TDWG Technical Architecture Group (TAG)
Online meeting
2023-07-10 13:00 UTC

Agenda/Meeting notes:

Participating:
- Steve Baskauf
- James Macklin
- Ben Norton
- Kit Lewers
- Raïssa Meyer
- John Kunze

Regrets:
- David Fichtmueller
- 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.

I. Set time and date for next meeting: 2023-09-11 13:00 UTC.
II. New items
    A. Formation of Mappings Task Group under the TAG.
        1. They will be working out a charter at their next meeting in September.
        2. Held a meeting 12 June, notes at https://docs.google.com/document/d/1EuqrWrOSCSio4yoaQkjhFm_Mw1hZ561XlBWLVXraJBY/edit#heading=h.9lsegj2qwwe1
        3. They will be reporting on their work at the session on Standards Mapping.
        4. See more info in the announcements section.
    B. Categorization of prior standards and retiring standards. (Kit, Steve)
        1. Background
            a) "Standards Status and Categories" document (https://www.tdwg.org/standards/status-and-categories/) defines 5 categories: Current Standard, Current (2005) Standard, Draft Standard, Prior Standard, and Retired Standard. There has been a long-standing interest in simplifying this system.
            b) Draft Standard is a temporary category, so we can disregard it.
            c) Current (2005) Standards should move into the Current Standard category by conforming with the SDS (formatting) and VMS (for vocabularies: have a Maintenance Group) and formally undergoing review. In this category: ABCD, TCS, Structured Descriptive Data (SDD), and DELTA. ABCD and

TCS are under active development within the TDWG community and should eventually move, DELTA and SDD are not.

d) There are currently no retired standards.
e) There are many Prior Standards. Some have the potential to move to Current Standards, some should probably be retired. It would be good, but probably difficult, to eliminate this category.
f) In the end we would like to have Current Standards, Draft Standards, and Retired Standards as the only categories.
g) Breakdown of existing standards by Kit's action categories: https://docs.google.com/document/d/17xfCJbz8Kb5MP5C2_9njibMikVUvbM0xrSi2EewMxa0/edit?usp=sharing

2. This has been an issue for years, old issues summarized in umbrella issue https://github.com/tdwg/tag/issues/37
3. Kit's categorization of Current (2005) and Prior Standards; research into use of standards
4. Usage survey to be administered at the annual meeting.

Notes: SDD could move into current standards, but does not have a champion, all authors are no longer active in TDWG or deceased. We will see how many people are using it when we do the survey. See the breakdown document for notes on how prior standards could be categorized.

Steve: What we really want in the end is just two categories: current standards and retired standards, with standards that are used and have "champions" being moved into current standards, and standards that are not being used (probably because they've been superseded by some modern standard being moved to the retired standard category.

There is actually a third category: former standards that have basically been turned into databases that are being managed by some group outside of TDWG. These former standards could be retired because basically TDWG isn't involved in their management. We could still list them as recommended for use, but take the TDWG label off of them.

Ben: Is it actually possible to determine that a standard isn't being used? Steve: little-used standards won't go away, they'll just get moved to another page so as to not confuse people.

Raïssa: It's hard to define a standard as not being used by people because that can change at any time. In contrast, it's very easy to say that a standard isn't being maintained any more. It's easy to tell when a standard is being worked on again.

Steve: the "Categories and Status" page doesn't define "Retired Standard", so we really could define it any way that we want. We

could define one as "nobody cares enough about it to maintain it". However, if there is a modern standard that has replaced an earlier one, we actually don't want people to use it or work in it. The last thing TDWG needs is duplication of effort. So it would make sense to essentially deprecate standards if they've been superseded by something else.

John: perhaps label them with something that corresponds to an action. so as to steer people who are doing new work to the actively maintained ones. If some old database marked up on an old standard suddenly became popular, then the old standard would be highly relevant, a third is ones that TDWG is actively looking at but not managing.

