TDWG Identifiers discussion

Meeting date: 21 March 2023
A suggested home for the proposed identifiers Task Group is https://github.com/tdwg/guid-as

Invitees:

- Rod Page rdmpage@gmail.com
- Jonathan Rees jar@mumble.net
- Rob Sanderson robert.sanderson@yale.edu
- Larry Lannom llannom@cnri.reston.va.us
- Paul Morris (2017) mole@oeb.harvard.edu
- Dimitris Koureas (2017) dimitris.koureas@naturalis.nl
- Dag Endresen (2017) dag.endresen@nhm.uio.no
- Joel Sachs (2017) joel.sachs@agr.gc.ca
- Pier Buttigieg (2017) pbuttigi@mpi-bremen.de
- Jeff Gerbracht (2017) jeff.gerbracht@cornell.edu
- Stan Blum (2017) stanblum@gmail.com
- Donald Hobern (2017) donald.hobern@adelaide.edu.au
- Roger Hyam r.hyam@rbge.org.uk
- Wouter Addink wouter.addink@naturalis.nl
- Tim Robertson trobertson@gbif.org
- Steve Baskauf steve.baskauf@vanderbilt.edu
- Debbie Paul dlpaul@illinois.edu
- John Weiczorek tuco@berkeley.edu
- Holly Little (standing in for colleagues)
- Also to broadcast on TDWG, iDigBio, SPNHC, and DiSSCO networks

Apologies:

- Rod Page
- Donald Hobern

Participants:

- Dimitris Koureas
- Wouter Addink
- Roger Hyam
- Ian Engelbrecht

Purpose:

The purpose of this meeting is to discuss the formation of a Task Group to review the TDWG guidelines on identifiers (GUIDs and LSIDs). We need to decide if we will form a Task Group, who will be the members, who will be the convenor, the specific purpose/goals of the Task Group, and to have some discussion of the key issues around identifiers. If the decision is to go ahead with a task group we can do some planning too.

Terminology:

When referring to identifiers we mean GUIDs, PIDs, LSIDs, UUIDs, DOIs, Handles, URIs, or anything people currently think of when talking about identifiers for entities and/or data records in the biodiversity data space.

Key Questions for the discussion:

- What exactly are we trying to do with identifiers (disambiguation vs linking)?
- There's been a lot of work on identifiers (see reading list below). Is this still a question, and if so, why?
- What do we need from a TDWG PIDs Guideline document?
- Where are the points of contention?
- Where have we got it wrong before, and what can we learn from this?
- If I want to store a PID in my database how do I do that?
- If I want to create a new PID, which one is best?
- How do I find out if a PID exists for something?
- What can a PID tell me about the thing it identifies?

Selected background reading (in no particular order):

Discussions already held recently in this regard, primarily through the TDWG TAG group are at https://github.com/tdwg/tag/issues/36

A valuable recent synopsis of the use of identifiers in biodiversity informatics: Agosti et al. (2022) Recommendations for use of annotations and persistent identifiers in taxonomy and biodiversity publishing. Research Ideas and Outcomes 8: e97374. https://doi.org/10.3897/rio.8.e97374

iDigBio GUID Guide for Data Providers (2013): https://www.idigbio.org/sites/default/files/internal-docs/idigbio-standards/iDigBioGuidGuide-2013-06-26.pdf

GBIF Beginners Guide to Persistent Identifiers (2011): https://doi.org/10.35035/mjgq-d052

CETAF Stable Identifier Guide: http://cetafidentifiers.biowikifarm.net/wiki/Main_Page

European Open Science Cloud PID policy: https://doi.org/10.2777/926037

Guralnick et al. (2015). Community next steps for making globally unique identifiers work for biocollections data. ZooKeys 494:133-154. https://doi.org/10.3897/zookeys.494.9352

Page, R. D. (2008). Biodiversity informatics: the challenge of linking data and the role of shared identifiers. Briefings in Bioinformatics, 9(5), 345-354. https://doi.org/10.1093/bib/bbn022

Nelson et al. (2018). Use of globally unique identifiers (GUIDs) to link herbarium specimen records to physical specimens. Applications in Plant Sciences, 6(2), e1027. https://doi.org/10.1002/aps3.1027

Jones et al. (2011) Identifiers in the Catalogue of Life: https://doi.org/10.1186/2041-1480-2-7

