

Meeting Minutes from the DDWG-2 telecons:

[DDWG Notes 2022-01-06](#)

[DDWG Notes 2022-01-20](#)

[DDWG Notes 2022-02-03](#)

[DDWG Notes 2022-02-17](#)

[DDWG Notes 2022-07-07](#)

[DDWG Notes 2022-07-21](#)

[DDWG Notes 2022-08-04](#)

[DDWG Notes 2022-08-11](#)

[DDWG Notes 2022-08-18](#)

[DDWG Notes 2022-09-01](#)

[DDWG Notes 2022-09-15](#)

[DDWG Notes 2022-09-29](#)

[DDWG Notes 2022-10-13](#)

[DDWG Notes 2022-10-27](#)

[DDWG Notes 2022-11-15](#)

[DDWG Notes 2022-12-08](#)

DDWG Notes 2022-01-06

title: DDWG Notes 2022-01-06

layout: default

date: 2022-01-06

January 6, 2022

Notes by Debra Kazden

Known Attendees: M. Drum, M. Gordon, E. Guinness, T. Hare, L. Huber, R. Joyner, D. Kazden, T. Lim, J. Mafi, S. McLaughlin, R. Simpson and M. Costa Sitja

Known Observers: P. Lawton

DDWG Agenda

(Included in email sent January 3, from T. Hare, DDWG agenda for Jan. 6th, 2022)

0) Front-matter

- CCB passed CCB-339 -- add Units of Power with SI watts as option
 - Mike Drum selected (in training) as new chair, starting after DDWG hiatus
 - Mitch updates for February DMSP (Data Dictionary) meeting
 - Vivian Tang taking over for Emily Law (EN), CCB coordination (and I assume Jira help). Thanks Emily!
- **(Discussed)**

I. Votes (maybe ready - again?).

2) CCB-335 -- Inventory Specification Allows Too Many Delimiters - - Steve to check out potential conflict.

**(Voted to pass - again.)

II. Status and New SCRs.

0) CCB-340 -- Local Internal Reference.local identifier reference cardinality - Mitch (please see new comments)

1) CCB-342 -- NASA, CODMAC, and PDS Processing Levels for Science Data Sets -- will discuss need from HQ.

2) CCB-336 -- Add a License Information class to the Identification Area - Mark B. (w/ comments)
-- NASA is landing on a CC0 license for data (from Jordan). But we still need to know where this info goes.

(Some Discussion)

2nd tier:

0) CCB-209 -- Correct Definitions of Tagged-*Object -- any updates?

1) CCB-332 - Header definitions are too loose for the validate tool
-- Marc C. to review for possible withdraw -- software issue, status.

2) CCB-324 -- still might need some support from DDWG members before vote (review current status).

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

(Not Discussed)

III. Discussions (order somewhat random, we will get to what we can and is ready to be presented)

ATM: 303

NAIF: 295, 269

CIS/IMG: 209/316, 164, CCB-325 (link) - Support for video as product observational

EN: status on 209 (Steve)

IPDA: 261, 263

PPI: 222, 282, 230

RMS: 314, 210/275, 212

SBN: 222 (helping PPI), 288, issuing new ticket for 211 - archiving XML

(New Ticket for 211 for pre-hiatus)

parking lot:

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

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

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

4) CCB-222 -- Add citation text to Citation Information -- elevate this? head to parking lot?

(Not Discussed)

Next Meeting: Thursday Jan. 20, 2022

DDWG

Front Matter

CCB-339 passed the CCB.

Mike Drum is in training as the new chair.

The DMSP F2F meeting has been moved to the last week in March to get past the Covid peek, so the DDWG hiatus will be later.

~ There was some discussion on how much the group should do. There are opportunities to start assigning things for the DMSP.

~ The organizing committee will have a virtual meeting for two days. Will so dry runs of activities, work on LDD namespaces. Don't see anything for regular tech people until March.

~ Sounds like the hiatus will not be in February, so we can get SCRs in.

~ We have a few months. By the end of March we will go into the hiatus.

There will be a new SCR for product metadata supplemental. The current restrictions prevent us from using it how we intended. RMS would like to put metadata for OPUS in a separate metadata bundle, use the NAIF paradigm and just update that, rather than every bundle that is potentially effected. SCR will ask to reduce some of the restrictions.

~ This will be a tough one. Make it as concrete as much as possible. A topic that can spiral very quickly.

~ Kept scaling the original proposal back.

~ Very important topic. Will be difficult to discuss.

Question: Any comments?

Answer: (Silence)

Votes

CCB-335 - Inventory Specification Allows Too Many Delimiters

See <https://gcc02.safelinks.protection.outlook.com/?url=https%3A%2F%2Fpds-jira.jpl.nasa.gov%2Fbrowse%2FCCB-335&data=04%7C01%7Cthare%40usgs.gov%7C2b6fb897607a406da91908d9db7d4c12%7C0693b5ba4b184d7b9341f32f400a5494%7C0%7C0%7C637782153461135258%7CUnknown%7CTWFpbGZsb3d8eyJWljojMC4wLjAwMDAiLCJQIjoiV2luMzliLCJBTiI6Ikl1haWwiLCJXVCi6Mn0%3D%7C3000&sdata=Yleq5dRbCLWHFklbygvqMib2zlCoD8w9ojhhiXMkmlw%3D&reserved=0>

Got wrapped around the axle on this one. Voted and then yanked it.

Question: Any comments? Should we vote?

Answer: Yes, let's vote.

Another Question: Are people comfortable?

Answer: We voted on this last time.

~ Yes.

~ Then there was an issue. It's okay now.

~ Person is not on the call, but gave agreement.

We voted to pass this last time.

~ Someone thinks we wanted an updated TA. Would like to vote again.

~ There were important questions. If no changes, not sure we need to vote again.

~ Let's just vote again, see if there are any no answers this time.

~ Yes, someone plans to be a no this time. It being non-backwards compatible outweighs the benefit. In the past we tried not to make things non-backwards compatible. Can see big issues for people.

~ Still an inventory. Only issue would be if people were using vertical bars or horizontal tabs.

Question: Inventory is not changing?

Answer: No. It's where it appears - will be parallel to table delimited, but unchanged. This is to prevent use of vertical bars and horizontal tabs.

~ In that case won't change vote.

**The Vote for CCB-335

ATMOS - Yes

IMG - Yes

EN - Yes

GEO - Yes

IPDA - Yes

NAIF - Yes

PPI - Yes

RMS - Yes

RS - Yes

SBN - Yes**

Passed.

