

**Meeting Minutes from the DDWG-2 telecons:**

[DDWG Notes 2023-01-19](#)

[DDWG Notes 2023-02-02](#)

[DDWG Notes 2023-02-09](#)

[DDWG Notes 2023-02-16](#)

[DDWG Notes 2023-02-23](#)

[DDWG Notes 2023-03-02](#)

[DDWG Notes 2023-03-16](#)

[DDWG Notes 2023-03-30](#)

[DDWG Notes 2023-04-13](#)

[DDWG Notes 2023-04-27](#)

[DDWG Notes 2023-05-11](#)

[DDWG Notes 2023-05-25](#)

[DDWG Notes 2023-06-08](#)

[DDWG Notes 2023-07-06](#)

[DDWG Notes 2023-07-20](#)

[DDWG Notes 2023-07-25 \(F2F\)](#)

[DDWG Notes 2023-08-03](#)

[DDWG Notes 2023-08-10](#)

[DDWG Notes 2023-08-17](#)

[DDWG Notes 2023-08-31](#)

[DDWG Notes 2023-09-14](#)

[DDWG Notes 2023-09-28](#)

[DDWG Notes 2023-10-26](#)

[DDWG Notes 2023-11-09](#)

[DDWG Notes 2023-11-14 \(F2F\)](#)

[DDWG Notes 2023-12-07](#)

## DDWG Notes 2023-01-19

---

title: DDWG Notes 2023-01-19

layout: default

date: 2023-01-19

---

January 19, 2023

Notes by Debra Kazden

Known Attendees: M. Drum, M. Gordon, E. Guinness, T. Hare, L. Huber, S. Hughes, R. Joyner, D. Kazden, T. Lim, J. Mafi, S. McLaughlin, B. Semenov and R. Simpson  
Known Observers: R. Alanis, A. Bailey, M. K. Crombie, P. Lawton, S. Loftin, M. McAuley, E. Schaefer, C. Seyb and M. Tiscareno

### ## DDWG Agenda

(Included in email sent January 17, from M. Drum, Agenda for January 19, 2023)

#### Front-matter/What's new

- Votes! We've got votes!
- Approved SCRs by March 2, 2023
- \*\* (Discussed) \*\*

#### Ready for Vote

- CCB-354 – Add new permissible values to Header/parsing standard id to support new versions of CDF – Joe Mafi
- CCB-355 – Add Funding Acknowledgement class to Citation Information Class – Myche McAuley
- CCB-359 – Allow Product Ancillary to include dictionaries – Anne Raugh \*\* (Did not vote) \*\*

#### Tier 1 issues

- CCB-357 – Create Product Annex – Now Product External – Lyle Huber
- CCB-325 - Support for video as product observational – Trent Hare
- CCB-211 – XML/schema based files as archival data – Jesse Stone, Mike Drum
- CCB-358 – Context object for Instrument Package – Anne Raugh - Counter proposal for Instrument Host amendments \*\* (Discussed 357, 325) \*\*

#### Tier 2 issues

- CCB-336 – Add a License Information class to the Identification Area – Tanya Lim - We know that we need to at least support CC0
- DOI-rich metadata into IM (ticket to be created), replacing CCB-222 and CCB-261 – Anne Raugh and any volunteers
- CCB-164 - Display Settings not required for images -- (IMG) - Need to think more about potential updates to Imaging and how Display might fit.
- CCB-353 – Fully label all bytes in data files – Anne
- CCB-356 – Update Constraints on Author/Editor Lists – Dick Simpson
- CCB-350 – Adopt more rigorous / stringent rules for leap seconds
- Product Metadata Errata? Create a supplemental label to live next to original label, that can add aliases and supersede metadata \*\* (Not Discussed) \*\*

Parking lot – need a driver!

- CCB-326 - Inconsistent Direction on How to Use Discipline Facets
  - CCB-324 – Create a data type for Type attributes that require a permissible value list.
  - RS: 259/281 (see new attachment to CCB-281) (on hold)
  - CCB-261 - Add an optional method for indicating that a product is a member of a larger collection that has a DOI -- Last chance before recommending withdraw.
- \*\*(Not Discussed)\*\*

Next meeting: 2/2/23

# DDWG

## Front-matter/What's new

There are a lot of people here today.

Voting today. TAs were completed. Hopefully people had a chance to look everything over.

For CCB-359 (Product Ancillary does not allow inclusion of either a Discipline Area or a Mission Area, see <https://pds-jira.jpl.nasa.gov/browse/CCB-359>) the TA revealed that we don't have to do it. Withdrawn.

# SCRs

## CCB-355 - Add Funding Acknowledgement class to Citation Information Class See <https://pds-jira.jpl.nasa.gov/browse/CCB-355>

There were comments in Jira on this.

~ There is a TA. Some definitions.

Question: Are people comfortable voting on this today? With liens? Want to discuss it?

Answer: Someone has a question based on the comments in Jira.

Another Question: Is there any reason this needs to be a class, and not just an attribute?

Another Question: How would we structure that?

Answer: Various aspects of attributes of things we plan to track. That's why it's associated with a class.

~ Separate attributes. To have them searchable.

~ Might want to search on all program elements.

~ Someone said they didn't know we could search on all attributes in the registry.

~ Yes, but not perfect for everything.

~ Still an example of a use. Not an SCR yet. Needs specification for each attribute in class and to show in the parent class that this is optional. Basically, still need to put in all the details.

~ Thought we did that.

Someone is looking at the Requested Changes. Not sure how this is not an SCR.

~ The implementer, Steve in the TA, typically implements exactly what's there.

~ TA did allude to that.

~ Reporter can fix it.

\*\*Action Item - Steve\*\* will send an example to \*\*Kate\*\* and she will fix the SCR.

If there's an enumerated list, it probably won't be useful outside NASA. ESA is not a funding source in Europe. Probably won't use it.

- ~ The example of NASA as a funding source was just an example. Could use free text.
- ~ Not great for search. Could have program. Could add European funding sources.
- ~ Usually each country is funded by their own sources.
- ~ Someone thought this is for research projects, not spacecraft data.
- ~ Someone else thought of it more broad. Could have it conditionally multiple, to put in everything.
- ~ Could be a network of funding for spacecraft. Could be useful for research projects.
- ~ Right now, most ESA data is associated with space missions. Just an observation. Not against this.

We should postpone this.

Question: Is under citation information where we want this?

Answer: For how to cite this.

Another Question: Yes, but isn't funding for the archive?

Answer: Never mind.

- ~ Someone would use citation information if they were citing the data. Having funding source would be good. Having enumerated list would be very helpful, but an enumerated list for NASA isn't helpful if something is being archived at JAXA or ESA or somewhere else. Could have enumerated list for funding source NASA, JAXA, etc., and sub-funding for programs. Complicated.
- ~ Funding source and program element would be useful for search. Civil servants would love this.

Question: Is program element also used outside NASA?

Answer: Program element is NASA speak, but we could use something similar. Plus, this is optional.

- ~ Difficult to put in an enumerated list. Too many and complicated.
- ~ And changes annually.
- ~ Once the list is too big no one would look and would just add what they want.
- ~ So if NASA, then program element and external source can be used as necessary. Can have enumerated list for NASA and not have to use it if not NASA.
- ~ Makes sense, but not sure we need an enumerated list for search the way Jordan is going anyway. On board with free text.
- ~ Could use award number and not search on funding source or program element.
- ~ Not sure how useful that would be.
- ~ We can leave it to the discretion of the SCR authors.

Not hard to make updates to enumerated list. Search works better with it. Tools will need to be more structured with out enumerated list.

- ~ Regarding the maintenance of an enumerated list - will need some more than others. Typically then would put in an LDD.
- ~ But if in an LDD it goes in a different part of the label.
- ~ Right. Not sure that matters. But if it needs frequent maintenance...

Question: If it's in an LDD can it be called from the citation area?

Another Question: Like with a reference, maybe?

- ~ Someone agrees it should be in the citation area.
- ~ Could be internal reference or guess maybe put an any clause to put it in citation. Never done it before, but it might work.

Question: How does instrument component work?

Answer: Only in context products.

Another Question: Is that something that could have different values? Would use enumerated list if consistent, but if well defined could use free text.

~ Could have how to form values in the DPH, like for CDAP etc.

~ Can update values if incorrect.

~ Check in peer review. How to form the values would be in the DPH, not an enumerated list.

~ Someone is beginning to think this should be in an LDD. The maintenance for an enumerated list would be a lot easier.

Question Is anyone signing up to be a dictionary maintainer for this?

Answer: No. This is just trying to put information together. Award numbers shouldn't be an enumerated list, but it's what you would want to search on. This is optional. Would vote for free text and leaving it optional.

~ Reporters should put what they want.

~ Someone agrees. Let them work on it.

~ Uncontroversial.

~ Just need to make it more clear and focused so Steve can cut/paste.

Question: What's wrong with just putting in a description? Seems like overloading us with attributes is getting out of hand.

Answer: Cutting this off here. We will vote in two weeks.

Another Question: Homework?

Answer: Steve will provide an example of the format that can be copied for the SCR. **\*\*(Action Items)\*\***

## CCB-354 - Add new permissible values to Header/parsing standard id to support new versions of CDF.

See <https://pds-jira.jpl.nasa.gov/browse/CCB-354>

Someone doesn't understand where the comments led us.

~ Reporter apologizes for not responding earlier. There was a document posted. Release notes of different versions of CDF back to 2.0. Most relevant after 3.4. Changes don't affect data format in a way to to not be able to create a valid CDF/A. Other issue was description. There are value meanings that are kind of generic. Follow format used for CDF 3.8. Unclear if something more is needed.

~ Someone has never been enthused about CDF and not a user, so have to trust that this is true.

~ Someone else is not sure that we need to address this. Definitions seem reasonable.

~ Reporter is unclear if something else is needed. The main question is if any changes to CDF after version 3.4 make it impossible to make valid CDF/A and the answer is no. Can describe the bytes and the location of the bytes.

Mostly okay now, but one change was to accommodate INF and NANs. PDS doesn't accept those in the SR.

~ That's a change we should put in the CDF requirement.

~ That's the essence additional changes we need to add to the SR. Don't want to go through that door.

~ The back door is open. INF and NANs were allowed in 3.4.

~ Independent of this. An open issue for 3.4. Shouldn't affect this. What we want to hear is that this doesn't re-open a door we already closed. No new doors to close.

~ Reporters understanding is that it has not. Welcomes hearing others review and point out red flags. CDF in general allows things we can't allow in PDS4. Key is to restrict them. We've done that with CDF/A. Need to add restrictions to NANs and INF to that, but think we can still make PDS4 compliant CDF/A.  
~ If it's documented that we are not opening new gateways, okay with that.

Question: Does anyone know someone more qualified to review this?

Another Question: Who uses CDF besides PPI?

Answer: ATMOS is just starting.

~ Technical issues would be the same regardless of the node.

~ That's what we need to monitor.

~ Right. Could be reviewed by anyone with tech expertise.

~ Someone spent a lot of time on this when it first came up, but it went through. Worried about back doors that allow people to put non-compliant stuff in. Will have to pass on this one.

~ Someone else doesn't see anything in SR about CDF.

~ CDF contains arrays. CDF/A requirements describe how to make them PDS compliant, like having no internal compression.

So for this SCR we should make sure it's not getting worse. Other back doors can be a different resolution.

~ Someone needs to check the list.

~ Most changes are to tools and libraries and don't affect the format. Some metadata attributes.

Welcomes someone with more technical knowledge to look at it.

Question: What level of tech expertise?

Answer: PDS4 expertise.

Another Question: To Steve - did you look at it?

Answer from Steve - Yes, but like others, did a lot of work with CDF early on. Understood the issue of back doors, but it's going to be beyond him.

~ CDF/A standard.

~ PPI wrote the CDF standard for PDS and SPASE to make CDFs that will work in those systems.

Not sure how this moves forward beyond just taking a vote. Can ask Jesse to look. Can defer until next time. If we haven't heard anything bad we can vote.

~ Reporter would be happy to have people look.

Question: Is there a way to verify CDF/A is compliant?

Answer: Yes. Have tools.

~ That sounds like the way to know nothing is going through the back door.

~ Someone likes the idea of having Jesse look at the list. Very concerned because CDF is a black box.

~ No more than FITs.

~ Right. There are restrictions on FITs. Have looked at CDF. Spent time learning.

Question: Can reporter add to the Proposed Solution that the intent is not to open any unknown doors?

Another Question: In the proposal, not as a comment?

Answer: Yes.

~ Okay. If someone is going to look, (reporter) can mark up his understanding of each item in the document.

~ Mike and Jesse will look at the document. **(Action Item)** ~ Someone would like reporter to mark up the document.

~ Reporter is happy to do that **(Action Item - Joe)** ~ Yeah, add that to the Jira ticket to clarify everyone's understanding.

**Action Item - Joe** will mark up the document, and **Mike and Jesse** will review it.

~ Explicit that it all has to be PDS4 compliant.

## CCB-357 - Create Product External

See <https://pds-jira.jpl.nasa.gov/browse/CCB-357>

Name has changed and there have been some clarifications.

~ Some things were made optional. Items 1 and 2 are good. Text was added to item 3. One question is where this goes in the SR. Thinks 9I. If need to flesh out external document more, let reporter know what to say. Need to cite the external document properly in the SR.

Question: Whose action is the external document?

Answer: Lyle will start it. If there are no strong opinions, will put it in section 9I in the SR. Could also be 2A.6.

~ Someone agrees with chapter 9. Chapter 2 is recommendations. People think it's requirements, but not.

The **Action Item** is on **Lyle** to start work on the external document. Can send the SCR for a TA.

~ We can vote in 2 weeks.

Question: Is there a brief description somewhere of why we went from annex to external?

Answer: No.

~ More generic description of what we are doing. Products that are external to archive, but want to include them in the IM.

## CCB-325 - Support for video as product observational See <https://pds-jira.jpl.nasa.gov/browse/CCB-325>

There will be a presentation for this next time.

Question: Is this for the next build?

Answer: That's the intention.

Another Question: For product observational or it's own product where we can watch changing technology and which codex are available?

Answer: Have codex specified for product browse. Given the salt and pepper results it might be horrifying. Will be making a presentation next time, but video without compression is basically a no-show. There's no video in the archive, which makes it sort of lame.

~ Agrees, but we need to segregate it as a separate product so people don't think compression is allowed in the archive.

We want to make this SCR happen, but the SCR isn't complete yet.

~ Reporter is working on it. Was first concerned with if it would be accepted at all.

~ Someone agrees this should be a separate product, maybe product encoded.

~ Reporter is doing testing.

~ We have several paths to go down... what product and what format. For documents we require text or PDF/A, but you can then add other formats. Could maybe do something like that here.

~ These are all good ideas. They should be discussed off-line. Need to get things into the ticket so we can have a formalized ticket.

**\*\*Action Item - Trent\*\*** will work on the SCR ticket.

We need to get this through. We have two meetings left. Shouldn't be thinking of new ideas now.

~ We should count on a third meeting in February. We usually do that.



## DDWG Notes 2023-02-02

---

title: DDWG Notes 2023-02-02

layout: default

date: 2023-02-02

---

February 2, 2023

Notes by Debra Kazden

Known Attendees: M. Drum, M. Gordon, E. Guinness, T. Hare, L. Huber, S. Hughes, R. Joyner, D. Kazden, T. Lim, J. Mafi, S. McLaughlin, B. Semenov and R. Simpson  
Known Observers: A. Bailey, D. Kahan, P. Lawton, S. Loftin, M. McAuley, E. Schaefer and M. Tiscareno

### ## DDWG Agenda

(Included in email sent January 31, from M. Drum, Agenda for February 2, 2023)

#### Front-matter/What's new

- Votes! We're still trying to get votes!
- New enumerated value issues: CCB-360, 361, 362
- Approved SCRs by March 2, 2023

#### Ready for Vote

- CCB-354 – Add new permissible values to Header/parsing standard id to support new versions of CDF – Joe Mafi - Nothing came up in further review
- CCB-355 – Add Funding Acknowledgement class to Citation Information Class – Myche McAuley - See attached documents for new class definitions and example **\*\*(Voted to Pass 354, did not vote on 355)\*\***

#### Tier 1 issues

- CCB-357 – Create Product External – Lyle Huber - Which SR section? Need external document
  - CCB-325 - Support for video as product observational – Trent Hare - Presentation Today
  - CCB-358 – Context object for Instrument Package – Anne Raugh - Further discussion on Instrument Host option
  - CCB-360 – Add 'Electrons' value to Units of Misc – Anne Raugh
  - CCB-361 – Add a permissible value for microradians to Units of Angle – Anne Raugh
  - CCB-362 – Add a permissible value of nm/mm to Units of Misc – Anne Raugh
- \*\*(Discussed)\*\***

#### Tier 2 issues

- CCB-211 – XML/schema based files as archival data – Jesse Stone, Mike Drum
- CCB-336 – Add a License Information class to the Identification Area – Tanya Lim - We know that we need to at least support CCO
- DOI-rich metadata into IM (ticket to be created), replacing CCB-222 and CCB-261 – Anne Raugh and any volunteers
- CCB-164 - Display Settings not required for images -- (IMG) - Need to think more about potential updates to Imaging and how Display might fit.
- CCB-353 – Fully label all bytes in data files – Anne
- CCB-356 – Update Constraints on Author/Editor Lists – Dick Simpson

- CCB-350 – Adopt more rigorous / stringent rules for leap seconds
- Product Metadata Errata? Create a supplemental label to live next to original label, that can add aliases and supersede metadata **\*\*(Not Discussed)\*\***

Parking lot – need a driver!

- CCB-326 - Inconsistent Direction on How to Use Discipline Facets
  - CCB-324 – Create a data type for Type attributes that require a permissible value list.
  - RS: 259/281 (see new attachment to CCB-281) (on hold)
  - CCB-261 - Add an optional method for indicating that a product is a member of a larger collection that has a DOI -- Last chance before recommending withdraw.
- \*\*(Not Discussed)\*\***

Next meeting: Special session 2/9/23

Regular session 2/16/23

#DDWG

## Front Matter

We've got a few different votes from last time. CCBs 354 and 355. Hopefully people had time to look them over. 355 has had some clean up.

There are three new issues. They are in Tier 1 because they are enumerated value additions. Hopefully uncontroversial.

# Ready for a Vote

## CCB-354 - Add new permissible values to Header/parsing\_standard\_id to support new versions of CDF See <https://pds-jira.jpl.nasa.gov/browse/CCB-354>

The action from last time was to make sure there's another stamp of approval on the changes and not a number and +/- infinity.

~ There's a CDF/A document which discusses the requirements for making a CDF that is compliant with PDS4 requirements. Need to add a note that NAN and INF are not allowed. Separate document. Not part of the SR. Needs to be specified in external document.

~ That will go along with it when this gets published in a few months.

**\*\*The Vote for CCB-354**

ATMOS - Yes

EN - Yes

GEO - Yes

IMG - Yes

IPDA - Yes

NAIF - Yes

PPI - Yes

RMS - Yes

RS - Yes

SBN - Yes**\*\***

## CCB-355 - Add Funding Acknowledgement class to Citation Information Class See <https://pds-jira.jpl.nasa.gov/browse/CCB-355>

Requested changes were made in attached documents. A bit confusing because the requested changes reflect the original requested changes. Would like the description to reflect the latest version. There is tracking in jira.

~ Will take that out now.

~ Someone has lots of comments. Finds it very confusing. The SCR puts funding acknowledgement class under the citation class, than award information class that does nothing. Should just repeat funding acknowledgement class. Another thing is that it would be easier to follow if all attributes started with funding acknowledgement and were specific for this class. Also, award number is alpha-numeric string, should be funding award, rather than award number. Funding acknowledgement text for the definition is flat out wrong. Willing to work on this. It needs clean up before we vote in two weeks.

~ Reporter said fine. Thought we had it, but spending more time cleaning this up is fine. Award information was a way to have multiples, but if there's another way to do it, that's fine.

~ Also, under citation information, but no information on what needs to happen there to accommodate that class.

~ Yes. Someone understands what's being proposed.

Question: If funding acknowledgement only occurs once, okay, but if more than once what's the problem?

Answer: Think it would be okay. Just collapsing it. Want to make sure all bases covered.

~ Useless layer of class built in that we can get rid of.

~ Makes sense.

~ Another Question: An example has funding acknowledgement class with multiple instances of award information. Nice and clear. Are you saying award information class should be in the funding acknowledgement class? Makes it a different thing from citation information.

Answer: Citation information is how you cite this data. It's how to refer to the bundle or collection.

Having multiple funding acknowledgements in that. Should have it plural.

Another Question: Just remove award information class?

Answer: Yes.

Someone can go either way.

~ Okay with someone else too. Another issue is that we didn't add citation class as parent class for this. Need to make implementation exactly clean.

~ A cookbook with all the what-ifs, so that what is passed to the CCB is very clean and then the CCB can just say Do It.

**\*\*Action Item - Dick\*\*** will work with **\*\*Mike and Kate\*\***. Likes this idea, but the SCR is still very confusing.

~ A last suggestion is that we shouldn't try to put the year in. We've got alpha-numeric award string. Some awards are for multiple years. As long as we have the award identifier it should be enough.

~ Award year is when granted. Useful information.

~ Someone is not sure they agree.

~ Reporters will discuss this.

~ It might not apply to every grant. Could be optional.

~ Could loosen up acknowledgement text, require text from the funding agency. Should be optional.

~ Proposal funding year could be selection year.

Question: Is that important?

Answer: It could be important. Track that sometimes.

We will move on. Will check back on this. Had to schedule a tentative special session next week. Seems like we need it. Can take two weeks on this or try for next week.

~ A first cut can be done today.

~ We can try for next week.

~ We will need to know it's ready ahead of time.

~ We want to get as much done as soon as possible.

# Tier 1

## CCB-357 - Create Product External

See <https://pds-jira.jpl.nasa.gov/browse/CCB-357>

An external document was provided.

~ The document was drafted and got comments from people. Need a TA and to finish the fine details so we can have a vote.

~ We can schedule a vote for next time.

Question: Did we decide which SR section?

Answer: Yes. It's in Jira. Third item of description.

~ Great. We will request a TA.

Question: Is everyone okay to meet next week?

Answer: (Silence)

~ We will schedule a vote for next week.

## CCB-325 - Support for video as product observational See <https://pds-jira.jpl.nasa.gov/browse/CCB-325>

Video presentation from Trent today. (see the video in jira)

Still slow on moving this ticket forward. Unclear if it should be product observational or product video. Concerned about having too many product things. Video product is allowed as document already. Had to guess about frames per second.

Quality discussion slide - uncompressed video exists, but not real useful in the real world.

~ The compromise is to make sure it's available in product native and allow compression for video with well known methods and recommend best practices for compression quality. Lossless, but very constrained.

~ Went through to show errors based on single frame. Shows color options too. Did a bunch of tests - reporting on a handful.

~ CRF of 25 is horrifying. A 25 compression is not recommended or on the verge of not recommended.

~ CRF 20 also is not recommended.

~ CRF 16 is recommended for documents. Pixels changing. Can see the most differences. Trending

towards this for science products. Still changing. Going from one form to another. Made several examples. (See the video) ~ Also compared size versus RMSE for the tiny little video of about 360 frames. Actually, get lesser size for lossless compression. Shown in the slides. Size is reasonable related to original frames. Most video at 1080P level. Very small, but when increasing time or resolution jumps up. So, instead of recommending the document quality want to do it at compression quality of 10 to 12. Still scary.

Question: Are we good with this? Should we move forward?

Answer: No choice. We need to move forward. Thinking about wording already. Could say maximum CRF is 10 to 12 range, but providers should consider lower if practical. Expect better CRF in a few years. ~ As technology improves we will need to move forward with different compression, which will require changes to what's recommended. Reason we're going with such old compression is popularity and it will be around.

Good presentation. Regarding product observational versus product video, we have invested a lot of time in product observational. Might be better off using that because of the legacy stuff. Product video would have a lot of duplication. Putting it in product observation is probably a better way to go.

~ Someone agrees. It conflates the IM.

~ The case for the other choice is so people don't try to archive other things - like an Excel spreadsheet - in product observational. Worried about tagging these with product observational and having to migrate software every decade. Making product video is a big red flag to remind us we need to migrate software every ten years.

~ Can't currently do encoded byte stream. Leaning towards keeping it out of product observational.

Could use product video only if there's a video or audio component.

~ File area. A lot in product observational could be problematic. Could do file area compressed to give it its own special grouping, but it could still be product observational.

~ Worth considering.

~ Audio will be problematic too. Maybe we want product encoded, but opening Pandora's Box.

~ Product observational encoded. Yes, it's opening Pandora's Box. We need to keep it under control.

~ If we call it encoded DPs might try to use it for Excel spreadsheets. If we call it audio video it points out it's not generic.

The short answer is yes, keep moving forward with this.

~ Reporter should pick what they think solves the most problems and move forward. We can fix in future if necessary.

~ Had a lot of information about codec, compression, etc. Not sure why not just keep it in original format it was downloaded in. More concerned about needing to migrate software every ten years.

~ That's why picking a conservative option. To try not to have to translate it every ten years.

~ Okay. More concerned about the repeat migration than the compression.

~ Would run into the same issues as JPEG2000.

~ Someone had similar concerns about this and if original compression stays in product native.

~ It will be more highly compressed coming down. Trying to set quality.

~ Better to accept what's coming down if it's not worse than what we're using. Seems like we're creating an artificial threshold.

~ If already using a well known codec allow as-is. Makes sense.

~ That's what someone else was trying to say too. Might be creating unnecessary work.

~ Can recommend specific compressors ahead of time too. Long wish. Will update the ticket.

~ Also consider file area audio video.

Regarding capturing the original formats in product native - just want to remind you that getting the MC to approve the additions is a gruelling process.

~ Had forgotten that.

We will check on this again next week.

---

Three new SCRs have been added:

\*\*CCB-360 - Add new permissible value to Units of Misc\*\* See <https://pds-jira.jpl.nasa.gov/browse/CCB-360>

\*\*CCB-361 - Add a permissible value for microradians to Units of Angle\*\* See <https://pds-jira.jpl.nasa.gov/browse/CCB-361>

\*\*CCB-362 - Add a permissible value of nm/mm to Units of Misc\*\* See <https://pds-jira.jpl.nasa.gov/browse/CCB-362>

If anyone has expertise in this area please make comments in Jira so we can move forward.

Question: Has anyone looked at these yet?

Answer: Someone looked. They are fine. Enumerated values.

~ For CCB-360, yes, we should have electrons.

~ Someone is a bit puzzled because we usually allow pre-fixes. Will have to read the SCRs.

## CCB-358 - Need a Context object for Instrument Packages See <https://pds-jira.jpl.nasa.gov/browse/CCB-358>

Question: What is the status of this one? The conversation on Jira seems unresolved.

Answer: Yes. Confusing. Made a counter proposal. Need to hear from Anne on the alternative. Ideas are in the comments.

\*\*Action Item - Pat\*\* will try to discuss this with Anne.

Question: Is this possible for the next build?

Answer: All urgent, but unable to make a promise at this time.

We will meet again next week on February 9.

DDWG Notes 2023-02-09

---

title: DDWG Notes 2023-02-09

layout: default

date: 2023-02-09

---

February 9, 2023

Notes by Debra Kazden

Known Attendees: M. Drum, M. Gordon, E. Guinness, T. Hare, L. Huber, S. Hughes, R. Joyner, D. Kazden, J. Mafi, B. Semenov and R. Simpson  
Known Observers: A. Bailey, M. Bentley (For T. Lim), M. K. Crombie, P. Lawton, M. McAuley, E. Schaefer and M. Tiscareno

## DDWG Agenda

(Included in email sent February 7, from M. Drum, Agenda for February 9, 2023)

Front-matter/What's new

- Votes! We're still trying to get votes!
  - Approved SCRs by March 2, 2023
- \*\*()\*\*

Ready for Vote

- CCB-355 – Add Funding Acknowledgement class to Citation Information Class – Myche McAuley - Requested Changes being clarified
- CCB-357 – Create Product External – Lyle Huber - TA will be completed before meeting \*\*(Votes to pass both, 357 with liens)\*\*

Tier 1 issues

- CCB-325 - Support for video as product observational – Trent Hare
  - CCB-360 – Add 'Electrons' value to Units of Misc – Anne Raugh
  - CCB-361 – Add a permissible value for microradians to Units of Angle – Anne Raugh
  - CCB-362 – Add a permissible value of nm/mm to Units of Misc – Anne Raugh
  - CCB-358 – Context object for Instrument Package – Anne Raugh - Awaiting feedback from author
- \*\*(Discussed 325)\*\*

Tier 2 issues

- CCB-211 – XML/schema based files as archival data – Jesse Stone, Mike Drum
- CCB-336 – Add a License Information class to the Identification Area – Tanya Lim - We know that we need to at least support CC0
- DOI-rich metadata into IM (ticket to be created), replacing CCB-222 and CCB-261 – Anne Raugh and any volunteers
- CCB-164 - Display Settings not required for images -- (IMG) - Need to think more about potential updates to Imaging and how Display might fit.
- CCB-353 – Fully label all bytes in data files – Anne
- CCB-356 – Update Constraints on Author/Editor Lists – Dick Simpson
- CCB-350 – Adopt more rigorous / stringent rules for leap seconds
- Product Metadata Errata?
- Create a supplemental label to live next to original label, that can add aliases and supersede metadata

**\*\*(No Discussion)\*\***

Parking lot – need a driver!

- CCB-326 - Inconsistent Direction on How to Use Discipline Facets
- CCB-324 – Create a data type for Type attributes that require a permissible value list.
- RS: 259/281 (see new attachment to CCB-281) (on hold)
- CCB-261 - Add an optional method for indicating that a product is a member of a larger collection that has a DOI -- Last chance before recommending withdraw.

**\*\*(No Discussion)\*\***

Next meeting: 2/16/23

#DDWG

## Front Matter

There's been a lot of activity in the last few days on product external. We will need to have a discussion before we are ready to vote.

Believe we are in a good place with 355.

# Ready for Vote

## CCB-355 - Add Funding Acknowledgement class to Citation Information Class See <https://pds-jira.jpl.nasa.gov/browse/CCB-355>

Some changes were made to make it clear what we are voting on. Hopefully, no one has any issues.

~ Someone made some changes about five minutes ago.

~ Someone else believes the issues were things decided to be left in the hands of the proposers.

~ We discussed taking out the extra layer of class. Don't remember discussing in identification area or citation area.

~ Originally it was under citation information, to be parallel with peer review publications. Okay with identification area if that's better.

~ Someone thinks citation area is how you cite this when you use it. Funding is separate from how others would cite it, but not a sword to die on. Sees this as serving a different purpose.

~ We can revisit this in the future when we look at citation area broadly. There's already a ticket to re-factor into more providence metadata.

~ Someone agrees.

~ Someone else says we should vote.

**\*\*The Vote for CCB-355**

ATMOS - Abstain

EN - Yes

GEO - Yes

IMG - Yes

IPDA - Yes

NAIF - Yes

PPI - Yes



RMS - Yes  
RS - Yes  
SBN - Yes\*\*

## CCB-357 - Create Product External  
See <https://pds-jira.jpl.nasa.gov/browse/CCB-357>

Want to discuss where we are at and if anyone has issues with moving forward with the vote.

~ When this was first put in it had some comments on LID construction and people said external to archive so we couldn't include that. Now back to that.

~ Someone thinks we should exclude LIDs that are external to the archive.

~ Reporter was asked to consult Jordan regarding resolving institutions, but there isn't such a beast.

Hopefully things will go to registries, but that's the best we can do. Agree we should forbid part, but it's not a show-stopper.

~ URNs used to be very popular. The prefixes make them unique. Nasa:pds:and... should be pretty safe.

Question: Do these forbidding statements go in the SR or the external document?

Answer: Maybe needs to be in both. SR might not be a reference for users of external.

~ Validate recognizes JAXA, ESA, etc., so confused.

~ Yes, so we exclude what's in the namespace registry.

~ Could be urn:ext and then go on, as long as it's specifically flagged it's safer. Think we can restrict by saying it's outside PDS but in the ecosystem, so controlling the beginning of the URN would be the way to go.

~ There are a lot of schematron rules to consider. Could be set up to require something specific for external.

~ Someone is troubled by that. Not sure the system of LIDs is used outside the PDS IM universe.

~ URNs are, but not as popular as they were 20 years ago. LIDVID comes from the EB XML standard.

Question: Do our LIDs need to be unique with stuff outside PDS? can that cause a problem for us?

Answer: It could.

~ So, if creating data sets on Zenodo or whatever might not want the LID to be branded non-PDS.

~ Someone agrees. Thinks starting with URN ext is too high. Could start with agency. Urn:nasa should be at the beginning of annex products. Maybe urn:nasa:pds-ama for ATMOS model annex, but a level too high for external.

Another Question: Black list versus white list, could it be hard wired to see a list that is legal for references?

Answer: We don't control that at the SR level.

~ We need to think about that.

~ Someone likes urn:esa and then what we need to distinguish the repository.

~ So we are down to having a black list in the SR for things to be used by non-archival products.

~ Reporter will add text to the external document and the SR.

~ Format should be the same number of colons.

~ Someone suggests we refer to the namespace registry to avoid duplications.

\*\*Action Item - Steve\*\* will send text on finding that to \*\*Lyle\*\*.

~ So, that's a lien.

There's another lien - text was just sent for a definition. We can vote with liens or wait.

Question: Someone has one more comment to discuss. If someone has an archive bundle, can't have an external product as a primary product in that, right? What about secondary? Think no, it has to be archival, but don't think it says that.

Answer: Another person agrees. Secondary has to be a primary member of some other primary bundle.

Another Question: What about external collections in a bundle?

Answer: The intent is that these are outside - not to be mixed.

~ The associations between them are made by membership.

Another Question: Do we need a level of validation built in to make sure LIDs start with the recommended prefixes?

Answer: Think it already does that.

~ We need to check with Jordan.

~ There are schematron rules to check LIDs.

~ Not sure if they check in the inventory.

Need to make sure we have addressed all the questions.

~ Someone has a question about who would manage this and the required changes.

~ Hopefully there are not a lot of required changes here. Could be handled by an LDD.

To poke back on the whole thing about having EN support DOIs for product external. Seems reasonable. Funded by NASA. EN should be able to provide DOIs for product external. There is precedence because they are getting DOIs for PSI. Want people comfortable with DOIs from PDS EN.

~ That's an MC level discussion. It's a funding issue, not necessarily something we can decide.

~ True. Outside SCR and DDWG, but external to PDS, and part of the ecosystem and not sure who else would give us a DOI.

~ DOIs are the responsibility of the repository holding the data, so with annex and PDS hats, if we want to ask EN to get a DOI while wearing our annex hat, fine, but outside archive, so should be on a case by case basis.

Question: Does ATMOS have separate money for their annex?

Answer: Over-guide.

~ So, it is PDS money.

~ Weird fit for under PDS, but it is funded that way.

~ USGS is different.

Question: What level of support do we expect for these? Validate structure? How are we going to do that? Still some holes in the description. Also, after reading the external document, worried it's wide open for abuse by people who don't want to do the archive work.

Answer: Send comments for the external document.

~ A lot of work.

~ Concerned about it saying the format is not PDS4 compliant.

~ That's why we are doing this.

~ So if someone is required to archive, but they don't want to use a compliant format they can shove it in here. Concerning that there are people who won't want to do the work.

~ There's a line that the use of this product is at the discretion of the curating node or organization.

~ It will work. If someone tries to give NAIF something external, they won't archive it.

Another item says it can be superseded, models, simulations, etc. Could get a lot of things, like

compressed data.

~ You are ignoring the paragraph that lists products not suitable for inclusion.

~ The point is there will be some DPs who will try to do that. Have to make sure there is enough in the requirement to say no, you can't do this.

~ Reporter needs wording for what needs to be said.

~ Regarding support for this, this really does come down to funding agencies discretion. If a person is funded by NASA and are supposed to archive, and they turn in product external, it's not our job to say it should be in PDS. Think we need to kind of let it be.

~ As far as validation goes, would like it to handle these products. Not a huge imposition. A few new classes.

Question: What about users being able to read the files?

Answer: That's encoding standard ID, and it's in there. External, so we say what the encoding standard is.

~ This is so PDS search returns products from the annex. We can say we just provide a service to find it and contact the annex if having trouble reading it. That should be the default. We are just providing the ability to find the products.

~ This group has enough expertise to say if a format is crazy. Hopefully we won't allow crazytown formats. We have in the past, like ISIS.

Moving on.

We have a couple liens, but prefer to vote, unless anyone is strongly opposed, so we can move on to video.

~ It depends on what's being sent to tighten up the wording in the external document.

~ This is a big black box. Have other things to do.

~ We could vote with liens if people are comfortable.

~ Someone is comfortable. If going to implement this we need to know what to implement. Sounds like most issues are in the documents.

~ Someone else agrees.

Question: Does anyone think their vote would change over wording in the documents?

Answer: (Silence)

**\*\*The Vote for CCB-357 with 2 liens**

ATMOS- Yes

EN - No

IMG - Yes

IPDA - Yes

NAIF - Abstain

PPI - Yes

RMS - Yes

RS - No

SBN - Yes\*\*

That's a majority.

Question: The liens have to be resolved before this goes to the CCB, right?

Answer: They would flag it and send it back to us.

~ Reporter can make changes to the document and the SR and then send it to people to review to make sure the liens are resolved.

The liens are the definition of file URL and black list of URNs they can't use.

~ Someone will try to send some words.

~ Black list words will come from Steve so point to the namespace registry. Then Mike has to be satisfied that the liens are resolved. Can comment in Jira and send it to the CCB.

# Tier 1

## CCB-325 - Support for video as product observational See <https://pds-jira.jpl.nasa.gov/browse/CCB-325>

Question: Any updates?

Answer: Reporter is working in the background. Moving forward with product observational.

Question: That is what we all decided, right? That's the last bit to clarify before tightening it up. Also want to get the white paper with recommendations out.

Answer: Proposed Solution needs a solution.

