Meeting Minutes from the DDWG-2 telecons:

DDWG Notes 2025-01-09
DDWG Notes 2025-01-23
DDWG Notes 2025-02-06
DDWG Notes 2025-02-13
DDWG Notes 2025-02-20
DDWG Notes 2025-03-06
DDWG Notes 2025-03-20
DDWG Notes 2025-04-03
DDWG Notes 2025-04-08 (F2F)
DDWG Notes 2025-04-17
DDWG Notes 2025-05-01
DDWG Notes 2025-05-15
DDWG Notes 2025-05-29
DDWG Notes 2025-06-12
DDWG Notes 2025-06-26
DDWG Notes 2025-07-10
DDWG Notes 2025-07-24
DDWG Notes 2025-07-31
DDWG Notes 2025-08-05
DDWG Notes 2025-08-14
DDWG Notes 2025-08-21
DDWG Notes 2025-08-28
DDWG Notes 2025-09-18

title: DDWG Notes 2025-01-09
layout: default
date: 2025-01-09

January 9, 2025
Notes by Debra Kazden

Known Attendees: A. Bailey, J. Blalock, B. Hirsch, L. Huber, S. Hughes, R. Joyner, D. Kahan, D. Kazden, P. Lawton, J. Mafi, S. McLaughlin, E. Schaefer, M. Seritan and M. Tiscareno (*'s indicate people who voted during this meeting - if applicable)

DDWG Agenda

(Included in email sent January 8, from J.Mafi, Agenda for 2025 Jan 09)

We will have a DDWG telecon this Thursday, January 9 at 9:30 PST (12:30 EST, 10:30 MST). Instead of discussing specific SCRs, we will review the status of all of the currently open SCRs, assess their status, next steps and priorities. Please be prepared to advocate for any issues that you would like to see be given high priority.

Upcoming meetings:
Jan 23 – DDWG telecon (regularly scheduled)

For the latest version of this agenda, see the Google Doc

Front-matter/What's new

1. Submitted to the CCB:
   - CCB#4 – Backwards Incompatibility policies
   - CCB#48 – Add JP2 as an allowed extension in the PDS tool Validate (approved)
   - CCB#13 – Ensure file name and directory path name adhere to SR
2. New issues in GitHub:
   - CCB#52 – Enable cross-namespace validations in sch files
   - CCB#53 – Improve cross-namespace validations in sch files
3. Ready for a vote (None)
4. Build 15.1 freeze date: 2/27/2025. The following SCRs are queued for implementation: CCB#48, CCB#49.
5. Withdraw by author (None)
   (Discussed)

Tier 1 issues

- CCB#3 – Adopt Semantic Versioning with version 2 transition The CCB returned this, requesting that it be resubmitted as part specific proposal to transition to IM V2.0.0.
- CCB#13 – Ensure file name and directory path name adhere to SR
- CCB#20 (JIRA-CCB-358) – Instrument Package Context Products Need to form tiger team

- CCB#4 – Backwards Incompatibility policies
  Presentation: Backwards Compatibility
- CCB#46 – Disallow multiple versions of the same LID from occurring in a single collection
- CCB#5 – Move to V.2.0.0 for June 2025 Release
- CCB#11 – Add a Funding Information class to the Identification Area
- CCB#32 – Add Host type to Observing System Component
- CCB#31 – Ensure that dictionary versions are specified in labels
- CCB#35 – Add Alias Index and Acronym Index to the Investigation class
- CCB#36 – Add Mission Phase to the Investigation class
- CCB#37 – Add List Author, List Editor, and List Contributor to the Modification Detail class
- CCB#39 – Fixed problem where LDDTool does not include all rules from IngestLDD
- CCB#40 – Change cardinality in Product Bundle to better support PDS3/4 hybrid bundles
- CCB#41 – Create a way for classes or attributes to be exposed to other LDDs but not to data providers
- CCB#44 – Issues caused by the failure to expose pds: attributes that are referenced in an LDD.
- CCB#45 – Add reference types values for *-to agency
- CCB#50 – Clarify Ambiguous Author Information in Product Document
- CCB#51 – Add an optional attribute in the Target Identification class to identify Primary/Secondary targets
- CCB#28 – As a user, I want validate to allow delimited tables with whitespace-only numeric fields
- CCB#12 – Document Nuances for Reusability of Generic LDD Classes/Attributes
- CCB#21 (JIRA-CCB-367) – Provenance LDD Using PROV-DM
- CCB#22 (JIRA-CCB-164) – Display Settings not required for images 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#23 (JIRA-CCB-211) – XML/schema based files as archival data
  (Discussed)

Tier 2 issues

- 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!

- JIRA-CCB-326 - Inconsistent Direction on How to Use Discipline Facets
- RS: 259/281 (see new attachment to JIRA-CCB-281) (on hold)
- JIRA-CCB-364 – Disallow bundles referencing collections by LID
  (Discussed)

Historical links:
Export of history: https:/github.com/NASA-PDS/PDS4-CCB/tree/main/data
Backup: https://nasa-pds.atlassian.net/jira/software/c/projects/CCB/issues (limited seats)
Next meeting: 23 Jan 2025

DDWG

Front Matter

The process today is going to be to check priorities and make sure we are getting things we need to accomplished. We will go through the SCRs as listed in the agenda, check status and discuss. We will try to prioritize in terms of urgency, ease of completion and any other considerations.

CCB#48 (Add JP2 as an allowed extension in the PDS tool Validate) has been approved by the CCB.

There are two new SCRs, CCB#52 (Enable cross-namespace validations in sch files) and CCB#53(Improve cross-namespace validations in sch files). They were submitted over the holiday break. Related to discussions on setting restrictions on one attribute in a dictionary that will effect another dictionary.

Question: Any comments?
Answer: Seems to be, no strong opinion about implementation, but seems like it would be good for cross references we were trying to promote in CCB#41 (Create a way for classes or attributes to be exposed to other LDDs but not to data providers). Think the consensus on CCB#41 is to expose things and trust people not to be bad citizens.
~ Someone agrees. Think cross referencing will be implemented. This is important for Mission Common LDD to enforce values for a specific mission. See these as important. Should get high priority.

CCB#3 (Adopt Semantic Versioning with version 2 transition) and CCB#4 (Backwards Incompatibility policies) went to the CCB. They rejected CCB#4. Mia sent Joe a detailed email with their concerns. He will forward it to everyone (**Action Item - Joe**). Basically, they see the transition to version 2 as a very complex change and they want to see a detailed plan for implementation that can be iterated on and reviewed, possibly at a F2F or something like the DMSP. That's where those two are. Think there's a lot of merit. Discussed for CCB#4 having the details written down. There's some work to do, a lot will need to be in coordination with the document writing group and the CCB. See this as a high priority task that needs coordination.

Tier 1 Issues

CCB#13 - Ensure file name and directory path name adhere to SR

Question: What's the status of this? We voted, it went to the CCB. Never heard back.
Answer: An email about our vote went to the CCB.
Action Item - Joe will check with the CCB on the status of CCB#13.

CCB#46 - Disallow multiple versions of the same LID from occurring in a single collection

There has been a lot of discussion. Think someone was working on this. Think the SCR still needs some updates.
~ It was suggested that we pivot to the SR language first.
~ We discussed this before the holidays. Will get back to it.
(**Acton Item - Matt**)

CCB#5 - Move to V.2.0.0 for June 2025 Release

June 2025 is out of date. We need to re-write this to resolve CCB#s 3 and 4 or close this and write a new one. Needs more work.

CCB#11 - Add a Funding Information class to the Identification Area

We haven't worked on this in a long time.

Question: Any advocates or thoughts?
Answer: Someone thinks there was a recent comment. T.C. was interested and Jordan said this is implemented. Confusing.
~ Sounds like we need to run this down.
Another Question: Do we know who T.C. is?
Answer: On JUICE for ESA. Think lead archivist for JUICE.
~ Probably with PSA to have access to GitHub.

CCB#11 was rejected/withdrawn and closed.
~ Oh. Okay.
~ It was withdrawn by author.

Action Item - Joe will remove it from the list.

CCB#32 - Add Host type to Observing System Component

This is linked to CCB#20.

CCB#20 - Instrument Package Context Products

The Tiger Team met and will meet again next week.

A comment on CCB#s 32 and 20 - We have a higher level ontology. Been asked lately to document it formally. Took all reference types and what they related to to start. Think we need to consider the plan to put them in the context products and think about how we're going to implement them and think about how we're doing it for instrument package and what it means for instrument package. Maybe I (Steve H.) should be on the Tiger Team. Hope that makes sense.
~ It does. In PDS4 everything is supposed to be explicit. If defining the ontology, it should be explicit.
~ The plan is to put them in the context products. Seeing observation area as ontology has been problematic. Needs to be defined explicitly.

Someone thought Steve already was on the Tiger Team.
Action Item - Matt will add Steve to the invite and send him all the Tiger Team emails. Meeting next Thursday.
Action Item - Steve will bring the draft ontology to throw darts at.

CCB#31 - Ensure that dictionary versions are specified in labels

Believe we need updates to documents.
~ Yes. We discussed it at the last meeting. Will check on the action items.
~ Someone thinks we just need to write up the Proposed Changes.
~ Some people discussed this. Think there's more to do. It gets into how references are discussed in the SR.

CCB#35 - Add Alias Index and Acronym Index to the Investigation class

Question: What's the status of this one?
Answer: Steve will try to adjust several SCRs. CCB#35 (this one), CCB#36 (Add Mission Phase to the Investigation class) and CCB#37 (Add List Author, List Editor, and List Contributor to the Modification Detail class) - the ones on ontologies are all requested by the system folks. All have high priority. Impact on search. About five of them. Need to try to get them done. They are all high priority. We could do one per meeting.

CCB#39 - Fixed problem where LDDTool does not include all rules from IngestLDD

Question: What is the status of this one?
Answer: Someone thinks this issue was from Jesse Stone.
~ Someone else thinks this was an implementation issue. Thinks it can be resolved. We should check with Jesse.
~ Best practice might be to combine the rules.
~ Could be manually done or done in LDDTool, but automatic wasn't preferred. Think document

and review approach is preferred.
~ Documenting could be important so people know they need to combine rules. Software wasn't throwing any comments so users would know it was failing.
~ LDDTool could validate this.

Question: Action item? Just leave it open? An LDDTool issue?
Answer: Think we can put the issue on GitHub.
~ Yes. Maybe we need to update the SCR.
Another Question: Who will make the update?
Answer: Steve will update it. (**Action Item - Steve**)

CCB#40 - Change cardinality in Product Bundle to better support PDS3/4 hybrid bundles

Think this was from Bob and that there has been some initial discussion on this.
~ Someone feels like Bob has mentioned that he doesn't want to do this anymore. Need to sit down with him and discuss it. Think it was initially an issue connected to MSL. Will take the action item to figure it out.
Action Item – John

CCB#41 - Create a way for classes or attributes to be exposed to other LDDs but not to data providers

This has been mentioned already. Related.
~ Discussed in context of Rings LDD. Looked at it last time when the DDWG discussed it in October and there was a request from SBN. Discussed cross references instead of duplicating. So yesterday reread the minutes from the conversation in October. Think if expose something it can be used anywhere, but not necessarily as a direct child. Would like some safe guards, but unclear if that will happen.
~ Someone agrees. We should discuss this and think about preferences and enforcement mechanism to not be used irresponsibly. There were some ideas about how to implement restrictions. We can discuss this.
~ Another person likes where this is going, but has some input on this. Would like to be included in the discussions.
~ Post a comment in GitHub. (**Action Item - Steve**)
~ Will do.

There are a number of issues related to sharing attributes.
~ The new ones are related. Think this has been desired for a long time. Think we can make everyone happy.

Question: Currently, if element flag equals true then it has to be a class - direct child. What if we said it has to be a child...? Never mind. Bad idea.
~ For those from the original development, the original plan was to expose all attributes. External flag has little to do with how the model works. Will explain in the comments.

Question: Do we need a Tiger Team for this?
Answer: Don't think we need one yet.

CCB#44 - Issues caused by the failure to expose pds: attributes that are referenced in an LDD.

This sounds like a related issue. Was also from Steve.
~ Yes.
~ This is the SPICE kernel one.
~ A hole that's existed for years. Definitely related. Should all be discussed at once.

CCB#45 - Add reference types values for *-to agency

We discussed this a lot before the holidays. There were action items. Don't think the updates have been made.
~ Someone thought we were pretty close on this one. Maybe ready to vote.
~ The SCR has to be updated before we can vote.
~ That sounds right. Will check. (**Action Item - Steve**)

CCB#50 - Clarify Ambiguous Author Information in Product Document

This was from Joe. The issue came up because in product document there are two places for author information. Purely a documentation issue. Need to have a discussion on author information in citation area versus document area. Not urgent, but we will move this forward at some point.

CCB#51 - Add an optional attribute in the Target Identification class to identify Primary/Secondary targets Think this is also from T.C. We discussed it when it was introduced. There were concerns about if this was needed.

Question: Any comments on this?
Answer: Lots of shades of meaning. Hard to see this being useful in a meaningful way. Maybe we need an advocate for this to attend a meeting.

Question: In ontology - a primary and secondary?
Answer: Assumed it means you observed an object and something else was in the frame.
~ Sometimes you mean to see two things. Lots of distinctions in different situations.
~ It happens a lot where you capture more than originally intended. Will set up a meeting with the requester. (**Action Item - Joe**)

CCB#28 - As a user, I want validate to allow delimited tables with whitespace-only numeric fields

Question: Advocates for this? Unclear what the need was and if it's proposing something that is allowed should be a validate ticket.
Answer: **Action Item - Joe** will take a look at this since he was involved in putting together the DSV spec.

CCB#12 - Document Nuances for Reusability of Generic LDD Classes/Attributes

This is being subsumed by the latest submissions, CCB#52 (Enable cross-namespace validations in sch files) and CCB#53 (Improve cross-namespace validations in sch files).
~ We should withdraw this.

Action Item - Steve will take a look first to see if CCB#12 is covered by CCB#s 52 and 53. If so, it can be closed.

Someone looked at this a long time ago.
~ That's why we are doing this today. To see if some SCRs can be closed or moved on. The action item is to take a look at this.
~ It seems like a duplicate. Has comments from Jordan.
~ The action is for Steve to decide if this can be closed.

CCB#21 - Provenance LDD Using PROV-DM

This is critical. Would like to change the title. It has evolved quite a bit.
~ Someone thinks it's desired to do that.
Action Item - Steve will update the SCR title.
~ Joe will update the agenda. (**Action Item - Joe**)

Put a comment so people will get an announcement. (**Action Item - Steve**) Edits don't trigger an alert, but comments do.
~ That makes sense.
~ Something we did work on and need to come back to.

CCB#22 - Display Settings not required for images

Question: Does anyone know the status of this?
Answer: Trent hasn't mentioned it. John will ask him. (**Action Item - John**)

CCB#23 - XML/schema based files as archival data

We haven't looked at this in awhile, but think it's still a concern.

Question: Any comments on this? It was Jira CCB-211.
Answer: Mike had a lot of perspective on this. It's why we're using LBL file extension. Would make sense to finish this. Jesse made a presentation on reasonable restrictions.
~ Sounds like an action item is to reach out to Jesse. Joe will do that. (**Action Item - Joe**)

Tier 2 Issues

Product Metadata Errata

Think there was a suggestion this could be dropped.

Parking lot

There are three in the Parking lot:

- JIRA-CCB-326 - Inconsistent Direction on How to Use Discipline Facets
- RS: 259/281 (see new attachment to JIRA-CCB-281) (on hold)
- JIRA-CCB-364 – Disallow bundles referencing collections by LID

Someone is unclear why the second one would be a good idea.
~ Someone else thinks the plan was LIDVID only.

Question: But why is that a good thing?
Answer: That's what PPI does. Provides a record of each product version included in a bundle, but other ways of doing that. It needs a driver. Think it was Jordan advocating for this. Joe will check with him. (**Action Item - Joe**)

It never got migrated to GitHub. Still on Jira.
~ Joe will talk to Jordan.

More Discussion on CCB#52 (Enable cross-namespace validations in sch files) and CCB#53 (Improve cross-namespace validations in sch files)

As far as I know, nothing in the IM precludes a DP from writing schema or schematron specific to a bundle. It's why we have a schema collection. Don't see why that would be a problem.
~ Someone tested and verified with the Mission Commons and MAVEN LDDs. Think the issue is there's no way to do it with LDDTool. Think it's not that you can't do this, but that we need LDDTool to know external context for the rules to apply.
~ LLDTool is a tool, not the model.
~ Depends on Ingest LDD, which is controlled by the model, so yes, a tool issue.
~ Both are right. The discussion needs to continue.

Action Item - Lyle and Joe post comments about this in GitHub.

Our next meeting is January 23.
Ethan gave Joe a reminder that we need to discuss NAN and INF.

title: DDWG Notes 2025-01-23
layout: default
date: 2025-01-23

January 23, 2025
Notes by Debra Kazden

Known Attendees: A. Bailey, J. Blalock*, B. Hirsch*, L. Huber*, S. Hughes*, R. Joyner, D. Kazden, P. Lawton, J. Mafi*, E. Schaefer*, B. Semenov*, M. Seritan and M. Tiscareno* (*'s indicate people who voted during this meeting - if applicable)

DDWG Agenda

(Included in email sent January 22, from J.Mafi, Agenda for 2025 Jan 23)

Apologies for the late agenda. We will have a DDWG telecon this Thursday, January 23 at 9:30 PST (12:30 EST, 10:30 MST). We will cover the following issues:
Brief reports on action items from last meeting:

- CCB#51 – J. Mafi to set up meeting with requestor (T. Cornet)
- CCB#45 – S. Hughes to update SCR
- CCB#40 – J. Blalock to check with B. Deen on withdrawing SCR
- CCB#39 – S. Hughes to update SCR
- CCB#28 – J. Mafi to research
- CCB#23 – J. Mafi to contact J. Stone/M. Drum/T. Hare
- CCB#22 – J. Blalock to contact T. Hare
- CCB#12 – S. Hughes to verify whether this is subsumed by CCB#52 and CCB#53
- CCB#45 – Add reference types values for *-to-agency: We will discuss and vote on this SCR.
- CCB#3 & CCB#4 CCB Feedback: We will review the CCB's feedback on these two SCRs and discuss approaches to addressing the issues raised by the CCB. I will resend the email from Mia Mace immediately following this agenda.
  For the latest version of this agenda, see the Google Doc

Front-matter/What's new

1. Submitted to the CCB:
   - CCB#13 – Ensure file name and directory path name adhere to SR
2. New issues in GitHub:
   - CCB#54 – Add modification history for classes and attributes
   - CCB#55 – Funding Acknowledgment class enhancement (funding source and funding acknowledgment text)
3. Ready for a vote (None)
4. Build 15.1 freeze date: 2/27/2025. The following SCRs are queued for implementation: CCB#13, CCB#48, CCB#49.
5. Withdraw by author:

- CCB#11 – Add a Funding Information class to the Identification Area
  (Discussed)

Tier 1 issues

1. CCB#53 – Improve cross-namespace validations in sch files
2. CCB#52 – Enable cross-namespace validations in sch files
3. CCB#51 – Add an optional attribute in the Target Identification class to identify Primary/Secondary targets
4. CCB#50 – Clarify Ambiguous Author Information in Product Document
5. CCB#46 – Disallow multiple versions of the same LID from occurring in a single collection
6. CCB#45 – Add reference types values for *-to-agency
7. CCB#44 – Issues caused by the failure to expose pds: attributes that are referenced in an LDD.
8. CCB#41 – Create a way for classes or attributes to be exposed to other LDDs but not to data providers
9. CCB#40 – Change cardinality in Product Bundle to better support PDS3/4 hybrid bundles
10. CCB#39 – Fixed problem where LDDTool does not include all rules from IngestLDD
11. CCB#37 – Add List Author, List Editor, and List Contributor to the Modification Detail class
12. CCB#36 – Add Mission Phase to the Investigation class
13. CCB#35 – Add Alias Index and Acronym Index to the Investigation class
14. CCB#32 – Add Host type to Observing System Component
15. CCB#31 – Ensure that dictionary versions are specified in labels
16. CCB#28 – As a user, I want validate to allow delimited tables with whitespace-only numeric fields
17. CCB#23 (JIRA-CCB-211) – XML/schema based files as archival data
18. CCB#22 (JIRA-CCB-164) – Display Settings not required for images
    - 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
19. CCB#21 (JIRA-CCB-367) – Provenance LDD Using PROV-DM
20. CCB#20 (JIRA-CCB-358) – Instrument Package Context Products
    - Tiger Team: M. Tiscareno (lead), B. Deen, B. Hirsch, S. Hughes, R. Joyner, J. Mafi, E. Schaefer.
21. CCB#13 – Ensure file name and directory path_name adhere to SR
22. CCB#12 – Document Nuances for Reusability of Generic LDD Classes/Attributes
23. CCB#5 – Move to V.2.0.0 for June 2025 Release
24. CCB#4 – Backwards Incompatibility policies
    - Presentation: Backwards Compatibility
25. CCB#3 – Adopt Semantic Versioning with version 2 transition
    - The CCB returned this, requesting that it be resubmitted as part specific proposal to transition to IM V2.0.0.
      (Some Discussions)

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!

- JIRA-CCB-326 - Inconsistent Direction on How to Use Discipline Facets
- RS: 259/281 (see new attachment to JIRA-CCB-281) (on hold)
- JIRA-CCB-364 – Disallow bundles referencing collections by LID **(Not Discussed)**

Historical links:
Export of history: https://github.com/NASA-PDS/PDS4-CCB/tree/main/data
Backup: https://nasa-pds.atlassian.net/jira/software/c/projects/CCB/issues (limited seats)
Next meeting: 6 Feb 2025

DDWG

Front Matter

There are two new issues:

CCB#54 - Add modification history for classes and attributes

CCB#55 - Funding Acknowledgment class enhancement (funding source and funding acknowledgment text)

CCB#54 is from RMS. They have been moving things around in LDDs, incrementing version numbers when they make changes. Annoyed there is no way to record what the changes are. May be that this doesn't need to be required, but would like it to be allowed as a child.

Question: This if for the ingest LDD?
Answer: Yes.

CCB#55 is from T.C. and is connected to the funding acknowledgment class. We will look at this in the future.

CCB#13 (Ensure file name and directory path name adhere to SR) has been passed by the CCB. It has been moved to implementation for the next release.

The next Freeze date is February 27, 2025. We will try to wrap up some of the SCRs as that date approaches.

~ A question was asked about what that date actually means. Unclear on if it is when things go to the CCB. We need a reasonable amount of time to get things done and for them to go to the CCB, and a week or two for them to respond. Then after that we should leave things to the next build.
~ At some point a deadline is a deadline.
~ Waiting for everything in the queue for the last build was problematic.
~ We need a clear deadline.

Action Item - Joe will talk to Jordan to get clarification on this issue.

CCB#11 (Add a Funding Information class to the Identification Area) has been withdrawn by the author. It was removed from our issues list.

Action Item Reports

CCB#51 (Add an optional attribute in the Target Identification class to identify Primary/Secondary targets) - this is from T.C. Joe made arrangements for him to attend the DDWG meeting February 20.

CCB#45 (Add reference types values for *-to-agency) - Steve made the updates and then started working on new updates from Joe and Ron. Updates continue. Will be discussed more today.

CCB#40 (Change cardinality in Product Bundle to better support PDS3/4 hybrid bundles) - John tried to check in with Bob but hasn't heard back. Likely because of fires. Will check in with him again.

CCB#39 (Fixed problem where LDDTool does not include all rules from IngestLDD) - Steve updated the SCR. Added comments.

CCB#28 (As a user, I want validate to allow delimited tables with whitespace-only numeric fields) Joe has no progress to report yet.

CCB#23 (XML/schema based files as archival data) - Joe emailed Jessie, Trent and Mike about this. No responses yet.

CCB#22 (Display Settings not required for images) - John got confirmation from Trent that he will start working on this again.

CCB#12 (Document Nuances for Reusability of Generic LDD Classes/Attributes) - Steve said that he believes that this is subsumed by the new SCRs, CCB#s 52 and 53. Was unclear if he should withdraw CCB#12 or if there is a process for it to be subsumed by the new SCRs. Joe said that we want information from CCB#12 and can create a dependency. Steve agreed that sounded

good. Joe said we do not want to withdraw it. We will figure something out. We will keep working on this and move forward.

Tier 1 Issues

CCB#45 - Add reference types values for *-to-agency

There has been a lot of discussion and edits for this one. Title has been updated to refer to agency and node. The Requested Changes have been updated to reflect the decisions made by DDWG votes. The attachments have been updated too. A note was added to explain the DDWG considerations in the approach to the implementation because the Problem Statement doesn't exactly match the implementation now. Think it's important to retain the old Problem Statement to make sure it was addressed.

Question: Does anyone think that should be edited?
Answer: Someone agrees. Has been wondering about that too. Thinks this is appropriate and reasonable. Likes keeping the history of the original statement.

Question: Any other comments?
Answer: (Silence)
~ We will keep the proposed statement as it was originally written, with a note on why the implementation is what it is.
~ Good points were raised. Also, the original Proposed Solution would be left and Requested Changes are what we do. Been moving to update the Requested Changes and Proposed Solution. Unsure why we need both.
~ We can look through the edits and revise to original state. Makes sense to retain the original Proposed Solution along with the DDWG notes. Maybe another guideline is to try to retain the original Requested Changes and Proposed Solution.

Question: Any comments?
Answer: (Silence)

This one is easier to deal with because the proposer is a member of the DDWG, but we do have SCRs that aren't from members, so it's important not to change the target after firing the arrow. We need to retain the original and explain any modifications.

Action Item - Joe will change the Proposed Solution back to the original. Implement is described in the Requested Changes.

Question: Any comments before we vote?
Answer: (Silence)

The Vote to approve CCB#45:
ATMOS - Yes
EN - Yes
GEO - Yes
IMG - Yes
IPDA - Not Present
NAIF - Abstain
PPI - Yes
RMS - Yes
RS - Not Present
SBN - Yes

This will be sent to the CCB.

CCB#3 (Adopt Semantic Versioning with version 2 transition) and CCB#4 (Backwards Incompatibility policies)

The email from the CCB was sent to the DDWG. Hopefully, everyone read it. Would like to hear what people thought about what the CCB is asking for. They want a detailed plan, a road map.

Question: Does anyone have any ideas?
Answer: (Silence)
~ The CCB mentions that maybe this could be discussed at the next F2F. We would need to do some work ahead of time, like with the DSMP, but don't think we want to do that for every major version. To create a procedure it might be worth that type of effort. Could form a Tiger Team to start the work on this.

Question: Thoughts?
Answer: (Silence)
~ If none then this will be a short meeting. Would like volunteers to work on an outline of a plan that we can throw darts at. Eventually we will need to work with the CCB and the Document Writing team to make sure we are developing a procedure that is feasible.

Question: Thoughts? Volunteers?
Answer: Ron volunteered to be on the Tiger Team.
Another Question: Anyone else?
Answer: (Silence)

Okay. Ron and Joe will have a lot of power. Excellent. They will start working on this and bring it to the group. The CCB asked for a very specific plan including for non-backwards compatible changes and for moving to version 2.0.0 and a process for future version transitions as well.

Question: Any final comments?
Answer: (Silence)

Next meeting in two weeks.

title: DDWG Notes 2025-02-06
layout: default
date: 2025-02-06

February 6, 2025
Notes by Debra Kazden

Known Attendees: A. Bailey, J. Blalock, K. Crombie, B. Hirsch, L. Huber, S. Hughes, R. Joyner, D. Kahan, D. Kazden, A. Knight, P. Lawton, S. Loftin, J. Mafi, S. McLaughlin, E. Schaefer, B. Semenov, M. Seritan and M. Tiscareno (*'s indicate people who voted during this meeting - if applicable)

DDWG Agenda

(Included in email sent February 4, from J.Mafi, Agenda for 2025 Feb 6)

We will have a DDWG telecon this Thursday, February 6 at 9:30 PST (12:30 EST, 10:30 MST). We will cover the following issues:

1. Action item reports:
    a. CCB#51 – J. Mafi to set up meeting with requestor (T. Cornet) b. CCB#45 – S. Hughes to update SCR c. CCB#40 – J. Blalock to check with B. Deen on withdrawing SCR d. CCB#39 – S. Hughes to update SCR e. CCB#28 – J. Mafi to research f. CCB#23 – J. Mafi to contact J. Stone/M. Drum/T. Hare g. CCB#22 – J. Blalock to contact T. Hare h. CCB#12 – S. Hughes to verify whether this is subsumed by CCB#52 and CCB#53 i. JIRA-CCB-364 – J. Mafi to contact J. Padams j. "Freeze Date" definition – J. Mafi to discuss with J. Padams
2. CCB#45 – This SCR was rejected by the CCB (see followup email). R. Joyner has updated the Schematron and test labels to address issues pointed out by A. Raugh. We will discuss the CCB's feedback and what (if any) the DDWG response should be.
3. CCB#52 & CCB#53 CCB Feedback: Please be ready to discuss these two, related SCRs.
4. Next meeting: The "Freeze Date" (Feb 27) is approaching quickly. I would like to consider holding an extra meeting next week (Feb 13), if there is a possibility that we have any SCRs ready for the DDWG by doing so. We will discuss this briefly at the end of the telecon.

Front-matter/What's new

1. Submitted to the CCB (none)
2. New issues in GitHub (none)
3. Ready for vote (none)
4. Build 15.1 freeze date: 2/27/2025. The following SCRs are queued for implementation: CCB#13, CCB#48, CCB#49.
5. Withdraw by author: CCB#11 – Add a Funding Information class to the Identification Area **(Discussed)**

Tier 1 issues

- CCB#55 – Funding Acknowledgement class enhancement (funding source and funding acknowledgement text)
- CCB#54 – Add modification history for classes and attributes
- CCB#53 – Improve cross-namespace validations in sch files
- CCB#52 – Enable cross-namespace validations in sch files

- CCB#51 – Add an optional attribute in the Target Identification class to identify Primary/Secondary targets
- CCB#50 – Clarify Ambiguous Author Information in Product Document
- CCB#46 – Disallow multiple versions of the same LID from occurring in a single collection
- CCB#45 – Add reference types values for *-to-agency
- CCB#44 – Issues caused by the failure to expose pds: attributes that are referenced in an LDD.
- CCB#41 – Create a way for classes or attributes to be exposed to other LDDs but not to data providers
- CCB#40 – Change cardinality in Product Bundle to better support PDS3/4 hybrid bundles
- CCB#39 – Fixed problem where LDDTool does not include all rules from IngestLDD
- CCB#37 – Add List Author, List Editor, and List Contributor to the Modification Detail class
- CCB#36 – Add Mission Phase to the Investigation class
- CCB#35 – Add Alias Index and Acronym Index to the Investigation class
- CCB#32 – Add Host type to Observing System Component
- CCB#31 – Ensure that dictionary versions are specified in labels
- CCB#28 – As a user, I want validate to allow delimited tables with whitespace-only numeric fields
- CCB#23 (JIRA-CCB-211) – XML/schema based files as archival data
- CCB#22 (JIRA-CCB-164) – Display Settings not required for images
- 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#21 (JIRA-CCB-367) – Provenance LDD Using PROV-DM
- CCB#20 (JIRA-CCB-358) – Instrument Package Context Products
- Tiger Team: M. Tiscareno (lead), B. Deen, B. Hirsch, S. Hughes, R. Joyner, J. Mafi, E. Schaefer.
- CCB#13 – Ensure file name and directory path name adhere to SR
- CCB#12 – Document Nuances for Reusability of Generic LDD Classes/Attributes
- CCB#5 – Move to V.2.0.0 for June 2025 Release
- CCB#4 – Backwards Incompatibility policies
- Presentation: Backwards Compatibility
- CCB#3 – Adopt Semantic Versioning with version 2 transition
- The CCB returned this, requesting that it be resubmitted as part specific proposal to transition to IM V2.0.0.

(Brief Discussion related to Action Items and Planning)

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!

- JIRA-CCB-326 - Inconsistent Direction on How to Use Discipline Facets
- RS: 259/281 (see new attachment to JIRA-CCB-281) (on hold)
- JIRA-CCB-364 – Disallow bundles referencing collections by LID **(Not Discussed)**

Historical links:
Export of history: https://github.com/NASA-PDS/PDS4-CCB/tree/main/data
Backup: https://nasa-pds.atlassian.net/jira/software/c/projects/CCB/issues (limited seats)

Next meeting: 13 Feb 2025

DDWG

Front Matter

No new SCRs were submitted this week.

CCB#45 (Add reference types values for *-to-agency) has been rejected by the CCB. We will discuss this more today.

Reports on Action Items
This includes the ones that someone requested be added to the list.

CCB#51 (Add an optional attribute in the Target Identification class to identify Primary/Secondary targets) - Joe - Talked to T.C. He is coming to the Feb 20 meeting.

CCB#28 (As a user, I want validate to allow delimited tables with whitespace-only numeric fields) - Joe - Nothing to report yet.

CCB#23 (XML/schema based files as archival data) - Joe - Heard from Jesse. They found a work around. Still an important issue. Unsure if it should be kept as an SCR for now.

JIRA-364 (Disallow bundles referencing collections by LID) - Joe discussed this with Jordan. This is the one to require the use of LIDVIDs in bundle, Lower priority, but we should keep it in the list.
Freeze date definition - Joe - discussed this with Jordan. The freeze date is the date that things need to be passed by the CCB, so the DDWG date is realistically two weeks earlier. That makes our deadline Feb 13, which is very soon.

Question: Any reports on the other action items?
Answer: Steve said no changes for his, except CCB#45. (Steve's are CCB#s 39 (Fixed problem where LDDTool does not include all rules from IngestLDD) and 12 (Document Nuances for Reusability of Generic LDD Classes/Attributes))

John reported that he spoke to Bob. CCB#40 (Change cardinality in Product Bundle to better support PDS3/4 hybrid bundles) can be withdrawn.
~ Bob needs to email Ron and cc Joe and ask that it be withdrawn.

CCB#22 (Display Settings not required for images) - John - Trent is working on it, but he is concerned about backwards incompatibility issues.

CCB#20(Instrument Package Context Products) and CCB#32 (Add Host type to Observing System Component) - Tiger Team is actively working.

CCB#31 (Ensure that dictionary versions are specified in labels) - Matt and Ron will begin discussing this.

CCB#21 (Provenance LDD Using PROV-DM) - This is not on the agenda for today, but we will begin discussing it after we resolve the issues for CCB#45. This is more work than we can finish before the freeze.

CCB#54 (Add modification history for classes and attributes) - Matt said he hadn't heard any objections and wanted to know if we could just do it.
~ We can move forward and start discussing the comments. It's in the queue.

Action Item - Joe is going to move the action items to a sub-category for the agenda. Many are closed now.

CCB#45 - Add reference types values for *-to-agency

This was rejected by the CCB. The CCB response was sent to the DDWG group this morning. Anne's comment was regarding complicated implementation, Lynn wanted use cases. Definitely more work to do if we want to get this past the CCB.

Question: Any thoughts?
Answer: Someone went through the responses. Tried to resolve them in his mind. Not seeing show-stoppers, but there are issues this team needs to resolve. Had understood that these would be added for use in product context and collection and maybe bundle.
~ The original Problem Statement mentioned more. There were concerns with use cases for some of them. Discussions led to thinking better use for context products.

Someone is seeing this for the first time.
~ Not everyone has received the CCB email yet.
~ Joe is re-sending it now.
~ People just got it.

To summarize, someone's take is that Anne complained about the implementation and Lynn questioned why we need this at all. Anne asked why the shorter list and made the point that's not how we have done things before.
~ Someone disagrees with that. We definitely did something to something else context products. The weird one was data, which is represented by product observational. Weird because no data context product. System side is asking it be broadened to product, which is the name of the class. Moving from data to product in this case. So, disagrees with Anne. Think we are following initial guidelines.
~ A very old argument, whether we should be generic or specific.
~ Jordan says we should be more generic, which was a surprise.
~ Anne has a good point we should be more consistent.
~ So would say product context to agency or node.
~ No, frankly that information is elsewhere. It doesn't need to be provided in the new type.

~ Agree, but Jordan is not parsing those. Can usually tell by name.
~ It's important to look at usage for the last decade. If ops is not finding the types useful then we should transition to more generic types, but that's a bigger discussion than this SCR.
~ It's been confusing that Jordan is changing to more generic. Anne hasn't been hearing those conversations.

Question to Ron: Another concern was Anne not seeing any use for node to context product references. Ron, have you addressed this in the updates?
Answer from Ron: Waiting for the final decision, but yes.
Another Question: Node to instrument might make sense, but not all of them. Any idea why we would need product to node for context products?
Answer: (Silence)

Someone proposes we vote on whether permissible values should be restricted to product and node.
~ Wait. We need way more discussion on the whole SCR before we vote on anything. My very first comment on the SCR was that I don't see what the use cases are. Believe it's the EDWG asking for something we don't need.
~ A person is on the EDWG and says we didn't ask for this.
~ Think it was required behind the scenes, but no one needs this. No real use case for this to ever be necessary.
~ The only use case someone else sees is if a user submits a question or request about a specific bundle to EN, so EN can forward it to the node.

Someone is sympathetic. Was irritated by the EDWG. Doesn't think they should have a say on how the IM or system works. Understanding is with the system people that the IM drives the system, so we capture all the information to drive this, but when system needs a something the question is if we can provide it. For this, Jordan needs it to drive the landing pages. Not from the EDWG. Jordan said he needs this. Thought about provenance. Sees this as a provenance or archive issue. Interested in requirements Jordan has.
~ If that's true it needs to be written in the SCR. Not just a vague statement. What we wanted when we asked for use cases. Need to make it more robust. Then maybe we can get it through.

The real problem depends on how far you want to extend it. It won't work without retrofitting everything in the system.
~ Should have been there from the start.
~ Think we rejected it.
~ It makes sense to someone that some context products should be able to (reference?) node to agency who archives them or products to who funds them, but if we want to argue that someone would download products without collection products we would need a schematron pointer to avoid duplicate information.
~ Not for individual products.
~ Then we should have bundle or collection to node/agency.
~ Back to general versus specific.
~ We voted on product to avoid the long list of all possibilities.
~ Someone disagrees with that.

Question: Would we allow people to use it if they came up with a product they wanted this for?

Answer: They would need to come to the DDWG.
~ That's why a person wanted use cases.
~ Maybe it doesn't matter. We need to consider why we would preclude things.
~ The question is why wouldn't we.
~ We can't close every loop hole. Trying to understand why so important.
~ The only legitimate use cases someone sees are for the system.
~ Agreement.

The next thing is that along the way Jordan said we can't include target because no one owns it, but the sample valid product was a target product. The one thing we said we would exclude was the example.
~ That makes the point that people might do things without good guard rails.
~ We should add documentation. Worried about the schematron rules and having an impossible set of requirements to manage. Can write up the guidance in the DPH.
~ The point in bringing it up is that it should never have gone to the CCB. We failed.
~ Someone agrees. Push-back from the CCB is fine.

Question: Was that cited by the CCB?
Answer: Lynn certainly did. Think Anne did too.
~ Clearly we need some additional work on this. Will reach out to Jordan for specific use cases on this. **(Action Item - Joe)**

Don't think it's our responsibility to tell the system what they need. We just provide what they ask for.
~ We need clarification on what they actually need.
~ And why.
~ Don't think anyone here has ever had to do a landing page. You have to constantly make changes.
~ We want use cases so we can try to implement in a way that will work for them.
~ In the EDWG, also want clarification. Just because something might help with the website doesn't mean it has to go in the archive. Just posted a comment proposing collection to agency, collection to node, bundle to agency and bundle to node, to agency for only those context products that are sensibly owned or funded by an agency.
~ We need use cases before we vote.
~ Jordan wanted us to keep it general. I was the one who proposed product.
~ Thought he prefers generic.
~ He does.
~ We can come back to this after Joe checks with Jordan and then vote on this. Agrees that text for the DPH should also be part of this.

These are just names for links. All Jordan really is asking for is the link. Don't think he cares what it is called. He wants product A to product B. Doesn't care what the name is.
~ We still need to know why he needs the link. Unclear why anyone would do some of these searches.
~ Products with DOIs would need landing pages.
~ Someone can imagine people wanting to know what products are curated by a specific node/agency.
~ Someone else can't. Bundle maybe.

_____

We have ten minutes left. We only have about a week to get things to the CCB to resolve before the freeze.

Question: Do we want to meet next week? Is there something we can finish?
Answer: Someone would like to meet more often. A lot on the agenda that isn't getting finished.
Another Question: That's fair. Is there any objection to meeting next week?
Answer: (Silence)
~ It will conflict with the Tiger Team meeting.

Question: Why this hard deadline for a build? Seems like the last build just came out.
Answer: It did. How it works. They need time to implement and test.
~ The real question is if we need two builds a year.
~ Jordan is trying to do a build when we want to, trying to move past this.
~ The IM has barely come out.
~ We really screwed up last time. Almost all the implementation happens after the deadline. Huge problem.
~ It was more than a month past the freeze date.
~ Not efficient. We don't even know what's in the IM model.

Question: What is driving this? A lot of time. Why do that?
Answer: A fair question. Think we have three queued for implementation this time. Don't see any others that will make it.
~ Someone thinks there are several we could do if we had a few more weeks. There are a bunch of loose ends, so not sure why we need a new build.
~ We need Jordan to answer. It's a system management question. Not sure.

Someone's opinion on the last one is that there is more content there than in a long time. Prefers builds with small changes. One a year would be horrendously difficult.
~ The question is why the hard schedule. We could do more without the hard deadline. It's what we have always done, but unsure what's keeping it to that. we can ask Jordan for clarification and if there is any flexibility in the build dates. **(Action Item - Joe)**

Don't think there's much we can do for this next build. Think our deadline for this current build is next week. Agree with what everyone has said. The last freeze dragged on.

We are out of time. We will discuss CCB#52 and CCB#53 next time.

Action Item - Joe will send a request to Vivian to set up a meeting next week.

Question: Any last thoughts?
Answer: For 52 and 53, unless we think they can be implemented right away, we should work on things that have a chance to pass. Suggest CCB#s 54, 31, 20 and 32.
~ CCB#52 does have a chance, CCB#53 is more complicated.
~ Someone agrees about CCB#52 and that we should work on things we have a chance to pass.
~ Trent is currently working on CCB#22. He's concerned about backwards incompatibility issues.
(Action Item - John will check in with Trent.)

Here is the list again for next week: CCB#s 52, 54, 31, 20 and 32.

~ Yes.

The Tiger Team members can stay online now and try to figure out our next meeting time.
~ Agreement.

Next DDWG Meeting February 13.

title: DDWG Notes 2025-02-13
layout: default
date: 2025-02-13

February 13, 2025
Notes by Debra Kazden

Known Attendees: A. Bailey, J. Blalock, K. Crombie, B. Deen, B. Hirsch, L. Huber, S. Hughes, R. Joyner, D. Kazden, A. Knight, P. Lawton, S. Loftin, J. Mafi, S. McLaughlin, E. Schaefer, B. Semenov, M. Seritan and M. Tiscareno (*'s indicate people who voted during this meeting - if applicable)

DDWG Agenda

(Included in email sent February 11, from J.Mafi, Agenda for 2025-02-13)

We will have a DDWG telecon this Thursday, February 13 at 9:30 PST (12:30 EST, 10:30 MST). For this meeting we will consider a series of the more straightforward SCRs, that there may be a possibility of getting through the full DDWG and CCB approval process prior to the freeze date (27 Feb). We will cover the following issues:

1) Action item reports.
2) CCB#52 – Enable cross-namespace validations in sch files
3) CCB#54 – Add modification history for classes and attributes
4) CCB#20 (JIRA-CCB-358) – Instrument Package Context Products
5) CCB#31 – Ensure that dictionary versions are specified in labels

Front-matter/What's new

1) Submitted to the CCB (none)
2) New issues in GitHub (none)
3) Ready for vote (none)
4) Build 15.1 freeze date: 2/27/2025. The following SCRs are queued for implementation: CCB#13, CCB#48, CCB#49.
5) Withdraw by author: CCB#40 – Change cardinality in Product Bundle to better support PDS3/4 hybrid bundles **(Brief Discussion)**

Action Items

a) CCB#39 – S. Hughes to update SCR
b) CCB#28 – J. Mafi to research
c) CCB#12 – S. Hughes to verify whether this is subsumed by CCB#52 and CCB#53
d) CCB#20 & CCB#32 – Tiger team formed and working the issue
e) CCB#31 – Ron and Matt working the issue

Tier 1 issues

- CCB#55 – Funding Acknowledgement class enhancement (funding source and funding acknowledgement text)
- CCB#54 – Add modification history for classes and attributes
- CCB#53 – Improve cross-namespace validations in sch files
- CCB#52 – Enable cross-namespace validations in sch files

- CCB#51 – Add an optional attribute in the Target Identification class to identify Primary/Secondary targets
- CCB#50 – Clarify Ambiguous Author Information in Product Document
- CCB#46 – Disallow multiple versions of the same LID from occurring in a single collection
- CCB#45 – Add reference types values for *-to-agency
- CCB#44 – Issues caused by the failure to expose pds: attributes that are referenced in an LDD.
- CCB#41 – Create a way for classes or attributes to be exposed to other LDDs but not to data providers
- CCB#40 – Change cardinality in Product Bundle to better support PDS3/4 hybrid bundles
- CCB#39 – Fixed problem where LDDTool does not include all rules from IngestLDD
- CCB#37 – Add List Author, List Editor, and List Contributor to the Modification Detail class
- CCB#36 – Add Mission Phase to the Investigation class
- CCB#35 – Add Alias Index and Acronym Index to the Investigation class
- CCB#32 – Add Host type to Observing System Component
- CCB#31 – Ensure that dictionary versions are specified in labels
- CCB#28 – As a user, I want validate to allow delimited tables with whitespace-only numeric fields
- CCB#23 (JIRA-CCB-211) – XML/schema based files as archival data
- CCB#22 (JIRA-CCB-164) – Display Settings not required for images
- 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#21 (JIRA-CCB-367) – Provenance LDD Using PROV-DM
- CCB#20 (JIRA-CCB-358) – Instrument Package Context Products
- Tiger Team: M. Tiscareno (lead), B. Deen, B. Hirsch, S. Hughes, R. Joyner, J. Mafi, E. Schaefer.
- CCB#13 – Ensure file name and directory path name adhere to SR
- CCB#12 – Document Nuances for Reusability of Generic LDD Classes/Attributes
- CCB#5 – Move to V.2.0.0 for June 2025 Release
- CCB#4 – Backwards Incompatibility policies
- -Presentation: Backwards Compatibility
- CCB#3 – Adopt Semantic Versioning with version 2 transition
- The CCB returned this, requesting that it be resubmitted as part specific proposal to transition to IM V2.0.0.
  (Some Discussions)

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!

- JIRA-CCB-326 - Inconsistent Direction on How to Use Discipline Facets
- RS: 259/281 (see new attachment to JIRA-CCB-281) (on hold)
- JIRA-CCB-364 – Disallow bundles referencing collections by LID **(Not Discussed)**

Historical links:
Export of history: https://github.com/NASA-PDS/PDS4-CCB/tree/main/data
Backup: https://nasa-pds.atlassian.net/jira/software/c/projects/CCB/issues (limited seats)

Next meeting: 20 Feb 2025

DDWG

Front Matter

There are a number of SCRs to review today. There have been developments on some, even the need for some of them may have changed.

Tier 1 issues

CCB#54 - Add modification history for classes and attributes

Basically, the request is to add a modification history class to DD class and DD attribute class so that changes in the ingest LDD file can be tracked.
~ Someone suggested we use comment. That's okay if it's easier for EN.
~ That would mean updating the documentation.
~ Comment is already allowed for DD class, but not DD attribute, so an SCR is still needed.

There is something at the file level. If we want it at the individual attribute level that can be useful in bigger LDDs.
~ Hard to keep it all in a change log. Would be better to capture in that class or attribute.

The SCR has been updated.

Question: Any comments?
Answered with Another Question: Can we update the Requested Changes, if it's already in one?
Answer: Yes.
~ Someone said it is already in attribute, not class.
~ A person is trying it in Oxygen now, Oxygen doesn't like it.
~ Schema has attribute, not class.
~ We should also have it in DD rule for completion, but they don't have version.
~ Should also put the definition for comment that we have been using all over.
~ That will be inherited.
~ Okay.

Someone is making the SCR changes in real-time.
~ The title should be updated too.
~ Done.