Status and New SCRs

CCB-340 - Local Internal Reference.local identifier reference cardinality

See <https://gcc02.safelinks.protection.outlook.com/?url=https%3A%2F%2Fpds->

[jira.jpl.nasa.gov%2Fbrowse%2FCCB-](https://gcc02.safelinks.protection.outlook.com/?url=https%3A%2F%2Fpds-jira.jpl.nasa.gov%2Fbrowse%2FCCB-)

[340&data=04%7C01%7Cthare%40usgs.gov%7C2b6fb897607a406da91908d9db7d4c12%7C0693b5ba4b184d7b9341f32f400a5494%7C0%7C0%7C637782153461135258%7CUnknown%7CTWFpbGZsb3d8eyJWljojMC4wLjAwMDAiLCJQIjoiV2luMzliLCJBTiI6Iik1haWwiLCJXVCi6Mn0%3D%7C3000&sdata=ZK8x%2FN71aZqGx9cEYsyNlvEmVfRk7VMXgT5%2Bci%2F4ZRI%3D&reserved=0](https://gcc02.safelinks.protection.outlook.com/?url=https%3A%2F%2Fpds-jira.jpl.nasa.gov%2Fbrowse%2FCCB-340&data=04%7C01%7Cthare%40usgs.gov%7C2b6fb897607a406da91908d9db7d4c12%7C0693b5ba4b184d7b9341f32f400a5494%7C0%7C0%7C637782153461135258%7CUnknown%7CTWFpbGZsb3d8eyJWljojMC4wLjAwMDAiLCJQIjoiV2luMzliLCJBTiI6Iik1haWwiLCJXVCi6Mn0%3D%7C3000&sdata=ZK8x%2FN71aZqGx9cEYsyNlvEmVfRk7VMXgT5%2Bci%2F4ZRI%3D&reserved=0)

Comments have been added in jira. It was suggested that a prefix could be used instead of wild-cards. That would allow them all to have the same reference. The second idea was a question back - which was answered.

~ A file was attached, to simulate it. Problem is you have to look two places to find display direction. Not as simple as what we have been using. Disinclined to go that route. The way the SCR is written is a better approach. Composite structure is more for the end user than for software.

~ Someone thinks we might just have to vote. Doesn't see the labels as something meant to be human readable, so thinks the structure as it exists now is the cleanest from a model perspective and it makes the most sense programmatically. Will always lean towards modular pieces that can be reused. But I don't work with labels, so if this is hugely inconvenient for them the vote can outweigh my vote.

~ It's too bad that Steve is not here today. This is beyond my expertise. Understand your point, but coming from the other perspective.

~ We have software that can parse labels, the registry. We are an interesting crowd because we are used to looking at XML. Most users glaze over when they look at XML.

~ You are killing me, but that's true.

~ It would be interesting to hear who has software to read labels and who actually reads them.

~ Thought we wanted it machine readable, but not incomprehensible to humans. More concerned about efficiency. Don't think the software will have a problem. TA did say the request is reasonable.

Anne said there is a generational divide here for this. The dinosaurs look at the labels. This will make it easier to look at the labels.

Question: So display settings change from one to one to one to many?

Answer: Yes. Would be one to many. Could reference multiple products.

Another Question: If the local Identifier reference is one to one, could there be one to many too, so we could have both worlds?

Answer: Decent idea, but very human solution. We have the ability in the LDDs to say how many references we want. Same intent behind the schema. I'm thinking about modeling.

~ Computer might be happier reading it once.

~ It comes down to code design.

~ The example is just display settings. It allows us to be as compact as possible. We have a TA. Not sure if we are ready for a vote.

~ Someone is comfortable voting.

~ Someone else is not comfortable voting, but would probably abstain.

~ Another person likes the most compact one.

~ Need more time to review this, but not something I'm likely to ever use.

****Action Item - Everyone**** should read this before we vote. Could vote next time.

There is one unresolved issue. If we remove the constraint, the constraint on LDD becomes useless. Not sure if that requirement can fit in a new solution.

Question: Like in display LDD how many times internal reference allowed?

Answer: Yes. Can just make the class one to many. Consequence would be no longer able to specify how many you want.

~ Good point.

Another Question: Could it be abused anywhere else?

Another Question: Is the only case where local internal reference is required the processing LDD?

Answer: Unsure.

~ It's in the Multi LDD for CDF files.

Another Question: See any possible problems?

Answer: Would have to think about it.

~ That's the homework. Think about where this could be abused. ****(Action Item - Everyone)****

~ Validate will have to make sure all the references exist.

~ Not sure validate has to do it. Think schema does it.

~ Think we use that in the Geometry LDD, but don't think this will cause problems.

~ Tool note is good. We are taking away knowledge by doing this.

~ PDS4 Viewer uses this.

~ GDAL will need updates.

~ Someone likes the argument when it's not visually presented.

Question: Who has taken over PDS4 Viewer?

Answer: Still contracting Lev.

~ We could ask him to chime in on this. Would be interesting.

~ The sample label doesn't have a corresponding label.

~ He could figure it out.

Question: Last comments?

Answer: (Silence)

CCB-342 - NASA, CODMAC, and PDS Processing Levels for Science Data Sets

See <https://gcc02.safelinks.protection.outlook.com/?url=https%3A%2F%2Fpds-jira.jpl.nasa.gov%2Fbrowse%2FCCB-342&data=04%7C01%7Cthare%40usgs.gov%7C2b6fb897607a406da91908d9db7d4c12%7C0693b5ba4b184d7b9341f32f400a5494%7C0%7C0%7C637782153461135258%7CUnknown%7CTWFpbGZsb3d8eyJWljojMC4wLjAwMDAiLCJQIjoiV2luMzliLCJBTiI6Iik1haWwiLCJXVCi6Mn0%3D%7C3000&data=17ay8et4%2BhxNq4c7fN0q4Mlyb4vLyuXNJTi08gArQJ4%3D&reserved=0>

Not sure why this needs to be an SCR. Maybe to babysit it though the path. It was born at an MC meeting when we discussed Malin's issue and terminology. Decided we needed to solidify these. This is a requirement from HQ. Probably on hold until Steve gets back.

Question: Any comments?

Answer: Someone thought about this a little bit. Made a comment on the PDS3 comments. Would help if there's a footnote that says the three letter abbreviations are examples and not used consistently.

Provide examples, like DDR - not always derived.

Another Question: Can you add that to the ticket?

Answer: Okay. ****(Action Item - Ed)****

Question: Anyone still feel strongly about hiding the old levels?

Answer: Footnote could say not to use them.

~ Could take the last two columns and make them the first two, and add the footnote. Maybe group to show what we do now and what's deprecated.

~ Not deprecated. Some are used in other systems. Can put lots of footnotes to be very clear about the terms for PDS4 and what used to be equivalent. Need to decide about telemetry.

~ Think it's in because ESA uses it.

~ Yes, but we never agreed PDS was doing it. Think MAVEN is.

~ IMG was asked to store telemetry. Declined.

~ Storing MAVEN telemetry as Safed.

~ Still not going away.

~ Can have a footnote about who stores telemetry data.

~ It's a complicated question for IPDA. The definitions have it as raw products.

~ It is too late to close that door.

~ Need a footnote that it's not commonly used.

Awkward SCR.

~ If successful in getting revised table, can offer it to the MC.

~ Can offer it to them just as Trent steps aside.

~ Mike would take all the heat.

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

See <https://gcc02.safelinks.protection.outlook.com/?url=https%3A%2F%2Fpds-jira.jpl.nasa.gov%2Fbrowse%2FCCB-336&data=04%7C01%7Cthare%40usgs.gov%7C2b6fb897607a406da91908d9db7d4c12%7C0693b5ba4b184d7b9341f32f400a5494%7C0%7C0%7C637782153461135258%7CUnknown%7CTWFpbGZsb3d8eyJWljojMC4wLjAwMDAiLCJQIjoiV2luMzliLCJBTiI6Ikl1haWwiLCJXVCi6Mn0%3D%7C3000&data=YM1jO%2FkcvxJLHS0spxdX4bHE04JYoamZw3hNCx1QUvw%3D&reserved=0>

This still needs discussion. People need to review backwards compatibility. There's been some discussion. If people have thoughts they can add them to the ticket.