~ Going forward with product observational.

Another Question: File area audio video?

Answer: Yes.

~ Someone likes that.

~ Reporter will move forward with that.

Someone hopes we will have one more discussion on this with something specific to throw rocks at and then vote.

~ We could evote.

Question: Other comments?

Answer: No.

---

The three enumerated value SCRs will be queued up for a vote next week. If there are any issues people should make sure to comment.

Question: Any chance of getting to CCB-350 next week? Not much is new. EN doesn't seem interested in doing much work on it.

Answer: It will be added for next week.

## DDWG Notes 2023-02-16

---

title: DDWG Notes 2023-02-16

layout: default

date: 2023-02-16

---

February 16, 2023

Notes by Debra Kazden

Known Attendees: M. Drum, E. Guinness, T. Hare, L. Huber, S. Hughes, R. Joyner, D. Kazden, J. Mafi, B. Semenov and R. Simpson  
Known Observers: A. Bailey, D. Kahan, P. Lawton, S. Loftin, M. McAuley and M. Tiscareno

### ## DDWG Agenda

(Included in email sent February 14, from M. Drum, Agenda for February 16, 2023)

#### Front-matter/What's new

- Liens resolved on CCB-357 Product External, sent to CCB
- Approved SCRs by March 2, 2023
- \*\* (Discussed) \*\*

#### Ready for Vote

- CCB-360 – Add 'Electrons' value to Units of Misc – Anne Raugh
- CCB-361 – Add a permissible value for microradians to Units of Angle – Anne Raugh
- CCB-362 – Add a permissible value of nm/mm to Units of Misc – Anne Raugh \*\* (Voted to pass all 3) \*\*

#### Tier 1 issues

- CCB-325 - Support for video as product observational – Trent Hare
- CCB-350 – Adopt more rigorous / stringent rules for leap seconds - Ready for final discussion before vote
- CCB-358 – Context object for Instrument Package – Anne Raugh - Awaiting feedback from author
- \*\* (Discussed) \*\*

#### Tier 2 issues

- CCB-211 – XML/schema based files as archival data – Jesse Stone, Mike Drum
- CCB-336 – Add a License Information class to the Identification Area – Tanya Lim - We know that we need to at least support CC0
- DOI-rich metadata into IM (ticket to be created), replacing CCB-222 and CCB-261 – Anne Raugh and any volunteers
- CCB-164 - Display Settings not required for images -- (IMG) - Need to think more about potential updates to Imaging and how Display might fit.
- CCB-353 – Fully label all bytes in data files – Anne
- CCB-356 – Update Constraints on Author/Editor Lists – Dick Simpson
- Product Metadata Errata?
- Create a supplemental label to live next to original label, that can add aliases and supersede metadata
- \*\* (Not Discussed) \*\*

Parking lot – need a driver!

- CCB-326 - Inconsistent Direction on How to Use Discipline Facets
  - CCB-324 – Create a data type for Type attributes that require a permissible value list.
  - RS: 259/281 (see new attachment to CCB-281) (on hold)
  - CCB-261 - Add an optional method for indicating that a product is a member of a larger collection that has a DOI -- Last chance before recommending withdraw.
- \*\*(Not Discussed)\*\*

Next meeting: 2/23/23

# DDWG

## Front Matter

We have three votes on enumerated values today, but they are not set up as enumerated type issues in jira. Hopefully the TA filled in the gaps. Unclear if Anne knows about the enumerated value tab that constrains what needs to go in there.

~ Pat will let her know about it. Pat also didn't know.

~ Units instead of regular attributes.

Waiting on CCB approval for CCB-357 (Create Product External, see <https://pds-jira.jpl.nasa.gov/browse/CCB-357>).

Question: Can we change the URN example?

Answer: Already done.

# Ready for Vote

## CCB-360 – Add new permissible value to Units of Misc See <https://pds-jira.jpl.nasa.gov/browse/CCB-360>

Comments were added. Says electrons and electron. Think we want to say electrons.

~ Person who did the TA believes the request was missing the Requested Changes. They wrote it. Said to blame them for electron/electrons.

~ No blame. Just want to be clear.

It should be electrons.

~ Yes, in other cases we use plural. Frames per second, pixels per... Could take a strawman on electrons versus electron.

~ Should be electron because already an electron per diam.

~ Yes.

Question: A definition or a unit?

Answer: An existing unit.

Another Question: Is it a non-backwards compatible change to fix it?

Answer: No, just a mixed bag. Needs to hear the rationale for electrons over electron.

~ In other places we use plural.

\*\*Strawman Pre-Vote - In favor of electronS in Requested Changes ATMOS - Yes EN - No GEO - Yes IMG -

No IPDA - Not Here NAIF - Abstain PPI - Yes RMS - Yes RS - Yes SBN - Abstain\*\*

Five yes votes, two no. Will change to electrons.

Question: Is it creating a huge amount of work?

Answer: No. No change to the technical side of the IM.

~ Someone was trying to go with precedence. Okay either way.

~ Original request was electrons.

~ Not sure Anne really cares.

~ Almost all units are singular. The exception is plural.

~ We're not consistent.

~ And this is more of a count, so okay with plural.

~ Group facet one says electrons.

\*\*The Vote for CCB-360 - electrons

ATMOS - Yes

EN - Yes

GEO - Yes

IMG - Yes

IPDA - Not Here

NAIF - Yes

PPI - Yes

RMS - Yes

RS - Abstain

SBN - Yes\*\*

## CCB-361 – Add a permissible value for microradians to Units of Angle See <https://pds-jira.jpl.nasa.gov/browse/CCB-361>

This has the same issue. Would rather this be singular. A real unit, not a count.

~ It is. Just the title is plural.

\*\*The Vote for CCB-361

ATMOS - Yes

EN - Yes

GEO - Yes

IMG - Yes

IPDA - Not Here

NAIF - Yes

PPI - Yes

RMS - Yes

RS - Yes

SBN - Yes\*\*

## CCB-362 – Add a permissible value of nm/mm to Units of Misc  
See <https://pds-jira.jpl.nasa.gov/browse/CCB-362>

The specific unit identifier for units of miscellaneous has values that can't easily be converted, so

basically a dumping ground. Could get rid of specific unit ID and remove the conflict, but then couldn't specify minimum and maximum. Not sure I understand how this is supposed to work. Think the ship has sailed. Could fix it with more subclasses. Unclear how onerous that would be.

~ Someone agrees. Think we could tackle separately. If someone wants to propose that.

~ Someone else agrees, but on the path to becoming not very useful. No time to put in on it. Will vote abstain.

~ Yes. Getting bigger.

**\*\*The Vote for CCB-362**

ATMOS - Yes

EN - Yes

GEO - Yes

IMG - Yes

IPDA - Not Here

NAIF - Abstain

PPI - Yes

RMS - Yes

RS - Abstain

SBN - Yes\*\*

Question: Back to CCB-360 - who has the responsibility to change the Requested Changes?

Answer: Someone just did it.

~ Good.

# Tier 1 issues

## CCB-350 - Adopt more rigorous/stringent rules for date/time strings See <https://pds-jira.jpl.nasa.gov/browse/CCB-350>

A new set of patterns that catch leap seconds was presented, except it can't catch leap years. Think this is still superior to what is in the system. The question is if there is a longer set of patterns if it will slow down validation. EN is most concerned about accuracy.

~ Testing has been done. Needs to be prototyped. A good step forward. Might slow things down a wee bit. Can discuss and vote soon.

Question: Any comments?

Answer: For implementation, glad testing is done. When make changes it will go into issue testing. Also, should be cut/paste for the actual changes.

~ Reporter will make sure the exact patterns are there.

~ That's the most complex changes for the entire model.

**\*\*Action Item - Dick\*\*** will make sure the exact patterns are in the Requested Changes.

Question: Performance concerns?

Another Question: The new formation rules would work with how we do date times now?

Answer: No change in how you put date time strings together. Better validation. Illegal values will now be caught rather than passes through.



We will vote next meeting. Probably next week.

## CCB-325 - Support for video as product observational See <https://pds-jira.jpl.nasa.gov/browse/CCB-325>

Doesn't look like the final Requested Changes and examples are there yet.

~ Still Alpha. Proposal would create new class encoded video audio... add encoded video audio under file area. It will change product observational to have encoded audio and video. Want to make sure we are not shooting this down.

~ Sounds like you want it shot down.

~ There are other ways to go. Listed in jira, but want this to be a primary science product.

~ Ancillary is not for primary science. More ready to adapt product observational.

~ Same if add at same level as product observational, but would be it's own new thing. Guess coming at this questioning what are the consequences of not putting this under file observational.

~ The purpose of ancillary is to support. It's a second class citizen to support the primary. Ancillary is not the place for it.

~ Someone agrees. Would cause problems for search.

~ That is a concern.

Question: Was the proposal done correctly? Number 3 in the list of proposed solutions goes down the path under encoded byte stream, but you don't get everything. If use encoded byte stream and image you don't get everything, but mix and match levels. Thinks the solution and model are correct. Need to test.

Answer: Subclass of encoded byte stream can be described in implementation.

~ If we allow it it allows a lot of things in there.

~ We already allow encoded header.

~ Yeah, but that's more of a throw away.

~ So this is the first encoded data object under product observational. Makes sense.

~ Makes sense. Can explicitly dis-allow things that aren't video or audio.

~ Not what we said last week. Product observational in file area audio video.

~ Reporter tried to go down that road. Number 2 in comments. Wasn't sure of the point of it. Same level. Went down the road to try to go full in and see what you all think.

~ If we do the same level, add file area encoded, should prefix with observational, but then should put under file area observational.

One point in favor of grouping in a dedicated class is that it is good for future archive when we need to migrate to a newer standard.

~ Good point. Could set a flag to check, but other things will need to be checked too.

~ Just trading semantics. The issue is do we want these separated out to flag that they need maintenance.

~ Someone doesn't understand. Thought we wouldn't have to migrate the archive in the future.

~ This is the safety boat. To revisit things that might need updating.

~ Not just worried about this as a flag, worried about the message to DPs.

Reporter also looked at file area supplemental.

~ Different use.

Question: Should we do a strawman poll between options 1 and 2? Fully under product observational

versus, option 2, file area audio video something that we all like, encoded audio video. Would that be helpful to move this forward?

Answer: Someone thinks so.

~ For option 2, strongly recommends observational in the class name.

~ Someone agrees. Can have supporting documents. Really focusing on products that are observational, so need observational in the class.

~ We will have strict control of what's allowed in.

~ Yes. Two small file formats.

Question: All three options. We'll go around. Say preference between the three options.

Another Question: Any reason audio/video before observational?

Answer: Product observational fine.

~ Same pattern as supplemental.

~ Right. Comments will be edited (in real time).

~ Someone is beginning to be sold on it.

Question: for Steve: Impact to the IM? Will the new class mess things up?

Answer: Good question, but don't think we are searching in file area for much. But there will be some impact.

~ Editing in real time.

Let's go around the room. 1, 2 or 3.

~ This is very helpful.

~ People can rank their choices.

\*\*Strawman rankings

ATMOS - 1, 3, 2

EN - 1

GEO - 2

IMG - Abstain

IPDA - Not Here

NAIF - Abstain

PPI - 1, 2, X

RMS - 2, 3 no X

RS - Abstain

SBN - 2, 1, X\*\* (see jira - this changed over the course of the conversation)

A very strong maybe.

Question: Thought we were leaning towards 3, so what is going on?

Answer: Thought it was 1 and 2. Absolutely against 3. If observational then observational. We can put in all the necessary hooks. Understands the issue of allowing encoded.

~ Someone agrees. Not sure what the objections are to 1.

~ Pandora's Box.

~ No.

~ Not enough of a signal to archive maintainers or users that it will need changes in the future.

~ If we want to make the point that this is different, eventually we'll have lots of flags. Should just put in object orientated model.

~ Regarding going back every ten years, think we absolutely have to look at everything. We have documents in Postscript that we have to migrate. Just a fact of life of trying to maintain an archive in the real world.

~ If people were always rational, fine, but we don't live in that world. Three is the safest option. Can make it less clunky.

GEO - Now leaning 1, 2, but if 1, would also use 2 underneath it.

~ We could.

~ Geo - Likes 2.

~ ATMOS - Amend to 1 XX.

~ You're killing me.

Not sure where we go from here. Need as close to a consensus next week as possible if any chance of getting this in.

~ Yes. Looks like it's going as written now.

~ We need to take the strawman and see if there's any feedback from Jordan. **\*\*(Action Item - Steve)\*\***

~ **\*\*Trent\*\*** will join the conversation with Jordan too. Feeling the pressure. Artemis will have a lot of video.

This won't be easy to finish in a week. A lot of documents.

~ SR needs work in section 4.

~ There's time for that.

~ Yes, but it needs to be in the SR.

~ It's unclear who does the document updates now.

~ The date for the document updates is further out.

~ A month. If we can find out what needs to be done.

~ Unclear how we get a policy vote. Plan is to go forward as it reads now. Will vote on it sort of as is.

~ Policy vote path is that we vote, send to the CCB, then it's ready for the next build and at the March MC it can be presented to the managers.

Reporter will try to get more samples. **\*\*(Action Item - Trent\*\*)**

So, assuming it's ready, we will vote next week.

~ Reporter thanks everyone for the help. Hard to wrap head around all the issues.

## DDWG Notes 2023-02-23

---

title: DDWG Notes 2023-02-23

layout: default

date: 2023-02-23

---

February 23, 2023

Notes by Debra Kazden

Known Attendees: M. Drum, T. Hare, L. Huber, S. Hughes, R. Joyner, D. Kazden, T. Lim, J. Mafi, S. McLaughlin, B. Semenov and R. Simpson  
Known Observers: D. Kahan, P. Lawton, S. Loftin, M. McAuley, V. Tang, M. Tiscareno and E. Schaefer (For E. Guinness)

### ## DDWG Agenda

(Included in email sent February 22, from M. Drum, Agenda for February 23, 2023)

#### Front-matter/What's new

- Liens resolved on CCB-357 Product External, sent to CCB
  - Approved SCRs by March 2, 2023
- \*\* (Discussed) \*\*

#### Ready for Vote

- CCB-325 - Support for video as product observational – Trent Hare - Straw Poll
- CCB-350 – Adopt more rigorous / stringent rules for leap seconds \*\* (Voted to Pass 350) \*\*

#### Tier 1 issues

- CCB-358 – Context object for Instrument Package – Anne Raugh - Discuss original proposal if there's time
- \*\* (Discussed) \*\*

#### Tier 2 issues

- CCB-211 – XML/schema based files as archival data – Jesse Stone, Mike Drum
- CCB-336 – Add a License Information class to the Identification Area – Tanya Lim - We know that we need to at least support CCO
- DOI-rich metadata into IM (ticket to be created), replacing CCB-222 and CCB-261 – Anne Raugh and any volunteers
- CCB-164 - Display Settings not required for images -- (IMG) - Need to think more about potential updates to Imaging and how Display might fit.
- CCB-353 – Fully label all bytes in data files – Anne
- CCB-356 – Update Constraints on Author/Editor Lists – Dick Simpson
- Product Metadata Errata?

Create a supplemental label to live next to original label, that can add aliases and supersede metadata

\*\* (Not Discussed) \*\*

#### Parking lot – need a driver!

- CCB-326 - Inconsistent Direction on How to Use Discipline Facets
- CCB-324 – Create a data type for Type attributes that require a permissible value list.
- RS: 259/281 (see new attachment to CCB-281) (on hold)

- CCB-261 - Add an optional method for indicating that a product is a member of a larger collection that has a DOI -- Last chance before recommending withdraw.

\*\*(Not Discussed)\*\*

Next meeting: 3/2/23

# DDWG

## Front Matter

We need to discuss things before we get to the votes today.

The CCB passed all five issues. The votes were almost all unanimous. There was one no vote for product external and one abstain on an enumerated value one.

# Ready for Vote

## CCB-325 - Support for video and audio as product observational See <https://pds-jira.jpl.nasa.gov/browse/CCB-325>

There have been a number of changes in the Proposed Solution and attached documents.

~ There was a lot of good feedback from everyone. Going to have the vote include audio. The goal is audio and video under product observational, not encoded byte stream. The idea is that audio and video go hand in hand. Audio has lossless, so pulls it into another category, but comments said to just do it. Tried to write up a next-gen policy statement, but was asked to update the old one. Will have the MC vote on it if this passes. Will still have the white paper. Took a first stab at SR, Concepts and other documents where things need to be updated. Hard part is stating that video is allowed and having a recommendation. Including a best practice has been confusing.

Question: Any outstanding issues or liens?

Answer: Someone went back through the supporting documents and would rather not vote today. Need to be careful about restrictions. Would rather wait another week or have a special meeting. Would rather go slower and feel better about accuracy. This has to go to the MC anyway.

~ This can go through the MC later. They can stop the implementation. The deadline is next week. We want it voted on so the CCB can take a pass, but agrees we aren't ready to evaluate all the liens today.

~ Right.

Question: Are we trying to get our ducks in a row regarding the documentation?

Answer: Yes, and to make sure we are not crossing wires on audio and video since we are going to have both. Need to clean up the SCR so it's clear what is intended.

Question: Is there any confusion in the SCR?

Answer: Based on the discussion last week and during the week, the implementation is straight forward, but there are policy issues. Okay with implementation based on the agreements so far.

Another Question: Says we are going to use H.264, but have we agreed on how we are going to decide on other standards? Are we going to re-encode things to H.264?

Answer: That's the intention.

~ Someone thinks that's okay. It talks a bit about stuff coming down from spacecraft - can use product

native, and then have the H.264 product too. Can get the weird spacecraft formats into H.264.  
~ If the spacecraft is using a weird format we can tell them to use H.264 and maybe they will. Did change attribute values so it's easy to update if we move to H.265 or something. Lots of changes in the last two weeks.

We can evote early next week if we have a list of liens and changes we are all okay with.

- ~ There's a lot to read, but willing to help get this ready for Thursday.
- ~ **\*\*Action Item - Mike\*\*** will let the CCB know this is coming in tight.
- ~ We have a meeting next Thursday.
- ~ The question is if that meets the deadline.
- ~ It's hard to spend quality time on something that can be voted down.

Question: It sounds like we all agree and just want to get it all in order. Can we vote and send it to the CCB with a note that we'll get the documentation done?

Answer: There are issues. Would like to see encoding standard definitions cleaned up. Can't leave it to someone else.

- ~ So, that's a second lien.
- ~ Someone agrees, we need two encoding standard options.
- ~ Yes.
- ~ The standards need to be specific. And there are other issues to clean up too.

Given how much documentation there is, giving the time to read it and come up with liens, we can call for an evote or vote next Thursday and beg the managers to let this through.

- ~ We can warn the CCB so they can look at this and maybe offer suggestions before we vote.
- ~ A soft submit to them.

Question: Other discussion on this today?

Answer: If questions of substance, like definition for encoding standard ID, people can email for straw votes as issues are found.

- ~ Sounds like a plan.

## CCB-350 - Adopt more rigorous/stringent rules for date/time strings.

See <https://pds-jira.jpl.nasa.gov/browse/CCB-350>

Not sure if this got a TA, but it has been around awhile. Mostly an internal thing.

- ~ Someone considered it as a bug fix, but some standard issues. Can wait to vote next week and get a TA.
- ~ Its fine to wait until next week, but we are pushing two SCRs to the deadline. Will need to talk to Jordan about the release deadline.
- ~ Mike can start that thread.

This isn't critical for the next build.

- ~ We will discuss it next time.

Question: Is the only issue that there is no TA?

Answer: Nebulous.

- ~ Someone thought it was a bug fix that has had a lot of testing. We could just put a lien in to paste the patterns in.

- ~ The patterns have changed a lot, but it is swapping one for another.
- ~ We can vote with a lien and do the TA.
- ~ Someone says they are ready to vote.

\*\*The Vote for CCB-350

ATMOS - Yes  
EN - Yes  
GEO - Yes  
IMG - Yes  
IPDA - Yes  
NAIF - Yes  
PPI - Yes  
RMS - Yes  
RS - Yes  
SBN - Yes\*\*

A final comment and question: Will need the test to run. Is everything attached?

Answer: Think it's all in the ticket, but let reporters know if test cases are needed.

# Tier 1

## CCB-358 - Need a Context object for Instrument Packages See <https://pds-jira.jpl.nasa.gov/browse/CCB-358>

People have been discussing this and think it should move forward as originally proposed. The issues have been pointed out. (Note taker missed something here) The observing system component references a context object by LID. In a different class. That has several different types, none of which are just host. Calling it a spacecraft is probably inaccurate, so we should re-visit this. If that makes sense.

~ Sort of. Will wait to see it in writing.

Question: To the last question in Jira - did you say no?

Answer: Right. Said to call it instrument host instead of instrument package. This is why I don't think that's possible.

~ Someone is unclear on why we can't do what APSS did with aliases for the multiple identifiers in there. The LID is for a single item. Can use different things to search for it and find it.

~ Someone else is putting up the context object to see why the alias list was special somehow.

~ Can see APSS Twins has a couple identifiers in there, so search gets it.

~ Discoverable in search. The question is if we want to add the structures for a package of instruments.

~ Don't think this precludes that.

Question: What if just APSS, not APSS Twins?

Answer: If that's desired.

~ More of a design decision. A number of SBN folks would like it to happen, but if there are disagreements about implementation we can... (Interrupted)

Question: Say again, what's the problem with instrument host?

Answer: An instrument host can be a member of instrument host in theory, not type. Actual

implementation is a type of host in label. Looking at a label, observing system component describes host. Reference to instrument host. So go to that context product to learn more about it. Arrive at the document that says it's an imager. We need to be able to have a value. Current values don't have instrument host or package. More specific.

~ Less specific. Deprecated. Maybe reversing that would fix it.

~ That was the decision.

~ Seems inconsistent.

~ Proposer would be happy with undeprecating it and expanding the list.

~ That's CCB-293 (Specific values used to support search have been deprecated from the type field of Observing System Component, see <https://pds-jira.jpl.nasa.gov/browse/CCB-293>).

~ To change it back. And it's also by Anne. Maybe we need to look into the history of why we did this.

Question: Does anyone have strong issues with the original proposal? Not sure if re-evaluating everything will help.

Answer: Would want to add instrument package, with member. Kicking the can down the road.

~ Many multiple layers.

~ In original label could reference package as a host from spacecraft object and from data label and also spacecraft, multiple hosts. Context product would say it's an instrument package, but not necessarily say that in the data label. Maybe we need to clarify this in jira and then discuss it.

Another Question: A number of options. Is she ultimately wanting a new context product? Would prefer a generic group for components. Could call it observing system components group. If all we want is to include the name, we can. Just need to reverse the deprecation.

Another Question: What about multiple spacecrafts in one - orbiter and lander? Is there a model for that already?

~ (Lots of people spoke at once)

~ This goes back to the beginning when we created observing system component that could be used many times. Grouping.

~ For things like orbiter and lander - grouped at the investigation, I think.

~ That works, but to try to clarify this - it makes sense to create a context product for named things. Should be explicit for search. This has been discussed since early on.

Question: If a generic grouping, would it be a type of instrument host?

Answer: A group of observing system components.

~ Failing to see how implementation of those classes would look. They have references and layers from multiple labels and different descriptions. Current solution is context product describing itself with a type. Confused about how a group would fit.

~ We would come up with names that make sense. Instrument group, etc.

Question: How was it solved in the example with the balloon, gondola, spacecraft with instruments? How was that created?

Answer: We would have to look it up.

~ That's the type of thing we're talking about - a way to group. Could define as a group and control it that way.

~ Sounds like there's some wiggle room. Original issue needs to be refined to clarify the problem.

~ Not doing something different, just using the xml to reference members of the group.

We will wrap up here today.



We will either evote or vote next week on CCB-325.

## DDWG Notes 2023-03-02

---

title: DDWG Notes 2023-03-02

layout: default

date: 2023-03-02

---

March 2, 2023

Notes by Debra Kazden

Known Attendees: M. Drum, T. Hare, L. Huber, S. Hughes, R. Joyner, D. Kazden, T. Lim, J. Mafi, S. McLaughlin, B. Semenov and R. Simpson  
Known Observers: R. Alanis, A. Bailey, P. Lawton, S. Loftin, M. Tiscareno and E. Schaefer (For E. Guinness)

### ## DDWG Agenda

(Included in email sent February 28, from M. Drum, Agenda for March 2, 2023)

#### Front-matter/What's new

- We have one job today: Video as Product Observational
- Unless there's anything else urgent, I suggest we adjourn afterwards to return time back to attendees
- Approved SCRs by today

#### Ready for Vote

- CCB-325 - Support for video as product observational – Trent Hare - Final adjustments made. Please review before the vote.

\*\*(Voted to pass)\*\*

#### Tier 1 issues

- CCB-358 – Context object for Instrument Package – Anne Raugh **\*\*(Not Discussed)\*\***

#### Tier 2 issues

- CCB-211 – XML/schema based files as archival data – Jesse Stone, Mike Drum
- CCB-336 – Add a License Information class to the Identification Area – Tanya Lim
- DOI-rich metadata into IM (ticket to be created), replacing CCB-222 and CCB-261 – Anne Raugh and any volunteers
- CCB-164 - Display Settings not required for images -- (IMG) - Need to think more about potential updates to Imaging and how Display might fit.
- CCB-353 – Fully label all bytes in data files – Anne
- CCB-356 – Update Constraints on Author/Editor Lists – Dick Simpson
- Product Metadata Errata?
- Create a supplemental label to live next to original label, that can add aliases and supersede metadata **\*\*(Not Discussed)\*\***

#### Parking lot – need a driver!

- CCB-326 - Inconsistent Direction on How to Use Discipline Facets
- CCB-324 – Create a data type for Type attributes that require a permissible value list.
- RS: 259/281 (see new attachment to CCB-281) (on hold)
- CCB-261 - Add an optional method for indicating that a product is a member of a larger collection that has a DOI -- Last chance before recommending withdraw.

**\*\*(Not Discussed)\*\***

Next meeting March 16

# DDWG

## CCB-325 - Support for video and audio as product observational See <https://pds-jira.jpl.nasa.gov/browse/CCB-325>

We are going to keep it very simple today and just have a vote on CCB-325.

~ People have been busy getting it ready.

~ It's been a calm few days.

~ There was lots of discussion about metadata tags. People suggested pushing that off until later. Things like video time or frame size. This vote needs to go through so the MC can vote.

The white paper is a supplement to one of the policies. The MC needs to agree to the changes.

~ A change to get on the books is to look at encoding every five years. Discussed ARTEMIS yesterday. There will be a lot of audio and video. Lots of Mars missions will also need this. It's timely.

~ Documents should emphasize the exceptionality of this and that it creates extra work for PDS. We need to be clear to DPs.

~ Yes. The only way to maintain that is to make sure the language gets to the documents the DPs will read, like the handbook.

~ Once we start having manned missions, back room science talk and video will be another potential data swarm, so yes, we need to advertise the capability to the community, but don't think everyone will start adding video since we now support it.

~ That wasn't my concern, it's other encoded things.

~ Data nodes should make sure people are not putting archives together unsupervised.

~ Problem of opening the floodgates.

~ The DDWG is poking the hole in this case.

~ Current instrument teams do understand, but we'll need to be careful going forward.

Someone wants clarification on reviewing every five years.

~ Review the list of encoded video and audio formats and see if they continue to be viable. Might need to be translated. Not happy about it, but we can't leave obsolete formats in the archive.

Question: Is telling them to do it themselves an option?

Answer: Some might not get translated. They might just sit there.

~ There are challenges, and have been challenges with migration, things like images that were in a funky format. Checking every five years is probably worth it.

Another Question: Is that part of the policy?

Answer: Yes.

~ This is an exception we need to treat very carefully.

Question: Other questions or comments before we vote?

Answer: (Silence)

**\*\*The Vote for CCB-325**

ATMOS - Yes

EN - Yes  
GEO - Yes  
IMG - Yes  
IPDA - Yes  
NAIF - Yes  
PPI - Yes  
RMS - Yes  
RS - Yes  
SBN - Yes\*\*

That's a strong signal that we are going forward.

We can end early today since we have had so many extra meetings.

Question: Now this goes to the CCB and then for MC approval. What are the chances of this being in this build?

Answer: The CCB is voting Tuesday. As far as I know we are going forward with this build. We will discuss it at the MC next week. It can go forward for now.

~ Steve will get started on it **\*\*(Action Item)\*\***

**\*\*Action Item - Mike, Trent and Dick\*\*** will get together about presenting this at the MC.

It sounded like there was a green light at the last MC F2F. Did a whole comparison of data loss because Becky asked for it.

~ It's good to show we have done the work.

Next meeting March 16.

## DDWG Notes 2023-03-16

---

title: DDWG Notes 2023-03-16

layout: default

date: 2023-03-16

---

March 16, 2023

Notes by Debra Kazden

Known Attendees: M. Drum, S. Hughes, R. Joyner, D. Kazden, T. Lim, S. McLaughlin, B. Semenov and R. Simpson  
Known Observers: R. Alanis, A. Bailey, D. Kahan, P. Lawton, S. Loftin, M. McAuley, J. Padams, M. Tiscareno and E. Schaefer (For E. Guinness)

### ## DDWG Agenda

(Included in email sent March 14, from M. Drum, Agenda for March 16, 2023)

#### Front-matter/What's new

- CCB approved Video issue, but recommends we create an additional issue to extend support to Product\_Ancillary and Product\_Browse
  - Jira is migrating to the cloud. EN soliciting input about the option converting it to Open Source
- \*\*(Discussed)\*\***

#### Ready for Vote

- None this week

#### Tier 1 issues

- CCB-358 – Context object for Instrument Package – Anne Raugh
  - CCB-211 – XML/schema based files as archival data – Jesse Stone, Mike Drum
  - CCB-336 – Add a License Information class to the Identification Area – Tanya Lim
- \*\*(Discussed)\*\***

#### Tier 2 issues

- DOI-rich metadata into IM (ticket to be created), replacing CCB-222 and CCB-261 – Anne Raugh and any volunteers
  - CCB-164 - Display Settings not required for images -- (IMG) - Need to think more about potential updates to Imaging and how Display might fit.
  - CCB-353 – Fully label all bytes in data files – Anne
  - CCB-356 – Update Constraints on Author/Editor Lists – Dick Simpson
  - Product Metadata Errata?
  - Create a supplemental label to live next to original label, that can add aliases and supersede metadata
- \*\*(Some Discussion)\*\***

#### Parking lot – need a driver!

- CCB-326 - Inconsistent Direction on How to Use Discipline Facets
- CCB-324 – Create a data type for Type attributes that require a permissible value list.
- RS: 259/281 (see new attachment to CCB-281) (on hold)
- CCB-261 - Add an optional method for indicating that a product is a member of a larger collection that has a DOI -- Last chance before recommending withdraw.

\*\*(Not Discussed)\*\*

Next meeting: 3/30/23

# DDWG

## Front Matter/What's New

Not all regular attendees are here today.

MC and CCB approved CCB-325 (Support for video and audio as product observational, see <https://pds-jira.jpl.nasa.gov/browse/CCB-325>). Liens from MC and recommendation for CCB to also support video in product ancillary and product browse. We will need a ticket for that. Should be relatively uncontroversial. Will be for the next build.

Jordan is here to discuss jira migrating to the cloud. The jira license expires in a month. The license is not cheap. Moving to the cloud. Open source. Not easy to change. We can moderate the comments.

Question to Jordan: Will the cost of your time to moderate this be less than the cost of the license?

Answer: If it becomes a burden we could get a license. They will require a log-in, so we can block bad actors. Like GitHub.

~ Bad actors are a big concern.

~ GitHub is a good example.

~ Free. Can try this out. Was part of the Roadmap to open up the PDS4 standards process.

Question: Will everything in jira be made public?

Answer: Will go through and scrub it first. Don't think it's worth a big effort to look for sensitive information. Doubt there's any social security numbers or other sensitive information. System paths are the only real security concerns - if any are there.

~ Good idea.

~ Someone is still concerned. We never paid attention when we submitted comments in the past. This is a big change now. There are issues for us to be aware of.

~ Like a Slack channel.

~ It might be a good idea to filter all the comments to find anything.

~ In reality, ChatGPT will go through it all in a month or two.

~ If it's open sourced it will get indexed by Google.

~ There are important internal deliberations. Not the face we want to present to the world. This is deep inside the sausage factory. Not what we want people to see.

~ Someone disagrees. That's the point of open science. We want people to see how the sausage is made. Could make everything up to now history.

There's a question about policy decisions - if we want to make this public, like GitHub, but don't think this is policy driven.

~ True. There have been complaints and the technology is making us ask the question now if we should open up jira. If we need to keep the cost we can figure it out. It's in the tens of thousands. For an actual cloud license is less.

~ An extreme paraphrase - not moving to PDS AWS cloud, this is to the jira cloud. Don't think we need to worry about performance. This just changes where it's installed. Goes from us to jira.

~ About 5K a year.

Question: If we take into account privacy and such, should we give it more thought?

Answer: We can move forward.

~ It's unfortunate we didn't get to bring this up to the MC.

~ We can do this. We can migrate to the cloud and get it protected for now.

~ A noble goal. We can keep this on the agenda and get it open sourced eventually.

Someone wants some clarification.

~ Our license expired in one month. Port (?) license expires in a year. On-site license is more expensive.

~ So, we can move to the cloud with protection and decide what to do when that license is up.

~ Yes.

Question: If our deliberations on jira should be public, then the next logical step is to make these meetings public. Do we want people we might talk about to be at the meetings?

Answer: When we are in the meetings or responding in jira it's sometimes hard to keep your cool.

~ Status quo for now.

#### # Tier 1 and Tier 2 Discussion

Some issues have been moved from Tier 2 to Tier 1 for this new cycle.

Question: Is there anything we really want in the next build, that's in approximately six months, that is not here yet? Something in the parking lot?

Answer: Someone asked that anything with his name on it moved up so he can focus on radio science. He needs to stop spending so much time on DDWG solutions. Needs time to do other things. This includes CCB-356 (Update Constraints on Author/Editor Lists, see <https://pds-jira.jpl.nasa.gov/browse/CCB-356>).

The resolution for some items could be that they are withdrawn.

In Tier 2, the first item, DOI-rich metadata into IM, was proposed by Anne. Now there's a requirement from EN to improve harvesting of metadata for DOIs. Expect this will get bumped to Tier 1 soon. People should let Mike know if anyone is interested in helping with the topic.

Mike will send out a question about shifts in our priorities with the next agenda so we can start to get the lay of the land for the next build.

In Tier 1 there are three issues. They are unchanged. We discussed CCB-358 (Need a Context object for Instrument Packages, see <https://pds-jira.jpl.nasa.gov/browse/CCB-358>) a few weeks ago. We tried to clarify the issue. People are still interested in trying to get it over the line.

Mike will continue to work on CCB-211 (XML/schema based files as archival data, see <https://pds-jira.jpl.nasa.gov/browse/CCB-211>) for this build cycle.

# CCB-336 - Add a License Information class to the Identification Area See <https://pds-jira.jpl.nasa.gov/browse/CCB-336>

There has been good clarification from Steve Crawford on the position that NASA takes on this. Everything is CC0, unless it isn't. It's a bit messy, but we should get the structure into the labels to support this. If we need to add more to it in the future, we can do it on a case by case basis. There was a recent comment in jira that sounds like a language disparity for us to consider. Also that CC0 might not meet ESA's needs.

~ That shouldn't stop us going forward here.

~ We will move forward can can add things as necessary as we move forward.

---

Question: Any other topics for today?

Answer: (Silence)

Next meeting March 30.



## DDWG Notes 2023-03-30

---

title: DDWG Notes 2023-03-30

layout: default

date: 2023-03-30

---

March 30, 2023

Notes by Debra Kazden

Known Attendees: M. Drum, E. Guinness, T. Hare, L. Huber, S. Hughes, D. Kazden, J. Mafi and S. McLaughlin

Known Observers: R. Alanis, A. Bailey, P. Lawton, S. Loftin and M. Tiscareno

### ## DDWG Agenda

(Included in email sent March 28, from M. Drum, Agenda for March 30, 2023)

#### Front-matter/What's new

- Start thinking about issues you'd like to complete this round. Since things are quiet, it's time to start culling the backlog.

\*\* (Discussed) \*\*

#### Ready for Vote

- None this week

#### Tier 1 issues

- CCB-358 – Context object for Instrument Package – Anne Raugh - Rewrote proposal, vote next time?

- CCB-211 – XML/schema based files as archival data – Jesse Stone, Mike Drum

- CCB-336 – Add a License Information class to the Identification Area – Tanya Lim

- CCB-356 – Update Constraints on Author/Editor Lists – Dick Simpson

- \*\* (Discussed) \*\*

#### Tier 2 issues

- DOI-rich metadata into IM (ticket to be created), replacing CCB-222 and CCB-261 – Anne Raugh and any volunteers

- CCB-164 - Display Settings not required for images -- (IMG) - Need to think more about potential updates to Imaging and how Display might fit.

- CCB-353 – Fully label all bytes in data files – Anne

- Product Metadata Errata? Create a supplemental label to live next to original label, that can add aliases and supersede metadata

\*\* (Discussed) \*\*

#### Parking lot – need a driver!

- CCB-326 - Inconsistent Direction on How to Use Discipline Facets

- CCB-324 – Create a data type for Type attributes that require a permissible value list.

- RS: 259/281 (see new attachment to CCB-281) (on hold)

- CCB-261 - Add an optional method for indicating that a product is a member of a larger collection that has a DOI -- Last chance before recommending withdraw.

\*\* (Discussed) \*\*

Next meeting: 4/13/23

# DDWG

## Front-matter/What's new

Lite today, so added an action item for everyone to start thinking about things we need to complete. We got a lot of big things out of the way. This is our opportunity to look back at things in the backlog. Want an early action item for all the nodes to start looking at jira. We will also do it as a group too. There are still some Tier 2 issues that are not really moving, and maybe some in the parking lot. If we can get all the way down the list maybe we don't need to meet every week.