Question: Other thoughts? Do we need a test schema or test labels or can we go forward like this?
Answer: There is a change to schema, so there will need to be a test.
~ And a proper TA.
Another Question: Any objections to being able to submit this to the CCB if we get this done?
Answer: (Silence)

The freeze is in two weeks. We could try to get this together and get it reviewed and do an evote early next week. Think this is fairly straight forward. Not a huge impact. Would be nice to get this into the next build. We will try to evote Tuesday. Hopefully that is enough time for the CCB to review it and vote.

Question: Comment is currently optional, assuming it will remain optional?
Answer: Yes. Adding that to the SCR in real-time now.
~ Excellent.

CCB#31 - Ensure that dictionary versions are specified in labels

This one was discussed about a month ago. There was a suggestion on documentation. Think this is becoming a documentation fix. A change so DPs and users know where to look in labels to find what dictionaries were used.

This fell off the table. It hasn't been worked on yet.
~ Then there is nothing to do for this today.

CCB#52 - Enable cross-namespace validations in sch files

Someone is going to speak from their car. There are two issues, this and CCB#53 (Improve cross-namespace validations in sch files). CCB#52 is the easy one, so it can be a rule, but it turns out, from the comments yesterday that there is already a way to do it in LDD ingest. Just need to add a few lines. It seems to work, so not sure we need this since there is that other mechanism.
~ Another person wasn't aware it was already possible.
~ It's useful to know that.
~ People are glad Jesse knew how to do it and described it for us (in GitHub).

Maybe we just need a document change to say how to do this. Don't think we need any substantive changes.
~ Good point, but not sure documents talk about LDDs much.
~ Don't know if they do, but maybe they should. CCB#53 is easier way to write the rules (Breaking up while driving…) With CCB#53, a better way of doing this. If we are contemplating doing that we could hold off on CCB#52 for now.

We will hold CCB#52 open for now, but it becomes a documentation issue.
~ Okay. Maybe we should discuss CCB#53.
~ We are trying to hold today for things we can finish today.
~ It's ready for a TA.
~ We can put CCB#53 on the agenda for an upcoming meeting.

Question: Is it worth doing a TA now to inform the discussion, to get ahead of the game?
Answer: People haven't looked closely yet to understand the issues.
~ On CCB#52, the question is if we want to keep it open, but the focus is changed, so we should change the title.
~ Yes, but not sure of the details yet.
~ It could be changed from enabling to documenting.
~ Yes.
~ The author can change it. **(Action Item - Bob D.)**

Question: Any other comments?
Answer: (Silence)

CCB#20 (JIRA-CCB-358) – Instrument Package Context Products

The Tiger Team had a fairly extended meeting yesterday. There are some questions to answer.
~ (Matt is sharing his screen) This has been discussed for a long time. It goes back to Jira 358. The original use case was New Horizons Ralph, with instrument group LEISA and MVIC. We didn't have a place in our hierarchy for that. Since then, more use cases have come to light. We need the hierarchy to be flexible. Not all of the discussion is captured in GitHub. One idea was to have several levels. It's something we can do, but it gets very complex, especially with all the pointers. On the other hand, we could just have instrument group point to instrument group. It could be recursive. But people didn't like that because we lose the hierarchy. People want to know that so people would know if Ralph is a child of New Horizons and not the other way around. What the Tiger Team is now suggesting - could name every level versus a flexible recursive approach. Suggest a compromise - extra levels, instrument group and instrument component. It's clear about relationships in context products.

If you have a group with a group, label doesn't know the relationship. Understand that's in the context product implemented by type names. Documentation would need them listed with parent and child relationships so they are implied in the label. Requested Changes would be to add instrument group and instrument component. SCR has draft definitions. Suggests that we add context to parent context and context to child context products and deprecate the previous instrument to instrument. A new approach, but not backwards incompatible. Also suggest add reference types is instrument group and is instrument component. That's the proposal from the Tiger Team, but they also want to allow for the possibility that people don't like the flexible approach, so the Tiger Team also produced a substitute motion in GitHub, called the Minority Report. It's not what the Tiger Team recommends, but it's an alternate so we don't have to sink the whole thing. the Tiger Team all came to consensus.

Question: Comments?
Answer: Someone is not sure where to start. Deprecating types makes it backwards incompatible. Full stop. Don't try to hide behind how the registry interprets it.
~ We don't have to deprecate it.
~ Whatever. By making a backwards incompatible change that makes this a hard no. A very convoluted set of solutions. A much simpler one was suggested in the past. This is far beyond the scope of what is necessary. MVIC and LEISA are instruments, Ralph is not, it's MVIC plus LEISA. We don't need all of this. Neither option here is good. We just need to point to Ralph in a way that can be found.

~ That may be all you need for Ralph, but there are other use cases. This is not a radical change, adding instrument group and component and group of group or component of component. Not a radical change.

New Horizon describes Ralph as an instrument.
~ This would allow DPs to decide about the children.
~ Someone understands.
~ Ralph is a simple example. Other examples are not simple. CADRE is a very complex example.
~ CADRE is an investigation with instrument host and group of three rovers, with a bunch of instruments and instrument groups. Several levels. Instrument hosts with instrument components.
~ Lunar Vertex is very similar. Rover and stationary part with very different instruments. CLPS examples are very complicated.
~ CLPS are a paradigm change.
~ Good point.

Back when we originally discussed this I was in the camp that said we didn't need this, but now, as a DP for LUCY, very hesitant in making this more complex, but getting into CLPS instruments might make me change my mind. Very complicated and different from what we have dealt with in the past.
~ Someone appreciates that. The way we did things in the past fits with this.
~ The context products noted the relationships, they just weren't machine actionable.
~ This would provide more clarity.

One Tiger Team member admits to pushing for the Minority Report because with like with CCB#45 think that trying to streamline reference types is a bigger issue and could increase possibility of rejection by the CCB. Don't like adding all the permissible values, but it reflects the standard approach.
~ Political expediency is reasonable, but it doesn't hurt to try.
~ Someone doesn't like rejection.

Question: Any last thoughts? Would like to get some straw-man polls on an approach.
Answer: Main thing is to have a straw-man to have EN do a TA. Should do it on both options.
Answer and Another Question: Okay. Steve, is that okay?
Answer and Another Question from Steve: Yes. Both should be discussed. Didn't attend the Tiger Team meeting. Will there be another meeting?
Answer: The Tiger Team is done.
~ At least until we get feedback from the DDWG.
~ The DDWG isn't done yet. Not ready to send this to the CCB.

The TA is not done until the DDWG agrees.

Question: If I put instrument Ralph in search, would I find it?
Answer: In product context there is a title. Can search on that, but can't wild-card.
~ If I read an old article on Ralph and go to PDS search I wouldn't find it.
~ Search would make use of aliases of instruments that are members of Ralph.
~ But we lose all of the hierarchy. Strongly against that. As far as searching - that was discussed. Think that the only practical way to support search is to have a combined search. It's not the ideal solution, but the best option I see.

~ Someone also wants to note that if someone says instrument group, that might confuse people.
~ That's something for nodes to consider. Don't think there will be a lot of retrofitting for what's been archived, but it will address complex relationships in CLPS missions.
~ Someone wants to re-emphasize that nobody is suggesting we re-do the context products. The diagram in GitHub is an example. Documentation says Ralph is an instrument. New ones would say it's an instrument group. It's less of an issue than people think.

We need to figure out what to do to unlock the TA process. Confused.
~ EN is not hearing a consensus. As soon as there is one then we can do the TA.
~ So we should do a straw-man poll to see if we can get consensus so the TA can be done.
~ There are two main approaches. Two new types of context products or specific list versus generic list of reference types.

Question: Can we take a straw-man poll now?
Answer: The DDWG is just seeing the Minority Report for the first time. We can't vote on this. We need more time.
~ Fair. We won't hold the straw-man poll now. We will do it in the upcoming weeks to give everyone a chance to think about this.
~ Sorry. We could do a straw-man poll to see if we think we are moving in the right direction.
~ Not sure that would improve anything. The Tiger Team is done until they get feedback from the DDWG.

Our time is up. We will try to get samples for CCB#54 and send that out and do an evote on Tuesday.

Someone has a last thought - feeling frustrated. EN is about to spend a bunch of resources with very little change on a new build when we have a bunch of stuff that is very close. Want to know how firm the deadline is. Want to begin a discussion on the build cycle.
~ That is above our pay grades, but we can discuss it in the future.
~ Nodes also have to put out a lot of effort. If the build is just going to be one small change we should put the breaks on.

We can discuss this again. Our time is up today.

The DDWG voted as follows:

CCB#54 - Item Passed: 7 Yes (ATM, EN, GEO, IMG, PPI, RMS, RS); 0 No; 1 Abs (NAIF); 2 NV (IPDA, SBN)

title: DDWG Notes 2025-02-20
layout: default
date: 2025-02-20

February 20, 2025
Notes by Debra Kazden

Known Attendees: J. Blalock, T. Cornet, B. Hirsch, L. Huber, S. Hughes, R. Joyner, D. Kazden, A. Knight, P. Lawton, S. Loftin, J. Mafi, S. McLaughlin, E. Schaefer, B. Semenov, M. Seritan and M. Tiscareno (*'s indicate people who voted during this meeting - if applicable)

DDWG Agenda

(Included in email sent February 19, from J.Mafi, Agenda for 2025-02-20)

We will have a DDWG telecon this Thursday, February 20 at 9:30 PST (12:30 EST, 10:30 MST). Thomas Cornet is scheduled to attend the meeting in order to discuss CCB#51. In addition, I've included CCB#55 (which he also submitted), in case there are any questions or concerns. We will cover the following issues:

1) Action item reports.
2) CCB#51 – Add an optional attribute in the Target Identification class to identify Primary/Secondary targets
3) CCB#55 – Funding Acknowledgement class enhancement (funding source and funding acknowledgement text)
4) CCB#45 – Add reference types values for *-to-agency: We will examine the feedback provided by the CCB (see forwarded email from Mia Mace of Feb 5) in rejecting this SCR and discuss how to respond to the issues identified. Please see Steve Hughes' comment of Feb. 12 with his thoughts.

Front-matter/What's new

1) Submitted to the CCB (none)
2) New issues in GitHub (none listed)
3) Ready for vote (none)
4) Build 15.1 freeze date: 2/27/2025. The following SCRs are queued for implementation: CCB#13, CCB#48, CCB#49.
5) Withdraw by author: CCB#40 – Change cardinality in Product Bundle to better support PDS3/4 hybrid bundles **(Very Brief Discussion)**

Action Items

a) CCB#39 – S. Hughes to update SCR
b) CCB#28 – J. Mafi to research
c) CCB#12 – S. Hughes to verify whether this is subsumed by CCB#52 and CCB#53
d) CCB#20 & CCB#32 – Tiger team formed and working the issue
e) CCB#31 – Ron and Matt working the issue

Tier 1 issues

- CCB#56 – Reconsider the PDS4 build cycle cadence
- CCB#55 – Funding Acknowledgement class enhancement (funding source and funding acknowledgement text)

- CCB#54 – Add modification history for classes and attributes
- CCB#53 – Improve cross-namespace validations in sch files
- CCB#52 – Enable cross-namespace validations in sch files
- CCB#51 – Add an optional attribute in the Target Identification class to identify Primary/Secondary targets
- CCB#50 – Clarify Ambiguous Author Information in Product Document
- CCB#46 – Disallow multiple versions of the same LID from occurring in a single collection
- CCB#44 – Issues caused by the failure to expose pds: attributes that are referenced in an LDD.
- CCB#41 – Create a way for classes or attributes to be exposed to other LDDs but not to data providers
- CCB#39 – Fixed problem where LDDTool does not include all rules from IngestLDD
- CCB#37 – Add List Author, List Editor, and List Contributor to the Modification Detail class
- CCB#36 – Add Mission Phase to the Investigation class
- CCB#35 – Add Alias Index and Acronym Index to the Investigation class
- CCB#32 – Add Host type to Observing System Component
- CCB#31 – Ensure that dictionary versions are specified in labels
- CCB#28 – As a user, I want validate to allow delimited tables with whitespace-only numeric fields
- CCB#23 (JIRA-CCB-211) – XML/schema based files as archival data
- CCB#22 (JIRA-CCB-164) – Display Settings not required for images
- 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#21 (JIRA-CCB-367) – Provenance LDD Using PROV-DM
- CCB#20 (JIRA-CCB-358) – Instrument Package Context Products
- Tiger Team: M. Tiscareno (lead), B. Deen, B. Hirsch, S. Hughes, R. Joyner, J. Mafi, E. Schaefer.
- CCB#13 – Ensure file name and directory path name adhere to SR
- CCB#12 – Document Nuances for Reusability of Generic LDD Classes/Attributes
- CCB#5 – Move to V.2.0.0 for June 2025 Release
- CCB#4 – Backwards Incompatibility policies
- -Presentation: Backwards Compatibility
- CCB#3 – Adopt Semantic Versioning with version 2 transition
- The CCB returned this, requesting that it be resubmitted as part specific proposal to transition to IM V2.0.0.
  (Discussed CCB#s 51 and 55)

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!

- JIRA-CCB-326 - Inconsistent Direction on How to Use Discipline Facets
- RS: 259/281 (see new attachment to JIRA-CCB-281) (on hold)

- JIRA-CCB-364 – Disallow bundles referencing collections by LID **(Not Discussed)**

Historical links:
Export of history: https://github.com/NASA-PDS/PDS4-CCB/tree/main/data
Backup: https://nasa-pds.atlassian.net/jira/software/c/projects/CCB/issues (limited seats)

Next meeting: 6 Mar 2025

DDWG

Front Matter

We have a guest today, Thomas Cornet, to discuss CCB#51 (Add an optional attribute in the Target Identification class to identify Primary/Secondary targets) specifically, and CCB#55 (Funding Ackowledgement class enhancement (funding source and funding acknowledgement text) if there is time, since he is in the house.

There are two new SCRs:
- CCB#57 - Allow Encoded Native in File Area Observational Supplemental
- CCB#58 - New Product Native values for MSL/Mastcam

We will not be discussing these today, but we will discuss them in the weeks to come.

Tier 1 Issues

CCB#51 - Add an optional attribute in the Target Identification class to identify Primary/Secondary targets

The proposal is to add the ability to identify primary versus secondary targets. The real question we had with this is what your teams, Thomas, mean when they're identifying primary and secondary targets because there are different ways of thinking about it. So, we wanted to clarify what was meant and why it is needed.
~ Thomas is here to clarify and answer questions. The request came when the teams started to discuss some instruments that usually measure planetary atmospheres or surfaces, scanned occultations or other measurements and essentially the need that they expressed was to have the ability to classify between the main target and the secondary target. At the very beginning it was basically to capture what was the main intended target at planning level when the observation was scheduled and what else was passing in the field of view, which was not captured at planning level. So that was the initial idea. What is below that is that the attributes are quite common for planetary missions that have a tool so this is something that can be used by other missions and this is why this is the solution that was proposed. Would rather not have to put this type of optional attribute to

describe the intended observation in a mission dictionary, but rather have it exist as an option that can be used or not in the core.

Question: Any thoughts on that? Can see the value of having it in the core, a single method that doesn't vary from project to project, but know this is something that Rings has done extensively - identifying secondary and sometimes unintended targets that are observed during some other observation.
Answer: Yes, that is definitely true. Cassini is an example where we dealt with this a lot. There are lots of different ways we can talk about what is meant by primary and secondary. For example, if you were going to look at the Rings, operationally, Saturn's rings were not a target that you could put in the pointing algorithms. It was always the center of Saturn with an offset. There are certain fields of metadata where Saturn is actually listed as the target because that's what gets entered into the pointing algorithm, even though that's not the scientific purpose of the observation at all. Similarly, sometimes you might be looking at a star or a particular coordinate in the sky and letting what's actually scientifically interesting to you pass through the field of view passively. That's another example of where what is operationally the target might not be the scientific target. There are other times where you're wanting to watch two objects pass by each other, with the mutual events, where one of them is eclipsing the other and so operationally you target one of them, but the other one is just as much the scientific target as the other. So those are just several examples of ambiguity and what a person might mean by primary target and secondary target. And then you have cases where you're meaning to stare at this one thing and completely unknown to us something else passes through the field of view and somebody might want to know it's there, so we talk about that and put that into the metadata as well. That might be what it meant by primary and secondary, but I'm not sure and that's why I was asking the question.
~ I think there are two ways to look at this. What the team expressed as of now is a view which may be a bit simplified of what could be done just by classifying as main and secondary. Yes, I think if there is the option to capture more scientific meaning to this attribute, like what you say. If I have two targets in my observation, and one is a star which is being orbited by a planetary body, than with this attribute and the two targets specification, I have the meaning of the observation, which is an occultation of a star by a planet or a planetary body. So, my view on this is that rather than just focusing on whether a target is primary or main or secondary, I would be keen to know a bit more from your side about what type of attributes would fit to expand a bit on the usage of these optional attributes. Like what would be an attribute that would fit as part of this target definition. Like order. As you say initial request of having some sort of classification is very broad and vague. I think it will, over time, be better defined, but already as of now we can see clear use cases and those could be used with this type of optional attribute with limitless possibilities.

It seems like there are a lot of potential uses for this. Think at a minimum it would be helpful to have a set of use cases to look at and consider as we are designing something to address the issue. It seems like a lot of these use cases are actually somewhat dissimilar. They're not really tracking the same information. Trying to capture all that within one approach is sometimes difficult.
~ For an example we are mentioning the observations of Saturn's rings. The target itself is the rings, but for pointing the target is Saturn. This type of information could be used to define the rings as main target and Saturn as the pointing target. Could use this approach to give a bit more consistency to the classification.
~ Someone is trying to look at a label to see how it was done, but you could. It would be different from what has been done already, but you could list all the targets in the target area. I guess you are wanting to designate different kinds of targets. Concerned that will lead to a precedent that

would be hard to maintain for certain situations, where as if the target just includes everything that we know to be in an image....(stopped talking)
~ Someone else thinks that suggestion for Thomas is probably the best one. If Thomas could provide us some label fragments to give us a target to shoot at, we probably would have a better idea of how this might actually work. But, as he is talking, it occurs to me that probably the best way for JUICE to handle this would be in a JUICE LDD. Others might disagree, and I am willing to consider the options.
~ But if it goes in a JUICE LDD this is something that might pop up for other missions later or when doing the migration from PDS3 to PDS4 or something else.
~ Yes. This might also be something that could go in the Mission LDD, which might be flexible enough to be used for lots of different possibilities.
~ That's not a problem. You can certainly re-use the name of an attribute and even the definition of an attribute from one one mission LDD to another. Not a big deal.
~ Not a big deal from the label design standpoint, but it is from the Harvest standpoint because you've got to search across what might not have the same name across all the different dictionaries and certainly not the same x-path.
~ Yeah, but you wouldn't do that search.
~ This is important information within the context of a particular product, but probably not for search. It's probably more for when you're dealing with a specific data product you want to understand.
~ That's right. That's why I think an LDD is perhaps the best choice, because if you're talking about search, all you are really going to search on is these targets in the target list in the target area. That's it.
~ Not entirely true. If I want to do a search to filter out observations of Ganymede which are surface observations and not atmospheric observations, if we have only as target identification the planetary body itself without optional attributes that can be used for search then I can't.

Question: Is Ganymede in the target area or not? You can list as many targets in the target area as are appropriate. Yes? Can't we?
Answer: There is no limit to the list. There is no limit to target identification class, but we can not do the differentiation between what would be an observation for a sub-fraction of the surface from an observation of a body that would be observing the exosphere and that might not contain an intersect with the plane of view.
Another Question: Someone doesn't see how this particular SCR solves that problem. Don't see that we are talking about exosphere versus surface here. Are we?
Answer: No. As was mentioned, this is, when it was raised, it was expressed quite generically between teams that want to identify and flag in their products when they have multiple targets passing through the field of view - which is the main target of the observation, the one that was intended at planning level and what came into the field of view which may not be intended at the time of the planned observation. What I was mentioning in this comment and also in the question is that in addition to this basic classification, let's say two types of targets, eventually, this could be an optional attribute that could be reused and refined a bit more for how targets can be classified depending on the purpose.
~ A person agrees and thinks this would be a good idea. A good example, a good case where we need examples of how this would be used. Believe that different disciplines have very different needs as far as targets go. This was an SBN example - with the surveys that we are doing, we thought awhile about how to do that because often the telescope is just pointing to a patch of sky and how you fit that into a target identification class because probably there are a bunch of stars

and asteroids and such in that field, but you don't know, so what SBN did eventually is we had a survey LDD that identifies the four corners of the field that is being observed. It specifically only makes sense in that survey context, but it is what works and I think every discipline could have their own little thing for how targets get identified. We see this with New Horizons and the plasma instruments. There is debate over over whether you're looking at the ISM or the solar wind or what the target really qualifies as. It seems like the sort of thing that should probably be handled individually by a node as needed, but I guess it depends on what other use cases there are.
~ That's a good point. Certainly it could be done as a discipline level LDD rather than just at mission level. I know that there are lots of Jupiter images that happen to have a moon or two in them, but if you're taking an image of Jupiter and you have a tiny little dot, which is Io, on it, you might want to filter out those images which may have been listed but aren't very useful if you're really trying to look closely at Io. Can see the use case in that case. It's out of my discipline. In fields and particles, but when we're in orbit around Mars we almost always include solar wind as a target just because Mars has such a small magnetosphere or ionosphere. The primary target is still probably Mars and that might be useful information. Anyway, that is some thoughts on the fly.

A suggestion is that we use the mission dictionary and we create an attribute there called intended science target, or something like that, that could be useful for search if you have it be in an enumerated list that guards against spelling errors and makes sure that everything is in there and is controlled and useful, although that would be a bit of work for the steward. Or you could have free text if you want it to be a little easier. It would allow you to designate among whatever is listed in target identification what is the scientific purpose of the science observation. Was just looking at a Cassini example. It was kind of ad hock. It was migrated over from PDS3. It is an example that helps you see what the issues are, but the way it was actually addressed in Cassini is not really a model. Think this might be useful across a range of missions. Interested to know if Thomas thinks this would address the concern that he's bringing, and if it would, then I think it might be useful to a lot of missions.
~ Thomas replied - Okay.That could be a solution. I would need to discuss it with the team to make sure that we could capture the needs there.
~ The advantage of that approach would be that it is quicker and easier to get updates into a discipline dictionary or any LDD. That would be an opportunity to play with it and to try to get it optimized, and if at some point the idea is to get something that's useful enough that it can get transitioned into the common model, than this is something that could benefit form that approach.
~ (The above suggestion has been added as a comment in GitHub.)

Someone thinks it would be faster for Thomas to do it in the JUICE LDD than the Mission LDD because then he has much more freedom to play with it.
~ Yes, but he is concerned about interoperability and for re-use of this so that we don't have to have every mission unnecessarily doing this. It's an old argument.
~ A person agrees.

It really makes sense to discuss LDDs, but another option is the provenance model that we have had for about a year. If you consider the observation an object, in a sense what you are doing is creating the "what happens" historically over that observation in terms of the targets that were encountered. The order and important steps would all be easily done in the provenance model. Just throwing that out. I just threw it into ChatGPT and it came back almost immediately. The point is that I'm hearing a lot of nuances here and it's exactly the way provenance is designed so that you can define all the nuances you want. On the issue of interoperability, Thomas could work on a

provenance for the observations. Also heard that maybe we could get some use cases and we maybe could come up with a team version of that. Maybe we could make it more interoperable. The idea would be to come up with a standard provenance for observations which include various options. Anyway, it's something to think about.
~ Alright.
~ Willing to share the existing provenance model, the existing LDD for provenance that's been discussed for about a year to various degrees, and I could send the response I just got from ChatGPT in terms of value and how it applies to an observation.
~ That would be useful. The only thing is that that would be to track the order in which targets might appear in observations, which is quite valid for time based observations, but maybe not so much for remote sensing at a given time in a given field of view and multiple targets in it. I believe that is correct.
~ Not a planetary scientist, but if you are looking at an object, what I can tell you is that if you define the object and you define the targets than attributes, in terms of importance, order, time, time stamps, all that is easily included in the provenance, so I would expect there probably would not be a problem to capture anything you wanted to the level of that deep in the model you can actually just start using keyword equals value strings for things that are not formally modelled so, almost nothing you couldn't put into it. I think the issue is interoperability. If you want something that other people are going to be using then you will have to make some type of generalities, general approaches, and agree on how things are going to be handled. So, on the question of what can be incorporated or what can be captured, I don't think there's an issue there. Just need to decide what you want to capture in a formal way. Could do a Tiger Team and work things out with those who would like it to be interoperable and those groups.

Question: Can Thomas repeat his question? I missed it.
Answer: What he was asking was if the provenance model was more fit to observations which are some sort of time series that track targets entering an observation, then another target entering and then another and so on, instead of more fit for remote sensing observations at very specific times that capture, in a single field of view, multiple objects, and one of them is the main one.
~ Right, so I confess I'm having a hard time understanding what is being proposed in terms of how it would look in a label. I don't know if you have any way to make it more clear, but if we go with the idea of an intended target science attribute in a mission LDD, we could make that something that you could have the option of having more than one instance of. So, if multiple intended science targets in the same data product, just have that attribute appear multiple times and that would distinguish between targets in the same product. If one is the intended target and if one just appears then you wouldn't have that be the intended science product. Seems very simple to me.
~ The example raised is very relevant to fields and particles data. When Cassini had a flyby of the Saturn satellites, they provided day files, and so, most of the day the target is Saturn, but for the hour or two closest to Enceladus then the target is Enceladus, so there is a lot of data at PPI where that is the case. Currently, PPI just throws both targets into the label and there's no differentiation between the fact that one flies through the entire day and one is a limited period.
~ Sounds like another case where having multiple instances of intended science target makes sense. Even though they aren't operating at the same time. It's a long product that covers a period of time where one thing is the intended science target and then something else is, all within the same data product. If you are wanting to designate which portions of time for when one is the intended science target I don't think any of the proposals on the table would help you with that.

~ Not part of this current SCR, but again, think this is a reasonable reason why it could be very useful to have a list of use cases to consider because that is a use case that I had not thought about, but it is a valid one.
~ We could also develop attributes for an LDD that would solve the problem.
~ Exactly. It would kinda help to identify some additional attributes that might be needed for each of these different use cases.
~ I like the idea of the use cases. Think that's critical here. We could get three or four on the table. Will volunteer that I **(Steve)** would be willing to show how the provenance model could be used to handle any use case you can come up with. Others can do it in a mission LDD and we could compare and contrast, but specific to the point, I don't think there is any use case that we couldn't provide as much nuance as we want. Again, just a matter of how specific or unique you want to have your particular example as opposed to something interoperable. That's the hard part, figuring out what to use to keep it interoperable, but I don't see any problem at all in terms of handling different use cases.

Several points. So, if several use cases could be written down, and then Steve and I **(Matt)** could both give some label snippets of how this would look in our given solutions that we are suggesting, that might be a way to go. Two larger points to make - one of them is I'm interested if everyone agrees that it seems like we are moving towards solutions that do not require a change to the IM and that these are LDD based solutions, which means that we do not have an impending deadline to get this into the next build. I'll then make my other point after I pause to see if people agree with that statement.
~ Good point. Agree with you that the issue of moving to LDDs as long as people agree that there is a way to allow LDDs to be interoperable, that we can have LDDs so they can be used across missions, as long as we all agree and make efforts to do that, then I think that's a great way to go.
~ Right now it seems like there's a lot here to try to handle and trying to have rapid turn around on updates and corrections would be a lot easier in an LDD the it would be in the common model.

Question: What is the usual time frame for requesting a change and implementing a change in an LDD based solution?
Answer: That could be done very quickly if the steward is engaged. You just make a GitHub change, make sure the tests run, send a message to EN and it can be done in a week or two. Sometimes it takes longer if there is a lot of discussion or if people kind of lose the thread, but in principal, I've seen it done in a matter of a week or two.
~ Okay. It will take time to gather proper use cases that are detailed enough that we can work with. That was a bit my worry with these first requests. I felt it was not precise enough. I tried to get a more detailed use case from the team. Some team members could help, but the main team members are staying quite top level at this stage, so between main and secondary, I think it will help to have a bit more details before hand. I can definitely try to get you some use cases. It will take a bit of time, but from that point I think what's proposed with LDDs if this goes with time table would probably be fine.
~ One addition about updates to LDDs. As soon as you submit one and it goes through the build process you actually have a schema and schematron that you can use to test and play with. The one to two week turn around is actually for getting it merged and released on the EN website, but you pretty much instantaneously have something that you can start to work with.

Question: To make sure I got it right, when we talk about an LDD based solution we are still talking about at LDD at the discipline level?

Answer: Yes.
~ The mission LDD is an LDD which means that it is not part of the core model. It has more flexibility in how it is maintained, but it is designed to be used by multiple missions. It's supposed to be generic attributes that any mission might want to use.
~ Okay. Got it.

Question: The second point was to check in on the archive development process going on for non-NASA missions, such as JUICE. I know that for Clipper and other NASA missions there's a large PDS involvement with the data archive working group, designing labels and creating a mission dictionary and also talking to the DDWG as necessary if changes are needed. Is IPDA or some other agency doing similar work for ESA missions such, as JUICE, or is there a way that we, as the PDS4 design working group, could be doing a better job of supporting European missions?
Answer: I only have a very biased view, as someone who started on JUICE in April of last year, but essentially, what we want to try to do is develop data products and labels and try to identify types of metadata to capture. There was, for some teams, quite some work already done to develop instrument data dictionaries which were sort of over lapping with existing discipline dictionaries. The reason why, I do not know, but for imaging instruments people by default tend to develop their own dictionaries instead of having a look at what has already been done in the Imaging LDD. This is something that we're trying to correct now by iterating with the teams. Apart from that, the process is exactly what we're doing now. We had these requests. I'm trying to see if there could be commonalities between instruments on the mission at some point. So, at some point I raised this ticket with you guys to discuss this with you. That would be the way, I think, in terms of archive development to possibly request things that would be useful in particular to multiple missions inside or outside of ESA.
Another Question: Are you the POC for JUICE?
Answer: I am one of two. I do remote sensing and my colleague (Sean?) covers the particle instruments.
Another Question: Very good. Do you work directly for ESA or PSA or…?
Answer: I'm part of the archive scientist team at PSA.
~ PSA. Okay. I hoped PSA was doing that work. Great to know that we have this connection and glad to have you here. We want to support you and that work that you are doing because it is really important.
~ Thank you.

Someone else has two points to make. Think this type of work would be great to report to the IPDA. There's an annual meeting and there is a time where things like this and this type of work and these types of capabilities are discussed. Seems like there is always a lot of interest in changes to the model. There is certainly a place in the meeting for that type of thing to be reported. This would be a great example of how it's done. The other point is I need clarity on the use cases. Unclear if we are getting them from people besides just Thomas. Would like to see them from others too. Would like to see three or four use cases.
~ Joe said that he can definitely product some use cases **(Action Item - Joe)**
~ Matt said he can produce them in the context of Cassini even though they have used a different solution at this point. **(Action Item - Matt)**
~ Someone's impression is that there are many different types of observations and the question is if we can classify them and start building these LDDs for those specific types of observations. Think the best bet is to get as many on the table as we can initially because it's better for the model in the long run, or five or six models in the long run.

~ Yes.

Question: Any additional comments?
Answer: (Silence)
~ The **Action Item** is for Joe and Matt and anyone else to produce use cases and post them to GitHub for us to discuss.

CCB#55 - Funding Acknowledgement class enhancement (funding source and funding acknowledgement text)

CCB#55 is making some fairly simple modifications to the funding acknowledgment class. Currently, only three funding sources are allowed. The proposal is to expand that. In addition, the funding acknowledgement text is limited to 255 characters. Again, the proposal is to expand the length of that.
~ Yes. This is for the funding acknowledgments and text, to expand the list of possible funding sources. Discussed this with Mark Bentley at ESA. Currently, ESA is not in the list. There could be a potential use for national agencies that fund research projects that produce data that can be archived as well as part of the main archive. So I was wondering if having a bit more flexibility than an enumerated list would be something that would be a possible solution, like some sort of free text entry.
~ Someone was surprised it isn't free text. Unsure if anyone knows why.
~ A person discussed it with Jordan. The origin of the enumerated list was to avoid multiple versions of the same funding source, to avoid many flavors of the same thing. But, at the same time, it is true that it restricts the users of the class and I think there is a lot of potential for that. One question from the PSA side is what would be the frequency at which we could request a new agency or funding source to be added to the list of allowed values and see from that point if this is a speed that is okay or if attribute could be a bit more free or free text.
~ Right now the enumerated list is NASA, NSF and RAS. We could add all the agencies that are part of IPDA off the bat.
~ Yes, but in France, for example, we have multiple ways to get activities funded. We have a national research agency and others and specific programs that can fund research activities. In other countries it might be a similar situation. If it's possible, we should have some sort of free text, something a bit more flexible, if the pace at which new values to be added cannot be as fast as the users need.
~ Historically, it's been locked into the build cycle, but there are discussions about trying to figure out a way around that, so this is something to consider.

Question: Maybe I misunderstood, but what about the current funding source being the primary funding source and those could be added as necessary? There's probably fewer of those. Then an additional funding text field where all the details that might need to be expressed could be. Not sure if I misunderstood what was being proposed, but would that be a solution?
Answer: There already is a funding acknowledgement free form field that can go in there.
Another Question: If that can be used for that purpose...?
Answer: But then there's an issue of also having the funding source being mandatory with the three values. When I tried to use the class I didn't have a way to use the class for data that were funded by an ESA project because ESA is not in the list of funding sources and you have to use a funding source to use the class.

~ Obviously we should add ESA to the enumerated list of funding sources.
Another Question: But then what about the many flavors of other funding sources that we can face in the future with people producing data to be archived?
Answer: NASA and NSF are both US sources of funding, so I guess what you are saying is you've got similar situations just in France itself.
~ Yes. Well, basically there are many agencies plus national funding programs. I feel that it would be extremely hard to capture them all in an enumerated list before people start to need them. My fear, what I wanted to discuss with this class and these possible changes was whether more flexibility is okay. Whether it is a no go, and if so, then how to make is so when users need the class they can use it in a time frame that is okay.

This sounds like an age old paradox that we are always dealing with - between free form and an enumerated list. In fact, it even came up in the intended science targets conversation we just had. If you want it to be easy for DPs to just put in what they're going to put in and move on, than free text is what you want. But if you do that then you are going to have spelling errors and people using different words to describe the same thing and so it's going to become useless for search. So, if you want this to just be something human readable so that people are able to dump information in and people will see that info, then free text is fine. If you want this to be useful for search than you want to have an enumerated list so that people will fit what they want to say into a certain number of objects that people will then search on. But, of course, Thomas is also talking about if you want to have an enumerated list than you need to be fast and flexible for adding necessary new values to the list.
~ Exactly. Yes.
~ Someone says let's do that. Give us a list of enumerated values that need to be added and let's add them. We can do it quickly.
~ One thing that strikes me with this list is it doesn't list programs. What we are seeing is that it lists NASA, not PDART or CDAP.
~ That's right. It is at the agency level. Not sure what RAS is doing there, but whatever.
~ So, think there may be an ability to add a new attribute to put stuff like specific programs that funding comes through from an agency.
~ But then that means the funding source itself would need to be optional when using the class otherwise. In this case, I want to list national funding for a project using the class, but if I use the class I must specify NASA, NSF or RAS, then I cannot.

We are all forgetting that ESA is a bunch of different countries and they all have something equivalent to NSF. They may not call it that, but at that level, even if it's the Greek NSF equivalent, it's not ESA, it's Greece or France or Germany, so we shouldn't combobulate all of Europe into ESA. I think that is part of what I am hearing.
~ So we might need a whole lot of national agencies that go into this. About the funding source being required if you have this class, the class is not required and if you don't have a funding source that you can specify here, unclear why would you want this class in the first place. You can leave this class out if you don't have a funding source you can acknowledge, but if you are going to put in the class it makes sense to have it without a funding source.
~ Hence the initial question of how to handle the additional funding sources that could be listed.
~ Give us a list and we will put them in.

Question: How fast can you put them in?
Answer: Maybe pretty fast.

~ We have a build coming up. Upcoming build and freeze point, but if you wait awhile, it will take longer.
~ This is the sort of thing we want to do very quickly. We know that these enumerated lists need to be enumerated and we really want to strip away the time and effort required to add to these lists.
~ Absolutely.
~ If you can get a us a list in the next week or two we can probably get that in before the build.
~ Yes.
~ The list would be very incomplete and biased. That's an issue.
~ Yeah, but it would be a start. As long as it covers what you need right now. In six months we can implement the next clean up.

Our time is up. We will have two last comments and then we need to wrap up.

Hesitating to bring this up, but on the point about what needs to be searchable, the issue here is that it's not clearly defined what needs to be searchable. Agency is only mentioned, I think, in one place. Just worried we will end up with things like our instrument list for PDS3 and then sometime later we'll say oops, the list is continuing to grow. Just wondering if there is a way to identify the primary sources which need to be added to the current list for search capability. Then we could have an additional text field that could list everything else if the definition is changed or we can add another text field to allow agencies or whatever might be involved in funding to be included. The current text field seems to be focused on the relationship as opposed to which ones exist. That's my input on this point.
~ Related to that, it seems that if there's a linkage here it should be between that alphanumeric number given in funding awards, so if you have a funding award you should be able to find it based on the source. So we wouldn't want to go higher or lower than the level of where that would apply. That seems logical to me.
~ Good point.

Thanks to Thomas for joining us today. It's been very helpful. Thomas will update the tickets.

We will meet again in two weeks.

title: DDWG Notes 2025-03-06
layout: default
date: 2025-03-06

March 6, 2025
Notes by Debra Kazden

Known Attendees: A. Bailey*, J. Blalock*, B. Hirsch*, L. Huber*, S. Hughes*, R. Joyner, D. Kahan*, D. Kazden, P. Lawton, T. Lim*, J. Mafi*, S. McLaughlin, E. Schaefer*, M. Seritan and M. Tiscareno* (*'s indicate people who voted during this meeting - if applicable)

DDWG Agenda

(Included in email sent March 4, from J.Mafi, Agenda for 2025-03-06)

We will have a DDWG telecon this Thursday, March 6 at 9:30 PST (12:30 EST, 10:30 MST). We will cover the following issues:

1) Action item reports.
2) CCB#55 Funding Acknowledgement class enhancement (funding source and funding acknowledgement text – A zip file containing example Schema and Schematron files with the requested changes implemented along with some test labels have been posted under "Applicable Attachments". Please review this SCR and be ready to vote.
3) CCB#45 Add reference types values for *-to agency – The "Requested Changes" section lists two options for how this SCR could be implemented. We will discuss and vote on which of these two implementations is preferable. Note that the attached Schematron and labels are out-of-date, and represent a partial implementation of Option 2. They will be updated following the vote.
4) CCB#20 Need a Context object for Instrument Packages (JIRA-CCB-358) – This SCR has been significantly revised following discussion in the tiger team. Please be ready to discuss and vote on the approach currently proposed under "Requested Changes". Note a test set of Schema/Schematron/test labels and revised documents have not been provided, nor has the Tech Assessment been updated to reflect the current proposal. These items will be produced and made available for review by DDWG members following approval of a specific implementation.

Front-matter/What's new

1) Submitted to the CCB: (none)
2) New issues in GitHub:
   CCB#57 – Allow Encoded Native in File Area Observational Supplemental
   CCB#58 – New Product Native values for MSL/Mastcam
3) Ready for vote: (none)
4) Build 15.1 freeze date: 2/27/2025. The following SCRs are queued for implementation: CCB#13, CCB#48, CCB#49.
5) Withdraw by author: (none)
   (Brief Discussion)

Action Items

a) CCB#39 – S. Hughes to update SCR
b) CCB#28 – J. Mafi to research

c) CCB#12 – S. Hughes to verify whether this is subsumed by CCB#52 and CCB#53
d) CCB#20 & CCB#32 – Tiger team formed and working the issue
e) CCB#31 – Ron and Matt working the issue **(Brief Discussion)**

Tier 1 issues

- CCB#56 – Reconsider the PDS4 build cycle cadence
- CCB#55 – Funding Acknowledgement class enhancement (funding source and funding acknowledgement text)
- CCB#54 – Add modification history for classes and attributes
- CCB#53 – Improve cross-namespace validations in sch files
- CCB#52 – Enable cross-namespace validations in sch files
- CCB#51 – Add an optional attribute in the Target Identification class to identify Primary/Secondary targets
- CCB#50 – Clarify Ambiguous Author Information in Product Document
- CCB#46 – Disallow multiple versions of the same LID from occurring in a single collection
- CCB#45 – Add reference types values for *-to agency
- CCB#44 – Issues caused by the failure to expose pds: attributes that are referenced in an LDD.
- CCB#41 – Create a way for classes or attributes to be exposed to other LDDs but not to data providers
- CCB#39 – Fixed problem where LDDTool does not include all rules from IngestLDD
- CCB#37 – Add List Author, List Editor, and List Contributor to the Modification Detail class
- CCB#36 – Add Mission Phase to the Investigation class
- CCB#35 – Add Alias Index and Acronym Index to the Investigation class
- CCB#32 – Add Host type to Observing System Component
- CCB#31 – Ensure that dictionary versions are specified in labels
- CCB#28 – As a user, I want validate to allow delimited tables with whitespace-only numeric fields
- CCB#23 (JIRA-CCB-211) – XML/schema based files as archival data
- CCB#22 (JIRA-CCB-164) – Display Settings not required for images
- 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#21 (JIRA-CCB-367) – Provenance LDD Using PROV-DM
- CCB#20 (JIRA-CCB-358) – Instrument Package Context Products
- Tiger Team: M. Tiscareno (lead), B. Deen, B. Hirsch, S. Hughes, R. Joyner, J. Mafi, E. Schaefer.
- CCB#13 – Ensure file name and directory path name adhere to SR
- CCB#12 – Document Nuances for Reusability of Generic LDD Classes/Attributes
- CCB#5 – Move to V.2.0.0 for June 2025 Release
- CCB#4 – Backwards Incompatibility policies
- -Presentation: Backwards Compatibility
- CCB#3 – Adopt Semantic Versioning with version 2 transition
- The CCB returned this, requesting that it be resubmitted as part specific proposal to transition to IM V2.0.0.
  (Discussed 55, 45 and 20 - voted to pass 55)

Tier 2 issues

1) 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!

1) JIRA-CCB-326 (link) - Inconsistent Direction on How to Use Discipline Facets
2) RS: 259/281 (see new attachment to JIRA-CCB-281) (on hold)
3) JIRA-CCB-364 (link) – Disallow bundles referencing collections by LID **(Not Discussed)**

Historical links:

- Export of history: https://github.com/NASA-PDS/PDS4-CCB/tree/main/data
- Backup: https://nasa-pds.atlassian.net/jira/software/c/projects/CCB/issues (limited seats)

Next meeting: 20 Mar 2025

DDWG

Front Matter

Starting with reports on action items. We will discuss CCB#20 today. No updates on CCB#28 yet.

Question: Any updates on any of the other action items?
Answer: (Silence)

There are three new SCRs:
CCB#57 - Allow Encoded Native in File Area Observational Supplemental

CCB#58 - New Product Native values for MSL/Mastcam

CCB#59 - funding year only allows the value "9999"

CCB#s 57 and 58 are both related to product native. The argument is that there should be a PDS compliant product. There's a lot to discuss, but not today.

The other new SCR, CCB#59, there is a problem with the pattern for the attribute for funding year. Probably just a bug fix. We will discuss it more in the coming weeks.
~ Bug fix because of the definition. There is already a policy, so maybe this is just a validate bug ticket.
~ No, it's in the IM. The pattern is wrong. It requires an update to schema.
~ An implementation issue. It won't need to be debated.
~ A good thing to fix and get in the next release if possible.

Action Item - Everyone look at CCB#59 in GitHub and leave comments.

Tier 1 Issues

CCB#55 - Funding Acknowledgement class enhancement (funding source and funding acknowledgement text)

The request was to add new permissible value for ESA for source and change the type to ASCII text collapsed, remove the restriction of 255 characters.

Question: Any comments?
Answer: (Silence)
~ No new comments in GitHub.

Adding ESA is obvious. Hadn't noticed the change to make it ASCII text collapsed. Unlimited. Don't want it unlimited, don't want people to be dumb, but unclear what the limit should be.
~ There is a maximum character we could use, but unsure where we would set that.
~ Free text, not for search.

We should remove the two liens from the TA.
~ Joe checking the liens in real-time.
~ Steve agrees they can be removed.
~ Ron agrees they can be removed.
~ That is the only change we would need to make.
~ Done in real-time.

The Vote for CCB#55
ATMOS - Yes
EN - Yes
GEO - Yes
IMG - Yes
IPDA - Yes
NAIF - (No Vote)
PPI - Yes
RMS - Yes
RS - Yes
SBN - Yes

No vote heard for NAIF, but we will move forward. This will be sent to the CCB after this meeting.

CCB#45 - Add reference types values for *-to agency

There have been a lot of comments on this. We need to discuss an approach to implement this. We will discuss concerns that people have with this.

Question: Comments?

Answer: All along have said there is no use case to demonstrate what this is important or needs to be done. Someone wanting it isn't enough. This is implementation driving the IM, not IM driving implementation and that is against all we have tried to do.
~ Someone supports that. As a member of the EDWG, unclear where this request came from. There was a request for use cases, but they have not been provided.

Schema and schematron labels have been shared. Understand this request came from Jordan. Thought he was able to show where this... (interrupted) ... not implementation driving the model. This is useful information. Says who is responsible for a product. Should have been included a long time ago. We do have curating node, so this is actually to provide what system is requiring. Tried to provide examples. More like a requirement that this information be provided.
~ Someone met with Jordan. Discussed this. No use case. Still haven't seen one. As far as being a system requirement, focusing on provenance, there are way better ways to do that that are less invasive or disruptive to the system.
~ Provenance is secondary reason for doing this. Who is responsible for this product - we need to take that seriously. In terms of how onerous, LID is a no-starter, funding area, not links. Need something for the system. How system is supposed to work.

NSSDCA understands, but would hesitate to put that in the bundle or collection because a node might have to hand over curation in a re-compete. Seems like it should be outside the bundle.
~ Useful perspective.
~ If provenance changes it changes the metadata, yes, but it needs to be captured. Important to know who was responsible for getting it into the archive.
~ For this speaker, to answer NSSDCA, it could be useful to have specific use cases that we can think about in terms of NSSDCA and other concerns.
~ Available in the test case.
~ Examples, not exactly use cases, and out of date. Partial rep of option 2. Need more information. While examples are helpful, they don't really explain how the information will be used by the system.
~ Someone withdraws the comment that there are no use cases. There are some from Feb 20. Four test cases with "valid" in the file name. One is a data collection with product to node, product to agency, but have funding acknowledgement class, so wouldn't need that and as NSSDCA points out, if we change the curation we need additional changes. Next example has investigation. Bundle to node doesn't seem like a good idea. Then Clipper, then planet Mars. Still missing a compelling argument. Investigation to agency, maybe.

Funding acknowledgment isn't sufficient. It isn't a link the system can pick up. Asking for links between products. Might need to split this into two SCRs. This was supposed to be done ASAP.
~ Not ASAP.

It's a red herring to be asked to find a search use case. If you look at the Problem Statement, this allows the system to find responsible node for a collection or collections of a node. System needs this. Those are the use cases. Search was never a use case. These are system requirements. Unclear why we don't provide this.
~ If it's a system requirement we should have use cases to show how the system will use this.

To the NSSDCA point, this is changeable metadata that we don't want to have live as part of the data. Seems like a better solution might be agency to investigation. We can use context products to point.
~ Good point. Can point the other way.
~ To push back, at PPI, we have a lot of VEX data, which is an ESA mission but our VEX data is funded by ROSES, so just knowing investigation and agency is not always enough information.
~ If we have the link in one direction, then the system can go the other direction. For provenance we need better arguments that this doesn't live in the archive.
~ The information is in bundle. Already there. Author list, curating node…

Unsure where to go from here.
~ There are issues to resolve still. Would like to see some specific use cases. Sympathetic with the argument that this proposal is not the best way to provide provenance. Sympathetic to NSSDCA concerns. Wonder if we need a product that lives outside the archive.
~ There is a difference between who curated when data first went into the archive versus into the future. Challenge the need for a pointer. We need additional layer between archive and products to have a website.