Discussion

The ones that are currently being targeted before the hiatus to wrap up are the ones we discussed today and license information and the new product metadata supplemental one. Don't think we will get to video. IMG needs the product metadata supplemental one, but concerned about it.

~ The SCR will be added to jira. RMS keeps getting users asking for PDS4 equivalent of a PDS3 index.

~ Could put it in an ancillary or miscellaneous collection.

~ Which would be a table.

Question: Why not a document?

Answer: It can be, but wanted tables of metadata. There are lots of workarounds, but don't want to do that.

~ Users won't like it being different node to node. IMG wants this to support the ecosystem and services beyond what PDS supports. Supporting that sort of thing will be important. Takes the ticket way out

there.

~ Also think we need to expand to meet the need to tie to outside resources to enable the archive to transition to the registry.

~ Interesting rabbit hole.

Closing this here. Next meeting will be January 20. We will vote on CCB-340.

~ And CCB-211 is on someone's plate. Will start fresh with that. Probably not nice and easy.

~ Good for pre-hiatus.

Action Item - Mike will issue a new ticket for CCB-211.

DDWG Notes 2022-01-20

title: DDWG Notes 2022-01-20

layout: default

date: 2022-01-20

January 20, 2022

Notes by Debra Kazden

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

Known Observers: P. Lawton and Mia(?)

DDWG Agenda

(Included in email sent January 18, from T. Hare, DDWG agenda for Jan. 20th, 2022)

0) Front-matter

- Susie requested help for a Tiger Team (She or I will introduce). Topic: revisit superseded data policy from 2011

- CCB requested clarification for CCB-335, which we passed last time. See their comments in Jira (Mark B).

- Quick question on who might know any Viking Orbiters' cruise trajectory data.

(Discussed)

I expect we will only get to "II. Status #4". If we need to talk to another issue, let me know.

I. Votes.

none

II. Status and New SCRs.

1) new CCB-344 -- Add data to partially processed product to reference types for Internal Reference - Susie

2) new CCB-343 -- Revise Product Metadata Supplemental - Mitch

3) CCB-340 -- Local Internal Reference.local identifier reference cardinality - Mitch (new comments from Lev, pds4 viewer author)

4) CCB-342 -- NASA, CODMAC, and PDS Processing Levels for Science Data Sets -- request is from HQ.

5) CCB-336 -- Add a License Information class to the Identification Area - Mark B. (w/ comments)

-- NASA is landing on a CC0 license for data.

(Discussed)

2nd tier:

1) CCB-209 -- Correct Definitions of Tagged-*--Object -- any updates?

2) CCB-332 - Header definitions are too loose for the validate tool

-- Marc C. to review for possible withdraw -- software issue, status.

3) CCB-324 -- still might need some support from DDWG members before vote (review current status).

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

(Not Discussed)

III. Discussions (order somewhat random, we will get to what we can and is ready to be presented)

ATM: 303

NAIF: 295, 269

CIS/IMG: 209/316, 164, CCB-325 (link) - Support for video as product observational

EN: status on 209 (Steve)

IPDA: 261, 263

PPI: 222, 282, 230

RMS: 314, 210/275, 212

SBN: 222 (helping PPI), 288, issuing new ticket for 211 - archiving XML

** (Not Discussed) **

parking lot:

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

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

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

4) CCB-222 -- Add citation text to Citation Information -- elevate this? head to parking lot?

** (Not Discussed) **

Next meeting: Thursday Feb. 3, 2022

DDWG

Front Matter

Susie may present on revisiting the superseded data policy from 2011 next time, but hoping to organize a tiger team before that. Maybe they can also attend. Unclear if this is a DDWG topic, but we can introduce it. It has to do with Mars 2020. They have lots of data coming in, and lots of changes. There's a policy on how to handle superseded data from 2011 - wondering how it stands up to PDS4 and our new way of thinking. Anyone interested can participate. This may be something for EN to lead.

Question: Any discussion? Volunteers?

Answer: Seems from this and CCB-343 where we are talking about supplemental metadata changes that we are changing the character of the archive. It's no longer permanent. Becoming very dynamic. We may need the MC to weigh in.

~ Someone agrees, the MC may need to approve of any changes to the policy. For Mars 2020, every time a product type has to be changed they redeliver the entire archive. Not many actual changes besides some dates. Storing tons of files that are basically the same thing. It becomes a question of what's practical. Agrees this is an MC issue. Doesn't want to be on the work group.

~ Someone else also does not want to be on the tiger team, but thinks we did something like this for High Rise because it was changing so rapidly.

~ Someone thinks that was also draft data.

~ Another person draws the line at that.

~ At SBN they have also been considering this. Accumulating bundles and versions. One possible solution is to off-line superseded data, but let people know it's available if they want or need it. Seems to balance the idea that we are an archive, but want people to use the newest version.

~ Could even put it in a deep backup and make users pay. We should document when users ask for

superseded data.

~ ATMOS does what SBN does. Don't recall anyone asking for superseded data.

~ PDS has always has a dichotomy - active archive or long-term. There really are two archives here. Issue is if the community wants historical data, especially as many are moving into DOIs, etc. The community has to trust that they can go back and get the older stuff. We can't betray that trust.

Status and New SCRs.

CCB-343 - Revise Product Metadata Supplemental

See <https://gcc02.safelinks.protection.outlook.com/?url=https%3A%2F%2Fpds-jira.jpl.nasa.gov%2Fbrowse%2FCCB-343&data=04%7C01%7Cthare%40usgs.gov%7C4b2a3cc9a3f04a2ec3de08d9e5db8fd5%7C0693b5ba4b184d7b9341f32f400a5494%7C0%7C0%7C637793553443389096%7CUnknown%7CTWFpbGZsb3d8eyJWljiMC4wLjAwMDAiLCJQIjoiV2luMzliLCJBTiI6I6k1haWwiLCJXVCi6Mn0%3D%7C3000&data=1iel8XB59iPoz1KL7qOrISv0CJegDiFyWscLzSRIS2M%3D&reserved=0>

When the old ticket was in jira, we tightened it so much that we made it impractical to use. Been discussing how to fix that. Propose removing that it can only be associated with one collection and a primary member. Trying to support OPUS. Want to do a bundle with separate collections for each instrument we are supporting. The products wouldn't be primary. Associations would be from metadata. It was pointed out that the changes make this very complicated. It was suggested that we just remove the restrictions.

~ Better off leaving primary and secondary members defined as they are for all other cases.

~ Will make that change in the Requested Changes. Concern is that metadata like this is dynamic. Users may want search parameters added. Purpose for putting this in OPUS bundle was so it could go in the deep archive annually. That allows us to be quick and dynamic.

~ Some metadata is dynamic. At IMG it feels like EN has a tough time trying to search across all the nodes, so now there's help. This is a good start for our users and maybe across nodes.

Question: Any comments?

Answer: One concept we've had a long time for search is that of the query model. So nodes can say what's important for search.

~ Good reminder. Once this is in place probably query model is a good way to let EN know.