# Tier 1 issues

## CCB-358 - Need a Context object for Instrument Packages See <https://pds-jira.jpl.nasa.gov/browse/CCB-358>

Mike met with Steve about this and they put together an new Problem Statement. It really boils down to adding new value to context product for instrument host, type instrument package.

(Note taker missed something due to an audio issue)

This looks like a better fit.

~ Great, but we still want approval from Anne.

\*\*Action Item - Pat\*\* will pass this issue to Anne.

Barring any drama, we will put this on the agenda for a vote next time.

Question: Does it need a TA?

Answer: Yes.

## CCB-211 - XML/schema based files as archival data See <https://pds-jira.jpl.nasa.gov/browse/CCB-211>

Nothing new. It will be worked on after licensing.

## CCB-356 - Update Constraints on Author/Editor Lists See <https://pds-jira.jpl.nasa.gov/browse/CCB-356>

There have been conversations, but no movement yet. Brought forward in the agenda. Dick (not here today) will speak on this next time.

# Tier 2 issues

If there are any updates for these let Mike know. DOI rich metadata is still on the table for Anne, EN and others who are interested.

## CCB-164 - Display Settings not required for images See <https://pds-jira.jpl.nasa.gov/browse/CCB-164>

People don't remember what this is about. Its an old issue. Doesn't seem complicated.

Question: Will someone take it up?

Answer: Matt will discuss it with Mitch. **\*\*(Action Item)\*\***

Someone believes it was an issue with how LDDs designed. Needed many instances to do references. Think we needed a change to internal reference. Memory is fuzzy, but it might not be a big issue, just some things to work out.

~ If people want to solidify their memories and leave comments than we can discuss this next time.

## CCB-353 - Anonymous bytes in data files are problematic and should be documented See <https://pds-jira.jpl.nasa.gov/browse/CCB-353>

This is relatively new. Marked urgent. Maybe we should move it up in the agenda. There's a lot here to chew on. Seems like a fine thing to get through or withdraw.

~ Someone thinks it will cause a real problem to implement. Would recommend that we withdraw this, but we can discuss it.

~ It might not be as scary as it sounds. Having to define bytes as data or whatever.

~ Someone else is not in favor of this. Sometimes there are things like gutter space.

~ It will be moved up in the agenda to review for a bigger discussion.

## Product Metadata Errata

This came up at the last F2F. Basically, the idea is that we want to create labels next to older labels with errors or whatever. Need to have something for the registry. If this is interesting to anyone we need to build it up. One thing it could be used for is LID changes. Believes there's an alias that can be used in the opposite direction.

Question: Any volunteers?

Answer: Matt will discuss this with Mitch.

# Parking Lot

We haven't looked at these in a long time.

## CCB-326 - Inconsistent Direction on How to Use Discipline Facets See <https://pds-jira.jpl.nasa.gov/browse/CCB-326>

This issue was reported by Dick, so we will wait until he's here to discuss this.

## CCB-324 - Create a data type for Type attributes that require a permissible value list.  
See <https://pds-jira.jpl.nasa.gov/browse/CCB-324>

This is a good data modeling one.

~ It would solve a problem. Unsure why we didn't do it a long time ago, but if people don't think it's that

important we could just vote.

- ~ There was a lot of discussion over type attributes.
- ~ Someone said it's been a long while, but thinks it had to do with attributes that end in underscore type. An issue with ingest LDD. Trying to figure it out, particularly for LDDs that have been around a while. Didn't want to change the name.
- ~ Need to make type attribute specific, not figure it out from the name.
- ~ Someone thought there was another issue too.
- ~ Everything should be declared and needs an enumerated list. Not sure why we didn't do that. Seems like a simpler fix.
- ~ Had lots of discussions about best practices for LDDs, but then saw what happened when implemented.
- ~ Can simply declare every attribute whether it needs a list or not. Fix should be fairly simple.
- ~ There was a fix with the DD attribute. Been awhile since it was discussed.
- ~ Shouldn't need that. Every attribute should be declared and the tool needs to be fixed.
- ~ Looking at the Requested Changes, the proposal was to add ASCII data type or UTF-8 data type, instead of ASCII short string.
- ~ Makes sense to someone, but still need to make sure it doesn't break old LDDs with underscore type.
- ~ Not backwards compatible. It is correct that we need to fix the tool. Maybe throw a warning.
- ~ Throwing an error.
- ~ Not working. If there's a single error it comes to a stop.
- ~ The simple solution is to make it a warning, but people will ignore it.

Moving this to Tier 1. We should at least figure out how we want to move forward with this for this build.

## CCB-261 - Add an optional method for indicating that a product is a member of a larger collection that has a DOI See <https://pds-jira.jpl.nasa.gov/browse/CCB-261>

Someone believes this is superseded by what Anne is doing.

- ~ It will be removed. We can add it back if need be.

DDWG Notes 2023-04-13

---

title: DDWG Notes 2023-04-13

layout: default

date: 2023-04-13

---

April 13, 2023

Notes by Debra Kazden

Known Attendees: M. Drum, E. Guinness, T. Hare, L. Huber, S. Hughes, R. Joyner, D. Kahan, D. Kazden, T. Lim, J. Mafi, S. McLaughlin, B. Semenov and R. Simpson

Known Observers: R. Alanis, M. K. Crombie, P. Lawton, S. Loftin and M. Tiscareno

## ## DDWG Agenda

(Included in email sent April 11, from M. Drum, Agenda for April 13, 2023)

### Front-matter/What's new

- EN is requesting input: Currently, the SR does not state any requirements that <Table-X> definitions in a label are not required to match the order of the file they are referencing, in the case of multiple tables. Validate enforces this already; should we adopt the restriction formally, or request Validate drop that validation?

- Let's talk IM 2.0

\*\*(Discussed)\*\*

### Ready for Vote

- CCB-358 – Context object for Instrument Package – Anne Raugh **\*\*(Did Not Vote)\*\***

### Tier 1 issues

- CCB-353 – Fully label all bytes in data files – Anne – Please review for a discussion

- CCB-336 – Add a License Information class to the Identification Area – Tanya Lim

- CCB-211 – XML/schema based files as archival data – Jesse Stone, Mike Drum

- CB-356 – Update Constraints on Author/Editor Lists – Dick Simpson

- CCB-324 – Create a data type for Type attributes that require a permissible value list. - Figure out where we want to go with this **\*\*(Not Discussed)\*\***

### Tier 2 issues

- DOI-rich metadata into IM (ticket to be created), replacing CCB-222 and CCB-261 – Anne Raugh and any volunteers

- CCB-164 - Display Settings not required for images -- (IMG) - Need to think more about potential updates to Imaging and how Display might fit.

- Product Metadata Errata?

Create a supplemental label to live next to original label, that can add aliases and supersede metadata **\*\*(Not Discussed)\*\***

### Parking lot – need a driver!

- CCB-326 - Inconsistent Direction on How to Use Discipline Facets

- RS: 259/281 (see new attachment to CCB-281) (on hold) **\*\*(Not Discussed)\*\***

Next meeting: 4/27/23

# DDWG

Jordan sent a request. SR has no requirement for tables, specifically for products with multiple products in them, that the order of products match what's in the label. Validate tool does. Jordan asked for a recommendation. Wanted to know if we should enforce what validate tool does. Would either have to update the SR or validate tool.

Question: Someone didn't think all data objects are required to be in the same file...?

Answer: Think we're talking about things in the same file.

~ We should provide a recommendation for multiple files too. A complication, but not in favor of this.

~ We do ask for an order of columns in the table with elements level metadata. It would be strange to require ordering without that called out.

~ We have offset. Wouldn't want to enforce this as a SR requirement.

~ There might be a reason a DP didn't do it.

~ Someone thinks this is hitting hard for Mars 2020. It could be a best practice. Computers are smart. Not sure why this would need to be a requirement.

~ It's for human readability. Think Jordan is worried we're not crossing boundaries. Extra work that might need to be done.

~ Seems like it works.

~ The issue is validating against the data. Easier when the offsets are in order.

~ Think there was a discussion of If byte versus start byte.

~ Should be doing it anyway.

~ It's software versus DPs. It would make it easier to validate if it's in order.

~ It could be a warning.

~ We are mentioning this for Jordan.

~ This is hitting for Mars 2020 in imaging products that used to validate. The project discovered an issue with recent files. Developers decided to do this without input.

It's a bug in validate. Question is if we want to keep it and enforce it in the SR.

~ The real issue is how comments will be collected. Maybe this should be an SCR.

~ Good point. We should update the SR either way.

~ Best practices go in the DPH.

~ Jordan wants us to decide if they are required to be in order or not. Jordan says developers go to the SR and in some cases it's not clear. If order doesn't count, we should state that.

~ Saying not a requirement is a requirement. We should tell developers there is no such requirement.

It would be helpful if the SR said byte ordering for arrays or tables is or is not required. Putting it as a best practice makes it hard. Hard to compare the DPH to the SR. Easier if the DPs are clear.

~ Someone ran into this with a PDART. It validated, then failed when re-validated. Unsure if we should tag it to the version of validate it passed. Seems like a validate bug. Awkward to have it suddenly not validate.

~ Validate breaking is a huge problem. NSSDCA requires us to use the latest version of validate.

Backwards compatibility issue. We need something in validate to fix this.

~ For developers, it's always true that some labels will pass one version of validate and fail another. A bug, but if it's preventing us from archiving data...

~ It's a huge request to make it required. Missions might have to redeliver. We need to think about this

carefully.

~ There are several issues. First, it's unclear if this is a bug in validate, so we want the SR to be clear if order counts or not. We need to close loop holes. Second, backwards compatibility issues. We should make this an SCR to capture all the comments.

We can do a straw poll to see where we are. We can create a document change request SCR and handle it in two weeks. It could also be a best practice for the DPH. Can provide a recommendation from this meeting.

~ There are multiple choices for the straw poll. Bug fix because software developers invented a requirement. We can handle documentation, but it's a can of worms. Developers were wrong. We should take it out.

Question: When was the change made?

Answer: In the last six months.

~ That's recent on the PDS scale.

~ Most data will pass.

~ It may be more pervasive than you think. There used to be advice that arrays should be in alphabetic order instead of byte order. It's a big issue in peer reviews, like for ORex in 2011, external reviewers. Don't care personally, but there should be a decision now. In a different review they wanted arrays ordered by importance. Behind all of this is a bigger can of worms than we expect.

~ Someone thinks this is a bug fix for validate.

~ Someone else is not hearing a lot of support for option 2.

The question is really if we want ordering to match byte order of elements.

~ Someone is also pushing for a vote on clarifying in the SR.

~ Our first choice would have to go in the SR. We should vote first.

~ That could be the recommendation we send to them.

**\*\*Strawman Poll of everyone on the call now Option 1 is do not adopt requirement that elements match order Option 2 is that we should adopt a requirement that label elements match the order in data objects.\*\***

**\*\*Ron - Abstain**

**Boris - Option 1**

**Danny - Option 1**

**Debra - Option 1**

**Dick - Bug Fix\*\***

Question: Point of order - every one on the call is voting even though not voting members?

Answer: Yes. This is not an official vote. Option 1 is do not adopt requirement that elements match the order in data object.

**\*\*Dick - Option 1**

**Ed - Option 1**

**Joe - Option 1**

**Steve - Abstain**

**Lyle - Abstain**

**Kate - Abstain**

Matt - Option 1  
Pat - Abstain  
Sheri - Abstain  
Stef (Silence)  
Tanya - Option 1  
Trent - Option 1\*\*

Sounds like the recommendation from the DDWG is that this is not a new requirement.

Question: Does anyone want to write up the language for the SR and DPH in an SCR?

Answer: Someone votes for Steve.

~ Someone else seconds that.

Answer from Steve - Okay. \*\*(Action Item)\*\*

Another person is not sure the SR needs to be updated. A non-requirement. Could update the DPH.

~ Someone agrees. Don't need to bring it up in the SR and accidentally encourage people to do it. Could put it in the DPH in best practices.

~ Someone else agrees with the concern, but declarative statements in negative are important. It's a practical document that people will read.

~ Best practice is to say that things are not required. SR would get very long if we included all that things that are not required.

Question: What's the best practice?

Answer: Put them in byte order.

~ Yes. Order doesn't count and the best practice is to put it in order.

~ Machine can read it. Human will have to figure it out.

~ Saying it counts or doesn't count.

~ The requirement is on the IM. DPs don't need to know it's not a requirement.

~ If someone is building labels they need to know how to do it.

~ It's natural to put in the order in the file, but we don't want to require that.

~ Analytics world will want to know if they can assume they are in order.

~ Good point. If writing software need to know so you don't make a wrong assumption like happened with validate tool.

~ People can test or ask DDWG.

~ That would mean everyone writing software would have to get on our agenda or write test cases for every single thing.

~ They can ask a node.

~ The nodes might not always know.

~ Eventually Chatgpt will absorb it and know.

~ It has to be written down.

~ A negative presence of a non-requirement would both be and not be a requirement.

~ This is a one liner footnote. Table order doesn't matter.

~ Fine.

We'll keep this on the agenda so we can get an SCR. Will send the straw poll as a recommendation.

Another Front Matter topic is SPD41.

~ Steve got pulled into a standards committee. The concern is what we need to address. Tom thought it



was a minor issue, but someone else said there's a mandate in Appendix E for missions, so bringing this up as a DDWG issue that should be an SCR so we can do the mapping and get feedback to the committee.

~ Someone just posted SPD41A.

~ The idea is to at least respond back to the standards group. If mandate then a whole other issue.

\*\*Action Item - Steve - will write an SCR that can become a formal recommendation.

There were about four more issues Jordan wants us to discuss. We can look at them next time. Someone let Jordan know the DDWG was looking for work.

~ We'll get them on the agenda.

IM Version 2.0.

Someone thinks it's a good idea to re-open the discussion on this and make it happen. We're not any closer than we were at our last discussion. We will at some point have a discussion of topics and what we need to do for 2.0, if we want a 2.0.

~ First we need to agree what it means to go to 2.0.

~ Someone agrees.

~ Someone else thinks it has to do with backwards compatibility.

~ At 2.0 that's not allowed. Then we would go to PDS5.

~ People are worried about that.

~ Non-backwards compatible changes happen.

~ They do happen, but would increase from first version to 2.0.

~ Someone does not want to be migrating into something that evolves every two years.

There is plenty to talk about at future meetings. We have an opportunity to have a tech session at the July F2F. We could discuss 2.0 there. Other possible topics are also being floated.

# Ready for Vote

## CCB-358 - Need a Context object for Instrument Packages See <https://pds-jira.jpl.nasa.gov/browse/CCB-358>

This has a TA now. It identified that it was missing the Requested Changes and added that.

Question: Any final thoughts?

Another Question: Is it already the case that it can be recursive?

Answer: Someone thinks the answer is yes.

~ It's essential to making this work.

~ Can reference product context by product context.

~ That needs to be added. We won't vote now.

Question: Is that what was meant - values for the reference type?

Answer: Need to be able to have everything in place.

~ Someone agrees. Need a new permissive value. Could be a lien.

Another Question: Is that okay?

Answer: We also need to edit the DPH, Appendix J, I think. Formation rules for context products.

~ That's another good point. We can punt on this so we can cross the Ts and dot the Is.

Question: Why do we need instrument package if nesting?

Answer: Instrument package is a type of host.

~ Someone needs to digest this.

Question: What technically constitutes an instrument package?

Answer: RALPH is an example. It shares a lot.

Another Question: But what specifically makes it a package?

Answer: Open to recommendations on this, but the intent behind this is that there are groupings of instruments that are referred to as one thing, but different instruments that each perform a different specific scientific function. Open to a different view.

~ Vagueness is important. Trying to write a context product for the instrument package. How it's constituted will vary from case to case. In Ulysses COSPIN the instrument was called an instrument, but it was really five sensors built and operated by separate teams.

~ Another example is Cassini MIMI.

~ All specified as instrument, but could put in the label with multiple context products to show the relationship.

~ Tried to differentiate how to reference in the context product and in the labels.

~ So in the RALPH case it's the instrument host.

~ Yes, and instrument host to instrument host would be in the context product.

The SCR will be clarified and we will vote next time.

---

That's it for today. We will discuss CCB-353 next time.

DDWG Notes 2023-04-27

---

title: DDWG Notes 2023-04-27

layout: default

date: 2023-04-27

---

April 27, 2023

Notes by Debra Kazden

Known Attendees: M. Drum, T. Hare, L. Huber, S. Hughes, R. Joyner, D. Kazden, S. McLaughlin, B. Semenov and R. Simpson

Known Observers: P. Lawton, S. Loftin and M. Tiscareno

## ## DDWG Agenda

(Included in email sent April 25, from M. Drum, Agenda for April 27th, 2023)

### Front-matter/What's new

1) Prospective Tech Session topics:

a) Context Product cleanup exercise

b) 2.0 - Define the meaning and goals

\*\*(Discussed)\*\*

### Ready for Vote

- CCB-358 – Context object for Instrument Package – Mike Drum **\*\*(Voted no)\*\***

### Tier 1 issues

- CCB-353 – Fully label all bytes in data files – Anne – Please review for a discussion

- CCB-336 – Add a License Information class to the Identification Area – Tanya Lim

- CCB-211 – XML/schema based files as archival data – Jesse Stone, Mike Drum

- CCB-356 – Update Constraints on Author/Editor Lists – Dick Simpson

- CCB-324 – Create a data type for Type attributes that require a permissible value list. - Figure out where we want to go with this

- Create a ticket that proposes changes to the SR and/or DPH to clarify that data objects described in a label do not need to match the order they appear in a file.

\*\*(Some Discussion)\*\*

### Tier 2 issues

- DOI-rich metadata into IM (ticket to be created), replacing CCB-222 and CCB-261 – Anne Raugh and any volunteers

- CCB-164 - Display Settings not required for images -- (IMG) - Need to think more about potential updates to Imaging and how Display might fit.

- Product Metadata Errata?

Create a supplemental label to live next to original label, that can add aliases and supersede metadata

\*\*(Not Discussed)\*\*

### Parking lot – need a driver!

- CCB-326 - Inconsistent Direction on How to Use Discipline Facets

- RS: 259/281 (see new attachment to CCB-281) (on hold) **\*\*(Not Discussed)\*\***

Next meeting: 5/11/23

# DDWG

## Front Matter

Prospective tech session can be scheduled. There are some looming appropriate questions that we could discuss. This includes the need to fix a lot of the older context products.

~ Since the DDWG task list is lite going forward, it seems like we could do something similar to what Mitch did for the LDD clean up for the context products. We need to spend some time as a group cleaning them up. All of us need to work on it. It will really improve the system and connectivity. It would make a start towards a meaningful version 2.0.

~ Someone said it feels like one or two experts could do this.

~ Someone else answered they suppose it could be farmed out, but don't think that's the right thing to do. Think this is the right group to do it. Unsure how long it would take. Haven't scoped out the changes, but maybe a three day workshop in the fall.

~ Another person said they definitely under-appreciated how important these are. Many were auto-imported. We should consult with EN on this.

~ Agreed. Did discuss this with Richard in email. He's unsure how much needs to be in the context products. Some of the PDS3 text should go away. What we really need are linkages going both ways.

~ One reference that someone wishes existed is target to target. A missing piece.

Question: Like satellite to planet?

Answer: Yes, and nested satellites and stuff like that. Hierarchies, parent/child relationships. Can create groupings. Need to scope it out.

~ We need to discuss the directional links so we can get the new SCRs.

~ Agree we should do this. Jordan has commented that he's all for it. There is work on target ontology being done. Jordan is especially interested because it will make search more robust. SBN is working on it too. Jordan also has a list of five critical things he wants done. We started discussing them last time. Submitted an SCR on the last one. Higher priority than this.

~ We will be sure to get them on the agenda.

Another SCR coming up on author list. Jordan is interested in that. Need to get support for DOI metadata sorted out. High priority.

~ Right. Still a lingering question on what we want to do. Don't want to copy/paste entire schema in IM.

~ Right. Need to figure out what metadata we need.

~ Been on the agenda awhile. Anne was planning to spearhead it.

~ Also, CCB 358 (see below) and five things Jordan wants addressed. Need to figure out what to do about LIDs changing too.

~ No problem, but unclear on what bundle reference by LID means.

~ Reference bundle by LID is causing issues. Jordan wants a discussion to clarify.

~ Someone can see one or two people doing a wonderful job on context products, but won't help with how they are referenced from products. Thinks would have to survey a lot of products. This this is a much bigger job that it appears.

~ Fair point.

Question: Someone remembers in St. Louis talking about not wanting LIDs to change. Sounds like this

isn't true. Did the DDWG approve an infrastructure for LIDs to change?

Answer: Some LIDs for context products have changed. There are LIDs pointing to context or other products that are deprecated and have newer products with different LIDs. A problem to solve.

Question: Should we submit SCRs for Jordan's concerns?

Answer: Yes. We should have tickets for them. Put them in and we can flesh them out and have discussions.

**\*\*Action Item - Steve\*\*** will put them in. He does not want to be the stuckie.

# Ready for Vote

## CCB-358 - Need a Context object for Instrument Packages.

See <https://pds-jira.jpl.nasa.gov/browse/CCB-358>

There has been plenty of Jira discussion in the last week. A few revisions were made due to the comments. A few new Requested Changes, including that the DPH needs to know how to use them for LID formation and the definition of type for instrument reference needs to specify /use lowest granularity, but doesn't always make sense, so changed to be more like other definitions. Open to feedback. Believe there are no open questions left.

~ Someone suggested inst host to inst host be limited so one end is a package, but people didn't want to do that.

~ Rewriting of the definition, saw might have use for things like spacecraft to spacecraft. Could see us not wanting to bake that requirement in.

~ Someone agrees. Doesn't quite cover all bases. Would need to discuss again. And not always an instrument package, could be a package of things.

~ Yes, but defining instrument package now.

~ Could have other packages later.

~ Keep this constrained to instrument packages. Inst host talked about mounting. Original SCR talked about more. A lot of possibilities. If we're talking about groups that operate in some way together that's a big change. Okay with instrument package. General package worries me.

~ Agrees. Since we're talking about context product we can be more fine brained.

Question: Any other comments?

Answer: (Silence)

Another Question: So, we're voting on instrument package?

Answer: Yes. As written. Instrument package, not general package. There was another possible lien. Think it was Requested Changes... Yes, it was instrument structure, LID include fragment for spacecraft or specify a host.

~ A unique identifier for human readability. Would just come up with a standard way.

Another Question: For guidelines?

Answer: Yes, a DPH issue. Regarding instrument LID formation.

Another Question: So how LID is formed? Wanting to allow options?

Answer: Yes. Question is update guidelines? And update LIDs? Hopes not.

~ Don't remember context LIDs being a DDWG issue.

~ Usually left to Richard.

~ LIDs just have to be unique/unusual.

**\*\*The Vote for CCB-358**

ATMOS – No  
EN – Yes and No  
GEO – Not Here  
IMG – No  
IPDA – Not Here  
NAIF Abstain  
PPI – Abstain  
RMS – Yes  
RS – No  
SBN – Yes\*\*

So, 3 yes, 4 no, but two votes for EN. 2 absent and 2 abstain.

~ EN doesn't get two votes – make it an abstain.

~ So, 2 yes and 3 no.

\*\*Action Item - Mike\*\* will look at next steps.

It doesn't pass.

~ It can be refined. Can try to convince folks this is the right thing to do.

~ Someone was convinced this is confusing and unnecessary and awkward by discussing with another person who had strong feelings.

~ Someone else can see that argument, but SBN sees this as a blocker, like in the case of Ralph. That's how these are referenced. Need to document in an intelligent way.

~ Just call Ralph an instrument. Can do it that way.

~ Don't want to shoehorn. Think the issue is discovery. It becomes critical for search. Need to know it's a package. Would recommend the goal is clarified to identify the need. Jordan is very supportive of this. Ontology. Being able to relate things to each other. Very important information. Will make discovery much easier.

Question: Not a fan of composite type, but isn't this a case for that?

Answer: It might work if we can get the relationships in. Think many people have issues with how observing system works. May be a difficult solution.

~ PDS doesn't have good search and discoverability yet. Okay to table this and let this fester as that improves.

~ If we don't solve it it will be solved elsewhere and we may not like the solution. Garbage in and garbage out. Shoehorn is one way of garbage in.

Question: So, if envision search and have references from instrument to instrument and will find they are related?

Answer: Won't be able to show differences between spacecraft with packages or instruments. Can lead to confusion.

~ Yes. Correct. Ralph isn't an instrument, but people will look for it. MVIC and LEISA are the instruments.

~ Need a hierarchy. The model is lacking that now. Unfortunate, but don't think we'll vote again today. Unsure what we will do. Will discuss it with Anne.

Maybe we can reconsider it next time. A piece here missing is how instrument products will point to each other, so, MVIC points to LEISA and they point to RALPH, so the connectivity is there. Still think

alias is the best solution.

~ EN uses alias to shoehorn things in.

~ They use an alias of name or identifier, not relationship.

~ For discoverability, system needs to recognize the other name.

~ Red herring. Don't think alias is the solution to the packaging problem. Different classification. Alias has nothing to do with class. Need to go back to modeling class.

~ EN is using alias as a way to mark a LID has been deprecated. LID instead of name. Overloaded feature.

~ On Jordan's list of things to address.

Question: Anyone opposed to revisiting this next week?

Answer: Someone would rather say this failed and have a new SCR.

~ Another person is willing to bring it back up. Record the vote, but we can reconsider it.

~ Not sure Jira supports multiple votes for a single issue.

~ Cleaner slate if we start over.

~ But no path forward for Anne.

~ She wanted it in the last build. It was marked urgent. Still some time before the next cut off. No huge rush. Might be a way to link the tickets together.

~ Lots of possibilities.

~ Reporter will look into it. Appreciates the willingness to reconsider this.

# Tier 1 issues

## CCB-353 - Anonymous bytes in data files are problematic and should be documented.

See <https://pds-jira.jpl.nasa.gov/browse/CCB-353>

Question: Have people had a chance to form opinions on this yet? Do we want to make sure every byte is accounted for?

Answer: Someone is opposed.

~ In a perfect world, this is awesome, but in the real world it's a can of worms.

~ Someone else is also opposed. Issue with validate too. Backwards compatibility is huge.

~ Ambiguous too.

Question: Is there a larger discussion we should have or should we send it back?

Answer: The argument is that it improves validation. Not sure that's true.

~ If there's a particular area with a hole maybe we can look at that.

~ VICAR and ISIS cubes have lots of blobs of stuff, but not sure what the point is.

~ If we are coming at this blind and don't know the format, just looking at a file and a label, might want to understand. Can see the value of allowing a good faith DP to describe every byte in a way that will help users down the line.

~ Allow is different from require.

~ Several people agree.

~ Someone does that already, but doesn't want it required.

Someone also doesn't think it matters for the MD5 stuff, which is why they don't think that should be in the labels.

~ A chicken and egg thing.

Question: Do you mean the MD5 in the label of the data product?

Answer: Of the file.

- ~ Someone disagrees. Has had files transfer with issues. Only caught because of the MD5s.
- ~ Someone else agrees we should do better than our laissez-faire MD5 stance.
- ~ Shouldn't be in a label, just for transfer.
- ~ There are tools, like Pooch, where you require MD5 and data. The MD5 issue with the cloud will bite us since we are not being consistent.
- ~ If we want to tackle the issue of end user quality of data we will have to look at this.
- ~ Cloud can potentially help us, but we are not there yet.

Reporter will try to restructure this ticket.



DDWG Notes 2023-05-11

---

title: DDWG Notes 2023-05-11

layout: default

date: 2023-05-11

---

May 11, 2023

Notes by Debra Kazden

Known Attendees: M. Drum, E. Guinness, T. Hare, L. Huber, R. Joyner, D. Kazden, J. Mafi and R. Simpson

Known Observers: R. Alanis, A. Bailey, M. Bentley (for T. Lim), M. K. Crombie, P. Lawton, S. Loftin, M. McAuley and J. Padams

## DDWG Agenda

(Included in email sent May 9, from M. Drum, Agenda for May 11, 2023)

Front-matter/What's new

- Deciding not to move forward with Instrument Package (358) or Fully Label Bytes (353)
  - This week we should spend some time discussing pressing issues from EN:
    - a) Bundles referencing their collections by LID instead of LIDVID - allow or disallow? (EN wants to disallow)
    - b) Metadata improvements for DOI metadata - our current Citation Information could be distinctly improved based upon our DOI Metadata Guidelines Document, e.g. make first/last names explicit, improve descriptions of various attributes, add ORCIDs, add funding organization, add RORs, etc.- Need to discuss if we want this in Citation Info, and need a redesign person/team.
    - c) What to do in the case of a Schematron/Schema version conflict
    - d) Implementation of a superseded\_by and supersedes reference type
  - Prospective Tech Session topics:
    - a) Context Product cleanup exercise
    - b) 2.0 - Define the meaning and goals
- \*\*(Discussed)\*\*

Ready for Vote

- None this week

Tier 1 issues

- CCB-336 – Add a License Information class to the Identification Area – Tanya Lim
  - CCB-211 – XML/schema based files as archival data – Jesse Stone, Mike Drum
  - CCB-356 – Update Constraints on Author/Editor Lists – Dick Simpson
  - CCB-324 – Create a data type for Type attributes that require a permissible value list. - Figure out where we want to go with this
  - Create a ticket that proposes changes to the SR and/or DPH to clarify that data objects described in a label do not need to match the order they appear in a file.
- \*\*(Minor discussion of 356 as related to topics from Jordan)\*\*

Tier 2 issues

- DOI-rich metadata into IM (ticket to be created), replacing CCB-222 and CCB-261 – Anne Raugh and

any volunteers

- CCB-164 - Display Settings not required for images -- (IMG) - Need to think more about potential updates to Imaging and how Display might fit.

- Product Metadata Errata?

Create a supplemental label to live next to original label, that can add aliases and supersede metadata  
\*\*(Not Discussed)\*\*

Parking lot – need a driver!

- CCB-326 - Inconsistent Direction on How to Use Discipline Facets

- RS: 259/281 (see new attachment to CCB-281) (on hold) \*\*(Not Discussed)\*\*

Next meeting: 5/25/23

# DDWG

## Front Matter

Following up from our last meeting. Mike spoke to Anne about instrument packages and bytes tickets. She was unhappy with how the vote went, but not interested in pushing it further. SBN will work around the DDWG. So, we can drop those SCRs and move forward.

~ Someone sympathizes with her, but still thinks the instrument package has merit. The issue was with the technical implementation.

~ Tried to get feedback on the implementation. Impression of the vote was that people didn't want it. Thinks it will come up again in the future, especially after the changes to context products, but doesn't see the urgency now.

~ PPI has treated Ulysses COSPIN as an instrument suite. Working on Cassini MIMI now. Will create three separate context products. Those are work arounds, but not a great solution. Doesn't really describe the relationships correctly.

~ Good point. Don't think the ticket addressed cases like that. Think how it will need to be structured will become clearer.

Steve is not here today, but Jordan is here. Wants to discuss some of EN's high priority topics. We discussed ordering a few weeks ago.

## Bundles referencing their collections by LID instead of LIDVID - allow or disallow? (EN wants to disallow)

The first question is about bundles being allowed to reference collections by LID instead of LIDVID. EN wants us to decide if we should allow or disallow.

~ For the system, just adding specificity. Lots of complications with provenance, re-usability and DOIs. Not having versions makes it hard to track past versions of snapshots in time. References by LID only can't do that. Argument was that we didn't want to have to update every time something changed, but it's not a good argument.

~ We always said the system would return the newest version, so someone is not sure they buy that, and if we do always update the bundle, not sure what we do with the older versions or if that invalidates the DOI.

~ No, but you lose the ability to track version in time without the LIDVID. If you have document and data collections, you can't tell which versions are related in time. Also, old bundle products can be moved off-

line or kept on-line with the version included.

~ EN seems very concerned with the previous products. Someone hasn't encountered much need to go back to previous versions.

~ This goes back to DOIs.

~ DOIs shouldn't be driving our model.

~ Reproducibility of science exactly. We need to be able to provide exactly what was used.

~ That's a low priority driver.

Question: Okay, but why not?

Answer: Backwards compatibility.

~ The system has to support the past, but future wise we can evolve and be better.

~ Someone would have to think on that, but the driver is not very compelling.

Question: If EN is going to enforce it with the bundle inventory, will they change the collection inventory? At the bundle level secondaries have a definite need for LIDVIDs.

Answer: Right.

~ At PPI, we have always used LIDVIDs, but there are instances where we might want LID.

~ Someone is still not seeing a compelling reason for this, especially if dealing with accumulating bundles. Adding new products, but the old ones are not changing. Not sure why we need to know what was there on a certain date. Also, having to keep old bundle products. Think most users won't care.

~ PDS has to support reproducibility of science.

~ Yes, we do. That's what an archive is all about. We can still track down exact versions. It might take longer than a click.

~ Defending EN - sees the need in the future. If all we have is registry and labels, we won't be able to guarantee that in some deliveries they didn't change old products, even accidentally. The feedback is a lot of friction in creating new and maintaining old products. NSSDCA didn't want us to do it, but if we can address the friction easily, we should.

~ Problem is at the product level. Collection inventories are by LIDVID. Sure, if you add a new collection to the bundle you increment the bundle and document the changes. If anything, might consider making modification history required.

Question: What if you add a document to a release of a data collection that enhances science? How can you know someone got the information?

Answer: This isn't in PDS's charter to answer.

~ A lot isn't.

~ That's not something we should be expected to answer.

~ If the document is updated, we would update the document collection and update the modification history. The VID doesn't tell you what happened.

~ It doesn't, but it adds traceability for the system. It will group them together. Persistent perspective of versions says things are related in time. Reason for not doing it is precedence. Seems like the reasons not to do it aren't good.

~ LIDVIDs in bundles doesn't solve the problem.

~ Also, while doing landing pages for past versions, can't show related collections.

~ Hopefully, we can keep discussing this. Seems obvious to someone. Unfortunate that we are allowing friction to be an issue.

Question: Does anyone besides SBN support this?

Answer: PPI would support this. We already do use LIDVIDs. It's helpful with ROSES to be able to say

what versions are allowed.  
~ IMG would be on board too.

Maybe this needs to be a ticket on the agenda that we can vote on. We will finish the discussion later.

## ## Metadata improvements for DOI metadata

It would be beneficial to gather metadata for DOIs. There's been a lot of discussion around the need for this, but no one has taken it up.

~ As EN is going through the DOI stuff there are some they have to manually curate. It would be helpful to have it there.

Question: Information that would be collected when getting a DOI?

Answer: It would be in the labels. Not have to ask the DP. It's multiple steps now. Credit for funding, curating, etc. Information that is useful to capture.

CCB-356 (Update Constraints on Author/Editor Lists, see <https://pds-jira.jpl.nasa.gov/browse/CCB-356>) was requested by Anne and Ron to increase information on the authors in the citation information. Maybe this can be combined with that SCR.

~ Someone said it should all go together in their opinion.

~ We need a list of things that should go in there.

Question: Is anyone interested in making up the list?

Answer: It sounds like EN already has a list, and Anne had a lot of information in her head.

~ Jordan can take the first stab at it.

~ Dick is the point person on CCB-356. The list can be integrated with that.

~ Dick said he wrote it, but he is not the owner. It really comes from Anne and Ron.

~ Jordan is volunteering Ron to work on this.

~ Ron said it's been in progress for awhile, just not solidified.

The SCR will be updated to reflect the job to be done.

## ## Schematron/Schema version conflict

Basically, this is in the top of a label. (Sharing screen to show XML at the top of a label) Schema reference, like version 1700. No where in the SR does it say it has to be the same version and not a mismatch.

~ Validation should catch that. Not sure it needs to be in a formal document.

~ Yes, but we have gotten in trouble in the past. Might want to add a thou shalt not. There is a warning in validation now.

~ This should be called out as an issue. Use of schema and schematron should be mentioned somewhere and validation should call it out.

Question: How does EN determine the version? By looking at the string?

Answer: Yes, for now.

~ Apparently, this is a problem, but if we make it required it could make things that used to validate not validate now.

~ We should document that explicitly.

The DDWG will adopt the ticket and figure out where the correct place to document that is.

### ## Implementation of a superseded by and supersedes reference type

Question: Is this just for the case of a new version over riding an older one or LIDs that change?

Answer: It was thought of for LID changes. It's overkill for versions.

~ Seems like this should be done in the registry. Hard to see how DPs would create new versions and do the updates. It should be done in the registry.

Another Question: What happens when the registry goes away? There are already ideas about how superseded should work. Registry can do it.

Answer: The concern is that if we create a tool, people will use it, but we want to discourage this.

Wonders if there's a god-mode only way to implement this so it can be new-able by end users, but not usable for DPs. Not sure how to do it.

~ Falls in line with derived product supersedes raw. A good way to say this is new.

Someone is working on this. Knows having LIDs change is frowned upon, but need a way to do it when it does happen. A way that doesn't depend on the registry.

~ EN has OPS keywords for internal system, registry, etc, maybe this should be adjacent. An undocumented feature and nodes can do the editing.

~ Someone agrees.

~ Someone else has a draft of the SCR. This discussion has been helpful.

~ We needs something to say superseded by versions. Need a way to say what supersedes what. An intermediate product to handle that use case.

Question: Is this something we could use product metadata supplemental for?

Answer: Yes, but creating a new version is better. More explicit.

~ Helps with the superseded by issue.

~ Someone would rather create a tool then have a separate table sitting in the archive. We can discuss the details, but it's more explicit then having an extra product around.

When the SCR is written it will need a lot of discussion and revision.

This leaves us with the next steps for each of the topics.

### ## Tech session

The time slot that we would have used for a tech session is being used for web modernization, so if we want to have a discussion it will be more ad-hock. Maybe in the evening.

~ Maybe we could do something informal. In favor of having a discussion.

Question: How many people won't be there?

Answer: Dick won't be there.

~ We'll have a hybrid set up.

~ We need to get our ducks in a row.

~ Let's have this in the back of our minds.