Option 1 is where Jordan and another landed. Think option 1 for product context makes the most sense, not for bundle or collection products. For the rest of this can't see a use case where this makes sense for search. Agree that EN and the registry need to show us why the current system doesn't work. Also, without a use case we will never get this by the CCB.
~ It will be divided into two SCRs. Context products and for bundles and collections. Hearing an issue of ambiguity between archiving and curation. A lot of nuance in curation and author list. Both are provenance information. Will focus on that, but for use cases, the current Problem Statement has three. Search is not one of them. Unclear if we need to expand.
~ If the system needs it, then someone would need the search. If no one will search, we don't need it.
~ Need information for landing pages. System needs information for users to see. Arguing for trying to provide that from the IM.
~ If we make this two SCRs, would like to hear from the EDWG or from Jordan. Also, curation is coming up a lot. It would be good to clearly describe curation or archive. Thinking this isn't curation.
~ That makes sense. Will propose new reference type values.

Someone sits in on the EDWG. No consensus decisions. They are being told to do this.
~ So no real authority to say what they need. So, will use the requirements from the system folks.
~ There is work to do. Would like to pursue the idea of a new product between the system and archive, but that is a major change.

CCB#20 - Instrument Package Context Products

The SCR has been updated after extensive discussion by the Tiger Team. Requested Changes include two new types of context products and all context products could have multiple instances and to add list of reference types to the list so relationships and hierarchy can be described.

Question: Any comments?

Answered with Another Question: There's a statement about EN preferring not to have underscores in the center?
Answer: That was written based on what EN said.
~ Back tracking, underscores are fine and probably preferred. Jordan can handle them. Having name of relationship in center seemed easier to parse.
~ Either way, what matters is a central list of central tokens that work. Just "to" for now. Can add more. Want to be respectful of EN's needs.
~ EN just needs a list of tokens.
~ Might put them back into is member of and proceed.

Question: Other thoughts?
Answer: (Silence)
~ There are a number of changes. No test implementation or test cases. Need feedback from the DDWG on the approach. A lot of work goes into making test cases, schema, etc. Would like input on this before going forward.

Underscores revised in real-time.

We will vote on the approaches described in SCR. Only change was underscores - which are now added back.
~ Another potential straw-poll on preference for "to" or "contains".
~ We can vote on that. If containing or contains is better.
~ Options put in the chat. (See below) Contains versus to member and to containing. Hopefully, that's not reversed.
~ That's correct.
~ Go to Additional information in the SCR, number 5. That's the one that is being suggested now, regarding internal underscore. To member and to containing.
~ Someone has no strong opinion either way.
~ Another person agrees.
~ "To" is what we have used in the past.
~ No strong opinion, just a slight preference.

Question: Example?
Answer: Someone does have a preference. "To" was used from the beginning, very general. "Contains" is a lot more information, has additional meaning. "Contains" provides more information, "to" is just continuing with a convention, which the CCB might like.
~ Fair point, but we are already using "to" in other places.

Someone put an example in Chat. (See below)

There are a number of comments here. The request is to provide feedback on the DDWG preference, if that is to continue with the "to" formation that was used in the past.

Option 1 - Current implementation, contains and is member of.

Option 2 - To use to, to containing and to member.

To be clear, option 1 is downward pointer instrument host contains. (Spoke very fast) Option 2 is an upward pointer, containing.

Straw Poll - Option 1 or 2:
ATMOS - Abstain "both are fine"
EN - Option 2 "to"
GEO - 2
IMG - 2
IPDA - Abstain
NAIF - Abstain
PPI - 1
RMS - Abstain
RS - Abstain
SBN - 2

The 2s have it. Will update to include the "to".

Question: Now?
Answer: We are over time now. Would like to approve the approach and than make test labels, schema/schematron and have a TA after this.
Another Question: Any objections to approving the SCR with these changes?
Answer: (Silence)

CCB 20 as currently proposed:
ATMOS - No ("Unnecessary level of splitting") EN - Yes GEO - Yes IMG - No IPDA - Abstain NAIF - Abstain PPI - Yes RMS - Yes RS - Yes RMS - Yes

Yes is the majority, so we will do the implementation and then look before final approval, before it goes.

Question: Any last comments?
Answer: (Silence)

Next meeting March 20

Meeting Chat

[10:25 AM] Schaefer, Ethan
"contains" vs "to member"
"is member of" vs "to containing"

[10:29 AM] Schaefer, Ethan
instrument host contains instrument host vs. instrument host to member instrument host

title: DDWG Notes 2025-03-20
layout: default
date: 2025-03-20

March 20, 2025
Notes by Debra Kazden

Known Attendees: A. Bailey*, J. Blalock*, T. Cornet*, B. Deen, B. Hirsch*, L. Huber*, S. Hughes*, R. Joyner, D. Kahan*, D. Kazden, A. Knight, P. Lawton, J. Mafi*, E. Schaefer*, M. Seritan and M. Tiscareno* (*'s indicate people who voted during this meeting - if applicable)

DDWG Agenda

(Included in email sent March 19, from J.Mafi, Agenda for 2025-03-20)

We will have a DDWG telecon this Thursday, March 20 at 9:30 PST (12:30 EST, 10:30 MST). We will cover the following issues:

1) Action item reports.
2) CCB#59 funding year only allows the value "9999" – Please review this SCR and be prepared to vote.
3) CCB#58 New Product Native values for MSL/Mastcam – I'd like to start the conversation on this SCR. There are two specific questions that I'd like to discuss. Is this data type an appropriate candidate for Product Native? Is there any additional information that the DDWG, CCB, and PDSMC needs to make a decision on this issue?
4) CCB#57 Allow Encoded Native in File Area Observational Supplemental – I'd like to start the conversation on this SCR as well. Specifically: Does this proposal violate the conditions under which Encoded Native products are allowed in the archive?
5) CCB#36 Add Mission Phase to the Investigation class – I'd like to restart the discussion on this SCR and identify next steps.

Front-matter/What's new

1) Submitted to the CCB:
   - CCB#54 Add modification history for classes and attributes
   - CCB#55 Funding Ackowledgement class enhancement (funding source and funding acknowledgement text)
2) New issues in GitHub:(none)
3) Ready for vote: (none)
4) Build 15.1 freeze date: 2/27/2025. The following SCRs are queued for implementation: CCB#13, CCB#48, CCB#49.
5) Withdraw by author:
   - CCB#45 Add reference types values for *-to agency

(Brief Discussion)

Action Items

a) CCB#39 – S. Hughes to update SCR
b) CCB#28 – J. Mafi to research
c) CCB#12 – S. Hughes to verify whether this is subsumed by CCB#52 and CCB#53

d) CCB#20 & CCB#32 – Tiger team formed and working the issue
e) CCB#31 – Ron and Matt working the issue **(Brief Discussion)**

Tier 1 issues

- CCB#59 - funding year only allows the value "9999"
- CCB#58 - New Product Native values for MSL/Mastcam
- CCB#57 - Allow Encoded Native in File Area Observational Supplemental
- CCB#56 – Reconsider the PDS4 build cycle cadence
- CCB#55 – Funding Acknowledgement class enhancement (funding source and funding acknowledgement text)
- CCB#54 – Add modification history for classes and attributes
- CCB#53 – Improve cross-namespace validations in sch files
- CCB#52 – Enable cross-namespace validations in sch files
- CCB#51 – Add an optional attribute in the Target Identification class to identify Primary/Secondary targets
- CCB#50 – Clarify Ambiguous Author Information in Product Document
- CCB#46 – Disallow multiple versions of the same LID from occurring in a single collection
- CCB#44 – Issues caused by the failure to expose pds: attributes that are referenced in an LDD.
- CCB#41 – Create a way for classes or attributes to be exposed to other LDDs but not to data providers
- CCB#39 – Fixed problem where LDDTool does not include all rules from IngestLDD
- CCB#37 – Add List Author, List Editor, and List Contributor to the Modification Detail class
- CCB#36 – Add Mission Phase to the Investigation class
- CCB#35 – Add Alias Index and Acronym Index to the Investigation class
- CCB#32 – Add Host type to Observing System Component
- CCB#31 – Ensure that dictionary versions are specified in labels
- CCB#28 – As a user, I want validate to allow delimited tables with whitespace-only numeric fields
- CCB#23 (JIRA-CCB-211) – XML/schema based files as archival data
- CCB#22 (JIRA-CCB-164) – Display Settings not required for images
- 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#21 (JIRA-CCB-367) – Provenance LDD Using PROV-DM
- CCB#20 (JIRA-CCB-358) – Instrument Package Context Products
- Tiger Team: M. Tiscareno (lead), B. Deen, B. Hirsch, S. Hughes, R. Joyner, J. Mafi, E. Schaefer.
- CCB#12 – Document Nuances for Reusability of Generic LDD Classes/Attributes
- CCB#5 – Move to V.2.0.0 for June 2025 Release
- CCB#4 – Backwards Incompatibility policies
- -Presentation: Backwards Compatibility
- CCB#3 – Adopt Semantic Versioning with version 2 transition
- The CCB returned this, requesting that it be resubmitted as part specific proposal to transition to IM V2.0.0.
  (Some Discussions)

Tier 2 issues

1) 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!

1) JIRA-CCB-326 (link) - Inconsistent Direction on How to Use Discipline Facets
2) RS: 259/281 (see new attachment to JIRA-CCB-281) (on hold)
3) JIRA-CCB-364 (link) – Disallow bundles referencing collections by LID **(Not Discussed)**

Historical links:
- Export of history: https://github.com/NASA-PDS/PDS4-CCB/tree/main/data
- Backup: https://nasa-pds.atlassian.net/jira/software/c/projects/CCB/issues (limited seats)

Next meeting: 3 April 2025

DDWG

Front Matter

For action item reports, Joe has nothing to report yet and the Tiger Team for CCB#20 is still working.

Question: Any updates on the others?
Answer: (Silence)

There are no new SCRs this week.

CCB#54 (dd modification history for classes and attributes) and CCB#55 (Funding Ackowledgement class enhancement (funding source and funding acknowledgement text)) have been submitted to the CCB. No word on final votes yet, but there seem to be some yes votes, so far.

Tier 1 Issues

CCB#59 - funding year only allows the value "9999"

Hopefully this is straight forward, although a possible issue with implementation has been pointed out. The SCR has been updated. Issue was originally asking for funding year to be updated, but that is deprecated. Another comment is date is restricted to a four digit year. Not sure if that's an issue.

Question: Additional thoughts or comments?
Answer: Regarding four digit year, can the first digit be a zero?
~ Someone else was also going to ask that.

~ There might be comet data in the archive with B.C. dates.
~ Yes, believe we used negative dates for those.
Another Question: Any three digit years?
Answer: We can look.
~ Think that's a separate SCR. Bigger than just funding year. We probably don't have to worry about funding by an ancient empire.
~ Agree. No need to change this SCR.

This SCR covers funding for a single year. Could need start and stop dates if the funding covers multiple years.
~ Good point. As written we would probably use the first year, but that's a good point.
~ For RNA projects we use the year the solicitation is selected. Then it's obvious that it takes a few years while the money is spent and the work is done. We should check the definition.
~ Funding year is the year the money is awarded.

Question: Any other thoughts or concerns?
Answer: (Silence)
~ No TA yet, but we are trying to make a change to how we do things. A lot of work gets done and then becomes invalid if we change the SCR, so we are trying to get proposals approved first and then test cases and such can be produced and then we can have a final consensus or evote. Think that's the direction we are trying to go.
Another Question: Any concerns with that? If not we will do a vote.
Answer: (Silence)

The Vote to approve CCB#59:
ATMOS - Yes
EN - Yes
GEO - Yes
IMG - Yes
IPDA - Yes
NAIF - Yes
PPI - Yes
RMS - Yes
RS - Yes
SBN - Yes

That passed. Steve will do the TA **(Action Item)** and then everyone can look again.

Someone said they were trying to unmute when comments were asked for. Have some feelings - thought that and CCB#20 were on track for the next build. Understand not having EN do extra work, but concerned about if this and CCB#20 will still be able to make the next build or if it's already too late.
~ We would need to hear from EN, but Jordan has been hesitant to discuss this. There's an SCR on this and we can discuss it at the F2F. Haven't heard a clear cut off date.
~ Frustrating.
~ EN just staged the IM, so it's ready to start.
~ We haven't heard a formal cut off. Last build was a nightmare for EN.
~ A nightmare is forming now. We could do CCB#59, but CCB#20 isn't ready yet.

~ The ball is in EN's court for CCB#20. Trying to understand the new procedure. Don't want to waste time, but we need to see something before we can approve it.
~ We are heading into a nightmare.
~ EN should check with their manager. It would be good if CCB#59 could get in.
~ Another person said their perspective is that three SCRs are ready. We can push CCB#59 in and then close it off for this build.
~ Someone else is looking for a date so they can say when implementation will be finished.
~ The freeze dates have been very flexible, but not the release dates. Maybe we need to build some realistic expectations into this.
~ We should discuss this at the F2F.
~ Good plan.

We will move on now. Maybe we can circle back to CCB#20 at the end today.

CCB#58 - New Product Native values for MSL/Mastcam

The format that all PDS3 data was delivered in for MMM EDRs - the PI was adamant about wanting the original stream available. For PDS4, they are delivering a PDS4 image with encoded byte stream. Not sure it passes validation. Trying to close the loophole by allowing them as product native. There's a program to decode and they agreed to write documentation if we get a straw vote.
~ Someone's concern is that this is opening product native to a lot of data. SEED data couldn't be made PDS4 compliant, but it is used by seismologists. Raw telemetry is unique, fine if we want to go there, nothing to prevent this usage, but it's a whole new usage.
~ Thought this was the purpose of product native, but it is also being delivered as an image.
~ Which is what is required for product native.
~ The PI was adamant about wanting the original stream out there.
~ Not sure we want to accept that argument.
~ It would have to go to the MC.
~ Someone always thought that particular DP wanted to keep his proprietary format proprietary, so confused, but willing to look at the documentation if they write it.
~ Another person was surprised too. The decompressing code is out there.
~ The line, someone else thinks, is that we don't archive telemetry. Not sure where this falls.
~ It was described to someone that there is a chance of better understanding it if we have the original data, not sure how true that is, but it's science information.

Question: The PDS4 product is a JPEG with artifacts?
Answer: Yes. Compressed data from the spacecraft has artifacts and we want to understand what they are.
~ Downsides are disk space, which is no big deal, or users finding this and spending time on this.
~ That would be mitigated by the next SCR, but as a stand alone, product native does have a pointer. Weak association, but that is why there is another SCR.
~ Someone agrees we need a way to let people know these are not the data they are looking for. Think it's a system issue.
~ Someone else thinks there's been some discussion on this.
~ Engineering data.
~ Think they are open to opening up the classifications of what they accept.

Back to the idea of opening up a product native flood gate - think we have enough hurdles to stop that.
~ And someone is not opposed to having raw telemetry in the archive.
~ Section 9D of the SR says product native can be refused for any reason.
~ Someone agrees. There is nothing that explicitly prohibits this. Plenty of escape valves for the nodes too.
~ We should state in the SR if we don't want to do this as a policy.

Question: Other thoughts?
Answered with Another Question: Isn't there the NSSDCA to store raw telemetry?
Answer: It certainly could go there, but we've been getting a lot more interest from missions to archive their raw telemetry. Usually it has gone to a data annex. This would open the door for an official way to put it in PDS.
~ There is a use case on JUICE with compression. Would get better results, but afraid of opening that door.
~ These data have bit fields, so they couldn't be archived as regular products. It's what product native was designed for.

Question: Would like to decide on the next step. Do we need additional information? Would like to do a straw-poll on if people are absolutely opposed or unwilling to consider this.
Answer: Someone agrees. If there's an openness to this it would give the team a reason to get the documentation written.

Non-binding straw-poll:
ATMOS - Yes
EN - Yes (to the concept in general)
GEO - Yes
IMG - Yes
IPDA - Abstain
NAIF - Abstain
PPI - Yes
RMS - Abstain
RS - Abstain
SBN - Yes

Seems like there is support for the idea. We will get more information and move forward. The next SCR is related.

CCB#57 - Allow Encoded Native in File Area Observational Supplemental

This is related to CCB#58, but also not related. It would apply to any native. The idea is regarding the pointing. The core of native is the encoded native. This is suggesting we put in file area supplemental to bind them. More clearly says what you want to use. Users wouldn't find product native alone, which could cause confusion. Binds them together. Advantage for users. Additional mechanism. It does not replace product native.
~ Someone would like to know how it was done by GEO for SEED.
~ GEO will have to look.
~ Would like to see how that was done and how this is different.

~ Can look that up and put it in the comments as an example. **(Action Item - Bob Deen)**
~ Someone assumes something was done to tie the GEO products together.
~ Someone else guesses it was done with an internal reference.

Question: Any other thoughts on this SCR?
Answer: If they are separate users would be able to download less. If they are together they would have to download both.
~ That's a system concern. They would get both.
~ There are examples of that. Would hope that download wouldn't force people to download everything. If so, that's a system issue.
~ People don't always read the label, so they could get confused if they get two products.
~ Not a killer.

Someone did a quick Google search and saw the SEED label for Apollo 11. It has an internal reference in the mission area, SEED equivalent, but it also has a PDS CSV file. It would look different here, but it shows what they did. Put the pointer in chat. (See below)
~ Against my usual position, but this could have impact on how the model could work, so we should get input from Jordan.

Question: Other comments? Would a straw-poll be useful?
Answer: We're not ready for that yet.
~ Bob will do the example for people to look at.
~ And we will reach out to Jordan for input from the system.

CCB#36 - Add Mission Phase to the Investigation class

There are only five minutes left today. The last item on the agenda is CCB#36. Would like to restart the discussion on this one and identify the next steps. There had been a lot of discussion on this.
~ There was a lot of discussion and comments. There was a cool idea that instead of having to form new products for every phase instead it could be done with unique local identifiers. Interesting implementation.
~ Add local reference to internal reference.
~ Pushes the envelope. Makes sense. Wish I thought of it. Solves a problem we haven't completely understood. We should look into it.

The next meeting is April 3.

Steve will do the TA for CCB#59 and then we will do an evote for final approval and pass it to the CCB.
~ Someone believes it is just a bug fix. If this is just fixing the implementation we can just inform the CCB. They don't need to vote.
~ Someone else agrees.
~ Think in the absence of complaints bug fixes just go through.

Meeting Chat

In our physical characteristics comet compilation, we have a table of apparitions from -466 to 1975. However, the dates are listed in a "Perihelion Number" column, which is just a 9-byte ASCII_String with a description of "Old IAU Comet Perihelion Number: Perihelion year followed by a roman numeral indicating the chronological order of perihelion passage within the year. When only one comet passed perihelion in a given year the numeral 'I' is omitted." So, not particularly relevant here. https://pds-smallbodies.astro.umd.edu/holdings/pds4-compil-comet-v1.0/phys_char/ by Benjamin Hirsch 9:49 AM

1 Like reaction. 1 yes by Thomas Cornet 10:00 AM

Can we see a SEED example label? by Lyle Huber 10:15 AM

<apollo:seed_file_name>xa.s11..att.1969.202.0.mseed</apollo:seed_file_name>
<Internal_Reference>
<lidvid_reference>urn:nasa:pds:apollo_pse:data_seed:xa.s11..att.1969.202.0::1.0</lidvid_reference>
<reference_type>ascii_to_seed</reference_type>
https://pds-geosciences.wustl.edu/lunar/urn-nasa-pds-apollo_pse/data/xa/continuous_waveform/s11/1969/202/xa.s11..att.1969.202.0.a.xml by Ethan Schaefer 10:19 AM

title: DDWG Notes 2025-04-03
layout: default
date: 2025-04-03

April 3, 2025
Notes by Debra Kazden

Known Attendees: A. Bailey*, J. Blalock*, B. Deen, B. Hirsch*, L. Huber*, R. Joyner*, D. Kahan*, D. Kazden, P. Lawton, J. Mafi*, E. Schaefer*, M. Seritan and M. Tiscareno* (*'s indicate people who voted during this meeting - if applicable)

DDWG Agenda

(Included in email sent April 1, from J.Mafi, Agenda for 2025-04-03)

We will have a DDWG telecon this Thursday, April 3rd at 9:30 PDT (12:30 EDT, 10:30 MDT). We will cover the following issues:

1) Action item reports.
2) CCB#59 funding year only allows the value "9999" – There were major concerns expressed about the implementation, which uses “ASCII Short String Collapsed” instead of an ASCII date type. We will discuss alternatives and decide upon a way forward.
3) CCB#57 Allow Encoded Native in File Area Observational Supplemental – Please review the example label excerpts posted by Bob Deen and be ready to continue the discussion on this SCR.
4) CCB#36 Add Mission Phase to the Investigation class – I’d like to restart the discussion on this SCR and identify next steps.
5) Preview of PDSMC F2F DDWG session agenda.

Front-matter/What’s new

1) Submitted to the CCB: (none)
2) New issues in GitHub:
   CCB#60 New permissible value for File Area Observational.Header.parsing standard id: "CDF 3.9 ISTP/IACG"
3) Ready for vote: (none)
4) Build 15.1 freeze date: 2/27/2025. The following SCRs are queued for implementation: CCB#13, CCB#48, CCB#49, CCB#54, CCB#55.
5) Withdraw by author:
   (Brief Discussion)

Action Items

a) CCB#39 – S. Hughes to update SCR
b) CCB#28 – J. Mafi to research
c) CCB#12 – S. Hughes to verify whether this is subsumed by CCB#52 and CCB#53
d) CCB#20 & CCB#32 – Tiger team formed and working the issue
e) CCB#31 – Ron and Matt working the issue **(Brief Discussion)**

Tier 1 issues

- CCB#60 - New permissible value for File Area Observational.Header.parsing standard id: "CDF 3.9 ISTP/IACG"

- CCB#59 - funding year only allows the value "9999"
- CCB#58 - New Product Native values for MSL/Mastcam
- CCB#57 - Allow Encoded Native in File Area Observational Supplemental
- CCB#56 – Reconsider the PDS4 build cycle cadence
- CCB#55 – Funding Acknowledgement class enhancement (funding source and funding acknowledgement text)
- CCB#54 – Add modification history for classes and attributes
- CCB#53 – Improve cross-namespace validations in sch files
- CCB#52 – Enable cross-namespace validations in sch files
- CCB#51 – Add an optional attribute in the Target Identification class to identify Primary/Secondary targets
- CCB#50 – Clarify Ambiguous Author Information in Product Document
- CCB#46 – Disallow multiple versions of the same LID from occurring in a single collection
- CCB#44 – Issues caused by the failure to expose pds: attributes that are referenced in an LDD.
- CCB#41 – Create a way for classes or attributes to be exposed to other LDDs but not to data providers
- CCB#39 – Fixed problem where LDDTool does not include all rules from IngestLDD
- CCB#37 – Add List Author, List Editor, and List Contributor to the Modification Detail class
- CCB#36 – Add Mission Phase to the Investigation class
- CCB#35 – Add Alias Index and Acronym Index to the Investigation class
- CCB#32 – Add Host type to Observing System Component
- CCB#31 – Ensure that dictionary versions are specified in labels
- CCB#28 – As a user, I want validate to allow delimited tables with whitespace-only numeric fields
- CCB#23 (JIRA-CCB-211) – XML/schema based files as archival data
- CCB#22 (JIRA-CCB-164) – Display Settings not required for images
- 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#21 (JIRA-CCB-367) – Provenance LDD Using PROV-DM
- CCB#20 (JIRA-CCB-358) – Instrument Package Context Products
- Tiger Team: M. Tiscareno (lead), B. Deen, B. Hirsch, S. Hughes, R. Joyner, J. Mafi, E. Schaefer.
- CCB#12 – Document Nuances for Reusability of Generic LDD Classes/Attributes
- CCB#5 – Move to V.2.0.0 for June 2025 Release
- CCB#4 – Backwards Incompatibility policies
- -Presentation: Backwards Compatibility
- CCB#3 – Adopt Semantic Versioning with version 2 transition
- The CCB returned this, requesting that it be resubmitted as part specific proposal to transition to IM V2.0.0.
  (Some Discussions)

Tier 2 issues

1) 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!

1) JIRA-CCB-326 (link) - Inconsistent Direction on How to Use Discipline Facets
2) RS: 259/281 (see new attachment to JIRA-CCB-281) (on hold)
3) JIRA-CCB-364 (link) – Disallow bundles referencing collections by LID **(Not Discussed)**

Historical links:

- Export of history: https://github.com/NASA-PDS/PDS4-CCB/tree/main/data
- Backup: https://nasa-pds.atlassian.net/jira/software/c/projects/CCB/issues (limited seats)

Next meeting: 17 April 2025

DDWG

Front Matter

There is a new issue - CCB#60 (New permissible value for File Area Observational.Header.parsing standard id: "CDF 3.9 ISTP/IACG"). It's a new permissible value for CDF. Just a standard update for the new version of CDF that's come out.

Action Item Reports - Joe looked at CCB#28 (As a user, I want validate to allow delimited tables with whitespace-only numeric fields) and posted a comment to GitHub. The proposal is to add space as a delimiter for DSV files, but that creates a potential conflict with the DSV spec if it were to be added. Could end up with files with different field delimiters according to the DSV spec and PDS4 label. Not ideal, but we can discuss it.

Question: Other action item reports?
Answer: (Silence)
~ For CCB#20 (Instrument Package Context Products) there has been some progress. Hopefully, the Tiger Team is getting ready to having more DDWG discussion. They are still working on it.

Note everyone is here to report on additional action items.

Tier 1 Issues

CCB#59 - funding year only allows the value "9999"

We started to hold an evote on this, but there were major concerns regarding the implementation chosen for this - using ASCII short string collapsed instead of ASCII Date YMD to restrict the use of z in the year. There's been a lot of discussion. Want to try to get a consensus today on how we want to move forward and implement this. There are two proposals. One is to change it to ASCII YMD and

live with potential abuse. Probably not a major concern for this attribute. The alternative is to create a new data type, ASCII Year and use the base of an integer to make it easier to restrict the values. Need to consider potential effects on publication year, which might have a bigger time span it needs to support, including three digit years or negative years.

Question: Any thoughts on this?
Answer: Someone posted a comment a little while ago wondering why this is being limited to a year. We could use a date.
~ Someone else said their understanding is that this is the program year, but not sure if a date is more applicable to other types of funding.
~ Hypothetically, it could be a date. Just an idea, but if year is more appropriate that's fine.

Another person is really sorry if it was unclear, but they think we should never have dates as strings with an ad hock format. If someone wants a whole date that won't harm anything. This was supposed to be easy. We should make this ASCII Date YMD and move on.
~ Someone wants to suggest we do something reasonable, obvious and simple.
~ Concerns were brought up while trying to implement this. Hadn't recognized possible problems. No problem with the suggestions here. It will be checked in peer review. No need to design something new or complicated.

Question: Other comments?
Answer: (Silence)

We will do a strawman poll on approaches. We can define ASCII YMD, we can put restrictions on the width of the field to limit it to just year. Lets vote on restricting to just four digits or full date and... (interrupted)
~ If doing plus or minus has to be more than four digits, and if add the z it could be up to six digits.
~ No one is funding in a negative year.
~ Could cause issues for publication year.

Question: Are we talking about changing the definition of ASCII Date YMD?
Answer: No. We would restrict the length.
Another Question: How would that affect publication date?
Answer: It wouldn't. The concern was with negative years, but we don't want that for funding year.

Question to Ron: Two weeks ago we talked about getting this into our current build. Is that still possible?
Answer: Don't have a good answer. It would depend on if the CCB pushes back.
~ It's safer to go with what rocks the boat less. ASCII Date YMD, without restrictions.
~ Someone agrees.
~ Another person said their preference is a string. Don't understand why the CCB got all wrapped around using YMD, it's just the Y.
~ And the other dates are based on string.
~ Exactly.

Let's have two votes on the preferred approach - restrict length to just year - will change back to ASCII Date YMD, so first restrict or not restrict length.

Strawman Poll - restrict or not restrict the length to year only?
ATMOS - Yes, restrict.
EN - Yes
GEO - No
IMG - No
IPDA - Not Here
NAIF - Abstain
PPI - Yes
RMS - No
RS - No
SBN - No

The vote is no. We will remove restriction and use YMD.

Next, the approach - consensus vote. See if there are any objections to implementing as YMD with no restrictions on length.

Question: Why aren't we restricting if it's just a year? Any advantage?
Answer: Restricting is more trouble than it's worth.
~ No, it's the opposite of that.
~ Someone thinks the concern is what is going to happen in the CCB.
~ If EN gets to weigh in, will tell the CCB it is a dumb decision to do free form.
~ Not free form.
~ If we want to define ASCII Date and restrict it that is more work for EN.
~ We might have funding class that are more than a year.
~ A use case. Doesn't harm anything to allow it.
~ Okay. We will move forward and try for another evote. It's easy if there is no restriction.

CCB#58 - New Product Native values for MSL/Mastcam

One comment added to CCB#58. A document was posted to describe the data format for the data. We can start moving forward on this soon.

CCB#57 - Allow Encoded Native in File Area Observational Supplemental

Concern about if this is needed. There are two points of view, at least, on if we will allow native data files in the archive.
~ The example of Apollo points out how easy it is to mess up. Supposed to have a pointer. They do, but it points to the documentation. Hard to validate pointer points to what you want. Possible and easy to screw up. Proposed mechanism is to package together. Enforces that you do have the observational date product.
~ Someone's concern, yes, Apollo is done incorrectly according to the SR, but thinking that was a problem that should have been caught in peer review. Would still have to verify that the product is equivalent. Doesn't get rid of the need for peer review. Prefer keeping non-compliant data much more separate.
~ True. It doesn't prevent gross negligence, but think it's better to keep them tightly coupled. The product is the product with a supplemental file. Philosophical difference of opinion.

Question: Is this to be a replacement to having a full product native label?
Answer: Yes.
~ Product native label should still be enforced. This can be an additional feature.
~ Duplicating native product twice.
~ No. No separate LID. Product native needs a separate LID and full information.
~ Would have file area with encoded native. Would describe the content as native, but there would be no separate LID. It would be part of the same LID. See that as an advantage.
~ Someone would need to think about that.
~ Another person agrees. For the Apollo example, this might have prevented the mistake. Think that coupling could be valuable. It makes sense intuitively.
~ Yeah. Want to emphasize that Apollo bug is a side effect. Did this for the tight coupling. Product native hasn't been used much.
~ InSight used it. Looked at those. Pointer is also to a document instead of to the equivalent.

Question: Another concern. If we have a label with both products, how do we get around delivering both? System issue, but if we don't deliver both than user is getting metadata describing something they don't have.
Answer: Unsure if that's a real concern. Not like these are mosaics.
~ We have gotten complaints from people having to download extra data.
~ That's a primary reason why this was framed they way it is, so novice users don't easily find things they don't want. An important reason to reject this. We did make a decision to keep non-archival products separate from the archive products. Nothing prevents archiving node from putting them together in the file structure.
~ It's annoying to have to repeat metadata. Multiple downloads is a node concern. Think having native as is own product makes it easier to find and not know there is something else.
~ For InSight, there was a way of getting everything, format, into PDS. Was a compromise.

One concern for me, as proposed, this is totally optional. It can be done either way. Concerned about inconsistency across the nodes. Unsure if we are worried about that. Unclear if it implies something about product native.
~ Someone doesn't see that as a concern. Things are done different across the nodes anyway, and they would all be clearly marked as native.
~ Another person agrees. Don't want novice users to find data this way, but there is a balance, we don't want useful data to be hard to find. We need a reasonable balance. Stuffing it into a product supplemental makes it a second class citizen.
~ Good point. It is a supplemental product and there are other supplemental products, like TIFFs or JPEGs, so a download system could just download observational data. Might mitigate the download problem. Would have to ask for the supplemental products.
~ That is done with allowing encoded native in File Area Observational Supplemental in the analyst notebooks.

It's kind of hard to discuss how this will look to users without the interface. We don't know what user facing registry will look like. It's currently set up in what we thought would make it clear, which is the primary product. If this is better we should deprecate the other way and not do it two separate ways. The whole point of PDS4 is to put this into a long term format. Believe everything in supplemental has to meet long term standards too. Interface needs to make it clear that this is a second class citizen.
~ We can request that for the interface.

This would differ from previous uses. Previous use for SEED was the primary user format. The other way was included as a secondary to comply with PDS4 standards. This is the opposite. Not what users will want, it's what the DP wants to include.
~ Right.
~ Apollo was encoded native, so kind of first class. Philosophical. This would be second class.
~ We could write the documentation to reflect all that.

Question: Any more comments?
Answer: Think Product Native is supposed to be the original format.
~ Correct.
~ Applies to SEED and this.
~ This is the data that was the telemetry converted into something archival. Not something people would want to use.

Question: This is regarding the Malin Cameras, where they downloaded as JPEGs?
Answer: There are a few possible formats, but yes, telemetry data stream comes from the spacecraft.
~ We should continue to think about this. It was a good discussion.
Question: Are there any action items? Do we want an opinion for EN about first or second class products?
Answer: And clarification on if can just download part of a product? Think the answer is yes.
~ More precise to say just download the observational product, not the supplemental. Could clarify the SCR if that would help.
~ Maybe we can discuss this at the F2F.

Question: Do we have a different product type for when people want it in something like an Excel spreadsheet?
Answer: Maybe Product External. Something to think about. A whole other discussion.
~ For HiRISE the archive products aren't the actual products that everyone uses. They use the browse product. For migration we converted them into GeoTIFFs.
~ Helpful.

On the other SCR,adding the new format is independent of this. Requires an MC vote. Maybe it can be next week.
~ It needs to go through the DDWG and CCB process first. Probably could be voted on at a monthly meeting.

Hopefully a lot of you will be at the F2F next week. A meeting link will be sent for Tuesday, for the DDWG open session. Current topics that mostly apply to the DDWG are the Provenance LDD and PDS4 version 2. That will be about what people want to see in version 2. It will be an informal, non-binding discussion. Please give this some thought. We will also discuss the PDS4 build cycle.

Question: Any last comments?
Answer: (Silence)

Next meeting in two weeks - April 17

title: DDWG Notes 2025-04-08
layout: default
date: 2025-04-08

April 8, 2025
Notes by Debra Kazden

Known Attendees: C. Aiken, A. Bailey, D. Hollibaugh Baker, T. Bar, J. Bauer, J. Chau, J Contreras, K. Crombie, B. Deen, L. Givens, K. Hansen, B. Harris, L. Huber, S. Hughes, S. Joy, R. Joyner, C. Juarez, J. Kang, D. Kazden, P. Lawton, S. Loftin, T. Loubrieu, Y. Ma, J. Mafi, J. Manweiler, I. Moon, T. Morgan, L. Neakrase, J. Padams, A. Raugh, E. Schaefer, M. Seritan, M. Staran, T. Stein, M. Tiscareno and R. Walker

DDWG F2F Discussion

There are a number of items to discuss today, including (not necessarily in the order they will be presented):

1) Reconsider the PDS4 build cycle cadence (Jordan and Matt)
2) PDS4 LDD for Provenance (Steve H.)
3) Collection Inventories and Versioning (Anne)
4) Brainstorming PDS4 Version 2 (Joe)
5) Search API training with Peppi (Thomas) (not covered by these notes)

Presentation: PDS4 Provenance LDD using the W3C PROV Data Model (Text under slide headings is conversation had at that time, not slide context.)

Presented by John (Steve) Hughes

Slide 2 - Topics
The problem is that these are Steve's use cases. He needs more from others to be able to move this forward.

Slide 3 - Provenance

Slide 4 - Previous Work
Early on, the decision was to use the W3C PROV DM. Across the nodes there are many different definitions of provenance and variations on how to use/share data. We need a common language and definitions. Tried to define a flexible model. Implemented as a PDS4 LDD. Will share some use cases.

Slide 5 - PROV-DM: The PROV Data Model W3C Recommendation 30 April 2013 Diagram with Entity, Activity and Agent. High level view. Abstract. Very general. Needs to be mapped to PDS4.

Slide 6 - 11 - Use Case 1 – Voyager Corrected Image
Comment: Since this is using Voyager ISS as an example, already in PDS3 with steps, unsure about IMG, but RMS is going to migrate. Wonder which version this is. Question is what was started with.

~ Old document on Voyager calibration. Just something Steve grabbed to help show a possible example.

Slide 12 - Use Case 2 – Relation Between Two Superseded LIDs Old LID and new LID for the same object. Want new label to say this supersedes that.

Slide 13 - Superseded LIDS - Use Case 3 - Uni-Directional New label, but can't put in metadata that says it's superseded.

Slide 14 - Superseded LIDS - Use Case 4 - Uni-Directional

Slide 15 - Superseded LIDS - Use Case 1 - Bi-Directional People didn't like this one. Too much work for this use case.

Slide 16 - Superseded LIDS - Use Case 2 - Bi-Directional

Slide 17 - Superseded LIDS - Use Case 5 - Uni-Directional This is about superseded LIDs, not products. Different. So, different use cases to support products. The term "supersedes" is too abstract. Many other possible terms could be used. Need to capture those.

Question: What's different? Derived from. How is that superseded? Just a different kind of product?
Answer: These are LIDs. Need to document it if a LID is different.
~ Last example changes the product to point out two different concepts. We are not discussing superseded products now.
~ The DDWG discussed this a lot. The normal way to supersede is to use the same LID with a higher VID, so unsure why we can't just say that's implied.
~ This is just providing a mechanism to capture information. Nodes, curators, have a deeper knowledge of how they want to capture this, but need to consider if the system wants the deeper knowledge.
~ We have Modification Detail.
~ That's text. This is a formal way of capturing provenance that can be used by the system. Captures nuances of why a change was made.
~ Someone recalls the DDWG had some use cases. Agree that superseded works with higher, better, newer versions, but there were a few use cases presented that showed a different use of version ID, like where products were processed differently, but had the same LID. By implication, those would be improved, new versions, but that was not actually the case.
~ Yes. We could re-hash all day. Jordan said the registry will give users products with the highest VID, so we might want to document it if we do something different.
~ Someone agrees that Provenance LDD is most useful in cases where things didn't go like in the usual process. Sees the huge utility in being able to describe unusual issues.
~ If software is archived, Provenance will be able to track versions over time. Could develop into something very powerful.

We need to get back on topic.

Slide 18 - Simple Implementation
This is an example of how to say this LID supersedes this one.

Question: Comments?
Answered with Another Question: Thinking outside the box, when you say that, can it be anywhere? Like if you deprecate an entire delivery - can there be a single product to list them all?
Answer: Good point. First we need the vocabulary and definitions. How it is implemented is step 2.
~ Problem when you start putting stuff somewhere else is that users might never see there's an impact. No clear description of why it's different so that's not recommended. It's not clear to users. They don't always get the global view.
~ If they get an old product, that doesn't say it's superseded, so the problem isn't any worse.
~ Having been on the IM and implementation side, think this is a bad thing. Should never happen. Should make this hard so people don't do it.

Slide 19 - 26 - Use Case 3 – Collection Stewardship Collection stewardship is related to CCB#45 (Add reference types values for *-to-agency), relates to the node who curates. It didn't make it through the CCB. This is another attempt. It came from ChatGPT. Hard to explain the difference from curation to archive. Management is curation.
~ Can do lots of associations for relationships. Can get pretty detailed or keep it very simple. Sort of a what-if.

Slide 27 - 34 - Use Case 4 – Planetary Science Observation There was discussion in the DDWG due to a request from PSA. They had more than one target and needed a way to identify the primary target.
~ This was an attempt to show that observations could be captured by PROV.

Slide 35 - Questions and Answers
The point of this was to say that we have an LDD for PROV. The priority is to capture superseded LIDs, but we can also create PROV LDDs for nodes and agencies or an observation itself. IM tried to capture that with Observation Area, but this can be a lot more specific.
~ Someone really liked the observation example, but a metadata product only. Almost need metadata only to describe a campaign.
~ Only three elements of PROV LDD - element, agent and activity, but could have a whole ontology and reference with provenance.

Question: In provenance, isn't it inherited? So if no changes, isn't it just inherited across? Seems like this is about exceptions.
Answer: Could be inherited, but this could become a whole cottage industry. Pandora's box.
~ It's been mentioned a few times that this might be for exceptions, but there are lots of possibilities. Seems like if exceptions are going to require this, we should define what exceptions are and be very explicit.
~ We can have several LDDs for different use cases. The DDWG tries to make solutions very broad, but that can cause problems. Can extend this to be more specific to individual disciplines. Like with the IM, first we did a common IM and then we extended it. The Provenance data model has been expanded into several areas.
~ Thinking about VID, we should have explicit understanding about upping a VID.
~ And we can document what the system understands that meaning to be.
~ Risk codifying current behaviour too.

Great discussion, but the scope is huge. Propose we make this happen for supersession. **Action Item is to EN**. This is a problem that needs to be solved.

Question: Can we have a proposal for how to use this before the next MC - for context products? Is that reasonable?
Answer: We need a Tiger Team. PPI has this problem. IMG too. We need someone to support Steve in this.
~ This is an LDD.
~ We should add documentation to the SR and DPH and have warnings that people should never do this.
~ Strengthening the SR and DPH on when to use or not use this is important.
~ This should be in the SR and a policy or standards document, not the DPH. That's just a handbook.

Question: Last comments?
Answer: (Silence)

Discussion: Reconsidering the PDS4 Build Cycle Cadence

Led by Matt Tiscareno and Jordan Padams

This is not that big an issue, but it is worth some discussion from the DDWG perspective. The DDWG works on a bunch of stuff, and then, all of a sudden, it's too late to get things in. Wanted to bring this up and ask about the deadlines, but also to support EN.

The schedule is for the system. Twice a year the process is done, but the cadence is flexible. We can do it more often. It's twice a year now, could be more often. Could do two or four times a year. Could skip it if nothing to add. Doing twice a year now because it has become an expectation, but it can be changed. There will still be specific dates for the DDWG.
~ That is a lot more than what the DDWG is suggesting. That would be a lot of work for EN.
~ When the DDWG is frozen the release is about two months later. That's a common complaint. Trying to condense that down.
~ Think the DDWG should try to be a little ahead of the curve for timing.

EN can do something to make dates and gates more transparent. Ideas welcome. EN is open to being flexible.
~ Someone has a concern - we shouldn't increase the number of builds. That would kill Steve and Ron, but it would be nice if small things could be handled.
~ That's where having more would be helpful, and it might be easier for Steve and Ron to get everything in. Or the DDWG could decide things like enumerations don't fit in the build cadence. Tough because version would go off cadence, but maybe that's fine.
~ Someone is getting a sense from the DDWG that there's a false deadline. They seem to wait until the last moment to get things in. Should make offset deadlines.

Question: Users expect two builds a year? Who are they? Seems like it's just us. Am I missing something?
Answer: IPDA too. Would need to make it clear if we are going to make a change, but it's fine to do that.

Question: Are we proposing once a year and as needed?

Answer: Yes. Would meet the DDWG needs. PDS4 is fairly mature at this point. Not a lot of large changes. Doesn't seem like things should be hung up on a must do by this date schedule anymore.

Question: As needed or once a year?
Answer: We always thought we could do point builds, but we never do. Once a year seems close to sufficient for our needs.

Question: Is release on demand positive or negative for developing missions?
Answer: A lack of consistency can be difficult for DPs, I think.
~ Would imagine they could ask for another release.
~ Maybe missions needs could help drive the schedule.
~ There is more that one mission operational at a time.

Sounds like we are agreeing to a more ad hock nature. As needed or not.
~ When someone thinks about this, there are things that need work, like interfaces, but others are very stable. Maybe we should not have them all on the same cadence. Vocabulary is very volatile. An active area of development. Need a graceful way to do those without needing a whole system build. Interface changes need to be on a longer schedule. Annual finally okay for data types too. Can separate the IM modules. I usually test in an LDD now, to test use cases and such. Discussed twenty years ago the idea of promoting things to the IM.
~ Having the same thoughts. If the structure matters in deadlines, then yes, different cadence for enumerations, so the questions is how easy that is.
~ Super easy. Fine by me (Jordan). DDWG can define requirements. Still have a build cadence and do point builds for things that need to be updated more often. Maybe quarterly instead of as hock. It's hard when there is a full release - there's a lot of testing to do, so knowing when that is coming is good. Ad hock could be hard on schedules.
~ Enumerations can be separated out, but it can get complicated. EN can support whatever we want to do. EN can put together a proposal so enumerated list updates would be off cadence - quarterly. No SR or DPH updates, just tag a new version of the IM. Can try the beginning of next year so we can discuss this with IPDA and our other colleagues.

Steve and Ron will be assigned the task of coming up with what we will do. **(Action Item - Steve and Ron)**

Presentation: PDS4 IM V.2.0.0

Presented by Joe Mafi.
~ Slide 1 created before the meeting. Slide 2 created during the meeting.

Slide 1 - PDS4 IM V.2.0.0 To-Do List Originally there was one slide in this presentation. A to-do list. Want us to brainstorm about what we want.

Slide 1 includes that we need to resolve CCB#s 3 and 4, make a Tiger team for 2.0 (Interested people can let Joe know. He is hoping to start work soon) and that we need to come up with an implementation plan.
~ The Tiger Team will work on the implementation plan. They will include whatever SCRs need to be resolved, what changes/deprecations we want in version 2 (which depends on CCB#4) and develop schema and schematron. They will also develop a procedure for changing versions in the future.

Question: Are there specific SCRs or updates that people want to see in version 2?
Answer: Someone has a list. Working with the CCB led to the first one. Finds with random spot checks that there are wasted CPU checks in schematron. Schema validation that is already happening is better, so since we have to pay for all CPU work, we should do an efficiency check. It's taking too long to validate large data sets and bundles. So, the bullet should be scrub schematron file to make it more efficient to execute and use.

Question: Why is this tied to version 2?
Answer: It doesn't have to be.
~ Should be an on-going effort.
~ A concern because of how long things take to validate. I end up writing a lot of test data.
~ EN would love some test data. They capture test data for every SCR.

Related to version 2 - we have no backwards compatibility in our labels so we can't have semantic versioning. We need to address that. The dependence of specific minor versions of IM schema has to be fixed before we go to version 2 if we want semantic versioning. Only the first digit is significant for validation.
~ The current idea is that we can validate to any specific schema, but this only would work if there are no backwards incompatible changes.
~ In version 1 we have allowed backwards incompatibilities, so this would only work if that is not allowed in version 2.
~ It goes back to the build schedule too.
~ Yes.

Feeling strongly, when we go to version 2, we should already be in good shape to work with our key partners so interfaces can be interoperable and work together (DOI interface, IPDA etc). The bullet is reconcile PDS IM with our key partners.

We need to independently do a reconciliation of our attributes to do open science.

Enumerations - we need to define language that can be re-used. This is non-trivial and non-backwards compatible if we want to do validation.

From these requests, a more comprehensive test suite is really the requirement for version 2.
~ Yes. And there are more things people want.

Someone's pet peeve is the MSN LDD - that most of it is mission specific and not in the core. Would like to consider that. So, bullet is promote MSN to the core.
~ Mission LDDs would still be important too.
~ Submitted an SCR - then a solution was pointed out, so it became documentation only.
~ That sounds like documentation that would go into the standards recipe book.
~ The DPH would be fine, but don't know mechanics of how to make that update.

Question: Are the LDDs stable? Would propose that we consider that. And shouldn't there be it's own standard for LDDs? Like for Geometry?
Answer: MSN could be used as-is, but isn't, and has a lot of things that should be in the core.
~ The discipline LDDs are standardized. They are hard to use because of the standards values.

~ We should look at the discipline LDDs. If something is core to how we do PDS than maybe it belongs in the core.
~ Common in common. Experts in that discipline should control discipline LDDs.
~ The disadvantage is that is tied to the build cycle.
~ Someone agrees that common is common. LDDs are supposed to be more responsive to changes.
~ If we put them in common than the whole DDWG gets to make decisions about whatever that discipline is, not just experts.

Disagreement about usefulness of mission phases in Common.

We will continue adding to the list, but wrapping this up now.

Someone recommends that the DDWG vote on the must-haves, then we can implement them and the should-haves can be done later. Let's not hold up the move to version 2.

Question: What is the purpose of the Version 2 Tiger Team?
Answer: The purpose is to not fight this out in the DDWG. The Tiger Team will design something and present it to the DDWG. Hoping to have CCB support too.

People will let Joe know if they want to participate.

Slide 2 - V.2.0.0 Implementation Specifics
Created during the meeting includes the following:

- Scrub of schematron file to make more efficient to make and use.
  Ongoing, not tied to a new build.
  Jordan asked for test data
- V2 (minor) semantic versioning
- Citation of LDD versions (IM attributes)
- Reconcile IM interfaces to external partners
- Reconciliation of IM attributes to Open Science + FAIR requirements (licenses, formats, etc…)
- Enumerations to vocabularies
- Promote MSN dictionary to common

Presentation: Collection Query Behavior

Discussion led by Anne Raugh

Slide 1 - Collection Query Behavior Very specific cases, possibly more relevant to curation than to data search.
~ Wants to know if people believe what she does.

