made the collection manager's lives easy. The issue is with the people who will actually be typing the stuff in.

Steve: many of the use cases are coming from new standards or additions, so there isn't a prior practice that needs to be changed. So we want fix this before each of the standards coming out of the pipeline come up with their own solutions.

James notes that for something like a region of interest in an image, the values will be coming out of software. But for something like the vernacular names use case, those will probably be typed in by a human. So we need to consider both the scripters and the inputters.

Ben mentions that we should also have recommendations for the properties that link to the arrays. If we have instances of the VernacularName class, the property should be hasVernacularName.

Steve notes that it's hard to type in any JSON more complicated than a list. For example, dictionaries nested inside lists are very tough to type by hand without error. But we also have a lot of smart people who can build tools.

Jonathan clarifies that this comes out of the desire to put this all in a single file. Because often what you do is create a second file (e.g. star schema).

Steve commented that once it gets complicated enough, the new GBIF model may just come into play. So the use of these recommendations would probably be only for the simplest cases where it makes sense to put stuff into one cell. Once you get a more complicated structure, then you basically need another table. So part of the charge for developing some recommendations would be to determine how simple or how complicated is appropriate to take this kind of approach.

Ben mentioned that there should be rules, such as if a species has no vernacular names, an empty array should be given as the value.

Conclusion: there seems to be a consensus that this is a problem and if we could come up with a solution, that would be a valuable thing.

## VII. Discuss future work for 2023

- A. Review standards status and recommend standards to be retired. Non-retired standards should have some mechanism for their maintenance. See:

1. <https://github.com/tdwg/tag/issues/21>
2. <https://github.com/tdwg/tag/issues/20>
3. <https://github.com/tdwg/tag/issues/16> (This is a 2017 task from the Executive committee. What does it actually mean? **One issue with retiring standards is to know what exactly does it mean to say that a standard is "retired"?**)

David made the point that there is a distinction between standards that aren't maintained any more and those that aren't used any more. Some standards may not be maintained because there isn't an active group to maintain them. But they may still be used. The criterion for retiring a standard should be when it isn't used any more.

Steve: not a high priority, but could be addressed if we have time. Perhaps combine into one issue.

- B. Other issues **Did not discuss due to lack of time.**

1. <https://github.com/tdwg/tag/issues/28> (anthropological/cultural knowledge terms needed across vocabularies - needs a champion)
2. <https://github.com/tdwg/tag/issues/22> (consider use of SKOS-XL for labels across TDWG vocabularies; maybe under consideration by TCS2?)

VIII. Any additional announcements or new business.

- A. Views controlled vocabularies are now in the public comment period. See <https://github.com/tdwg/ac/issues/245> and comment.
- B. Revise TAG membership for 2023.
1. Constitutional limit of "fewer than 15 members"
  2. One year commitment.
- C. Report from the TDWG Unconference Group about Standards Mappings (David) **David: this has come up multiple times and involves many of the standards. Mapping standard SSOM was mentioned. There is now a channel in the regular TDWG Slack (not the annual meeting Slack): #mappings-between-standards for discussion.**

Holly Little suggested using the paleo group as a use case because there are a lot of different standards in that field. There was also some question about what kind of group would be needed.

1. Is this something that should/could be pursued under the umbrella of the TAG? **We should discuss this in the next meeting since there is a lot of interest and it falls within the TAGs area of interest.**

Steve: Is there enough interest to charter a task group. Interested people should discuss who could be core members and identify a convener.

- D. Status of the old TDWG-TAG Mailing List: <https://lists.tdwg.org/mailman/listinfo/tdwg-tag> should we start actively using it or retire it for good? (suggested by David)

1. Comment by David: This is just something I noticed recently. It is not urgent, so feel free to postpone this to a future meeting if time runs out.
  2. This group decided that Slack would be it's primary communications method, so we can just ignore this mailing list for now.
- E. There is currently an incompatibility between ABCD and ABCDEFG, how to handle this with ABCD not being a current standard?  
<https://github.com/tdwg/abcd/issues/15>  
Comment made after meeting (Sam Leeflang): Was discussed later on the day in the ABCD maintenance group. We will add types to the recordBasis to make ABCD (3.0) more useful for geocollections data.

IX. Action items for next meeting (or before):

- A. Ian to convene a meeting to identify participants to form a task group for revising the GUID and LSID Applicability Statements.
- B. Those interested in the mapping issue should try to identify potential core members who might also serve as a convener. We can then try to make progress in chartering a task group.
- C. Steve will clean up the issue tracker as discussed in numerous places above.
- D. Steve and Ben will work on a recipes document and revision of Ben's white paper on booleans. Steve to investigate possible places on the website where TAG recommendations would live.
- E. Additional discussion of a course of action on complex values.