### # Meeting Chat

from Mark Bentley to everyone: 9:43 AM

We tried to be strict with versioning at the start, but since we add new products every day, versions got into hundreds, and it got silly. So now we only increment the collection numbers for "big" changes. So if we need to link to a given collection, we can, but in PSA we would not have this tracking in reality anyway

from Mark Bentley to everyone: 9:49 AM

Agreed with Ed - we haven't found a reason to keep this tracking (tracking from a bundle VID to a collection VID and hence collection inventory and contents)

from Mark Bentley to everyone: 9:49 AM

And for DOIs we would issue for product LIDVIDs and not go through collections - reproducibility is not really at risk here?

from Dick Simpson to everyone: 9:57 AM

Jordan said there was no need to reference context products by LIDVID; why not?

from Lyle Huber to everyone: 10:03 AM

Nothing is \*stopping\* you from referencing context products by LIDVID.

from Jordan Padams (he/him/his) to everyone: 10:03 AM

+1

from Mark Bentley to everyone: 10:07 AM

An amusing upshot of referencing context products by LID was that when one was deprecated (and the name included "DEPRECATED"), one of our regular context product updates resulted in the PSA user interface displaying "DEPRECATED" in a menu ;-)

from Mark Bentley to everyone: 10:09 AM

Sounds good - our tool which wraps validate does this kind of check, it would be great to move it to validate itself

from Jordan Padams (he/him/his) to everyone: 10:09 AM

[https://urldefense.us/v3/https://github.com/NASA-PDS/validate/issues/629;!!PvBDto6Hs4WbVuu7!PFuTcuhiyatH71Zcx8BxFupcQbhBmrDLsEEsOFDcUJEmRdLr90WkDE0DYXy5dMu7xMwslXTrR9M8pvrWDNk7vVs9R2pO3Hp\\$](https://urldefense.us/v3/https://github.com/NASA-PDS/validate/issues/629;!!PvBDto6Hs4WbVuu7!PFuTcuhiyatH71Zcx8BxFupcQbhBmrDLsEEsOFDcUJEmRdLr90WkDE0DYXy5dMu7xMwslXTrR9M8pvrWDNk7vVs9R2pO3Hp$)

from Jordan Padams (he/him/his) to everyone: 10:10 AM

+1

from Mark Bentley to everyone: 10:14 AM

We have this problem with LID changes, but I think it may be tricky to get the instrument teams to build their pipelines to support this - but in principle it's definitely needed

from Trent Hare to everyone: 10:16 AM

I see "supersedes" is great optional attribute. but "superseded\_by" implies updating something...

from Jordan Padams (he/him/his) to everyone: 10:19 AM

superseded\_by would require a new intermediate product to be created to link the 2. e.g. if productY::1.0 supersedes productX::1.0, you would need a productX::2.0 that includes `superseded\_by`

from Jordan Padams (he/him/his) to everyone: 10:19 AM  
for those cases where `productY::1.0` already exists

from Mark Bentley to everyone: 10:26 AM  
I don't think there will be IPDA/PSA representation

## DDWG Notes 2023-05-25

---

title: DDWG Notes 2023-05-25

layout: default

date: 2023-05-25

---

May 25, 2023

Notes by Debra Kazden and Dick Simpson

Known Attendees: M. Drum, E. Guinness, L. Huber, S. Hughes, R. Joyner, D. Kazden, T. Lim, J. Mafi, S. McLaughlin, B. Semenov and R. Simpson  
Known Observers: R. Alanis, P. Lawton, S. Loftin, E. Schaefer, C. Seyb and M. Tiscareno

### ## DDWG Agenda

(Included in email sent May 23, from M. Drum, Agenda for May 25, 2023)

#### Front-matter/What's new

- Anything new with Context Product initiative?
  - We have a morning slot available for F2F Tuesday 7/25
  - SCR Freeze date 8/14/23
- \*\*(Discussed)\*\*

#### Ready for Vote

- None this week

#### Tier 1 issues

- CCB-364 – Disallow bundles referencing collections by LID - Further discussion or clarification before vote?
- CCB-336 – Add a License Information class to the Identification Area – Tanya Lim
- CCB-211 – XML/schema based files as archival data – Jesse Stone, Mike Drum
- CCB-356 – Update Constraints on Author/Editor Lists (and add other metadata) – Ron Joyner
- CCB-324 – Create a data type for Type attributes that require a permissible value list. - Figure out where we want to go with this
- CCB-363 – Changes to the SR and/or DPH to clarify that data objects described in a label do not need to match the order they appear in a file.
- CCB-366 – Schematron/Schema Version Conflict
- CCB-367 – Superseded by / Supersedes reference type **\*\*(Discussed CCB-364)\*\***

#### Tier 2 issues

- CCB-164 - Display Settings not required for images -- (IMG) - Need to think more about potential updates to Imaging and how Display might fit.
  - Product Metadata Errata?
  - Create a supplemental label to live next to original label, that can add aliases and supersede metadata
- \*\*(Not Discussed)\*\*

#### Parking lot – need a driver!

- CCB-326 - Inconsistent Direction on How to Use Discipline Facets
- RS: 259/281 (see new attachment to CCB-281) (on hold) **\*\*(Not Discussed)\*\***



Next meeting: 6/8/23

#DDWG

## Front Matter

The next SCR freeze data is August 14. It's time to start narrowing in on what we want to do for build 14.0.

Laura has let us know that there is a morning slot available on Tuesday at the July F2F. We should probably schedule something, to discuss whatever needs discussing. Not everyone will be able to be there. We can send a poll to see if people can call in or attend to see if it's worth doing.

We started talking about context products about a month ago.

~ Nothing has been done yet.

~ Mostly we need a kick off to the work we'll need to do to modernize the PDS3 to PDS4 context products. It's being discussed at PSI. They are needing to rebuild a set of products. Doing it on the side, but aware of the need for this to happen. Eventually, we will probably need to assign people to work on this.

# Tier 1 issues

We discussed a number of issues with Jordan last time. The most contentious was bundle referencing collection by LID, so we will start there.

## CCB-364 - Bundles referencing by LID See <https://pds-jira.jpl.nasa.gov/browse/CCB-364>

Lots of strong, conflicting opinions.

Question: What is NSSDCA's point of view on this?

Answer: NSSDCA has discussed this internally. They don't expect this to affect typical deliveries at this time, but for Catalina Sky Survey if the bundle version changes each time that will affect them, but they can handle it. Would need to make some changes for Catalina Sky Survey if they start incrementing the minor part of the VID for the bundle. They have been discussing the possibilities with Jesse Stone.

~ Catalina Sky Survey is an ongoing mission with nightly deliveries. It's expected to go on forever. The model can support it, but it's a giant growing pain for NSSDCA to support.

~ The delivery process for Catalina Sky Survey is different from other deliveries. Jesse tweaked the deep archive delivery to get a weekly snapshot. They have been delivering for about a year.

~ Keyed off the LIDVID.

~ Yes. Master key can be searched. NSSDCA assigns IDs when they get deliveries. The short answer is there is no huge impact for them.

Question: Would this be implemented in the fall?

Answer: It would be part of build 14.0, which would be released in December. Don't think PSI would adopt it for Catalina Sky Survey.

(Note taker's power went out. Gone approximately 2 minutes. Meeting was recorded from this point

and Dick Simpson provided notes, which are included at the end of these notes.)

Question: So, creating a collection and each day is a version, so what does NSSDCA get? A new version a day? Is any single collection versioned?

Answer: NSSDCA receives the bundle product with a LIDVID that doesn't change. Data is added on a nightly basis. They get a weekly update. A snapshot on day 7. The bundle product - they get version 7. Never version 1 to 6 of collection products.

Another Question: Do they ever get a version 8? Something that replaces a previous version?

Answer: Haven't received any replacements. Corrections are made further up.

~ Each collection gets a new LIDVID daily with new products. In some cases, if there are corrections, will change the products to 2.0 and include them in the next day's increment.

~ And the latest version of the collection supersedes the previous.

~ If there's ever a case where you prefer a LID reference to a LIDVID this is it.

Question: Someone is interested in use cases. Wants to completely understand use cases. Are we open to discussion on the SCR in general?

Answer: Someone says yes. We brought this up to show the possible hurdles.

~ The issue from Jordan is in essence that we're breaking provenance.

~ Someone disagrees.

~ It seems like more of a blanket. Might want to allow it. Reverse the philosophy. Allow it where necessary. As far as provenance, we can discuss what it means to have an archive. Linking a bundle to a collection at this time, that's good provenance. With LID you no longer have that relationship.

~ Totally wrong. Still linking to the collection. The only change is that you added more stuff. There's no change to science.

~ Someone disagrees.

~ The bundle doesn't change. It has the same collections. But if you add or remove a collection then you have to increment the bundle label. User doesn't care about version because they are just taking products from the collections. Versions are tracked in the inventory. Accumulating collections are the same use case as Catalina Sky Survey.

~ That's the point. There are use cases where LID should be allowed, but currently allowed everywhere.

~ Would be fine, but what Jordan said last week didn't make sense. If updating a collection, we use modification history to explain the changes. A more robust use of modification history might solve some of the provenance issues.

Someone is still unsure how we can move forward before we understand all of the use cases. We can then generalize it and make policies. Concerned about Jordan's concerns. We don't want problems in five years.

~ Someone else would like to hear more about use case provenance for a good archive.

~ Catalina Sky Survey is one. Accumulating is another.

~ All SPICE kernel archives are accumulating. Specifying by LID doesn't work. Have to increment the bundle version and collection by LIDVID.

~ So when there are no changes in accumulating products LID works. That's one use case.

~ Catalina Sky Survey is a type of accumulating bundle.

~ No changes to the collections.

~ Someone wouldn't use those words. There are different types of changes. Products changing or siblings being added.

~ In Catalina Sky Survey, not all collections go in the deep archive.

~ Correct.

- ~ Like they don't exist.
- ~ Right. As collection number 7 is deep archived, 1 to 6 can be deleted.
- ~ In principal, for this use case, there's no difference if you use LID or LIDVID.
- ~ Someone doesn't think that's the correct way to put that.

Someone's concern is with how you enforce anything. If you take any collection and have to access which use case it is there is no accumulating tag. Not sure how you would enforce this other than expecting nodes to do the right thing.

- ~ Good point. We need to be able to validate.
- ~ Someone else thinks we would have to see if it's a new version.

Question: So validate would have to look at previous versions? Seems very complex.

Answer: We deal with the IM. Jordan and staff are paid to make it work. Provenance is a big issue. Getting rumbles from SMD about how trustworthy we are.

- ~ The SCR as written makes it hard to know what we're discussing. Needs to be better written. Also, regarding the lack of guidance - totally disagree. We said we don't want to reference by LIDVID when we were designing this. Using LID so always get the latest version.
- ~ Someone disagrees. Thinks we're getting feedback now that this is an issue. Agrees the SCR should be beefed up a bit.
- ~ Yes, guidance was more broad in the beginning.
- ~ When we discussed this last time, it was pointed out that this is a major non-backwards compatible change. The primary use case Jordan was bringing up was to do with scientific reproducibility. Don't see that as a driver for us. If that's going to be an issue going forward nodes will need to increase their storage.
- ~ True. A huge impact on the budget.
- ~ Someone has a problem with the idea of people expecting previous versions of a bundle with a few clicks.
- ~ This is about loss of information. Losing the relationship.
- ~ Someone agrees. It's not tracked in the model.
- ~ Jordan could track with a table, but than it wouldn't be in the archive.
- ~ The information is there and it is tracked.

Question: You are documenting it?

Answer: Modification history in the collection label.

- ~ Someone is unclear on what information is being lost.
- ~ So, if you have science with a DOI attached to a paper from 15 years ago and you go to the node they give you the latest version of the data, but you don't know when the bundle DOI was generated or what version of the collection you need to reproduce the science for that paper. In the intervening time things like the calibration could have changed.
- ~ You also don't know when they got the data they used.
- ~ That's a different use case. If the data is superseded because of a change to the calibration.
- ~ If you are making a magnetic field model you will get different answers.
- ~ A scientist will dig deeper than a LIDVID.

Another Question: What if you don't know that the data you were given is wrong?

Answer: A scientist would figure it out.

- ~ That's the issue. Science re-usability. We need to justify why we lost the information.
- ~ Someone said you have to show me.
- ~ A lot of information is necessary to figure out when the data was accessed.

~ System needs to say this is the relationship at a point in time. Can lose it with LID instead of LIDVID.  
~ EN is saying they can't do it without LIDVID. They are making that claim. We would need to prove them wrong if they are. Maybe modification history is good enough, but in fifty years when no one is around to answer questions about why things have changed we lose reproducibility.

Someone suggests we use modification history to address this.

~ Important questions to address. LIDVID doesn't explain what happened. You have to explain what changed. Have to capture that somewhere.

~ Fair.

~ We have an SCR for superseded too. Just trying to get the use cases down.

Don't forget that the NSSDCA archive will have the different versions for years. Preserving the data is their mission. Might be able to help.

~ Someone is interested in the method used to retain data. If someone came to NSSDCA looking for Catalina Sky Survey data from 15 years ago, unclear how much information NSSDCA would need.

~ Someone thinks if they had the SIP LIDVID they could pull it. Or a bundle or collection LIDVID.

If the bundle has a LID that's different.

~ If version 1 was requested they would give you that bundle product and all collections. Would need to ask for a specific version.

~ Jordan says if we want to reference by LID then we no longer retain reproducibility if we do this. So, if a bundle LID isn't good enough by itself, NSSDCA isn't going to check modification history. So not as useful as collection LIDVID if looking for a snapshot of data.

Someone has a bundle. The current bundle uses collection LIDVIDs. The older bundle just used LIDs, so now it points to the newest collections. Not sure if that made sense.

~ Yes. Pointing to the current version. If you point to version 12, then a user knows there are 11 previous versions. Bundle LIDVID is not as useful as collection LIDVID for reproducibility.

~ LID points to a set of collections. No information about the time of ingestion. That's the problem. System returns the latest, but the bundle, if using LID, is pointing to the set of collections. Not to a specific collection.

~ Someone is not sure they agree. It points to the latest.

~ That's implementation. A nicety. But the logic is a LID points to a set of collections.

~ Someone agrees that bundle that references LID points to a set of collections, but it can also point to products. Modification history is optional. Has turned in a few with vague comments. If going to try to be specific we have quite a few things to deal with.

~ Collection inventories require LIDVIDs.

~ For primary products, secondary can be by LID.

We are out of time. We are narrowing in on ways to move forward. Will update the ticket and go from there.

Question: NAIF uses LIDVIDs in the bundle, so if they add data, what do they do with the old collection products?

Answer: Accumulating, so update accordingly. Increment versions. That way they can always reconstruct what was there at a given time. Mark previous kernels in collection inventory as secondary.

Another Question: What happens to the old collection products?

Answer: They stay there. NAIF doesn't remove anything.

~ Someone will have to look.

~ Look at MAVEN. It's now at about version 33.

~ SBN also recently created a tool for superseded - to join older and newer versions.

## ## Meeting Chat

(captured by J. Mafi.)

Even if Catalina Sky Survey won't be updating its version of the information model, other data deliveries like it will come along, so we should be sure to put together a system that works for CSS and data deliveries like it.

## ## Dick Simpson's notes

Called to order by Mike Drum at 0930 PDT.

Attendees: Rafael Alanis, Mike Drum, Ed Guinness, Lyle Huber, Steve Hughes, Ron Joyner, Debra Kazden, Pat Lawton, Tanya Lim, Sheri Loftin, Joe Mafi, Stef McLaughlin, Ethan Schaefer, Boris Semenov, Casey Seyb, Dick Simpson, and Matt Tiscareno

Joe said Debra had lost power and requested that the meeting be recorded. In the meantime Dick offered to take notes.

## DDWG Agenda

### Front-matter/What's new

1. Anything new with Context Product initiative?: We started discussion on this topic a month ago; has anything changed? Lyle said he has not done anything more.

Mike said this would mostly be a kickoff for work needed to modernized context products that have been (or need to be) migrated from PDS3 to PDS4. SBN needs to rebuild a set of context products, so this is a driver for them. We may need to assign context product revisions to DDWG members.

2. We have a morning slot available for F2F Tuesday 7/25: We could use this time for a Tech Session; this could be as simple as a face-to-face DDWG; but we could do more. Mike will poll DDWG to see who can attend and who might be able to call in if not physically there. We can discuss goals for IM v2.0.

3. SCR Freeze date 8/14/23: We should narrow what we want for the next build (Build 14.0) and focus on those issues.

### Ready for Vote

1. None this week

### Tier 1 issues

1. CCB-364 (link) – Disallow bundles referencing collections by LID - Further discussion or clarification before vote? Mike expects this to come to a vote that will not be unanimous; there are strongly, conflicting opinions.

Stef said that requiring LIDVID references is not likely to affect typical deliveries to NSSDCA at this time. And data users can always request recovery of old bundles/collections from NSSDCA; of course, those recoveries are limited to what DNs have delivered to NSSDCA and most DNs deliver at a fairly slow cadence.

A possible exception is the Catalina Sky Survey (CSS). Future deliveries to NSSDCA might be changed since the bundle version is currently being held at

v1.0 while collections are accumulating. There are five collections; each can be augmented as often as nightly and the collection version\_id is incremented when that happens (some version\_id values are already as large as 500). New data are delivered to NSSDCA weekly, so as many as six of seven versions of each collection exist only at the DN. There have been no cases where data have been removed or replaced (necessitating a new version), so CSS represents a use case in which there is accumulation only. Mike views this as perhaps the most common and important use case favoring reference by LID.

Steve said CCB-364 would break provenance by requiring that all references be by LIDVID. He offered a compromise in which LIDVID references would be strongly recommended but not required — this would allow exact reconstruction of bundles in most cases. Ed said that adding more products to a collection does not change the science value of the bundle/collection, and Steve agreed. Ed suggested that more robust use of Modification\_History could document situations in which data have been removed or changed. However, Modification\_History is not easily integrated into reconstruction procedures that are machine based). If we rely on Modification\_History, Dick said we need to make sure that it is required and that it contains the necessary information for reconstruction.

Boris said NAIF bundles only accumulate. Some kernels are superseded; collection version\_id is incremented but the old versions are not removed. Joe's concern is how to enforce selection of use case — that is, we don't have an attribute that says a bundle/collection is accumulating, so how will we know whether members have been modified or removed? Mike suggested comparing collection Inventories to distinguish cases in which there are only additions from cases where products have been replaced or removed.

Lyle noted that the SCR is very vague about what is being requested; it is difficult to discuss CCB-364 without having specific change requests. He also said there has been guidance on whether to reference by LID or LIDVID, even though that guidance may seem insufficient today. The guidance was that users would receive the most recent version when reference was by LID. Steve said this tells the user what is happening but doesn't necessarily allow reconstruction of a previous bundle, and that strict 're-usability' is what SMD and others are seeking.

Lyle countered that reconstructing older versions of bundles/collections has not been a major driver at ATMOS; if instant re-usability is required (that is, exact reconstruction and real-time delivery of a previous bundle), ATMOS would have to increase its storage by two orders of magnitude and GEO and CIS will probably have to expand by more. ATMOS can reconstruct earlier versions; but data users have to ask and then wait until staff can do the necessary work. The information to reconstruct exists, but the older versions can't be downloaded with just a few clicks.

Mike said we should update the ticket so that it is more specific about what is being requested.

DDWG Notes 2023-06-08

---

title: DDWG Notes 2023-06-08

layout: default

date: 2023-06-08

---

June 08, 2023

Notes by Debra Kazden

Known Attendees: M. Drum, E. Guinness, T. Hare, L. Huber, S. Hughes, R. Joyner, D. Kahan, D. Kazden, T. Lim, J. Mafi, McLaughlin and R. Simpson  
Known Observers: B. Hirsch, S. Loftin and M. Tiscareno

### ## DDWG Agenda

(Included in email sent June 7, from M. Drum, Agenda for June 8, 2023)

#### Front-matter/What's new

- An issue came up with Validate not knowing how to treat bit patterns in Special Constants. We should create a ticket to handle this situation (more info at

[https://urldefense.us/v3/https://github.com/NASA-PDS/validate/issues/611\\_!!PvBDto6Hs4WbVuu7!P5HVTqmfHbADMWKj4wNGQffeGBtOMoicEUdUIBDyvZ4TjDJCNDkhgLAeD9oRZMESKyphcFsAu9mvtDnEyE00WwwFL9uk2EB\\$<https://urldefense.us/v3/https://github.com/NASA-PDS/validate/issues/611\\_!!PvBDto6Hs4WbVuu7!P5HVTqmfHbADMWKj4wNGQffeGBtOMoicEUdUIBDyvZ4TjDJCNDkhgLAeD9oRZMESKyphcFsAu9mvtDnEyE00WwwFL9uk2EB\\$>](https://urldefense.us/v3/https://github.com/NASA-PDS/validate/issues/611_!!PvBDto6Hs4WbVuu7!P5HVTqmfHbADMWKj4wNGQffeGBtOMoicEUdUIBDyvZ4TjDJCNDkhgLAeD9oRZMESKyphcFsAu9mvtDnEyE00WwwFL9uk2EB$<https://urldefense.us/v3/https://github.com/NASA-PDS/validate/issues/611_!!PvBDto6Hs4WbVuu7!P5HVTqmfHbADMWKj4wNGQffeGBtOMoicEUdUIBDyvZ4TjDJCNDkhgLAeD9oRZMESKyphcFsAu9mvtDnEyE00WwwFL9uk2EB$>) )

- SCR Freeze date 8/14/23

\*\*(Discussed)\*\*

#### Ready for Vote

- None this week

#### Tier 1 issues

- CCB-356 – Update Constraints on Author/Editor Lists (and add other metadata) – Ron Joyner - See latest updates. Ready for vote?

- CCB-364 – Disallow bundles referencing collections by LID - Updates made to ticket

- CCB-363 – Changes to the SR and/or DPH to clarify that data objects described in a label do not need to match the order they appear in a file. - Let's make a decision

- CCB-336 – Add a License Information class to the Identification Area – Tanya Lim

- CCB-211 – XML/schema based files as archival data – Jesse Stone, Mike Drum

- CCB-324 – Create a data type for Type attributes that require a permissible value list. - Figure out where we want to go with this

- CCB-366 – Schematron/Schema Version Conflict

- CCB-367 – Superseded by / Supersedes reference type **\*\*(Some Discussion)\*\***

Tier 2 issues

- CCB-164 - Display Settings not required for images -- (IMG) - Need to think more about potential updates to Imaging and how Display might fit.
  - Product Metadata Errata?
  - Create a supplemental label to live next to original label, that can add aliases and supersede metadata
- \*\*(Not Discussed)\*\***

Parking lot – need a driver!

- CCB-326 - Inconsistent Direction on How to Use Discipline Facets
  - RS: 259/281 (see new attachment to CCB-281) (on hold)
- \*\*(Not Discussed)\*\***

Next meeting: Skipping a week for ACM, so 7/6/23

# DDWG

## Front Matter

A quick note. It's at the bottom of the agenda. Because of the meeting in Flagstaff, our next meeting will be July 6. Then we will have a F2F in July at the MC.

The new thing this week is a discussion between Mark Bentley and Jordan about validate not knowing how to treat bit patterns in special constants. There are some missing, all are ASCII. Should match the type. They want us to create a ticket.

Question: Anyone with special knowledge of these issues?

Answer: IMG has dealt with this because of VICKAR and ISIS. It can be painful.

~ This came up because validate is treating it as a string. Problem in the programming area. We should provide some guidance.

Someone said that in GitHub all the early stuff seems to be about valid min and max. Thinks in the wrong field. Special constants with valid min and max shouldn't be flagged, it just tells the value range. Seems like people wanted validate to point to values outside the range. That was never how it was supposed to work.

~ Someone else thinks that was missing. Thinks Jordan expected that. If there are no constraints on that than it reduces the urgency, but it seems implied. Should match the class.

~ Not sure about that.

~ Says it should be the same data type as the elements in the object or field.

~ But ASCII short string collapsed. Can represent everything with that. Doesn't work with binary data.



- ~ Someone thought this came up because of floating point numbers.
- ~ Very hard to represent.
- ~ Unsure how you deal with this with byte swapping.

Question: Has anyone else dealt with this?

Answer: SBN.

- ~ Holdover from PDS3. Binary was problematic for special constants too. Worded as best as possible because it says 'should'. Could redesign to make more sense.
- ~ Someone says there will be a lot of problems if we give it the same name.
- ~ Agree. Not necessarily an error, just a range. A recommended set of values.
- ~ Don't have to flag it. Special constant is the limit, not illegal. As user, wouldn't want to include it in your calculations. Not necessarily an error. Just not a legitimate data value.

Someone is up for redesigning the whole thing to be more functional.

- ~ That would be very useful.
- ~ We will keep this on the agenda and create a ticket.

# Tier 1

## CCB-356 - Update Constraints on Author/Editor Lists

See <https://pds-jira.jpl.nasa.gov/browse/CCB-356> <<https://pds-jira.jpl.nasa.gov/browse/CCB-356&nbsp;&nbsp; &nbsp;  &nbsp;  &nbsp;   >>

The ticket has been updated. It's on the agenda to see if there are any more issues before we vote.

Someone says it seems highly non-backwards compatible and can't figure out what's being said about what it would actually look like.

Question: If embedded in a label?

Answer: Yes. Too many comments.

- ~ No where near ready for a vote. Needs to be fleshed out before we can really discuss it.
- ~ Reporter wanted to discuss it before fleshing it out.
- ~ A lot of work just to get citation into a label. There's value in that, but not sure we really want to do that to our DPs.
- ~ Someone sees this as imperative. Probably not as complicated as it looks. The value we get is much higher than the burden.
- ~ It's hard for a human to read. Might be better for DP.
- ~ It's impossible to parse without all this.

Question from reporter: What would people be interested in seeing?

Answer: Examples. It's a lot of overhead.

~ Anywhere we have author or editor list is being replaced with this.

~ A lot of possibilities.

~ Example includes everything, even what is optional.

Question: Seeing person is a class. So, for every author we would have to fill out all this stuff?

Answer: In authors list we delineate with a semi-colon.

~ We used that to be compatible with Astro database.

~ This author class can be used with any syntax, in any number of formats.

~ Someone asked if this solves the concern about being able to put in organizations, and said they are not sure it does.

~ Can define person, organization or both.

~ Author list is and/or.

~ Hard to understand.

~ This brings in orcidIDs. We need to include that in whatever implementation we do.

~ Was waiting to flesh out with the definitions until we discussed the structure.

Someone would like some sense of the backwards incompatibility of this.

~ It's not backwards compatible, but you could run a script, put another attribute, still include author/editor list and it would still be backwards compatible.

~ This would have to be optional.

~ If this is optional, people would just do that.

Question: Where is this going?

Answer: Anywhere you see author or editor list.

Another Question: Should this only be in the DOI?

Answer: Not all citations have DOIs. Not using this to describe bundles, for things like literature references.

~ Thought citation information.

~ Yes, citation info.

~ Sounds like not everyone is on the same page.

~ Someone thought this supplants anywhere with author/editor lists. Fairly obvious stuff.

Question: Where did people think this was going?

Answer: Anywhere there's a literature reference. Thought that was the original problem statement, and I wrote it.

~ Someone didn't understand it that way. This is their implementation of how they interpreted the concern. Sees the benefit to doing it. Problematic to stay with author lists with values delineated with semi-colons. Need a way to define contributors.

~ Someone else agrees. This is not meant to duplicate functionality of a DOI. Datacite is fully capable of doing that. This is bare bones, but makes it useful. Semi-colon separated list of names is hard to work with. Shouldn't be relying on the DOI standard for our standard.

~ All the information needs to be collected from the DPs for the DOI. Having it in the IM collects it together so we can get the DOI. It's a manual process.

Question: Is ADS still doing that?

Answer: It was rational at the time. Unclear how they do it now.

Another Question: Does someone want to check up on what ADS is doing now? Seems like this is a good thing to do, even if it's not backwards compatible. Shouldn't be a huge number of products.

~ ADS is still being pretty basic with their author list. Can check another one.

~ Bibtec. Have to have the metadata somewhere, but couldn't offer a separate link for each author without more information.

Question: Is there an ADS format on the export list?

Answer: Yes. Using semi-colons.

~ No who belongs to what.

~ Not clear.

~ Could check with ADS to see why they have this format. Maybe there's a reason.

~ Someone is not convinced we have to do it their way just because.

~ True, but it might be relevant to us as we make a decision.

~ Could be annoying.

~ OrclIDs for everyone could be very complicated.

~ The structure looks like what would be in a context product. Unclear we would want more context products.

~ Good lord.

~ Could rely on orclIDs for author/editor information.

~ That's the right track if we can support it with tools. Would be the right way to do it. Maybe someday.

Next steps. If there are any fundamental complaints with the structure put them in a comment.

**\*\* (Action Item - Everyone) \*\*** The ticket will be updated to include a justification for the backwards compatibility break.

~ Can throw in another semi-colon list to keep it backwards compatible.

~ Should break the backwards compatibility. Need people to store the metadata we need.

Question: Why is the editor class different from the author class?

Answer: Fonts in jira.

~ The same information is in both.

Another Question: Organization can be an unofficial collaboration. Is that allowed? Like if we write a paper with the Cassini Mag team?

Answer: Interesting use case.

Someone wants to go back to if this is for citation information or more general.

~ It could be it's own provenance class. Would deprecate citation information class. That would solve the backwards compatibility issue.

~ Someone else thought we were replacing author/editor list with something more functional where ever they exist. Not sure primary use is always for the DOI.

~ IM specification shows author list in software, citation information and document. Presumably editor

list too.

~ In that case, maybe it would just replace.

~ It's an attribute. We're talking about turning it into a class.

~ The solution is beyond the scope of the problem. Was originally looking for a way to include organization.

~ Someone disagrees. Anne wanted structure.

~ Couldn't augment the old one to include all the information. Could de-scope, but think this is the right way forward.

~ Looked at Anne's wiki when designing this. Included contributor. Could throw it out.

People should put comments in jira. Otherwise, we need to define organization, author, editor, orcid, etc. Maybe can attach an XML example too.

~ That can be done (\*\*Action Item - Ron\*\*)

Moving on.

## CCB-364 - Bundles referencing by LID

See <https://pds-jira.jpl.nasa.gov/browse/CCB-364> <<https://pds-jira.jpl.nasa.gov/browse/CCB-364&nbsp;&nbsp; >>

Some updates were made, but it's still mostly a problem statement.

Question: Anyone have any new thoughts?

Answer: NSSDCA just updated. Their concerns are pretty limited. The big hang ups were to define the use cases.

~ Someone had a thought. In the discussion last time, about users looking at a publication and wanting the scientific data that was used and the DOI - It was said that the LID points to a set of collections. If a single DOI and multiple versions of bundle, not sure how a user goes from a DOI to a specific version. Should be considered as one of the use cases.

~ One suggestion to solve this is to include access date with citations so can trace back to the state of a particular bundle or collection at that time.

~ Unless using it along modification history description, unsure if the granularity would help.

~ Nodes do it. LIDVIDs definitively give you a list of products.

~ Someone said they tried and failed to explain why they think this is so important. Can't get to the right collections/bundle with just a DOI if no LIDVID. Can't guarantee the snapshot of a given time.

~ So we would be expected to keep every version of every collection online in perpetuity.

~ Someone doesn't think landing pages work for that. Did look back at NAIF's example, but they are using such an old IM that there are no DOIs. Not sure it's a good example for us. Still work to do here.

Someone suggested that we could issue a new DOI for every new version of bundle. Build that into DOI metadata.

~ Right, but if DOI only points to collection by LID then it still will point to the same collections. Sounds

like the biggest concerns are accumulating data sets and not having all the data online at all times.  
~ Someone has been here 30 years. Only has to look for old versions once a decade. Solving a problem that isn't a problem.

~ LIDVIDs make it easier, even if only necessary once a decade. Will work on enumerating use cases.

Next meeting July 6.

DDWG Notes 2023-07-06

---

title: DDWG Notes 2023-07-06

layout: default

date: 2023-07-06

---

July 06, 2023

Notes by Debra Kazden

Known Attendees: M. Drum, M. Gordon (for M. Tiscareno) E. Guinness, T. Hare, L. Huber, R. Joyner, D. Kahan, D. Kazden, J. Mafi, McLaughlin and R. Simpson  
Known Observers: A. Bailey, P. Lawton, S. Loftin and M. Mace

## DDWG Agenda

(Included in email sent July 5, from M. Drum, Agenda for July 6, 2023)

Front-matter/What's new

- New issue: CCB-369 Add has top plane and has side plane to id reference type
  - An issue came up with Validate not knowing how to treat bit patterns in Special Constants. We should create a ticket to handle this situation (more info at [https://urldefense.us/v3/https://github.com/NASA-PDS/validate/issues/611;!!PvBDto6Hs4WbVuu7!M2R3x7oSuP9LeWYulIZogRrUkaarF6nEkVJgKlokGH\\_zVUGFL3mhTxNIGgHeY3PcjAFiXkFr7VbQMfvDzKtUvjPDtbkWSFRx\\$](https://urldefense.us/v3/https://github.com/NASA-PDS/validate/issues/611;!!PvBDto6Hs4WbVuu7!M2R3x7oSuP9LeWYulIZogRrUkaarF6nEkVJgKlokGH_zVUGFL3mhTxNIGgHeY3PcjAFiXkFr7VbQMfvDzKtUvjPDtbkWSFRx$))
  - SCR Freeze date 8/14/23
- \*\*(Discussed)\*\***

Ready for Vote

- None this week

Tier 1 issues

- CCB-356 – Update Constraints on Author/Editor Lists (and add other metadata) – Ron Joyner - Further revisions made, comments from Dick.
- CCB-364 – Disallow bundles referencing collections by LID - Still needs some work enumerating use cases
- CCB-363 – Changes to the SR and/or DPH to clarify that data objects described in a label do not need to match the order they appear in a file. - Let's make a decision
- CCB-336 – Add a License Information class to the Identification Area – Tanya Lim
- CCB-211 – XML/schema based files as archival data – Jesse Stone, Mike Drum
- CCB-324 – Create a data type for Type attributes that require a permissible value list. - Figure out where we want to go with this
- CCB-366 – Schematron/Schema Version Conflict
- CCB-367 – Superseded by / Supersedes reference type **\*\*(Some Discussions)\*\***

Tier 2 issues

- CCB-164 - Display Settings not required for images -- (IMG) - Need to think more about potential updates to Imaging and how Display might fit.
- Product Metadata Errata?

Create a supplemental label to live next to original label, that can add aliases and supersede metadata

\*\*(Not Discussed)\*\*

Parking lot – need a driver!

- CCB-326 - Inconsistent Direction on How to Use Discipline Facets

- RS: 259/281 (see new attachment to CCB-281) (on hold) \*\*(Not Discussed)\*\*

Next meeting: 7/20/23

# DDWG

##Front Matter/What's New

JPL is going to teams, so WebEx might go away.

CCB-369 (Add has topline and has sideplane to id reference type, see <https://pds-jira.jpl.nasa.gov/browse/CCB-369>) is a new issue from Tanya. Wanted to get it on people's radar. Will get in onto the agenda.

Question: Has anyone looked at this yet?

Answer: Someone thought this is outside the scope of anything for PDS4. All rejected for PDS3.

Surprised to see this. It might not get far.

~ Someone assumes they would be stored separately and have software to put back together, like ISIS.

~ Even ISIS dropped the side and back planes because it's too convoluted.

Question: It probably is too convoluted. Is this group even willing to think about this?

Answer: (Silence)

~ We should put a comment in jira. Maybe the scope or history are unknown.

New addition to the agenda - a new issue from NSSDCA. Xman discussion.

~ NSSDCA provides software to generate a PDS manifest at the nodes. It's called xman. Now some nodes are storing data in the cloud and NSSDCA wants to enhance xman. There are a few options. Could enhance to generate a manifest from data stored locally or in the cloud or could have two software options - one for the current local file system and a new web interface for data in the AWS cloud. Wants opinions, but thinks option one is preferred. Will send a summary of this to the nodes. If people have opinions, NSSDCA would like to hear.

IMG is deep in the cloud now, so cloud support is critical. There are tools, like Python for S3 that can probably help with the transition for xman. Another good one is Our Clone. It does every kind of copy and wraps. Maybe could make xman our clone. Our clone is becoming critical for IMG.

~ People weren't aware of that.

~ All open source. Supports so much. Many uses. Also use Python routine Boto3 to wrap.

~ Stef will pass that on to the programmers at NSSDCA. They may already know. So, sounds like USGS has no preference regarding if they keep current xman if can do the wrapping.

~ No. Being able to copy from S3 to S3 would be amazing.

Question: Any other opinions?

Answer: Someone would like the information sent out. The person who handles this is not here.

**\*\*Action Item - Stef\*\*** will send the information and options out to the nodes. People can email her if they have opinions.

Another issue is validate and special constants. Think we need a proposal.

Question: Did anyone volunteer to take this one?

Answer: Someone doesn't remember where we left it, but had some concerns for binary. Went back and forth with EN folks over byte swapping issue. If we always view special constants as observed value, then binary numbers can be hexadecimal strings. Think EN is going ahead with that. Don't think the SR needs to be changed. We need the DPH to say if you have binary data or bit strings it needs to be expressed in zero X notation. Don't think we need to do much more.

Another Question: Just update the documents to say the values?

Answer: Could give examples. Special constants is a class. Validate will just look to see it's formatted correctly. Doesn't really do anything if it exceeds the max value. User will have to figure out what to do if over/under the min or max. Special constants just tells you the range.

~ Someone agrees. What they loath about special constants is that there are so many of them that all sound the same.

~ Someone else agrees. A lot that seem the same.

Question: Are there any other types we need to worry about besides binary?

Answer: Think this came up because of bit strings and binary.

~ There's still the NAN issue.

~ Recognized by IEEE754. Had an SCR to deprecate NAN and INF, but no one was interested. There's a dev version of validate that handles NANs now.

Another Question: So, NANs are valid now?

Answer: Not flagged as errors.

~ Can't describe in special constants area, but recognized by the IEEE standard.

~ Probably could put them in special constants.

Sounds like they are happy. Just need to kick off the document writing.