Slide 2 - Perfect world assumptions

Slide 3 - Question 1 - When an agent requests a list of all LIDVIDs that are members of version X.y of collection A, what should be returned? For questions 1, 2, and 3, Anne is assuming the answer is the first statement. So, 1, every LIDVID goes in an inventory table.

Question: What should be in an inventory table?
Answer: That will be answered at the end.
Another Question: What's the point of a collection? That impacts the answers.
Answer: Asking mechanical questions, not philosophical.

Question: Should there be multiple versions of LIDVIDs in a collection?
Answer: That's at the end.
~ A collection is a product, and that conveys information. Should include multiple versions of LIDs.
~ You are jumping ahead.
Another Question: Someone is confused by the question. If asked for a specific version, shouldn't you get that?
Answer: That's collection version. Products could have different versions.
~ Technical answer - product is an inventory table. So, collection product is the table, so a query is ambiguous. The question is if you want the inventory.
~ Not in this example. Everything in an inventory is a member.

Someone always thought that if there were multiple versions in an inventory that that's an error.
~ End.
~ The definition of agent allows too much. Would write code.
~ We are discussing the base action of the registry now first.

Question: To clarify, collection version 3 could have products version 5 and then could be 6 later, so question is what should be returned?
Answer: This was a very simple question. We are spending too much time on it.
~ Confused the issue by putting option 2 in there. Superseded LIDs should be removed. Only should be one version of a LID in an inventory table. You confused and irritated everyone.

Slide 4 - Question 2 - When an agent requests a list of all LIDVIDs in no-version-specified Collection A, what should be returned? (Anne's answer: Every LIDVID in the inventory table for the highest archived version of Collection A.)
~ We need to be thinking about certified data.
~ Depends on who the agent is, who has permission to look.
~ If a node requests everything for a collection, the question is what they get.
~ A DP might get a different answer from a user.
~ If not authorized to see the data won't see the collection.

Question: Why isn't there more than one answer?
Answer: That's not the question. The question here is vanilla.
~ It's not specific enough.
Another Question: Is there a distinction between asking the registry or the inventory table?
Answer: For this, assuming all queries are to the registry.
~ Most users wouldn't think about options. They would use the default usage. Most users won't care. We don't want access to the registry to be complicated.

Slide 5 - Question 3 - Is it valid to have multiple versions of the same product in the inventory table for a Collection?
For example:
P, urn:nasa:pds:bundle:collectionA:product_1::1.0
P, urn:nasa:pds:bundle:collectionA:product_2::1.0
P, urn:nasa:pds:bundle:collectionA:product_2::2.0
P, urn:nasa:pds:bundle:collectionA:product_2::3.0
P, urn:nasa:pds:bundle:collectionA:product_3::1.0
P, urn:nasa:pds:bundle:collectionA:product_4::1.0

The question is if it's valid to have multiple versions of the same product in an inventory table for a collection. The example is a collection with several versions of the same product in the collection. It's valid, but unclear if it should be.

In the room, 2:1, people think it should be invalid.
~ It is currently valid, but the question is if it should be. One node does this, maybe it's not a great idea, but they have it.

CCB#46 (Disallow multiple versions of the same LID from occurring in a single collection) was filed to forbid this. The standards are unclear. The DPH and Concepts document seem to say this is not allowed, but it passes validation. The SCR did not pass. If we think this should be allowed, should give it a different LID.
~ This is a practical problem for DPs when calibrations change over time. You get into a peer review can of worms when it's called a new product. As a DP, don't want to have to keep going through peer review process every time I change numbers, so I have done what's in Anne's example. Disagree that the SR and DPH say this is prohibited. Using it. Maybe it's not the best way, but validate should have clear rules.
~ That's helpful.

Question: Accumulating data sets don't get peer reviewed every time, right?
Answer: In accumulating, generally, if doing the same thing each time, don't need a new peer review.
~ People are confused about why this is necessary.
~ Reviewers start from scratch every time they get a new file.
~ We used this method when we started using PDS4. If need to change, okay, but need to put rules and expectations in for what gets a new LID or version. If going to make this invalid need to be specific on what that means.

Question: What do we gain by making this not allowed?
Answer: More philosophical questions. Most users want the most recent version, and multiple versions could lead to confusion.
~ Someone understands that, but users ask for what they want based on a results list.
~ CCB#46 was concerned with data products. Calibration products are a different case. Don't want users confused by a bunch of superseded old products.
~ GEO doesn't do a new peer review because they get a new calibration product.

Someone doesn't understand why version 1, 2, 3 of same product exist. Think users wouldn't want to see anything besides the best one.

~ Calibration and other things change during a mission. Changes to things that apply at parts of time, and we need contemporary files.
~ Suggest we call them cal-alpha, cal-beta, etc.
~ Ah, the terrors of putting information into the file names. This is a design decision by missions.
~ Instead of looking at a LID, would have to look at LIDVID.
~ Calibration files are linked to the data files at some nodes.
~ The question seems to be in context of LIDVID. The question is what that means.
~ Divide the question. If there's a new VID that means there's a new version, it doesn't mean the previous version was superseded. Being able to reference earlier versions is important.
~ There are two dimensions to versioning calib files. In one case, the new version is better. The other way is for a different time. As a DP, one regret I had was not putting a time stamp on my calib files. Conflating two issues.
~ We have nuance of version, but we haven't come to consensus on what version constraints are versus what it means to change a LID. Maybe we need some direction in the DPH.
~ Like how we need Provenance LDD when there is an exception. We need to decide on the typical way this gets done and what the default version is versus what to do when there's an exception.

Question: When saying that version 1 is older, then what happens if I miss a version?
Answer: That's a grenade. The question is if version numbers have to be contiguous.
~ No.

Version number just means there's a difference, which is the whole point of the Provenance LDD. Can't capture nuance in the SR.
~ Discussion is what is the nature of versions. Superseded can be taken off-line, but I wouldn't do that. Think if user wants to reproduce old work they need to be able to get that.
~ Someone agrees, but that doesn't have to be online.
~ Superseded should be accessible.
~ Yes.

This topic wasn't originally supposed to be about calibration. Even SPICE kernels have the newest version online.
~ Need definition of version.
~ Need definition of superseded.
~ Need to agree to disagree and move forward.

Question: Is there something we can vote on or that can be actionable?
Answer: The first decision to make is the purpose of incrementing a version, then we can move forward.
~ Consequences of incrementing a version.
~ Sounds like some nodes think incrementing means superseded. Others don't.
~ Also don't agree on what superseded definition is. Need to be explicit. Need a clearly stated answer.
~ The supersession question should be a DDWG discussion. For supersession, treating them like bugs or mistakes. Thinking about superseded data users and would like to know what we want the API to return. The default now is the latest, but we can also build in possibility to request more.

Lunch Break
Back to Anne.

A quick recap - For question 3, getting the idea that people think this shouldn't be allowed.

Slide 6 - Question 4 - When new data are added to make version (X+1).0 of Collection A, does the inventory table: 1) Contain all the LIDVIDs from the X.0 version of Collection A, plus the new LIDVIDs; 2) Contain only the new LIDVIDs; or 3)Something else ? Question 4 is only true in narrow case when only added new data.
~ Yes.

Slide 7 - Question 5 - When the intent is to drop products that are in version X.0 of Collection A from version (X+1) of Collection A, then: 1) The LIDVIDs that appeared in version X.0 inventory table will not appear in the version (X+1).0 inventory table; or 2) Something else? If you supersede a product, it won't appear in the next version of the collection.

Question: Edge case. What if you have a data product version 1, and version 2 is then produced in a subsequent collection, then the PI says version 2 is ridiculous, only version 1 should be there - what happens? Do you make a version 3?
Answer: Should wait for this to come up at a review. Difference is superseding data versus withdrawing data. Part of curatorial process. Maybe you create version 3, maybe you withdraw version 2. Unclear.
~ It was an edge case.

Question: Can a product that is version 2 become version 3?
Answer: It stays the same version unless it actually changes.
Another Question: What if there's a previous version?
Answer: If we screw up then we'll fix it.

Slide 8 - Question 6 - Are the answers to questions 1-5 any different for "S" entries in the inventory table than it is for "P" entries?

Question: Secondary members of collection that are primary elsewhere, are they required to be returned with the collection?
Answer: Yes. Secondary is treated as primary as much as possible. Prohibit more than one version of a secondary product.
~ Someone is worried about things in the archive now.
~ Still working out details. Trying to plan for 100 years in the future, not 20 years in the past.
~ A lot to deal with. Fine with this going forward, concerned about what we already have.
~ We use that for documentation and we use LIDVID. Seems like a good idea. Unclear how that should work for calib.
~ Specific files referenced with internal references.

(End of Anne's Presentation)

Question: Is there anything that people want to discuss?
Answer: Maybe we can come up with some action items.

Over lunch, someone thought about this - as a user and DP, there are a lot of collections of software that won't work with different versions. Edge case, but there are times when you might want to make the multiple versions available.
~ Someone else totally supports that. Another reason superseded shouldn't go off-line. People doing research projects might need to continue with specific versions. Got to keep things available.
Question: EN wants feedback. Superseded and versions aside, but if there are multiple versions, what would people expect when asking the API for products? Default is to provide the highest, most recent version.
Answer: Someone would expect to get what the VID is. Without the VID would expect to get the latest.
~ The question is when there are multiple LID products with multiple VIDs.
Another Question: Most users, what would they want?
Answer: Latest version by default.
Another Question: If there are three versions, would they want all three or just the latest?
Answer: Need to decide what mechanism we should follow to make sure there aren't multiple versions in a collection, could be historical documentation, but newer versions should be used. Want to make that rule.
~ Someone agrees, but in some cases things are already archived with multiple versions in one inventory. So, for that, we want to know what should be returned.

Question: Do we need an SCR?
Answer: This is a policy.

If a user is asking for a LIDVID, they should get that, especially if it's associated with a DOI.
~ Someone feels like we need to document why the weird ones are weird and then also not do it anymore going forward.
~ There's a question online about users asking for a specific LIDVID, but maybe they don't know there is a better product available.
~ Maybe the API could inform them.
~ EN says the API can't do that. Can give people what the majority would want. The default. Will support the corner cases and have a mechanism for specific versions requested.
~ Not everyone is in agreement, but sounds like most people are.

Question: Do we need an SCR?
Answer: We have one, CCB#46.
~ We need one for software.

If I ask for a LIDVID, should get that, but if ask for a LID, get the newest. Should follow the provenance chain. Have software find the new LID.
~ That should be the case.
~ Challenging, but if you get an old product there's nothing to say it's old.

The DDWG will move forward with CCB#46.

Question: And the accompanying changes for the DPH? And still another document too? Just want a definite answer.
Answer from Jordan: As keeper of the document team, please no. DPH is already insanely large. Would like How-Tos broken out.

~ We need to start using Confluence.
~ Yes.
~ Requirements need to go into the SR. We need to do a better job of clarifying in both.

Question: Does this now state the fundamental that Anne's Question 3 (Is it valid to have multiple versions of the same product in the inventory table for a Collection?) is not allowed?
Answer: It needs to go through the DDWG process, but that's the proposal.

We need to start developing the Provenance LDD and move forward with superseded.
~ Think we just decided that incrementing a version number means supersession, so we don't need the Provenance LDD, but we do need it for edge cases.
~ Feeds together well. Will clearly define superseded and point users to Provenance LDD for edge cases, which we don't want people to do.

Question: Any Action Items from the build cadence discussion?
Answer: The **Action Item is to EN** - they will draft up splitting out enumerations versus build cycle cadence and see what's possible, check out the technical hiccups.

Question on the Provenance LDD: Going ahead, is this to support supersession?
Answer: That part will be developed first. The first priority is to support supersession. Can expand later.
~ Steve will be going forward with simple implementation, at least first. **(Action Item)**
~ As long as it supports what we need.

Action Item - Joe will get the version 2 Tiger Team formed and start working.

Question: Timeline?
Answer: Could set a target date. Unsure if there are any good dates to try to be ready by, to do with a mission.
~ The first target would be to identify for common data intended for us all and then think about changes.
~ Someone believes it will be an SCR. A monstrous SCR that will include tests, documentation, etc. It will take a while to review.
Another Question: Any concerns with that approach?
Answer: Someone is trying to think if there are any mission drivers.
~ A timeline would help move things forward.
~ Missions are Clipper and Dragonfly.
~ There are Clipper people in the room. Producing labels now.
~ Files are supposed to be released this calendar year.
~ Pipeline gets locked in early in the time line.
~ It is also too late for Dragonfly. This was supposed to be PDS2010.

Question: How about June 2027?
Answer: Based on the current cadence, would freeze February 2027. That would give us a little less than two years. Probably do-able. So we will target June 2027.

Question: Ancillary to a lot of the conversation, but as a user and software DP, how does the use of AI fit into all of this? A consultant just told me he could build an app in two months.

Answer: From EN's perspective, Steve's been using it extensively for IMs, and we are using it for generating test cases. Repos use it for code reviews, but think it's an unreached area of the PDS. Fascinating world. Maybe can simplify templates or use to develop code to use the data. An evolution we are still trying to work out. Biggest problem for PDS is that the detailed structure of the IM makes it hard for AI to use it. It just flattens it.
~ Some AI is being used to build new AI.
~ Yeah, many fascinating things we can do. Figuring out how to expose our rich data to AI and machine learning is very exciting.

(The next talk is a Tech Demo on the PDS API. DDWG notes stop here.)

title: DDWG Notes 2025-04-17
layout: default
date: 2025-04-17

April 17, 2025
Notes by Debra Kazden

Known Attendees: A. Bailey*, J. Blalock*, B. Hirsch*, L. Huber*, R. Joyner*, D. Kahan*, D. Kazden, A. Knight, P. Lawton, J. Mafi*, S. McLaughlin, E. Schaefer, B. Semenov, M. Seritan and M. Tiscareno* (*'s indicate people who voted during this meeting - if applicable)

DDWG Agenda

(Included in email sent April 16, from J.Mafi, Agenda for 2025-04-17)

We will have a DDWG telecon this Thursday, April 17 at 9:30 PDT (12:30 EDT, 10:30 MDT). We will cover the following issues:

1) Action item reports.
2) CCB#59 funding year only allows the value "9999" – This SCR has been updated to reflect the changes discussed during the last DDWG telecon. Please review and be ready to vote.
3) CCB#46 Disallow multiple versions of the same LID from occurring in a single collection – This issue was discussed during the DDWG Session at the PDSMC F2F meeting and a consensus was achieved on how to proceed. Please review this SCR and be prepared to discuss.
4) CCB#57 Allow Encoded Native in File Area Observational Supplemental – We will continue the discussion on this SCR with the goal of identifying any questions or concerns that need to be resolved before it can come to a vote.

Front-matter/What's new

1) Submitted to the CCB: (none)
2) New issues in GitHub:
   CCB#61 Disconnect between SR and Validation of ASCII Real
   CCB#62 Transfer Manifest class should enforce data types for its two mandatory columns
3) Ready for vote: (none)
4) Build 15.1 freeze date: 2/27/2025. The following SCRs are queued for implementation: CCB#13, CCB#48, CCB#49, CCB#54, CCB#55.
5) Withdraw by author:
   (Brief Discussion)

Action Items

a) CCB#39 – S. Hughes to update SCR
b) CCB#12 – S. Hughes to verify whether this is subsumed by CCB#52 and CCB#53
c) CCB#20 & CCB#32 – Tiger team formed and working the issue
d) CCB#31 – Ron and Matt working the issue **(Brief Discussion)**

Tier 1 issues

- CCB#60 - New permissible value for File Area Observational.Header.parsing standard id: "CDF 3.9 ISTP/IACG"
- CCB#59 - funding year only allows the value "9999"

- CCB#58 - New Product Native values for MSL/Mastcam
- CCB#57 - Allow Encoded Native in File Area Observational Supplemental
- CCB#56 – Reconsider the PDS4 build cycle cadence
- CCB#53 – Improve cross-namespace validations in sch files
- CCB#52 – Enable cross-namespace validations in sch files
- CCB#51 – Add an optional attribute in the Target Identification class to identify Primary/Secondary targets
- CCB#50 – Clarify Ambiguous Author Information in Product Document
- CCB#46 – Disallow multiple versions of the same LID from occurring in a single collection
- CCB#44 – Issues caused by the failure to expose pds: attributes that are referenced in an LDD.
- CCB#41 – Create a way for classes or attributes to be exposed to other LDDs but not to data providers
- CCB#39 – Fixed problem where LDDTool does not include all rules from IngestLDD
- CCB#37 – Add List Author, List Editor, and List Contributor to the Modification Detail class
- CCB#36 – Add Mission Phase to the Investigation class
- CCB#35 – Add Alias Index and Acronym Index to the Investigation class
- CCB#32 – Add Host type to Observing System Component
- CCB#31 – Ensure that dictionary versions are specified in labels
- CCB#28 – As a user, I want validate to allow delimited tables with whitespace-only numeric fields
- CCB#23 (JIRA-CCB-211) – XML/schema based files as archival data
- CCB#22 (JIRA-CCB-164) – Display Settings not required for images
- 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#21 (JIRA-CCB-367) – Provenance LDD Using PROV-DM
- CCB#20 (JIRA-CCB-358) – Instrument Package Context Products
- Tiger Team: M. Tiscareno (lead), B. Deen, B. Hirsch, S. Hughes, R. Joyner, J. Mafi, E. Schaefer.
- CCB#12 – Document Nuances for Reusability of Generic LDD Classes/Attributes
- CCB#5 – Move to V.2.0.0 for June 2025 Release
- CCB#4 – Backwards Incompatibility policies
- -Presentation: Backwards Compatibility
- CCB#3 – Adopt Semantic Versioning with version 2 transition
- The CCB returned this, requesting that it be resubmitted as part specific proposal to transition to IM V2.0.0.
  (Some Discussions)

Tier 2 issues

1) 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!

1) JIRA-CCB-326 (link) - Inconsistent Direction on How to Use Discipline Facets

2) RS: 259/281 (see new attachment to JIRA-CCB-281) (on hold)
3) JIRA-CCB-364 (link) – Disallow bundles referencing collections by LID **(Not Discussed)**

Historical links:
- Export of history: https://github.com/NASA-PDS/PDS4-CCB/tree/main/data
- Backup: https://nasa-pds.atlassian.net/jira/software/c/projects/CCB/issues (limited seats)

Next meeting: 1 May 2025

DDWG

Front Matter

Question: Any action item reports?
Answer: The CCB#20 Tiger Team is still making final touches to the SCR. Think they are getting close to ready to have a discussion.

There are two new SCRs this week:
CCB#61 - Disconnect between SR and Validation of ASCII Real, and
CCB#62 - Transfer Manifest class should enforce data types for its two mandatory columns

CCB#61 is from Ron. A lot of hoopla over a very little thing. Someone noticed the current pattern allows a value that's not allowed in the SR. Need to fix the SR, amend the pattern or tell them not to do that.

CCB#62 is from Mark Bentley. He's not here today.

We will look at both of the new SCRs in the weeks to come. Both are marked as bugs. Hopefully they are fairly straight forward.

There are a few agenda updates for today. Lyle would like to have a discussion about the location of the context products and Boris wants to discuss CCB#46. We will start with CCB#46.

Tier 1 and Other Issues

CCB#46 - Disallow multiple versions of the same LID from occurring in a single collection

This was discussed at the F2F. Basically, it would make it so that validate reports multiple instances of the same LID, but with a different VID, as an error. Boris posted a comment that this will cause big problems for NAIF.
~ The basic principal on which all SPICE archives are built on is that there is a single kernel collection, the latest version of which lists every single LIDVID in the archive. NAIF does not version kernels individually, but they do version meta-kernels. That's intentional, so that the latest one can be identified. The fact that the latest supersedes others in most cases does not mean that others

are not needed or desirable. In fact, they are, for two very specific cases. 1) An earlier calculation was done using a previous version of a meta-kernel. It should be available for immediate search. 2) Absolutely, experts would like to see the evolution of the meta-kernels in the stream from version 1 to version n, by comparing them to see what data was added or may have been replaced and so on. So, that's why the latest collection has every version. That means multiple versions of the same lead meta-kernel. All the archives NAIF has done to date and all the archives NAIF plans to do, because the tools rely on this and other things do too, so we will have to continue this into the future. There are ten archives, maybe twelve, and three quarters of them are continually accumulating and the ones that are not accumulating on a regular basis will be accumulating in the future because, by definition, there's a single SPICE archive at NAIF for every mission. So, if we prohibit multiple VIDs for a single LID in a collection that will break all of the SPICE archives and it will not give us anything that will be a replacement for that. Provenance LDD is a no-go for NAIF because, right now, we have minimal labels that serve the purpose and are the way the community has been dealing with SPICE for way longer before PDS4. The same exact principal was applied to PDS3 archives and all the tools that deal with PDS3 archives could be easily migrated to PDS4 because this is preserved. So this will all go down the drain, and to me, all of this hinges on one philosophical question - why listing all product LIDVIDs in an archive should be prohibited for the collection of that archive. Not sure who came up with this idea that it is something that should never be done, because that is exactly what this item is calling for. I'm done.

Question: Any responses?
Answer: Someone absolutely, 100% agrees that superseded products should remain accessible indefinitely, but think that the difference here is that there are other ways of doing it besides keeping it in the inventory continuously. I guess that is kind of where the breakdown comes in because, frankly, there was an agreement reached, although you weren't part of that conversation as far as I know, but other nodes who have been doing this purposely, and there are some, said that they would find a different way of doing what they have been doing.

Question: Why can't both be allowed? I am trying to understand why listing all LIDVIDs should be disallowed. Why not both ways?
Answer: Part of the reason is that if that happens in a another node inventory then it's an error. It is not intentional. So that is something that some other nodes would like to see flagged. It could be a warning instead of an error, but it should be flagged if it happens.
~ For NAIF inventory it is not just intentional, it is required. So, we have complete disagreement here and I just don't see how this being a way to deal with inventories in some nodes should completely invalidate inventories at other nodes.
~ So far, someone agrees. Suggest it could be a best practice.
~ This is absolutely the best practice. It is the only way SPICE archives were done. They are done by NAIF, US SPICE providers and by three international agencies. All SPICE archives are like this and tools rely on this and we expect future tools to rely on this.
~ Someone says they weren't very clear. What they were suggesting is that SPICE might be an exception, but the best practice would be to not allow the same LID from occurring in a single collection.

Question: How about we make a validate flag that does it and then everyone can be happy? NAIF would be allowed by the standard, but then people can enforce it and that's going to be best practice for particular nodes. What prevents us from doing that?

Answer: SCRs under discussion and right now, the SCR just says exceptions to this may be allowed. I think NAIF falls under that category.

Question: Did the MC take any of this as an action item? Heard kind of a brief summary of their discussion and it sounded like they touched on this, but I wasn't sure if they are waiting on our action or discussion on this.
Answer: Someone hasn't heard anything officially in terms of anything out of the MC.
~ Someone else says that they guess for most data products, they think, the principal that we are talking about definitely makes sense. If you have, for example, a Cassini image that has multiple versions, nobody wants to look at the evolution of this data product because the old versions had an error that has now been corrected, and we don't want people looking at the old version and thinking there's something in it that they should look at. That's the primary reason motivating this change. Recognize that that might not apply equally to everything and so we should make sure that it applies to what we want it to apply to and doesn't apply to what we don't want it to, but that doesn't mean making it optional. Think we need to be clear about what it applies to and what it doesn't apply to because if it's optional the DPs will do what is convenient for them at the expense of future users who will be confused and that's what we want to try to avoid. I happen to think that even for meta-kernels if old versions are available with a few extra clicks to our expert users who want them, and not there to confuse novice users, my opinion is that would be best practice. NAIF clearly disagrees and if the consensus is that we want to allow an exception then I also think that NAIF's arguments are similar to what people were saying about certain approaches to calibration products. If we want to make an exception for all calibration products and SPICE products I would not be vehemently opposed to that if we want to say there are various ways we want to accommodate other use cases and still get the control that we need for the primary purpose of this change.

One thing that came up at the F2F discussion last week was how EN treats these things and basically, when people are accessing data through EN it treats multiple versions of the same product as an error and only delivers the latest version. So if somebody downloads a full collection EN will only deliver the latest. Believe it will only offer the latest if it's providing a list of what's in the specific collection.

Question: Right. That's important and I want to expand on it a little bit. NAIF said they understand that when people are using the PDS4 registry to look at a SPICE archive - and if they go for a specific meta-kernel, they are only going to see the latest version of that meta-kernel. The tool might have a switch that they can turn on to see all versions. I expect that will be the case, but that will take some extra doing and you want to trust our expert users to do whatever we tell them they need to do to get what they want. But why not, if you think all versions of a meta-kernel are of interest to everybody, why not give them different LIDs and that way everybody will see them?
Answer: Because giving them a different VID serves the purpose, so you can always look through a list of files for the registry like you discussed and it will give you the latest at a glance, or output from whatever API, so we don't have to give different LIDs for that. In fact, this idea of LID and VID giving you the version works perfectly, so unclear why we would have to change it. Not sure why we would break the process that works so well.

Question: So NAIF agrees that novice users should only ever need to see the latest version of a meta-kernel?

Answer: Sure. That's what the WebGeocalc tool does. It always uses the latest version to compute. But one could go and load manually the previous versions if they want to repeat calculations.
~ I think we agree and the question is how to implement this. We can carve out an exception for NAIF's preferred way of implementing it.
~ Again, NAIF's biggest philosophical concern is that this particular change will prohibit listing everything that has ever been added to an archive in the collection of the archive. That's what it is doing, so if I have an archive of a million things and two of them are obsolete, I cannot list them anymore, which to me makes the listing for this archive incomplete and that's just wrong, because if you have an inventory of anything that has been archived, whether it's your old mail or passports or anything and you want everything there no matter if superseded, you want everything, this approach will prohibit that forever.
~ The old versions of a collection are still available so people who want to dig for it can find it.
Another Question: Why do they have to dig for it? If the latest version of the collection could contain everything, why?
Answer: Could also equally say why does one have to swim through the morass of a lot of obsolete and un-useful products to try to reach what they want. We definitely have heard the feedback from users too, so I understand where you are coming from, I 100% do, and I agree on some level this is philosophical, but there's a question about whether as a standard we should adopt a "universal" (with carved out exceptions) standard as a standard. We should probably be consistent across nodes because we want our users to have a consistent experience and reasonable expectations that are well documented and explicit, so I think that's the underlying motivation here. Consistency and the dominant philosophy here right now is that we shouldn't subject users to having to weed through and have extra knowledge about how to identify obsolete products and then weed through them to find the right product because that will and has created confusion in genuine, concrete use cases.
~ It was already stated that the API only returns the latest products. Maybe the API just needs a flag which will say return all products and the default will be to return the latest. Nobody will be confused. Everybody will get the latest LIDs and they will happily chug along.
~ A switch can be available and an expert user can use it.

Question: Well, but if I want it to be available with a switch, not with a whole bunch of clicks that I don't need to do, and I want my collection to enable it by listing everything in it, how is this not a viable approach?
Answer: I feel like I just answered that. Viable in the sense that it could be done that way I suppose, but it's not preferable and as a standard we probably should be consistent and that was the consensus that was reached at the F2F. I'm not saying I don't understand where you are coming from, I just...(Interrupted)
~ I apologise I wasn't at the F2F, it was scheduled after I scheduled my vacation, so it's too bad that people need to vacation once in a while, but what I am trying to say is that if we take my approach, then someone who doesn't want to use the registry and wants to just download that inventory file can do it with a simple, unique command, with one option. What the registry is going to do or they can see everything and there is a physical record of it. This registry is going to be there or not. As I see with cloud or various flight projects where clouds disappear and then maybe come back, and then something else happens. But when I have my physical file that has everything, that's something tangible. Maybe we are going away from this. Maybe they won't exist in the wild, but I want my file to have every single product in my archive to be listed. That's really important. Important to many things - for providers and users alike. And this will prohibit this.

Someone thinks the DDWG understands the objection here. For this to go forward we need to figure out how to handle this exception. And, frankly, there are other nodes that have used this approach of having multiple versions of a single product in an inventory. Their use is slightly different where each of the versions can be valid and not necessarily superseded. Valid over different times or in different situations.
~ One more thing - in SPICE archives NAIF is applying exactly the same logic to our document collection. We have a single document that occasionally changes and gets a new version. The LIDVID of that document is versioned and NAIF will want those versions to be available as well because sometimes changes between simple HTML files involve removing things and having the traceability and having this ability to compare these versions is very important, at least to NAIF, the archive and possibly some users, so if this becomes an exception, NAIF would like this exception for documents too.
~ The concerns are understood. A lot of this is philosophical. Don't think anyone disagrees that superseded versions of products need to continue to be available. It's just how we do that.
~ The NAIF approach makes them available with ease, so that should be taken into account.
~ And a lot of us think that's a problem because then it creates confusion, but anyway, we've been through this.

The NAIF method works in part because NAIF has a small minority of their data products that they version at all. That works for NAIF, and that's fine, and we can allow that.
~ Fine. Thanks.
~ We will have to think about how we can carve out exceptions for cases where... We'll have to consider this further.
~ Maybe an obvious accommodation would be product SPICE and maybe also product calibration should maybe be exempt from this requirement.
~ NAIF also wants this for product document.

Someone wouldn't support this for document. Believe that GEO wouldn't. Just wanted to be up front about that.
~ Well, that will break SPICE archives. All of them. This is exactly why NAIF is not doing any transitions because non-backwards compatible changes like this are just an enormous impact on the archives that are functional, in place, accepted by the community and so on. Still cannot believe this "let's put everything on the chopping block" fashion now fourteen years after PDS4 was started. That blows my mind.
~ To some extent, some of us have been surprised that this wasn't already standard because this is the plain faced meaning of what a version means - that old versions are available, but not what is primarily returned, so to some extent there's a different perspective.
~ NAIF hears that. Wants to thank everyone for listening and understanding.
(Boris had to leave)

Someone wants to point out that there is no product calibration. It gets called lots of things, but that's not a category in terms of a carve out.
~ Well, alright. Then it won't be that easy.
In the collection file, you can indicate it's a calibration collection or data collection or whatever, so possibly can trigger off that.
~ Another person says they hesitate to make a carve out unless there is a very specific request for it, and it seemed like there' wasn't push back at the F2F. Understand that Boris wasn't there, but aside from SPICE, don't know that there are lingering concerns about calibration that weren't

addressed by just using LIDs instead of LIDVIDs, but I could be wrong. Would want to wait until there was such a request until we consider the details for a carve out.

Unfortunately Boris is no longer here to yell at me for saying this, but I thought last week that the Provenance LDD was the solution, the mechanism for these carve outs. If for any reason the same LID with a different VID that does not supersede the previous version, then we can say so using the Provenance LDD. Don't think we will need all that many extra lines in the label. That seemed like a viable solution.
~ Someone agrees, but that doesn't necessarily effect NAIF's concerns. Whether multiple LIDVIDS are allowed for some LIDs in the collection is not necessarily the same question as what versioning represents and when exceptions to that versioning philosophy have to be documented. A separate question, tried to bring it up in a GitHub comment before this meeting started, but it's worth considering too.
~ Right and a person is okay with having a complete carve out for SPICE, but not for product document, which NAIF was also asking for. It might not be too onerous if NAIF wants to have multiple product documents with the same LID in their collections, to ask them to do Provenance LDD to document the reason for it, rather than making an exception for product document, as requested.
~ Another person thinks it was said that NAIF will not consider using a carve out. No way to compel someone to use an LDD.
~ Well, NAIF never uses the current version of the IM anyways, so not sure it makes any difference.

Someone from NAIF said - Oh, we do use the current version sometimes. I guess I also don't really understand why we have VIDs if we can't have multiple versions, and I think it's really valid to have multiple versions in a collection where you need to potentially reference previous versions. In our inventory files NAIF doesn't have all the LIDVIDs as the primary member in the inventory. Only have the latest as primary in the current version.

Question: So all are secondary?
Answer: Yes. In version 1 everything is primary. In version 2 everything that was in version 1 is now secondary and then everything new for version 2 is primary.
~ That might actually solve the problem.
~ Yeah, it might.
Another Question: So, suggesting that secondary members have a bundle that could be same LID as the previous version, compared to primary members of the same bundle or collection?
Answer: Yes.
~ Someone needs to think about that.
~ Another person too. Interesting approach to the problem because it would be primary to a previous version, but secondary to the later versions, which might be a solution.

Someone agrees with NAIF's comments. Having used SPICE archives, the way they are doing it makes sense. As they make a new version of a kernel, that becomes the primary in a newer collection and earlier versions become secondary. Like the way they are doing that and think it makes sense. I know my two cents don't mean much, but I would like to see SPICE allowed to go forward doing this and not be restricted and need to change at this point. It was well thought out, how they are doing this, and they handled secondary products as originally intended.
~ Thank you for sharing that.

Question: Anyone else want to comment?
Answer:(Silence)

We definitely have some things to think about and need to read through the comments that have come up in the last 24 hours and see if we can figure out a way to move this forward.

Question: Someone seemed to be suggesting that using the Provenance LDD to explain an exception would also allow multiple LIDVIDs for the same LID to be in the same collection. Is that what it is being suggested because that's definitely what NAIF wanted to do with documents, as I understood?
Answer: Yes. We haven't really written this out, but what I kind of have in mind after last week's discussion is that one exception to this rule would be if the label contains an entry from a Provenance LDD that explains why the same LID with a different VID doesn't actually supersede the previous version. That seems like something we could incorporate into the system.
~ That's interesting. I would like to hear EN's opinion on how difficult that would be. The first thought by me is that would be difficult.
~ Thought it's what we were talking about last week.
~ Right, and maybe I'm mistaken, maybe we didn't all come to the same conclusion. My conclusion was that I thought we were talking about two separate things last week and maybe that's why there is confusion. My understanding was that we were talking about what a VID means, what it means when we increment a VID and also whether we should allow multiple VIDs for the same LID in a collection. And that those were conceptually related, but might not be technically related in any way. You're suggesting that they are technically equivalent. That's what I did not gather from the F2F.
~ Think we were primarily talking about them as separate issues, but we did cross the streams at one or two points.
~ My understanding was that the Provenance LDD would be useful to explain this is why this VID doesn't actually mean superseded in some conventional sense, not that that would have anything to do with excluding it from the application of this rule. That was my understanding.

There's a lot to sort through here. We will look at this again in the coming weeks.

Location of Context Products

Someone has a story to share, that will probably turn into a bit of a rant. They believe this affects everybody here so think people need to hear it. Node was working on a peer review for a recent submission and realized that we missed adding an instrument context product, so we sent the context product off to Richard Chen to put in the system and we weren't sure if he had done it. We were using the same version of validate that we had used before - 3.5.1, and we used the validate -U option to reload the context products, and that didn't work, so we asked Jordan why this didn't work and he said that it's been refactored and you should download a new version - 3.6.3. Then a day later 3.7.0 came out, so we got both of those. So, the reason I bring that up is that whatever change this was broke backwards compatibility. So I would have thought it should have been version 4, since that's what I understood semantic versioning to mean. But anyway, moving on. We ran 3.7.0 and the context product in question was still not loaded, but more deeply concerning than that, two other context products were now missing from the context system. Disappeared. So now we're missing three context products. Thought that maybe it wasn't updated, so I ran the -U again. Same thing. I went back to 3.6.3, which hadn't actually run before, and the two missing context

products were actually there. So, something has occurred where context products have been removed from the system and I am absolutely scared for what has gone on, especially considering what was going on last week when we were being asked how safe our systems are. Context products should not be lost. Further, it seems as though the bundle location that we have been using for a long, long time is no longer being updated and context products are apparently being archived somewhere else, where I don't know where they are. I can't seem to find them and Richard doesn't know either. He has not given us this information and it sounds like control of the context bundle has passed over from Richard to somebody else. This is very scary to me. If EN is moving the context bundle they need to tell us. They need to tell us where it is and it needs to absolutely not be residing in GitHub, if that's where it is, because GitHub is not a data repository. I will stop the rant there.

Thank you for pointing this out. Did not realize this might be going on. Say this because the NSSDCA front end process that accepts PDS4 bundles for deep archiving actually keys off of mission, inst host, and instrument context LIDs in bundle and collection products to associate, or link, those deep archives to the NSSDCA master catalog.
~ That's what context products are for.
~ Yeah, so if the location of those is no longer that public website that we have been using for years, then some of us might want to know where the new one is.
~ Agreement.
~ Actually went to look for something recently that I thought Richard had said was registered, but couldn't find it in the public repository at EN. Didn't question it. Glad you did. Would like to hear more about this.
~ We should know what's going on.

Question: Someone has a question that overlaps with that. You mentioned whether different versions of validate are able to recognize the existence of certain context products, but you didn't say whether or not they actually exist in the browsable archive of context products?
Answer: Yes. They are there in the old system that is no longer being updated.
Another Question: Including the one that was missing in the first place, that you had to add?
Answer: No. That one is not there.
Another Question: Not there, but validate can find it now?
Answer: No. Validate still can't find it. Unclear if Richard has put it in or not. Trying to understand what is going on.
~ Okay. Thanks for clarifying.

Someone else said that they recently corresponded with Richard about adding some context objects and he mentioned during that that they wouldn't be updated on EN's website, but they did end up in validate. So there is some conscious decision there, but I don't know what the rational is.
~ Yeah, that's essentially my understanding too and I'm not okay with this.
~ Agreement.
~ I don't want to just search for LIDs, I want want to see the product. Just like I would want to see a data product at a node. If I'm looking for a data product, I want to find it.
~ Agreement.
~ Another person agrees. They also make data bundles as NSSDCA restores or repackages NSSDCA holdings into PDS and often go to the repository that EN has for context products to get an idea of what to do. Realizes that context products are changing, but it is a great template to at least

get an idea of what needs to be done. Anyway, would like to see the entire repository somewhere available at least to PDS or people who make archives.

I guess the next person who needs to say something is Ron.

Question from Ron: Why would you think I would say anything?
Answer: Because you are the only EN person on this call and so if all you have to say is that you're going to convey our concerns to Jordan and Richard, fine.
Action Item - Ron will convey the concerns to Jordan.
Another Question to Ron: Do you actually know where the context products are being stored?
Answer: No. That really falls to Richard.
~ Yeah, but I think, what I'm gleaning from Richard is that he doesn't have control over it anymore. It sounds like the EDWG has control over it and I am not okay with that.

Another person has tried to get things updated and the process that they thought happened is that the updates are on GitHub but if you want to get them released into what they thought was the active version available online, you would have to submit a ticket to get it, can't remember where you submit that, to get it released at some point. Sounds like maybe there's a disconnect between validate and at least what's online and maybe getting things online, but it's not as simple as updating what's in GitHub. In terms of getting things released.
~ GitHub is not supposed to be a data repository.

As an EDWG member, also not sure what's going on here. When you say the EDWG, you really mean the EN web designers.
~ Yes.
~ Just wanted to be clear about that.

We will convey the concern to Jordan and try to get some clarification on what's going on.
~ Yeah. Maybe Joe can summarize the DDWG's concerns. **(Action Item - Joe)**

We have ten minutes left.

CCB#59 - funding year only allows the value "9999"

There's been a lot of additional discussion on this. Concerned about where we are now. Anne has raised some concerns about the approach that the DDWG has decided upon - thinking that it would be the most acceptable approach for the CCB. Now, that said, think that the DDWG needs to do what we think is the best approach and not what we think the CCB is going to agree to necessarily. Wonders if there are any additional thoughts on the approach that we discussed last time, which is basically updated in the SCR, to use ASCII Date YMD as the type with no restrictions. It would be able to not just be a year, it could be a full date and it could include the Z.

Question: Any thoughts on that?
Answer: Yes. A thought. As one of the people who was a fan of the ASCII Date YMD, just because the publication date used it before, it seemed to make sense and part of it was the politics. I wanted to make sure those went through. But now having heard Anne's clarification, which, I guess, I misunderstood originally, I guess I'm now inclined to like the idea of just going with an

ASCII non-negative integer for that instead. If I'm in the minority that's fine, but I've changed my pre-vote from the last meeting.
~ Someone else thinks they had misunderstood what her proposal was as well. What her proposal basically is is instead of creating a new type to just use the existing integer type.
~ Specifically, we would probably use what she called a base type, which confused people, but I think the equivalent would be ASCII non-negative integer. The same thing that is used for Mars Year. That's what confused me when she said base type. I thought she wanted us to find a new data type, but she really just meant to use an already existing data type that makes sense and already has that base.
~ Agreement.
~ Realize this is similar to what has been suggested originally. I think that was to use short string.
~ Yes. We weren't using an integer type, but we were restricting to integers using a pattern. I think this is a better way of doing that. It comes out to the same thing. Makes sense.

Question: Any additional thoughts?
Answer: I spoke pretty strongly against using a free form string with ad hock formatting to make it look like a year. I have a lot less concern about just using an integer, and rather than restricting it to four characters, since it's a integer, we could define a maximum value, like 2100 or something.
~ Someone else was going to suggest the same thing. Minimum of 1950 and a maximum of 2100, and when it gets to 2100 I won't be around to worry about it.
~ (Laughter)
~ Another person wouldn't object to that. Just wasn't sure, always concerned about putting us in a corner, besides there is no where else we have made that sort of restriction on dates, including the publication year, so it just seems reasonable. But wouldn't object if somebody wants to put more restrictions on it.

There are already simple types defined for all of the ASCII dates. They all use string, followed by one or more patterns and can understand why someone might want to think about using ASCII non-negative integer, but really, all the dates are all strings.
~ That's only because those dates have to accommodate non-digits, hyphens and what not. That's not really relevant. If you look at Mars Year, I think it does use ASCII Non-negative integer. Kind of the most analogous in this sense, other than publication year, which is unconstrained.
~ We might need to reconsider how publication year is handled after this, although, I think we need a different range. As pointed out, we don't need to track funding back before the formation of NASA. Not sure how far back we would have to go, but I think the formation of NASA is probably a decent starting date and 2100 is plenty far into the future

Question: Any more thoughts on this?
Answer: Feeling more comfortable. Understand that fundamentally every character is a character, but once we have defined a type I feel more comfortable if the limitations are consistent with that type, so if we restrict this to four digits we are basically saying it has to be under ten thousand and since we're defining it as an integer I would rather use an integer as the limitation. I'm also curious about Mars Year. I don't see that attribute in the specification. I would expect maybe that's in an LDD somewhere that's being referred to.
~ Yes. Actually, it's Start Mars Year and Stop Mars Year in the observation area in the Mars 2020 parameters, so never mind. I just looked up PDS4 and didn't look more closely. But they use an integer and that seems like a reasonable place to start.
~ Someone else also likes consistency and would do integer with no limits.

~ Have a feeling Anne wouldn't like that either. She said something about four digits. Think I looked at Datacite because she wanted to be compatible with them, which is not the only concern, but think they require four digits. I'll double check that. **(Action Item - Ethan)**

Someone proposes between 1950 and 2100 or 1900 and 2100.

Question: Any other comments?
Answer: (Silence)

If everyone feels ready, would like to do a straw-poll to vote on the approach before we make any more changes. The approach would be to change the type to ASCII non-negative and then set the min/max values to the year NASA was founded, 1950 something and 2100.

Question: Why not higher than 2100?
Answer: Because it will be someone else's problem long before 2100. For one, I would guess that by 2100 we won't be in PDS4 anymore.
~ We could make it 2200. 1900 and 2200. And also if we are placing these limits then there's no point to using non-negative integer instead of just using integer, which is simpler. I would just go with integer.
~ No objection about that. Just don't want us to get bogged down on bounds. I am not objecting. I will support any reasonable bounds.
Another Question: 1900 to 2200, any objections?
Answer: (Silence)

The approach will be to change the type to ASCII integer with minimum of 1900 and maximum of 2200. We'll just hold a straw-poll on updating the SCR to reflect that requested change.

The Straw-Poll:
ATMOS - Yes
EN - Yes
GEO - Yes
IMG- Yes
IPDA -Not Here
NAIF - Yes
PPI - Yes
RMS - Yes
RS - Yes
SBN- Yes

All yes votes, so we will go ahead and update the SCR to reflect this approach. Hopefully we will vote on it next week.

Someone asked Joe to scroll down on the screen showing the SCR - and to do a refresh. Believes that is what everyone is asking for.
~ No.
~ Changed this to ASCII integer.
~ It would be just ASCII integer and instead of patterns and minimum and maximum length it would be, not sure of the keywords, but a minimum and maximum value.

~ Min inclusive and max inclusive.
~ It was updated in real time.

Questions: Any objections?
Answer: (Silence)

_____
Question: Any last comments?
Answer:(Silence)
~ Hearing nothing, we will meet again in two weeks.

title: DDWG Notes 2025-05-01
layout: default
date: 2025-05-01

May 1, 2025
Notes transcribed by Debra Kazden

Known Attendees: A. Bailey, J. Blalock, B. Deen, B. Hirsch, L. Huber, S. Hughes, R. Joyner, A. Knight, P. Lawton, J. Mafi, S. McLaughlin, E. Schaefer, B. Semenov, M. Seritan and M. Tiscareno (*'s indicate people who voted during this meeting - if applicable)

DDWG Agenda

(Included in email sent April 29, from J.Mafi, Agenda for 2025-05-01)

We will have a DDWG telecon this Thursday, May 3rd at 9:30 PDT (12:30 EDT, 10:30 MDT). We will cover the following issues:

1) Action item reports.
2) CCB#59 funding year only allows the value "9999" – This SCR has been updated to reflect the changes approved by the DDWG during the last telecon. Please be prepared to discuss and vote on approval.
3) CCB#46 Disallow multiple versions of the same LID from occurring in a single collection – Please review the comments from the last two weeks on this SCR and be prepared to discuss. We will work on identifying a way to move forward on this issue.
4) CCB#58 New Product Native values for MSL/Mastcam – A document describing the proposed new product native format has been added to the SCR, along with sample Python code for reading data in this format.

Front-matter/What's new

1) Submitted to the CCB: (none)
2) New issues in GitHub:
   CCB#63 - International characters / Unicode support for organization name and display full name
3) Ready for vote: (none)
4) Build 15.1 freeze date: 2/27/2025. The following SCRs are queued for implementation: CCB#13, CCB#48, CCB#49, CCB#54, CCB#55.
5) Withdraw by author:
   (Brief Discussion)

Action Items

a) CCB#39 – S. Hughes to update SCR
b) CCB#12 – S. Hughes to verify whether this is subsumed by CCB#52 and CCB#53
c) CCB#20 & CCB#32 – Tiger team formed and working the issue
d) CCB#31 – Ron and Matt working the issue **(Not Discussed)**

Tier 1 issues

- CCB#62- Transfer Manifest class should enforce data types for its two mandatory columns
- CCB#61 - Disconnect between SR and Validation of ASCII Real

- CCB#60 - New permissible value for File Area Observational.Header.parsing standard id: "CDF 3.9 ISTP/IACG"
- CCB#59 - funding year only allows the value "9999"
- CCB#58 - New Product Native values for MSL/Mastcam
- CCB#57 - Allow Encoded Native in File Area Observational Supplemental
- CCB#56 – Reconsider the PDS4 build cycle cadence
- CCB#53 – Improve cross-namespace validations in sch files
- CCB#52 – Enable cross-namespace validations in sch files
- CCB#51 – Add an optional attribute in the Target Identification class to identify Primary/Secondary targets
- CCB#50 – Clarify Ambiguous Author Information in Product Document
- CCB#46 – Disallow multiple versions of the same LID from occurring in a single collection
- CCB#44 – Issues caused by the failure to expose pds: attributes that are referenced in an LDD.
- CCB#41 – Create a way for classes or attributes to be exposed to other LDDs but not to data providers
- CCB#39 – Fixed problem where LDDTool does not include all rules from IngestLDD
- CCB#37 – Add List Author, List Editor, and List Contributor to the Modification Detail class
- CCB#36 – Add Mission Phase to the Investigation class
- CCB#35 – Add Alias Index and Acronym Index to the Investigation class
- CCB#32 – Add Host type to Observing System Component
- CCB#31 – Ensure that dictionary versions are specified in labels
- CCB#28 – As a user, I want validate to allow delimited tables with whitespace-only numeric fields
- CCB#23 (JIRA-CCB-211) – XML/schema based files as archival data
- CCB#22 (JIRA-CCB-164) – Display Settings not required for images
- 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#21 (JIRA-CCB-367) – Provenance LDD Using PROV-DM
- CCB#20 (JIRA-CCB-358) – Instrument Package Context Products
- Tiger Team: M. Tiscareno (lead), B. Deen, B. Hirsch, S. Hughes, R. Joyner, J. Mafi, E. Schaefer.
- CCB#12 – Document Nuances for Reusability of Generic LDD Classes/Attributes
- CCB#5 – Move to V.2.0.0 for June 2025 Release
- CCB#4 – Backwards Incompatibility policies
- -Presentation: Backwards Compatibility
- CCB#3 – Adopt Semantic Versioning with version 2 transition
- The CCB returned this, requesting that it be resubmitted as part specific proposal to transition to IM V2.0.0.
  (Some Discussions)

Tier 2 issues

1) 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!

