

Topic: Minimum Information about a Digital Specimen - MIDS

Date: 2024 July 11

Time: 12 PM Eastern

In Attendance

Holly Little, NMNH	Lindsay Walker, Symbiota	Erica Krimmel
Roger Burkhalter, Retired Sam Noble Museum	Talia Karim CUMNH	Natalia Lopez Carranza, KUMIP
Suzanne Mills, Carnegie MNH IP	Amanda Millhouse, NMNH	Jessica Utrup, YPM
Alex Cano, NMNH	Sylvia Orli, NMNH	
Cat Chapman, iDigBio	Jessica Bird, NMNH	
Chris Maves, FMNH	Simon Goring, Neotoma	
Alejandra Martinez-Melo, ANSP	Jacob Van Veldhuizen, CUMNH	
Adam Mansur, NMNH	Wai-Yin Kwan	

Resources

- Paleo Data Working Group [Google Folder](#)
- MIDS information on the TDWG website: <https://www.tdwg.org/community/cd/mids/>
- MIDS GitHub: <https://github.com/tdwg/mids>
- MIDS Calculator: github.com/AgentschapPlantentuinMeise/MIDSCalculator

Agenda

- Information sharing from Cat Chapman, co-convenor of the MIDS Task Group

Action Items

-

Notes

- MIDS = Minimum digital specimen completeness
 - Level 0-3, where 0 is bare minimum (see screenshots below)
 - Not intended to measure or enforce data quality, but does assess data richness
 - Meant to complement existing data in DwC format

- Ideal would be to have MIDS integrated into CMS to help assess the completeness of your record → use as an evaluation tool
 - Requirements may vary by collection type
- <https://github.com/AgentschapPlantentuinMeise/MIDSCalculator>
- <https://github.com/tdwg/mids>
- Discussion
 - Some specimens may not have a physical voucher, e.g. observations, eDNA, but this may be a separate issue from minimum data required for a record
 - Some things may change once material sample is fully hashed out
 - MIDS could be used as justification for internal data improvement, e.g. in grant writing justification for funding or strategic planning
 - How do you distinguish records that can be improved vs. cannot?
 - No dependencies specified yet, e.g. if you have decimalLat vs no long
 - Could different CMSes give you a different MIDS score?
 - Need to be careful of comparing collections using metrics due to “apples vs oranges” phenomenon
 - Could be run via the GBIF IPT as an implementation option, within a CMS, or otherwise. Multiple implementation methods may be necessary since not everything can go to GBIF, e.g. Neotoma’s data is not in GBIF
 - Ways to provide feedback:
 - MIDS task group meets every Thursday → email Cat to participate/get added to mailing list: cchapman@floridamuseum.ufl.edu
 - Comment on GitHub issues

MIDS 0

Skeletal record

MIDS information element	Example (Darwin Core terms)
PhysicalSpecimenID	occurrenceID, catalogNumber, materialSampleID
Organization	institutionID, institutionCode, ownerInstitutionCode

- Not all terms are required to satisfy the element, e.g. you could just use occurrenceID to satisfy PhysicalSpecimenID

MIDS 1

everything in MIDS-0, plus...

MIDS information element	Example (Darwin Core terms)
Name	scientificName, vernacularName, scientificNameID, organismName, title
SpecimenType	basisOfRecord
ObjectType	preparations
License	license
Modified	modified

MIDS 2

Everything in MIDS-0 + MIDS-1, plus...

MIDS information element	Example (Darwin Core terms)
qualitativeLocation	locality, county, country, municipality, verbatimLocality, island, verbatimCoordinates, verbatimLongitude, verbatimLatitude, continent, countryCode, islandGroup, stateProvince, waterBody, higherGeography
quantitativeLocation	locationID, decimalLatitude, decimalLongitude
GeologicalAge	earliestAgeOrLowestStage, earliestEonOrLowestEonothem, earliestEpochOrLowestSeries, earliestEraOrLowestErathem, earliestPeriodOrLowestSystem, latestAgeOrHighestStage, latestEonOrHighestEonothem...
CollectingAgent	recordedBy, recordedByID
CollectionDate	eventDate, Year, verbatimEventDate
CollectorNumber	fieldNumber, recordNumber
Media	mediaType, associatedMedia, accessURI

- GeoAge obsv only applies to paleo, not other collection types, e.g.
<https://github.com/tdwg/mids/issues/61>

MIDS 3

Everything in MIDS levels 0-2*, plus...

MIDS information element	Example (Darwin Core terms)
GeographicalLocalityID	geodeticDatum, coordinateUncertaintyInMeters, coordinatePrecision, footprintSRS, locationID, decimalLatitude, decimalLongitude, footprintWKT
InstitutionID	institutionID
CollectorID	recordedByID
ScientificNameID	scientificNameID
IdentifiedByID	identifiedByID
AssociatedMediaID	identifier
????	under development/discussion

*requirements will vary based on collection type

- <https://github.com/tdwg/mids/issues/81>

Chat Copy

11:59:57 From Erica Krimmel to Everyone:

Shared notes for today

12:02:01 From Erica Krimmel to Everyone:

For anyone joining in the last couple minutes, here are our shared notes for today

12:02:09 From Alex Cano to Everyone:

Reacted to "For anyone joining i..." with 👍

12:04:18 From Alex Cano to Everyone:

shared notes for today

12:04:18 From Holly Little to Alex Cano(Direct Message):

I didn't take host from you so can you make Cat host

12:04:24 From Holly Little to Alex Cano(Direct Message):

sorry co-host

12:05:21 From Lindsay Walker to Everyone:

I like it!