**\*\*Action Item - Ron and Dick\*\*** will work on examples. **\*\*Trent\*\*** will help too.

# Tier 1 Issues

## CCB-356 - Update Constraints on Author/Editor Lists See <https://pds-jira.jpl.nasa.gov/browse/CCB-356>

We had a lively discussion about this last time. Jira was updated after. Now is the time to air grievances and then we will schedule a vote.

~ Reporter wanted consensus before making changes to the schema and schematron so that we can vote yes.

This was brought up at the IPDA meeting. No complaints. Clarification good. Can move forward.

Needs a TA.

~ It's not ready.

~ Then we are not voting. Looks awfully non-backwards compatible.

~ That was addressed by author/editor list still being there.



- ~ Need to edit the text.
- ~ The question is if we want author/editor list to be a class and attribute.
- ~ That gives someone the heebies. Would like to brainstorm a different term for the class.
- ~ Then it's no longer backwards compatible. We can constrain with a choice statement.

Question: Is that backwards compatible?

Answer: As close as we can get.

- ~ Choice gives us that.
  - ~ Yes, but not promoting it. There was a lot of push back.
- Another Question: Can you do a choice between a class and attribute?
- Answer: Think so.
- ~ Someone is not sure it's worth making it backwards compatible.
  - ~ Currently, author list is an attribute in the IM, not a class.

Question: Would creator class replace citation information?

Answer: No. It would replace author list.

Another Question: So, author list would go away?

Answer: Yes. It would vanish.

- ~ Not exactly backwards compatible, but close.
- ~ Yes, but would prefer to remove author list and force people to use creator class.
- ~ Could deprecate citation information and use creator class.
- ~ The point is there's information in citation information that isn't in the creator list, like DOI, that we need to capture.
- ~ If we allow citation information to allow author list or creator list, think it's better to encourage people to use creator list.

Question: Are there other concerns that people want addressed?

Answer: We will be getting together in person.

Reporter wants to know if they are removing author list.

- ~ Someone likes the idea. Could have a choice between author list element or creator list class.
- ~ Someone else thinks we'll get flack if this is not backwards compatible.
- ~ It's not backwards compatible either way.
- ~ Another person won't complain anymore. We can vote for this as-is.

Reporter still wants to know if they should remove or retain author list in creator list.

- ~ Someone says take it out. Either author list attribute or creator list. Can't have both.
- ~ So it's a choice, and it's backwards compatible. That makes someone happy.
- ~ Reporter will consult with Steve on the choice statement. It's complicated.
- ~ Might require a schema rule.

Question: Anything else?

Answer: That class as written is optional, so there's nothing to prevent having author/editor list, so completely backwards compatible.

Another Question: Is that the intention?

Answer: This makes the DOI easier to capture.

Another Question: Is it valid if you don't have any of the three?

Answer: All optional. Author list, editor list, creator class. But we could require creator class.

- ~ Someone is in favor of requiring.
- ~ Someone said that requiring it has the potential to be bad, but they like it.

If creator list is required, with author/editor list zero to many, would need to write something.

- ~ Required for other metadata standards, but tough for derived.
- ~ Someone is always surprised by how many requests he gets for changes to authors in DOIs.
- ~ Changes to authors is political. Sometimes it's easier to leave it out instead of adding 25 people.
- ~ It's a one bit change. Could be a lien.
- ~ Trivial.
- ~ We have a way forward.

While we're thinking about this, regarding special constants and how many. They are from PDS3 where we had definitions for them all. Probably should put definitions for them in PDS4. Probably in Anne's wiki already.

- ~ A lot of the definitions are it's not a good number.
- ~ The values N/A, unk, and null were for metadata, not data. They had specific meanings, but they weren't for data, so think having them in special constants isn't necessary.
- ~ Good point. Similar to how folks want to use nil.
- ~ So, a no.
- ~ It's up to Ron and others to clarify the definitions. People are happy to help.
- ~ We should take a whack at it.
- ~ Backwards compatibility issue is they've been used.
- ~ Someone is frustrated because people use different definitions.
- ~ Leaving ambiguity to keep it backwards compatible doesn't serve the long term goals of PDS. Can put qualifying language into the definitions.
- ~ Push for best practice, which isn't a great PDS thing.
- ~ It will be at the discretion of the work group.
- ~ Definitions of attributes and classes can be changed in new versions.
- ~ Someone didn't think we were good at incrementing versions.
- ~ Good reminder.

## CCB-364 - Bundles referencing by LID See <https://pds-jira.jpl.nasa.gov/browse/CCB-364>

We are waiting on Steve to work out use cases we want to enumerate. Will push to get this moving forward. Only a small handful of meetings left.

## CCB-336 - Add a License Information class to the Identification Area See <https://pds-jira.jpl.nasa.gov/browse/CCB-336>

This just got a new comment. Believe that feelings were pretty mixed on this. Unclear if we should just vote or if people want to hash this out.

- ~ The Problem Statement was updated this morning. It was ambiguous. Wants to see objects ordered by offsets in the same order in the label.
- ~ Someone is still against this. When we have RGB layout it hits bands in different order, could hose the ability to use it.
- ~ Should add that as a comment to jira.
- ~ We need to make sure we know what problems it would cause. I'm against it. Heavily non-backwards compatible. Enforces linear order to things that might not be linear.

Anyone with opinions, please get them into the ticket. Will leave this on the agenda for next time and then schedule for a vote.

Meet in two weeks and then we will have a session at the F2F. We will decide on topics to tackle.

## DDWG Notes 2023-07-20

---

title: DDWG Notes 2023-07-20

layout: default

date: 2023-07-20

---

July 20, 2023

Notes by Debra Kazden

Known Attendees: M. Drum, E. Guinness, T. Hare, L. Huber, S. Hughes, R. Joyner, D. Kazden, J. Mafi, S. McLaughlin and R. Simpson  
Known Observers: A. Bailey, P. Lawton, S. Loftin, M. Mace and E. Schaefer

### ## DDWG Agenda

(Included in email sent July 18, from M. Drum, Agenda for July 20, 2023)

#### Front-matter/What's new

- New Schematron rules for Product External

a) Rules provided by Ron

b) Do we wish to review this as DDWG or is this just an implementation detail?

- F2F Open Session next week! Preliminary Agenda is a discussion around these two items:

a) CCB-364 – Disallow bundles referencing collections by LID - Still needs some work enumerating use cases

b) CCB-363 – Changes to the SR and/or DPH to clarify that data objects described in a label do not need to match the order they appear in a file

- SCR Freeze date 8/14/23

\*\* (Discussed) \*\*

#### Ready for Vote

- None this week

#### Tier 1 issues

- CCB-356 – Update Constraints on Author/Editor Lists (and add other metadata) – Ron Joyner - Further revisions made, comments from Dick.

- CCB-364 – Disallow bundles referencing collections by LID - On agenda for F2F

- CCB-363 – Changes to the SR and/or DPH to clarify that data objects described in a label do not need to match the order they appear in a file. - On agenda for F2F

- CCB-336 – Add a License Information class to the Identification Area – Tanya Lim

- CCB-211 – XML/schema based files as archival data – Jesse Stone, Mike Drum

- CCB-324 – Create a data type for Type attributes that require a permissible value list. - Figure out where we want to go with this

- CCB-366 – Schematron/Schema Version Conflict

- CCB-367 – Superseded by / Supersedes reference type \*\* (Some Discussion) \*\*

#### Tier 2 issues

- CCB-164 - Display Settings not required for images -- (IMG) - Need to think more about potential updates to Imaging and how Display might fit.

- Product Metadata Errata?

a) Create a supplemental label to live next to original label, that can add aliases and supersede metadata

**\*\*(Not Discussed)\*\***

Parking lot – need a driver!

- CCB-326 - Inconsistent Direction on How to Use Discipline Facets

- RS: 259/281 (see new attachment to CCB-281) (on hold) **\*\*(Not Discussed)\*\***

Next meeting: 7/25/23 for MC F2F

Normal meeting 8/3/23

# DDWG

## Front Matter

We'll start with something new and exciting. Schematron rules for product external. Hopefully, everyone saw this. From Ron. This is to support product external. It's something we might or might not review.

We need to decide if we want to review this or if it's an implementation detail.

~ Someone says it's an implementation detail. No real changes to the model. It just improves the schematron rules.

~ Another someone agrees. Just thought people might want to discuss this because product external is so new.

~ This is a ten thousand foot look level. Wanted it to validate, so it needed changes to the schematron rule since LIDs won't conform to usual rules because of the number of things separated by colons, segments. Modified. Still testing. Need to start urn: and then it counts segments.

~ Sounds good. If anyone has concerns they should connect with Ron directly.

Next topic is that the F2F is next Tuesday. There is a link from Laura for information. We will discuss CCBs 364 (Bundles referencing by LID, see <https://pds-jira.jpl.nasa.gov/browse/CCB-364>) and 363 (Clarification on ordering of data objects in label, see <https://pds-jira.jpl.nasa.gov/browse/CCB-363>) and anything else anyone wants to suggest. For CCB-364, we need examples. And we need to decide about CCB-363.

~ Someone will be traveling Tuesday morning. Hopes to arrive in time for at least part of the session. Other node reps will attend remotely.

~ Regarding CCB-363, someone isn't against it, but if it's not a non-backwards compatible change it should go through smoothly. Worried about breaking backwards compatibility. Don't want to interfere with people who might have used it.

~ Someone agrees. As long as it doesn't apply retroactively. It would affect Mars 2020.

Question: When we worry about backwards compatibility, generally, are we worried about when people have to change IM or validation breaking?

Answer: Worried about people keeping up. Validate should keep up.

~ Someone else is worried about user tools keeping up.

~ At one node they usually don't change IM for new deliveries.

~ It's probably rare, but concerned about people with home grown software. Not sure if it's a big deal.

~ In this case it would just be label creating software. Getting off track. We can discuss this next week. If anyone won't be attending let Mike know.

~ Someone really wants to hear why this is a problem. Wonders if we are being engineers.

Tier 1 items might not make the freeze data. Hopefully, we can do something with CCB-336 (Add a

License Information class to the Identification Area, see <https://pds-jira.jpl.nasa.gov/browse/CCB-336>. CCB-211 (XML/schema based files as archival data, see <https://pds-jira.jpl.nasa.gov/browse/CCB-211>) will probably drop off. CCB-366 (Schematron/Schema Version Conflict, see <https://pds-jira.jpl.nasa.gov/browse/CCB-366>) and CCB-367 (Provenance LDD Using PROV-DM, see <https://pds-jira.jpl.nasa.gov/browse/CCB-367>) are from EN. They want us to address those.

~ CCB-367 is probably closely related to CCB-364. If we pass that then we need to do something with CCB-367. Tracking products that get superseded is becoming a real issue for some nodes.

~ This gets into a philosophical discussion very quickly.

~ Concerned with how we track it. If there's a way to link it it would be very helpful. Unsure how we would do it.

~ On CCB-367, the reason it is there is because of LIDs being changed. We need to link the new one to the previous one. It's much broader now than was when originally submitted. We should get the discussion going.

~ Good for derived. If we're smart about this it could solve that.

~ A lot to cover. Provenance.

Question: Can we discuss this on Tuesday?

Answer: Yes. We will include this.

~ Someone wants a general solution.

~ Another person agrees. The sooner we start discussing this the better.

Question to Steve: Can you edit jira so we have a place to start?

Answer: Yes. **\*(Action Item - Steve)\*\*** ~ We will discuss this Tuesday.

# Tier 1 issues

## CCB-356 - Update Constraints on Author/Editor Lists See <https://pds-jira.jpl.nasa.gov/browse/CCB-356>

There has been some work on this since last time. Definitions were attached.

~ Documents were redone.

~ Yes, edited versions. No real show stoppers. Was concerned with the hierarchy. Concerned about revisiting authors while adding contributors. Text needs clarity. Want people to prioritize the listings. Also need to clarify family name and given name and suffixes, and what to do if two given names. Need to do some clean up. It's a bit complicated, but workable. Might be able to vote before the freeze. Still queasy on this, but think we can meet the August deadline.

~ Someone pulled up the Datacite schema. It doesn't address that.

~ Anne had said that she's quoting them, so her wiki is probably not our best resource.

~ Something we can try to address. Want us to be able to vote at the next telecon in two weeks.

Question: What are our blockers? Concerns about author names, anything else?

Answer: The Requested Changes are not filled out in a meaningful way. Would like it to say where it will be used. Haven't looked in real detail. Unsure how it will apply in documents. Need to spend some time on this.

~ Someone assumes the Requested Changes will have new attributes and say what changes to make.

Regarding Datacite, that's where we are going. It makes sense to conform to their standard. Think it's okay to be more strict, but we need to use Datacite as the general case or guideline. Regarding orcidIDs, etc, if we are going to use those unique identifiers, which is a great opportunity, will need to start

validating them. It affects the system side, which is fine. Benefits us to use those. It's where the world is going.

~ We do have a responsibility to have information in our labels. Ideal to have context products for every person, but too much overhead.

~ No need to re-invent the wheel.

~ Sounds like we need to compare schema to capture as much as we need for when we are creating DOIs.

~ Should make clear in DPH that we are using a more global standard.

~ If add attributes for suffix we will need a way to fit that in.

Question: Has anyone done that while creating DOIs?

Answer: Someone has, but doesn't remember how. Will get the information to Mike **\*\***(Action Item-Pat)**\*\*** ~ Hopefully, we can vote in two weeks and get this out the door as our big task for this build.

## CCB-211 - XML/schema based files as archival data See <https://pds-jira.jpl.nasa.gov/browse/CCB-211>

This is getting punted.

## CCB-324 - Create a data type for Type attributes that require a permissible value list.  
See <https://pds-jira.jpl.nasa.gov/browse/CCB-324>

This is assigned to Emily Law. Haven't thought about this in a long time.

Question: Any comments?

Answer: Important. We did discuss getting rid of the rule that required an enumerated list for type attribute.

~ Providing a data type was to clarify, but we then got into issue of backwards compatibility. Something we should have done.

Question: This is pretty fleshed out. Was there push-back? Can we vote?

Answer: Someone thinks there was push-back.

~ The concern was a flag. Overlap of purpose. What value would be used to validate.

~ Sounds like this would not be backwards compatible.

~ Think the position was that it would be required and could have a warning. Could change current error to a warning.

~ So, if not using this data type, you should. Probably would be backwards compatible. Will raise this on the list. Sounds like we are close. Everyone should review it. Hopefully we can get it passed by the end of this cycle.

----

That's about it. Wrapping up early today. Most of us will see each other next week.

DDWG Notes 2023-07-25 (F2F)

---

title: DDWG Notes 2023-07-25

layout: default

date: 2023-07-25

---

July 25, 2023

Notes by Debra Kazden

\*There are also notes by Laura Givens.

# DDWG F2F

## Topics for Discussion

# CCB-364 – Bundles referencing by LID See <https://pds-jira.jpl.nasa.gov/browse/CCB-364>

# CCB- 367 – Provenance LDD Using PROV-DM See <https://pds-jira.jpl.nasa.gov/browse/CCB-367>

# CCB- 363 - Clarification on ordering of data objects in label See <https://pds-jira.jpl.nasa.gov/browse/CCB-363>

CCB-364 has been discussed. Need valid use cases for referencing by LID. Need to enumerate these cases. Asked people to bring examples.

Question: Why is it necessary to change from what's there?

Answer: Lots of value in encouraging LIDVID values given registry inner workings and tying the data together. No version will cause problems.

~ All of one nodes context references are LIDs only so they can grab the latest version of a context product. They do this always. Finds the explanation not compelling.

~ The case of context objects is different from observational data.

~ Someone disagrees. No difference for how the registry has to work.

~ Yes, but context products don't describe data that changes, data might. Unlikely to be a bug in a context product.

~ Someone said LOL.

~ Not going to need to know the previous. Cases where we need to reproduce a bundle with exact versions of collections are more reasonable and frequent than where we would need to reproduce context products.

~ Someone said prove it. Only had to do it once a decade.

~ Disaster cases.

~ Not a disaster. I can find it.

Question: What about after you?

Answer: Will leave enough instructions.

~ LIDVID would help. IM has metadata. Should be in the IM.

Someone is confused.

~ We need to start over. This SCR is from Jordan. He is worried about provenance. Recreating bundle



provenance to archive. Need to be able to get back to a bundle as it was submitted. Context products are outliers. It doesn't make sense to force LIDVID there. Other components should have LIDVIDs required unless there is good reason not to. Let's find examples where it needs to be done. That is why the SCR was submitted.

Question: EN is fully for this, right? They believe it's important to reproduce bundles through this method?

Answer: Yes. Jordan says he currently can't do it. Huge red flag. Trustworthy digital repositories should be able to recreate.

Another Question: So we can reproduce it as is?

Answer: No. We are not keeping all the old versions online.

~ Not because of practical reasons, but retrievable.

~ Online is irrelevant. Archived. Need to add additional identifiers so we can get back to it. It involves NSSDCA too. Same issue. Need to be able to recreate it.

Someone from NSSDCA said they can definitely reconstitute a bundle with SIPs if no VID is available.

~ SIP is a mechanism they use to do that. Question is if we are going to register SIPs. Sounds like we would be going to the swamp.

Question: Has anyone here had actual reason to go find a previous PDS4 product that wasn't on the archive?

Answer: Yes (from NAIF)

~ Can reconstitute DOIs.

Another Question: Are we changing DOI with new version?

Answer: Yes.

~ Not everyone does for accumulating bundles.

~ That's one use case. It's good or practical for an accumulating bundle.

~ Someone wants to argue against that. Different meta kernels if LID only.

~ Another person agrees. Prefers VIDs.

~

Someone said that outside NAIF and DOIs not seeing the use cases. This is non-backwards compatible in a severe way. If there's no use cases not sure why we should do this.

~ It will be too late if we wait.

~ The issue is provenance and FAIR. Not a use case. It's a principal.

~ There is a way to do this in the IM. Big issue is that EN is concerned they will get a request to rebuild a bundle and not be able to do it.

~ EN shouldn't be concerned. The node would do that.

~ We are moving towards a central node.

~ The issue is if we are preserving provenance for the archive for the long term.

Let's do an around the room vibe check.

~ For ROSES proposals where they are limited to specific versions it's very useful to have the VID information, even when it's accumulating.

~ Yes. Let's see if people want to disallow, leave as is with LIDs allowed.

Someone says disallow. Discourage a free for all? So 97% disallow and 3% discourage, but tricky because

there are exceptions.

- ~ Someone else is against this. Backwards incompatibility scares them.
- ~ Someone else agrees completely – disallow.
- ~ Another node is having an internal discussion. They are between a free for all and discouraging.
- ~ NSSDCA leans with that node. Free for all and discourage for accumulating. Allow LIDVIDs.
- ~ Someone says it's always been all.
- ~ Don't do number 1. Totally disallow for accumulating. Would be crazy for Catalina Sky Survey.
- ~ That is a specific discussion, have been exceptions made for how to handle the version.
- ~ PSA only has accumulating bundles and don't increase VIDs.

Question: Free for all?

Answer: No restrictions.

Another Question: Has PSA ever gotten a request to reconstruct?

Answer: Never had that request. People care about products.

There are other ways to think about this. Maybe the way we version things could change, like for accumulating data. Some accumulates constantly. Could pick a time scale for versioning. That could be considered, maybe once a month or year so some sort of scale of chunks of time. Think that makes a lot of sense.

Question: What would happen where there's a new calibration? Re-version?

Answer: Would be a new version, new DOI.

- ~ Needs new LIDs.
- ~ For what is referencing the bundle file.
- ~ Collections would have to change if want to reference the bundle file. Have to know "more". We're conflating bundles and collections.

The whole point of collection products is a user could just grab that. They might not want everything in a bundle. Whole point of why we designed this the way we did.

- ~ Then bundles are just a thing that floats. Maybe they don't even need DOIs. Maybe only collections need DOIs.
- ~ Bundles just hold collections together.
- ~ It makes sense to break up accumulating every year. Unclear how else to do it. Agrees with discouraging or disallowing.

The software guys are leaning towards disallow. They want things as baked into the system as possible. Won't always have people with the necessary knowledge over time. Should bake this into the model to assure reproducibility into the future, even though it's onerous to do so.

Question: So the modification history class doesn't matter?

Answer: Nobody reads that.

- ~ Nobody wants old versions.

DOIs didn't exist until three years ago and reproducibility became a pillar for PDS. Sees the benefit.

- ~ We should finish going around the room.
- ~ Wait, sees the benefits, but afraid of the non-backwards compatibility.
- ~ Someone hears that.
- ~ NSSDCA can reconstruct versions of bundles given to them. Collections too.

Question: Is that true of accumulating where the nodes don't deliver the older products that aren't delivered with it?

Answer: Good question. Think NSSDCA just returns what is given to them.

~ For accumulation, someone hopes to get to the point where we don't have to redeliver everything, every time.

~ NSSDCA needs to check associations, thinks they keep records. Can look into it.

Continuing around the room...

~ Someone prefers the status quo, free for all.

~ Someone else is fine with it as it is. Doesn't see strong use cases.

~ One node has always followed disallow, so would be happy to say discourage, but we will still do what we think is best.

~ Someone else said it depends.

~ Another person said discourage.

Someone has an alternate proposal. Suggests this is an issue for 2.0. It would make sense to implement at that time, then and only then.

~ Someone is shaking their head. Says we always have to be backwards compatible. Version one doesn't go away. Need to make versions correctly.

~ In other data models they only version collections and items.

~ Maybe we shouldn't give DOIs for bundles. Could change how we do it.

~ We are handling DOIs all wrong.

~ We can discuss this later.

Not a full buy in from the DDWG. We probably won't make any changes this build cycle, at least.

Question: Ramifications of not changing?

Answer: Someone thinks any bundle without LIDVID references could be trouble in the future.

~ We could make a policy that if you want old data you have to go to NSSDCA to get it.

~ That's not a good idea.

CCB-367 seemed related to this, but the ticket has been overhauled.

~ We seem to have the same discussions. We still need to capture the use cases and add them to the SCR.

~ Use case is accumulating archive, especially one that accumulates daily.

~ That's collections. Another use case is context products.

~ Also typically true of a lot of documents.

Question: Any observational data situations that should always be LID only?

Answer: (Silence)

~ Should say desirable not necessary.

~ The ticket will be updated. We will circle back to CCB-367.

CCB- 363

The question here is if data objects in label should match same order as in a file. Previously we came to a soft agreement that they should not and that we should update the SR and DPH, but then there were

comments. People have opinions. We want to discuss it now.

Question: Does anyone have any new thoughts?

Answered with Another Question: Is there an example of where it's not in order?

~ This came from a validate issue ticket from GEO. Validate had to sort before reading.

Question: If the tools are already updated why are we discussing this?

Answer: To clarify.

We can stick with backwards compatibility. Say it's recommended to be in order, but not in order is fine.

~ PDS4Viewer could be an issue.

~ Someone can't see why it would fail or why we would need for it to be in order.

~ Cleaner, but XML, so already not human readable.

~ Someone's counter is there are often times with images where RGB bands are out of order. Could hose the use of that.

~ Someone agrees, but think what was just said is that there's no policy or requirement. We need a statement in the SR or DPH.

It doesn't sound like much to do on this. Clear case where we do need to have things out of order.

~ Treating it as a format. Need it to have capability to read out of order.

~ The work has been done already for validate tool. Can see issues for software folks, but no strong case for this. So the decision is to clarify that this is not required.

~ Recommended, but not required.

~ Someone likes recommended.

We can circle back to CCB-367, but first we need a stuckie for this.

~ The outstanding issue is in the SR or DPH.

~ Someone believes it's the DPH and maybe the SR.

~ Maybe we can comment that it is not required in the SR and then say recommended in the DPH.

CCB-367

From system. When LIDs change for a product we need a systematic way to show they are from the same product. Goes in the category of provenance again. The problem is that provenance is new and we have had recent requests for more sophisticated information. Other provenance requests want to make it general, so we use standard from WW(?). They are making standards for provenance. We can adopt that. Attachments (in Jira) show the basic model. Very general. Can apply anywhere, in many ways. Need PDS use cases. Proposal is to do LDD provenance. See chart in SCR. A lot of effort lately. Apply best practices. SCR includes example for calibrated Voyager image. Several drafts to throw darts at.

~For some folks we have a processing info LDD. This is taking that and standardizing or generalizing to provenance as a capability to be tracked. Product by product basis. Very important that the IM follows this.

Question: We need this, but if have calibrated images from a single image, new product that points to the old product - but how do you do the opposite?

Answer: The system registry handles that. Can use the LDD to find it. Can version products, but then it's chicken and egg.

~ Someone is thinking of prism. Could be a nightmare to point back to.  
~ Source product wouldn't know what's there. Registry shows that. System can do that as long as we have the connections. Need to come up with the cookbook. Need the use cases.

Back to Jira, comment in Additional Information is unclear. You need both supersedes and superseded by.

~ The reverse link.

~ That's the issue. We can't do that. Need more examples. Once you do that new version, can't glue together.

~ Problem with downloading the old version without knowing it's old. We have the provenance, which is important. ~ Worrying. Doesn't seem like there's a solution to the forward pointing thing and need software.

~ The info is there. Need to know to get it. Same issue for context products. There is a level of complexity.

Question: So, if I request a version, you'll give me the latest product?

Answer: That's how LIDs work, but the system doesn't know what superseded what. There was an issue with that for Mars 2020.

Another Question: This is an LDD, so not required for everything. Just specific individual products that need that traceability?

Answer: Could be useful in many use cases. Use cases are being worked on.

~ Someone is working on additional examples.

~ We can expand over time.

Question: Why is this an LDD and not part of the IM?

Answer: It should have been in from the beginning. Provenance is core to any archive.

~ Backwards compatible. Would support core. Has to be optional.

~ Like an LDD that can work for any type of product.

Question: How close is this to ready for build 14?

Answer: Not ready for build 14.

Another Question: Who controls what superseded what? What if a DP decides to supersede something by saying their product is better?

Answer: Curating node will have to do their job and curate.

Another Question: Derived by and derived from?

Answer: Someone said that Anne is trying to make those connections.

~ Maybe we should enumerate some of the use cases.

~ Someone will talk to Anne.

~ Issue with provenance. Model has standard terms to use. Can add more for our system too.

## DDWG Notes 2023-08-03

---

title: DDWG Notes 2023-08-03

layout: default

date: 2023-08-03

---

August 3, 2023

Notes by Debra Kazden

Known Attendees: M. Drum, E. Guinness, T. Hare, L. Huber, S. Hughes, R. Joyner, D. Kazden, T. Lim, J. Mafi, S. McLaughlin, B. Semenov and R. Simpson  
Known Observers: M. Gordon, P. Lawton, E. Schaefer and M. Tiscareno

### ## DDWG Agenda

(Included in email sent August 1, from M. Drum, Agenda for August 3, 2023)

#### Front-matter/What's new

- We're out of time, so we will likely need to have an emergency extra meeting on August 10th
- SCR Freeze date 8/14/23
- \*\* (Yes - Meeting Aug 10) \*\*

#### Ready for Vote

- CCB-363 – Clarification on ordering of data objects in label Suggested lien: Introduce warning in Validate tool \*\* (Did not vote) \*\*

#### Tier 1 issues

- CCB-356 – Update Constraints on Author/Editor Lists (and add other metadata) – Ron Joyner - Edge cases to work through
- CCB-364 – Disallow bundles referencing collections by LID - Any parting thoughts?
- CCB-336 – Add a License Information class to the Identification Area – Mike Drum - Current plan: Support CC0, everything else goes into new Context Product class
- CCB-324 – Create a data type for Type attributes that require a permissible value list. - Vote by end of cycle
- CCB-366 – Schematron/Schema Version Conflict
- CCB-367 – Superseded by / Supersedes reference type \*\* (Some Discussed) \*\*

#### Tier 2 issues

- Revisit Instrument Package Context products – Matt Tiscareno
- CCB-164 - Display Settings not required for images -- (IMG) - Need to think more about potential updates to Imaging and how Display might fit.
- CCB-211 – XML/schema based files as archival data – Jesse Stone, Mike Drum
- Product Metadata Errata?
- Create a supplemental label to live next to original label, that can add aliases and supersede metadata \*\* (Not Discussed) \*\*

#### Parking lot – need a driver!

- CCB-326 - Inconsistent Direction on How to Use Discipline Facets
- RS: 259/281 (see new attachment to CCB-281) (on hold) \*\* (Not Discussed) \*\*

Next meeting: 8/10/23 for final votes

# DDWG

## Front Matter

(There was a pre-meeting discussion that began with the question of if anything had been resolved at the F2F. People had thought we had for CCB-363, but then a new question was raised for discussion today. Regarding CCB-364, there didn't seem to be enough buy in from the DDWG to move forward at this time. It's on the agenda for last thoughts today.)

The freeze date is August 14, which is before our next regularly scheduled meeting, so depending on if we need a vote on anything we might schedule a meeting for next week, August 10.

Mitch is going to call in today to continue the discussion on CCB-363 (Clarification on ordering of data objects in label, see <https://pds-jira.jpl.nasa.gov/browse/CCB-363>).

# Tier 1 Issues

## CCB-356 - Update Constraints on Author/Editor Lists See <https://pds-jira.jpl.nasa.gov/browse/CCB-356>

It's unclear how much of our discussions have made it to jira, but there have been a lot off-line about suffixes, middle initials, non-western names, etc. Hopefully, we can resolve soon and get ready for a vote. Want to solve the issues ASAP.

Still some disagreement about ordering proposed given family name attribute. Issue is that some non-American names have different ordering. In favor of just putting the name in, but Anne had other thoughts. Also, the official is the orcid, but that's not required. Anne thinks we're over engineering.

Question: Would it help to require orcid?

Answer: If it's easy to get one.

- ~ It would get rid of all kinds of concerns with names.
- ~ But not everyone has one or could get one. People could be deceased.
- ~ Seems impossible for us to require that.
- ~ And it seems wrong to require an ID to archive data.
- ~ There are issues.

Question: Is the whole module optional?

Answer: Yes. Could use author list.

- ~ Than not required. It's hard to buy that it's too hard to get an orcid.

Question: What about when we are migrating old data?

Answer: Maybe we could make it nillable.

- ~ That might work.
- ~ Someone is uncomfortable with requiring it, although understands wanting a way to tell it's the same person. Would think that DPs would make sure their name is properly represented.

~ Not always. It could vary. With and without middle initial, etc. Think it should be required but nillable.  
~ Worth investigating. Will discuss this with Anne. **\*\*(Action Item - Dick)\*\***

Regarding non-western names - given name and family name is pretty international. In multiple data systems. Probably something people are used to running into. Will know how to wrangle their names into the fields. Still prefer optional and human readable name, if that's useful.

~ Seems like we would want one or the other to not conflict.

~ Maybe it would conflict for good reason, like display name versus name for search. Extra optional field called display full name is what someone would do.

~ Someone else sent a link for everyone to see now. It includes the example from Datacite on Zenodo. They call it creator name. Unsure if it's optional. Also has orcid.

We're over thinking this. Older, wiser institutions have already solved this. We should get this to a vote.  
~ It will be discussed with Anne.

Another example has a more Asian name. Has creator name. Shows orcid is optional.

~ The point is that it's okay not to fill in everything.

~ Family name and given name are standard. Creator name might address concerns with display name. Has specific format. Anne is aware of this.

~ We need to consider what creator name actually means.

~ Yes. There's also contributor name to think about.

~ We could use creator full name or something.

~ Need to get something that works for Anne and makes sense to the rest of us.

We can vote next time. Should push this forward.

## CCB-364 - Bundles referencing by LID See <https://pds-jira.jpl.nasa.gov/browse/CCB-364>

We've spent too long discussing this without enough yeses. Punting to next term.

## CCB-336 - Add a License Information class to the Identification Area See <https://pds-jira.jpl.nasa.gov/browse/CCB-336>

A new proposal of the class structure was finished yesterday. Would like people to review it. Maybe we can vote next week. The license information class structure is pretty similar to original description with required name and internal reference. Think we should be able to create a document that describes the license and have it widely available, referenced by LID. Name would be CC0. Reference would be the document with full legal description. In other cases we would allow you to archive your own document, so it could be available and be distributed at the same time. For ESA, they need more of a terms and conditions. Did a different example for that. Will need enforcement at the node level.

We will need some text to vote on for what's allowed at NASA sites. Pretty sure we can't archive anything with a restrictive license.

~ Becky, Tim and others worked on a policy. Mainly, all PDS is CC0, with exceptions. Doesn't have to be tied to this.

~ Someone is worried the license tag might be too hard core for ESA and worried about this as a context product. Sometimes it's necessary to point to an external document.

~ Had changed the name to usage, but circled back to license. Maybe too scary.



~ Maybe it could just be terms so we don't mix license with terms.

Question: Terms as a child attribute of license?

Answer: Yes.

~ Someone is thinking that's okay, but worried they would be long and overpower the label.

We can discuss this and put it up for a vote next week. Please, everyone, take a look. **\*\*(Action Item - Everyone)\*\***

## CCB-363 - Clarification on ordering of data objects in label See <https://pds-jira.jpl.nasa.gov/browse/CCB-363>

Someone feels that we should have it as a requirement that order in a label is the same in a data object. When we were designing PDS4, we wanted PDS4 to be prescriptive. Haven't heard any argument that there's an advantage to this. When looked at CDF for PPI had trouble with order. Making things simple seems like a good idea.

~ Someone else had a use case, but never proved we needed this because PDS4 is a format, not pointing to a format, so have to allow lots of things. The example was an image with bands that are out of order. Would have a section of the file telling how to display it.

~ What we're talking about is what goes in the file area class of the label. Asking they are in the read order.

~ Feels like over engineering. People haven't used this. Don't know why.

~ Yeah, it wasn't required. Would like to ask DPs why they did that. There should be a rational for it. For the sake of users, would like to be clear and not have a layer of confusion.

Someone understands the concern, but validate has been updated to allow out of order.

~ Want as much clarity for end user as possible.

~ No concrete example, but there are lots of digital file formats that are out of order to logically make sense. Images are an example. XML itself is multi-dimensional. It can jump all over the place. A restriction like this might end up making it more confusing for users. Seems like an onerous restriction that could bite us in the future.

~ Can't have overlap. Not sure I accept that we need this. Should correct this in the SR. Don't see that we need to allow labels that don't describe what's in the data file.

Someone would like to know the impact to validate if we make this rule. Not schema based. Unclear how validate will know only for certain IM and above.

~ That's what started this. Validate broke.

~ Yeah, but it wasn't tied to a certain version. Think we need to be careful not to break things.

~ Agrees, this is outside of versioning.

Someone said that another reason they are uncomfortable is that it assumes that labels are being read by linear reading humans, but they are read by computers. When loaded into memory there's an order.

~ That sums up why someone else is against this. Hopes computers are reading the labels.

~ Humans write the programs.

~ Yeah, put in start and end byte. We've made PDS4 as simple as possible. This is outside that scope. PDS4 is not simple.

Someone recognizes we don't want to shoot ourselves in the foot in the future, so this should be

generic, not prescriptive. Needs change over time.

Another objection to the SCR as written is that someone doesn't like putting best practices into the SR. They should go on the DPH. Wording there should be a specific order, but we mean a sequential order. Directions for best practice isn't clear.

~ We decided at the F2F the SR would say this is not required and the DPH would say the best practice.

~ Wording needs to be revised.

~ Someone is surprised we were going to put a not requirement in the SR.

~ Another person agrees. The Problem Statement also isn't true. SR 4B1 says fields must be in ascending order.

~ Good points. Good reason not to vote today.

Question: How are people feeling? Do we want to revise this again? Vote for requirement?

Answer: (Silence)

~ Steve is not here now. Mike will discuss this with Steve. Others can jump in too. **\*\*(Action Item - Mike)\*\***

What would really facilitate this is would be a real example that shows why not to require it. If there's a valid example of why to have it in other order, not sequential, that would help.

~ Trent will look for an example. **\*\*(Action Item)\*\***

Digital objects can have different byte types. We allow that, but not different order. Strange.

~ Someone thinks there's a reason of clarity when they are not in sequential order.

Matt said he will help with this SCR.

We have a week. We can vote next week on whatever state the ticket is in.

~ We need confirmation from the software guys at EN that this won't be a problem.

~ Jordan said it was a bug.

Question: If it's a requirement will that mess things up?

Answer: Don't think so.

~ If data is being produced this will affect that. Could be a problem.

~ Specify IM every time.

~ This is non-backwards compatible. Some old models won't work if validate rule is applied now.

Two ways to look at this. Required or not and is validate okay either way. We know it's okay if it's not required. Need to ask.

~ Mars 2020 was delivering data where objects in the label weren't in the same order in the label.

Validate didn't flag, but then a new version of validate got errors because the rule was included. That happens a lot with validate.

~ While tweaking to add constraints they sometimes add ones we didn't agree on.

~ People can use the previous version of validate.

~ That's not always possible.

~ Configuration control issue.

~ We need to find out how this will affect validate.

The TA says there is no impact on PDS tools. Not sure I believe that.

~ That's for best practices and validate already updated. Now we're talking about requiring it and validate being able to check based on version number. Might need to put this on GitHub.

~ A problem.

~ Someone will find examples of some not in sequential order in the past. If they are not convincing enough we will change to required and get a new TA.

People can comment over email and on jira. Watch the ticket.

-----

We have three votes for next week.

Question: There are three more tier one items. Are we voting on them next week? And the first item on tier two?

Answer: Unlikely we will have them done by next week. We might need to beg EN for more time.