1) JIRA-CCB-326 (link) - Inconsistent Direction on How to Use Discipline Facets
2) RS: 259/281 (see new attachment to JIRA-CCB-281) (on hold)
3) JIRA-CCB-364 (link) – Disallow bundles referencing collections by LID **(Not Discussed)**

Historical links:

- Export of history: https://github.com/NASA-PDS/PDS4-CCB/tree/main/data
- Backup: https://nasa-pds.atlassian.net/jira/software/c/projects/CCB/issues (limited seats)

Next meeting: 15 May 2025

DDWG

Front Matter

Question: Any action item reports?
Answer: (Silence)

There is a new SCR this week:
CCB#63 - International characters / Unicode support for organization name and display full name

There are a few additional attributes that probably need the same kind of support. Believe this was a kind of bug in the way these were implemented, but we will discuss this in future weeks. Think it's a fairly straight forward issue, although there have been others like that. Hopefully, we'll discuss and be able to resolve this fairly quickly in future weeks.

Question: Someone is curious whether the people who raised this issue agree. Thought that the response in GitHub was really good. We just need to change ASCII short string collapsed to UTF-8 short string collapsed, or whatever, and that would solve all the problems. Are there concerns about that or is that something we can discuss going forward?
Answer: That is the fix, I think, yeah, it should be straight forward.
~ Cool.
~ Think that was just an oversight. Think that should have been the intended implementation all along because there's been discussion right from the beginning that we need to be able to support foreign names and things like that.
~ Good job for articulating that.
~ Actually, someone thinks we need to bring the tool people in on this one because they have run into having to change UTF-8 to non-UTF-8 when they use some of the tools. They get all unhappy about it.
~ Okay. That's a good point and we should at least let them know well ahead about the potential impacts or potential change and let them access the impacts.

Question: How will it affect users who don't understand Chinese, Japanese, Korean, Arabian? When I see a name that's made up of foreign characters, what am I to do? Is there going to be an

equivalent name provided in English? A secondary for search for those who don't speak other languages? Is there already a place for that?
Answer: That's a good point. It's almost like you want to have something added, could be for any field that we have now, something that says the more extended name and allow that, like if a Japanese person wanted to have their name in Japanese that should be a different field then the full name field that would be used normally. Think that's the way we should be going. Want to make an additional comment too, a point that is coming up and we all need to come to grips with, I don't think there's any such thing as a simple change anymore. As EN mentioned, we need to bring the system people in on that. It might even be bigger than that. What I am seeing now, especially with the requests from CCB to do complete testing, is that we almost need to implement this in a test sandbox or something before we implement any particular change. Need to test any individual change in a test sandbox before before there's a TA done. It's getting that complicated in terms of making sure we aren't impacting something down the line. So, this is something that I'm just throwing out, that we should start thinking about.

It's worth noting that we do have for person a display full name that could potentially be used, as we have discussed, but I don't see anything equivalent for organization, and since display full name could mean a variety of things, we could consider adding something like ASCII full name or something like that. Not a deep dive, but a first thought.
~ The issue here is with accented Latin characters, like for Jose (pretend there's an accent typed) for example. There are other fields out there that are UTF-8 and any of them could be a Chinese character or something, but unclear if that means we need a mostly Latin accent allowed sort of type. Don't know.
~ Yeah, we have to think about how and where we draw the line as part of the discussion.
~ There seems to be a difference between a name in the native characters in something translated into something that is normally readable on the web, in an English sense, so I do think we're talking about another field of some kind.

Someone would like to point out that they believe the test sandbox is part of the TA, not in advance of the TA. That is part of the tech assessment.
~ Another person agrees. It has come to that.

Alright. Thanks for those comments. Sounds like we've got more work ahead then anticipated, but it's good to have this input up front and to be thinking about these issues.

Question: Any additional comments?
Answer: The issue we are talking about, being Chinese or Latin or whatever, is nothing new for these fields or the existing fields that use UTF-8, so it's almost a separate question. We already allow that in other fields, so these fields that are supposed to be people's names, to do them properly is a small change, but then there's a concern that any of these fields that are UTF-8 can be filled with emojis or something. They are part of UTF-8, and so maybe there is something we should do to prevent that.
Another Question: And do we need a requirement for a presumably English equivalent of anything that is non-English?
Answer: What I was trying to say is that I think it's a separate issue. We support UTF-8 names. These are additional instances of that, so I think it's that simple, but a separate issue is that it could be misused and we need to decide how we deal with that.
~ That's fair.

~ Someone didn't realize that UTF-8 goes beyond the kind of extended Latin alphabet, so obviously, maybe ASCII is to restrictive and UTF is perhaps too expansive, but we could maybe implement it along with a note in the SR saying that every character that's used should be readily understandable to an English speaker, or something like that. It would be hard to implement, but if it's in the SR curators should notice such things.
~ It's unclear how you would check besides doing a spell check.
~ Soon we will have AI enforcement of the standards that can notice things like this.

It also comes down to search because not all people will be able to type international characters to search for a particular organization, so unless there is a parallel name in English, and we can say to search for it, we will never find it.
~ That's absolutely something. Was going to bring up the same concern. Google works that way, if you search for "J" "o" "s" "e" you will get Jose with the accent on the e, so it's possible, but I think if we do that I suspect we would have to implement it in the tool itself. This discussion is becoming larger than I anticipated. No easy change doesn't come back to haunt us. Maybe we should turn back to the schedule for now.

There have been some good ideas and they will be included in the notes. Prefer to circle back around to this when we have all had more time to think about it. Appreciate the comments. Think these are definitely some considerations that we want to take into account when we go into the discussion. I do like the suggestion of separating the issue of how do we manage things that are unreadable to English speakers and how we provide an equivalent or make sure there's an equivalent somehow, because right now, not sure there's anywhere in the IM to capture an equivalent of the non-English version. Kind of think of that like archiving the raw data - it can be a very important resource but we don't have prefect translations.
~ Maybe the current attribute should remain ASCII and a new one for if people really want to spell their thing correctly and do it in UTF-8.
~ Something like organization name native.
~ Eh.

Someone agrees with a lot of what has been said, but what we seem be pushing towards is a policy whether or not the PDS, in general, is in English. For example, we write descriptions in English and maybe that needs to be addressed by the MC. A policy to say anything operational has to be English because that's how the system is designed, but to allow for name native fields where necessary for someone who wants to add an emoji to describe something.
~ If we differentiate between things that are essentially Latin, but have accents and other similar markings as opposed to different alphabets.

Thanks to everyone for this discussion.

Tier 1 Issues

CCB#59 - funding year only allows the value "9999"

Think we all felt like there was good agreement on this and there have been updates made to implement the changes, basically to change the data type to ASCII integer and set a fixed range on it. Then a comment was posted yesterday questioning that approach.
~ Someone apologized and said they missed the last meeting and didn't realize how far we got on this issue. There's been a lot going on to try to get the build out, but it always bothered me, the concept of ASCII integer. It seems like we are kind of coming in the back door. It just doesn't seem consistent with the object orientated modeling approach that we have taken. Tried to describe that in the comment. Was trying to say that it seems that the four character year, based on what I read and the requirements for this field and the two use cases, is that it's just simply the first four characters of our standard YMD date and if that's not the case then we should just model it as such. Suggesting a new data type for two reasons. One is that it could be used in at least two use cases now and could be used in other cases. It makes sense as long as we're not talking about something that actually is converted to an integer and used in arithmetic that we should just limit it to what was originally designed. The PDS adopted the ISO-8601 and made modifications, and basically, that is what we adopted. The question is why we don't continue with the model as it is and simply define the component we're talking about as a new data type. It's an opinion. Open for questions.

Question: Any thoughts on this approach? Anne had made an argument against full date or anything more than a year in terms of compatibility with what DataCite requires.
Another Question: Someone missed that. What did she say about DataCite?
Answer: Another person believes it was DataCite.
~ Don't mind taking that into consideration. If DataCite is someone we've got to interface with, but as the comment was put in, in general, if we look at Dublin Core or any other of these standards, it's generally just the four character string.
~ Her comment in GitHub says DataCite has to deal with complaints about inconsistent metadata. The main problem there would be if our metadata includes full dates as we had discussed obviously. What they are looking for is just a year, so then that creates problems. If we determine a way of restricting date to just a year, then, think that resolves that issue. One of the issues that was talked about was compatibility with other data types, but unclear if there are any situations in which we need to compare a funding year to another date.
~ Response is that on the system side if you were using funding date in some type of report and you wanted to order it then you would just say okay, I'll just order it on the four character string and that's kind of dealt with in an interoperability section in the JSON LDD and schema.
~ Can see a lot of the conversation in GitHub. Anne did not actually call in, so we were considering what was suggested and Anne made a strong argument against that. It seems, to some extent, that the fundamental issue is what the point of a data type is. Have to decide if it is more to categorize everything according to an ontology or if it's more to just make sure that we have the validation being done in a convenient way. Anne is pointing out that a year really is nothing more than an integer, so having ASCII date year is redundant and unnecessarily complicated because we could just call it an integer. Someone specifically asked if it is important for the data type to point out this isn't just any integer, this is a year, and if that makes it special somehow. Anne's point was no, it's just an integer and the only reason we have things like ASCII date YMD is as a shorthand for

consistency and ease of way to have more complicated dates. When we have them we have an identified way to do that, but if all we are doing is year, Anne's argument is that it should just be an integer. That seems reasonable to me and nobody has argued against it, but maybe that is what is being done now.
~ When we first started working with date time we decided to adopt ISO 8601, which is much broader than we ever needed, so we constrained it down to what we have now. So, 8601 basically is how we view date time. Think there were some questions about why we chose it, and somebody said it was a bug that we chose early on to say ASCII string in this particular case - funding year as ASCII string and constrained down to nine numeric digits. I don't see the difference between ASCII string and nine numeric digits or ASCII integer and nine numeric digits. We had ASCII string, limited to nine digits. That should have been sufficient. I really don't see a significant difference in moving to ASCII integer with four digits. If we want to get into the more philosophical perspective of saying okay why would this be ASCII string with nine digits - we could create a date type. Think that's a different discussion and we don't necessarily need to go there. Guess I'm coming back to the position of wanting to know what was wrong with what we had other than that I might have missed someone saying it didn't work. If so, that's a bug and we need to make it work. Not sure if this all made sense. Unsure what's wrong with making it ASCII string with four digits and those four digits having to be a numeric digit. Unclear what was wrong with that or how this is different from ASCII integer.

I think if you had read Anne's comment, and I hate to speak for her, so interpreting the best I can, I think she comes from the perspective like you did, that it really depends on what you are comparing it to. So, if you are comparing a year to a duration, for example, and we had another field that was duration, with maybe a float representation of a year in seconds, for example, then it would be useful to have them both start out as integer or float as appropriate numeric types. She even mentioned float as a possibility for a year in a hypothetical sense. It really does come down to what we expect to be comparing to. A numeric value to other numeric values or s string to other strings. I think there's an argument for both. This is a special case of date string. Date strings have to be strings because they have non-digit characters in them. On the other hand, one can see it as a special case of a unit full scalar that measures a particular year. I don't know that one is better than the other, but I am inclined to think it makes sense to follow the standards when in doubt here, so I would like to hear what the standards say. I know we have already heard a little bit, but I don't know if there are other standards to be considered here.
~ Think that's a key point and it's one we haven't gone into. Need to consider what the requirements for this field are. Unclear if there is a stated need for a scalar so you can compare. If that is the case, it changes the argument totally, but I hadn't heard that we needed to do that for this particular attribute.
~ It was something that Anne brought up in a hypothetical. Unclear how common it would be to do that.
~ That is someone's understanding. Seems more hypothetical, but think going back to what we have been doing - modeling from day one, it's object orientated. Need to think about what this is, what it should look like and how it will be used. As far as I can tell, it might be used for sorting or search, but other than that it seems primarily just for documentation for users. So, let's define it as an object. Either it's a string or we could create a new data type for it. Making it ASCII integer implies additional requirements which haven't been laid on the table. Basically just going with the flow of what we've always done. Unclear on why we are taking the jump onto the other road.

There is another argument that's totally different but important to bring up. Anne pointed out that the unit of year could change. Might not always be Earth year. We do have Mars year elements, for example, or we could be talking about civil years versus astronomical years or whatever. So that is a fair point that in other contexts, not in publication year, but in other cases, it could be that we are talking about different units for year and that would certainly lend itself to a scalar interpretation because we are applying a unit to a scalar rather than a string. Again, it becomes a question of whether this is a specific case of a date string or a specific case of a unit full scalar, and I don't know if there's a clear option other than resorting to the standards here. I would be curious about what Anne would say in light of the points that have been made by EN.
~ That's a good point and we should question what 8601 focused on. Maybe we should look back and see if 8601 was in Earth years and allows for the ability to express years in other systems. Good point, but if it's 8601 and that was designed for Earth years, I'm still asking why we are leaving that model.

Question: It is uncertain where this all leaves us. Know that the possibilities mentioned here were suggested previously and rejected for one reason or another. We could vote on it as it is, but I'm not sure… Is there any objection to voting on the current implementation? As opposed to walking these back and trying to reconsider them? Again, one of the options discussed previously was to create a new data type that would consist of just year. I don't recall the details, but that was rejected, I think, by Anne.
Answered with Another Question: Quick Question, did you vote last time or was it just the first consensus?
Answer: We voted multiple times on this. We voted last time to approve this approach of using ASCII Integer, and this week was just to approve the implementation.
~ Someone said okay, I don't mind a vote, but I want to be clear, I'm not insisting on a new data type. I would prefer using ASCII string and limiting it to four characters if we are not creating a new data type. To me, it's basically that we're raising an issue of semantic inconsistency by saying this is an integer, unless it's going to be used as an integer, and I don't hear that requirement.
~ The initial implementation was ASCII string restricted to four characters and some things like that, but it was the implementation that drew the most complaints.
Question: What are the complaints against the existing implementation?
Answer: Don't recall specifically, but that was the first implementation. The second discussed was a new data type. Then there was actually a third, I think, which was just to use ASCII date YMD. Then there were complaints about that being incompatible with DataCite, which just uses year for things like funding year. That's how we arrived at ASCII integer.
~ All for keeping it simple. Maybe that's why want to know what was wrong with the existing implementation because still not understanding that and why we don't just keep it the same for simplicity.
Another Question: What do you mean by the existing implementation? The existing implementation is restricted to the year being 9999. That's the problem.
Answer: That's a bug.
~ Yes.
~ That should have been fixed so that it is restricted to four numeric digits. That's what was meant by 9999.
~ That's where we got off track to begin with because if restricting it to four digits, it could be a three digit year with a z on the end, using one of the data types.
~ But ASCII integer precludes the z.
~ Yes, exactly, but the current implementation, can't remember if ASCII short string or date…

~ It's ASCII short string. Again, my point is maybe the implementation was wrong and maybe what's currently in there doesn't work, but the intent was to have ASCII string with 4 numeric digits and I think that's do-able. I think this is just a bug and so I am asking the question. It sounds like we have the same restrictions, but we are saying we should go to ASCII integer. We are complicating things when we could just fix what we have.
~ We are far past complicating things.
~ Don't mind a vote, was just trying to make my point clear.

For something that was functionally so simple to implement we have gone round and round and round with a number of different implementations. I personally would rather that we create a new data type because as it was pointed out earlier on, publication year also uses year and the current SCR doesn't address all the other attributes. It's simply funding year, so to me, we need to address the broader category and I prefer ASCII year. I won't hold up the vote. If this is going to push this thing through, bless everybody.
~ (Laughter)
~ Joe shared his screen to show the current implementation. It's ASCII short string collapsed, from 1 to 4 characters, but the pattern value of 9999 forces it to just be that value.

Question to EN: Do you know what's wrong with that? Why would it force 9999 when it should just require four numeric digits? If it's a bug it should be fixed.
Answer: It restricts the value to 9999.
~ It's not giving a character range. It's giving an exact value.
~ Then it would need a schematron rule to do that.
~ No.
~ It's done with a pattern.
~ Yes, and what we have done is remove the pattern and replaced it with an integer with minimum inclusive and maximum inclusive.
Another Question: If we have a string isn't there a way to use schematron to enforce a particular pattern of characters?
~ There is, but ...(interrupted)
~ If pattern isn't going to work, then let's create a schematron rule.
~ Can make the pattern work. The implementation that was rejected by Anne and the CCB had a perfectly good pattern in it.
~ So we just need to check the pattern.
~ The CCB rejected it. Philosophically based rejection. It was a perfect pattern. It worked for all the test cases. Life was good. It got rejected and then we went down these rabbit holes.

One of the issues that has been raised with funding year and with other year attributes is that we may want to restrict the date ranges. That would be really complicated to do using patterns or other things like that.
~ Someone is going to give their opinion and back off, but it just sounds like we are just not clear on the requirements for this field and until we are sure we are simply saying lets make it an integer and I'm still lost on the requirements and trying to figure out why we are doing that.

Question: Did we actually pass this? And officially send it to the CCB? And they rejected it or were on the way to passing it and then people chimed in?
Answer: We were on the way to passing it and then Anne pointed out some concerns and issues and that led to a bunch of additional discussion.

~ Think we should respect previous decisions. If this is just a regex error, that it was supposed to be four numerical digits and instead it accidentally restricted it to those four digits being 9999, the simple fix is to just make it four non-numerical digits and if people, for philosophical reasons don't like that it's four non-numerical digits, they should file a different SCR. This SCR is "I want to use a funding year that is not 9999 and please fix it." My initial comment was that this is clearly a bug fix that we shouldn't even have to get approval from the CCB because clearly the original intent was to have four non-numerical digits and it was just implemented incorrectly. It's not actually a change in policy, but somehow this has been derailed to become a big argument about changing the policy instead of just fixing the implementation of the original policy.
~ Yeah. The original implementation, I think, was ASCII string or short string collapsed. The one thing that caused complaints was that it allowed things like negative years and years with less than four digits, which for funding year, doesn't make any sense.
~ For all ASCII YMD we don't always restrict them to years that are within the space age.
~ Yeah.
~ That's an argument for a philosophical change, which is beyond the scope of the original request.
~ Fair point.
~ People agree.

Question: Any other thoughts?
Answer: Really do think the CCB is extending an overreach on rejecting things that are working functionally just because they have in mind an alternate implementation.
~ I don't remember at what state the initial complaints about the implementation of the pattern and use of ASCII String came from the DDWG. Don't think it was the CCB. Don't think they stepped in until later in the discussion. I frankly like this approach, just basically what we would do here is just fix the pattern, minimum length and maximum length because we are restricting it to a four digit value which could at least get us closer to the space age, which is really the only date ranges that make sense for funding year. Not sure if there are any complaints. Sounds like there's at least some support for keeping it ASCII short string collapsed and changing the pattern values and minimum and maximum length to be something reasonable. If no objections we can hold a straw poll for that approach. I apologise for doing this again, but we should make sure we have consensus because all of the implementation takes time and we don't want to have issues that are known at this point to come up after the work is done.

Someone is adding a GitHub comment in real-time. This is what they believe, based on all the history of edits made, pretty sure, this is what we supplied to the CCB that was rejected and we can all see it had a negative sign in the pattern and we included a maximum value of 5 to include the negative sign.
~ Someone else believes the implementation we would want is to get rid of the negative and make it 4 and the min and max could go away.
~ Or we could make them both 4.

Question: Does that slow validation at all? If we have a redundant specifier? Curious if in practice that helps or hurts.
Answer: Probably a few more cycles.
~ Get rid of the redundancy.
~(New GitHub pattern typed in real-time)

Question: What happens when the Egyptian space agency joins us and they want to specify funding year for Pharaohs to fund priests in B.C?
~ (Laughter)
Answer: Someone was thinking about that, during the golden age of Islam when…
Another Question: Are you telling me that the movie Stargate wasn't real?
~ (Laughter)

Question: Going back to my point, is there a policy in PDS where we are using English and we are using standard western dates? At some point we should have the MC set that up for us.
Answer: A person thinks we are spending too much time on a joke.
~ We just spent 40 minutes on this. Come on.
~ (Laughter)

Question: The proposed implementation is in GitHub. Any complaints? Please speak up, otherwise we will implement it and vote on it next time.
Answer: (Silence)

CCB#46 - Disallow multiple versions of the same LID from occurring in a single collection

The issue is that we are trying to find a way to move forward. Everyone acknowledges the issue this would cause NAIF if it started being reported as an error if there were multiple versions in a single collection inventory, so think we have stepped back form that approach, but the question is if there's a way for this to move forward. Think we want to decide that before we ask the MC to decide anything. Unsure if anyone has thoughts on a way to work this. Would like to see this reported at least as a warning, it doesn't have to be an error. For a PPI collection, if we've got two versions of identical LIDs in a collection than that would be an error for a PPI data set.

NAIF's main point is that this is the kind of change that should raise unbelievably high red flags as soon as it comes on the radar because we are not in 2012 anymore. Then it was a free for all and we could follow any principals we wanted. We have a very well functioning PDS4 system that includes nodes, providers and users, none of whom have used the registry before, but everybody has dealt with inventories and the current practice, obviously, has worked, so any non-backwards compatible change of this scale or magnitude should automatically generate a no response and only then when every corner of the system has been looked at should we consider that maybe on some ideological or philosophical reason we could change it. I think the process was broken here and I think the DDWG needs to make this consideration of what we have and how to keep it backwards compatible a much higher priority, especially for changes like this. Another thing that is different about this change, I may be mistaken, but I believe this is not going to be enforced by the IM. It's not going to be part of schema or schematron or anything like that. It's going to be enforced by tools, validate, etc., and I don't know how well validate obeys changes made in a particular IM version or particular release. The reality was that there was at least one change four years ago that invalidated some of the European SPICE archives because validate added a check that was expanded back to older versions. This is a situation that needs to be taken with the utter most care and any restrictions put on the system that functions well better not be put on it. It can be an option. I completely agree if you want an option for this, but as a mainstream, no. I don't think so.
~ It may be sort of expensive to implement in validate so I don't know. Think that's an issue that Jordan has mentioned before.

~ It's worth saying that this is actually an issue that's been bubbling around at GEO for a long while, a few years, so we were happy this was brought up. I wouldn't say that we've rushed into this. This was brought back in September of last year, so this has been a lengthy, if not continuous, conversation. I agree, we shouldn't rush into this, but not sure we have, but we shouldn't. We had a long discussion at the F2F. I know not everyone was there, but there was an attempt to hear from all the stakeholders there. I'm not trying to dismiss anything, I think we don't have to be too ashamed about how we've approached this. There's a lot of care that has been taken and we should continue to take care. It's too bad we didn't realize some of the NAIF concerns until late, but I am very happy they have been voiced now. I agree with what was said earlier that is does seem reasonable that we would carve out an exception for NAIF. I've been thinking about how that should be done at GEO. GEO is still mulling it over and not quite decided internally, but one suggestion that was made was that we would still keep this an error for primary products, but when there's a duplicate LID in a collection as long as it's secondary that would be acceptable. It can be only for primary LIDs. That was one thought. I'm open to that. Believe that's how NAIF works now, so that would perhaps not create an issue and be a solution and would help us to resolve this long standing issue that GEO had observed and had concerns about. I don't think this is simply "leave well enough alone" because GEO hasn't felt it for a long time. Think there does need to be some serious discussion about the best way forward and it needs to consider all of the node concerns, not just err on the side of doing nothing. A more balanced approach is probably preferable.

Question for Boris: Do you think that approach, and you can change your answer if you look at it more carefully later, would that approach just described work for NAIF?
Answer: Yes, but thinking about the system as a whole and what I would like to ask the nodes is to see how many inventories they have in their archives that they curate where this would not work. I know secondary versus primary as done by NAIF is an exception. I know the majority of nodes or at least some, have never labeled any products that came in an earlier version of a collection as secondary in the later version of the collection, so that's not a common practice. Maybe I'm wrong here. If this is not a consistent practice across everybody then the suggestion that was just made won't work. I am also thinking about cases where it's totally appropriate to have multiple versions like a document describing calibration algorithms. A previous version of this document is very important and the version before as well, and you probably want to include all of them and that's the case of having them all be listed in the inventory. Again, that's my view and I think care needs to be taken for any of that and good analysis needs to be done.
~ Absolutely.
~ IMG has a lot of archives that would not validate with this because we have all the old versions in there. We were thinking about changing that, however, that doesn't help old archives. I kind of like the idea of making it an option in validate, like check versions, however, the point that was raised about it being technically difficult is not to be dismissed lightly because it is one thing if you have a self contained bundle, you can just look in the inventory file, but if it's an accumulating bundle and you have an inventory file that you are merging with, normally you don't have the full inventory file when you do validation because you're validating what you are adding to the bundle, and so validate has to be able to find the fill inventory in order to be able to check the versions and that's likely to be quite expensive, so I'm not sure if we want to do this, do that check, even by default. Maybe explicitly it can be turned on or something because it may be very slow. There's value in doing a best practice in the DPH. Can say here's what versions are generally supposed to be and here are some exceptions to that, just as a guideline for people. I don't think we should be mandating one way or the other right now.

Question: Interesting. Thought at the F2F IMG had sort of come around to this way of thinking. This surprises me slightly. I know IMG had not followed this in the past, but I thought that you had sort of come around. Am I wrong?
Answer: No. I think that we will probably. We are planning on moving to this paradigm. Not there yet, we've got some tool things. We're planning on doing it, but we are sensitive to NAIF's concerns and other concerns about calibration documents and things like that. It's hard to say one size fits all, that's why I was suggesting that maybe as a best practice or guideline saying this is what versioning usually means, but then here are some cases where it doesn't. It's hard to make a case for actual data products where you don't version it normally, except in NAIF's case, but for things like documentation and calibration there are good use cases for not doing that, so that might be part of the guidelines.
~ Yeah, and I think you made a key point because this issue, I believe, arose a lot out of the fact that early on the standards were made intentionally vague in some areas to allow those of us who had to use that standard to create archives to do something reasonable and to have the flexibility to do that instead of having to come back and make changes every time we bumped into a wall where we needed a hallway. I think what has happened because of that is that we have gotten different understandings of how things should be used and what they mean within PDS itself. Again, the way PPI has always thought about it is if we've got two versions of the same LID in a single inventory then that's an error, but that isn't documented anywhere. It's just an assumption that PPI has always made. I think we need to document what the expectation is and what things really mean. That's one of the things that did come out of the F2F.
~ One small quibble with the way something was said. Said we need to decide what version means and I actually see that as a separate issue. I guess philosophical, but very important in practice, whether we allow LIDs to be duplicated with different VIDs in a collection is not the same thing as what versioning incrementation means. I do think we need to have a separate discussion about what a version incrementing means by default and when that doesn't apply because if it doesn't always mean supersession we should be able to document that. We shouldn't rely on the user to derive the meaning and we should also require it to be documented in my opinion, but that's a separate question than what should be allowed in a collection. Nit picking, but I think it's important in practice. To answer the earlier question, I know SBN and IMG have both used the multiple LIDVIDs per collection approach consciously and intentionally, unlike most other nodes, but both of them appreciated where we were coming from and both expressed a willingness to converge on this mode, so this is somewhere that we have consulted other nodes that do this and they decided that this was a reasonable approach going forward. This is not something ad hock. I also think that we shouldn't be so reticent to force a standard because we are a standards committee and there are times when standards are called for. We should bear that in mind, but I also want to echo the concern about validate performance. I am especially concerned about the primary/secondary solution. Even though it would work well for NAIF and seems like a possible solution, it could be prohibitively expensive. We have to consider if we should think about a different approach or if that sort of forces our hand in some way.
Question: One question - when you said you consulted with the nodes, have you consulted with the user base? The producers base? I think it's very important at this point to not just constrain these discussions to just the node. Of course, we don't want to go to PDS as a whole, but each node could have done that. I would imagine that after ten years of usage and expectations and having only physical files in hand, users have developed some processes that need to be honored and acknowledged.
Answer: They may, but I think it's at the discretion of the node to decide what is best. I fully agree that that would make sense. The reason GEO has had an issue with this is that we have had users

contact us confused by why there were multiple versions in a single collection, so we have been consulted by users rather than consulting users directly, and that is part of the reason why we are pretty strong on this particular point.
~ Then the question is how many users said absolutely I want to keep it this way. So, I know that when somebody is confused they will reach out and ask, but people who are happy aren't going to reach out and ask, they are just going to keep using it and be building their tools.
~ Someone thinks the difference is that it's built into your system. That is how your system is designed to work. This was, at best, an exception, and at worst, an error, when this happened with the data sets that GEO has been dealing with, so that's a very important distinction here that we shouldn't over look.
~ Yes, and for NAIF it was a big part of the design. It was a design principal on which archives were built. Totally the opposite from what happened at GEO.

It looks like there is one last comment and then we are out of time so we will cut off after that, but we do need to continue this discussion and we will in future meetings.

Just a quick thought is that we could put a flag in the collection saying version is interpreted as one way or the other if we wanted to document this potentially. As I'm saying, IMG wants to move in this direction. I'm just uncomfortable about forcing others and that's why I'm trying to make it optional.
~ Okay.

Next Meeting May 15

title: DDWG Notes 2025-05-15
layout: default
date: 2025-05-15

May 15, 2025
Notes by Debra Kazden

Known Attendees: A. Bailey, J. Blalock*, B. Deen, B. Hirsch*, L. Huber*, S. Hughes*, R. Joyner, D. Kahan*, D. Kazden, A. Knight, P. Lawton, T. Lim*, J. Mafi*, S. McLaughlin, E. Schaefer*, B. Semenov*, and M. Tiscareno* (*'s indicate people who voted during this meeting - if applicable)

DDWG Agenda

(Included in email sent May 13, from J.Mafi, Agenda for 2025-05-15) We will have a DDWG telecon this Thursday, May 15th at 9:30 PDT (12:30 EDT, 10:30 MDT). We will cover the following issues:

1) Action item reports.
2) CCB#59 funding year only allows the value "9999" – This SCR has been updated to be a simple bug fix of the original implementation of funding year so that it allows year values other than “9999”. Please be ready to vote on this proposal.
3) CCB#58 New Product Native values for MSL/Mastcam – A document describing the proposed new product native format has been added to the SCR, along with sample Python code for reading data in this format. Please review these materials and be ready to discuss the merits of this proposal.
4) CCB#21 (JIRA-CCB-367) Provenance LDD Using PROV-DM – While LDD’s are not subject to DDWG approval, there has been a request that the DDWG review and discuss the reasons for developing the Provenance LDD.

Front-matter/What’s new

1) Submitted to the CCB: (none)
2) New issues in GitHub:
   CCB#63 - International characters / Unicode support for organization name and display full name
3) Ready for vote: (none)
4) Build 15.1 freeze date: 2/27/2025. The following SCRs are queued for implementation: CCB#13, CCB#48, CCB#49, CCB#54, CCB#55.
5) Withdraw by author:
   (Brief Discussion)

Action Items

a) CCB#39 – S. Hughes to update SCR
b) CCB#12 – S. Hughes to verify whether this is subsumed by CCB#52 and CCB#53
c) CCB#20 & CCB#32 – Tiger team formed and working the issue
d) CCB#31 – Ron and Matt working the issue **(Not Discussed)**

Tier 1 issues

- CCB#62- Transfer Manifest class should enforce data types for its two mandatory columns
- CCB#61 - Disconnect between SR and Validation of ASCII Real
- CCB#60 - New permissible value for File Area Observational.Header.parsing standard id: "CDF 3.9 ISTP/IACG"

- CCB#59 - funding year only allows the value "9999"
- CCB#58 - New Product Native values for MSL/Mastcam
- CCB#57 - Allow Encoded Native in File Area Observational Supplemental
- CCB#56 – Reconsider the PDS4 build cycle cadence
- CCB#53 – Improve cross-namespace validations in sch files
- CCB#52 – Enable cross-namespace validations in sch files
- CCB#51 – Add an optional attribute in the Target Identification class to identify Primary/Secondary targets
- CCB#50 – Clarify Ambiguous Author Information in Product Document
- CCB#46 – Disallow multiple versions of the same LID from occurring in a single collection
- CCB#44 – Issues caused by the failure to expose pds: attributes that are referenced in an LDD.
- CCB#41 – Create a way for classes or attributes to be exposed to other LDDs but not to data providers
- CCB#39 – Fixed problem where LDDTool does not include all rules from IngestLDD
- CCB#37 – Add List Author, List Editor, and List Contributor to the Modification Detail class
- CCB#36 – Add Mission Phase to the Investigation class
- CCB#35 – Add Alias Index and Acronym Index to the Investigation class
- CCB#32 – Add Host type to Observing System Component
- CCB#31 – Ensure that dictionary versions are specified in labels
- CCB#28 – As a user, I want validate to allow delimited tables with whitespace-only numeric fields
- CCB#23 (JIRA-CCB-211) – XML/schema based files as archival data
- CCB#22 (JIRA-CCB-164) – Display Settings not required for images
- 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#21 (JIRA-CCB-367) – Provenance LDD Using PROV-DM
- CCB#20 (JIRA-CCB-358) – Instrument Package Context Products
- Tiger Team: M. Tiscareno (lead), B. Deen, B. Hirsch, S. Hughes, R. Joyner, J. Mafi, E. Schaefer.
- CCB#12 – Document Nuances for Reusability of Generic LDD Classes/Attributes
- CCB#5 – Move to V.2.0.0 for June 2025 Release
- CCB#4 – Backwards Incompatibility policies
- -Presentation: Backwards Compatibility
- CCB#3 – Adopt Semantic Versioning with version 2 transition
- The CCB returned this, requesting that it be resubmitted as part specific proposal to transition to IM V2.0.0.
  (Some Discussions)

Tier 2 issues

1) 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!

1) JIRA-CCB-326 (link) - Inconsistent Direction on How to Use Discipline Facets
2) RS: 259/281 (see new attachment to JIRA-CCB-281) (on hold)
3) JIRA-CCB-364 (link) – Disallow bundles referencing collections by LID **(Not Discussed)**

Historical links:
- Export of history: https://github.com/NASA-PDS/PDS4-CCB/tree/main/data
- Backup: https://nasa-pds.atlassian.net/jira/software/c/projects/CCB/issues (limited seats)

Next meeting: 29 May 2025

DDWG

Front Matter

Question: Any action item updates?
Answer: (Silence)

Tier 1 issues

CCB#59 - funding year only allows the value "9999"

This has been updated with the changes that were discussed last week - just fixing the pattern value to make it so the full set of numeric values are available - any four digit year. Decided not to address other issues since this is basically a bug fix.

Question: Any comments before we vote?
Answer: (Silence)

The Vote to approve or not approve CCB#59:
ATMOS - Yes
EN - Yes
GEO - Yes
IMG - Yes
IPDA - Yes
NAIF - Yes
PPI - Yes
RMS - Not Here
(Arrived after the vote, asked to be counted as a Yes)
RS – Yes
SBN - Yes

That passed. It will be pushed to the CCB.

Question: Is it clear that all we did is a bug fix?

Answer: Good question. Joe will put a note into the original SCR that that is what was done before sending it to the CCB. **(Action Item – Joe)**
~ They may not notice.
~ We can email them too.

CCB#21 - Provenance LDD Using PROV-DM

This is not something that needs DDWG approval, but Steve wanted DDWG comments and review on some of the ideas being proposed here.
~ The big picture is that we have a provenance LDD that we have developed. Four use cases were proposed. The first to be implemented is going to be superseded LIDs. Use cases were presented at the F2F. We won't go through that all now. One new thing that wasn't discussed at the F2F was the concept of reasons. There are five: replacement, duplication, merge, obsolete and policy change. They are all in the latest attachment to the SCR from Steve. There is one comment.
~ It's an example of one where a LID changed. Hopefully it is done right.
~ More examples will be done. So far we just have product context and product observational.
~ Those will be moved to the applicable documents **(Action Item - Joe)**
~ These could be in any product. Think that's the proposal.

CCB#58 - New Product Native values for MSL/Mastcam

Question: The proposal is to add some new types for product native. In the SCR, MSL/Mastcam and a few others are specifically mentioned, but there are additional cameras. Would those be covered too? Same format?
Answer: No proposal at this time to include the M20s, but they are the same format.

Question: Any additional questions or comments?
Another Question: Malin Space Science Systems (MSSS) is okay with this?
Answer: They asked for this and wrote the documentation that describes the formats.
Another Question: Any questions on the documentation? The next step is that we should review the documentation and make sure it is sufficient for product native.
Answer: The intent is to have it in the bundle, but it won't have a LID until after this is approved. Future bundles can point to it.

Question: Any further discussion or comments on this?
Answer: (Silence)
~ If no comments, would like to vote on this next time. Need to review the documents first, before we pass them to the CCB. We will vote in two weeks.

Back to CCB#21 and CCB#59 discussions

Question: Regarding due diligence on CCB#21, can we do a straw poll on the list of reasons? If there are comments, please supply those.

Question on CCB#59: Did we vote on CCB#59? RMS would like to change their vote from absent to yes.

Answer: Okay.

Someone looked at the example "reason" InSight label posted. We can hold a straw poll or provide recommendations on the list next week.

Question: Is there anything else people want to discuss?
Answer: On that label - it has attribute values. Concerned about the keyword names. Hard for search.
~ Good point. A design decision based on W3C model. In this case, was trying to minimize what was being required here. Could have used a providence relationship, like derived from, but just used supersede reason. The short answer is there's no other place to give this information.
~ Could do prov:reason. It would simplify search.
~ Will take a look at that to see if it can be made to work. In provenance you just define entities, relationships and such. Will see what we can do.
~ Maybe could look at different ways to handle search for this.

Question: Other comments?
Answer: Back to CCB#46.

CCB#46 - Disallow multiple versions of the same LID from occurring in a single collection

NAIF added a comment in GitHub and is asking for nodes to access impact, see how many collections are violating the proposed changes and if it was an error or has merit. We haven't seen any real examples of impacts.
~ Very helpful to have that information. Maybe EN has a way to do that.
~ Think the nodes would know better.
~ EN should be able to do it for all ingested collections. If no tool now, maybe it's one they could provide.
~ Someone can think of a way to do it. Interesting question. Can probably do it for us. Think it probably does exist at PPI as an error, but might be surprised. A good thing to check.

Question: IMG just wrote a tool that will check for duplicate LIDs. People can run this. Are we still considering this as more than a best practice guideline?
Answer: There was a presentation at the MC yesterday by Tom Stein. We need to examine the PDS policy on versioning. We will start in the DDWG by looking at different use cases. Will write a summary. SBN had mentioned using version a different way where it would make sense to have multiple versions of the same LID in an inventory, and NAIF wants to gather all examples to try to come up with recommendations for a policy that would go to the MC.
~ Someone doesn't think we will have one policy. Wouldn't vote for anything that would be a problem for NAIF.
~ Starting point of the discussion should be to carve out an exception for SPICE.
~ NAIF also needs it for ancillary and documents, which kills it for everyone else.
~ And some for calibration.

That takes us back to this request. Let's all take a look.

Question: Is this more of a validate issue? If yes, could have a flag to check or not for multiple versions of a LID, but if it's about what the registry delivers that is a different issue.

Answer: Need to do some reconnaissance to figure out what is going on. Need a tool from EN to search the registry.
~ A tool could probably be set up to do this. Could be an interesting project. We might already have the framework for this.
~ API might already be able to do this.
~ But we would still need the nodes to check. Knowing their content.
~ There is some work to do here.

Question: Any last comments?
Answer: IMG will make their script available if anyone wants to use it.

Next meeting May 29.

title: DDWG Notes 2025-05-29
layout: default
date: 2025-05-29

May 29, 2025
Notes by Debra Kazden

Known Attendees: J. Blalock*, L. Huber*, S. Hughes*, R. Joyner, D. Kahan*, D. Kazden, A. Knight, P. Lawton, S. Loftin, J. Mafi*, S. McLaughlin, E. Schaefer*, B. Semenov*, M. Seritan and M. Tiscareno* (*'s indicate people who voted during this meeting - if applicable)

DDWG Agenda

(Included in email sent May 28, from J.Mafi, Agenda for 2025-05-29 (Revised)))

Below please find a modified agenda for tomorrow's DDWG telecon. Matt pointed out that the CCB#20 vote outlined for tomorrow had already by conducted on the March 6 telecon. My apologies for the mix up.
We will have a DDWG telecon this Thursday, May 29th at 9:30 PDT (12:30 EDT, 10:30 MDT). We will cover the following issues:

1. Action item reports.
2. CCB#58 New Product Native values for MSL/Mastcam – Please review the attached documentation, including the descriptions of the mini header and payload data formats and be ready to discuss and vote on this proposal.
3. CCB#60 New permissible value for File Area Observational.Header.parsing_standard_id: "CDF 3.9 ISTP/IACG" – This request is for a new permissible value for an existing attribute. We will discuss and, as possible, vote on this SCR.
4. PDS4 Versioning Use Cases Survey – We will set aside some time to discuss any questions or concerns associated with this project.

Front Matter

1) Submitted to the CCB:
   - CCB#59 funding year only allows the value "9999" – Passed
2) New issues in GitHub:
   - CCB#64 Add additional values to investigation as an allowable reference type
   - CCB#65 Modify description attribute to be required in Telescope and Facility Context Products
3) Ready for vote:(none)
4) Build 15.1 freeze date: 2/27/2025. The following SCRs are queued for implementation: CCB#13, CCB#48, CCB#49, CCB#54, CCB#55.
5) Withdraw by author:(none)
   (Some Discussion)

Action Items

a) CCB#39 – S. Hughes to update SCR
b) CCB#12 – S. Hughes to verify whether this is subsumed by CCB#52 and CCB#53
c) CCB#20 & CCB#32 – Tiger team formed and working the issue
d) CCB#31 – Ron and Matt working the issue
   (CCB#20 Mentioned)

Tier 1 issues

- CCB#63 - International characters / Unicode support for organization name and display full name
- CCB#62 - Transfer Manifest class should enforce data types for its two mandatory columns
- CCB#61 - Disconnect between SR and Validation of ASCII Real
- CCB#60 - New permissible value for File Area Observational.Header.parsing standard id: "CDF 3.9 ISTP/IACG"
- CCB#59 - funding year only allows the value "9999"
- CCB#58 - New Product Native values for MSL/Mastcam
- CCB#57 - Allow Encoded Native in File Area Observational Supplemental
- CCB#56 – Reconsider the PDS4 build cycle cadence
- CCB#53 – Improve cross-namespace validations in sch files
- CCB#52 – Enable cross-namespace validations in sch files
- CCB#51 – Add an optional attribute in the Target Identification class to identify Primary/Secondary targets
- CCB#50 – Clarify Ambiguous Author Information in Product Document
- CCB#46 – Disallow multiple versions of the same LID from occurring in a single collection
- CCB#44 – Issues caused by the failure to expose pds: attributes that are referenced in an LDD.
- CCB#41 – Create a way for classes or attributes to be exposed to other LDDs but not to data providers
- CCB#39 – Fixed problem where LDDTool does not include all rules from IngestLDD
- CCB#37 – Add List Author, List Editor, and List Contributor to the Modification Detail class
- CCB#36 – Add Mission Phase to the Investigation class
- CCB#35 – Add Alias Index and Acronym Index to the Investigation class
- CCB#32 – Add Host type to Observing System Component
- CCB#31 – Ensure that dictionary versions are specified in labels
- CCB#28 – As a user, I want validate to allow delimited tables with whitespace-only numeric fields
- CCB#23 (JIRA-CCB-211) – XML/schema based files as archival data
- CCB#22 (JIRA-CCB-164) – Display Settings not required for images
- 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#21 (JIRA-CCB-367) – Provenance LDD Using PROV-DM
- CCB#20 (JIRA-CCB-358) – Instrument Package Context Products
- Tiger Team: M. Tiscareno (lead), B. Deen, B. Hirsch, S. Hughes, R. Joyner, J. Mafi, E. Schaefer.
- CCB#12 – Document Nuances for Reusability of Generic LDD Classes/Attributes
- CCB#5 – Move to V.2.0.0 for June 2025 Release
- CCB#4 – Backwards Incompatibility policies
- Presentation: Backwards Compatibility
- CCB#3 – Adopt Semantic Versioning with version 2 transition
- The CCB returned this, requesting that it be resubmitted as part specific proposal to transition to IM V2.0.0.

(Some Discussions)

Tier 2 issues

1) 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!

1) JIRA-CCB-326 - Inconsistent Direction on How to Use Discipline Facets
2) RS: 259/281 (see new attachment to JIRA-CCB-281) (on hold)
3) JIRA-CCB-364 – Disallow bundles referencing collections by LID **(Not Discussed)**

Historical links:
- Export of history: https://github.com/NASA-PDS/PDS4-CCB/tree/main/data
- Backup: https://nasa-pds.atlassian.net/jira/software/c/projects/CCB/issues (limited seats)

Next meeting: 12 June 2025

DDWG

Front Matter

There is work for the CCB#20 Tiger Team to do before we have a final vote on that one.
~ EN needs to test the implementation.
~ Schema and schematron and edits to the SR and DPH need to be done.
~ Fair.
~ We should decide who is going to do what so we can move this forward.

Tier 1 Issues

CCB#58 - New Product Native values for MSL/Mastcam

Question: Hopefully everyone has looked at the documents from the provider for the data products. Any comments on the documents or the proposal to approve this as a new type of product native? We have discussed this in the past.
Answer: GEO thought the documents deserve a full peer review.
~ That is not a bad a idea. It would delay this, but a good point with regards to something of this nature, to make sure the documentation really describes the data.
~ This is only the second type we have added. Not a universal type, like SEED.
~ True, but the number of missions with Malin instruments on them is a large number.
~ Agree, but as a standard… (Interrupted) Comment and Another Question: Good point. We need an idea of reviewers for this document. People from the DDWG would be okay if they have the right expertise. But first, are there any other concerns?
Answer: (Silence)

We need volunteers for a peer review panel for this document - to give it a full review.

~ (Deathly silence.)
~ We will need two or three reviewers, including someone from IMG.
~ IMG will figure out who can do that. Working with a skeleton crew. Will let Joe know who will participate. **(Action Item - John)**
~ It could be a data user.
~ IMG needs to find someone with the time.
~ And we need a couple of others.

Question: Any nodes know someone with the expertise?
Answer from Ethan (GEO): Not an expert, but will throw my hat in.
~ Joe will also be on the panel, but not an expert.
Another Question: Anyone else?
Answer: (Silence)
~ We will get this going. Bob Deen is not here, but Joe and John will let him and the Malin folks know about this so they can participate. **(Action Item)**

CCB#60 - New permissible value for File Area Observational.Header.parsing standard id: "CDF 3.9 ISTP/IACG"

This is just a request to add a new permissible value to support CDF version 3.9.

Question: Any discussion or concerns?
Answer: (Silence)
~ If there are no concerns, would like to vote on this one.

The Vote for CCB#60:
ATMOS - Yes
EN - Yes
GEO - Yes
IMG - Yes
IPDA - Not Present
NAIF - Yes
PPI - Yes
RMS - Abstain
RS - Yes
SBN - Not Present

Passed. It will go to the CCB.

CCB#60 needs a TA and test cases.
Action Item - Joe will put those together.

The CCB did approve **CCB#59 (funding year only allows the value "9999")** as a straight forward bug fix.

Use Case Survey

Joe sent everyone an email request a week ago. Matt pointed out that some nodes may need some help from EN to see if multiple LIDs appear. PPI is working on their own script.

Question: Any questions on the request? Unclear if we really need software. It will take time to go through holdings and identify different uses of version. SBN has already started. They want to get an idea of how many of their holdings are included. The desire is to understand the frequency of use in the nodes an PDS. It may not be frequently used, but it's key to how things run. Any questions or comments about this task?
Answered with a Question: How do you want this submitted? In the CCB notes? In a white paper?
Answer: More of a white paper write up. This is policy, not specifically for CCB#46 (Disallow multiple versions of the same LID from occurring in a single collection). Joe will collate and provide results the the MC. Currently due June 13. People can let Joe know if they need to extend that.
~ Someone already thinks we do. This request is something the registry should be able to do. It's not a good use of time to do this in an ad hock way. EN should provide a tool.
~ Joe can talk to Jordan **(Action Item - Joe)**
~ RMS is not doing any work on this. Waiting for EN.
~ IMG agrees. Without clear instructions and tools, can't justify the time to do this. Bob has said there are multiple versions, but we don't have time to figure out the specifics without tools to do it. Also, unclear how much context is wanted on why something was done.