~ Yes, but it will be complicated.

CCB-335 - Inventory Specification Allows Too Many Delimiters

See <https://gcc02.safelinks.protection.outlook.com/?url=https%3A%2F%2Fpds-jira.jpl.nasa.gov%2Fbrowse%2FCCB-335&data=04%7C01%7Cthare%40usgs.gov%7C4b2a3cc9a3f04a2ec3de08d9e5db8fd5%7C0693b5ba4b184d7b9341f32f400a5494%7C0%7C0%7C637793553443389096%7CUnknown%7CTWFpbGZsb3d8eyJWljiMC4wLjAwMDAiLCJQIjoiV2luMzliLCJBTiI6I6k1haWwiLCJXVCi6Mn0%3D%7C3000&data=t%2B96%2B730sZ4pwtV%2FJH97vM6hAgxT9zVYZ8yvONJmwzE%3D&reserved=0>

We voted to pass this. The CCB has posted questions in jira from Mark.

~ Someone commented on the questions. Tried to answer, but wasn't entirely sure. Asked others.

~ There's some confusion.

~ Basically, there are two options. Define inventory to be parallel to table delimited, which is what the SCR proposes and we passed, but in modeling can't define an attribute two ways, so the choice is to

make a new attribute, can't use the original field delimiter. Need a new one. Can't use inventory as we used to. Inconvenient for us, users and software. The other way is to leave inventory as subclass of table delimited and restrict the delimiters. Violates rules in modeling, but that solution has a smaller impact. We need to discuss this more. May leave inventory as it was and add some schema restrictions. Read the questions in terms of the original SCR - need to reread the answers, but think we agree with what the CCB has raised.

~ Someone talked to Jesse, on the CCB. Concluded that it wasn't just that they want us to use schematron rule. In number 4 these rules aren't visible in the model, which, I think, is why we wanted to do it in the model. Think they want to address that concern.

~ Unclear which is preferred way to go forward. Maybe we should take off-line to decide if we will follow model principals or do something else.

~ That could change my vote.

~ Makes it a non-backwards compatible change. Should have done this at the start. Hopes the CCB accepts our first answer, which was why do this at all - people should look and see if they agree. Beyond that, unsure what to recommend to the DDWG. Might want to go back to an earlier version.

~ The issue is only in a namespace, in common. Not so much the model. It's implemented in XML schema, so theoretically, it works. Not sure of the solution.

Mike D. will join in the off-line discussion with Dick, Steve, Mitch and Jesse.

This ties a bit to CCB-209, so we can say we discussed it, even though we made no real progress.

Viking Orbiter cruise trajectory data?

Question: Does anyone know if Viking Orbiter cruise trajectory data exists?

Answer: Stef at NSSDCA is scanning through documents to see. Marc might have some.

~ Marc: No, don't. Had a request - data was curated 25 plus years ago. Unsure if the data exists.

Stef will let Marc know if she finds it.

In the late 1990s JPL had a data preservation project for old CDs and such. NSSDCA has that. Includes some Vikings. Also old radio science. Data may be hard to read. Will see what can be found.

CCB- 344 - Add data to partially processed product to reference types for Internal Reference

See <https://gcc02.safelinks.protection.outlook.com/?url=https%3A%2F%2Fpds-jira.jpl.nasa.gov%2Fbrowse%2FCCB-344&data=04%7C01%7Cthare%40usgs.gov%7C4b2a3cc9a3f04a2ec3de08d9e5db8fd5%7C0693b5ba4b184d7b9341f32f400a5494%7C0%7C0%7C637793553443389096%7CUnknown%7CTWFpbGZsb3d8eyJWljiMC4wLjAwMDAiLCJQIjoiV2luMzliLCJBTiI6I1haWwiLCJXVCi6Mn0%3D%7C3000&sdata=AqAhS%2FPNsFCwq1W%2BJ06tQCXeSxr8y1yvZkkn0uGa%2Bi0%3D&reserved=0>

Question: Any issue with adding this enumerated value?

Answer: The point is that we have values for three of four processing levels. We just don't have partially processed.

~ Looks pretty straight forward.

~ Maybe we can vote next time. It ties to what is being worked on for CCB-342.

CCB-342 - NASA, CODMAC, and PDS Processing Levels for Science Data Sets

See <https://gcc02.safelinks.protection.outlook.com/?url=https%3A%2F%2Fpds-jira.jpl.nasa.gov%2Fbrowse%2FCCB-342&data=04%7C01%7Cthare%40usgs.gov%7C4b2a3cc9a3f04a2ec3de08d9e5db8fd5%7C0693b5ba4b184d7b9341f32f400a5494%7C0%7C0%7C637793553443389096%7CUnknown%7CTWFpbGZsb3d8eyJWIjoiMC4wLjAwMDAiLCJQIjoiV2luMzliLCJBTiI6IjEhaWwiLCJXVCI6Mn0%3D%7C3000&data=EZOd%2BXDdxcAigmZ5tgHMKJdW5cy4%2FBJ3xEQ3le8%2BgKg%3D&reserved=0>

Steve did see the comments from last time. Will get to it.

~ Jordan said there's a comment to clarify some of it.

~ One cycle of changes was done, but the new comments are not applied yet.

~ Recommendation was to highlight the PDS4 values. Need a good way to show that and PDS3, but sort of hidden, not so obvious.

~ The idea was to put PDS4 column on the left.

~ And highlight it over the rest.

~ And note PDS3 wasn't always used consistently. Might want to try a few variations. PDS4 only, PDS4 on the left...

~ Someone looked at our current documentation. It says telemetry is not archived in PDS. We are changing it to say it is not commonly archived.

Someone is curious about the original task.

~ It's no longer clear.

~ Someone thinks one confusing item was what to do with the PDS3 lines. Putting "Hey Dummy! don't use this!" in red is probably not okay. Maybe a cross hatch or something to show deprecation.

~ Not really deprecated.

~ Can make a few versions. Issue is people want a mapping.

~ Two things are going on here - translating PDS3 to PDS4 and using words instead of aberrations, but missions are still using the aberrations. Mapping okay, but like the hatch mark not to use these idea.

Question: So, the action item is a few versions of this and bring it back next time. Don't think we need to vote.

Answer: Let's see if we can get any clarification from Jordan.

~ He just wants our recommendation.

~ Recommend you look at a recent SIS for PDS4.

~ There are some Mars 2020 ones with the table. ***(Action Item - Ed - will forward one to Steve)****

CCB-340 - Local Internal Reference.local identifier reference cardinality

See <https://gcc02.safelinks.protection.outlook.com/?url=https%3A%2F%2Fpds-jira.jpl.nasa.gov%2Fbrowse%2FCCB-340&data=04%7C01%7Cthare%40usgs.gov%7C4b2a3cc9a3f04a2ec3de08d9e5db8fd5%7C0693b5ba4b184d7b9341f32f400a5494%7C0%7C0%7C637793553443389096%7CUnknown%7CTWFpbGZsb3d8eyJWIjoiMC4wLjAwMDAiLCJQIjoiV2luMzliLCJBTiI6IjEhaWwiLCJXVCI6Mn0%3D%7C3000&data=vnai7PECjUrLorYw4fu0i6922r9o2kjD%2FouUNXrCujQ%3D&reserved=0>

Lev responded.