## DDWG Notes 2023-08-10

---

title: DDWG Notes 2023-08-10

layout: default

date: 2023-08-10

---

August 10, 2023

Notes by Debra Kazden

Known Attendees: M. Drum, T. Hare, L. Huber, S. Hughes, R. Joyner, D. Kazden, J. Mafi, S. McLaughlin and R. Simpson  
Known Observers: A. Bailey, M. Bentley, P. Lawton, S. Loftin, E. Schaefer and M. Tiscareno

### ## DDWG Agenda

(Included in email sent August 9, from M. Drum, Agenda for August 10, 2023)

#### Front-matter/What's new

- Jordan is willing to give us a short extension, so if we don't get things done today we can try next week
- SCR Freeze date 8/14/23
- \*\* (Discussed) \*\*

#### Ready for Vote

- CCB-336 – Add a License Information class to the Identification Area – Mike Drum
- CCB-363 – Clarification on ordering of data objects in label Suggested lien: Introduce warning in Validate tool
- CCB-356 – Update Constraints on Author/Editor Lists (and add other metadata) – Ron Joyner - Unsure on status, may delay \*\* (Only voted on 336 - it passed) \*\*

#### Tier 1 issues

- CCB-324 – Create a data type for Type attributes that require a permissible value list.
- CCB-366 – Schematron/Schema Version Conflict
- CCB-367 – Superseded by / Supersedes reference type
- \*\* (Discussed) \*\*

#### Tier 2 issues

- Revisit Instrument Package Context products – Matt Tiscareno
- CCB-364 – Disallow bundles referencing collections by LID
- CCB-164 - Display Settings not required for images -- (IMG) - Need to think more about potential updates to Imaging and how Display might fit.
- CCB-211 – XML/schema based files as archival data – Jesse Stone, Mike Drum
- Product Metadata Errata?
- Create a supplemental label to live next to original label, that can add aliases and supersede metadata
- \*\* (Not Discussed) \*\*

#### Parking lot – need a driver!

- CCB-326 - Inconsistent Direction on How to Use Discipline Facets
- RS: 259/281 (see new attachment to CCB-281) (on hold) \*\* (Not Discussed) \*\*

Next meeting: 8/17/23 unless we get everything done. If we do, 8/31

# DDWG

## Front Matter

We might not have everyone today. Jordan gave us a bit more time for tier one issues. We can meet next week if need be. It's on the agenda. We will need to let Jordan and the CCB know what we're working on if we do need more time.

There was a request to discuss problems with product external today.

# Ready for Vote

## CCB-336 - Add a License Information class to the Identification Area See <https://pds-jira.jpl.nasa.gov/browse/CCB-336>

We haven't spent a lot of time on this recently, but there's not much to discuss or worry about. Reporter signed off that this meets their needs. Centralized bundle isn't a concern. Nodes will need to make sure what they need is included. Will be CCO for most. Got a TA.

Question: Any last minute concerns?

Answer: While doing the TA, couldn't find a definition for the license information class. It should be there to be perfectly perfect.

Answer: It's in the Additional Information section. Can be put in the Requested Changes.

~ It should be in the Requested Changes, but fine.

Another Question: The reference is to document or context product. Can't understand why to a context product. Is it a new kind of context product?

Answer: Wanted to leave it open. This is really for document. Mentioned briefly in Proposed Solution.

Question: Are we good to go?

Answer: Someone said okay, they are content on this one.

\*\*The Vote on CCB-336

ATMOS - Yes

EN - Yes

GEO - Yes

IMG - Yes

IPDA - Yes

NAIF - Yes

PPI - Yes

RMS - Yes

RS - Abstain

SBN - Yes\*\*

Yay! We got a ticket through!

## CCB-363 - Clarification on ordering of data objects in label See <https://pds-jira.jpl.nasa.gov/browse/CCB-363>

As of yesterday, someone withdrew their reservations about making this a requirement.

Suggested Changes... Can vote with a lien. SR will not be updated with this text. DPH will be updated with this change.

It's been watered down so much that it doesn't seem useful anymore.

~ Someone agrees, but it makes it more official if it's written down.

The document writing people want to know where in the DPH this should be added.

~ Good point. We can decide after the vote or wait to vote.

~ Someone will be voting no. The Requested Changes wording needs to be re-worked, especially the mention of specific order. Don't think this is worth doing.

~ We can punt. This is not urgent.

This no longer requires any EN action.

~ We can punt.

Document writers can see section 4.2.3, data storage constraints in the DPH. It might make sense as a footnote.

~ Wishy-washy area. Sounds perfect.

Okay. Let's move on.

## CCB-356 - Update Constraints on Author/Editor Lists See <https://pds-jira.jpl.nasa.gov/browse/CCB-356>

No updates on this in the last week.

~ Someone proposed putting display full name attribute in, someone else thought we were going to drop given name and full name. If we leave those we still have parsing issues. The new issue is that these all go in the citation information area. Now in author/editor list, which is in multiple places. Doesn't solve the problem as originally stated. We need more than a week on this.

Author/Editor lists were brought back in for backwards compatibility. Looked at schema - only used in the class.

~ Also in document and software, but we only changes citation information.

Question: Aren't they still in the same parent class?

Answer: Requested Changes does state they become children. Solves the problem. As far as all the name options, in favor of having all three. Part of the Problem Statement is to gather metadata for DOIs, and Datacite uses all those names. We'll still need to solve that. Think it's better to do that up front.

~ Someone agrees and says we should use what they have. Guidelines.

~ Someone else said that Anne says there are no requirements. They default to orcidIDs. They don't get into the details of how to parse names.

~ That's the point. If they don't care, we don't need to. We're talking about unique identifiers. We don't need to add constraints when Datacite doesn't.

~ This was brought up by Anne. It was to clarify the soup. It's gotten much muddier instead.  
~ It won't get clearer over time.

Question: Did we check in with ADS? They do this all the time.

Answer: We did look at an example from Zenodo.

~ If we are trying to get this for DOIs we can follow the guidelines that are already out there. Question is what Anne says about given and family name.

~ She brought them back.

~ She did, but she is not following the DDWG week to week.

~ She's not answering email on this.

~ We will have unique IDs, so given and family names are not important.

~ Required fields. Only identifier is required. Name, type, etc are not required in the Datacite schema.

~ Someone made a good point that we should go with what ADS uses.

~ Identifier and identifier type required, but don't have to have any names.

Another Question: Is it true that ADS requires those fields?

Answer: Someone suggested several meetings ago that we should contact ADS. Can do that since no one else did. **(Action Item - Matt)** ~ What they have is related to our author/editor list with semi colons.

~ Probably why we did that.

Question: What if we take out the semi colons and break it into classes? Create author name, editor name and let it be?

Answer: We've already asked Mom (ADS) and Dad (Datacite). Time to grow up.

Another Question: Did we ask ADS?

Answer: We should ask Anne. This has gotten muddier.

~ She wouldn't push this through. We are stalling. Thinks we have a path forward with display full name.

Unclear how we decide. We put a comma, so maybe we are already solving this.

~ We should pass this and let Anne start using it. We can tighten it up as needed.

**Action Item - Mike, Ron and Dick** will work on this over the coming week. Think we are close, at least as close as we will be. We will discuss this next time.

# Tier 1 Issues

## CCB-367 - Provenance LDD Using PROV-DM See <https://pds-jira.jpl.nasa.gov/browse/CCB-367>

Moving this to tier two. It's been renamed. Will discuss this later.

## CCB-324 - Create a data type for Type attributes that require a permissible value list.

See <https://pds-jira.jpl.nasa.gov/browse/CCB-324>

We need a point person to help us get this over the hump.

~ Someone thinks they are the point person. Early on we said we should have had a data type, but we didn't. Just a list of enumerated values. The point is we should have done it. If we want to do it now, needs to be typed. Will still need to discuss backwards compatibility.

Someone said this feels like deja vu. Worked to enumerate every underscore type. Made several updates.

- ~ Someone agrees, yes, this is deja vu. Thought someone had done updates.
- ~ They were mostly done for the Geometry and Imaging LDDs.
- ~ If this is resolved we can remove the issue.
- ~ Not exactly resolved, found a work around.
- ~ The concern was from the error/warning messages. Work around fixed the necessary fields. Maybe the end result is we should put another entry in the DPH for best practices and withdraw this.
- ~ We might have already done that.

Question: Does anyone remember?

Answer: Someone wonders if our memory is discussing this two years ago and not doing it.

- ~ Sounds like people made the changes, but didn't close this out. We could vote or withdraw.

Another Question: Does anyone need this now?

Answer: Someone is uncomfortable with anything requiring data types with human readable names. Should be a dead issue. Narrow minded, not forward looking. Don't like tying the data type to name.

- ~ We could punt to version 2.0
- ~ Someone else wouldn't mind. Agrees that name is just a string, not in the model. From the very beginning we have had workarounds, but as a purist, it does kind of hurt.

Sounds like we need to do some checking to see if errors are still thrown.

- ~ Someone is uncertain.
- ~ Let's get some facts. This doesn't sound urgent.
- ~ This was a request from Anne to clean up things. Was triggering errors. Will check on it. **\*\*(Action Item- Steve)\*\***

## CCB- 366 - Schematron/Schema Version Conflict See <https://pds-jira.jpl.nasa.gov/browse/CCB-366>

Question: Is this just a document update? Think this is something we can punt on.

Answer: (Silence)

## Product-external

ATMOS is working on product external labels and encountering issues. Knows Rings, IMG and others are also working on this. There have been three issues. First, the LID formation issue. If it's not six segments it gets thrown out, and if it doesn't have a resolving institution it gets an error. There's been some push to remove NASA or PDS from the LID, but funding is coming from NASA, so would like to have NASA and PDS in there.

- ~ Someone said they don't think anyone said NASA wouldn't be in the LID.
- ~ Someone else thought it was a tech issue.
- ~ The concern was for non-NASA.
- ~ Anything from the United States should have the funding agency.
- ~ We should have guidance for the nodes and we need to address six segments for LIDs.

Ron just made updates for this. Thought we agreed to make the six segment LID standard.

- ~ It just had to start with URN. Someone didn't want NASA, and it had to have three fields.
- ~ It's unclear what we are actually using.

Someone wants to clarify - they said PDS shouldn't be in the LID, but it should be NASA, and then the name of the annex. Discussed this with Nancy and she agreed, but it's a minor point. At Rings, just want



a four segment LID because treating as bundles. Can discuss that.

~ Bundles only have four segments LIDs in PDS. Products have six. Unclear how to have a bundle without products.

~ To have it in the registry. At the bundle level is kind of elegant for search.

~ Someone wants to use it the way we do for PDS4.

~ Someone else is pretty sure we just implemented something that already includes this.

People should check with Mitch on the formation of the logical identifier.

~ Fine, but not everyone was in the loop on this.

~ Took the proposed LID and created rules. Checked with Mitch, then made some changes.

To someone this is a bridge off the PDS island. We want to say at the bundle level that we are going off the island.

~ Ultimately, we need to be able to register these, so they need to be able to validate.

~ Communicated with someone at ATMOS about this. Concerned about the need for four segment LIDs at Rings, but other than the LID their product external is currently validating. May need an SCR to expand agencies.

Question: NASA PDS in the LID?

Answer: Means it's registered in PDS.

~ That's a whole lot of intelligence in LIDs. Pandora's Box. Maybe we need an extra attribute besides the LID.

~ The first three segments of LID have always been to point to the resolving institution.

~ URN, not URL. Now might be a good time to decide if URNs are for resolving.

~ Good topic for a F2F.

~ Need to consider what URN implies.

~ The goal is to resolve this so proper registry is identified.

~ That's the issue going forward. Should discuss it with Jordan.

There are two more LID based questions. If want a reference list to point into the PDS archive, unclear if it should be internal or external. External doesn't have a LID option.

~ Someone says internal.

~ Someone else agrees.

~ Good. Similar to context products. Want to use those too.

~ This is why we argued about the name product external.

~ It's easier if we are allowed to use internal references.

Regarding resolving institutions, if putting something in the ATMOS annex, putting that in a LID is a problem because ATMOS exists in PDS.

~ It was a very optional idea. Like the argument that ATMOS doesn't stand on their own.

~ Sounds like we are setting standards for annexes. Don't think we can set all of this.

~ ATMOS should decide what they want to do.

~ They want to be able to register products so they can be found.

~ Someone agrees about URNs. If we always have to add in a new system it gets hurtful.

~ We need to talk to the API work group.

~ Don't think we need any new SCRs for now. Think we'll be okay with how Ron's schemas work. Need to talk to Jordan about resolving institutions.

DDWG Notes 2023-08-17

---

title: DDWG Notes 2023-08-17

layout: default

date: 2023-08-17

---

August 17, 2023

Notes by Debra Kazden

Known Attendees: M. Drum, E. Guinness, T. Hare, L. Huber, S. Hughes, R. Joyner, D. Kahan, D. Kazden, J. Mafi, S. McLaughlin and R. Simpson  
Known Observers: R. Alanis, A. Bailey, P. Lawton and M. Tiscareno

## ## DDWG Agenda

(Included in email sent August 15, from M. Drum, Agenda for August 17, 2023)

### Front-matter/What's new

- Anything we don't finish today will slip to next build
  - F2F November - Tuesday session on Context Objects?
- \*\* (Discussed) \*\*

### Ready for Vote

- CCB-363 – Clarification on ordering of data objects in label - Updated based on DDWG recommendations
- CCB-356 – Update Constraints on Author/Editor Lists (and add other metadata) – Ron Joyner - Possible changes coming before meeting, we will vote on whatever makes it in by then \*\* (Only voted on 363 - Voted to pass) \*\*

### Tier 1 issues

(everything else moved to Tier 2 until next meeting)

### Tier 2 issues

- Revisit Instrument Package Context products – Matt Tiscareno
  - CCB-324 – Create a data type for Type attributes that require a permissible value list.
  - CCB-366 – Schematron/Schema Version Conflict
  - CCB-367 – Superseded by / Supersedes reference type
  - CCB-364 – Disallow bundles referencing collections by LID
  - CCB-164 - Display Settings not required for images -- (IMG) - Need to think more about potential updates to Imaging and how Display might fit.
  - CCB-211 – XML/schema based files as archival data – Jesse Stone, Mike Drum
  - Product Metadata Errata?
  - Create a supplemental label to live next to original label, that can add aliases and supersede metadata
- \*\* (Not Discussed) \*\*

### Parking lot – need a driver!

- CCB-326 - Inconsistent Direction on How to Use Discipline Facets
- RS: 259/281 (see new attachment to CCB-281) (on hold) \*\* (Not Discussed) \*\*

Next meeting: 8/31/23

# DDWG

##Front Matter

We have slipped past the deadline. We will wrap up anything we can today. There are two tickets to work on - CCBs 363 and 356. Doubt 356 is ready for a vote, but we'll get to that.

For the November MC we have an opportunity for a Tuesday F2F meeting. We could discuss context objects.

~ Someone is not ready to say we need to do that yet.

~ Someone else said F2F meetings are very good and productive. A good discussion of the issues.

~ Another person agrees - productive.

Question: Anymore discussion on going to version 2.0 of PDS4? Is that worth discussing?

Answer: It's on us. Another good thing to hash out.

~ Might not be ready in November, but could hash out a plan.

~ Someone agrees. About a 30 minute discussion.

Another Question: Is there a holding pattern for tickets pushed to 2.0?

Answer: Someone has never checked jira for tickets tagged for that.

~ Others haven't either.

We should have at least a half day meeting. We can formalize some of the things we've talked about. Will get that scheduled. Will discuss with Jordan. Could do afternoon if Jordan doesn't need time so that people can fly in in the morning.

~ Maybe we could discuss the five SCRs from Jordan too.

# Ready for a Vote

## CCB-363 - Clarification on ordering of data objects in label See <https://pds-jira.jpl.nasa.gov/browse/CCB-363>

Some new comments have been added. The Requested Changes were accidentally put in the Proposed Solution. That has been fixed. Think we have landed where we want to be.

~ Someone was concerned about the Problem Statement. Made suggestions to make it more clear.

~ Can propose language of clean up after we vote.

~ Will try later today.

Question: Are people okay with that?

Answer: Fine with that, but concerned about what's now in the Requested Changes.

~ Update DPH 4.2.3 with first line of second paragraph in Requested Changes... "When ordering the elements..."

~ Thought we were talking about order in label. Parts of it is what is bothering them.

~ Edited in real time.

~ Good call.

Question: Other comments before we vote?

Answer: Someone is going to vote no. A lot of best practices in the DPH. Not sure we need this in the

DPH. It's in jira and the meeting notes.

~ Fine, but typical person won't look in jira or meeting notes, so no problem putting this in the DPH.

**\*\*The Vote for CCB-363:**

ATMOS - Yes

EN - Yes

GEO - Yes

IMG - Yes

IPDA - Not Here

NAIF - Abstain

PPI - Yes

RMS - Yes

RS - No

SBN - Yes\*\*

Question: is there a lien to rewrite something?

Answer: The lien is the Problem Statement needs to add correct context for current requirement and best practices.

~ As a point of order, a lien should be written down before we vote.

~ A good practice.

~ It could go in the Additional Information place too.

**\*\*The lien: Rewrite the Problem Statement to add context around the current practices and requirements related to ordering elements.\*\***

That should be added to the ticket.

~ We don't want to stop the meeting to edit the ticket.

~ It can be put in the chat.

**## CCB-356 - Update Constraints on Author/Editor Lists** See <https://pds-jira.jpl.nasa.gov/browse/CCB-356>

This was updated two days ago. Matt reached out to ADS and got a snapshot of their schema. We still haven't changed the Proposed Solution on this. People have been discussing it. They have concerns.

~ The most important is that for documents without a person as author it's unclear how to parse and get them in the system. That's what the Problem Statement focused on. Discussion focused on person names, not organizations. Need to address that. Validation errors for organization now if no author.

~ Warning or error. Parses weird, especially if that organization name has a comma in it. Wanted to make parsing more reliable and identify all potential author organizations and names.

This is not solved. Blocking a good solution for a perfect solution.

~ Someone agrees. This is good enough.

~ Someone else is in favor of getting this in the system and tightening it later.

This would need a giant lien. Could do an evote after we update the Requested Changes or keep this on the agenda or make a new ticket to address further concerns and add new fields to this new class.

~ Someone had a message from Jordan on this. The information needs to be available to support ADS, Datacite, etc, without parsing. We're a modeling organization. If it's all there, part of the complication

for backwards compatibility is the lists were left in. Issue is wanting to model sufficiently so we don't need to parse.

~ Someone else agrees. We make non backwards compatible changes all the time. Shouldn't be holding this back.

~ Not arguing to remove the list. Would like to ignore them for some purposes. Need to be able to support our new requirements.

Let's go around the room to say if we want to address this in a future build or try to pass as an evote over the next week.

~ Someone wants to know if it will get in the next build.

~ Someone will ask Jordan. **\*\* (Action Item - Mike) \*\*** Can't promise, but we did warn them about what we are working on. Probably it will be fine. Been discussing this for months already.

If people want to type their thoughts in chat:

Chat captured:

"from Alyssa Bailey to Everyone: 10:00 AM

I have to drop off - vote a

from Lyle Huber to Everyone: 10:00 AM

Atm votes a.

from Dick Simpson to Everyone: 10:00 AM

Simpson prefers 2a

from Trent to Everyone: 10:00 AM

A

from Ed Guinness to Everyone: 10:01 AM

Vote for A. I need time to see the updated schema and test it.

from matt to Everyone: 10:01 AM

Prefer to implement a solution for this build"

Most feedback seems to be to work on this for the next build.

Question: Does Matt want to speak to the information from ADS?

Answer: Limited usefulness. Should go with what we have. If imperfect we can improve it later. Was a list of attributes. Unsure if they are all optional. Not clear salvation will come from that quarter.

Around the room - there were three for B and five for A. We will cut it off here. We will discuss this in two weeks for the next build.

## DDWG Notes 2023-08-31

---

title: DDWG Notes 2023-08-31

layout: default

date: 2023-08-31

---

August 31, 2023

Notes by Debra Kazden

Known Attendees: M. Drum, E. Guinness, T. Hare, L. Huber, S. Hughes, D. Kahan, D. Kazden, J. Mafi, S. McLaughlin, B. Semenov and R. Simpson  
Known Observers: A. Bailey, P. Lawton, E. Schaefer and M. Tiscareno

### ## DDWG Agenda

(Included in email sent August 29, from M. Drum, Agenda for August 31, 2023)

#### Front-matter/What's new

- New build cycle!
  - F2F Tue November 14th- Afternoon session scheduled to discuss 2.0 wishlist
- \*\* (Discussed) \*\***

#### Ready for Vote

None today

#### Tier 1 issues

- CCB-356 – Update Constraints on Author/Editor Lists (and add other metadata) – Ron Joyner
  - Revisit Instrument Package Context products – Matt Tiscareno
  - CCB-324 – Create a data type for Type attributes that require a permissible value list.
  - CCB-366 – Schematron/Schema Version Conflict
  - CCB-367 – Superseded by / Supersedes reference type
  - CCB-164 - Display Settings not required for images -- (IMG) - Need to think more about potential updates to Imaging and how Display might fit.
  - CCB-211 – XML/schema based files as archival data – Jesse Stone, Mike Drum
- \*\* (Discussed) \*\***

#### Tier 2 issues

- Product Metadata Errata?
  - Create a supplemental label to live next to original label, that can add aliases and supersede metadata
- \*\* (Brief Discussion) \*\***

#### Parking lot – need a driver!

- CCB-326 - Inconsistent Direction on How to Use Discipline Facets
- RS: 259/281 (see new attachment to CCB-281) (on hold)
- CCB-364 – Disallow bundles referencing collections by LID **\*\* (Brief Mention) \*\***

Next meeting: 9/14/23

# DDWG

## ## Front Matter

We have a new build cycle. No news from the CCB yet on the tickets we passed. Hopefully they all cleared. The issues that were left over are in tier one now. None seem very difficult. Hopefully we can tackle them. People should let Mike know if there is anything new to discuss.

There is a F2F scheduled November 14. Scheduled for the afternoon. An open discussion on ideas for the 2.0 release.

Question: On that topic, can we start a list before hand?

Answer: Good idea. People can start submitting things and we can hash them out.

~ It would be helpful.

\*\*Action Item - Mike\*\* will start a document with requirements, steps, blockers, etc. People can add to it.

Someone is technically on vacation today. Would like to bring up an issue to discuss and then drop off.

~ Fine, but first we will have a different discussion, on CCB-336, that they will probably want to participate in.

EN is trying to implement CCB-336 (Add a License Information class to the Identification Area, see <https://pds-jira.jpl.nasa.gov/browse/CCB-336>). The issue is implementing the Requested Changes, as stated, but it's unclear if the DDWG discussed that the identification area is used by all products, so all products will have the license. Will need to write schematron rules. Need to consider if this was only intended for product observational. Could move it out of the identification area. Right now, it would be in every product and optional.

Question: To clarify, because it's in the identification area it will be in every product?

Answer: Yes. Because it's in the identification area it will be in every product.

~ Someone is okay with that.

Another Question: But always optional? Not required for product bundle or product collection?

Answer: The class is available for every product type, but the reference type won't be available.

~ People might want it there.

~ People might want to include open access articles that are published. That would be product document so there are definitely other places we would want to use it.

~ Could be useful for browse products too.

~ Someone agrees.

Question: It will be added to the identification area and will be in every product. Do we validate that reference type has to be product to license?

Answer: That was chosen to be as generic as possible, to just have one reference type.

~ So we will validate it inside every product. No problem. Releasing on Saturday.

A topic someone wants to bring back up is instrument package context products.

Wanted to bring this back up again. We voted No, which surprised people. Wanted to know if it was due to imperfections in the ticket that should be fixed or if people really didn't want it. Seemed like an elegant solution to a problem that nodes are having. Wondered if it could be revived and fixed.

~ PPI would like to see it revived. Recently made some context products that really should be pointing to a parent. Thought the issue was with how the proposal was written. Didn't see anything wrong with with it, it just needed additional work.

Question: Anyone not want this no matter how good the fix is?

Answer: Someone wasn't in favor of it, but okay going forward until they see the new SCR.

~ Another person agrees this should be brought up again.

**\*\*Action Item - Mike and Matt\*** will take another look at the ticket and try to clean it up.

# Tier 1 Issues

## CCB-164 - Display Settings not required for images See <https://pds-jira.jpl.nasa.gov/browse/CCB-164>

Someone doesn't have anything to say about display settings for images, but if people want to discuss it, will listen.

~ Someone else doesn't remember the details of this. It does have Requested Changes. It's an old ticket.

~ Another person would have to look at it again.

The original issue was that there were cases when products could have many image types, particularly back planes, and needed separate instances for each object. Thought we passed something to point to any number of image objects. If we did, this should be easy to clean up.

~ Someone thinks that's right.

~ We can see if that has been implemented and then clean this up quickly.

**\*\*Action Item - Trent\*\*** will check on this.

Someone thinks it will require a schematron rule.

Its good that we have a plan.

## CCB-356 - Update Constraints on Author/Editor Lists See <https://pds-jira.jpl.nasa.gov/browse/CCB-356>

This is the SCR that slipped the last build. People have been looking at it. Unclear if there have been any updates yet.

~ Not yet. Concerned the solution doesn't address the proposed problem. Will discuss it more.

## CCB-324 - Create a data type for Type attributes that require a permissible value list.  
See <https://pds-jira.jpl.nasa.gov/browse/CCB-324>

Think we are close on this one. Think the framing included a historical question about what has already been passed.

~ Someone had an action item to report back on this.

~ We should be able to make a decision after we get the background information.

Question: Was the issue the fatal warning?



Answer: Yes. An error was thrown and the LDD doesn't get published. Discussed adding additional parameter type type.

~ **Action Item - Steve** will check to see if the error was changed to a warning.

Someone thought we did that.

~ Some else agrees.

---

The next two SCRs CCB-366 and CCB-367 were left over from EN requests. CCB-367 was renamed. Assumes nothing has happened yet.

## CCB-366 - Schematron/Schema Version Conflict See <https://pds-jira.jpl.nasa.gov/browse/CCB-366>

This doesn't have much of a description yet. Can't remember what the DDWG needs to decide.

Question: Is this related to schematron and schema having to be the same version?

Another Question: Is that a question?

Answer: Think it was Jordan's question to us. If yes, validate tool would need to check for that. It should be in the SR and DPH.

~ So a documentation update ticket. Needs to be fleshed out. Can give it a shot. It doesn't seem like a ton of work. **(Action Item - Mike)**

Someone is having trouble getting their head around this. Schematron and schema files - version equivalent.

~ Think in the product label if you use IM 1.20.0.0, then schematron file reference should say 1K00 and schema location should also be 1K00.

~ In a label.

~ You could have schematron 1B00 and schema 1K00 and validate won't complain.

~ So this is only a validate issue.

~ That's probably why it's not in our documentation yet. Jordan asked us if validate should be checking this.

~ It should be.

~ Makes sense. They go together.

We should check with Jordan to make sure we understand the issue.

~ **Action Item - Mike** will send the new description to Jordan to make sure it's what he was talking about.

## CCB-367 - Provenance LDD Using PROV-DM See <https://pds-jira.jpl.nasa.gov/browse/CCB-367>

There's been a lot of progress on this by another group that does target ontologies. Discussed this with them this morning. The idea was to use formal W3C and use that for provenance for PDS. Very generic. Did that this morning. Made an example with Voyager. The group will review it and give feedback. Kate Crombie is in that group. The bottom line is that the SCR will be updated and then we can start discussing it.

~ Fantastic.

## CCB-211 - XML/schema based files as archival data See <https://pds-jira.jpl.nasa.gov/browse/CCB-211>

This hasn't been discussed in awhile. It will be fleshed out and revisited. Hopefully it can be extended beyond just XML and provide a structure where you can archive a product and all it's schema files. Just need to take another pass to make sure the T's are crossed.

Question: Any high level concerns?

Answer: (Silence)

~ Will keep this on the agenda and discuss it again in the future.

# Tier 2 Issues

We still have product metadata errata idea. Can try to get a champion for that.

# Back to CCB-211

Someone had files that validated in the past and now don't. Extraneous XML files. Jordan put a ticket in to pass extraneous XML files that aren't labels.

~ Someone else just had a similar problem in Extras.

~ The issue should go away. Would hate to have to label them or to have to throw them away.

~ Issue is if a file isn't logged - don't think we should throw it away.

~ Right. It won't be part of the bundle, but historical.

~ The issue should go away.

Someone thinks that goes away when we adopt LBLX extension too, but it's a chicken and egg issue.

~ Jordan said to go to LBLX.

Question: If using LBLX and have an XML data file in a bundle that should have a label. Will validate check to make sure it has a label?

Answer: Unsure.

~ Sounds like Jordan's ticket would cause a problem in this case. Would think it's an extraneous XML file and ignore it.

~ There should be a flag for LBLX or XML for label extension.

~ Someone is not sure how validate would look at an extra file and see it doesn't have a label.

~ Validate validates bundles and collections and products. If there are other files they should be ignored, but we do get errors now.

~ Over reaching.

In migrations there will be a lot of stuff that's not covered in the bundle.

~ Definitely an issue here.

~ A version 2.0 problem.

~ Laughter.

Validate should report unlabeled files.

~ That would be a good thing. Doesn't sound like we need a ticket for that now. Jordan is putting in a ticket for the extraneous files.

~ Right. It will slow down validate because it will have to crack it open.

Work will continue on CCB-211.

# Parking Lot

## CCB-364 - Bundles referencing by LID See <https://pds-jira.jpl.nasa.gov/browse/CCB-364>

This was moved to the parking lot to indicate it's not going anywhere for now.

Product metadata errata might belong there too. Would be akin to product metadata supplemental. Compelling idea, but not enough use cases to push it forward.

Next meeting in two weeks.

DDWG Notes 2023-09-14

---

title: DDWG Notes 2023-09-14

layout: default

date: 2023-09-14

---

September 14, 2023

Notes by Debra Kazden

Followed by Notes by Richard Simpson

Known Attendees: M. Drum, E. Guinness, R. Joyner, T. Hare, L. Huber, S. Hughes, D. Kazden, T. Lim, J. Mafi, S. McLaughlin, B. Semenov and R. Simpson

Known Observers: A. Bailey, P. Lawton, E. Schaefer and M. Tiscareno

## DDWG Agenda

(Included in email sent September 13, from M. Drum, Agenda for September 14, 2023)

Front-matter/What's new

- Jira is currently down?
  - F2F Tue November 14th- Afternoon session scheduled to discuss 2.0 wishlist
  - Document: DDWG Nov F2F Topics
- \*\*(Discussed)\*\*

Ready for Vote

- None today

Tier 1 issues

- CCB-356 – Update Constraints on Author/Editor Lists (and add other metadata) – Ron Joyner - See new comments and updates
  - Revisit Instrument Package Context products – Matt Tiscareno - We will take another look at the ticket and revive it
  - CCB-324 – Create a data type for Type attributes that require a permissible value list. Need to determine what current requirements are around type attributes. Either just turn error into warning, or create this new data type OR add flag to ldd tool
  - CCB-366 – Schematron/Schema Version Conflict - Send to Jordan to make sure that our understandings match
  - CCB-367 – Provenance LDD Using PROV-DM - Will discuss in the next 3-4 weeks
  - CCB-164- Display Settings not required for images -- (IMG) - Trent will take another look at this. One issue was the one-to-many reference to make sure it made sense for multiple image objects.
  - CCB-211 – XML/schema based files as archival data – Jesse Stone, Mike Drum
- \*\*(Discussed)\*\*

Tier 2 issues

- Product Metadata Errata?
  - Create a supplemental label to live next to original label, that can add aliases and supersede metadata
- \*\*(Not Discussed)\*\*

Parking lot – need a driver!

- CCB-326 - Inconsistent Direction on How to Use Discipline Facets
- RS: 259/281 (see new attachment to CCB-281) (on hold)
- CCB-364 – Disallow bundles referencing collections by LID **\*\* (Not Discussed) \*\***

Next meeting: 9/28/23

# DDWG

## Front Matter

Jira was down - unclear how long - but it seems to be fixed now. Hopefully it won't be a recurring problem.

The document we discussed last time to gather ideas and expectations for version 2.0 discussion at the November F2F has been created. So far, it just has wish-list with full backwards compatibility. Would like people to throw ideas into the document. Keep them short. **\*\* (Action Item - Everyone) \*\***

# Tier 1 issues

## CCB-356 - Update Constraints on Author/Editor Lists See <https://pds-jira.jpl.nasa.gov/browse/CCB-356>

There was a comment added after our last meeting about the parts of citation information used in other cases. A few changes were made in the attachments too. Unclear if we need to pull them into the ticket. All the good stuff is in the attachments.

Someone believes this is ready for a vote. More examples might be nice, but most things are optional, so there are quite a few possible configurations a DP could use. Includes a few examples - a minimal 100% backwards compatible one with author list and a more specific example with all classes used, so yes, potential for overlap and misinformation. Would need to catch those issues (like different spellings of author name) in peer review. It's ready for a vote.

~ Someone disagrees. Thinks people working on this need to get onto the same page. Agrees that it's all being phrased for citation information, but author or editor list are in documents too. Need to decide if we want to keep those or replace or allow mixes - like author list as an attribute, editor list as a class, and we need to hash out how to use contributor list. Need DDWG guidance, especially before beginning the writing for the documentation. Can't edit the DPH before we hash this out.

~ Months ago people pretty much approved this, except they wanted author and editor lists added back. Would like a vote, even if the people working on this are on different pages.

~ We have things in examples that don't make sense.

~ Someone agrees some things seem odd. Also has concerns about how this will work with documents and all optional, so think users will be confused. There's also the potential issue of this going in multiple places and not matching up. Think there are a lot of details to consider for implementation in real bundles.

~ We need a decision tree.

~ Someone doesn't think what's here is wrong. We just need to flesh out the details on how it will be used.

Question: Where are the DPH changes?

Answer: Appendix K has a section on how to flesh out citation information, but that's not the only place author or editor lists show up. Could miss the boat for documents and software.

~ Someone is not sure that software uses author or editor lists, but there's a schematron rule in the IM schematron that you have to use it for documents, so have to include citation information, which is where you start to get the overlap.

~ Someone else thinks there is also a document area.

Question: Is there a reason they both have to exist?

Answer: In document class there's author or editor list, not under citation information. In citation information they show up somewhere else.

~ If using product document, in the identification area you can have citation information, but you never can tell with product document that citation information is required.

~ Not required.

~ But in a backhanded way, because of the schematron rule, it is. Unclear why.

~ Think you could just have author or editor list. Citation information is us telling users how we want it referenced, so it shouldn't be required.

~ Sounds like a separate issue.

~ Someone wouldn't put citation information in a document that someone else put together.

~ Seems like we just need provenance information.

~ Very philosophical.

Question: We have the classes we have. Are we saying we want to replace author and editor lists? If we are keeping these around can they also be in documents and software?

Answer: We could try to put something together - take out the lists, substitute the classes, but it won't be backwards compatible.

Another Question: Would that solve the original problem?

Answer: It solves the issue of them being parallel.

Question: How do we solve the issue with old documents by the Magellan team?

Answer: That would still exist. We don't have a great way to handle that yet.

~ But there's an organization class. Could use that, but wouldn't have an orcid.

~ Taking class structure and putting in an attribute. Won't work. Would need to use the new class structure.

~ Thought that's why the SCR was written.

~ Anne wanted this for Datacite, but we also had issues too. Software was misinterpreting things like Google, Inc.

~ Someone thought this would help with that problem - using organization.

~ It should solve all known problems. Not there yet.

Question: Is there a list somewhere?

Answer: Not a systematic list.

Another Question: Can you put them in the ticket so we will know what the remaining concerns are?

Answer: Yes. **\*\*(Action Item - Dick)\*\***

---

EN needs to take over for a discussion about some issues before the build.

CCB-336 (Add a License Information class to the Identification Area, see <https://pds-jira.jpl.nasa.gov/browse/CCB-336>) says to include the full license document, but unclear what that is.

~ Someone thought that would be worked out with HQ and lawyers. Thinks the license information is downloadable, but is not an expert.

~ The other option is to wave this as a deliverable for for this SCR.

~ Yes, but will effect referential integrity. Anne might be able to provide something.

~ The MC decided we would use CCO to cover most types of data but we were to decide about implementation. Anne seems like a good person to ask about creating the product.

~ Someone thinks we do need to create the product.

~ Another person is confused about why if we are just creating the structure. Unsure why it's required for the implementation now. This is a simple internal reference in a new class. Not sure why we need it now. It's like saying we need all context products created before we can create a context product class.

~ The SCR says to create a document at EN that includes the full CCO legal document.

~ So we need a product document.

~ Someone is still confused about this. Not sure what needs to happen now.

~ The IM has been updated. Doing testing now. During that also doing the build integration and testing. Need to have that reference for referential integrity.

Question: Why can't it be a dummy product?

Answer: We need a real product.

Another Question: Will there be a test class looking for the LID?

Answer: Yes. For testing. Would be a hole if not.

Someone is worried about the implication that this is the only product reference that can be included because that's not true.

~ Someone else agrees, but the issue on the table for the build is to create a product at EN. Can do a waiver or do something that says product to license. Also need to update the DPH to use this class as part of this SCR, so someone needs to write the instructions. Also for updating the SR.

~ New classes, so should be in the SR.

~ Not all classes are in the SR, and we would need additional information.

We can work this offline. Don't think this group is qualified to do this.

~ The document writing team doesn't know what to say.

~ We will work this offline.

Back to Tier 1

## Revisit Instrument Package Context products

No updates on this yet. It will be worked on for the next meeting.

## CCB-324 - Create a data type for Type attributes that require a permissible value list.

See <https://pds-jira.jpl.nasa.gov/browse/CCB-324>

This is from CCB-204 (Define and enforce best practices for discipline and project dictionaries, see <https://pds-jira.jpl.nasa.gov/browse/CCB-204>) from Anne to clean up the IM and validation. She provided a list of things to clean up. In 2021 there was discussion about changing error to warnings if get underscore type without permissible value. A decision tree was provided of the options on the table.

- ~ Someone doesn't remember this.
- ~ Another person is sharing their screen to show the decision tree.
- ~ It's probably from at least two years ago.

Question: Should this be added to the SCR?

Answer: If that's how we want to implement it.

- ~ Currently it just fires a warning.
  - ~ This does seem to match what someone thinks it should say, so yeah, would add that to the ticket. It might soon be ready for final discussion and a vote.
- \*\*Action Item - Steve\*\*** - will add it to the SCR.