Someone else thinks the context will be helpful. One of the statistics that we want is each node's way of using versioning. We shouldn't worry about the June 13 deadline. That is off the table. Joe will discuss this with EN.
~ Regarding context, IMG will try to understand what was done and why.
~ Someone agrees with what is being said and thinks Jordan will agree too.

Action Item - Joe will discuss having a tool from EN with Jordan.
~ For consistent results.

Question: Did all nodes really register all old versions?
Answer: For PPI, believe PPI has some older versions of things that are not registered, but think the current active stuff is more important. Superseded might not be a good example, at least for PPI.
Another Question: Understand that some early versions are superseded - any requests for them? Should we predict requests for reproducibility?
~ We don't get a lot of requests, but we should be thinking about that. The registry should be able to serve older versions.
~ NAIF put all SCLK kernels in the archive because people, including PPI, wanted the old kernels. SPICE isn't the only place where this will be applicable.

Question: Other concerns?
Answer: (Silence)

Action Item - Joe will contact Jordan an send out a revision of the email.

Thank you to everyone who worked on **CCB#21 (Provenance LDD Using PROV-DM)** comments. The input is appreciated.

Question: That's it for the agenda today. Are there any other SCRs that people want to look at?
Answer: Yes. **CCB#64 - Add additional values to investigation as an allowable reference type**.

This is a new one.
~ Yes. Another enumerated value. Not much of a difference from the one we just voted on. This SCR just suggests adding two new values. Schematron has a long list. Would like to know if there are any objections to these.
~ Not sure people have looked at this yet. We can put it on the agenda for next time and have a vote.
~ Sounds good.

Someone is looking now. Context product types that already exist, need to add references.
~ Wanted specific reference from instrument to investigation.
~ That's kind of weird.
~ Reporter got a request to write the SCR, so did.

Question: Can Reporter find out more about the use case? Telescope is not usually bound to an investigation.
Answer: Can ask. **(Action Item - Ron)**
~ They may think they want it because they aren't understanding this correctly. We don't want to do something that doesn't make sense.

Another Question: Who requested this?
Answer: Reporter doesn't remember, but will look for a use case.
~ Someone will write a comment.
~ Someone else was communicating with Jordan a few weeks ago and understands that EN is revising context products, adding instrument LIDs, trying to do some high up linking. Maybe that's where this came from.
~ Knowing the use case would be useful.

title: DDWG Notes 2025-06-12
layout: default
date: 2025-06-12

June 12, 2025
Notes by Debra Kazden

Known Attendees: A. Bailey, J. Blalock*, L. Huber*, S. Hughes*, R. Joyner, C. Juarez, D. Kahan*, D. Kazden, A. Knight, P. Lawton, S. Loftin, J. Mafi*, S. McLaughlin, E. Schaefer*, B. Semenov*, M. Seritan and M. Tiscareno* (*'s indicate people who voted during this meeting - if applicable)

DDWG Agenda

(Included in email sent June 10, from J.Mafi, Agenda for 2025-06-12)

We will have a DDWG telecon this Thursday, June 12th at 9:30 PDT (12:30 EDT, 10:30 MDT). We will cover the following issues:

1) Action item reports.
2) PDS4 Versioning Use Cases Survey status update.
3) CCB#28 As a user, I want validate to allow delimited tables with whitespace-only numeric fields – The "Proposed Solution" and "Requested Changes" have been filled in for this SCR. I would like to discuss these and vote on whether they are sufficient. If so, a marked up version of the SR with the requested changes will be produced for a final vote.
4) CCB#64 Add additional values to investigation as an allowable reference type – The following link is to some slides outlining a number of use cases: https://docs.google.com/presentation/d/1xeZvGo46BzyTatrTBcAQut-dQY3gJEXtmrv7o0f_7yk/edit?slide=id.g1139100b8a0_0_0*slide=id.g1139100b8a0_0_0 We will go over these slides during the DDWG telecon. Please be prepared to discuss and possibly vote.

Front Matter

1) Submitted to the CCB: (none)
2) New issues in GitHub: (none)
3) Ready for vote:(none)
4) Build 16.0 freeze date: 8/??/2025. The following SCRs are queued for implementation: CCB#59
5) Withdraw by author:(none)
   (Brief Discussion)

Action Items

a) CCB#39 – S. Hughes to update SCR
b) CCB#12 – S. Hughes to verify whether this is subsumed by CCB#52 and CCB#53
c) CCB#20 & CCB#32 – Tiger team formed and working the issue
d) CCB#31 – Ron and Matt working the issue **(Not Discussed)**

Tier 1 issues

- CCB#65 Modify description attribute to be required in Telescope and Facility Context Products
- CCB#64 Add additional values to investigation as an allowable reference type

- CCB#63 - International characters / Unicode support for organization name and display full name
- CCB#62 - Transfer Manifest class should enforce data types for its two mandatory columns
- CCB#61 - Disconnect between SR and Validation of ASCII Real
- CCB#60 - New permissible value for File Area Observational.Header.parsing standard id: "CDF 3.9 ISTP/IACG"
- CCB#58 - New Product Native values for MSL/Mastcam
- CCB#57 - Allow Encoded Native in File Area Observational Supplemental
- CCB#56 – Reconsider the PDS4 build cycle cadence
- CCB#53 – Improve cross-namespace validations in sch files
- CCB#52 – Enable cross-namespace validations in sch files
- CCB#51 – Add an optional attribute in the Target Identification class to identify Primary/Secondary targets
- CCB#50 – Clarify Ambiguous Author Information in Product Document
- CCB#46 – Disallow multiple versions of the same LID from occurring in a single collection
- CCB#44 – Issues caused by the failure to expose pds: attributes that are referenced in an LDD.
- CCB#41 – Create a way for classes or attributes to be exposed to other LDDs but not to data providers
- CCB#39 – Fixed problem where LDDTool does not include all rules from IngestLDD
- CCB#37 – Add List Author, List Editor, and List Contributor to the Modification Detail class
- CCB#36 – Add Mission Phase to the Investigation class
- CCB#35 – Add Alias Index and Acronym Index to the Investigation class
- CCB#32 – Add Host type to Observing System Component
- CCB#31 – Ensure that dictionary versions are specified in labels
- CCB#28 – As a user, I want validate to allow delimited tables with whitespace-only numeric fields
- CCB#23 (JIRA-CCB-211) – XML/schema based files as archival data
- CCB#22 (JIRA-CCB-164) – Display Settings not required for images
- 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#21 (JIRA-CCB-367) – Provenance LDD Using PROV-DM
- CCB#20 (JIRA-CCB-358) – Instrument Package Context Products
- Tiger Team: M. Tiscareno (lead), B. Deen, B. Hirsch, S. Hughes, R. Joyner, J. Mafi, E. Schaefer.
- CCB#12 – Document Nuances for Reusability of Generic LDD Classes/Attributes
- CCB#5 – Move to V.2.0.0 for June 2025 Release
- CCB#4 – Backwards Incompatibility policies
- Presentation: Backwards Compatibility
- CCB#3 – Adopt Semantic Versioning with version 2 transition
- The CCB returned this, requesting that it be resubmitted as part specific proposal to transition to IM V2.0.0.
  (Some Discussions)

Tier 2 issues

1) 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!

1) JIRA-CCB-326 - Inconsistent Direction on How to Use Discipline Facets
2) RS: 259/281 (see new attachment to JIRA-CCB-281) (on hold)
3) JIRA-CCB-364 – Disallow bundles referencing collections by LID **(Not Discussed)**

Historical links:
- Export of history: https://github.com/NASA-PDS/PDS4-CCB/tree/main/data
- Backup: https://nasa-pds.atlassian.net/jira/software/c/projects/CCB/issues (limited seats)

Next meeting: 26 June 2025

DDWG

Front Matter

There was some discussion on the PDS4 version use case survey at the MC on Wednesday. They are asking each node to use their own method to figure it out. Jesse Stone provided a command line that can be used and a number of others have methods as well. Believe Jesse's is for Linux - can use that or come up with your own.

Question: Any questions or comments?
Answer: For IMG the issue isn't that they can't write a script, the problem is that they don't have the time. Can't justify the time. Knows we want this, but don't think they can justify the time to go through their two sets of holdings - at USGS and JPL. Can try, but can't promise. Can see what the team thinks, but not expecting a warm reception.
~ Understood.
~ Also, IMG is not understanding the context here. We know it's being used, so to us, we need to accept it and do a work around.
~ The ultimate goal is an MC policy on versioning. If you don't submit anything you won't be considered.
~ Seems very cart before the horse.
~ The discipline nodes already agreed. The only node that strongly objects has already turned in their survey.
~ Yes, the other nodes answered, but without due diligence. Need to do due diligence before coming up with a policy. Understand time constraints, but if you want to say we accessed the impact, you need to actually do that.
~ Thanks. We can take that back to the MC.

Tier 1 Issues

CCB#28 - As a user, I want validate to allow delimited tables with whitespace-only numeric fields

There has been some confusion on this one. The proposal is to allow values with spaces only. Been some discussion. The proposal is to clarify the SR on how spaces need to be considered in DSV. See GitHub comments. It will say can be space padded. What's there now is ambiguous.

Question: Any thoughts or comments?
Answer: It needs a TA.
~ Agreement.
~ Someone says they think the revision to max will affect tools. Actually working on it.
~ Yes, it will affect validate. Good point.
~ The whole SCR will affect validate.
~ That will need to be listed in the impact statement and we will need to submit a validate ticket if this gets approved.

Question: Any other comments?
Answer: (Silence)

The Vote to approve this approach:
ATMOS - Yes
EN - Yes
GEO - Yes
IMG - Yes
IPDA - Not Present
NAIF - Yes
PPI - Yes
RMS - Yes
RS - Yes
SBN - Not Present

The approach is approved. It will be implemented, impact statement and TA will be done and we will bring it back for a final approval.

CCB#64 - Add additional values to investigation as an allowable reference type

Hopefully everyone has looked at the slide deck provided by EN on why this is being requested. There have also been comments in GitHub.

This doesn't seem necessary. The only argument seems to be the number of clicks. Redundant information, which can create issues, based on the use cases presented.
~ Someone completely agrees. Think this has been a disagreement in philosophy for awhile. The web design people want to change the archive to fit their design. Seems like the wrong way around. This is all about making web design easier, but it would make archiving harder.

~ Someone else is starting to come around to the other point of view. Wonders if it's not legitimate to think of web designers as users. If we can make changes to support them without harm to other users or DPs, we should. This reminds (me) of the early days of PDS4. PPI had made a design decision to only include references to instrument because you could get the rest of the information, then PPI later came around to agree that it doesn't hurt anything to include all the information. Yes, potential links would need to be resolved, but think CCB#20 (Instrument Package Context Products) will help with some of this.

Another person is going to try to straddle the fence. It's been an issue for a long time. First, would like to know who proposed this.
~ Think it came from the web designers.
~ Don't want to contradict Jordan if it is from him, but if we view the IM as an ontology in a tool, which it is, in Protegee, then there are tools called reasoners that can be run. Unfortunately, we aren't there yet. Cleaning up the ontology for context products now. Ultimately, they should contain the few necessary relationships so that all other relationships can be inferred. Not there yet, but we will get there, so think we could pass this for the web designers to continue their work, but concerned about ending up with spaghetti code.
~ Someone mostly agrees, but still confused about adding this if going to have reasoners. (Spoke very quickly.) Daisy chaining is the more elegant solution. As a programmer, want the minimal information so we can derive other information dynamically. EN can do that since they offered to retrofit the context products. Think that it has risk of harming current structure because it can introduce discrepancies.
~ Someone else agrees.

Question: Other comments?
Answer: Pretty much agree as far as these particular use cases, but think there's a significant difference between directed graphs and ontologies. The point is that we are trying to make it so web designers can get what they need by extracting it from the ontology. Hope we have all the basic relationships we need to get from A to B and B to C.
~ Yes, A to B and B to C, but we don't need to include A to C. Can infer that.
~ Yes, we need the core basic information to do what we need to do.

CCB#20 is an important part of this current SCR. Ontology is described. CCB#20 resolves that by defining the relationships.
~ Someone pretty much agrees. Also thinks the necessary information is already in the IM. This would repeat it for the web designers convenience. Wish they had come to us to discuss this and other issues. This would need to be required, so a lot of effort. Seems larger then the effort to do the daisy chain.
~ Yes, and validation would be difficult because these context products are what we would want to validate against.

Question: Any additional comments? Any desire to look at the slide presentation?
Answer: Someone else wants to weigh in. We've got the web group who's requests come through Jordan, so Jordan wants this. He asked for it and there must be some basis for why he wants this.
~ People think this could cause harm.
~ It can cause discrepancies.

~ Yes, so don't see any reason to open this up to more discrepancies. Jordan may have a good reason for requesting this, but not seeing that in the use cases. Think CCB#20 will exacerbate this too and create potential problems in validate in the future without added archive value.
~ When the TA is done and the impact statement is written we need to point out the very large number of context products that will need to be retrofitted. Not insignificant.

Question: Any other comments? EN provided a slide deck. It might be useful to talk to who ever is behind the SCR to discuss why they think this is necessary and the potential harm we see.
Answer: Sounds good.
Action Item - Joe will reach out to Jordan to try to have the web design folks attend a meeting to discuss this issue. Direct conversations can be very helpful.

Question: Last comments?
Answer: (Silence)

This is all that was on the agenda for today. We have time if there is anything else people want to discuss. We have a F2F in August. We have an hour and a half of DDWG session time on the first day. People can post any recommendations for topics of discussion to the online spreadsheet. Joe will send a link to that at the end of the meeting.

Someone wants to do a quick look at CCB#61.

CCB#61 - Disconnect between SR and Validation of ASCII Real

They wanted to use a value that didn't fit the pattern. Can't remember the details, maybe we should wait to discuss this. Never mind.
~ If the pattern is causing a problem, we should try to fix it, but they are a lot to deal with.

Question: Is the value valid under the standard? If it isn't we should ask them to change it.
Answer: We need to look at the specific value they are trying to support.
~ The value is posted. The SR requires zero to precede the decimal point. The value doesn't have a zero.
~ Looks like a disconnect from the SR to the pattern.

Action Item - Ron will update the SCR more.

Our next meeting is in two weeks. June 26.

The new version of the IM, 1.24, is now released.

title: DDWG Notes 2025-06-26
layout: default
date: 2025-06-26

June 26, 2025
Notes by Debra Kazden

Known Attendees: A. Bailey, J. Blalock*, R. Deen, L. Huber*, S. Hughes*, R. Joyner, C. Juarez, D. Kahan*, D. Kazden, A. Knight, P. Lawton, T. Lim*, J. Mafi*, S. McLaughlin, E. Schaefer*, B. Semenov* and M. Tiscareno* (*'s indicate people who voted during this meeting - if applicable)

DDWG Agenda

(Included in email sent June 24, from J.Mafi, Agenda for 2025-06-26)

We will have a DDWG telecon this Thursday, June 26th at 9:30 PDT (12:30 EDT, 10:30 MDT). We will cover the following issues:

1) Action item reports.
2) PDS4 Versioning Use Cases Survey status update.
3) CCB#28 As a user, I want validate to allow delimited tables with whitespace-only numeric fields – A document containing the proposed revisions to the SR has been posted under Applicable Attachments. Please review and be ready to vote.
4) CCB#57 Allow Encoded Native in File Area Observational Supplemental – Please review the latest comments on this SCR and be prepared to discuss and possibly vote.
5) CCB#62 Transfer Manifest class should enforce data types for its two mandatory columns – Please review and be ready to discuss this SCR.

Front Matter

1) Submitted to the CCB:
   - CCB#60 New permissible value for File Area Observational.Header.parsing standard id: "CDF 3.9 ISTP/IACG"
2) New issues in GitHub:
   - CCB#66 Improve documentation for List Author, List Contributor, and List Editor
3) Ready for vote:(none)
4) Build 16.0 freeze date: 8/??/2025. The following SCRs are queued for implementation: CCB#59
5) Withdraw by author:
   - CCB#64 Add additional values to investigation as an allowable reference type

   (Brief Discussion)

Action Items

a) CCB#39 – S. Hughes to update SCR
b) CCB#12 – S. Hughes to verify whether this is subsumed by CCB#52 and CCB#53
c) CCB#20 & CCB#32 – Tiger team formed and working the issue
d) CCB#31 – Ron and Matt working the issue

   (Mentioned)

Tier 1 issues

- CCB#66 Improve documentation for List Author, List Contributor, and List Editor
- CCB#65 Modify description attribute to be required in Telescope and Facility Context Products
- CCB#64 Add additional values to investigation as an allowable reference type
- CCB#63 - International characters / Unicode support for organization name and display full name
- CCB#62 - Transfer Manifest class should enforce data types for its two mandatory columns
- CCB#61 - Disconnect between SR and Validation of ASCII Real
- CCB#60 - New permissible value for File Area Observational.Header.parsing standard id: "CDF 3.9 ISTP/IACG"
- CCB#58 - New Product Native values for MSL/Mastcam
- CCB#57 - Allow Encoded Native in File Area Observational Supplemental
- CCB#56 – Reconsider the PDS4 build cycle cadence
- CCB#53 – Improve cross-namespace validations in sch files
- CCB#52 – Enable cross-namespace validations in sch files
- CCB#51 – Add an optional attribute in the Target Identification class to identify Primary/Secondary targets
- CCB#50 – Clarify Ambiguous Author Information in Product Document
- CCB#46 – Disallow multiple versions of the same LID from occurring in a single collection
- CCB#44 – Issues caused by the failure to expose pds: attributes that are referenced in an LDD.
- CCB#41 – Create a way for classes or attributes to be exposed to other LDDs but not to data providers
- CCB#39 – Fixed problem where LDDTool does not include all rules from IngestLDD
- CCB#37 – Add List Author, List Editor, and List Contributor to the Modification Detail class
- CCB#36 – Add Mission Phase to the Investigation class
- CCB#35 – Add Alias Index and Acronym Index to the Investigation class
- CCB#32 – Add Host type to Observing System Component
- CCB#31 – Ensure that dictionary versions are specified in labels
- CCB#28 – As a user, I want validate to allow delimited tables with whitespace-only numeric fields
- CCB#23 (JIRA-CCB-211) – XML/schema based files as archival data
- CCB#22 (JIRA-CCB-164) – Display Settings not required for images
- 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#21 (JIRA-CCB-367) – Provenance LDD Using PROV-DM
- CCB#20 (JIRA-CCB-358) – Instrument Package Context Products
- Tiger Team: M. Tiscareno (lead), B. Deen, B. Hirsch, S. Hughes, R. Joyner, J. Mafi, E. Schaefer.
- CCB#12 – Document Nuances for Reusability of Generic LDD Classes/Attributes
- CCB#5 – Move to V.2.0.0 for June 2025 Release
- CCB#4 – Backwards Incompatibility policies
- Presentation: Backwards Compatibility
- CCB#3 – Adopt Semantic Versioning with version 2 transition

- The CCB returned this, requesting that it be resubmitted as part specific proposal to transition to IM V2.0.0.
  (Some Discussions)

Tier 2 issues

1) 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!

1) JIRA-CCB-326 - Inconsistent Direction on How to Use Discipline Facets
2) RS: 259/281 (see new attachment to JIRA-CCB-281) (on hold)
3) JIRA-CCB-364 – Disallow bundles referencing collections by LID **(Not Discussed)**

Historical links:
- Export of history: https://github.com/NASA-PDS/PDS4-CCB/tree/main/data
- Backup: https://nasa-pds.atlassian.net/jira/software/c/projects/CCB/issues (limited seats)

Next meeting: 10 July 2025

DDWG

Front Matter

Regarding action items and versioning case studies the assignments have been made. Please submit when you can. No deadline anymore, but please submit the versioning study so the MC can make a decision regarding a versioning policy.

New SCR: CCB#66 - Improve documentation for List Author, List Contributor, and List Editor.
~ The issue is that there are no real clear guidelines on how to assign personnel to these.

Question: Any comments on this? Just discussing as an introduction for now.
Answer: (Silence)

Tier 1 Issues

CCB#28 - As a user, I want validate to allow delimited tables with whitespace-only numeric fields

There has been quite a lot of discussion in GitHub. Some proposed updates to the SR were posted, followed by suggested updates to make the proposed updates less ambiguous and to correct typos.

Question: Any thoughts on this? Hope everyone has read the edits. Any concerns?
Answer: Someone liked it. Ready to go forward.
~ We will vote to approve the SCR as written and send it to the CCB.

The Vote for CCB#28:
ATMOS - Yes
EN- Yes
GEO - Yes
IMG - Yes
IPDA - yes
NAIF - Yes
PPI - Yes
RMS - Yes
RS - Yes
SBN - Not Here

There is no TA.
~ We will get a TA and a system impact statement and move this to the CCB.

Seems like we are voting twice on everything now. We need a TA before the final vote.
~ Yes. First we vote to approve the approach, then to approve. TA and system impact should be part of the final approval. Will try to integrate the steps better. Can try to have an evote after the TA.
~ That sounds reasonable, as long as the TA is straight forward.
~ With any evote if anyone wants to discuss something further in a meeting, we can.

CCB#62 - Transfer Manifest class should enforce data types for its two mandatory columns

There were questions in GitHub. The DPH talks about a different checksum file that is part of the delivery. One issue is that the checksum file doesn't have any restrictions on the format, but that would be a separate SCR. This was originally from Mark B.

Question to IPDA: Do you know if the delivery packages that IPDA uses include checksums?
Answer: Unsure.
~ Know that what is described in the DPH is used in a lot of MD5 tools. Would be a good structure to adopt.
~ IPDA agrees.

This change makes it a little more similar to an inventory. If we want to make more changes to manifest we can.
~ Someone wants to point out that manifest and checksum manifest are two separate things. We don't want to confuse them.
~ Someone else believes there are different intents.

Question: Other thoughts? On adding restrictions?
Answer:(Silence)

Action Item - Joe will follow up with Mark to see if he wants to add any additional restrictions.

The other issue mentioned in the GitHub comments is that this adds requirements on the transfer manifest class, so we will need to update the documentation. A good template is the inventory class.

Question: Has NSSDCA commented on this? Would like to hear from them.
Answer: Their representative is here.
~ Just want to cover all bases.
~ NSSDCA says they use the PDS4 Deep Archive Tool. The tool uses manifests (checksum and transfer) to make the SIP, which they use. The tool chews through them to make the SIP. NSSDCA only looks at the SIP.

Action Item - Joe will check with EN on how this could affect them and the PDS4 tool and the SIP. And **Stef** will add a comment in GitHub.
~ Someone agrees that it is a good idea to talk to EN.
~ This SCR wouldn't change the file, but would make sure they are more consistent with each other.

To sum up, the **action items** are that **Joe** will follow up with EN to make sure no impacts on the Deep Archive Tool and **Stef** will put a comment in GitHub on the use of the files to produce the SIP.

Would like to vote to approve the approach next time. Following that we will create the implementation package and get the TA and impact statement in place so we can do a final vote.

Question: What SCRs would people like to see prioritized for future meetings?
Answer: CCB#21- Provenance LDD. Think we are converging on a solution. Would like people to look at the comments. This is limited to superseded LIDs. An OPS LDD that you can insert in superseded labels.
~ It's an LDD, so not really under the DDWG purview.
~ Yes, but appreciate the node inputs. There were some questions and concerns about searchability. Would like to move forward.

Comments were added to CCB#39 (Fixed problem where LDDTool does not include all rules from IngestLDD) to update the documentation and add an error message. Think we should do it and close that one out.
~ Someone agrees. Think we can go ahead and fix LDDtool to produce a warning.

The Tiger Team for CCB#20 (Instrument Package Context Products) and CCB#32 (Add Host type to Observing System Component) is done working. Waiting on EN. Want those in the queue for the next release.
~ We will need updates to the documentation. Unclear who will make all the updates.
~ Maybe a similar issue for CCB#31 (Ensure that dictionary versions are specified in labels). Think EN doesn't have everything they need.

Action Item - Ron, Matt, Joe and Megan will email to discuss CCB#s 31, 20 and 32.

We may also want to highlight CCB#66 (Improve documentation for List Author, List Contributor, and List Editor)

Someone is surprised CCB#37 (Add List Author, List Editor, and List Contributor to the Modification Detail class) is not closed. The documentation is unclear. Trying to finish up a project, unsure who is a contributor and how we fill that out.
~ Another person agrees and thinks we are probably all doing something different.
~ One person suggests it might be someone in the acknowledgments.
~ Or maybe node personnel. It's unclear.

Question: We had a long discussion on this a while back. There was a Power Point presentation. Maybe by Dick. Does anyone else remember this?
Answer: No, but it would be good to look at what we discussed before.
Action Item - Ron will look for the Power Point.
~ Someone suggested that it might be in another SCR, maybe CCB#15 (Update Constraints on Author/Editor Lists). **(Action Item - Joe)**
~ We can ask Dick.
~ CCB#15 has several Power Points.

The DDWG, with instrument context scrubs has identified a number of instrument types for the LDD. Would like to request they are in soon so we don't have to update them so often.

CCB#57 - Allow Encoded Native in File Area Observational Supplemental

There has been a lot of discussion, and some parallel discussions too. about how these products are served out. Concerned, but unsure how relevant that is to this exact issue.
~ Someone posted something yesterday that was spot on - very clarifying. The quick summary is that this is a simple proposal to take the current encoded native and use that as file area supplemental. Not suggesting we get rid of product native. There are cases where both are useful.
~ Someone disagrees. With this approach it is automatically returned with search. Would have to filter out.
~ Search gets the product observational. It returns the observational product, not the native product.

Question: Other thoughts?
Answer: It's returning the product, so product and encoded native. This will make people download the encoded native file, which in many cases they don't want. We can direct the registry to do what makes sense. We should maintain a firewall between product observational and native. That just makes it clear. We shouldn't duplicate what product native already does.
~ When you search, you get the observational product. One component is the encoded product. Not a huge download, not big products. Think it walls it off better. You won't get a native product without a product observational. You can't get it without the observational product. Won't stumble across it. It could be a philosophical difference. Just an option. Provides flexibility for different use cases.
~ Someone thinks the firewall makes sense. GEO discussed this and thinks it's useful to explicitly subordinate a native to a main product. Philosophically there is an argument to just have one way of doing things, like in Python, although, that's often violated. GEO thinks it should be an option.

~ Yeah, product native is supplemental by it's nature, so disagree this wouldn't be clear. Registry will have instructions to make it clear that product native is always supplemental. Seems like we are re-doing product native. Giving options on this to DPs could cause confusion.
~ Not explicitly supplemental. Not everybody reads the documentation. Useful to mark it as a subordinate.
~ A person says their thinking is that they understand the concern of making it supplemental to the primary product is implementation, but making the relationship clear is a system issue. Don't see it as an IM issue, see it as a system issue.
~ Another person tends to agree. We designed the model with lots of discussion. What we did makes sense in the model. It provides information the system can use. Agree we need to link the products, but think we are too worried about what the system will do. They can make it work once we give them the requirements.
~ Unsure why saying it's not clear it's supplemental. Maybe we need to amplify that in the documentation.

Native encoding proposed to be used in both places is a data format. That could be Ping, text, could be supplemental, could be a number of things. This is using a format in two different ways. One is official and one is to provide subordinate format.
~ The relevance is that in this case the format is non-PDS4 compliant.
~ Ping isn't a compliant format but is allowed in certain cases - just not as an observational product.
~ You are pointing out this can be used for anything.
~ Product native is based on encoded native. Have to go through the same approval process as all product natives.
~ Good. Still want to maintain the firewall.
~ Not arguing against product native, see the need, but want it explicitly marked as supplemental because we will always have users who won't read the documentation. Should be explicit. Philosophical point. Think it makes sense to provide both options.

Seems like there are multiple philosophical points here. For the model and science products. Think would both be useful stand alone with the correct tools. Supplemental metadata provides additional information that is useful for science, but not stand alone. Stand alone might be part of the issue here.
~ From a technical standpoint, would have trouble creating a label for product native. It would be very minimal label without much information. Not really stand alone without all that rich metadata. If attached to the observational product it gets that rich metadata from the observational product.
~ Think it's being provides because the DP believes it provides... (interrupted)
~ Can look at the original JPEG stream and determine things you can't without it. Additional information that a handful of super experts can use. That's why it's supplemental. It does provide additional information, but not something that most users would want.

We shall make final comments for today and vote at the next meeting.

Can see how this is supplemental now. Native format could be used by someone with proper tools and documentation.
~ Yes.

Firstly, appreciate the point that it's not always clear that product native is always supplemental. Might need another SCR to address that. Also thought encoded native would never exist without product native. Unclear why it would only have a minimum label.
~ Technical issue on how the software is built. And yes, encoded native was only created for use in product native, but it's philosophically compatible. It's restricted. We can't help that people don't always read documentation, but we have to acknowledge people sometimes don't.

Time is up. We will vote on this approach in two weeks. Will start the review of the documentation next week.

title: DDWG Notes 2025-07-10
layout: default
date: 2025-07-10

July 10, 2025
Notes by Debra Kazden

Known Attendees: J. Blalock*, R. Deen, L. Huber*, S. Hughes*, R. Joyner, C. Juarez, D. Kahan*, D. Kazden, A. Knight, J. Mafi*, S. McLaughlin, E. Schaefer* and B. Semenov*
(*'s indicate people who voted during this meeting - if applicable)

DDWG Agenda

(Included in email sent July 8, from J.Mafi, Agenda for 2025-07-10)

We will have a DDWG telecon this Thursday, July 10th at 9:30 PDT (12:30 EDT, 10:30 MDT). We will cover the following issues:

1) Action item reports.
2) CCB#28 - As a user, I want validate to allow delimited tables with whitespace-only numeric fields – TA and System Impact statements have been added for this SCR. We will allow time to discuss any concerns identified by these statements followed by another vote if needed.
3) CCB#62 Transfer Manifest class should enforce data types for its two mandatory columns – We will discuss the proposal (jmafi comment dated 8 Jul 2024) to expand this SCR to enforce multiple formatting requirements on Transfer Manifest objects to ensure that they match the format described for them in the DPH.
4) CCB#57 Allow Encoded Native in File Area Observational Supplemental – Please be prepared to discuss and vote on this SCR.

Front Matter

1) Submitted to the CCB:
   - CCB#60 New permissible value for File Area Observational.Header.parsing standard id: "CDF 3.9 ISTP/IACG"
2) New issues in GitHub:
   - CCB#67 - Special characters not allowed in citation information fields
3) Ready for vote:(none)
4) Build 16.0 freeze date: 8/??/2025. The following SCRs are queued for implementation: CCB#59, CCB#60
5) Withdraw by author:
   - CCB#64 Add additional values to investigation as an allowable reference type
   (Brief Discussion)

Action Items

a) PDS4 Versioning Use Cases Survey – All
b) CCB#39 – S. Hughes to update SCR
c) CCB#12 – S. Hughes to verify whether this is subsumed by CCB#52 and CCB#53
d) CCB#20 & CCB#32 – Tiger team formed and working the issue
e) CCB#31 – Ron and Matt working the issue
   (Mentioned)

Tier 1 issues

- CCB#66 Improve documentation for List Author, List Contributor, and List Editor
- CCB#65 Modify description attribute to be required in Telescope and Facility Context Products
- CCB#64 Add additional values to investigation as an allowable reference type
- CCB#63 - International characters / Unicode support for organization name and display full name
- CCB#62 - Transfer Manifest class should enforce data types for its two mandatory columns
- CCB#61 - Disconnect between SR and Validation of ASCII Real
- CCB#58 - New Product Native values for MSL/Mastcam
- CCB#57 - Allow Encoded Native in File Area Observational Supplemental
- CCB#56 – Reconsider the PDS4 build cycle cadence
- CCB#53 – Improve cross-namespace validations in sch files
- CCB#52 – Enable cross-namespace validations in sch files
- CCB#51 – Add an optional attribute in the Target Identification class to identify Primary/Secondary targets
- CCB#50 – Clarify Ambiguous Author Information in Product Document
- CCB#46 – Disallow multiple versions of the same LID from occurring in a single collection
- CCB#44 – Issues caused by the failure to expose pds: attributes that are referenced in an LDD.
- CCB#41 – Create a way for classes or attributes to be exposed to other LDDs but not to data providers
- CCB#39 – Fixed problem where LDDTool does not include all rules from IngestLDD
- CCB#37 – Add List Author, List Editor, and List Contributor to the Modification Detail class
- CCB#36 – Add Mission Phase to the Investigation class
- CCB#35 – Add Alias Index and Acronym Index to the Investigation class
- CCB#32 – Add Host type to Observing System Component
- CCB#31 – Ensure that dictionary versions are specified in labels
- CCB#28 – As a user, I want validate to allow delimited tables with whitespace-only numeric fields
- CCB#23 (JIRA-CCB-211) – XML/schema based files as archival data
- CCB#22 (JIRA-CCB-164) – Display Settings not required for images
- 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#21 (JIRA-CCB-367) – Provenance LDD Using PROV-DM
- CCB#20 (JIRA-CCB-358) – Instrument Package Context Products
- Tiger Team: M. Tiscareno (lead), B. Deen, B. Hirsch, S. Hughes, R. Joyner, J. Mafi, E. Schaefer.
- CCB#12 – Document Nuances for Reusability of Generic LDD Classes/Attributes
- CCB#5 – Move to V.2.0.0 for June 2025 Release
- CCB#4 – Backwards Incompatibility policies
- Presentation: Backwards Compatibility
- CCB#3 – Adopt Semantic Versioning with version 2 transition
- The CCB returned this, requesting that it be resubmitted as part specific proposal to transition to IM V2.0.0.

(Some Discussions)

Tier 2 issues

1) 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!

1) JIRA-CCB-326 - Inconsistent Direction on How to Use Discipline Facets
2) RS: 259/281 (see new attachment to JIRA-CCB-281) (on hold)
3) JIRA-CCB-364 – Disallow bundles referencing collections by LID **(Not Discussed)**

Historical links:
- Export of history: https://github.com/NASA-PDS/PDS4-CCB/tree/main/data
- Backup: https://nasa-pds.atlassian.net/jira/software/c/projects/CCB/issues (limited seats)

Next meeting: 24 July 2025

DDWG

Front Matter

No one here from RMS today.

New SCR - CCB#67 - Special characters not allowed in citation information fields.
~ This is related to CCB#63 (International characters / Unicode support for organization name and display full name). We need to discuss these soon. We need to support international characters. Need to figure out how to navigate this issue. We will discuss it at a future point.

Question: Any comments?
Answer: Someone doesn't know much about this, but non-Roman characters are a concern, but probably not for CCB#63. We already allow the type that lets these in. Chinese characters are a concern, but we have support. The question is if we want to support Chinese characters.
~ The issue is search.
~ Another person thought we discussed an additional attribute for this and then having English across the whole set of data.
~ Yes. We will discuss this at a future meeting. We might need another SCR.
~ We should vote it down if we are going to discuss it at a future meeting.

Regarding action item reports, someone had a comment to make about CCB#20, but not here yet. We can come back to that.

Tier 1 Issues

CCB#28 - As a user, I want validate to allow delimited tables with whitespace-only numeric fields

In SR, to clarify use of space padded CSV fields for validate. Updated the SR and system impact statement and TA done. Need to make sure that hasn't changed votes. A question in the TA. Initially it was a question about if this would impact the DSV standard.
~ Someone thinks they said it might.
~ Yes, true, as it applies to PDS4. Edits made in SR.
~ Someone said their concern is that they remember at one time having a DSV document, separate from the SR. Wondered if those need to stay in sync.
~ This is a separate document (named a document - 4180?)
~ Not that. The standard written by PPI.
~ Not referenced here.
~ The question is if the PPI document is still in force and would need to be updated.
~ We would need to research that. It's separate from this SCR. Could be an action item for Joe. He will make a note to do some research and report on it at a future meeting. **(Action Item - Joe)**

Question: Can the lien be removed or does the action item need to be completed first?
Answer: EN is okay with removing the lien.
~ EN does not have any additional concerns.

Lien removed in real-time.

Question: Given this TA, system impact statement and the removed lien, any concerns with voting to approve this? Anyone want to discuss this further? Would like to do a consent vote if not.
Answer: (Silence)

The Vote to Pass CCB#28 will be recorded as it was last time - using the results from the last meeting.

CCB#62 - Transfer Manifest class should enforce data types for its two mandatory columns

Joe tried to contact Mark about this. Mark submitted the SCR. Hasn't heard back. Went ahead and added a table to this SCR. Suggestion is that we add similar requirements to transfer manifest class as inventory class, as described in the DPH. Transfer manifest files, the DPH suggests, they don't need a label and aren't official PDS4 products, but there is a transfer manifest class in the IM, and NSSDCA and others use them. Format is important. Unsure what is done for LIDs.

Someone wants to reiterate that NSSDCA doesn't use the transfer manifests, but the deep archive tool does bounce from it. They do crack open the inventory files. Changes to that will affect the NSSDCA. Probably others too.

~ Another person wants to mention a few things. The information component is OPS, SIP, etc. All OPS. At one point we designed product deep archive to go to NSSDCA, it is labeled, but not sure it is being used. We have DIP distribution package and SIP submission package. NSSDCA uses SIP. The way the SCR is written, this is a change to the IM. Not sure that's true. Believe this change would be a change to schematron rules to validate the inventory.
~ Yes. Rules would be enforced by schematron. Mark specifically asked for the data types. The question is if we want to do something similar to what we do for inventory files.

Question: Why is this being proposed? Nodes are directly involved with inventories. Transfer manifests are OPS. Unclear why this is being proposed, if there has been a problem.
Answer: Addressed in the SCR. IPDA uses transfer manifests. Variability causes some issues. No rules. Unsure if rules will fix that.
~ We should get the system folks involved. Really more of an OPS thing.
Another Question: Anyone else use the transfer manifest? Besides EN and IPDA?
Answer: (Silence)
~ So maybe that's the extent of the use. Was created for PDS4. Not required, but if there are problems, it's valid to follow up.

Question: Any objections to expanding the SCR to include format?
Answer: (Silence)

Action Item - Joe will continue to try to make contact with Mark.

We will try to discuss the full proposal at a future meeting.

Question: Any last comments?
Answer: (Silence)

CCB#57 - Allow Encoded Native in File Area Observational Supplemental

Question: There has been some discussion on this. We have had a few weeks to think about it. Any last discussion before we vote on moving forward?
Answer: One of the dissenting people is not here.
(Joe tried to ping him in real-time.)

Someone is still unsure. Can see the logic, but it goes against the original idea of what product native was supposed to be.
~ Yes, but not sure that's a problem.
~ That is the right question. It creates two ways of doing the same thing. There are pros and cons.

Question: Comments?
Answer: Worried, if from modeling perspective, if we are going from one approach to the other. That would affect the system people. That's a concern.
~ Adds a new component to the IM. The question is what you put in that slot. Less of a con.
~ Everything under the hood uses the same encoded native.
~ Confused about what comes up for search. Now the system has to worry about that. It doubles the amount of work.
~ It can be treated the same regardless of which one it is in.

~ The download question is going to persist if this is accepted or not. This SCR is not creating extra work from the download perspective or for search. We can decide if we want it marked separate. Don't see potential extra work.

Internally, what is proposed here is a supplemental product. It shouldn't matter that it is native. Don't think download is much of a concern.
~ The concern is the logic. If have an observational product that could lead to a native product, now would have to look to see if there's an associated supplemental product. Two ways to look at this now. It will impact system people. Different ways things are associated.
~ A JPEG file can be a browse file or a document or an associated file now. So, looking at native encoding as a file format, like JPEG, which can be used three different ways.
~ Hearing EN say they want a link to the native data. Don't think we need to abstract it out at that level.

It might be useful to get a clarification of impacts from the system people. See if this really would cause a big impact and a lot of work.
~ If we decide on one approach and then allow two, that's additional complexity. Concerned about complexity for search. Think we need to get input from the system folks. Jordan might say it's no problem.
~ This proposal is one file. Product and supplemental product. Nothing too link to.
~ Concern is abstract associations in the model being implemented in different ways.
~ Someone sees it a different way. If included as file area supplemental, it's part of the product, not just associated. Think the Malin folks are arguing this is part of the product.
~ Yes. That's one of the reasons for this. The history here, Mike Malin argues that if you have the original data stream you can infer things. One of the main reasons that's what they delivered for PDS3. That doesn't fly here, but it provides information people can use. It's a little different from seismic data because that's in a format no one is ever going to use. This is different because people will use the observational product, except in rare cases.

Someone's opinion still stands. If we can only allow file supplemental where additional useful information then I am okay with that. We implement internal references to get associations, so adding complexity. Need to hear from dissenter.
~ Someone thinks a lot of his concern was philosophical and that it violates original intent.
~ The fact that it would be supplemental says it's additional useful information.
~ Still worried about system and the download.

Question: A question to help with this and CCB#58 (New Product Native values for MSL/Mastcam)… the ODR product, is the conversion to EDR non-reversible? Can't get back? Decode?
Answer: Yes. Once it is decoded you can't get back. Lose potential additional information about the image quality.

Question: Are there any last comments? Think we need to hold off on the final vote.
Answer: We need a TA.
~ Yes, and an impact statement.
~ Someone said that Jordan did comment in GitHub. Download isn't part of this yet. Kind of punted. How this works will be set up with future discussions. Think EN will be able to work with anything reasonable. Not so much value from hearing from them.

Question: Last comment for today?
Answer: (Silence)

We will meet again in two weeks.

title: DDWG Notes 2025-07-24
layout: default
date: 2025-07-24

July 24, 2025
Notes by Debra Kazden

Known Attendees: A. Bailey*, J. Blalock*, R. Deen, S. Hughes*, R. Joyner, C. Juarez, D. Kahan*, D. Kazden, A. Knight, P. Lawton, J. Mafi*, S. McLaughlin, E. Schaefer*, M. Seritan* and M. Tiscareno* (*'s indicate people who voted during this meeting - if applicable)

DDWG Agenda

(Included in email sent July 22, from J.Mafi, Agenda for 2025-07-24)

We will have a DDWG telecon this Thursday, July 24th at 9:30 PDT (12:30 EDT, 10:30 MDT). We will cover the following issues:

1) Action item reports.
2) CCB#57 Allow Encoded Native in File Area Observational Supplemental – Please be prepared to discuss and vote on this SCR.
3) CCB#62 Transfer Manifest class should enforce data types for its two mandatory columns – Please review the latest comments on the SCR, including Jordan's response (11 Jul 2025) to questions coming out of the last DDWG telecon, and an additional document (posted 21 Jul 2025) showing the proposed additional requirements.
4) CCB#68 ASCII File Name and ASCII Directory Path Name are using malformed regex expressions – This is hopefully a simple bug fix. The attachments include a sample implementation and test labels. Please review this SCR and be ready to vote.
5) DDWG Open Session Topics – Please think of potential topics for the DDWG Open Session at the PDSMC F2F (5 Aug 2025). We will discuss/brainstorm potential topics for discussion.
6) Next meeting – The next regular DDWG telecon is cancelled since it falls on 7 Aug 2025, the last day of the PDSMC F2F. However, since the freeze date is approaching, we will try to identify issues that we would like to have included in the next build and decide if we need to meet more frequently in the coming weeks in order to try to complete them. Please be prepared to identify any SCRs that you would like to see included, and to your availability to support weekly meetings.

Front Matter

1) Submitted to the CCB:
   - CCB#28 - As a user, I want validate to allow delimited tables with whitespace-only numeric fields
2) New issues in GitHub:
   - CCB#68 - ASCII File Name and ASCII Directory Path Name are using malformed regex expressions
3) Ready for vote:(none)
4) Build 16.0 freeze date: Aug 2025. The following SCRs are queued for implementation: CCB#59, CCB#60
5) Withdraw by author:
   - CCB #64 Add additional values to investigation as an allowable reference type **(Discussed)**

Action Items

a) PDS4 Versioning Use Cases Survey – All
b) CCB#39 – S. Hughes to update SCR
c) CCB#12 – S. Hughes to verify whether this is subsumed by CCB#52 and CCB#53
d) CCB#20 & CCB#32 – Tiger team formed and working the issue
e) CCB#31 – Ron and Matt working the issue **(Not Discussed)**

Tier 1 issues

- CCB#67 Special characters not allowed in citation information fields
- CCB#66 Improve documentation for List Author, List Contributor, and List Editor
- CCB#65 Modify description attribute to be required in Telescope and Facility Context Products
- CCB#63 - International characters / Unicode support for organization name and display full name
- CCB#62 - Transfer Manifest class should enforce data types for its two mandatory columns
- CCB#61 - Disconnect between SR and Validation of ASCII Real
- CCB#58 - New Product Native values for MSL/Mastcam
- CCB#57 - Allow Encoded Native in File Area Observational Supplemental
- CCB#56 – Reconsider the PDS4 build cycle cadence
- CCB#53 – Improve cross-namespace validations in sch files
- CCB#52 – Enable cross-namespace validations in sch files
- CCB#51 – Add an optional attribute in the Target Identification class to identify Primary/Secondary targets
- CCB#50 – Clarify Ambiguous Author Information in Product Document
- CCB#46 – Disallow multiple versions of the same LID from occurring in a single collection
- CCB#44 – Issues caused by the failure to expose pds: attributes that are referenced in an LDD.
- CCB#41 – Create a way for classes or attributes to be exposed to other LDDs but not to data providers
- CCB#39 – Fixed problem where LDDTool does not include all rules from IngestLDD
- CCB#37 – Add List Author, List Editor, and List Contributor to the Modification Detail class
- CCB#36 – Add Mission Phase to the Investigation class
- CCB#35 – Add Alias Index and Acronym Index to the Investigation class
- CCB#32 – Add Host type to Observing System Component
- CCB#31 – Ensure that dictionary versions are specified in labels
- CCB#28 – As a user, I want validate to allow delimited tables with whitespace-only numeric fields
- CCB#23 (JIRA-CCB-211) – XML/schema based files as archival data
- CCB#22 (JIRA-CCB-164) – Display Settings not required for images
- 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#21 (JIRA-CCB-367) – Provenance LDD Using PROV-DM
- CCB#20 (JIRA-CCB-358) – Instrument Package Context Products

- Tiger Team: M. Tiscareno (lead), B. Deen, B. Hirsch, S. Hughes, R. Joyner, J. Mafi, E. Schaefer.
- CCB#12 – Document Nuances for Reusability of Generic LDD Classes/Attributes
- CCB#5 – Move to V.2.0.0 for June 2025 Release
- CCB#4 – Backwards Incompatibility policies
- Presentation: Backwards Compatibility
- CCB#3 – Adopt Semantic Versioning with version 2 transition
- The CCB returned this, requesting that it be resubmitted as part specific proposal to transition to IM V2.0.0.
  (Some Discussions)

Tier 2 issues

1) 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!

1) JIRA-CCB-326 - Inconsistent Direction on How to Use Discipline Facets
2) RS: 259/281 (see new attachment to JIRA-CCB-281) (on hold)
3) JIRA-CCB-364 – Disallow bundles referencing collections by LID **(Not Discussed)**

Historical links:
- Export of history: https://github.com/NASA-PDS/PDS4-CCB/tree/main/data
- Backup: https://nasa-pds.atlassian.net/jira/software/c/projects/CCB/issues (limited seats)

Next meeting: TBD (July 31, 2025)

DDWG

Front Matter

We are missing a few node representatives today.

No news on CCB#28 (As a user, I want validate to allow delimited tables with whitespace-only numeric fields) yet. It was sent to the CCB.

New SCR: CCB#68 - ASCII File Name and ASCII Directory Path Name are using malformed regex expressions.
~ Hoping to discuss this today. Hopefully it is straight forward and we can vote today.

Tier 1 Issues

CCB#57 - Allow Encoded Native in File Area Observational Supplemental

Thought this was ready for a vote last week, but we wanted a representative from RMS present since they are vocal opponents of this.

Question: Any last comments? There has already been a lot of discussion on this. Is there anything new to say? If not, we will vote to approve this as written. It needs a TA and a system impact statement before the final vote.
Answer: Speaker is sorry to have missed last week and is actually on vacation now. Continues to think we should keep native data out of product observational. Native has a name. It is not designated as supplemental.
~ To respond, the summary is that there are two different use cases for product native. In the seismic case it's the file you actually want users to use, and in this case it's not.
~ Binary non-archival data should never be archived if it's not what you want people to use. Can direct them to another location in the website.
~ A person is not sure they agree or disagree that they should be kept separate.
~ Another way product native could be handled is as product external, but want this as part of the archive. In seismometer case it's what the community expects to use, so that is different.