Question: Any comments?

Answer: Someone was impressed with Lev's response.

~ Doesn't necessarily solve anything.

~ Someone read it through a lens that he agrees with them, but not sure what next. Sounds like composite structure solution is too complicated. Still in favor of the programmatic solution.
~ Someone's take away was don't do the composite structure.
~ Shouldn't make change to cardinality. Would still have multiple local internal references for LDDs.
~ Thought it said the change was trivial and composite structure not a viable solution. Programming perspective.

Question: So, to clarify his third paragraph?

Answer: Someone read the comment in the label. That hadn't been removed. So that comment is meaningless now. Can ignore that third paragraph.

In the second paragraph, where it points our change would break restrictions on cardinality in LDD, want it added that this is known, but won't be a big deal. Would like to consult with LDD stewards. If this is a non-issue fine, just would like to see that addressed in the SCR.

Someone would like to go back to composite structure. Lev makes good points, but the main concern about this being the wild west isn't true. Problem with the model, once we describe a structure, it's hard to change over time. Composite structure was to have a dynamic way to create new structures based on need. Never got there, but there was a rhyme and reason to it.

~ This is a shiny object. Was never proposed for this SCR or this kind of problem. Unnecessary part of this discussion.

~ Someone disagrees. We implemented a very cut down version, but the original concept was very broad.

~ Someone else is not so scared of composite structure. Sounds like we can move forward.

****Action Item - Mitch**** will send emails to the LDD stewards. Would like ****Mike**** to email him exactly what to ask.

No volunteers to work on the superseded policy.

We will review CCB-335 next time.

We will probably vote on CCB-344 next time.

Next meeting February 3.

DDWG Notes 2022-02-03

title: DDWG Notes 2022-02-03

layout: default

date: 2022-02-03

February 3, 2022

Notes by Debra Kazden

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

Known Observers: P. Lawton and S. Slavney

DDWG Agenda

(Included in email sent February 2, from T. Hare, DDWG agenda for Feb. 3rd, 2022)

0) Front-matter

- Topic: revisit superseded data policy from 2011 - Offering Susie to re-introduce (optional).
- CCB requested clarification for CCB-335, which we passed last time. move back to discussion? (Mike, Dick, Steve, Mitch and Jesse).
- ****(Discussed. Susie attended for this discussion)****

I. Votes (needs T/A).

- 1) CCB-344 - Add data to partially processed product to reference types for Internal Reference
****(Voted to pass 344 and 342)****

II. Status and New SCRs.

- 1) CCB-343 - Revise Product Metadata Supplemental - Mitch (see updates)
- 2) CCB-340 - Local Internal Reference.local identifier reference cardinality - Mitch (Mitch requested feed-back from outside)
- 3) CCB-342 - NASA, CODMAC, and PDS Processing Levels for Science Data Sets -- request is from HQ (please review layout).
- 4) CCB-336 - Add a License Information class to the Identification Area - Mark B. (w/ comments)
-- NASA is landing on a CC0 license for data.
****(Discussed)****

2nd tier:

- 1) CCB-209 - Correct Definitions of Tagged-*Object -- any updates?
- 2) CCB-332 - Header definitions are too loose for the validate tool
-- Marc C. to review for possible withdraw -- software issue, status.
- 3) CCB-324 - still might need some support from DDWG members before vote (review current status).
- 4) CCB-326 - Inconsistent Direction on How to Use Discipline Facets
****(Not Discussed)****

III. Discussions (order somewhat random, we will get to what we can and is ready to be presented)

ATM: 303

NAIF: 295, 269

CIS/IMG: 209/316, 164, CCB-325 - Support for video as product observational
EN: status on 209 (Steve)
IPDA: 261, 263
PPI: 222, 282, 230
RMS: 314, 210/275, 212
SBN: 222 (helping PPI), 288, issuing new ticket for 211 - archiving XML, new 3D OBJ ticket?
(Not Discussed)

parking lot:

- 1) CCB-164 (link) - Display Settings not required for images -- (IMG) - Need to think more about potential updates to Imaging and how Display might fit.
 - 2) RS: 259/281 (see new attachment to CCB-281) (on hold)
 - 3) CCB-261 - Add an optional method for indicating that a product is a member of a larger collection that has a DOI -- Last chance before recommending withdraw.
 - 4) CCB-222 - Add citation text to Citation Information -- elevate this? head to parking lot?
- **(Not Discussed)**

Next meeting: Thursday Feb. 17, 2022. Last chance to vote(?)

DDWG

Front Matter

Susie is here. She brought slides for her presentation on the superseded data policy.

~ The issue is how to manage older versions of revised older data. We have conflicting policies, both written and unwritten. Keep the older versions awhile versus keep older versions forever. The old policies were written for PDS3. Have been operating under the idea that once a product has a LIDVID it exists forever, but that's not documented anywhere. Something needs to be done. The use case that brought this up is Mars 2020. Their image products are re-released with higher VIDs even though the data doesn't change. Saving all the previous versions seems unnecessary and a waster of storage. We need a work group to consider how long we should keep older versions of revised products, if policies need to be revisited and how this will work. A work group is proposed. Seems logical for the DDWG to be the driver for this to consider the questions and make recommendations to the MC. In the meantime, keeping everything from Mars 2020. Maybe there could be a representative form each node, led by EN, but we'll see.

~ Great intro summary.

~ Susie is sending her slides to the DDWG.

~ Someone agrees that EN needs to be integral to this. There are a few volunteers already. Will see if we can find more.

~ Anne would probably want to be involved.

~ Pat will make sure she gets the slides ****(Action Item)****

~ Dick will volunteer, not lead. Seems pretty cut and dried. Surprised Mars 2020 is allowed to do this.

~ Someone agrees. In the policy we need to make it clear that if nothing has changed but the processing date that we should keep the same version.

~ Agreed. We need to revise the policy.

~ Ron is interested, but doesn't want to lead.

~ Joe will volunteer for PPI.

~ Lyle for AMOS.

~ If Steve promises to get the TA in today we could vote with a lien. **** (Action Item - Steve) ****
~ Someone thinks it's fine to go ahead with a lien.

**** The Vote for 344 - with a lien**

ATMOS - Yes

EN - Yes

GEO - Yes

IMG - Yes

IPDA - Yes

NAIF - Yes

PPI - Yes

RMS - Yes

RS - Yes

SBN- Yes ******

Status and New SCRs

CCB-343 - Revise Product Metadata Supplemental

See <https://gcc02.safelinks.protection.outlook.com/?url=https%3A%2F%2Fpds-jira.jpl.nasa.gov%2Fbrowse%2FCCB-343&data=04%7C01%7Cthare%40usgs.gov%7C5bf5437f3d7d4094630908d9f1a26df0%7C0693b5ba4b184d7b9341f32f400a5494%7C0%7C0%7C637806502200254223%7CUnknown%7CTWFpbGZsb3d8eyJWljiMC4wLjAwMDAiLCJQIjoiV2luMzliLCJBTiI6I1haWwiLCJXVCi6Mn0%3D%7C3000&sd=34LNxL%2BHL71VAEpA6XyKTrBPA%2B%2F5QnikK3eE2mEnIE%3D&reserved=0>

There are some updates.