## CCB-366 - Schematron/Schema Version Conflict See <https://pds-jira.jpl.nasa.gov/browse/CCB-366>

No updates yet.

## CCB-367 - Provenance LDD Using PROV-DM See <https://pds-jira.jpl.nasa.gov/browse/CCB-367>

On pause.

## CCB-164 - Display Settings not required for images See <https://pds-jira.jpl.nasa.gov/browse/CCB-164>

Not discussed.

## CCB-211 - XML/schema based files as archival data See <https://pds-jira.jpl.nasa.gov/browse/CCB-211>

No movement yet.

## Schema Rules for Product External Discussion

This is also part of the build. It's a good time for people to state any objections.

The SCR was closed, then schematron rules were cobbled together after people started using it.

~ Not sure how much people would have reviewed it. We did discuss it in a meeting. Maybe people can respond to the group and if there is no feedback we can go ahead. Mainly about LID formation.

~ Yes, because rules didn't apply to product external.

~ There were some discussions. Someone didn't have strong opinions. Mainly concerned about first three segments of LID.

~ Right. Urn:nasa:pds. Think that didn't validate.

~ Someone else thinks validate tool and schematron limit the choices for product external.

~ Yes. Validate wants nasa:pds, but ATMOS used annex.

~ Right. Wanted something besides nasa:pds .

**\*\*Action Item - Ron\*\*** will resend out the rules and wait for violent agreement.

## Github Discussion re: space as a delimiter in tables

There has been discussion in Github about allowing space to be a delimiter in product tables. Originally



we didn't want to allow that. Not an SCR yet, but it could be coming up again.

Someone is concerned about Github/jira being open to users.

~ When that was discussed there was push back.

~ We can discuss it again in the future.

---

Richard Simpson also provided the agenda with notes for this meeting.

From: [mdrum@psi.edu](mailto:mdrum@psi.edu)

To: [pds-ddwg2@list.jpl.nasa.gov](mailto:pds-ddwg2@list.jpl.nasa.gov)

Subject: Agenda for September 14, 2023

Date/Time: 2023-09-13T13:40:00 PDT

Hi all,

We will have a DDWG telecon this Thursday, September 14th at 9:30 PDT (12:30 EDT, 9:30 in Arizona). Call-in information is listed at the bottom of the agenda. Notes are built around the agenda distributed by e-mail; but there have been some formatting changes. Notes from meeting discussion are in red; action items are in bold red.

For the latest version of this agenda, see the Google Doc

DDWG Agenda

Meeting called to order by Mike Drum at 0932 PDT.

Attendees: Alyssa Bailey, Mike Drum, Ed Guinness, Trent Hare, Lyle Huber, Steve Hughes, Ron Joyner, Debra Kazden, Pat Lawton, Tanya Lim, Joe Mafi, Stef McLaughlin, Ethan Schaefer, Boris Semenov, Dick Simpson, and Matt Tiscareno.

Front-matter/What's new

Jira is currently down? It's back up now. Mike doesn't know why it went down or for how long.

F2F Tue November 14th- Afternoon session scheduled to discuss 2.0 wishlist Document here: DDWG Nov F2F Topics. The document is pretty bare bones right now; Mike would like to get a good list of v2 desires before the DDWG F2F in November. List topics, don't discuss.

Ready for Vote

None today

Tier 1 issues

CCB-356 (link) – Update Constraints on Author/Editor Lists (and add other metadata) – Ron Joyner

See new comments and updates. Ed commented in JIRA after the last DDWG that (1) we need to figure out what to do with other classes (Document and Software) in addition to Citation\_Information and (2) we need more examples. Mike said that the 'good stuff' is in JIRA for people who want to read it. Ron thinks CCB-356 is ready for a vote; more examples might be nice, but they are not necessary. A lot of what is proposed is optional, so the range of examples could be quite broad. Dick disagrees that we are ready for a vote. Among other issues, the following have not been resolved: (1) do we want attribute and class author/editor lists in parallel (this makes the SCR backwardly compatible) or are we replacing attributes with classes? (2) do we allow mixes? (3) is Contributor\_List in parallel with Creator\_List or is it subsidiary? (4) How do we handle organizations when using attribute lists, and have we handled organizations in the class hierarchy successfully? Dick offered to compile a list of organization names that have caused problems.

CCB-336: Ron needs help creating the accompanying document for implementing CCB-336. Steve and Ron don't believe there can be an implementation without the document; but the document involves 'policy' which is beyond Ron's expertise, Mike's pay grade, and the authority of DDWG. Some suggested this needs to be kicked to MC; others thought Anne Raugh might be able to draft something. Steve said the IM has already been updated without the document; but Integration and Test will fail unless the document (or a dummy document) exists.

Instrument Package Context products – Matt Tiscareno

We will take another look at the ticket and revive it. Matt had nothing new.

CCB-324 ([link](#)) – Create a data type for Type attributes that require a permissible value list.

Need to determine what current requirements are around <type> attributes. Dick created a truth table in 2021 that summarizes what happens in various situations. Steve will add the table to the ticket in JIRA.

Either just turn error into warning, or create this new data type. Currently a WARNING is issued.

OR add flag to ldd tool

CCB-366 ([link](#)) – Schematron/Schema Version Conflict

Send to Jordan to make sure that our understandings match. Nothing new.

CCB-367 ([link](#)) – Provenance LDD Using PROV-DM

Will discuss in the next 3-4 week. Nothing new.

CCB-164 ([link](#)) -- Display Settings not required for images -- (IMG) Trent will take another look at this. Nothing new.

One issue was the one-to-many reference to make sure it made sense for multiple image objects

CCB-211 ([link](#)) – XML/schema based files as archival data – Jesse Stone, Mike Drum. Nothing new.

Product\_External guidelines for LID formation (follow-up to CCB-357) will be going into the next build, if there are no objections. Ron will e-mail current guideline text to DDWG; members should comment.

Dick said that there has been discussion in GitHub about allowing <SP> as a field delimiter. This has been discussed by DDWG previously and the decision was not to allow it; but there may be a new SCR. Steve noted that there is a parallel discussion in GitHub (and e-mail) among people who are not affiliated with DDWG who would like access to PDS JIRA.

Tier 2 issues

Product\_Metadata\_Errata?

Create a supplemental label to live next to original label, that can add aliases and supersede metadata

Parking lot – need a driver!

CCB-326 (link) - Inconsistent Direction on How to Use Discipline Facets

RS: 259/281 (see new attachment to CCB-281) (on hold)

CCB-364 (link) – Disallow bundles referencing collections by LID

Next meeting: 9/28/23

Adjourned 1029 PDT.

## DDWG Notes 2023-09-28

---

title: DDWG Notes 2023-09-28

layout: default

date: 2023-09-28

---

September 28, 2023

Notes by Debra Kazden

Known Attendees: M. Drum, E. Guinness, R. Joyner, T. Hare, D. Kazden, T. Lim, J. Mafi, S. McLaughlin and R. Simpson  
Known Observers: A. Bailey, B. Hirsch, M. Mace and E. Schaefer

### ## DDWG Agenda

(Included in email sent September 27, from M. Drum, Agenda for September 28, 2023)

#### Front-matter/What's new

- F2F Tue November 14th- Afternoon session scheduled to discuss 2.0 wishlist

- Document: DDWG Nov F2F Topics

\*\*(Discussed)\*\*

#### Ready for Vote

- None today

#### Tier 1 issues

- CCB-356 – Update Constraints on Author/Editor Lists (and add other metadata) – Ron Joyner - New comment from Dick on organization names

- Revisit Instrument Package Context products – Matt Tiscareno

- CCB-324 – Create a data type for Type attributes that require a permissible value list. Vote next meeting

- CCB-366 – Schematron/Schema Version Conflict - Talking to Jordan

- CCB-367 – Provenance LDD Using PROV-DM - Work in progress

- CCB-164 - Display Settings not required for images -- (IMG) - Trent will take another look at this. One issue was the one-to-many reference to make sure it made sense for multiple image objects.

- CCB-211 – XML/schema based files as archival data – Jesse Stone, Mike Drum

\*\*(Discussion Focused on 356, some brief discussion on 324 and 366)\*\*

#### Tier 2 issues

- Product Metadata Errata?

- Create a supplemental label to live next to original label, that can add aliases and supersede metadata

\*\*(Not Discussed)\*\*

#### Parking lot – need a driver!

- CCB-326 - Inconsistent Direction on How to Use Discipline Facets

- RS: 259/281 (see new attachment to CCB-281) (on hold)

- CCB-364 – Disallow bundles referencing collections by LID **\*\*(Not Discussed)\*\***

Next meeting: 10/12/23 (Changed to 10/26/23)

# DDWG

## Front Matter

Some people did not get the latest agenda. Hopefully, it's not a persistent issue, but the link in all agendas will always work and go to the current google document.

Reminder - there is a document for everyone to put ideas, input, hopes and expectations for version 2.0 into. If no one puts anything in the document then the push will be to have a 2.0 with a backwards compatibility lock.

# Tier 1

## CCB-356 - Update Constraints on Author/Editor Lists See <https://pds-jira.jpl.nasa.gov/browse/CCB-356>

Updates were added to jira. Dick created a PowerPoint that we will go over now.

Slide - What we have now. The question is how to get more information into author and editor lists in a consistent way. Author/editor list is under document and software and in identification area.

Question: Left box - product document, identification area. Is that just a list? Confused. Usually dictionary is at the top.

Answer: Right. Identification area should be before document.

Slide 3 - Definitions for the author list and editor list. There has been some confusion, but instructions are in the IM.

Slide 4 - Problems. Has examples for if author/editor is an organization or has a non-standard name and how we pass this to Astro data system and DataCite.

~ Someone wanted to take these out and have classes.

Slide 5 - Possible Solution - Now we have classes, instead of attributes. Includes some bugs to the names.

Slide 6 - Possible Solution - Exploding classes, almost everything is optional. (Three boxes - Author list, Editor list and Contributor list.) All include orcidIDs. Contributor also has contributor type.

Slide 7 - Possible Hybrid solution (with use rules). Includes five fairly high level use rules.

Slide 8 - Possible Hybrid Solution Details.

Skipping the rest of the slides now to look at the example in Jira (not part of the PowerPoint).

~ Now looking at test file attachment - testfile1.xml. Example - two characters in text file "x" and "CRLF". Label is an example of how we do it today for product document.

~ Testfile2.xml - same example without attribute author list. Uses class author list. Includes two person classes with display full name, given name, family name and orcid. Document still has author list as an attribute.

- ~ In third test file - substituted author list class in document.
- ~ This is a hybrid solution. People get a choice to use author list attribute or class. More work, but people like Anne think it's a better way of doing it. Allows for contributors and organizations. Someone has questions and comments. Obvious for DP it's easier to use the author list. The author list class could be very useful for harvesting information, so the question is how we encourage people to use the class versus the simple author list.
- ~ Wish list for version 2.0.
- ~ Fine with the structure, but unclear why there is duplicate information in the citation information and document classes. Should think about fixing that.
- ~ This is a legitimate question. Think the simple answer is can save work using attribute in the document class and spelling it out in document.
- ~ This is only for putting new product documents in the public domain that we want cited correctly. Wouldn't do it for a paper by Carl Sagan. Author should put the citation they want.
- ~ Someone said there is a schematron rule that says if you have product document have to have citation information. Think we should fix that too.
- ~ the IM spec says citation information has cardinality zero to one, but there is an error in validation without it. Unclear where that comes from.
- ~ Schematron rule.
- ~ We shouldn't have schematron rules that aren't backed up in the SR.
- ~ Someone likes where this is going. Still needs some clean up.
- ~ Got rid of the creator list. Think this all makes sense. Unsure how to encourage DPs.

Leaving in the attributes should make this backwards compatible. Maybe we should remove them at version 2.0.

- ~ Someone agrees. Added that to our version 2.0 notes. We have some things that we added for backwards compatibility that we should remove at 2.0. It's easier to use the attribute list, but already making giant nested classes anyway. Would be doing ourselves a favor to remove the attributes.
- ~ We could make Ron make an app to create this.
- ~ Someone has written those sorts of programs. Would just need a set of rules and make the best fit. Names are definitely messy, won't be perfect, but display full name helps.
- ~ Unless Elon Musk's son starts writing product documents.
- ~ We should just never archive anything by celebrities.
- ~ Or their offspring.

Question: Is there any way to look up orcidIDs?

Answer: No.

- ~ Someone tried to find examples. No real luck.
- ~ Another person said orcid.org does allow search. Just found my name. Probably best to look at recently published papers.
- ~ Can be used for references. Note they aren't always numeric. Mine has an X in it.
- ~ X might be an exception.
- ~ For extra special.

So, for moving this forward, think some of this needs to get into the Proposed Solution. We need to decide which fields should be required. I would say none or just one, like family name. And we need to decide if we should allow author attribute to remain for backwards compatibility.

- ~ Some people made lists of problems. Not sure we should start monkeying with display name. Examples were trying to address some of this. Could provide more samples.

~ Originally we wanted to remove the attributes, deprecate them. They were put back for backwards compatibility. Would deprecate them.

~ We could do a quick poll in the chat to see if we want to deprecate them. (See chat at the end of the notes.) ~ It's not that simple. Could deprecate them in citation information, but there could be problems in document and software. Could get back to issues with where to put additional information, like the reference.

~ If DPs can fill out author list names they can use the classes.

~ Still would rather not remove from document or software.

~ Could deprecate everywhere and get a warning if used.

~ Someone would still rather leave the simple options at least until we get to version 2.0.

~ If we could push this forward for product observational it would be huge progress. Rarely seen document, and never seen software.

~ We don't really put software in PDS.

Four options were written on the agenda:

1. Yes, keep everywhere for backwards compatibility
2. Yes, keep in Documents/Software but remove in Citation Information
3. Deprecate everywhere (WARNING if used, but you are allowed to archive with it)
4. Remove everywhere

Someone prefers option 3 or 4. Asks that everyone please enter their preference in chat so we can try to make a call. (See chat at the end of these notes.)

Question: Deprecate everywhere means it goes away at version 2.0?

Answer: Yes.

~ Someone added to the 2.0 wishlist that anything deprecated will go away at version 2.0.

If we go with option one, keep them everywhere, maybe we should remove them at 2.0.

~ We need to get people used to the idea that this is coming. Once we get to 2.0, think we will get to 3.0 and 4.0 quickly. We're not in the sandbox anymore.

~ Yes. Would be restricting ourselves from having nice things until we get to the next version.

There's not much in the chat. Seems to be a dead heat between options 1 and 3.

We should just deprecate it now if we are going to do it. To warn people.

~ Deprecating seems like a good compromise.

~ To make things simpler, someone will switch from option 1 to 3.

Let's write the ticket like we are deprecating everywhere. We can update the SCR. The discussion will continue.

Question: What happens in validate if you use a deprecated value? Is there an error in oxygen?

Answer: Validation gives a warning.

~ Someone is unsure what Oxygen does. Or other software.

~ Someone else used a deprecated value recently. It was explicit.

~ As far as documenting our intent, you are warned.

~ Someone else does see it a lot for host versus spacecraft. In Oxygen it comes up yellow instead of green or red.

~ That will be the call for that. Deprecate everywhere, but people can use it.

~ Right. Yellow and underlined in Oxygen.  
~ Will deprecate everywhere, but people can use it and will get a warning.

We're getting closer.

## Revisit Instrument Package Context products

No progress yet.

## CCB-324 - Create a data type for Type attributes that require a permissible value list.  
See <https://pds-jira.jpl.nasa.gov/browse/CCB-324>

Don't think there has been any progress on this. It won't be on the agenda for a vote yet. People should look at it and raise any issues.

## CCB-366 - Schematron/Schema Version Conflict See <https://pds-jira.jpl.nasa.gov/browse/CCB-366>

Jordan updated the description of this to be a better description of the issue. The main concern is that it's not explicitly documented anywhere. Validate does throw a warning. We should document it somewhere like the SR. We need to cross our Ts.

Question: Anyone want to put together the language?

Answer: (Silence)

~ Mike will try. **\*\*(Action Item)\*\***

## CCB-367 - Provenance LDD Using PROV-DM See <https://pds-jira.jpl.nasa.gov/browse/CCB-367>

This is being worked on.

## CCB-164 - Display Settings not required for images See <https://pds-jira.jpl.nasa.gov/browse/CCB-164>

Nothing new on this. We will circle back to it.

# Meeting Chat

from Trent Hare to Everyone: 10:04 AM  
Musk's son: "X Æ A-12 Musk."

from Trent Hare to Everyone: 10:09 AM  
already changed to "X Æ A-Xii". Easier to say... :-)

from Ed Guinness to Everyone: 10:16 AM  
Keep them for backwards compatibility.

from Joe Mafi to Everyone: 10:16 AM  
iii

from Dick Simpson to Everyone: 10:16 AM  
I prefer (i) for now.

from Alyssa Bailey to Everyone: 10:16 AM  
1

from Trent Hare to Everyone: 10:18 AM



iii

from Dick Simpson to Everyone: 10:22 AM

I'll change to (iii) to keep our path simpler. It's easy to add the attribute choice later.

## DDWG Notes 2023-10-26

---

title: DDWG Notes 2023-10-26

layout: default

date: 2023-10-26

---

October 26, 2023

Notes by Debra Kazden

Known Attendees: M. Drum, E. Guinness, L. Huber, S. Hughes, R. Joyner, D. Kazden, T. Lim, S. McLaughlin, B. Semenov and R. Simpson  
Known Observers: A. Bailey, M. K. Crombie, P. Lawton, S. Loftin, M. McAuley, J. Padams, E. Schaefer and M. Tiscareno

### ## DDWG Agenda

(Included in email sent October 24, from M. Drum, Agenda for October 26, 2023)

#### Front-matter/What's new

- F2F Tue November 14th- Afternoon session scheduled to discuss 2.0 wishlist

Document: DDWG Nov F2F Topics

- EN is going to drop the private Jira due to budget constraints. We need to discuss alternatives.

- Private JIRA Cloud with New Funding Source - Discuss with MC / PO / HQ a funding source for the ~\$4000 annual licensing fee for a new, private JIRA Cloud instance to migrate our data. EN does not have resources available for this.

- Public, Open Source JIRA Cloud with history - All past, present, and future issues migrated to JIRA Cloud. This was shot down last time, but migrate it all and get it all for free. All history is world readable.

NOTE: Any user who wishes can be anonymized prior to

migration. [https://urldefense.us/v3/\\_https://confluence.atlassian.com/adminjiraserver0822/anonymizing-users-1142237248.html\\_!!PvBDto6Hs4WbVuu7!Nn4MIAYNaNNIhew-siEzbtndW7bPnWMHiBYy-VDIP09Ba-EhoA5-nmRjxVAtzMAii2rpiDwPd9nwk4LD32vNO7JmDExfoRj\\$](https://urldefense.us/v3/_https://confluence.atlassian.com/adminjiraserver0822/anonymizing-users-1142237248.html_!!PvBDto6Hs4WbVuu7!Nn4MIAYNaNNIhew-siEzbtndW7bPnWMHiBYy-VDIP09Ba-EhoA5-nmRjxVAtzMAii2rpiDwPd9nwk4LD32vNO7JmDExfoRj$)

- Public, Open Source JIRA Cloud without easily accessible history - All future SCRs will be world readable, history only accessible to 10 or less users. We can migrate history to a free, closed JIRA, but only 10 users can access it.

- Alternative ticketing system, like GitHub Issues

\*\*(Discussed)\*\*

#### Ready for Vote

- None today

#### Tier 1 issues

- CCB-356 – Update Constraints on Author/Editor Lists (and add other metadata) – Ron Joyner Updated attachments and proposal from last month's discussion Plan for vote in two weeks?

- CCB-358 – Instrument Package Context products – Matt Tiscareno/Mike Drum Reviving this issue, please add comments to ticket about your current position

- CCB-324 – Create a data type for Type attributes that require a permissible value list.

#### Vote next meeting

- CCB-366 – Schematron/Schema Version Conflict Ticket updated, just need to provide language for the standards

- CCB-367 – Provenance LDD Using PROV-DM – Steve Hughes Ready to discuss

- CCB-164 - Display Settings not required for images -- (IMG)

Trent will take another look at this

One issue was the one-to-many reference to make sure it made sense for multiple image objects

- CCB-211 – XML/schema based files as archival data – Jesse Stone, Mike Drum

\*\* (Discussed) \*\*

Tier 2 issues

- Product Metadata Errata?

Create a supplemental label to live next to original label, that can add aliases and supersede metadata

\*\* (Not Discussed) \*\*

Parking lot – need a driver!

- CCB-326 - Inconsistent Direction on How to Use Discipline Facets

- RS: 259/281 (see new attachment to CCB-281) (on hold)

- CCB-364 – Disallow bundles referencing collections by LID \*\* (Not Discussed) \*\*

Next meeting: 11/9/23

# DDWG

## Front Matter

There is a lot to discuss today. We won't get to the whole agenda.

We have two meetings left before the F2F and our version 2.0 discussion. People need to get topics into the list. Mike will compile it and organize the conversation before the F2F.

EN won't support paying for a private Jira anymore. It's 4 or 5 thousand annually. There are a few options.

~ EN won't be paying for a private Jira. Been paying for it for years. We can ask the PO. That's option 1, to ask them to fund it. Option 2 is to make it open source, which we have been approved to do. The idea was shot down in the past, but have since learned that we can anonymize people. The last options, C and D are some sort of clean open source Jira or Github. History would only be accessible by a handful of users. Future would be open source Github or Jira.

~ Someone would like to be reminded why people were opposed to option 2.

~ Conversations and such that people thought were private would become public. Won't be scrubbed.

~ If we go the anonymous route...

~ Comments in new system would still have name history. History would be anonymized, not future.

Question: Could we trace back if necessary?

Answer: Maybe. Unsure. Fifty, fifty.

~ Someone likes the anonymization of history, but afraid we would lose expertise and knowledge.

Question: The DDWG was set up by EN. What is EN's preference?

Answer: Whatever is easiest. B or C. Making it open in general is preferred. Either system.

~ Someone is not worried about it being anonymous, but would like to be able to see the history. Not worried about the names. Nice to be able to get the expertise.

~ Someone agrees. Not C.

Question: Was part of the concern about making things public that some examples might have used unreleased data?

Answer: That's tricky to scrub.

~ Yeah, but unachieved data is all over the internet. Just because it's open source doesn't mean it's easy to find. People have to look for it.

Question: Can we migrate Jira to Github? Would we be starting over? We should assign all comments to Boris.

Answer: There are a few software capabilities to migrate, but attachments and such might not work. In the last two options, history is limited to a few people. Clean slate going forward.

Another Question: Better integration if we go to Github?

Answer: Either way.

Question: Are we currently paying anything to Jira?

Answer: JPL has an on site license for Jira. EN has been paying for and hosting it on a server that is going away. Unsure how much it was, but in the thousands, but cost will go up with new way they are doing licensing.

~ So, we need a couple thousand a year.

~ We would need someone else to fund it. EN won't.

Question: Open cloud Jira will be open to the whole world? Any topics we wouldn't want people to see, like IT security, could be a part of DDWG discussions.

Answer: That should never be a DDWG discussion.

~ Still, a good reason to have private Jira.

~ With IT security or anything private, that all goes in a private Github. Wouldn't be in Jira. Hears the concern, but it doesn't apply here.

~ If we know it's open source that will affect what we talk about. We don't use the full power of Jira. We could always email each other. Less concerned about security leaks.

~ Option D does have private Github. Could use D and B.

~ Someone does not like declassifying what was previously a private forum. Don't want DPs sniping at us as we make decisions. Seems like something that should be taken to the MC to decide. Between options A and D.

~ That's reasonable.

We can do a straw poll or consider our options.

~ Can do a strawman poll for the DDWG recommendation and then have the MC consider it.

~ Someone agrees.

~ The straw poll will be next time so nodes can discuss the options.

Question: What if we vote that the MC should find the money?

Answer: It doesn't seem like a huge amount.

~ One reason EN brought this up is that it's tax payer money. We could spend it doing other things.

~ Could use Github for issues with LDDs. Maybe a more cost effective solution.

~ Could do it now, issue is not being able to keep the history accessible. Attachments and such wouldn't work. Only ten or so users could still access it.

Someone is worried about Github being hard for people, but we aren't using the full capabilities of Jira.

~ Option D is revised. It should be a combo. Language will be provided for the straw poll.

# Tier 1 issues

## CCB-358 - Need a Context object for Instrument Packages.

See <https://pds-jira.jpl.nasa.gov/browse/CCB-358>

Some people are interested in reviving this ticket. Would like to learn why people voted no last time.  
~ It seems like the solution here is elegant and simple. Would allow for layers, for packages, like Ralph. If people would like modifications it would be helpful to know what. Was surprised this got voted down.

We can discuss this today, but if people voted no last time, please put a comment in Jira so we can know why. The no vote was a bit of a blind-side. Would like to to hear the concerns. Homework for everyone  
\*\*(Action Item - Everyone)\*\*

If this goes to a vote without hearing from people it will just waste DDWG time.

## CCB-324 - Create a data type for Type attributes that require a permissible value list.

See <https://pds-jira.jpl.nasa.gov/browse/CCB-324>

We will vote on this next time.

Question: Any updates?

Answer: None in awhile.

Another Question: Concerns?

Answer: People haven't looks at this in awhile.

Someone said this is ready, but it's been off the radar. Will be put higher on the agenda.

## CCB-367 - Provenance LDD Using PROV-DM See <https://pds-jira.jpl.nasa.gov/browse/CCB-367>

We kind of discussed this for several months. Someone thinks it is ready for serious discussion. Wants it in the next build. Want to use the W3C standard that has common terms and relationships to tag. It can be applied many ways. Collection uses cases, two so far, are superseded LIDs (relating new to previous) and when there's a derived product, like a Voyager image that was processed. Think there's now an LDD that works for our use cases. Wants people to look at this and see if it makes sense. Would like to vote in the next few meetings. Wants it for the next build.

Question: Replaces processing?

Answer: Yes. The idea is to come up with a standard provenance process. Maybe a deprecate process. It would improve interoperability.

Another Question: For reference list?

Answer: Yes. For all possible provenance. A side note is that the W3C website has a lot more than we have used.

Question: Are there specific things in Jira that people should look at?

Answer: (Sharing screen to look at Jira) The LID superseded use case. The idea in provenance, the W3C high level view, the entities, provenance, activities, agents and relationships. Made up a superseding activity for this one. Shows how to relate them all in various ways. Includes seven examples in the

attachment. Shows the model that is strictly defined, but could use additional attributes. Derived from is the key one in this example.

~ Would like ASAP to be able to firm this up in the LDD. Spent a lot of time on this. Used ChatGPT. Gave several approaches. Then need to define the process/activity. Need DDWG agreement on how to describe it. Opening up a new type of LDD. Still need more use cases. For now, want people to look at superseding as our first use case.

Question: Someone has questions. This sounds like a enumerated named list for images or need for superseded and superseded by is being driven by a disjoint. Not what we are talking about, right?

Answer: Had assumed all superseded LIDs were one case. Interesting that that is not the case. We need to capture all the use cases. Can try to lay out a framework, focus on LIDs and use cases and move forward.

Someone likes the structure. Clever, flexible, but different from the traditional way. We designed the submodel. A deviation in style. Might have a higher learning curve. Not sure anyone cares about style.

~ Interesting question. Wondered what ChatGPT would come up with. Went with one of them, but fine using what the team wants.

~ Someone wasn't sure how much of this was informed by W3C.

~ Everything is informed by the W3C standard. Slightly different way to define the model. Tag names are defined in the model. Key point is that the terms are defined in the standard.

~ It might be interesting to see a more traditional PDS4 style mock up.

~ Another person also did a version. Can post that too so people can compare and contrast. The examples have obvious differences.

## CCB-356 - Update Constraints on Author/Editor Lists See <https://pds-jira.jpl.nasa.gov/browse/CCB-356>

A new PowerPoint was added after our last discussion. At the last meeting we agreed to swapping list classes for attributes and adding list contributor. Would like feedback. Hopes to vote next time, but complex, so expecting feedback.

Question: Comments?

Answer: Let's put it up for a vote next time. Then we will get comments.

---

We will vote next time on CCBs 356 and 324. People should put comments and concerns in Jira.

Start looking at 358 and 367.

Lots of homework.

---

The following email was sent after the meeting:

[DDWG] Action Items due November 9 2023

Hi all,

There were five action items for the entire group to prepare for before next meeting:

- 1) Review the options we have for JIRA so we can provide a recommendation to the MC a. Private JIRA Cloud with New Funding Source - Discuss with MC / PO / HQ a funding source for the ~\$4000 annual licensing fee for a new, private JIRA Cloud instance to migrate our data. EN does not have resources available for this.
  - b. Public, Open Source JIRA Cloud with history - All past, present, and future issues migrated to JIRA Cloud. This was shot down last time, but migrate it all and get it all for free. All history is world readable. NOTE: Any user who wishes can be anonymized prior to migration. <https://confluence.atlassian.com/adminjiraserver0822/anonymizing-users-1142237248.html>
  - c. Public, Open Source JIRA Cloud with private read-only SCR history accessible by 10 or less people - All future SCRs will be world readable, history only accessible to 10 or less users. We can migrate history to a free, closed JIRA, but only 10 users can access it.
  - d. Public or Private GitHub Issues with private read-only SCR history accessible by 10 or less people - We can migrate history to a free, closed JIRA, but only 10 users can access it. All future SCRs will be managed in either GitHub Issues (can be private or public), or privately by only the 10 members of the DDWG in JIRA.
- 2) Add comments to CCB-358 about any concerns you have with implementing Instrument Package Context Products
- 3) Review CCB-356 (Author/Editor lists) final proposal for a vote
- 4) Review CCB-324 (Type attribute) proposal for a vote
- 5) Review examples and provide additional use cases for CCB-367 (Provenance LDD)

DDWG Notes 2023-11-09

---

title: DDWG Notes 2023-11-09

layout: default

date: 2023-11-09

---

November 9, 2023

Notes by Debra Kazden

Known Attendees: M. Drum, E. Guinness, L. Huber, S. Hughes, R. Joyner, T. Hare, D. Kahan, D. Kazden, J. Mafi, S. McLaughlin, B. Semenov and R. Simpson  
Known Observers: A. Bailey, M. Bentley, P. Lawton, D. Politte, M. McAuley, J. Padams, E. Schaefer and M. Tiscareno

## DDWG Agenda

(Included in email sent November 7, from M. Drum, Agenda for November 9th, 2023)

## Front-matter/What's new

- F2F Tue November 14th- Afternoon session scheduled to discuss 2.0 wishlist

Document: DDWG Nov F2F Topics

- EN is going to drop the private Jira due to budget constraints. Straw poll on the following options which we will use to make recommendation to MC

- Private JIRA Cloud with New Funding Source - Discuss with MC / PO / HQ a funding source for the ~\$4000 annual licensing fee for a new, private JIRA Cloud instance to migrate our data. EN does not have resources available for this.

- Public, Open Source JIRA Cloud with history - All past, present, and future issues migrated to JIRA Cloud. This was shot down last time, but migrate it all and get it all for free. All history is world readable.

NOTE: Any user who wishes can be anonymized prior to

migration. [https://urldefense.us/v3/\\_https://confluence.atlassian.com/adminjiraserver0822/anonymizing-users-1142237248.html\\_!!PvBDto6Hs4WbVuu7!K2dGE\\_oRvM4yhGFnx46T\\_M76vXnjlx2O\\_YcemG4bqIUeIzPpgCU3HwEKcv7buCXcuzRkE3l3Dq1ihH3JXeD0z3kTelylP47\\$](https://urldefense.us/v3/_https://confluence.atlassian.com/adminjiraserver0822/anonymizing-users-1142237248.html_!!PvBDto6Hs4WbVuu7!K2dGE_oRvM4yhGFnx46T_M76vXnjlx2O_YcemG4bqIUeIzPpgCU3HwEKcv7buCXcuzRkE3l3Dq1ihH3JXeD0z3kTelylP47$)

[1142237248.html\\_!!PvBDto6Hs4WbVuu7!K2dGE\\_oRvM4yhGFnx46T\\_M76vXnjlx2O\\_YcemG4bqIUeIzPpgCU3HwEKcv7buCXcuzRkE3l3Dq1ihH3JXeD0z3kTelylP47\\$](https://urldefense.us/v3/_https://confluence.atlassian.com/adminjiraserver0822/anonymizing-users-1142237248.html_!!PvBDto6Hs4WbVuu7!K2dGE_oRvM4yhGFnx46T_M76vXnjlx2O_YcemG4bqIUeIzPpgCU3HwEKcv7buCXcuzRkE3l3Dq1ihH3JXeD0z3kTelylP47$)

- Public, Open Source JIRA Cloud with private read-only SCR history accessible by 10 or less people - All future SCRs will be world readable, history only accessible to 10 or less users. We can migrate history to a free, closed JIRA, but only 10 users can access it.

- Public or Private GitHub Issues with private read-only SCR history accessible by 10 or less people - We can migrate history to a free, closed JIRA, but only 10 users can access it. All future SCRs will be managed in either GitHub Issues (can be private or public), or privately by only the 10 members of the DDWG in JIRA.

\*\*(Discussed)\*\*

Ready for Vote

- CCB-356 – Update Constraints on Author/Editor Lists (and add other metadata) – Ron Joyner

- CCB-324 – Create a data type for Type attributes that require a permissible value list.

\*\*(Voted on both)\*\*

Tier 1 issues

- CCB-358 – Instrument Package Context products – Matt Tiscareno/Mike Drum New comments being added



- CCB-366 – Schematron/Schema Version Conflict Ticket updated, just need to provide language for the standards
  - CCB-367 – Provenance LDD Using PROV-DM – Steve Hughes Any new use-cases?
  - CCB-164 - Display Settings not required for images -- (IMG) Trent will take another look at this One issue was the one-to-many reference to make sure it made sense for multiple image objects
  - CCB-211 – XML/schema based files as archival data – Jesse Stone, Mike Drum
- \*\* (Brief Discussion of CCB-358) \*\*

Tier 2 issues

- Product Metadata Errata?

Create a supplemental label to live next to original label, that can add aliases and supersede metadata

\*\* (Not Discussed) \*\*

Parking lot – need a driver!

- CCB-326 - Inconsistent Direction on How to Use Discipline Facets
  - RS: 259/281 (see new attachment to CCB-281) (on hold)
  - CCB-364 – Disallow bundles referencing collections by LID
- \*\* (Not Discussed) \*\*

Next meeting: 12/7/23

# DDWG

There's a lot today.

~ The F2F is next week. The topics document is slowly accumulating agenda ideas. It will be compiled later today to see what we have time for. A new idea is to review all the issues in Jira to see what we want for version 2.0. Concerned that might be too long of a discussion, but will take a look. The 2.0 wishlist did not grow much beyond that.

~ Someone is curious about full backwards compatibility until version 3.0 transition time.

~ Another person thinks it will be quicker than our 1.0 to 2.0. Our version 1.0 was more like a 0.0. We can talk about plans for updates.

~ Someone else has been discussing this with Jordan.

~ Think when we go to 2.0 we will have to go to 3.0 if there are non-backward compatible changes.

~ Yes. EN wants to go to semantic versioning for 2.0. Would move to 3.0 or 4.0 when non-backwards compatible changes.

~ Okay with that, but surprised that computers are okay going from four number version to a three number version.

~ We can talk about that, but basically using a known version scheme instead of making one up.

For Jira, let's discuss the options. Someone added a new option to the list - to have EN make the decision.

~ That proposal is based on the fact that EN picked the tool to manage us and they should pick the new tool. Proposal to go through all the SCRs is part of that. Might not be practical to do at the F2F, but should be done.

~ Someone agrees. It is important to preserve the information in the current Jira. Important from a historical perspective to be able to look at past decisions we made and have that information.

~ Someone thought option C and we could give shared passwords to the nodes to view the history.

~ That can be done, but don't say that to EN because of IT security.

SBN wants to make two points. One, practicality. The move to Github would be very difficult because we use a lot of attachments, and two, this is the year of open science, so we should be moving towards that. At least the new discussion should be open.

Question: Anne says future discussion should be open?

Answer: Yes.

~ It's the year of open science, but also of strict IT security.

~ Someone hears that, but all of our software development is in open source world. We just have to deal with it. Also, regarding open communication - there are ways to shut this down if there are bad actors or threats. We could even have a rolling window where only people we approve can comment. We will keep the history no matter what, in the cloud, available to a smaller group of people.

~ We're already doing development on Github, which gives lots of opportunities to comment to us. We are not a closed book, but it's also okay for us to have a private space. Would be good to keep our conversations private and access to our history. Favouring option A. We should discuss it with the MC and PO. Think seamless access to our history and keeping it private are the preferred option.

For the straw poll it seems like we will be spread out over the options. We can rank our choices and present it to the MC or let EN make the call.

~ These options get into the weeds of implementation. Interested in knowing how much agreement there is for what was just suggested.

~ Someone completely does not understand why this needs to be private. Unsure what we are saying that needs to be so private and who is trying to find it. We shouldn't be afraid of people seeing what we are saying. Option B.

~ It's complicated. Boils down to private versus public and Jira versus Github. Private versus public is the real debate.

~ Someone believes option A is the most cost effective. There is no transition cost. The other options will need additional management and fees. Whatever we are currently paying, it will be more, and will include unknown headaches.

~ There will be a cost no matter what. Labour cost, because it's on a machine at JPL.

~ Regarding private versus public, think it's bad form to take anything people understood was private and make it public.

We will go down the attendee list by node and see what people say:

~ ATMOS - B, D, C

~ EN - Private history, future open. D or C. D,C, B.

~ GEO - D, A. Prefers private DDWG deliberations.

~ IMG - Only option to be private going forward is A. Confusing. B, C, D.

~ IPDA - D, C, A

~ NAIF - A, D private

~ PPI - A, D private, but not sure.

~ RMS - A, D private, C, D public.

~ RS - E - let EN decide or A - least disruptive. After that, don't care.

~ SBN - C, A, B

Feels like this needs to go into a ranking algorithm. Seems relatively split.