Steve: The modern standards are typically about how to transfer data, while the old ones are more about how to set up database records. Since our current focus is on interoperability and data sharing, then we don't care so much if someone has their own personal database set up according to an old standard. What we care about is that if they want to share those data, they should use one of our modern standards.

Steve: in the database "standards" there isn't really control by a community process as there is in our modern standards. So they aren't really "standards" as we use the term, although they are sort of like controlled vocabularies.

Steve: there isn't an established process for retiring standards – can the executive committee just do it? Probably there should be a 30 day public comment period just like there is for ratifying a standard.

Proposed course of action: We recommend
1. definitions for the categories
2. which standards should fall into those categories
3. that there be a 30 day public comment period managed by the TAG and look for consensus in the same way we manage the other standards processes. After that, make a recommendation to the Executive Committee.

III. Follow-up on items from previous meetings
    A. Followup on 2023-03-13 meeting, topic of feedback/suggestions for standards without maintenance groups: they all have GitHub repos with issue trackers. However, some of the trackers are rarely used and may not be monitored. But suggestions can be recorded there.
    Monitoring these standards and potentially starting up a task group if changes are necessary really seems like something that the Executive Committee should manage, not us.
    B. Status of the Identifiers Task Group (Ian, Steve)

1. Dave Bloom has agreed in principle to convene the group.
2. We have been working on drafting a charter.
3. It is somewhat unclear whether there should be an actual Interest Group (in which case they would sponsor the TG) or only a Task Group (sponsored by the TAG).
The opinion was expressed that in the interest of keeping down the number of groups that need to be managed, it would be better to just have a task group. But Dave and recruited core members would need to scope the work of the group while writing the charter.

C. Best practices for borrowing terms from non-TDWG vocabularies (held over from previous meeting).
1. Two meetings ago we said we could continue discussion on this at today's meeting. See notes from meeting, new items B.
2. Steve's final comment: The solution that most people seemed comfortable with was: if "same as" relationship: borrow, if not: mint. Audiovisual Core didn't really do that and therefore has "usage guidelines" for a number of borrowed terms. But that was a decision made at that time.
3. Is this something that the TAG wants to make a statement about with respect to a best practice? Or should we just let groups do what they think is best? For example, the CamtrapDP specification mints an entirely new set of terms, some of which have sameAs relationships with other terms.
Ben: sameAs means exactly the same definition and scope.

James: what happens if you change the definition? Answer: then you need to mint your own term. The Humboldt Extension had to do that with a term that means mostly the same thing as a DwC term, but not exactly the same thing.

Steve: sometimes the organization from which we've borrowed the term makes a change. AC has a policy that if a borrowed TDWG term changes, AC updates to the new version automatically. For non-TDWG borrowed terms, the version is frozen at what it was at the time it was borrowed unless there is an explicit change to update it. Unfortunately, some of the external vocabularies don't have robust versioning.

Some of this could get fleshed out better by the mapping group, since it's related.

There was sentiment for recommending the Latimer Core approach (borrow if exact match, mint and map if not) only. However, Steve thought that some people might make a strong argument for the Camtrap DP approach. Frictionless doesn't support namespaces, but TDWG uses them. One possibility is to follow a JSON-LD-like system where a context document defines the property names by referring to

   D. Recommendations for expressing complex values
     1. This needs more work, but Steve hasn't focused on it since the last meeting. Notes at:
https://github.com/tdwg/tag/blob/master/meetings/2023-05-08-tag-meeting-notes.pdf
     2. Unless someone has something to say, we will probably need to defer on this to a future meeting.

 IV. Any additional announcements.

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

A. Camilla: volunteered to do the Spanish translations for the Boolean vocabulary.
B. Steve and Kit will get together to work on a proposal for categorizing the standards.
C. Steve will turn Ben's boolean reference doc into Markdown so that it can go with the other docs.