~ Basically, it was pointed out that some complicated stuff was added. Took some out. Needs a TA. Could vote next time.

~ Someone thinks the changes are good, but this has the potential to be under the same superseded argument.

~ The last paragraph will be expanded to be more specific. Think otherwise everything is there that the document team will need.

Question: Any comments? Can Steve do the TA?

Answer: SR references section 9.G.

~ Yes. Restrictions are enumerated there.

Question: More comments?

Answer: (Silence)

CCB-340 - Local Internal Reference.local identifier reference cardinality

See <https://gcc02.safelinks.protection.outlook.com/?url=https%3A%2F%2Fpds-jira.jpl.nasa.gov%2Fbrowse%2FCCB-340&data=04%7C01%7Cthare%40usgs.gov%7C5bf5437f3d7d4094630908d9f1a26df0%7C0693b5ba4b184d7b9341f32f400a5494%7C0%7C0%7C637806502200254223%7CUnknown%7CTWFpbGZsb3d8eyJWljiMC4wLjAwMDAiLCJQIjoiV2luMzliLCJBTiI6I1haWwiLCJXVCi6Mn0%3D%7C3000&sd=BpFtU0fRc1a8ALR6wJk8E9Wamn3zD2ml3GDP%2FgrSY6w%3D&reserved=0>

We have a presentation from Joe, but a brief introduction first.

~ At the last meeting Jesse had concerns about possible impact on LDDs. Sent email to the stewards list. Got eight responses. Comments have been put in jira. Joe had concerns about the impact on the Multi LDD, so that's what his presentation is about.

A few issues came up while reviewing the Multi LDD. This could cause issues for the LDD. This is used for CDF data. CDFs include multidimensional arrays, support data arrays that contain values associated with part of the data. How these relate to each other is captured in the Multi LDD. There are three classes in the Multi LDD to discuss. Primary value, axis value and face value. The issue is that for this LDD there is a one to one association. Allowing multiple instances of local identifier reference would obscure the information intended to be conveyed. Could end up with labels that don't make sense. Not a deal killer, but would be a concern if CCB-340 is to go forward.

So, primary axis attribute has required cardinality of one instance, so only one local identifier reference. That occurs with a lot of things. If very concerned about this can write a schematron rule.

~ Brought it up because other LDDs might have this issue too.

~ Goes along with what Jesse suggested. Having schematron rules to prevent abuse.

~ Yes. Sounds like that's our work around. Think there's potential for confusion.

~ For the Multi LDD, the steward should write it so it can only do the certain thing. Unclear where the problem is.

~ The problem is local identifier is inherited from common.

~ So, someone could use it multiple times, even though that would be silly.

~ Yes, but easier if the validator will tell them it's wrong.

~ We should document all the rules. Thirty years down the road the model will be all there is to guide people.

Question: How many other places are we leaning on schematron? Are we opening Pandora's box here?

Answer: Someone still thinks this is a mistake.

Another Question: To move forward, do we need to include a schematron rule in the SCR? Is it good enough to put a lien for the DMSP to write rules and address how to get them into the LDDs when they get going?

Answer: Someone likes having them provide examples and documentation.

~ Someone else thinks that would help to clarify our perspective.

~ This is a great opportunity for Jesse to do a regression test.

The action items are to update the SCR with specificity with change to the SR, lien to DMSP to restrict cardinality to one in documentation.

Question: Do we need to revisit the TA?

Answer: Yes because now there is an impact to the SR.

Another Question: Vote next time? That might be it for awhile.

~ Someone put some questions into jira to discuss off-line. Trying to understand this.

CCB-342 - NASA, CODMAC, and PDS Processing Levels for Science Data Sets

See <https://gcc02.safelinks.protection.outlook.com/?url=https%3A%2F%2Fpds-jira.jpl.nasa.gov%2Fbrowse%2FCCB-342&data=04%7C01%7Cthare%40usgs.gov%7C5bf5437f3d7d4094630908d9f1a26df0%7C0693b5b44b184d7b9341f32f400a5494%7C0%7C0%7C637806502200254223%7CUnknown%7CTWFpbGZsb3d8ey>

JWljoiMC4wLjAwMDAiLCJQjoiV2luMzliLCJBTiI6I6k1haWwiLCJXVCI6Mn0%3D%7C3000&data=XeqzXMcgvhUqRTf%2BFTR2QhvwSQiiQVzsuYRO1k%2Fn%2BtU%3D&reserved=0

Someone really appreciates all of the editing. The group has done a good job. Need everyone to take a look. Think we are getting as close as we can.

~ It's unclear if we need to vote on this.

~ A vote would be a way to formally recommend it to the MC.

Question: Which version are we leaning towards?

Answer: The Word document.

~ It needs to go to the MC because we made a change in the definitions.

Question: The MC needs to approve the changes?

Answer: Yes.

~ Needs to go to them anyway. This would go with the policy document to them. This assumes the changes would be put in the policy.

We can vote with a lien that the policy be updated.

~ No objection. No need for a TA.

~ Someone thinks it looks great.

~ Someone else agrees.

~ We will vote next time.

~ It shouldn't take too long.

~ Could vote now.

~ Should add as resolution to make sure the policy needs to be changed too.

~ So we should vote next time. Would like to get this to the next MC.

~ Could do an evote.

~ The next MC is next Wednesday. Propose we vote now. For the MC we could provide the table and policy.

~ But that policy is referenced in one of the footnotes.

~ That's fine.

Question: Does it have to go to the MC in February instead of March?

Answer: There's an action item on EN. Would like to vote today to recommend with the caviots.

~ We can write something for EN to take to the MC.

****The Vote**

ATMOS - Yes

EN - Yes

GEO - Yes

IMG - Yes

IPDA - Yes

NAIF - Yes

PPI - Yes

RMS - Yes

RS - Yes

SBN - Yes**

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

See <https://gcc02.safelinks.protection.outlook.com/?url=https%3A%2F%2Fpds-jira.jpl.nasa.gov%2Fbrowse%2FCCB-336&data=04%7C01%7Cthare%40usgs.gov%7C5bf5437f3d7d4094630908d9f1a26df0%7C0693b5ba4b184d7b9341f32f400a5494%7C0%7C0%7C637806502200254223%7CUnknown%7CTWFpbGZsb3d8eyJWljiMC4wLjAwMDAiLCJQIjoiV2luMzliLCJBTiI6I6k1haWwiLCJXVCI6Mn0%3D%7C3000&odata=AJdC5lhLnEOn0Vav6NZi2vRTzjBR7qhMvOAawfNOnwY%3D&reserved=0>

Nothing new on this.

Discussion

Question: Is there anything we need to start addressing? We have one last meeting in February. No new real issues yet. Anne is asking maybe to bring up an issue for OBJ files.

Answer: OBJ doesn't seem like a file extension that should be reserved. Not sure it's a good idea.

~ My reaction was asking if we really need to reserve it. Describing in in a PDS4 label is really the key.

Unclear if needing to reserve an extension is the issue. Unclear. Maybe if she wants to move forward she could clarify her worries on this. User guide for 3D formats would be great.

Open Floor.

~ (Silence)

Our next meeting is our last chance to vote. We will vote on three. Probably the last meeting for a few months.

Next meeting February 17.