~ Someone has been on the public side. Don't think we'll get much interruptions. We can barely get people to look now. Would love to have input from more people.

~ Yeah, only participation is from the nodes. Our software is in Github and only gets looked at by the nodes. No one cares. A and D seem more popular. Fine with D private.  
~ Someone else is unclear on why we need to keep this private, but agrees we should keep legacy comments private because we thought it was private.  
~ Another is on board with that.

Question: So, people are on board with history being private and public going forward?  
Answer: People seem okay with Github.

EN proposes Github and private. So, D, private. All history there, a few with access. Private Github going forward to manage or can propose A and bring it to the MC. Actually, we will bring both to the MC - D private and A.  
~ A lot of people prefer Jira.  
~ We'll present the options to the MC.

# Ready for Vote

CCBs 356 and 324 are ready for a vote. there has not been much activity of either.

## CCB-356 - Update Constraints on Author/Editor Lists  
See <https://pds-jira.jpl.nasa.gov/browse/CCB-356>

It seems like there are two proposed schemas. Unclear if something should have been removed.  
~ Someone thought we are going with swapping classes for attributes and schemas can follow. The TA went that way.  
~ Someone else reviewed this yesterday. Can't tell which attachments because there are so many. Wants to know if Steve knows what he is going to implement.  
~ Steve thought he did. This should be cleared up.  
~ Someone wants to know what we would be voting on.  
~ We will go with the one that was uploaded a minute after - the second one. Uploaded at 9:59. Will remove the other attachments. Can be a lien.  
~ We will need to resolve any differences.

Question: Are we voting on the October 10 attachment?  
Answer: PowerPoint presentation.  
Another Question: Based on a document above?  
Answer: The PowerPoint assumed we would do the swap.  
~ The Proposed schema made sense to someone.

The text file is out of date based on changes proposed by Dick. It comes down to a cut and paste for Steve. A draft of the changes to the DPH is here too.  
~ Really old.  
~ We're not ready to vote. We want all this to be explicit.

Question: To clarify, is this all changes to the DPH or just for this issue?  
Answer: This issue.  
~ Compared the schemas - different.  
~ Out of date with the most recent proposed changes.

- ~ Could look at the PowerPoint pseudo schemas, but not cut and paste.
- ~ Steve said he could implement based on the PowerPoint.

Question: Has everyone looked at the PowerPoint?

Answer: A lien to do the changes to the DPH.

- ~ Someone is okay with that.
- ~ If we know what to implement and no one has any issues, it's okay to vote.
- ~ Someone asked to be reminded of what we are voting on.
- ~ Vote to implement based on the October 10 PowerPoint and liens to revive the draft for the DPH

**\*\*The Vote for CCB-356:**

Atoms- Yes

EN- Yes

GEO - Abstain

IMG - Yes

IPDA - Abstain

NAIF - Abstain

PPI - Yes

RMS - Yes

RS - Yes

SBN - Yes\*\*

That was a big one.

**## CCB-324 - Create a data type for Type attributes that require a permissible value list.**

See <https://pds-jira.jpl.nasa.gov/browse/CCB-324>

Question: No activity in the last weeks. Any final comments before we vote?

Answer: Two people will abstain. Unclear what was being said.

- ~ Someone else is also confused. Unclear anymore what the problem is.
- ~ Another person likes it.

Someone looked at the old comments. One from Mitch that at some point the LDD flag changed this to a warning. Enumerated flag solves the problem with out adding complexity.

~ CCB-204 (Define and enforce best practices for discipline and project dictionaries, see <https://pds-jira.jpl.nasa.gov/browse/CCB-204>) said it was an error.

~ Yes, until Jordan said he wouldn't post with errors, so we made it a warning. Probably should have made this a warning too.

~ This is a warning, more of a formality. Confusing to people.

Question: Do we want to withdraw this or vote?

Answer: We should vote and see how people feel. We can't move forward without a vote.

- ~ Okay, let's vote.
- ~ We can vote on if we want to withdraw.
- ~ Someone would vote to withdraw. The bang is not worth the buck.

**\*\*The Vote on CCB-324**

ATMOS - No

EN - No  
GEO - No  
IMG - Abstain  
IPDA - Abstain  
NAIF - Abstain  
PPI - Abstain  
RMS - No  
RS - No  
SBN - Yes\*\*

This won't be pushed to the CCB.

~ We probably should have done it way back in the beginning of PDS4.

# Tier 1 issues

We have a few minutes left. Will broach the subject of instrument context.

~ Someone was hopping, if Jordan still here, wanted to ask about LDD updates.

## CCB-358 - Need a Context object for Instrument Packages.

See <https://pds-jira.jpl.nasa.gov/browse/CCB-358>

There has been some discussion on instrument packages and new comments in Jira.

~ Someone thinks it has been resolved. Need to make it clear that it's an instrument package.

~ Larger point was if we needed to more carefully define instrument host instrument packages.

~ Instrument package is more adaptable.

~ Also a suggestion to make it nestable.

~ That's exactly what this is. Every instrument host has to have a type. This is adding instrument package type.

~ Can be a host of a host.

~ Documents already say we can do this, just trying to.

~ We can clean up the definitions.

~ This is about other use cases, besides just Ralph and Cassini MIMI.

~ Radio Science could be a package.

~ Don't know if open and closed loop have distinctions, different instruments or teams. If not, might not be a package.

Adding this option gives more flexibility to the DP. The more information we can gather, the better.

~ Someone agrees.

---

The F2F is next week. Our next meeting after that will be December 7. We will probably only have one meeting in December. Only one more meeting in 2023.

DDWG Notes 2023-11-14 (F2F)

---

title: DDWG Notes 2023-11-14

layout: default

date: 2023-11-14

---

November 14, 2023

Notes by Debra Kazden

\*There are also notes by Laura Givens.

## DDWG Agenda

(Included in email sent November 13, from M. Drum, DDWG Open Session Agenda - Tuesday Nov 14)

\*\*Part 1: IM 2.0 – Jordan Padams and Steve Hughes\*\*

12:00 to 1:30 PST

- 1) Full backwards compatibility until 3.0
- 2) Removing deprecated items
- 3) Timelines and future major version cadence

\*\*Part 2: Other DDWG topics\*\*

2:15 to 3:45 PST

- 1) Product External best practices: ~30 minutes
- 2) Provenance LDD: ~15 minutes
- 3) Cleaning up context products: ~15 minutes
- 4) A plan for the 56 unresolved tickets

\*\*Part 3: Review Decisions/Action Items\*\*

3:45 to 4:00 PST

# DDWG F2F Discussion

## Part 1

\*\*Full backwards compatibility until 3.0\*\*

The goal of the 2.0 discussion is to understand what 2.0 is and when we will transition.

Starting with a presentation by Jordan, Steve, EN

PDS4 IM 2.0

We need to decide if backwards compatibility is allowed after version 2.0 and if non-backwards compatibility forces us going to version 3.0.

~ Someone says no, backwards compatible changes are allowed.

~ Another person says no non-backwards compatible changes from now on. Going to use versioning the way we're supposed to.

Proposal to use semantic versioning. Three numbers, separated by dots. m.n.p. Can have two digits for m. We will deprecate the use of letters. Namespace info won't change.

m - major  
n - minor  
p - patch  
Well defined rules exist.

Question: Does anyone remember why we started with four number versions? Maybe there was a reason. Want to know what we are abandoning.

Answer: Someone is not sure. Can check notes.

~ Someone remembers arguing against four at the time.

Another Question: How many people use all four now? Would be happy with three. Think the fourth was for bug fixes.

Question: Will there be any trouble with parsing?

Answer: Might impact some custom software ~ Can't throw anything away.

~ Tools were built to handle what makes sense. Better to do this now.

Question: Why the tick tock release cycle?

Answer: INT (integration and testing) life-cycle. Required. Every six months. Costly to do a fully integrated test. Can choose cadence.

Another Question: AB - fiscal year? Calendar year? arbitrary?

Answer: Arbitrary.

Question: Why A & B at all?

Answer and Another Question: That's why asked about the tick tock cycle. Why not 13, 14, 15? Every release would be minor.

~ Right. System builds increment over time.

~ Not sure how 13A, 13B maps...

~ Unsure why not just have mnp number?

~ Not good for bug fix, unplanned hot fix. We have system build and separate builds.

People are trying to understand the steps. Now they are every six months. In future, 2.N.0. Every six months update - N.1, N.2, etc.

~ Yes. PDS hasn't used the third number much in the past year.

Someone would like to go back to the question about why we had four numbers, but if people aren't really using all four, maybe it would be better to do what is more common.

~ We still haven't heard a good reason to change yet. It's a big impact.

~ Software engineering best practice. Only places we really need to change are rare - like schematron version number. Software typically does a sub-string check.

~ The question is totally reasonable. Primary reason is to go to something standard and use versioning correctly going forward. More in common practice.

In our m.n.o.p version the o was ignored. It was intended to increment when enumerated values changed.

~ That hasn't been a problem. Adding enumerated values is backwards compatible. Don't think we need to worry about that.

~ Standard values really only change on a build.

Question: What is the impact on user tools, especially non-EN tools? Is this a concern? Like for ISIS?

JMars? There could be substantial impacts. We could stick with four and say the last one doesn't matter.

Answer: Each node could check with their DPs and users to check impact.

~ This is for 2.0, which will be advertised well in advance. There could be a problem for PDS-Viewer. A lot of people do depend on that. It does need to be version aware.

~ Most bugs or problems would be when you are checking for backwards compatibility, from a software point of view. Believes that it would be good to do this, despite the minor pain involved. Version 2.0 is the right time to do it.

Regarding backwards compatibility - there are always non-backwards compatible changes. Software has to know and adapt, but we can introduce this and move forward. At version 2.0, this won't be the biggest change.

Someone is still not sold on this. People have written scripts. Having stuff change is a big impact. Not convinced that this being common practice is a sufficient reason to change.

~ It's a benefit to our customers. It can be well advertised. Goes hand in hand with non-backwards compatibility.

This can be finalized in a future vote in the DDWG.

There is a question that ties into this, a major question to address if we push for version 2.0. Need to decide if we push for version 2.0 if that means we have to migrate to 2.X. that's a major funding issue. If not, need to know that impact.

~ No. The system has to support the past versions. Migrating is up to the nodes.

~ Recently passed an SCR for author lists and FAIR, but if that's not in our holdings, not sure we are meeting FAIR.

~ FAIR is how it appears in the archive. Software support for versions prior to version 2.0 is not going away. Can add supplementary metadata in the registry or can rewrite labels. Can do what makes the most sense. Long term planning. Maybe get supplemental metadata into the archive too.

~ Looking forward to that tool. Someday.

~ It will be essential, at least for SBN.

~ Half of our labels don't contain DOIs. The future landing pages will have them, even if they are not in the label. Systems always has to move forward. The archive can't be perfect, we just have to keep augmenting it.

Before we go to a 2.0 vote we need things spelled out in writing so we know and understand what we will be voting on.

**\*\*Removing deprecated items\*\***

The first question is why remove deprecated items at all. The original design was to allow for change. Deprecated things are flagged, so the IM is staying self consistent. Has complete history all the way back to version 1.0. We can roll back at any time.

Some reasons to continue with the current scheme:

~ Compatibility issues

~ Loss of historical context

~ Resource requirements



~ Reduced maintenance.

For system, in general, when you just keep removing a method, why bother if when you upgrade / deprecate, it provides a more seamless transparent way to move forward. Removing deprecated items becomes dicey. Things blow up.

~ It used to be in software development that things were deprecated for an entire release to warn people it would be removed, so it wouldn't have to be supported forever, and it was removed at the next major version. Don't want to have to continue to support deprecated features forever.

~ But that puts a requirement on the nodes to update. If data is the same as version 1.0, and they move to version 3, there's no history of what they need to support.

Question: So, we want to go to semantic versioning and say that people don't need to check version?

Answer: Maybe we need to be less lenient.

~ We want to clean up the IM as we move forward.

~ Someone agrees.

~ We have been including deprecated things for documents and schema.

~ Software for version 3.0 shouldn't have to support deprecated things.

~ Software can say what version it is designed to work with.

~ That's too much of a burden. We want to clean it up.

~ Someone would like an example of it being a burden.

~ When to find display settings.

~ Yes, but there were lots of standard values eliminated that no one ever used.

~ Standard values don't affect software. Don't understand why that's a problem.

Question: Could we segregate documentation from output? Provenance is important to capture. If we could keep it in the system, that would balance the two.

Answer: People will validate to a certain IM, that should be fine.

~ This makes sense for EN.

~ The database IM is stored, no need to write things out. We can decide where the complete history is written out.

~ We want to encourage the behaviour that we want to see in the community.

~ Need provenance.

~ LDDs need this, need to use provenance LDD in display LDD.

~ We can auto-generate warnings for users.

A proposal - The IM database will remain the same, won't write out deprecated stuff to LDDs, etc.

Someone can see himself wasting time searching for deprecated attributes.

~ We can document the changes.

~ Deprecate in version X etc. Only Steve sees the data base.

~ In Github. Hard to find.

This proposal and semantic versioning changes will be voted on in a future DDWG meeting.

Question: This backwards compatibility thing, where we keep going to new versions, will this come with a desire to no longer make non-backwards compatible changes? We were in the sandbox before. We need to agree to be very careful about non-backwards compatible changes.

Answer: Non-backwards compatible changes happen fairly often, like the SCR for author/editor list.

~ The concern is that we want to evolve. Agree, we want to avoid them, but there is evolution.  
Another Question: Was there a proposal to keep backwards compatibility? Do people look for that?  
Answer: That's the third topic for today.

### **\*\*Time-lines and future major version cadence\*\***

The question is when we are moving to 2.0 and how much friction we want to add to the process. Think we would have a default number of minor versions and save up major changes (non-backwards compatible) for the next major version.

- ~ Someone likes the idea of trying to figure backwards compatible ways of doing things. Concerned with other time-lines driving necessary changes.
- ~ The development process evolution, the way someone assumed it would work, would be like in other standard organizations. (Transition process was described) It's not a sudden switch that is flipped. Also, we have a registry which allows us to fudge how things are presented to the public. Major version is to clean out the stuff that needs to be cleaned out. Takes a long time.
- ~ Someone thinks the pattern the DDWG used when designing author/editor list is probably how we want to go in the future. Evolve the model, but deprecate and know deprecated things will eventually be removed.
- ~ Missions can use things that are eased into the IM, optional in this IM, required in the next.
- ~ We have versioning at the class and attribute level.

Question: If we don't drop deprecated things there are a lot of inefficiencies?

Answer: Schematron takes forever to validate. XSD validation is better, but schema is very complex. A major problem. Also, we need to maintain the documentation, even for deprecated things. Users will continue to use deprecated things. Everything new has to work with old things - lots of overhead.

- ~ Someone remembers speaking with Jordan a while ago, he said he would like to support IMs forever, but tools will get very expensive.
- ~ The registry and update the metadata.
- ~ For community tools they will decide. EN will support it all going back. The every version system.
- ~ Point well taken, but we do have to care about community tools.
- ~ That is part of the life of software.
- ~ If you don't know it's coming it's annoying. Our engaged community should always know about changes coming in the IM. They should never be surprised.

Question: What are we doing to communicate updates?

Answer: That's a whole can of worms itself. A conversation we should have. All nodes talk to their communities.

- ~ At the system level there are no clear lines of communication. No clear user list. We blast out emails to several lists. Need a better way to communicate.

The conclusion is that we will need a proposal about major version cadence, but we have a good way to make changes going forward, despite timing concerns.

Implement in current IM, deprecate in next, then remove in next major version. Can handle it if things get in on a minor version.

The last point to decide is when we will go to version 2.0.

- ~ It doesn't seem like there is much of a hurdle blocking us, except getting the information out.
- ~ There will be some impacts on EN to support it.

Question: Move to version 2.0 in six months?

Answer: Just a build. Would need to plan. We are in 14.1 now. Could plan it for next December.

~ Version 2.0 will be in December 2024. We will work out the specific details.

Another Question: Is one year enough time?

Answer: People should rush in their non-backwards compatible changes now.

~ People haven't been holding back, and we do have a handle on how to do it now. LDDs will be harder.

If people have issues with this, now is the time to air them.

Question: Will the rules for versioning change for LDDs?

Answer: If versioning changes, it will change for everything.

Another Question: Is the desire to keep the same version for major version of IM and LDDs?

~ That is already broken. Went away a long time ago.

~ We will use versioning properly.

The DDWG will be having some very lively discussions.

Break

Question: Why go to a new version every two years by default?

Answer: Part of the reason for the cadence is so we can clean things up over time. We just had an SCR where we created new classes, deprecated attributes. Backwards compatible, but we want to eventually clean that up.

~ The definitions of major and minor are laid out. Don't see why a time line is laid out. We should just do it when we need to.

~ As we adopt semantic versioning and proper versioning we need to make sure the IM is clean and make sure that we are not moving too fast. So we will deprecate things and adopt changes at a quicker pace. That's why a cadence. It's important for those reasons.

~ Someone disagrees. Fine with a cadence for builds, once a year, but still doesn't see a reason to change the major version every few years.

~ The DDWG should evaluate the schedule. Want phasing. Don't want major version, major version, major version...

~ Right. We haven't had many non-backwards compatible changes. No objection to evaluating every few years.

~ We don't want to declare a cadence where people rush to get untested stuff rushed into the next version.

~ But we could get people dragging their feet the other way.

~ Requirements from missions come in a rush. DDWG should review every year.

Question: Change comes from the missions?

Answer: Missions don't know what they will need until they try to run their pipelines and then they start demanding immediate changes. Also, as the PDE becomes a thing there will be pressure on PDS and the IM from that.

~ Over the last few years, not so many changes were driven by missions, which use LDDs. There have been more from FAIR, DOIs, etc.

~ That's going to ramp up.

~ Someone is on board with rewording evaluate the need for new versions every year to watch the cycle

of when we deprecate things so it's not done too quickly.

## ## Part 2: Other DDWG topics

### \*\*Product External best practices\*\*

Not all nodes are using product external yet, but some nodes are and there are concerns.

~ If using a product external, it should be in an external bundle. And there are remaining concerns with LID formation for product external.

Question: External bundles and collections make sense for organizing principal. Is it just a matter of adding another type?

Answer: It should be there already.

Another Question: Is the question if there should be a requirement?

Answer: RMS made a stand alone product. It's not a member of a collection or bundle. Unclear if that is okay for the registry.

~ Registry can because no checks for product external. Standard can say product external is excluded. Better to have everything labeled.

Question: How are we sure LIDs are unique for product external?

Answer: We wouldn't know until harvest.

~ LID formation has been gone over.

~ Grouping is reasonable, but some nodes don't want to label every product, so the question is what to do when we have a data set of files. RMS made a table of references. Kind of like a collection inventory. It gets the data searchable. User only cares about the set as a whole.

~ The registry only allows findability at the bundle level.

~ No. At the product level.

Someone is unsure what an external bundle is.

~ Someone else thought it was a bundle with product type external.

Question: Is it in the IM?

Answer: Yes. Bundle and collection type external.

Another Question: Did the DDWG decide this?

Answer: Yes.

~ Someone thought the point of product external was to attach data to a label/metadata even though it's not an archival product.

~ Someone doesn't understand an external data set. No inventory with LIDs. It has filenames.

Another Question: If product external has labels for products?

Answer: It's overhead. Can have that, but RMS didn't need the individual products.

Another Question: Why is there a different way of handling this?

Answer: Different use cases.

At PPI, we use product external with LIDs, collection, bundle. We do it the regular way.

~ That's ideal for findability. We are talking about if you want findability but don't have time to do it all.

~ A different way of doing it. Can have a single label with as many files as you want.

~ The difference for RMS was they used a list like in a collection inventory. The other way would be a giant label, referenced by path.

- ~ A quirk is that it would need to be in the same directory.
- ~ We could tweak it like documents. Don't see why we need a different way of doing it when we already have a way that works.

This sounds like a new SCR.

- ~ Maybe we should bring this up tomorrow. People might want a quick way for things to be found. Quick and easy for the PDE. Product external is a bridge. It should be quick and painless. This data set at RMS has a whole bunch of data products listed. It's not practical to have all the paths.
- ~ Could be done an ad hoc way each time. Not a good idea.
- ~ Sounds like the PDS IM is ad hoc for DOI, but not validation.
- ~ Not if validation is checking metadata requirements. Trying to figure out ways to get things in the registry. Can get a DOI.
- ~ But this could be a link to repositories that are not part of PDS.
- ~ Unsure why PDS is labelling things in Zenodo or other repositories without constraints. We can just get a DOI and give it rich metadata.
- ~ Minimum curating.
- ~ This is very disturbing.

Product external label is part of the archive, but what it labels is not, so it's a bridge off the island.

- ~ Doesn't make sense to include that in the PDS archive. Should have DOIs.
- ~ That's what we are doing. Connecting... (Interrupted) ~ There are better mechanisms to make the connections. That's what DOIs are for.
- ~ Across data repositories, beyond DOIs.
- ~ Our data is under curation. This will at least erode our reputation.
- ~ Someone disagrees. It will enhance it.
- ~ Unclear why it would erode our reputation.

This is a deep philosophical discussion.

- ~ We have product external.
- ~ We have spent a lot of time on philosophy instead of best practices.
- ~ Need more examples. Can even see using this for the PDS3 archive.
- ~ Someone is still fuzzy on this.
- ~ It makes sense for things that aren't translatable.
- ~ And the idea is to not short circuit the metadata. Hopefully, archives outside PDS will have metadata and the ecosystem is allowing us to link to that. Want to use it more like a product native.
- ~ Right. That's why we spend so much time on the name.

**\*\*Provenance LDD\*\***

The DDWG was all assigned homework to look through provenance model and the two use cases and add more to address individual discipline node needs. We now have CCB-367 (Provenance LDD Using PROV-DM, see <https://pds-jira.jpl.nasa.gov/browse/CCB-367>).

The model is accepted world wide, with accepted keywords for relationships and activities. Our uses cases are Voyager corrected image and superseded LIDs. We have an example of the manual design Voyager use case in XML. The second use case is the relationship between two superseded LIDs. The problem is that the way it's applied depends on the domain and we just want a general solution. Steve went to ChatGPT to play with this.

~ Phase 1 - ChatGPT came up with a few iterations. He chose one to write up in XML with a time stamp. Eventually, got an XML schema.

~ Phase 2 - Used the schema to merge in the superseded LID. Got a draft schema and then a cleaned up schema and reverse engineered a PDS4 LDD/ingest LDD.

~ He found some interesting things - it told him to define the relationship and the activities. Saved several months of DDWG discussion for the first draft. Also interesting was that it added things he hadn't thought about - attributes. It extended the model to add additional information.

Someone is concerned that this takes over the processing LDD.

~ This is just one of the use cases. It shows superseding the processing LDD. Unsure how to map it. Under the hood could use that model still.

~ A lot of people use the processing LDD.

~ We should think of a way to map that.

~ Can try to merge it in

Steve showed his XML - his example was discussed.

Part of ChatGPT is that it will help us get drafts on the table earlier for the DDWG.

Question for EN: Think we could pay to train on the IM?

Jordan: It already is for PDS3. It was crashing our servers. It's just doing it.

~ You pay to use it. They grab all the free information they can get.

~ EN looked into it. Training for planetary. The latest ChatGPT version knows about PDS. There are ways we can potentially use it or ask questions or add information, but we are still on the outskirts. It's fascinating.

~ Someone believes it is picking up PDS labels, but we should be able to train it to use the IM, give it a mapping and get drafts of labels.

~ Have to pay to use the latest version.

Question: On the provenance LDD side of it - where would this go? In the existing product?

Answer: Superseded LIDs, like when LIDs are mismatched, need a new product to put that in. Need some new product to make it happen. Other cases might need to be thought about. Unclear where they would go. Provenance of the IM versus the products.

~ We will need to hard code something into the IM.

~ Someone was hoping to get something like this. The model of the model that describes every change from every version and then can do translations etc. This is a very critical piece. More valuable for provenance of the model itself.

~ There are so many possibilities. Need a cook book of examples. Can build it up over time.

The provenance LDD is structured different from what we are used to. It's interesting. Still need more use cases.

~ Supersedes is easy versus superseded.

~ Need to think about making it specific.

~ We can work on real stuff. It will help us decide if this is useful.

**\*\*Action Item - Everyone\*\*** we need more use cases.

~ Radio science might be an example.

~ Alternate version, PDS3 to 4 migration, might be helpful with compressed.

- ~ It comes down to reproducibility.
- ~ Mosaics. Would need a reference of something.
- ~ Already supported by the current IM.

### \*\*Cleaning up context products\*\*

Many of the context products were cut and pasted from PDS3 and need to be examined. They are not in good shape. Missing information. Teams of us need to go through them, examine their completeness. Should be pointers at both ends. Need a focused, coordinated effort like the DSMP.

Question: How many are there - order of magnitude?

Answer: Targets alone are in the thousands.

- ~ Someone is not sure this needs to be done before version 2.0. Seems untied.
- ~ Content instead of structure. Should at least start to see if there are structural changes that would improve the system.
- ~ The amount of work, if it's in the thousands, will cause concerns of who does what. It's complicated. No authority to make people do the work. Might need more budget.

Question: What would the general process be?

Another Question: And what is the minimum requirement for context product and what is the point of them?

Answer: Need ideas on how to approach this.

- ~ Someone thinks it's probably true that more structural than anything else for connections for search.
- ~ Maybe we should start with hierarchies and work down to see how it works.
- ~ Someone agrees.

ATMOS led the creation of the context product guide. A pretty solid start. Need a plan for how to get this started. Maybe each node could try five or ten hierarchies.

- ~ That would be the point of doing a two or three day DDWG session, like the DSMP. But yes, we should have each node do a hierarchy of a mission first, see what needs to be done, and the time scale.
- ~ Trying to curate this. It's not an in person thing, more of a homework assignment. People will need to investigate stuff.
- ~ We need to make a plan to make a plan.
- ~ Would still lack a holistic view.
- ~ Richard Chen.
- ~ Yes. We would need someone to create the structure we need.

Someone is wondering what we actually want to solve.

- ~ Referential integrity.
- ~ There are two problems. The content of the context products and attributes and relationships between context products. Nodes have the content.
- ~ Yes.
- ~ Good. We want to keep it simple.
- ~ This is an important enough type of product to have a consistent way of creating them.
- ~ It would be useful to put references to papers and DOIs in them and get the walls of text out.
- ~ It might be useful to separate legacy ones from new ones that follow guidelines. Concerned about referential integrity.
- ~ Can do it, but they go in all kinds of directions. Complicated.

- ~ Someone has ideas about new references that are needed, but many are missing. Many don't follow the handbook.
- ~ Many were created early.
- ~ We need to resolve the relationships and be very clear.
- ~ We need to figure out the right answer and named relationships.

Sounds like we need to start a requirements document with needs, concerns, etc.

- ~ Someone will create a spreadsheet. (Jordan?)

Context products can bind products together or provide a lot of information.

- ~ Not sure we can do the second part.
- ~ Should resist overloading the context products. They should lead to another source of information. Context product would be the hub to connect things.
- ~ Someone agrees. Likes the term hook. Unique identifiers. No limit on what to reference.
- ~ Links to what is there is critical to the system.
- ~ Information would be in the documentation. The whole point is to facilitate search.
- ~ Exactly.
- ~ In web world we need a sufficient definition of the concept. Should be able to get to the documentation.
- ~ So make a Git website for every context product.
- ~ We already have a way to do that.

### ## Part 3: Review Decisions/Action Items

Agenda review

- 1a - proposal
- 2a - proposal
- 3a - proposal

Someone wants more stringent requirements for making non-backwards compatible changes. Added as 3c proposal.

- ~ Someone else will be interested to see this. Will need proof things can't be done in a non-backwards compatible way.
- ~ We can often do something - but it might not make sense.
- ~ It would be good to have a set of principals.

### Part 2

- 1a - need more product external use cases.
- 2a - more use cases
- 2b - more in PDS4 style. (Jordan will help Steve) 3a - Before 2.0?
- 3b - Action Item - EN
- 3c - holistic view. Maybe a Tiger team with Richard.
- 3d - new structure?
- 3e - task nodes
  - i - references
  - ii - content



Someone made a non-backwards compatible change to a context product last week. Same LID, new name.

4 - At some point we need to go through the DDWG list and figure out which SCRs we need to work on that will be non-backwards compatible.

Almost time for a new DDWG chair. Mike will be looking for a replacement.

DDWG Notes 2023-12-07

---

title: DDWG Notes 2023-12-07

layout: default

date: 2023-12-07

---

December 7, 2023

Notes by Debra Kazden

Known Attendees: M. Drum, E. Guinness, L. Huber, S. Hughes, R. Joyner, D. Kahan, D. Kazden, S. McLaughlin and R. Simpson  
Known Observers: A. Bailey, B. Hirsch, M. McAuley, E. Schaefer and M. Tiscareno

## DDWG Agenda

(Included in email sent December 5, from M. Drum, Agenda for December 7th, 2023)

## Front-matter/What's new

1. We're moving to GitHub: [https://urldefense.us/v3/\\_https://github.com/NASA-PDS/PDS4-CCB/issues\\_!!PvBDto6Hs4WbVuu7!NYbpZpNsS\\_7qS68YiejVg5CnTR2OcExSSw0de3LMLb9SRjTK7dyyo151V5Z5Yy2TdTLb3dhCUanSiAzD-lhki1LXORIXn-ub\\$](https://urldefense.us/v3/_https://github.com/NASA-PDS/PDS4-CCB/issues_!!PvBDto6Hs4WbVuu7!NYbpZpNsS_7qS68YiejVg5CnTR2OcExSSw0de3LMLb9SRjTK7dyyo151V5Z5Yy2TdTLb3dhCUanSiAzD-lhki1LXORIXn-ub$)

a. Issue Template created for all new issues b. New issues should go there; active ones may need to be copied c. Proposals from the F2F have been created there  
2. CCB isn't going to pass Author/Editor lists a. Missing explicit way to indicate first author/editor/organization b. Angst about compatibility  
\*\*(Discussed)\*\*

Ready for Vote

- None this week

Tier 1 issues

- CCB-358 – Instrument Package Context products – Matt Tiscareno/Mike Drum Ready for vote?
- CCB-366 – Schematron/Schema Version Conflict Ticket updated, just need to provide language for the standards
- CCB-367 – Provenance LDD Using PROV-DM – Steve Hughes Still looking for use cases
- CCB-164 - Display Settings not required for images -- (IMG)

Trent will take another look at this

One issue was the one-to-many reference to make sure it made sense for multiple image objects

- CCB-211 – XML/schema based files as archival data – Jesse Stone, Mike Drum

\*\*(Some Discussion)\*\*

Tier 2 issues

- Product Metadata Errata?

Create a supplemental label to live next to original label, that can add aliases and supersede metadata

\*\*(Not Discussed)\*\*

Parking lot – need a driver!

- CCB-326 - Inconsistent Direction on How to Use Discipline Facets
- RS: 259/281 (see new attachment to CCB-281) (on hold)
- CCB-364 – Disallow bundles referencing collections by LID **\*\*(Not Discussed)\*\***

Next meeting: 1/4/24

#DDWG

## Front Matter

We will be missing some people today. This is our only meeting in December. There are a lot of changes coming.

The MC voted to get us to GitHub. EN has set up an area/template to set up issues. It's a private group, so it will look like a 404 error. Doing a soft transition. Still using Jira for now. People can get access to the GitHub through Jordan or Mike for now. The link will take you to issues tab on the PDS4/CCB repository main page. It's a bit confusing. Very blank. Click the issues tab. The new issues template is very nice. People can create an issue or feature enhancement request. It's pretty straight forward. The first step is to create your own GitHub account. Then it can be linked.

Question: Will the CCB members be able to get to it?

Answer: Eventually, if not yet.

~ Will evolve over time. Will be where we're going. Added a few proposals from the F2F. Will discuss them soon, possibly today.

Other news, that is not official yet, is that the CCB voted one yes, three no and two abstains on author/editor lists. A lot of their concerns are things we worked through.

~ Non-backward compatible.

~ There were two camps - one concerned with first author or editor and another concerned with non-backwards compatibility.

~ Who the first author is is not clear.

~ Right. No longer obvious. May need to make another design pass on that. As far as compatibility, working with Jordan on that. Thinks there is a misunderstanding about compatibility. Might go to the MC.

~ It's not backwards compatible.

~ It is.

~ Not. Think the issue is to do with deprecation.

~ Yes. Don't have to make changes to your software, so it's backwards compatible.

Question: What happens in validate if you use a deprecated value?

Answer: Warning. So it passes validation.

Question to NSSDCA: To submit to NSSDCA we have to submit a validation report. Will it be accepted with warnings?

Answer from NSSDCA: Validation reports are required by EN, not NSSDCA. Believes it's just looking for errors, but not sure. It's a question for EN.

~ Okay, so we can still submit, but the question is really for EN. Wondering because just learning the process now.

~ NSSDCA doesn't use the reports.

~ Just trying to understand if there are any gotchas when using a deprecated value.

~ NSSDCA uses the bundle and collection products, links in documentation. Mostly concerned about

URLs in submission manifest table and checksums.

~ So, if there's an issue it's probably related to the registry.

~ Regarding registry and validate, Jordan is aware of this and says this is compatible, so that's why he thinks there's a misunderstanding. Another person thinks it's better to do this now, rather than deprecate at version 2.0. It gives people a full year to make changes if we do it now.

~ Great argument for any change request with backwards compatibility issues.

~ It's awkward if we deprecate and move into the next version quickly. Think the requirement should be class or attribute. Should be deprecated a full year before it's removed.

We need to be careful. People don't have to move to the next IM.

~ A speaker agrees.

~ Not a good general strategy.

~ Someone is currently migrating with version 1.13. They don't want to jump to 1.20 or 1.21. That's a whole new ball game.

~ When you start a new project, you should use the most current model.

~ 1.13 was most current when the work started.

~ We need to encourage people to use the newest IM, but we have to support past versions. We can wait for an official response.

Question to Steve: Would like to know why this is not backwards compatible. Would like it explained to those who think it is.

Answered with Another Question: Is this an action item for Steve to the CCB?

Answer: No, explain it now.

~ Can't.

Another Question to Steve: This is backwards compatible, even though the TA says otherwise?

Answer: Used the team view in the TA, but personally prefers the other view.

~ A new definition of backwards compatibility.

~ If an old label can't validate, then it's not backwards compatible.

~ Right. An old label would still validate, but with a warning, so backwards compatible.

~ But this is cruising to not validating. So not backwards compatible.

Someone disagrees with that definition of backwards compatibility.

~ Someone else understands the concerns.

~ People would like to see examples in the common model.

~ They are more in the LDDs, but will look in the common model **\*\*(Action Item - Mike)\*\*** ~ We have removed stuff that no one ever used to convince ourselves things were backwards compatible.

~ Can't say that because it could still be used in the future if someone is making labels with an old IM.

Obviously, the impact of non-backwards compatible change on something no one was using is less than what we are saying here about requirements on DPs and fair timelines for deprecated items. We can call things backwards compatible, but we need an actual understanding of the usage and impact. Backwards compatibility really makes a difference when we go from optional change to required change.

~ We've made lots of non-backwards compatible changes. We do know, and think most of us agree with the need for this change. We need to discuss a timeline.

~ We can call this backwards compatible classic.

~ Don't have the authority to do that.

~ Regardless of what we call this, it's better to do this before major version.

~ See that and agree in principal, but would implement a major version.

~ That is what we will do. 2.0 will come along next to 1.21.

Question: Release 1.22 at the same time as 2.0?

Answer: No. Would be at least a six month gap.

Some people on the CCB think they shouldn't vote for this because it's non-backwards compatible.

~ Could make it backwards compatible, but it muddies the model.

~ Someone would vote against that.

~ We should move on.

~ Yes, the CCB decision will be revised by them.

# Tier 1 issues

## CCB-358 - Need a Context object for Instrument Packages

People have talked and agreed to a few changes. Need to update a few of the definitions. There is interest in moving this forward.

~ We discussed this last time.

~ Mike has an action item to address concerns in the comments. Would also like a better sense of where people are on this.

Question: Wouldn't show up in regular labels, just context products?

Answer: Yes.

Another Question: Would be in the observing system component?

Answer: It's confusing. Tried to be vague on data product labels. Most changes are just to context products. A new change for data products.

Another Question: Would imply if you define an instrument, it would be it's own instrument, under instrument host?

Answer: Yes. Mimi would be an instrument host. Lemms would be instrument. Cassini would also be instrument host.

~ Okay. Clever, but still possible confusion on what constitutes a package.

~ That's the part that needs work and changes.

Probably an okay SCR if we clarify some stuff.

~ Someone thinks a lot of what to call an instrument is up to the DP. We just need to provide tools and let DPs do what makes sense.

~ Could say that radio science is a package of open loop and closed loop. We haven't before.

~ Someone still thinks we provide tools and let DPs do something reasonable.

Question: Would anyone here still vote no?

Answer: Someone would prefer to nest instrument host.

~ We are nesting.

~ We have to have types of host.

~ Just providing instrument package as a type of host.

~ Taking this as a good sign. Will try to get this updated. Hopefully we can vote in January, if we are ready.

## CCB-366 - Schematron/Schema Version Conflict

We need to clarify that this is an inappropriate thing to do.

Question: Is anyone able to write good documentation on this?

Answer: (Silence)

~ We will ask Jordan what he thinks.

## CCB-367 - Provenance LDD Using PROV-DM

This was discussed at the F2F. Looking for more use cases.

~ Someone provided some info for a draft. A good example. Would still like more. Someone pointed out that the Voyager product was never archived, so it's not a valid use case. Would like more valid use cases. There was some concern about how this was laid out. Will look at that, but will just reorganize a few things. Will be worked on.

~ Someone else has feedback to provide.

## CCB-164 - Display Settings not required for images

Skipping this. Trent is not here.

## CCB-211 - XML/schema based files as archival data

Nothing yet. Maybe when Mike is no longer chair he will have time to work on it.

---

Next meeting January 4, 2024.

~ Happy Holidays.