Question: Can we vote in reverse order?
Answer and Another Question: Yes. Too close to the top?
Answer: Always.

The Vote to approve CCB#57 as written. It still needs a TA, system impact statement and a final vote.
RS - Yes
RMS - No
PPI - No
NAIF - Abstain
IPDA - Not Here
SBN - Not Here
IMG - Yes
GEO - Yes
EN - Abstain
ATMOS - Not Here

That's 3 yes, 2 no, 2 abstain and 3 not here. Will have to check the rules. We have votes from the majority of the group, so probably enough to send this forward. Will see if any of the no votes want to vote. **(Action Item - Joe)**

CCB#62 - Transfer Manifest class should enforce data types for its two mandatory columns

A table was posted that shows what changes could be made, then an additional document to clarify what is being proposed to add to make it more like inventory.

Question: Any comments?
Answer: (Silence)

Would like to vote on going ahead to update the SCR to include what's proposed, to make it more like inventory and regulate the format. A yes vote would add the yellow items as required values in a

transfer manifest object. No would stick with the green (see these in GitHub). The question is if we want to add these fields.
~ Not all the nodes are here.

Question: Transfer manifest is not part of the archive, right?
Answer: Right. Jordan says they do use the labels, but the DPH says it's not part of the archive.
~ Don't think we should be doing this. There's no real requirement for the labels.
~ Someone wants to know if there is a reason not everything is required.
~ Max field length value is easy to figure out, it's not delimited. Want flexibility. Same for file specification.
~ Unclear about field location.
~ Probably should be fixed. Should be part of the proposal.

NSSDCA wants to note that transfer manifest is used by the Deep Archive Tool to generate the SIP, product SIP deep archive. Think there's a restriction on LIDVID field to 256 and it triggers things. Think there are restrictions.
~ Yes, but someone else thinks there's a limit on data type. Could be additional limits, but since it's fixed width, don't necessarily want to force people to use 256 byte field if they don't need to.

Question: Any additional comments?
Answer: (Silence)

We will continue with the vote to add these values. This was proposed by IPDA. They use these and wanted to make sure the format matched what is in the DPH, even though they are not archive products. We will need to describe this in the SR.

The Vote to update CCB#62:
EN - No
GEO - Yes
IMG - Yes
IPDA - Not Here
NAIF - Abstain
PPI - Yes
RS - Yes
RMS - Yes
SBN - Not Here

That's 5 yes, 1 no. We will discuss and vote on this again next time.

CCB#68 - ASCII File Name and ASCII Directory Path Name are using malformed regex expressions

This is a new SCR. It is unusual to discuss and vote so quickly, but think this is just a bug fix. One of the tools caught this. The update would be to the common schema.

Question: Any questions or comments?
Answer: Someone using a tool besides Oxygen or validate found this.

~ Someone else did some testing with Oxygen and didn't find any issues.
~ Don't think it's possible for there to be a backwards non-compatibility issue.
~ Agreement.

Question: Other issues?
Answer: (Silence)

Would like to vote to move forward.

The Vote to pass CCB#68:
SBN - Not Here
RS - Yes
RMS - Yes
PPI - Yes
NAIF - Yes
IPDA - Not Here
IMG - Yes
GEO - Yes
EN - Yes
ATMOS (Voted Yes by email to Joe before the meeting)

This does have a TA and a system impact statement. Will go to the CCB.

Meeting Planning

Our next meeting is supposed to be August 7, but that is the last day of the F2F. The August 7 DDWG meeting is cancelled because several members of the DDWG will be at the F2F. The concern is the next freeze date, which is some time in August. Unknown what the specific date is, but if there are SCRs we want to get into the next build we might want to move to weekly meetings. Know we want to see CCB#20 move forward.

Question: Are there other SCRs to work on before the freeze?
Answer: Think we should meet more often in general.
~ Yes, but it's hard to get other work done.
~ Someone would like to see CCB#53 (Improve cross-namespace validations in sch files) move forward. There is a lot of technical work that has to happen on the EN side. Would like to get it moving in the DDWG.

Someone else suggested CCB#58 (New Product Native values for MSL/Mastcam), which will need DDWG and MC approval.
~ People would be happy with that.

People can let Joe know if there are others. **(Action Item - Everyone)**

Question: Is there any chance of the MC discussing CCB#58 in two weeks?
Answer: No. The DDWG will need more time to discuss it.
~ And the (documentation) peer review is ongoing.

~ The due date for feedback is tomorrow. Gotten some. Hope to wrap that up soon so we can move forward and vote on this.

Think we should meet more in the next month.

Question: Any conflicts with July 31?
Answer: (Silence)
Question Continued: August 7, no meeting. August 14, 21 and 28 - any problems?
Answer: Not everyone can make August 21.

Action Item - Joe will email a schedule to everyone. Joe has travel in September, probably a break for a few weeks after that.

DDWG Open Session at MC F2F

The last topic is that we have two items for the open session. A CTLI discussion added by Lyle and a Provenance LDD discussion added by Steve.

Question to Steve: Can we remove the Provenance LDD discussion?
Answer from Steve: Would like input.
~ Someone doesn't know that it needs to be part of the discussion.

Question: Are there other topics? We have time available.
Answer: The one that was brought up, CCB#53 (Improve cross-namespace validations in sch files), about LDD ingest, a proposal to make changes to LDD ingest language to support validation. Maybe would benefit from a more general discussion.
~ Good point.
Another Question: Do people think it would be useful to discuss that?
Answer: Yes. Excellent idea.

Action Item - Bob and Steve will have slides ready for a discussion on CCB#53 for approximately 10 to 15 minutes.

If there are any additional topics people can post them to the MC shared drive for the August F2F meeting or email Joe to add them.

title: DDWG Notes 2025-07-31
layout: default
date: 2025-07-31

July 31, 2025
Notes by Debra Kazden

Known Attendees: A. Bailey*, J. Blalock*, R. Deen, B. Hirsch*, L. Huber*. S. Hughes*, C. Juarez, D. Kahan*, D. Kazden, A. Knight, P. Lawton, J. Mafi*, S. McLaughlin, E. Schaefer*, B. Semenov*, M. Seritan* and M. Tiscareno*
(*'s indicate people who voted during this meeting - if applicable)

DDWG Agenda

(Included in email sent July 29, from J.Mafi, Agenda for 2025-07-31)

We will have a DDWG telecon this Thursday, July 31st at 9:30 PDT (12:30 EDT, 10:30 MDT). We will cover the following issues:

1) Action item reports.
2) CCB#57 Allow Encoded Native in File Area Observational Supplemental – Please be prepared to discuss and vote on this SCR.
3) CCB#62 Transfer Manifest class should enforce data types for its two mandatory columns – We will discuss and vote on the proposed implementation for this SCR.
4) Review Open SCRs – We will review open SCRs to determine current status, next steps, and priorities.

Front Matter

1) Submitted to the CCB:
   - CCB#28 As a user, I want validate to allow delimited tables with whitespace-only numeric fields
   - CCB#68 ASCII File Name and ASCII Directory Path Name are using malformed regex expressions
2) New issues in GitHub: (none)
3) Ready for a vote: (none)
4) Build 16.0 freeze date: Aug 2025. The following SCRs are queued for implementation: CCB#59, CCB#60
5) Withdraw by author:
   - CCB#41 Create a way for classes or attributes to be exposed to other LDDs but not to data providers

Action Items

a) PDS4 Versioning Use Cases Survey – All
b) CCB#39 – S. Hughes to update SCR
c) CCB#12 – S. Hughes to verify whether this is subsumed by CCB#52 and CCB#53
d) CCB#20 & CCB#32 – Tiger team formed and working the issue
e) CCB#31 – Ron and Matt working the issue **(Some Discussion)**

Tier 1 issues

- CCB#67 Special characters not allowed in citation information fields
- CCB#66 Improve documentation for List Author, List Contributor, and List Editor
- CCB#65 Modify description attribute to be required in Telescope and Facility Context Products
- CCB#63 - International characters / Unicode support for organization name and display full name
- CCB#62 - Transfer Manifest class should enforce data types for its two mandatory columns
- CCB#61 - Disconnect between SR and Validation of ASCII Real
- CCB#58 - New Product Native values for MSL/Mastcam
- CCB#57 - Allow Encoded Native in File Area Observational Supplemental
- CCB#56 – Reconsider the PDS4 build cycle cadence
- CCB#53 – Improve cross-namespace validations in sch files
- CCB#52 – Enable cross-namespace validations in sch files
- CCB#51 – Add an optional attribute in the Target Identification class to identify Primary/Secondary targets
- CCB#50 – Clarify Ambiguous Author Information in Product Document
- CCB#46 – Disallow multiple versions of the same LID from occurring in a single collection
- CCB#44 – Issues caused by the failure to expose pds: attributes that are referenced in an LDD.
- CCB#41 – Create a way for classes or attributes to be exposed to other LDDs but not to data providers
- CCB#39 – Fixed problem where LDDTool does not include all rules from IngestLDD
- CCB#37 – Add List Author, List Editor, and List Contributor to the Modification Detail class
- CCB#36 – Add Mission Phase to the Investigation class
- CCB#35 – Add Alias Index and Acronym Index to the Investigation class
- CCB#32 – Add Host type to Observing System Component
- CCB#31 – Ensure that dictionary versions are specified in labels
- CCB#28 – As a user, I want validate to allow delimited tables with whitespace-only numeric fields
- CCB#23 (JIRA-CCB-211) – XML/schema based files as archival data
- CCB#22 (JIRA-CCB-164) – Display Settings not required for images
- 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#21 (JIRA-CCB-367) – Provenance LDD Using PROV-DM
- CCB#20 (JIRA-CCB-358) – Instrument Package Context Products
- Tiger Team: M. Tiscareno (lead), B. Deen, B. Hirsch, S. Hughes, R. Joyner, J. Mafi, E. Schaefer.
- CCB#12 – Document Nuances for Reusability of Generic LDD Classes/Attributes
- CCB#5 – Move to V.2.0.0 for June 2025 Release
- CCB#4 – Backwards Incompatibility policies
- Presentation: Backwards Compatibility
- CCB#3 – Adopt Semantic Versioning with version 2 transition
- The CCB returned this, requesting that it be resubmitted as part specific proposal to transition to IM V2.0.0.
  (Some Discussions)

Tier 2 issues

1) 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!
1) JIRA-CCB-326 - Inconsistent Direction on How to Use Discipline Facets
2) RS: 259/281 (see new attachment to JIRA-CCB-281) (on hold)
3) JIRA-CCB-364 – Disallow bundles referencing collections by LID **(Not Discussed)**

Historical links:
- Export of history: https://github.com/NASA-PDS/PDS4-CCB/tree/main/data
- Backup: https://nasa-pds.atlassian.net/jira/software/c/projects/CCB/issues (limited seats)

Next meetings:
Jul 31 – Extra Meeting
Aug 7 – Cancelled (PDSMC F2F)
Aug 14 – Extra Meeting
Aug 21 – Regular Meeting
Aug 28 – Extra Meeting
Sep 4 – Regular Meeting (Cancelled? IPDA Steering Committee)

DDWG

Front Matter

This is the first of the weekly meetings in August, except for next week during the MC.

Question: Does anyone have an action item report?
Answer: (Silence)
~ A number of the issues are being worked. We will hear progress soon.

CCB#57 - Allow Encoded Native in File Area Observational Supplemental

We voted on this last time. The vote was 3 yes, 2 no, 2 abstain and 3 not voting. It passed with the votes as they were cast. Since then the TA was written up. It seems that and the system impacts were fairly minimal. Unsure if it takes into account product observational with supplemental objects that aren't readable.
~ Expecting some modifications to software, so that should have been included in the implementation. Remember Jordan saying there would be minor changes.

Would like to have a consent vote. If anyone wants to discuss this, we can, or we can go with the vote from last time, or we can discuss it and re-vote.
~ We should re-vote.

Question: Any discussion?
Answer: (Silence)
Another Question: Can someone summarize the pros and cons from last week?

Answer: Unsure if Bob will be here today.
~ We can try to summarize. People can jump in. Pros are that it doesn't undo the ability to archive as product native. Got the sense the way these products are described in CCB#58 that they really are considered by the DPs to be supplemental, so allowing them to be supplemental here accurately describes the relationship. The con is that it creates two ways to do it. There was also concern about the original intent of product native. Another concern was how the system would support it.

Question: Anything to add?
Answer: A small clarification. Yes, believes some data is supplemental, the contrast is the seismic data we already have.That is the distinction. That was a good summary.
~ A very fair summary, but there are concerns.
~ Bob is here now.

Question: A summary was requested. We had a quick recap of pros and cons. Any last comments you want to add?
Answer: We would find it very useful so hope we get a positive vote.

The Vote to approve CCB#57:
ATMOS - Yes
EN - Abstain
GEO - Yes
IMG - Yes
IPDA - Not Here
NAIF - Abstain
PPI - No
RMS - No
RS - Yes
SBN - Yes

That's 5 yes, 2 no. The item passes. It will be sent to the CCB for their review. Appreciate all the work and discussion that we have had on this matter.

(Bob has to leave. Will work on his presentation for the F2F.)

CCB#62 - Transfer Manifest class should enforce data types for its two mandatory columns

The Requested Changes and the Proposed Solution have been updated to include the additional fields. Test package is included. Unsure if people have looked at this yet.
~ There are typos in the Requested Changes and Proposed Solution.
~ People can let Joe know what the typos are and he will fix them **(Action Item - Joe)**
~ "Manifest" in one class level item is spelled wrong.
~ Thank you for catching that.

Question: Are there any other thoughts or discussion on this? Has anyone looked at the test labels or test implementation?

Answer: (Silence)
~ Everyone should look. **(Action Item - Everyone)** We will get a TA and system impact statement and vote in two weeks.

Question: Someone thought a look at the checksum manifest was also suggested?
Answer: Yes, as a separate SCR. That can use this as a guide. Concerned about the label being completely optional and the product not being in the archive. Think we need to add requirements to the SR.
~ These were planned from the beginning, four of them, as OPS products, to do transfers to the deep archive. Moving in the right direction. A note from Jordan says this has use for OPS. It's been a product since the beginning, maybe not as an archival product.
~ Yes, and it was submitted by Mark Bentley because they use them. They need this SCR to ensure uniformity.
~ Again, archive information package is in the archive, when distributed it's the DIP. Submission into archive is the SIP. PSA is more consistent with the OSA model than we are. Jordan would like us to move that way.

Question: Any last comments on this one?
Answer: (Silence)
~ We will try to get the updates in place and vote in two weeks.

Review of Open SCRs

We will review and set priorities for open SCRs and identify any that need F2F discussion.

A new one yesterday, **CCB#69 (HDF5 as an accepted "observational" and "supplemental" format)** from Trent.
~ John thinks this is related to the THEMIS migration. Will discuss it with Trent to better understand it. It is not expected anytime soon. Not critical. Want to do it correctly. **(Action Item - John)**
~ PPI is happy to see this. Have had requests to have data in HDF5. Definitely want to be part of this discussion. Think other nodes will want it too.
~ We have discussed this. Heard the HDF5 system was moving towards PDS4 compliance with certain constraints, from a few years ago.
~ Yes, this is similar to the GeoTiff situation. Need to understand the guardrails and how to make it work and make sense without doing bad things. Will discuss it with Trent.
~ We should look at CDF and CDF/A regarding guardrails.
~ The lesson is that it is possible to make PDS4 compliant CDF even though the specification allows a lot more.

Question: Other comments? Agree we need to do this right and understand the potential pitfalls.
Answer: (Silence)

CCB#3 - Adopt Semantic Versioning with version 2 transition, CCB#4 - Backwards Incompatibility policies and CCB#5 - Move to V.2.0.0 for June 2025 Release These are all related to the transition to version 2.0. Joe has some action items on these. Think that's it. Some time ago we discussed having a work group to identify what we need to move to 2.0.

CCB#12 - Document Nuances for Reusability of Generic LDD Classes/Attributes From Ron.

Question: Any comments on this one? Was it subsumed by another one?
Answer: Possibly. Need the heritage on this.
~ Last comment in GitHub is that it's on the back burner until CCB#53 (Improve cross-namespace validations in sch files) is done.

CCB#20 - Instrument Package Context Products Someone believes the Tiger Team has delivered this to EN for a full TA and impact statement and then it will come back to us for a vote.

~ Someone else sent an email around outlining tasks we still need to do. Not quite ready for EN. Might need another Tiger Team meeting.
~ There was a meeting. It's all worked out. Working on implementation.
~ Someone still needs to do a set of test labels. We also need documentation changes. Not done yet.
~ Another person said they were supposed to produce some sort of context products, but it would be easier to make sure they are accurate with new schema and schematron.
~ Those are done, but still think another meeting is in order. EN is waiting on test labels.

The Tiger Team will discuss this Tuesday at the F2F.
~ Ethan will be attending virtually, Ron will be on vacation.
~ If Ron won't be available it should be discussed another time.

The Tiger Team will schedule a time to discuss this.

CCB#21 - Provenance LDD Using PROV-DM

Question: Does this need to stay open? Think Steve got his feedback, and this is an LDD, so no vote is necessary.
Answer: It can be closed.
~ We can keep it open, but there is nothing more for the DDWG to do.

Action Item - Steve will discuss this with Jordan.

CCB#22 - Display Settings not required for images IMG believes that Trent was working on this, but it wasn't finished. John will talk to Trent so it can be wrapped up. **(Action Item - John) **

CCB#23 - XML/schema based files as archival data Unclear who this is from. Still an open issue. Think it still needs work, but think there is a work around.

~ Someone remembers Jesse doing a lot on this one from the Jira days.
~ Yes, it's from Jesse.
~ We should move this forward. PPI has gotten requests for this.
~ GEO too.
~ A big part of the issue is allowing what is okay to allow and not what is dangerous to allow.

Question: Is there still a use case?
Answer: Someone is trying to remember... might have worked around with different file names. Rusty.
~ Someone else's guess is that they were labeled as string text.
~ We have LBLX, but this is about allowing XML as a data file. Question is who wants to do that and how do we allow it while limiting anything dangerous.
~ Another person wants to know what would be dangerous.
~ Jesse had a presentation. Can get things to execute. Basically hacking.
~ People need to see the presentation from Jesse.
Action Item - Joe will find that presentation.
~ Think it's in the zip file in GitHub.

Ethan said he is willing to help on this one in a Tiger Team.
~ Joe said he will defiantly help on this. It feels like something to resolve before going to 2.0.

CCB#28 - As a user, I want validate to allow delimited tables with whitespace-only numeric fields This was passed by the DDWG and is under consideration by the CCB. They have concerns.

~ Someone heard that. Asked that the concerns be put on GitHub, but they haven't been yet.
~ Another person believes the concerns are from Anne.
Action Item - Joe will talk to her.
~ We should put this on the F2F agenda for Tuesday.
~ Good idea.

CCB#31 - Ensure that dictionary versions are specified in labels Someone discussed this with Ron. Believe Ron has what he needs to make updates to the proposal. Decided that what is currently required is sufficient and we need to fix the documentation. The **Action Item- is on Ron**.

~ Ron agreed.

CCB#32 - Add Host type to Observing System Component

Question: Is this tied to another SCR?
Answer: We should retire this one. Was a different solution to the problem in CCB#20 (Instrument Package Context Products) and this isn't the direction that we are going. But maybe we should think about this.
~ Someone remembers this as possibly connected to CCB#20.

Action Item - Matt will think about this. We can discuss it at the F2F.
~ Okay.

CCB#35 - Add Alias Index and Acronym Index to the Investigation class This has been overcome by events and is being used but not finalized. We can get back to it.

~ The next step is to think about it.
~ Yes. Something is being used. Steve will try to get everyone up to date. **(Action Item)**

CCB#36 - Add Mission Phase to the Investigation class There has been some discussion, but not in awhile.

~ The status is the same as the previous one. A few SCRs were added at once from Jordan. This and CCB#35 (above), and **CCB#37 (Add List Author, List Editor, and List Contributor to the Modification Detail class)**

CCB#39 – Fixed problem where LDDTool does not include all rules from IngestLDD

Question: Status?
Answer: This is from Jesse. It was resolved. Rules are in separate contexts. System prefers combined. Jesse understands. We should close this.
~ It fails silently.
~ Yes. Think should add a check to LDDtool.
~ In the final suggestions. Two suggestions.
~ Both would be useful additions. Need to keep this open for the DPH.

Action Item - Steve will write a ticket for LDDtool and leave this open for the DPH.

Question: Is there someone who can write the DPH update?

Answer: **Action Item - Ron** will take a look at it.

CCB#44 - Issues caused by the failure to expose pds: attributes that are referenced in an LDD

Question: Status?
Answer: RMS was using kernel type and it turned out it hadn't been exposed in common, so it wasn't correctly parsing for RMS to use it. Asked why this is an SCR in the comments. There was some discussion and the matter dropped.
~ Someone thinks CCB#52 (Enable cross-namespace validations in sch files) and CCB#53 (Improve cross-namespace validations in sch files) are somewhat related.
~ Issue number 807(?) links to this. Unsure if that is fixed or still an issue.
~ Someone is trying to remember. The question is when to use xs:reference. Should use it anytime we expose an element, anything that's going to be referenced. If we exposed everything people would see differences. Changes wouldn't affect operations.
~ A person would need to see more about what that means.
~ Yes, it has been an issue for years.

Question: Can Steve put something together to show the impact of this?
Answer from Steve: Yes, but not high priority. **(Action Item- Steve)**
~ It will be nice to resolve.

Our time is up, but would like to continue with the list. Thanks for attending. If you need to drop off, go ahead. We will meet on August 14.
~ There's a lot of work to continue with this list. Can't stay. Maybe we should continue with this on Tuesday morning. Think we should review CCB#32.
~ Okay.

Homework - Everyone look at any SCRs you are interested in or associated with to see the next steps and priority.

title: DDWG Notes 2025-08-05
layout: default
date: 2025-08-05

August 5, 2025
Notes by Debra Kazden

Known Attendees: C. Albers, A. Bailey, D. Hollibaugh Baker, J. Bauer, J. Blalock, A. Gicquel-Brodtke, K. Crombie, B. Deen, M. Drum, R. Fergason, L. Givens, K. Hansen, T. Hare, M. Hartke, L. Huber, S. Hughes, S. Joy, D. Kazden, C. Kingston, A. Knight, L. Kolokolova, P. Lawton, S. Loftin, Y. Ma, J. Mafi, J. Manweiler, T. Morgan, L. Neakrase, J. Padams, L. Paliashchuk, A. Raugh, E. Schaefer, B. Semenov, M. Seritan, T. Stein, J. Stone, V. Tang, M. Tiscareno and R. Walker

DDWG F2F Discussion

Topics to discuss today:
1) CTLI Discussion - Camera/Imager - L. Huber
2) CCB Object to Proposed Revisions to SR 4C.1 (CCB#28) - J. Mafi
3) CCB#53 - Multimission Validation of Common Attributes - R. Deen
4) CCB#69- HDF Introduction - T. Hare
5) Review of Open SCRs - J. Mafi

CTLI Discussion - Camera/Image

Discussion led by Lyle Huber

The history is that the first version of PDS4 wasn't perfect. Very shocking. Not populating the instrument types fully was a big mistake. Doing it later was unpleasant and a mistake. Instrument types are dynamic values. Needed lots of updates to the IM, so the CTLI LDD was made to handle instrument types and make it easy to update context products. The stewards have questioned the differences between camera and imager. In PDS3 it was camera. In PDS4 first we had camera, then some moved to imager. It's unclear why or what the difference is. The people at IMG node don't know. The distinction might be denoting types - close matches. It's the plan, but not available yet. Search shouldn't care between camera and imager. Camera and imager should both return everything that is camera or imager. We need to get the definitions right, but not invalidate anything already in the archive. That's where we need to go. There has been a lot of discussion in GitHub.

Question: Are there any data products that use this LDD or just context products?
Answer: Type list is in the IM, not in the CTLI namespace. Not allowed in the IM. Only for context products.
~ The issue is search.
~ If we wanted to deprecate a value it would mean fixing the context products, not the data.

Someone said synonymous would be better. Don't care if it's camera or imager. Maybe the original distinction was film versus electronic. Can just pick one.
~ Someone else thinks imager, like push broom, isn't a camera.
~ Synonymous.

~ A person from IMG seconds that an imager is a more generic term. Think it would be very helpful if they were made synonymous.
~ Another person objects to two values that mean the same thing.
~ There are already many of those. Not an issue.
~ EN would have to confirm that.
~ It will be run by Jordan before the next version of the LDD is published.

We need a definition in the description of the terms, so if imager it would say also see camera and vice versa. **(Action Item - Lyle)**

We can have multiple instrument types and can allow multiple instruments for data to be discovered.

CCB Object to Proposed Revisions to SR 4C.1 (CCB#28) Discussion led by Joe Mafi

See Presentation: CCB Objections to Proposed Revisions to SR 4C.1 (CCB#28) (Text under slide headings is conversation had at that time, not slide context.)

Slide: Background
This SCR was to clean up issues with validation. Delimited tables often have space padding. Empty field also sometimes have that.
~ Not empty.
~ Yes, but we need to agree on how validation will work and how to clarify confusion over how validation currently works. In validate tool if there is a field with only spaces that is supposed to be numeric, validate flags it as invalid. There is a statement in the SR about trailing spaces. Might be why. The proposal passed from the DDWG to the CCB, but the word "null" had been used for empty, and it was pointed out that that is not empty.

Slide: CCB Objection Summary
The DSV spec was written years ago, so didn't want to edit too much, but can clean up.
~ Someone has never seen another delimited table with leading and trailing spaces as part of the value, part of the field.
~ Ambiguous. That's what needs cleaning up.
~ Changes it. Same problem is other data fields in fixed width tables with empty string. Can't add quotes.
~ We need clean up in 4C.1 and changes in 4C.2. Just want to clarify, we do need clean up.

Useful exercise would be to make BNF or flow diagrams.
~ The question is if spaces matter between commas. We need to clarify that space isn't blank. Can't say what space is.
~ Need to re-read the SR, unsure about embedded fields.
~ Can say field with only space character is empty, can update to that.
~ Someone likes that solution.

Question: Can't we go to an open CSV standard if there is one?
Answer: Our DSV standard is copy/paste with revisions. Can keep it as is. Don't need to go crazy fixing it.

Another Question: Maintainable long term?
Answer: We won't use Java forever. Once AI can re-write our code.

Action Item - Joe, Anne and Jordan will work on a re-write. Will follow up so description captures it in 4C.2.

Slide: Backup
The back up slides have the original documents and Anne's email.

Slide: Proposed Edits to SR 4C.1

Slide: CCB Objections (original)

CCB#53 - Multimission Validation of Common Attributes Discussion led by Robert Deen

See Presentation: Multimission Validation of Common Attributes (Text under slide headings is conversation had at that time, not slide context.)

Slide: What are we talking about?
This was originally discussed at the November F2F. Discussed multiple mission value of common attributes. Discipline LDDs are supposed to encapsulate common terms to be reusable. Mission LDD is more specific. Need to restrict for Mission if in Discipline LDD.

Two SCRs are related to this discussion. CCB#52 (Enable cross-namespace validations in sch files) thought a modification was needed to make it work, but it turns out you can already do it now, so it's a documentation fix. CCB#53 (Improve cross-namespace validations in sch files) proposes a better way to do it.

Slide: Use cases
There are several use cases.

Slide: What does this look like in practice? Mission phase name is a favorite example. No common definition across missions, but it is used in all missions. If you don't have a Mission LDD in the (XML label) prologue you never see the validations.

Slide: Example
Mission phase name is Cruise.

Slide: DDWG #53 Proposal
New class to Ingest LDD - restriction. Three supplemental classes. Definitions are in the backup slides in the shared drive.

Slide: Examples of DD_Restriction

Slide: Case 1: Restricting a valid value list using DD_Restriction Have to tell it the namespace value domain to list permissible values.

Slide: Case 2: Restricting a valid range using DD_Restriction Same. Threw in a nuance to show exactly how it applies to that mission. Doesn't change the basic definition.

Slide: Case 3: Requiring optional attribute using DD_Restriction Product type name is used. Changes the minimum cardinality to 1.

Slide: Case 4: Restricting cardinality using DD_Restriction Also restricts cardinality. Has a minimum and maximum.

Slide: Case 5: Restricting format via regex using DD_Restriction Validate that it only matches that format, conditional tells you when it applies.

Question: How is this implemented?
Answer: It's in the slides.

Question: Any additional questions or comments?
Answer: Someone sees very little that can't already be done by the Mission LDD. Don't understand why against that LDD. Thinks this will cause validate CPU problems to validate all these restrictions.
~ That is another person's concern too. Been digging, trying to find slides for this.
~ DDWG Ingest LDD in the shared drive.

This is a system and software engineering problem. Shouldn't affect the end user. Mission LDD is fine. This is improvements for the system or developers. Users don't need to know.
~ The point is Mission phase name has no value mission to mission, so it belongs in a Mission LDD, not above it.
~ At EN, would need to create a tool that knows all the mission phase names for all the missions for search. On the tool side, need to know all the nuances.
~ Better to use the Mission LDD. It has the right information for that specific mission.
~ Mission phase name is not a realistic search across PDS. When it's a facet for search the user has already made some decisions about what to search on. A typical search is for mission or a set of missions, then users might look for phases.

Question: If a tool for everyone, how does EN know about the mission phases?
Answer: Trying to solve the problem with abstracts.
Another Question: Why would EN need to do that? OPUS does it with attributes defined for different missions. Made available for cross search, but not for mission phase name. It doesn't do that. Uses mission specific facets.
~ Someone is interested in what is necessary to maintain OPUS. The concern is for when all the missions are included.
~ OPUS is significant to maintain. It will never be a substitute for discipline experts looking at data and figuring out what goes together.

Question: Why not have a keyword that is restricted for a specific mission?
Another Question: Are all mission phase names in namespace?
Answer: Yes.
~ We know the context object for mission can be restricted to be in common namespace so we can validate with schematron. Requires a re-write.

~ Someone wants to echo that. A lot of the use cases look like standard schematron stuff. Curious about what the benefit is.
~ The point is that it becomes schematron. CCB#52 talks about how to do it now. Not very user friendly. This proposal makes it easier to use.
~ Helpful maybe with driving new DPs to make sure they are reusing attributes in a common way. Having something that helps new developers would be helpful.

Think we got off topic. Can use the context products. Think this presentation is trying to show us something we might want to use from Discipline LDDs across missions.
~ The example - each mission is defined. This would allow us to pull common keywords with different value sets out of Discipline LDD and put into Mission LDD.
~ Use Case 4, min and max values can go in the Mission LDD. Pushing it up a level doesn't buy you anything.

Question: Then why do we have Discipline LDDs?
Answer: We don't have time for that discussion.
Another Question: Why do we have Mission LDDs?
Answer: Lots of history.
~ Abstract versus development flag. Both being true. This is exactly the same thing. This does it.

Someone has been looking at this too. Standard value lists. There are real reasons to restrict things to do validate without duplications. Been reading schematron and XSD manuals trying to understand. Think we don't want to change namespaces. Believe we would need to re-work logic for ingest logic. What we do now is not long term sustainable.
~ We did this seven to ten years ago.
~ We have been discussing it since the original design.
~ Proper approach to object orientated modeling.
~ Disagree. Don't want multiple pass schematron or to change namespace.
~ This is just DD Pile.
~ Multi-conditions after require multi pass schematron. There are things we should avoid. Working on a solution so we are not wasting CPU cycles when we run validate.
~ This is very easy to implement and it is optional.
~ Trying to find a good solution for every situation. Long term. It won't change cloud costs.
~ Maybe this should be discussed off-line.

This seems like the old age philosophical debate between lumpers and splitters. It has been put to me that sometimes I am a lumper and sometimes I am a splitter.
~ We all are.
~ The request here is to lump, but a lot of us are splitters. The concern is cost versus benefit. Should look out for things with commonalities that we might want to search on.
~ Not for or against this, but the solution has to be implemented by DPs in a reasonable way. We need to try to make the solutions reasonably implementable.
~ The solution needs to work for all the nodes and IPDA.

An example is the Imaging LDD. Cameras vary a lot, but a lot of common ways to express things. The Imaging LDD has almost no validations because it's so open ended. With this we could. Philosophy aside, it's something that we always intended, but extremely cumbersome. This makes it easy. Need to decide where to apply it.

~ Someone agrees. Might be useful for some things. Seems like it add complexity, but fine.
~ Someone else thought schematron, but if that's too difficult, then okay.
~ This is schematron rules in Ingest LDD.

CCB#69- HDF Introduction

Discussion led by Trent Hare

Trying to allow as a product observational and supplemental, like CDF. It's a self describing format. PDS4 would only have raw. Like FITS, etc., it's an open standard. It's used by NASA, recommended by the Library of Congress, and an ISO standard. Already in PDS4 for LUCY and ORex. This is to formalize it like CDF or FITS. It could be controversial. Documentation is on GitHub with example code and some interesting scripts that can write a PDS4 array or binary table. Need this for Europa Clipper.

Question: Why not have this as a product observational?
Answer: It would be either observational or supplemental.

It's a software defined format, like CDF, FITS and others. Can be archived in PDS4 compliant files, but not all FITS files are compliant.
~ This would be a PDS4 compliant HDF5. Non-tiled, no compression.
~ Have to hammer it home that it's PDS4 compliant.
~ Scripts are included to make it PDS4 compliant. DPs have to know it has to be PDS4 compliant.
~ Reviewers have to know it has to be the PDS4 compliant versions of those softwares.

Someone remembers when HDF5 was first considered. We had been promised a PDS4 compliant HDF5 with constraints.
~ Yes. It will be simple to convert. It will be easy for users. Understand the reviewer concern about reviewers reviewing the format and not the PDS4 compliant version of the format. Just like GeoTiff, we have to be careful about what is allowed.
~ We will need to document.
~ It will be easy to tell that it is untiled and uncompressed.
~ Can see it with the PDS4View tool which uses the PDS4 label.
~ Think we need to be careful with DPs. Often find metadata in ODL that doesn't make it into the new label.
~ Can't hide the metadata. It has to be in the PDS4 label. It's allowing users and scientists to use the format they are used to using, with tools that support them and they have a PDS4 label. Allowing formats people use to be in the archive in a PDS4 compliant way. People can make comments in GitHub.

In Earth science, HDF5 allowed it to be easier to make readily available analysis ready data sets. That will be good across SMD for readiness. Forward thinking.
~ This gets you 80% of the way there with the scripts. Gets the part you want. Cloud optimal.

Review of Open SCRs

Discussion led by Joe Mafi

See Presentation: PDS4 CCB Open SCRs

We will skip through a lot of the slides for this discussion.

CCB#20 (Need a Context object for Instrument Packages) and CCB#32 (Add Host type to Observing System Component)
CCB#32 is now it's own thing. No longer tied to CCB#20.

CCB#21 - Provenance LDD Using PROV-DM This will be worked off-line. LDD issue.

CCB#22 - Display Settings not required for images John and Trent will work on the next steps.
~ This might be the oldest one in the system.

CCB#23 - XML/schema based files as archival data This has been around awhile. The original reason has been worked around. Need to fix this before version 2.0. Ethan is working on it. Anne has use cases.

CCB#28 - As a user, I want validate to allow delimited tables with whitespace-only numeric fields Already discussed today.

CCB#31 - Ensure that dictionary versions are specified in labels This is a documentation issue. Ron has produced documentation.

CCB#35 (Add Alias Index and Acronym Index to the Investigation class), CCB#36 (Add Mission Phase to the Investigation class), CCB#37 (Add List Author, List Editor, and List Contributor to the Modification Detail class)
Steve is working on these.

CCB#39 - Fixed problem where LDDTool does not include all rules from IngestLDD LDDTool ticket and update to the DPH, so LDD stewards don't think their rules.... (Note taker didn't hear the end)

CCB#44 - Issues caused by the failure to expose pds: attributes that are referenced in an LDD Steve will do a presentation on this. Not high priority. It is strongly related to CCB#s 52 and 53.

CCB#46 - Disallow multiple versions of the same LID from occurring in a single collection There are strong opinions on this one. There is a request from the MC for a use case survey at the nodes. We need an MC policy on versioning. Work is continuing.

CCB#50 - Clarify Ambiguous Author Information in Product Document Two parts - relationship and documentation.

CCB#51 - Add an optional attribute in the Target Identification class to identify Primary/Secondary targets Lots of discussion online. This was requested by PSA. We need to figure out the next step.
~ This seems like it came from the Editorial Work Group.
~ To improve the archive, and will also improve the website.

CCB#52 (Enable cross-namespace validations in sch files) and CCB#53 (Improve cross-namespace validations in sch files) Already discussed today

CCB#56 - Reconsider the PDS4 build cycle cadence This should be withdrawn. We were supposed to discuss and we did.

CCB#57 - Allow Encoded Native in File Area Observational Supplemental Has been passed to the CCB.

CCB#58 - New Product Native values for MSL/Mastcam The documentation is in peer review now. We will discuss this when that is complete. This will go to the MC ultimately.

CCB#59 - funding year only allows the value "9999" This is queued for implementation.

CCB#60 - New permissible value for File Area Observational.Header.parsing standard id: "CDF 3.9 ISTP/IACG" This is queued for implementation.

CCB#61 - Disconnect between SR and Validation of ASCII Real Update to the pattern. Complex update.

CCB#62 - Transfer Manifest class should enforce data types for its two mandatory columns DDWG is actively working on this one. Needs a TA, impact statement and values for the SR.

CCB#63 (International characters / Unicode support for organization name and display full name) and CCB#67 (Special characters not allowed in citation information fields) There are concerns with these. Could impact search. Need to figure out how to make it work.

CCB#65 - Modify description attribute to be required in Telescope and Facility Context Products This may be a bug fix.
~ It will look like asteroids. Think maybe something lost in the PDS3 translation. Unsure if it is worth having it there. A question of what's useful.
~ The DDWG will discuss this. It is currently optional.

CCB#66 - Improve documentation for List Author, List Contributor, and List Editor Need to clarify. No clear directions on how to use these.
~ The new contributor category was added. Need to clarify what that means. Needs discussion.
~ It was originally to auto generate the DOI information, but because of the way it was implemented it can't be used for that. It does at least give us parseable information.
~ That's an independent concern.

CCB#68 - ASCII File Name and ASCII Directory Path Name are using malformed regex expressions The DDWG already voted on this and passed it to the CCB. Now there is an extended discussion going on.

CCB#69 - HDF5 as an accepted “observational” and “supplemental” format Already discussed today.

CCB#70 - file name is not being validated for an extension as required by StdRef 6C.1.1 This might be a bug fix.
~ No. This is related to file name thing. Solution to this should be part of the other one.

Question: Does the SR have a requirement that file names must have an extension?
Answer: Think something to do with periods.
~ Someone is unsure. Directory names not allowed to.... (interrupted)
~ Yes. File names require a period. Can't have Readme, has to be a Readme.txt.

This is the end of the DDWG portion of the meeting.

title: DDWG Notes 2025-08-14
layout: default
date: 2025-08-14

August 14, 2025
Notes by Debra Kazden

Known Attendees: A. Bailey, J. Blalock*, T. Cornet, B. Hirsch*, L. Huber*. S. Hughes*, R. Joyner, C. Juarez, D. Kahan*, D. Kazden, A. Knight, P. Lawton, T. Lim*, J. Mafi*, S. McLaughlin, E. Schaefer*, B. Semenov* and M. Tiscareno*
(*'s indicate people who voted during this meeting - if applicable)

DDWG Agenda

(Included in email sent July 12, from J.Mafi, Agenda for 2025-08-14)

We will have a DDWG telecon this Thursday, August 14th at 9:30 PDT (12:30 EDT, 10:30 MDT). We will cover the following issues:

1) Action item reports.
2) CCB#28 As a user, I want validate to allow delimited tables with whitespace-only numeric fields – The Requested Changes to the SR have been revised to address concerns expressed by the CCB. Please review the revised document and be prepared to vote on the revised SCR.
3) CCB#62 Transfer Manifest class should enforce data types for its two mandatory columns – We will discuss and vote on the proposed implementation for this SCR.
4) CCB#32 Add Host type to Observing System Component – We will begin discussion of this SCR.
5) Review Open SCRs – We will review open SCRs to determine current status, next steps, and priorities.

Front Matter

1) Submitted to the CCB:
   - CCB#68 ASCII File Name and ASCII Directory Path Name are using malformed regex expressions
2) New issues in GitHub: (none)
3) Ready for a vote: (none)
4) Build 16.0 freeze date: Aug 2025. The following SCRs are queued for implementation: CCB#59, CCB#60
5) Withdraw by author:
   - CCB#41 Create a way for classes or attributes to be exposed to other LDDs but not to data providers

Action Items

a) PDS4 Versioning Use Cases Survey – All
b) CCB#39 – S. Hughes to update SCR
c) CCB#12 – S. Hughes to verify whether this is subsumed by CCB#52 and CCB#53
d) CCB#20 & CCB#32 – Tiger team formed and working the issue
e) CCB#31 – Ron and Matt working the issue **(Some Discussion)**

Tier 1 issues

- CCB#67 Special characters not allowed in citation information fields
- CCB#66 Improve documentation for List Author, List Contributor, and List Editor
- CCB#65 Modify description attribute to be required in Telescope and Facility Context Products
- CCB#63 - International characters / Unicode support for organization name and display full name
- CCB#62 - Transfer Manifest class should enforce data types for its two mandatory columns
- CCB#61 - Disconnect between SR and Validation of ASCII Real
- CCB#58 - New Product Native values for MSL/Mastcam
- CCB#57 - Allow Encoded Native in File Area Observational Supplemental
- CCB#56 – Reconsider the PDS4 build cycle cadence
- CCB#53 – Improve cross-namespace validations in sch files
- CCB#52 – Enable cross-namespace validations in sch files
- CCB#51 – Add an optional attribute in the Target Identification class to identify Primary/Secondary targets
- CCB#50 – Clarify Ambiguous Author Information in Product Document
- CCB#46 – Disallow multiple versions of the same LID from occurring in a single collection
- CCB#44 – Issues caused by the failure to expose pds: attributes that are referenced in an LDD.
- CCB#41 – Create a way for classes or attributes to be exposed to other LDDs but not to data providers
- CCB#39 – Fixed problem where LDDTool does not include all rules from IngestLDD
- CCB#37 – Add List Author, List Editor, and List Contributor to the Modification Detail class
- CCB#36 – Add Mission Phase to the Investigation class
- CCB#35 – Add Alias Index and Acronym Index to the Investigation class
- CCB#32 – Add Host type to Observing System Component
- CCB#31 – Ensure that dictionary versions are specified in labels
- CCB#28 – As a user, I want validate to allow delimited tables with whitespace-only numeric fields
- CCB#23 (JIRA-CCB-211) – XML/schema based files as archival data
- CCB#22 (JIRA-CCB-164) – Display Settings not required for images
- 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#21 (JIRA-CCB-367) – Provenance LDD Using PROV-DM
- CCB#20 (JIRA-CCB-358) – Instrument Package Context Products
- Tiger Team: M. Tiscareno (lead), B. Deen, B. Hirsch, S. Hughes, R. Joyner, J. Mafi, E. Schaefer.
- CCB#12 – Document Nuances for Reusability of Generic LDD Classes/Attributes
- CCB#5 – Move to V.2.0.0 for June 2025 Release
- CCB#4 – Backwards Incompatibility policies
- Presentation: Backwards Compatibility
- CCB#3 – Adopt Semantic Versioning with version 2 transition
- The CCB returned this, requesting that it be resubmitted as part specific proposal to transition to IM V2.0.0.
  (Some Discussions)

Tier 2 issues

1) 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!
1) JIRA-CCB-326 - Inconsistent Direction on How to Use Discipline Facets
2) RS: 259/281 (see new attachment to JIRA-CCB-281) (on hold)
3) JIRA-CCB-364 – Disallow bundles referencing collections by LID **(Not Discussed)**

Historical links:
- Export of history: https://github.com/NASA-PDS/PDS4-CCB/tree/main/data
- Backup: https://nasa-pds.atlassian.net/jira/software/c/projects/CCB/issues (limited seats)

Next meetings:
Aug 21 – Regular Meeting
Aug 28 – Extra Meeting
Sep 4 – Regular Meeting (Cancelled? IPDA Steering Committee)

DDWG

Front Matter

New SCRs: **CCB#69 (HDF5 as an accepted "observational" and "supplemental" format)** and **CCB#70 (file name is not being validated for an extension as required by StdRef 6C.1.1)**.
Posted two weeks ago.
~ CCB#69 was discussed at the F2F. Proposal to identify rules for creating PDS4 compliant HDF5 files.
~ CCB#70 would be a modification to the pattern for file names that require a file extension. Maybe worth considering it with CCB#68 (ASCII File Name and ASCII Directory Path Name are using malformed regex expressions).

Another new SCR: CCB#71 - Add Leap Second Flag
~ We have talked about this over and over again.
~ Anne is very thorough. This looks well done.
~ We will discuss this in the next few weeks.

Another new SCR: CCB#72 - New enumerated values for funding source
~ There was an additional thread where IPDA considered additional funding sources.
~ Yes. Tried to compile a list of funding sources for ESA missions. Not a full list yet.

Question: Does IPDA have actual data products for all of these?
Answer: Yes.
~ That's the difference between the US and Europe.
~ We can discuss this and change it, but this has been discussed in the past. Most of these are national space agencies, but some are universities.
~ Someone's eye fixated on the universities. The question is if this shouldn't be an enumerated list, but that would lead to typos. We need to consider why funding source is even here.

Question to Tanya: Do you foresee one or two universities per country or many more?
Answer: Hard to say. Going in the direction of smaller missions. Could be like instruments on CLPS. Funding could come from outside space agencies. Personal point of view because seeing universities participate in small missions, like CLPS.
~ Some funding comes from the national level. NASA side has less sources, but for European, every country can be different.

We will discuss this in a few weeks.
~ IPDA tried to capture what they know.
~ Yes, and it is likely to grow.
~ We need to consider the purpose of this. It was created with CCB#11(Add a Funding Information class to the Identification Area), need to decide if it should be an enumerated list. Need to ask ourselves what the point of this is.

Question: Comments?
Answer: This will never bring all the results. A bogus search that won't work.
~ If it's optional that's a point in favor of not having an enumerated list.
~ This would be a huge non-backwards compatible change that is not worth it.

Tier 1 Issues

CCB#28 - As a user, I want validate to allow delimited tables with whitespace-only numeric fields

There was concern in the CCB about wording for updates in the SR. There has been some discussion about the updated wording.
~ Someone liked the suggested change, moving the last statement in 4C.1(format description) to 4C.2 (describing how to put info in a PDS label). There is some ambiguity that needs clean up.

Action Item - Joe will move the statement after the meeting.

Question: Would like to vote now on the approach and the lien on the wording. Any comments or concerns on the new wording?
Answer: Comment about the purpose. We have table character that is human readable, so wondered if table character should be preferred.
~ That's been done in the past. There are plenty of fixed width tables in CSV in PDS. Think this is happening for the sake of backwards compatibility. This came up because a delimited table with empty spaces was flagged as an error by validate. We have stuff in the archive that needs this.
~ Cleaning up ambiguity in the standard. Cleaning house.
~ Yes. Cleaning up what validate was triggering on.

Question: Any other thoughts or comments?
Answer: (Silence)

The vote will be to approve the new wording with one lien - that statement - unless quoted, leading and trailing spaces are not part of the field - will be moved to 4C.2 .

The Vote:
ATMOS - Yes
EN - Yes
GEO - Yes
IMG - Yes
IPDA - Yes
NAIF - Yes
PPI - Yes
RMS - Yes
RS - Yes
SBN - Yes

All yes votes. Hopefully the new wording will help with the CCB.

CCB#62 - Transfer Manifest class should enforce data types for its two mandatory columns

There was an update to the test package. Described in GitHub. We now have a TA and an impact statement. If there's no comments, would like to vote on this. The one change is that in the SCR Mark had pointed out a typo in the description. The change was to correct that in the schema file.

Question: Any comments?
Answer: (Silence)

The Vote on the current implementation of CCB#62:
ATMOS - Yes
EN - Yes
GEO - Yes
IMG - Yes
IPDA - Yes
NAIF - Yes
PPI - Yes
RMS - Yes
RS - Yes
SBN - Yes

This will be passed to the CCB. It was pointed out that we should also examine the checksum class, but we should see what the CCB thinks about this before we do an SCR for that.