DDWG Notes 2022-02-17

title: DDWG Notes 2022-02-17

layout: default

date: 2022-02-17

February 17, 2022

Notes by Debra Kazden

Known Attendees: M. Drum, M. Gordon, E. Guinness, L. Huber, S. Hughes, R. Joyner, D. Kazden, T. Lim, J. Mafi, S. McLaughlin, R. Simpson and M. Costa Sitja
Known Observers: P. Lawton

DDWG Agenda

Included in email sent February 15, from T. Hare, DDWG agenda for Feb. 17th, 2022 (last one before hiatus)

0) Front-matter

- recall we passed CCB-344 and CCB-342 last meeting.
- Mike Drum will be leading meeting (thanks Mike). Trent will e-vote

I. Votes :

- 1) re-vote CCB-335. Please review new comments - maybe a quick discussion prior to voting.
- 2) CCB-343 -- Revise Product Metadata Supplemental. getting T/A.
- 3) CCB-340 -- Local Internal Reference.local identifier reference cardinality. Last chance for a quick discussion (see new comments).

(Voted on all three)

II. Status and New SCRs.

- 1) CCB-336 -- Add a License Information class to the Identification Area - Mark B. (w/ comments)
-- NASA is landing on a CC0 license for data.

2nd tier:

- 1) CCB-209 -- Correct Definitions of Tagged-*Object -- any updates?
- 2) CCB-332 -- Header definitions are too loose for the validate tool
-- Marc C. to review for possible withdraw -- software issue, status.
- 3) CCB-324 -- still might need some support from DDWG members before vote (review current status).
- 4) CCB-326 -- Inconsistent Direction on How to Use Discipline Facets

(Discussed)

III. Discussions (order somewhat random, we will get to what we can and is ready to be presented)

ATM: 303

NAIF: 295, 269

CIS/IMG: 209/316, 164, CCB-325 (link) - Support for video as product observational

EN: status on 209 (Steve)

IPDA: 261, 263

PPI: 222, 282, 230

RMS: 314, 210/275, 212

SBN: 222 (helping PPI), 288, issuing new ticket for 211 - archiving XML, new 3D OBJ ticket?

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

parking lot:

- 1) CCB-164 -- Display Settings not required for images -- (IMG) - Need to think more about potential updates to Imaging and how Display might fit.
 - 2) RS: 259/281 (see new attachment to CCB-281) (on hold)
 - 3) CCB-261 -- Add an optional method for indicating that a product is a member of a larger collection that has a DOI -- Last chance before recommending withdraw.
 - 4) CCB-222 -- Add citation text to Citation Information -- elevate this? head to parking lot?
- **(Very brief mention)****

Build 12.1, end of Feb. we are locked-down

DDWG - Last Meeting Before the Hiatus

Mike Drum will be leading the meeting today. Trent will e-vote.

I. Votes

There are three issues up for a vote today. We will go least to most controversial.

CCB-335 - Inventory Specification Allows Too Many Delimiters See <https://pds-jira.jpl.nasa.gov/browse/CCB-335>

Voting on this again.

Question: There were questions in the comments. Have they been resolved?

Answer: Someone believes they are. The last comments were endorsements. Limits field delimiter to comma. It's what everyone wants.

~ With schematron rules.

We will vote if there are no objections.

~ (Silence)

****The Vote to pass CCB-335:**

ATMOS - Yes

EN - Yes

GEO - Yes

IMG - Yes (e-vote)

IPDA - Yes

NAIF - Yes

PPI - Yes

RMS - Yes

RS - Yes

SBN - Yes******

CCB-340 - Local Internal Reference.local identifier reference cardinality See <https://pds-jira.jpl.nasa.gov/projects/CCB/issues/CCB-340>

This is the local internal reference cardinality one. There has been some discussion and clarification.

Question: Does anyone have any more comments?

Answer: Nope.

Another Question: Vote?

Answer: (Silence)

**The Vote to pass CCB-340:

ATMOS - Yes

EN - Yes

GEO - Yes

IPDA - Yes

IMG - Yes (e-vote)

NAIF - Yes

PPI - Abstain

SBN - Abstain** (Not comfortable, but don't want to block it.) **RMS - Yes RS - Abstain**

Three abstains.

CCB-343 - Revise Product Metadata Supplemental

See <https://pds-jira.jpl.nasa.gov/projects/CCB/issues/CCB-343>

This got a T.A. and there has been some discussion.

~ Someone says we are ready to vote.

Question: Does anyone want to raise any concerns?

Answer: If everyone has read the comments we can vote. Think this undermines what we originally wanted to do and undermines the original intent. Will vote no, but we can vote.

~ It's a tricky problem. Can see it needs to change if going to do what RSM needs and seems to follow along with what EN needs. Excited about the changes, but there's been a lot of scope creep. Ties into bigger discussion of what PDS4 is for. We could discuss this or just vote.

~ Valid goals, but don't think this is the way to do it. Maybe should pick a different product if going to do it different.

Question: Comments?

Answer: Some of the Jira comments are only 38 minutes old. Not comfortable voting. Probably would abstain.

Another Question: Does anyone else feel strongly? A lot of this is new, but pointed out incentives for metadata to not get updated along side the supplemental might cause problems for users.

~ Yes. No requirement that updates be anywhere near the products.

~ The problems is that there is still no working registry. As long as we have the references, we should have the associations. When raw comes in first, then browse or calibrated products are made, the raw products aren't revised to point to them, just the browse or calibrated point to the raw. OPUS or nodes would let users know there's additional metadata. Have to provide the associations. Would like us to vote to get this in the next build cycle.

Question: Could we pass and keep some restrictions until we have a registry?

Answer: Pretty unworkable.

~ Agrees.

~ Hamstringing it.

~ From a modeling perspective, the link is there. Registry will turn it into something useful. Nodes could take care of the hole until the registry is up to speed, but the model is fine.

~ The intent behind this is to supplement the data. The actual data doesn't change.

~ Whether this is important or not is subjective.

CCB-192 (Add Product Metadata Supplemental, see <https://pds-jira.jpl.nasa.gov/projects/CCB/issues/CCB-192>) was the original SCR. This came up because RMS has been generating enhanced geometric metadata. For some instruments the geometry used to be based on predictions. RMS uses the most recent SPICE info. They don't update the labels, but do point to the enhanced metadata. PDS4 registry will do that. This is to get the users the best values/metadata. Capability within the model is there. Until the registry is there, nodes point users to the best information. GEO also produces better metadata.

~ Maybe for the analyst notebook. Not sure for Rover locations.

Question: So, it seems like we have some clarity. Is it okay to vote?

Answer: (Silence)

**The Vote to pass CCB-343:

ATMOS - Yes

EN - Yes

GEO - Yes

IMG - Abstain (e-vote)

IPDA - Abstain

NAIF - Abstain

PPI - Abstain

RMS - Yes

RS - No

SBN - Yes**

We'll see what the CCB does with that.

II. Status and New SCRs

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

See <https://pds-jira.jpl.nasa.gov/projects/CCB/issues/CCB-336>

Question: Any new updates?

Answer: Trent was carrying it along.

~ We will leave it.

Question: Any updates? This is the last DDWG until after the DMSP.

~ Data Modeling Standardization Project.