12:05:21 From Erica Krimmel to Everyone:

I love the shifting colors!!

12:05:34 From Amanda Millhouse (she/her) to Everyone:

Subtle pink hues. Very cool!

12:08:00 From Simon Goring (he/him) to Everyone:

Will we be able to get a copy of this presentation?

12:08:08 From Alex Cano to Everyone:

Reacted to "Will we be able to g..." with 👍

12:08:28 From Erica Krimmel to Everyone:

Reacted to "Will we be able to g..." with 👍

12:08:52 From Alex Cano to Everyone:

Replying to "Will we be able to g..."

We will link to it in the shared notes for today

12:09:28 From Simon Goring (he/him) to Everyone:

Replying to "Will we be able to g..."

Thanks :)

12:10:30 From Suzanne Mills to Everyone:

Reacted to "Will we be able to g..." with 👍

12:13:55 From Talia S. Karim (She/Her) to Everyone:

So its more of an evaluation tool

12:16:13 From Simon Goring (he/him) to Everyone:

github.com/AgentschapPlantentuinMeise/MIDSCalculator

12:18:31 From Erica Krimmel to Everyone:

Probably hardly rent free in your mind tho 😊

12:18:44 From Cat Chapman - (she/her/hers) to Everyone:

Reacted to "Probably hardly rent..." with 😊

12:18:46 From Cat Chapman - (she/her/hers) to Everyone:

Reacted to "Probably hardly rent..." with 😊

12:19:42 From Simon Goring (he/him) to Everyone:

We've got over 10 million observations, so it's not quite fringe :p

12:19:57 From Talia S. Karim (She/Her) to Everyone:

I was just thinking the same thing about observations

12:20:35 From Simon Goring (he/him) to Everyone:

We could show completeness at each level:

Level 0: 2/2

Level 1: 3/5

Level 2: 2/6

12:24:41 From Talia S. Karim (She/Her) to Everyone:

Also very useful for strategic planning in collections

12:25:38 From Talia S. Karim (She/Her) to Everyone:

I had that same question Adam

12:25:41 From Suzanne Mills to Everyone:

good point, Adam

12:25:41 From Erica Krimmel to Everyone:

Reacted to "I had that same ques..." with 🏆

12:25:45 From Lindsay Walker to Everyone:

+1

12:25:46 From Simon Goring (he/him) to Everyone:

Reacted to "Also very useful for..." with 👍

12:31:59 From Adam Mansur to Everyone:

NMNH has something similar in place for calculating record quality scores in EMu. Not based on MIDS but it's the same idea.

12:33:05 From Cat Chapman - (she/her/hers) to Everyone:

<https://github.com/tdwg/mids>

12:33:33 From Cat Chapman - (she/her/hers) to Everyone:

<https://github.com/AgentschapPlantentuinMeise/MIDSCalculator>

12:34:58 From Cat Chapman - (she/her/hers) to Everyone:

<https://github.com/tdwg/mids/issues/61>

12:39:45 From Adam Mansur to Everyone:

This is the issue for GeographicalLocalityID: <https://github.com/tdwg/mids/issues/81>

12:41:10 From Cat Chapman - (she/her/hers) to Everyone:

[ccchapman@floridamuseum.ufl.edu](mailto:cchapman@floridamuseum.ufl.edu)

12:45:48 From Cat Chapman - (she/her/hers) to Everyone:

we are in talks w/GBIF about this and they are onboard with the idea of MIDS... so maybe implementation will happen, eventually :)

12:47:27 From Adam Mansur to Everyone:

MIDS does have keywords for missing data, including withheld

12:47:38 From Cat Chapman - (she/her/hers) to Everyone:

at the end of the day, the way I see it: implementation (see: software development) will be the biggest roadblock to MIDS being used, simply because dev work isn't easy.

12:50:46 From Simon Goring (he/him) to Everyone:

https://github.com/AgentschapPlantentuinMeise/MIDSCalculator/blob/main/data/schemas/DwC-GBIF_schema.json#L111

12:54:02 From Simon Goring (he/him) to Everyone:

If we choose to move forward with this, who would we talk to?

12:54:32 From Erica Kimmel to Everyone:

<https://paleo-data.github.io/grscicoll>

12:56:03 From Cat Chapman - (she/her/hers) to Everyone:

Replying to "If we choose to move..."

moving forward with using that tool? Mathias would be a good contact for that.

12:56:11 From Simon Goring (he/him) to Everyone:

Replying to "If we choose to move..."

Thanks!