CCB#32 - Add Host type to Observing System Component

This is an older one. It had been connected to CCB#20 (Instrument Package Context Products), but at the F2F it was pointed out that it's not necessary for this to be tied to that. The issue is that a

while ago there were a number of permissible values for observing system component type that were deprecated.

Question: Any thoughts on this?
Answer: (Silence)

Someone's concern with the SCR is that some of the deprecated things for instrument wouldn't include instrument type. Removed spacecraft and just had host. That made sense, but other changes, like removing facility, didn't all make sense.

Question: Are there proponents for this change?
Answer: The concern is unclear.
~ (Opening the IM release, observing system in real-time.) A list of types. Quite a few are deprecated. Spacecraft was deprecated, now you would use host. We can also look at context product types. If we treat it similar to how we treat an instrument. Another concern is that facility is deprecated, but it's the class.
~ The proposal isn't to resurrect the list, but to set up rules to have an enumerated list that is identical to the class. Unsure what we should do , but when people only look at the product, not the context product, it's hard to know what kind of data is being looked at.
~ That's why only proposing for host.
~ Should be a schema rule that they must be identical. The same information in a different place.
~ Part of someone wonders how many users who download data don't already know what they have. System should deliver to people what they want. Unsure how much of a problem this is.
~ It's been discussed for awhile. Context products are supposed to be under the hood.
~ Someone agrees. People will know what they are downloading. This particular area in the IM is an artifact of PDS3. It's been problematic since day 1. Contextual information is very complex, that's the whole point of the context products. Where the information for search is. It's up to the system folks to make it work. We can address it if users are complaining, but don't think they are. We have been discussing this for over a decade.

Question: Other thoughts?
Answer: (Silence)
Another Question: What's the next step to move this forward?
Answer: A hint. Not intended to describe everything that we can describe. Should vote and move on. We should do the initial vote and see if EN should spend time on this.
~ We can do a straw poll.
~ Someone would have to abstain for now, but this does cause problems with more frequent changes to the IM.

Depending on implementation, it can be very dangerous for backwards compatibility.
~ It could invalidate a lot of released products if it's required, but if it's not then it's not very useful.
~ Someone wishes we could make it required for new products, but agrees it's not necessary to update all the old stuff.
~ If we don't we can't build the software that would use this.
~ But the primary purpose is human readability.
~ We will let this simmer for a week and vote next week.

IPDA Announcement from Tanya Lim

Tanya will be handing her DDWG role to Thomas Cornet. He will be the new ESA representative from IPDA, the new voting member. Tanya will back him up when necessary.
~ People thanked Tanya for her service and welcomed Thomas.

Question: Are there any updates on CCB#46 (Disallow multiple versions of the same LID from occurring in a single collection) or CCB#68(ASCII File Name and ASCII Directory Path Name are using malformed regex expressions)?
Answer: No update on CCB#46.
Another Question: Is GEO collecting the responses?
Answer: Yes. They have received a few.
~ PPI will do theirs. **(Action Item - Joe)**

CCB#68 was delivered to the CCB. Nothing back except for the suggestion that the new SCR, CCB#70 (file name is not being validated for an extension as required by StdRef 6C.1.1) should be considered at the same time while the hood is open.
~ The CCB is talking about updates. We should let them know we appreciate their work and are willing to spruce it up if asked. Hoping to expedite this. Want it in the next build.

Question: Are people happy with Anne's solution?
Answer: People need to read her comments. Know she wanted to contact who ever maintains the external standard. Will read what she wrote, but agree.

Question: When is the deadline?
Answer: The end of September.
~ So we have two meetings in August, then ESPC and IPDA, then September 18 meeting. So really, there are only two more meetings.
~ Yes, and the CCB needs time.

We have two weeks to get things through the DDWG. We will try to figure out CCB#68 and to get CCB#20 ready. People interested in CCB#20 should stay on this call after to work out a meeting.

People can let Joe know if there are other SCRs that we are trying to get into the next build so they can be on the agenda in the next few weeks.
~ IPDA would like CCB#72 ASAP, please.

title: DDWG Notes 2025-08-21
layout: default
date: 2025-08-21

August 21, 2025
Notes by Carlos Juarez and Joe Mafi, some reformatting by Debra Kazden

DDWG Agenda

(Included in email sent July 19, from J.Mafi, Agenda for 2025-08-21)

We will have a DDWG telecon this Thursday, August 21th at 9:30 PDT (12:30 EDT, 10:30 MDT). We will cover the following issues:

1) Action item reports.
2) CCB#32 Add Host type to Observing System Component – Please be ready for a preliminary vote on the proposed approach for implementing this SCR.
3) CCB#20 Need a Context object for Instrument Packages (JIRA-CCB-358) – This original SCR has been updated to reflect the proposed implementation, including updated Schema and Schematron files, sample labels, and proposed updates to the SR and DPH. Please review these materials and be prepared to discuss and vote.
4) CCB#72 New enumerated values for funding source – We will begin discussion of this SCR if we have time following the previous two topics.

Front Matter

1) Submitted to the CCB:
   - CCB#28 As a user, I want validate to allow delimited tables with whitespace-only numeric fields
   - CCB#62 Transfer Manifest class should enforce data types for its two mandatory columns
   - CCB#68 ASCII File_Name and ASCII Directory Path_Name are using malformed regex expressions
2) New issues in GitHub:
   - CCB#73 Support named locations on targets
3) Ready for a vote: (none)
4) Build 16.0 freeze date: Aug 2025. The following SCRs are queued for implementation: CCB#59, CCB#60
5) Withdraw by author:
   - CCB#41 Create a way for classes or attributes to be exposed to other LDDs but not to data providers

Action Items

a) PDS4 Versioning Use Cases Survey – All
b) CCB#39 – S. Hughes to update SCR
c) CCB#12 – S. Hughes to verify whether this is subsumed by CCB#52 and CCB#53
d) CCB#20 & CCB#32 – Tiger team formed and working the issue
e) CCB#31 – Ron and Matt working the issue

Tier 1 issues

- CCB#71 - Add Leap Second Flag

- CCB#72 - New enumerated values for funding source
- CCB#70 - file name is not being validated for an extension as required by StdRef 6C.1.1
- CCB#69 - HDF5 as an accepted “observational” and “supplemental” format
- CCB#68 - ASCII File Name and ASCII Directory Path Name are using malformed regex expressions
- CCB#67 - Special characters not allowed in citation information fields
- CCB#66 - Improve documentation for List Author, List Contributor, and List Editor
- CCB#65 - Modify description attribute to be required in Telescope and Facility Context Products
- CCB#63 - International characters / Unicode support for organization name and display full name
- CCB#62 - Transfer Manifest class should enforce data types for its two mandatory columns
- CCB#61 - Disconnect between SR and Validation of ASCII Real
- CCB#58 - New Product Native values for MSL/Mastcam
- CCB#57 - Allow Encoded Native in File Area Observational Supplemental
- CCB#56 – Reconsider the PDS4 build cycle cadence
- CCB#53 – Improve cross-namespace validations in sch files
- CCB#52 – Enable cross-namespace validations in sch files
- CCB#51 – Add an optional attribute in the Target Identification class to identify Primary/Secondary targets
- CCB#50 – Clarify Ambiguous Author Information in Product Document
- CCB#46 – Disallow multiple versions of the same LID from occurring in a single collection
- CCB#44 – Issues caused by the failure to expose pds: attributes that are referenced in an LDD.
- CCB#41 – Create a way for classes or attributes to be exposed to other LDDs but not to data providers
- CCB#39 – Fixed problem where LDDTool does not include all rules from IngestLDD
- CCB#37 – Add List Author, List Editor, and List Contributor to the Modification Detail class
- CCB#36 – Add Mission Phase to the Investigation class
- CCB#35 – Add Alias Index and Acronym Index to the Investigation class
- CCB#32 – Add Host type to Observing System Component
- CCB#31 – Ensure that dictionary versions are specified in labels
- CCB#28 – As a user, I want validate to allow delimited tables with whitespace-only numeric fields
- CCB#23 (JIRA-CCB-211) – XML/schema based files as archival data
- CCB#22 (JIRA-CCB-164) – Display Settings not required for images
- 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#21 (JIRA-CCB-367) – Provenance LDD Using PROV-DM
- CCB#20 (JIRA-CCB-358) – Instrument Package Context Products
- Tiger Team: M. Tiscareno (lead), B. Deen, B. Hirsch, S. Hughes, R. Joyner, J. Mafi, E. Schaefer.
- CCB#12 – Document Nuances for Reusability of Generic LDD Classes/Attributes
- CCB#5 – Move to V.2.0.0 for June 2025 Release
- CCB#4 – Backwards Incompatibility policies
- Presentation: Backwards Compatibility
- CCB#3 – Adopt Semantic Versioning with version 2 transition

- The CCB returned this, requesting that it be resubmitted as part specific proposal to transition to IM V2.0.0.

Tier 2 issues
1) Product Metadata Errata?
   a) Create a supplemental label to live next to original label, that can add aliases and supersede metadata

Parking lot – need a driver!
1) JIRA-CCB-326 - Inconsistent Direction on How to Use Discipline Facets
2) RS: 259/281 (see new attachment to JIRA-CCB-281) (on hold)
3) JIRA-CCB-364 – Disallow bundles referencing collections by LID

Historical links:
- Export of history: https://github.com/NASA-PDS/PDS4-CCB/tree/main/data
- Backup: https://nasa-pds.atlassian.net/jira/software/c/projects/CCB/issues (limited seats)

Next meetings:
Aug 28 – Extra Meeting
Sep 4 – Regular Meeting (Cancelled? IPDA Steering Committee)

DDWG

Front Matter

CCB#28 (As a user, I want validate to allow delimited tables with whitespace-only numeric fields) and CCB#62 (Transfer Manifest class should enforce data types for its two mandatory columns) have not been heard back on from the CCB yet.

CCB#68 (ASCII File Name and ASCII Directory Path Name are using malformed regex expressions) was rejected as written. See discussion on GitHub. Anne had written recommendations including proposing a new pattern. They want to get it into the next release.

We will talk about CCB#68 and CCB#70() next week, lots to catch up on since a lot of recommendations made.

New SCR - CCB#73 - Support named locations on targets
Summary: targets have been large objects, historically. Surface missions have named objects on those targets, those names should be available for search in the metadata. Idea to add a new class called location that is a member of target, which provides the name and whatever body (ex. IAU) created that name. Not trying to add geolocation, just want to add name for search purposes.

Tier 1 Issues

CCB#32 - Add Host type to Observing System Component Want to hold a vote today but need a tech assessment and system impact to make sure vote is final.

Request to Add host type to Observing System Component Specifically for Host, to be able to specify spacecraft, lander, etc.
Ethan Schaefer has concerns how performance will be impacted to validate. He agrees with the motivation for the request but it might be problematic to performance of validate tool.
Bob says could make the validation for this options by using options of validate. Bob does not see it as a problem, like tech being a hindrance Steve says he agrees with Bob, says it is just a hint, that the impact against performance would be worth it Need it specified in the SCR, that shall be validated?
Ethan says he does have concerns with referential integrity, type referential integrity, says he would vote no if can't validate, at least optionally Joe says text should change so that it can be added that validate can check Steve says reference integrity should be limited to internal references, if we say the names have to match, duplicate metadata causes issues, since request has been to better document info in the label, it almost has to be a hint. Doing the referential integrity is taking it too far.
Joe disagrees
Bob does see it as referential integrity checking and should be enforced with validate Ethan says it's a semantics thing
Joe says verifying that the type matches is more important than the names matches, and it comes up as warning if don't match
Ethan says, this might not be a big as to engineering, since it is using metadata that is already downloaded Bob points out that validate complained that names don't match, so not much of a performance hit Joe asks if Ethan is ok with Bob that it should be supported with validate if feasible or what?
Ethan says he is not sure but he is willing to proceed with vote since it's prelim vote Steve asks, what is used for searches. Assumes type in context prods. what is ramification of having wrong type in labels. Will it impact the system? he says users not really impacted Ethan says whatever is in context prods will be used. User will be impacted if wrong type in label so he fundamentally disagrees with Steve.
Tiscareno updated wording in GitHub
Tiscareno also not a fan if can't validate Steve asking if context prod changes, do you have to do ver 2.0 in label?
Tiscareno doesn't see that happening but Joe points out that there have been changes, primarily instruments Bob doesn't think host type to change but once validate changes in the future and validates something at the time, so doesn't see it as a concern.
Joe is concerned, maybe misunderstood Matt, his impression is the Matt expects this is reliable and the metadata can be harvested from metadata. Joe says the system will rely on context prods for this type of info

Vote on CCB#32:
ATMOS No
EN - Yes
GEO - Yes

IMG - Yes
IPDA - no vote
NAIF - Abstain
PPI - No
RMS - Yes
RS - Yes
SBN - Yes

PASSES
Steve when can we have it in place, says today In impact statement will reference what has been stated. Asked something but it is something the people that voted yes would have to address. Next time will ask for a consent vote and if so, will hold another vote Giving people more days if they want to re-vote Steve will contact Jordan

CCB#20 - Instrument Package Context Products updated proposal introduces inst group and component, also allows context products to be used recursively.

highlighted changes in CCB#20
Lyle says system impact says it is backwards compatible, he doesn't think it is Joe points a case were it is not, Cassini MIMI is defined as an instrument, Mimi Inca, Lemms. etc. Changing MIMI to inst group without allowing recursions would mean changing lids and making product invalid. Allowing recursion makes it backwards compatible. Joe thinks that we should live with what has been done so far, and use the new context products types going forward.
CLPS has complex inst host and other arrangements so current system needs to be updated to adapt for this Bob says everything in this CCB#20 is an addition. SO any existing thing is compatible. If you go and change stuff already there, then that is not backwards compatible but that is not a requirement. He is confident that it is back comp The SCR specifies that it allows us to define the hierarchy of MIMI and its sub-instruments, which you could not before Bob says it is a change in context prods, not prods themselves Joe says don't even need to change the context prods, so it is "great". Joe asks if people have had sufficient time to view this thing or if they want more time to read it over. He says the sooner they can vote on this the better but is giving them time. It has a tech assessment and system impact so having a final vote.
Tech assessment needs to be rewritten :(so no vote, hold off until next Thursday

CCB#72 - New enumerated values for funding source

Enumerated values for funding source He wants to talk about CCB#68 too 72, quite a few proposed additions to funding source name enumerations. Do we want to not have it enumerated, that always carries some hazards. Point made that the way Europeans funding works is diff than USA, so need additional funding sources to be able to reference Steve says if having a standard list helps, but what is the taxonomy, how do the text fit together? proposes leaving funding source as is and label it as national agency, for example, and the other use a type list like used for instruments. the type list (?) third is use type list for both (?) Joe concerned about second option, since institutions directly funding source and should be acknowledge as main funding source.
Steve is saying if they are of equal funding stature, they should be acknowledge then, but if different, use type list Bob says use something like CTLI, to make it easy to add changes, like in an instrument dictionary. Don't enumerate in core, takes to long to update Lyle CTLI people are willing

to search for spectrometers, meteorology, plasma inst. For this attribute, the reason for doing this is for documentation or do we expect people to really search for everything that ESI put up in PDS (?) He thinks this is more for documentation or search Boris said something, finds all of this is baloney??
Lyle recommends just making it simple text field, Boris agrees Bob says middle ground, making a table of suggested value, keep free text. So not enforced but a good data provider would use the same terminology Boris says Jordan promises AI will find stuff, he sounds skeptical about the search benefit for this
Joe likes Bob's idea of having examples available to use. Joe maybe wants some instructions that can be applied and give some examples, agrees with Boris, don't want to update the list constantly

Joe says he views the different CCB approaches as follows:
1) Add requested enumerated values in PDS4 Common
2) Add requested enumerated values in external LDD (similar to CTLI)
3) Make Funding Source/name a free text field

Joe would want input from nodes, thinks they should approach in a way
Tiscareno:
Steve had a 2a version, divide the lists into separate lists, one in core and another in CTLI.
Tiscareno says no one would be able to search on this Ethan says all of the existing products don't have this info Bob says in terms of search, never will be able to search on everything, but make it the best they can going forward. He sees it as not too important, improve were you can. Bob asks what is the time frame for this, is it urgent from IPDA, do they need it for this build? We might have to do number 1, and separately consider 2 or 3. If not needed right away, can do something else.
Joe will ask Mark but they said they want it in next build.
Bob worries about 2 or 3 to be able to get in this build.
Joe says 3 could do, but 2 likely not implemented before next build.
Ethan agrees with Bob to not give up just because some products don't have this. Ethan who would want to search for this, like where these products came from, but would Europeans find that useful?
Bob, responding to Steve, doesn't see the value in splitting the list Steve, CTLI approach can be types or names or anything, no implication of any type of element being listed.
Bob misunderstood, just keep what is there and add something CTLI, mechanism based, and old one is deprecated Steve says yeah, that is what he is saying, and names listed can be added Lyle to Ethan, Bentley, says he is fine not using enumerated list.
They disagree the use of searching for Europeans.

Straw Poll on 1,2,3:
ATMOS - 3,2,1
EN - 2,3,1
GEO - 2,1,3
IMG - 3,2,1
IPDA - Not here
NAIF - 3,2,1
PPI - 3,1,2
RMS - 3,2,1
RS - 3,2,1
SBN - 3,2,1

3 preferred approach, Mark Bentley will be contacted, give him feedback, let him know DDWG feedback and let him know SCR being updated will vote next week on it

title: DDWG Notes 2025-08-28
layout: default
date: 2025-08-28

August 28, 2025
Notes by Debra Kazden

Known Attendees: A. Bailey*, J. Blalock*, B. Hirsch*, S. Hughes*, R. Joyner, C. Juarez, D. Kahan*, D. Kazden, A. Knight, P. Lawton, T. Lim*, J. Mafi*, E. Schaefer*, M. Seritan and M. Tiscareno*
(*'s indicate people who voted during this meeting - if applicable)

DDWG Agenda

(Included in email sent July 26, from J.Mafi, Agenda for 2025-08-28

We will have a DDWG telecon this Thursday, August 28th at 9:30 PDT (12:30 EDT, 10:30 MDT). We will cover the following issues:

1) Action item reports.
2) CCB#32 Add Host type to Observing System Component – The TA and System Impact statements have been added for this SCR. The System Impact includes a lien regarding verifying the impact of the SCR on validation. Please review the most recent comments regarding system impacts and be ready to discuss any concerns. If appropriate, we will also conduct a consent vote on this item.
3) CCB#20 Need a Context object for Instrument Packages (JIRA-CCB-358) – This SCR includes Schema and Schematron files, sample labels, proposed updates to the SR and DPH, as well as some supplemental materials. Please review the proposal and attachments and be prepared to discuss and vote.
4) CCB#72 New enumerated values for funding source – This SCR has been updated to reflect the implementation approach approved by DDWG vote last week. Supplemental materials including sample Schema, Schematron, test labels, and proposed updates to the DPH have also been added. Please review the proposal and attachments and be ready to discuss and vote.
5) CCB#68 ASCII File Name and ASCII Directory Path Name are using malformed regex expressions & CCB#70 file name is not being validated for an extension as required by StdRef 6C.1.1 – Please review and be ready to discuss and possibly vote on these SCRs.

Front Matter

1) Submitted to the CCB:
   - CCB#28 As a user, I want validate to allow delimited tables with whitespace-only numeric fields
   - CCB#62 Transfer Manifest class should enforce data types for its two mandatory columns
   - CCB#68 ASCII File Name and ASCII Directory Path Name are using malformed regex expressions
2) New issues in GitHub:
   - CCB#73 Support named locations on targets
3) Ready for a vote: (none)

4) Build 16.0 freeze date: Aug 2025. The following SCRs are queued for implementation: CCB#59, CCB#60
5) Withdraw by author:
    - CCB#41 Create a way for classes or attributes to be exposed to other LDDs but not to data providers

(**Some Discussion)**

Action Items

a) PDS4 Versioning Use Cases Survey – All
b) CCB#39 – S. Hughes to update SCR
c) CCB#12 – S. Hughes to verify whether this is subsumed by CCB#52 and CCB#53
d) CCB#20 – Tiger team formed and working the issue
e) CCB#31 – Ron and Matt working the issue

(Discussed CCB#20)

Tier 1 issues

- CCB#71 - Add Leap Second Flag
- CCB#72 - New enumerated values for funding source
- CCB#70 - file name is not being validated for an extension as required by StdRef 6C.1.1
- CCB#69 - HDF5 as an accepted “observational” and “supplemental” format
- CCB#68 - ASCII File Name and ASCII Directory Path Name are using malformed regex expressions
- CCB#67 - Special characters not allowed in citation information fields
- CCB#66 - Improve documentation for List Author, List Contributor, and List Editor
- CCB#65 - Modify description attribute to be required in Telescope and Facility Context Products
- CCB#63 - International characters / Unicode support for organization name and display full name
- CCB#62 - Transfer Manifest class should enforce data types for its two mandatory columns
- CCB#61 - Disconnect between SR and Validation of ASCII Real
- CCB#58 - New Product Native values for MSL/Mastcam
- CCB#57 - Allow Encoded Native in File Area Observational Supplemental
- CCB#56 – Reconsider the PDS4 build cycle cadence
- CCB#53 – Improve cross-namespace validations in sch files
- CCB#52 – Enable cross-namespace validations in sch files
- CCB#51 – Add an optional attribute in the Target Identification class to identify Primary/Secondary targets
- CCB#50 – Clarify Ambiguous Author Information in Product Document
- CCB#46 – Disallow multiple versions of the same LID from occurring in a single collection
- CCB#44 – Issues caused by the failure to expose pds: attributes that are referenced in an LDD.
- CCB#41 – Create a way for classes or attributes to be exposed to other LDDs but not to data providers
- CCB#39 – Fixed problem where LDDTool does not include all rules from IngestLDD
- CCB#37 – Add List Author, List Editor, and List Contributor to the Modification Detail class
- CCB#36 – Add Mission Phase to the Investigation class

- CCB#35 – Add Alias Index and Acronym Index to the Investigation class
- CCB#32 – Add Host type to Observing System Component
- CCB#31 – Ensure that dictionary versions are specified in labels
- CCB#28 – As a user, I want validate to allow delimited tables with whitespace-only numeric fields
- CCB#23 (JIRA-CCB-211) – XML/schema based files as archival data
- CCB#22 (JIRA-CCB-164) – Display Settings not required for images
- 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#21 (JIRA-CCB-367) – Provenance LDD Using PROV-DM
- CCB#20 (JIRA-CCB-358) – Instrument Package Context Products
- Tiger Team: M. Tiscareno (lead), B. Deen, B. Hirsch, S. Hughes, R. Joyner, J. Mafi, E. Schaefer.
- CCB#12 – Document Nuances for Reusability of Generic LDD Classes/Attributes
- CCB#5 – Move to V.2.0.0 for June 2025 Release
- CCB#4 – Backwards Incompatibility policies
- Presentation: Backwards Compatibility
- CCB#3 – Adopt Semantic Versioning with version 2 transition
- The CCB returned this, requesting that it be resubmitted as part specific proposal to transition to IM V2.0.0.

(Some Discussions)

Tier 2 issues

1) 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!

1) JIRA-CCB-326 - Inconsistent Direction on How to Use Discipline Facets
2) RS: 259/281 (see new attachment to JIRA-CCB-281) (on hold)
3) JIRA-CCB-364 – Disallow bundles referencing collections by LID

(Not Discussed)

Historical links:
- Export of history: https://github.com/NASA-PDS/PDS4-CCB/tree/main/data
- Backup: https://nasa-pds.atlassian.net/jira/software/c/projects/CCB/issues (limited seats)

Next meetings:
Sep 4 – Regular Meeting - Cancelled due to conflict with IPDA Steering Committee meeting
Sep 18 – Regular Meeting

DDWG

Front Matter

No new SCRs this week. No news from the CCB.

Tier 1 Issues

CCB#32 - Add Host type to Observing System Component

We voted to pass this last week, but it still needed a TA and system impact statement. There is a lien on the system impact statement for validate. Jordan weighed in on the impact to software. His comments were posted by Steve in GitHub. Hopefully people have read them.

Question for Steve: Does that resolve the lien for this?
Answer: Not sure.
~ It wasn't clear to people.
~ The question is what the actual Requested Change is.
~ The key Requested Change is for it to be enforced by validate if technically feasible.
~ Yes. There were messages by Jordan, Steve and Joe. One last week. Hopefully people read them. The concern is the host relationship. Proposing a text description as a solution. Has some benefits - it's simple to implement, human readable, avoids enumerations... (more). The suggestion is a simple text field. The question is if we want to discuss that.

That change would take this off the table. We would need to re-do the SCR. We would need to discuss that with the submitter.

Question: We already have a description under observation. Would that work?
Answer: It should.
~ Not strongly advocating for that description. We could ask the submitter.
~ True, that would suffice. The question is based on the conversation last time. Unsure why we would go beyond that. But as a point of order, we would derail this.

This was approved as it is. If there is support for this change we can take it off the table, update the SCR and re-vote. It would need to be an e-vote if we are trying to get this in the next release. We could do a straw poll.

Question: Is there a second to this?
Answer: (Silence)
~ Since there is no second, would prefer to keep this as-is with enumerated list. That brings us back to the question.
~ Someone repeat the question.
~ The current proposal is for a new attribute host type with permissible value list that matches inst host types in the system. The suggestion is to change that to free text where DPs could give a hint about it without being constrained by a permissible values list.
~ Seems like the actual information is in the context product, so all that is needed is a hint in the observing system.

~ It would be more difficult for search, but maybe easier for users with descriptions.

There is no reason for it to be an enumerated list in any case. Just needs to be the same as type. Sounds like validation step is onerous. If Jordan says it's fine then would prefer to do it.
~ We already have the information. If the goal is a hint than this would be completely redundant if checked against host in context product.
~ It violates the duplicate data principal from a modeling perspective.

It doesn't sound like there is support to change this. Need to close the lien before we send it to the CCB.

Question: Can the lien be closed?
Answer: To clarify, the Requested Changes include validation against the context products, so the answer is we clarified. Considered the issues. The lien can be removed.
~ Joe is doing it in real-time.

Question: Do we want to do a consent vote? Given that this has already been approved, we could do another vote or just send it to the CCB.
Answer: (Silence)

The previous vote stands.
~ It passed 6 to 2, with one abstention.

CCB#20 - Instrument Package Context Products

This is a big one. It's been worked on quite a bit. Big changes to the IM.

Question: Any discussion? Questions? Comments? Concerns?
Answer: (Silence)
~ We had a week to review this. The TA was updated. The system impact statement is done.

The Vote on CCB#20 as currently written:
ATMOS - No (ATMOS is not here, but Lyle emailed his vote to Joe before the meeting)
EN - Yes
GEO - Yes
IMG - Yes
IPDA - Yes
NAIF - Abstain
PPI - Yes
RMS - Yes
RS - Yes
SBN - Yes

This will be passed to the CCB for their consideration.

CCB#72 - New enumerated values for funding source

We voted last week on the different approaches from the original proposal. The materials in GitHub reflect the changes.

Question: Comments? Questions?
Answer: (Silence)

There is a lien that we need wording for the DPH. Unsure if Steve has had a chance to look at it yet.
~ Steve says the lien was not his.
~ Ron is not sure why we need a lien. It went from a long list to a much shorter list. Not sure what would go in the DPH.
~ We should put some guidelines in the DPH so we don't get the Wild West. Guidelines are along the lines of what is posted, to get more consistent values from one to another.

Question: This is posted?
Answer: Yes. Under attachments. Went up Tuesday. Guidelines for DPs to get a little more uniformity. One of the concerns is about the guideline that limits this to the Latin alphabet. Could make it hard to put names in native language.
~ If we allow diacritical marks people can spell names correctly. Don't see a problem.
~ There are examples. For Korean to English unsure what the translation would be.
~ Someone is reading the proposed DPH updates now. It's great. Explains the impact. Would modify the impact statement to applicable impacts.

We could ask people at the IPDA meeting next week.
~ That is a great suggestion. Joe has a talk on the coming changes to the IM.
~ Yes, that would be useful. A good idea.

It seems like some of these are more than examples. Recommendations.
~ The concern is that we don't want to have to update the DPH every time we add a new value. We can ask people to follow the examples. Free text, so we can't enforce.
~ Someone agrees, but we can make recommendations.
~ That's an enumerated list in the DPH.

Question: Other comments? Would like to vote on this. We can keep it open based on input from IPDA members, but we should vote while we are all on the line now.
Answer: (Silence)

The vote will be with Steve's Suggested Change to change the wording to say recommendations instead of examples. Changed in real-time.

Question: Any other comments?
Answer: A side comment. It seems like for funding agency we could have included a RORID, but we would need a new SCR. It's just a thought.
~ A good thought, but agree, it's a separate SCR.
~ Another person also agrees.

The Vote on CCB#72:

ATMOS - Yes
EN - Yes
GEO - Yes
IMG - Yes
IPDA - Yes (with discussion next week)
NAIF - Yes
PPI - Yes
RMS - Yes
RS - Yes
SBN - Yes

This passed with two liens - post updated DPH text to the SCR and to discuss with the IPDA folks to make sure we are capturing their preferences.

CCB#68 (ASCII File Name and ASCII Directory Path Name are using malformed regex expressions) and CCB#70 (file name is not being validated for an extension as required by StdRef 6C.1.1)

These are connected. Both are updates to patterns.

There have been lots of suggestions, including Anne's suggestion that we include a pattern for directory names. EN did some testing. Seems to be working as expected.
~ EN updated this morning. The pattern doesn't catch it if it ends with an underscore or dash, but it works.
~ Anne had suggested that we rewrite the SCR. Some edits were done. Reworded what was there. Updated rules for file names.
~ Someone added a comment and tested Anne's file names. Comments include test results. It works. Has one question about following characters, like CRLF, but it passes. (See GitHub for more)

Someone likes the idea of changing to xs:token and allowing extra white space. Ultimately, what's implemented is if the white space is in a file name. Assumes validate will check. Think it would catch mismatches. Unsure if it does actually check file name independent of what it says in the label. Need to check rules for file names.
~ Someone else asked Jordan if XML parser is used by validate. The answer is yes. More basic question is that this issue was because we didn't have a standard. Now we do. Unsure if that is addressed. We could write it. IM could validate strictly to be consistent with the standard.
~ Anne has a nice example with very long file name, not on a single line. So there may be reasons to have that.

We will vote on xs:token versus xs:string. Would like to give everyone a chance to state their preference.

Question: Does everyone understand the difference? Anne outlines it well in her comments. We currently have xs:string, so we can't have white space in the attribute values. Xs:token would allow a value to be multiple lines. It was pointed out that there is a difference between IM specification

and the actual schema file. In the IM spec it says token, but it is implemented as string in the schema.
~ Yes, the raw technical input specifies token.
~ We will do a quick poll on token or string.

Someone's last comment in GitHub was that they planned to ask Jordan a question.
~ That is a "never mind" Not what we are talking about.

Someone is still unclear on what we are going to vote on.
~ Anne's comment gives the basic information. She tested a long file name with space padding and CRLF. Validate flagged it as a problem because xs:string. If the base type were xs:token that would allow white space. It wouldn't be part of the value. Only difference is that if it remains xs:string there is no option to break up a long name.

Question: Is there any reason not to change to token?
Answer: It could have white spaces and be wrong and that wouldn't be flagged.
~ Someone tested that.
Another Question: Leading or trailing white spaces would be ignored?
Answer: Yes.
~ Leading white space would be flagged as an error.
~ Been strict up to now in IM. IM controls string with rules and schematron, so leading spaces would be an error. Talking about loosening it to allow blanks. As a purist, would want specification correct and to match the standard that says you can't have leading or trailing spaces.
~ Someone did a lot of testing and thinks token is okay.

Someone thinks leaving it as string does provide better validation, but if file name has spaces it will be flagged and a mismatch. It depends on what validate checks for.
~ Jordan said it won't check for leading spaces. It shouldn't make any difference to the system people.
Question: Any last comments?
Answer:(Silence)

Poll - The Preference Between String and Token:
ATMOS - Not Here
EN - String
GEO - Token
IMG - Token
IPDA - Abstain
NAIF - String
PPI - Token
RMS - Token
RS - Abstain
SBN - Token

The preference is for token. Five for token, 2 for string and 2 abstain. Based on that we will put that in the current proposal.

Question to tester: You said you tested Anne's pattern. Which one?

Answer: It was the pattern in the re-write. Unsure where it came from. Posted the test results.
~ The pattern that Anne proposed was tested by Ron too. Separate. Tokens here make it less efficient. That is why was curious about which pattern was used for testing.
~ It was pulled from her comments.
~ Someone is confused about the origin of that pattern too. Testing it now.
~ Another person is disinclined to worry too much about efficiency. Not sure we have a good handle on that and there is no way to verify which is faster in reality. If Anne preferred this then that is good enough for this speaker. Personal opinion is that we shouldn't get lost in the weeds. Shouldn't get bogged down.
~ Sensible. It really only becomes a problem if huge scales, if running hundreds of thousands of labels.
~ Speaker likes both patterns from Anne's document. EN is testing them now.

We have some updates to make and issues to sort out. Need to see the results of the new testing versus the previous testing. Hope to vote in time to include this in the next build. We will e-vote if possible.

Next meeting is September 18.

Meeting Chat

Joyner:
Can you post the exact pattern / regEx that you used please. I am getting different results for 'foobar_.jpg' and 'foobar-.jpg'
gracias

Schaefer:
Done. (On GitHub.)

title: DDWG Notes 2025-09-18
layout: default
date: 2025-09-18

Sept 18, 2025
Notes by Debra Kazden

Known Attendees: A. Bailey, J. Blalock*, T. Cornet*, B. Hirsch*, L. Huber*, S. Hughes*, R. Joyner, C. Juarez, D. Kahan*, D. Kazden, A. Knight, P. Lawton, S. Loftin, J. Mafi*, S. McLaughlin, E. Schaefer*, B. Semenov*, M. Seritan and M. Tiscareno*
(*'s indicate people who voted during this meeting - if applicable)

DDWG Agenda

(Included in email sent September 17, from J.Mafi, Agenda for 2025-09-18
I hope you've enjoyed the last few weeks with no DDWG telecons. We will have a DDWG telecon this Thursday, September 18th at 9:30 PDT (12:30 EDT, 10:30 MDT). We will cover the following issues:

1) Action item reports.
2) CCB#72 New enumerated values for funding source – The IPDA Steering Committee provided the requested feedback on how funding sources should be identified (please see Anne's comment of 5 Sep 2025 in GitHub and my suggested updates to the Proposed Changes).
3) CCB#63 International characters / Unicode support for organization name and display full name & CCB#67 Special characters not allowed in citation information fields – Please review these SCRs and be ready to discuss.
4) CB#50 Clarify Ambiguous Author Information in Product Document – Please review the options outlined in my comment of 31 Oct 2024 and be ready to discuss.
5) CCB#66 Improve documentation for List Author, List Contributor, and List Editor – Adjacent to CCB#50. Please review and be ready to discuss.

Front Matter

1) Submitted to the CCB:
   - CCB#28 As a user, I want validate to allow delimited tables with whitespace-only numeric fields – Approved
   - CCB#62 Transfer Manifest class should enforce data types for its two mandatory columns – Approved
   - CCB#20 Need a Context object for Instrument Packages (JIRA-CCB-358)
   - CCB#32 Add Host type to Observing System Component
   - CCB#57 Allow Encoded Native in File Area Observational Supplemental
2) New issues in GitHub:
   - CCB#74 Allow non-archival data in popular contemporary formats to appear alongside archival counterparts
   - CCB#75 Need new value for object type ("Interstellar Object")
3) Ready for a vote: (none)
4) Build 16.0 freeze date: Aug 2025. The following SCRs are queued for implementation: CCB#28, CCB#59, CCB#60, CCB#62
5) Withdraw by author:

- CCB#41 Create a way for classes or attributes to be exposed to other LDDs but not to data providers
  (Discussed)

Action Items

a) PDS4 Versioning Use Cases Survey – All
b) CCB#39 – S. Hughes to update SCR
c) CCB#12 – S. Hughes to verify whether this is subsumed by CCB#52 and CCB#53
d) CCB#20 – Tiger team formed and working the issue
e) CCB#31 – Ron and Matt working the issue **(No Updates)**

Tier 1 issues

- CCB#75 Need new value for object type ("Interstellar Object")
- CCB#74 Allow non-archival data in popular contemporary formats to appear alongside archival counterparts
- CCB#73 Support named locations on targets
- CCB#72 New enumerated values for funding source
- CCB#71 - Add Leap Second Flag
- CCB#72 - New enumerated values for funding source
- CCB#70 - file name is not being validated for an extension as required by StdRef 6C.1.1
- CCB#69 - HDF5 as an accepted “observational” and “supplemental” format
- CCB#68 - ASCII File Name and ASCII Directory Path Name are using malformed regex expressions
- CCB#67 - Special characters not allowed in citation information fields
- CCB#66 - Improve documentation for List Author, List Contributor, and List Editor
- CCB#65 - Modify description attribute to be required in Telescope and Facility Context Products
- CCB#63 - International characters / Unicode support for organization name and display full name
- CCB#62 - Transfer Manifest class should enforce data types for its two mandatory columns
- CCB#61 - Disconnect between SR and Validation of ASCII Real
- CCB#58 - New Product Native values for MSL/Mastcam
- CCB#57 - Allow Encoded Native in File Area Observational Supplemental
- CCB#56 – Reconsider the PDS4 build cycle cadence
- CCB#53 – Improve cross-namespace validations in sch files
- CCB#52 – Enable cross-namespace validations in sch files
- CCB#51 – Add an optional attribute in the Target Identification class to identify Primary/Secondary targets
- CCB#50 – Clarify Ambiguous Author Information in Product Document
- CCB#46 – Disallow multiple versions of the same LID from occurring in a single collection
- CCB#44 – Issues caused by the failure to expose pds: attributes that are referenced in an LDD.
- CCB#41 – Create a way for classes or attributes to be exposed to other LDDs but not to data providers
- CCB#39 – Fixed problem where LDDTool does not include all rules from IngestLDD
- CCB#37 – Add List Author, List Editor, and List Contributor to the Modification Detail class

- CCB#36 – Add Mission Phase to the Investigation class
- CCB#35 – Add Alias Index and Acronym Index to the Investigation class
- CCB#32 – Add Host type to Observing System Component
- CCB#31 – Ensure that dictionary versions are specified in labels
- CCB#28 – As a user, I want validate to allow delimited tables with whitespace-only numeric fields
- CCB#23 (JIRA-CCB-211) – XML/schema based files as archival data
- CCB#22 (JIRA-CCB-164) – Display Settings not required for images
- 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#21 (JIRA-CCB-367) – Provenance LDD Using PROV-DM
- CCB#20 (JIRA-CCB-358) – Instrument Package Context Products
- Tiger Team: M. Tiscareno (lead), B. Deen, B. Hirsch, S. Hughes, R. Joyner, J. Mafi, E. Schaefer.
- CCB#12 – Document Nuances for Reusability of Generic LDD Classes/Attributes
- CCB#5 – Move to V.2.0.0 for June 2025 Release
- CCB#4 – Backwards Incompatibility policies
- Presentation: Backwards Compatibility
- CCB#3 – Adopt Semantic Versioning with version 2 transition
- The CCB returned this, requesting that it be resubmitted as part specific proposal to transition to IM V2.0.0.
  (Some Discussions)

Tier 2 issues

1) 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!

1) JIRA-CCB-326 - Inconsistent Direction on How to Use Discipline Facets
2) RS: 259/281 (see new attachment to JIRA-CCB-281) (on hold)
3) JIRA-CCB-364 – Disallow bundles referencing collections by LID **(Not Discussed)**

Historical links:

- Export of history: https://github.com/NASA-PDS/PDS4-CCB/tree/main/data
- Backup: https://nasa-pds.atlassian.net/jira/software/c/projects/CCB/issues (limited seats)

Next meeting:
Oct 2 - Regular Meeting

DDWG

Front Matter

Welcome back. We had a bit of a break. Think the freeze happens in the next week or week and a half. We are now working on the build after next. Believe we got everything we hoped to CCB. CCB#s 28 and 62 are queued for implementation. We are waiting to hear back on CCB#s 20, 32 and 57. There's a CCB meeting planned to discuss CCB#s 20 and 32.

There are two new SCRs:

- CCB#74 - Allow non-archival data in popular contemporary formats to appear alongside archival counterparts
- CCB#75 - Need new value for object type ("Interstellar Object")

CCB#74 came out of conversations. We have always had an ethos in PDS4 that product observational is reserved for archival data, but we also have people who want non-archival products. Easy in PDS3 to put archival format next to contemporary format. No way to do it in PDS4. Could use product external or define something new so can have them next to archive products, but clear that they are not part of the archive.
~ This will be discussed in the future.

CCB#75 is because there is a weird ambiguity in our target types in how they are defined. Need to add interstellar products.

Tier 1 Issues

CCB#72 - New enumerated values for funding source

We discussed the proposed guidelines with the IPDA. Anne posted comments describing the recommendations from them. Based on that, proposing changes to the Recommended Changes and implementation. There's also a comment that funding source be zero to many, but there are issues with that. Joe will double check that it's zero to many.
~ Someone already checked, but thought it meant there might be multiple awards.
~ IPDA has instruments funded by multiple agencies. Co-funded. Multiple sources for the same award.
~ If there are use cases we can fix it to support that. See no harm in making funding source zero to many. Will edit in real-time.

Question: Any questions or comments?
Answer: (Silence)

Question: Would people be okay to vote on the approach so we can work on the implementation? Any objections?
Answer: Someone doesn't have any objections but wants it written down, exactly what we are voting on.

~ Joe is adding it in GitHub in real-time.
~ This is a non-binding vote on implementation recommended in Joe's comments from two days ago, to use RORID to facilitate search and to update the DPH.

Question: What if a person doesn't have an ID? How did that become required?
~ Someone seconds the question.
Answer: Currently, funding source is optional. Could just put text to identify an individual. Could make organization name or RORID. Could run into the same issue with ORCID.
~ It's very easy to get an ORCID or RORID. Not an issue.
~ It's a matter of privacy.
~ This is optional. Don't have to use the acknowledgement. Can be anonymous.
~ So, if I want to be acknowledged I have to have an ID. Don't understand why that became necessary. We should use name and make ID optional.
~ A person agrees.
~ Then this is useless for search.
~ Describing an edge case.
~ There are people and organizations that won't want an ID because then they will get tracked.
~ Two people agree.
~ This is optional.

Then we are back to an enumerated list.
~ It is unclear why text wouldn't be sufficient.
~ Someone agrees completely. This isn't a searchable class, so no need for an enumerated list. Agree that it is reasonable to have it for documentation, but people won't search on it.

Question: Is there any response to the question why text wouldn't be sufficient?
Answer: Text would be sufficient.
Another Question: Would it be searchable?
Answer: There is text based search, but we can't find funding source everywhere.

Question: Is the current design, where this is optional, is that sufficient or do we want another option for RORID or ORCID?
Answer: Organization name is very clear.
~ Exists to include associated text. Funding acknowledgement text is up to the organization and what's determined appropriate by PDS peer review.
~ Someone just thinks this is asking too much and can lead to inconsistency.
~ A person laughed.

We will give everyone two weeks to think about this and vote on an approach next time.

CCB#63 (International characters / Unicode support for organization name and display full name) and CCB#67 (Special characters not allowed in citation information fields)

These are related. Both request the ability to add specific characters, non-ASCII characters for certain attributes. One comment on this was that we need to check with the system folks to determine impacts. Joe contacted Jordan and asked him. Jordan wasn't certain, but said we need to support this and they will figure out how to do it.

Question: What exactly does Jordan say we have to support?
Answer: We have to support what is in these two requests and allow diacriticals and things like that. UTF-8 types.
~ Neither SCR has much written under the Suggested Changes.
~ Changing data types for the list of attributes to UTF-8 types is what the change request would be.
~ Change of type of text field to apply this.
~ Trent provided some examples. Proposed UTF-8.

Question: If we allow this, what is and is not in UTF-8?
Answer: Good question. Can use all Latin characters with diacritical marks. Unsure of what else.
Another Question: UTF-8 has a million characters in it. Is there a way to limit it?
Answer: Someone agrees with the concern. It opens it up to character sets which would be problematic for search. The question is if and how. Presume we could do some pattern limits, but unsure.
~ While writing tools someone ran into patterns. They do exist. We could decide which is best for us.
~ That is part of the question, the limitations.

We do constrain already. Basic Latin, Extended Latin, but then there are so many others - Arabic, Emoji, Chinese… We need to decide what we want to constrain to before we decide.
~ We should reach out to our international partners to see if Basic Latin and Extended Latin is enough.
~ We need to gather IPDA perspectives, and I would suggest that English is the common language for science that is recognized.
~ Good point.
~ Thomas will ask IPDA members to see if Basic Latin and Extended Latin would meet their needs. **(Action Item - Thomas)**

Question: What about a parallel attribute name native? Has that idea died?
Answer: Not sure, but this is the direction in the proposals.
~ Just thinking if we use English it has to be ASCII for name and then UTF-8 for parallel.
Another Question: Can you post that in GitHub?
Answer: Will do. **(Action Item - Boris)**

Someone for EN said that in working with all the metadata, whenever people use UTF-8 it is very problematic for the tools. Have to find and convert to ASCII. Even have a Python to search and convert. Preference is no UTF-8.

Question: What's the context?
Answer: XML labels. Even the DOI service. UTF-8 characters have to be stripped out. Tools can't handle it now. UTF-8 is scattered throughout the IM.
~ We need to be careful about where we will allow UTF-8. Don't see a problem with allowing it for name.
~ It makes life hard for EN because the tools can't actually handle UTF-8 at all yet.
~ Jordan said we need to support this, so EN will have to upgrade the tools.
~ EN will need to figure out the scope. Need to be careful about changing letters from UTF-8 to ASCII because that can change meanings. Need to be careful.

Question: Other comments on this?
Answer: (Silence)

The action item is on Thomas to check with the IPDA for their feedback on limitations on UTF-8 and if they want more than extended Latin.
~ He understands the problem.
~ We have people who speak many languages who use PDS. Can ask them how to write their own names.

So, in the IM we have terminological entry to allow alias. Early in PDS we tried to translate into Russian Cyrillic. Maybe there is a need for the native, so the model does allow it, but it's complicated.

Question: How would that work? Would it turn all attributes into classes?
Answer: Unsure. Maybe just a native area. Wouldn't be for search. Probably not classes, just a way to add native information. Maybe in citation area.
~ Someone wouldn't want to move too far from the area they were used in.
~ The point is that we have done it.

Question: Any last comments on this?
Answer: Thomas will get feedback and Boris will post a comment.
~ We can come back to this after people have thought more about it.

CCB#50 (Clarify Ambiguous Author Information in Product Document) and CCB#66 (Improve documentation for List Author, List Contributor, and List Editor)

These are kind of related. In CCB#50 the question is in product document there is author related information in two places, but no guidelines about how to use them. Another question is if they are identical if we need it in both places. A possible change to the IM. Definitely a change to the documentation.
~ A comment - this has been discussed at a node. Now have a contributor type, so the need for different use in document versus citation information should go away. If contributor type is properly defined.
Question: Is there any potential use for having two lists of authors? Maybe production of label versus document? But not seeing much value in separate information on who made the labels. That can lead to confusion.
Answer: Someone still thinks contributor type helps with this.
~ Just having one area in citation area makes the whole thing more uniform. Leaning towards option 3 in GitHub.

When this was talked about in the past, someone discussed it with Dick. Asked why he thought it should be in two places. He can't remember. Maybe it came from Todd King. Unsure. Maybe we should ask Anne if she remembers.
Action Item - Joe will ask Anne.
~ It's probably something we need to clean up.

Question: Do people here remember anything? Why author information is duplicated?
Answer: Someone doesn't think citation information was in every product originally. Think that's why document product required author information. Thought it was only in bundle and collection products.
~ Citation information is optional in most products, but required in bundle, collection and documents.
~ That sounds right. Maybe.
~ We will try to figure this out. We will reach out to Anne.

In field reports example (in GitHub) could be different. Maybe contributor list would help.
~ Someone thinks that might be a better way to handle it.

Question: Is citation information commonly used in product observational?
Answer: PPI uses it for every individual product because it is where the product description is.
~ RMS uses it for bundles and collections only, but makes sense in document.

Question: What if we set up a schematron rule that if author information exists in a label then the citation information can't have it?
Answer: That could be done, but would only be for product document.
~ That is less drastic than deprecating stuff.
Action Item - Matt - put that comment in GitHub.

Our time is up.

Our next meeting is October 2.
~ Thomas will try to have a sub attend for him.