Martin et al. 2005. The impact of Life Science Identifiers on informatics data. https://doi.org/10.1016/S1359-6446(05)03651-2 (not open access unfortunately)

McMurrey et al. (2017). Identifiers for the 21st century. PLOS Biology https://doi.org/10.1371/journal.pbio.2001414

Juty et al. (2020). Unique, Persistent, Resolvable: Identifiers as the Foundation of FAIR. https://doi.org/10.1162/dint_a_00025

Hardisty AR, Addink W, Glöckler F, Güntsch A, Islam S, Weiland C (2021) A choice of persistent identifier schemes for the Distributed System of Scientific Collections (DiSSCo). Research Ideas and Outcomes 7: e67379. https://doi.org/10.3897/rio.7.e67379

See also the references in the above articles for more, there is a lot that's been written about this topic.

LSIDs

For reference, Tim Roberston provided the following screenshot from the TDWG Executive Committee Github repo regarding the decision on LSIDs, and indicated that GBIF continues to maintain lsid.info. TDWG owns and pays for the domain name.

ask TAG to retire GUID and Life Sciences Identifiers Applicability Statements standard? #90

Edit   New issue

⊘ Closed   **wouteraddink** opened this issue on Jul 14, 2017 · 4 comments

---

wouteraddink commented on Jul 14, 2017

If I remember correctly, at the 2016 conference LSIDs were officially deprecated by TDWG. However, on our website we still recommend them?? The GUID and Life Sciences Identifiers Applicability Statements is still listed as current standard. I think we should ask TAG to retire this standard. A new one should be created, and preferably should refer also to CETAF identifiers.

---

dkoureas commented on Jul 18, 2017

We have decided to retire the resolution service, but I am not sure whether we took the decision to deprecate the Standard.

⇥ **stanblum** added this to the **August 2017** milestone on Jul 24, 2017

---

ghwhitbread commented on Aug 29, 2017 · edited ▾

From TAG works in progress

#.2 Update the TDWG GUID applicability standard documents at
https://github.com/tdwg/guid-as or retire standard and replace with GUID recommendation web pages.

- ⊘ **Reimburse Darwin Core Group travel costs as approved at Jonkoping** #7 Revise https://github.com/tdwg/guid-as/blob/master/guid/applicability_statement.doc
- ⊘ **Role of former chair of conference organisation in TDWG 2015** #8 Deprecate https://github.com/tdwg/guid-as/blob/master/lsid/applicability_statement.doc

#.7 Revision of guid-as

- review implementation recommendations and principles
- revise recommended GUID lists
- update meta-data serialization specification (currently RDF+XML, XML)

#.8 Deprecate lsid-as

- Remove from guid-as
- Annotate as deprecated
- Incorporate principles into guid-as ((authority, namespace, resource-identifier), GET meta-data)

---

gkampmeier commented on Sep 4, 2017

Please note implications of the use of our service in Darwin Core documentation e.g.,
http://rs.tdwg.org/dwc/terms/guides/text/index.htm

⇥ **dkoureas** modified the milestones: **October 2017 (@TDWG17)**, **August 2017** on Sep 11, 2017

↗ **ghwhitbread** mentioned this issue on Sep 22, 2017

Revise GUID standard tdwg/tag#9   ⊘ Closed

⊟ 3 tasks

---

dkoureas commented on Oct 18, 2017

The XC has agreed (XC1709) to deprecate the GUID AS.

Sidebar

Assignees ⚙
No one—assign yourself

Labels ⚙
None yet

Projects ⚙
None yet

Milestone ⚙
October 2017 (@TDWG17)

Development ⚙
Create a branch for this issue or link a pull request.

Notifications   Customize
🔕 Unsubscribe
You're receiving notifications because you're watching this repository.

3 participants

🔒 Lock conversation
⚲ Pin issue ⓘ
→ Transfer issue
🗑 Delete issue

Meeting notes

Participants:

Wouter Addink
Roger Hyam
Dimitri Koureas

Discussion points:

- There were quite strongly differing opinions voiced about identifiers during the discussion, ranging from the question of identifiers having been resolved and that no further meetings or discussions are required, to the question of identifiers not being resolved and a viable solution that data managers and publishers can implement on the ground still needs to be identified.
- There are also still strongly divergent opinions about the value and ongoing use of LSIDs.
- As attendance of the meeting was limited, no task group convenor was identified, and the question of how to proceed with revising the GUID and LSID Applicability Statements was returned to the TDWG TAG.