The DMSP meeting will be the last week in March. The DDWG will be on hiatus for a few months.

CCB-209 - Correct Definitions of Tagged-*Object

See <https://pds-jira.jpl.nasa.gov/projects/CCB/issues/CCB-209>

No update. Will try to get something in the next couple of months.

CCB-332 - Header definitions are too loose for the validate tool

See <https://pds-jira.jpl.nasa.gov/projects/CCB/issues/CCB-332>

Comments were added in October that suggested an update to the SR. Need feedback. Could withdraw.

Someone said they hadn't looked at it in a while and asked if anyone else had.

~ Another person asked for a summary of the issue.

~ Software issue though a mistake concerning the definition of header. Found a bug in validate tool.

That was solved. Can provide a summary next time.

This is not time sensitive. We can let it spill over to next time.

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

See <https://pds-jira.jpl.nasa.gov/projects/CCB/issues/CCB-324>

Need a flag bearer for this.

~ Someone was the flag bearer, but was going up the wrong hill. Trying to be a purist and wanted a proper data type, but won't fall on a sword for this.

~ We can revive this later.

~ It's fine to let this ride.

We will revisit this after the hiatus.

~ We have a number of SCRs that have been approved and issues too. Fine.

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

See <https://pds-jira.jpl.nasa.gov/browse/CCB-326>

Someone doesn't remember this, so it isn't urgent yet.

~ Anne needs to be involved in this one.

~ Mike can bring it up to her and have a discussion. ****(Action Item)****

~ Let her know it's rated urgent.

Parking Lot

There are issues in the parking lot that are assigned to people. Something to look forward to.

Last Comments

Question: Any comments before we wrap up?

Answer: Would like to get Danny Kahan onto the mailing list for this meeting when we re-start. Not sure who to contact.

~ Vivian tang.

Question: What do we do if something critical comes up?

Answer: We can have an emergency telecon. The idea was to give the DMSP members a chance to focus on that. There is homework for everyone. People can email Mike if we need to spin up a meeting.

~ Maybe a subset of the group can work on emergencies.

~ Will see if we have emergency meetings every two weeks.

~ Mike will be in touch with everyone.

Congratulations to Mike for a good start. A forty-five minute meeting and three SCRs voted on.

DDWG Notes 2022-07-07

title: DDWG Notes 2022-07-07

layout: default

date: 2022-07-07

July 7, 2022

Notes by Debra Kazden

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

DDWG Agenda

(Included in email sent July 6, from M. Drum, DDWG Agenda for Thursday, July 7th 2022)

Front-Matter

- DMSP update from Mitch
- New DDWG season - Approved SCRs by 8/26/22
- Any new issues that teams want to bring up? ****(Discussed)****

Ready for Vote

(None this week)

Tier 1 issues

- CCB-348 – Add Units of Density as a unit of measure – Trent Hare
 - CCB-336 – Add a License Information class to the Identification Area – Tanya Lim
 - CCB-347 – Importance of Reserved File Name Extensions vis a vis Other Requirements – Dick Simpson
 - CCB-211 – XML/schema based files as archival data – Jesse Stone, Mike Drum
 - CCB-349 – Rename Dictionary to Submodel – Mike Drum
- **(Discussed)****

Tier 2 issues

- Higher order products (CCBs to be created) – Lyle Huber
- CCB-345 – Target Identification attributes name and type allow nillable – Trent Hare ****(Some Discussion)****

Parking lot – need a driver!

- CCB-209 – Correct Definitions of Tagged-*Object
 - CCB-326 - Inconsistent Direction on How to Use Discipline Facets
 - CCB-324 – Create a data type for Type attributes that require a permissible value list.
 - CCB-164 - Display Settings not required for images -- (IMG) - Need to think more about potential updates to - Imaging and how Display might fit.
 - RS: 259/281 (see new attachment to CCB-281) (on hold)
 - CCB-261 - Add an optional method for indicating that a product is a member of a larger collection that has a DOI. Last chance before recommending withdraw.
- **(Not Discussed)****

Next meeting: 7/21/22

DDWG

Front Matter

Everyone is welcomed back. Mike Drum is the chair now. The goal for the new DDWG season is to get new SCRs approved by August 26. That gives us 3 or 4 meetings. The SCRs are beginning to be sorted into Tier 1 and Tier 2. Some of them have been discussed with EN. The big ones we will be looking at are archiving XML schema and licensing, but there are others to look at too. The parking lot SCRs need drivers. We probably won't get to them until someone gets into the car, picks up the flag and drives off. They need some effort.

DMSP Update

Some of the DDWG members were at the DMSP close out meeting last week. Everyone's participation, effort and the knowledge they brought were appreciated. There used to be about six people who understood LDDs and now there are about 20 people. Got a lot done, but not everything. One issue is the term dictionary. Submodel is a possible alternative for that term. It's the least onerous. CCB-349 (Rename Dictionary to Submodel) will affect the IM, SR and DPH and might not be approved. Not as much progress on documentation was made as was hoped, but did decide what should be present, including a new handbook for stewards. Focused on the output of LDDTool and changes. There were seven teams, each with 2 or 3 discipline dictionaries. Stewards were not on the team reviewing their dictionary. All received recommendations. The plan was for assessment teams to help the stewards fix issues, but not all made the build. The DMSP officially ended, but some people are interested in continuing the work. Ron offered to help stewards with schematron rules and advice. Looked at issues including cross referencing, focusing on classes. Need to implement the solutions that were come up with. The final report and lessons learned will be discussed at the MC F2F. One key to running a major project like this is to have Anne on the team.

Question: As the larger issues, someone doesn't remember all the solutions. Will we be addressing things like cross referencing and regression tests here?

Answer: Discussed that. Using the registry. It needs to be put in place by EN. Don't think the DDWG needs to be too involved yet. LDDTool needs to be updated and we need the registry. We are much further off on regression testing. EN needs to look at a way to compare changes from older versions, an automated review to look for non-backwards compatible changes.

~ Someone still has some updates to make for the handbook. May ask for help.

Another topic is documentation. Maybe some could be harvested out of the LDDs. Wonder if we have defined standards for each submodel.

~ It needs to be written up.

~ Someone is unsure. Jesse showed some automatic stuff, but we probably need manual updates too. Not sure if we landed on an answer about automating.

~ Thought we agreed it would be on GitHub, but need documentation.

~ What Jesse generated is available, but hesitate to say it can all be automated. Some is more complicated. It's something for us to help support.

~ Someone would like to help on the document team. Has some template ideas.

Tier 1

CCB-348 - Add Units of (Mass) Density as a unit of measure See <https://pds-jira.jpl.nasa.gov/browse/CCB-348>

This is for a need from OReX. Doing it with density allowed in the labels. Need units of density class. Unclear if we should go beyond SI density and there was a comment about mass density and others, like number density. Unsure if we want these in or not. Probably should allow for the main values.

Question: Would that fit under units of mass density?

Answer: Might need a different one. No need to put it in until someone needs it.

~ PPI uses it more in data fields.

****Action Item - Trent**** will update the jira comments.

Question: Any other comments on this? Answer: (Silence)

~ This needs a TA. All in the class.

~ Will need to make sure it works.

~ Someone hasn't read it yet. Wants to read it before we vote.

Another Question: Based on the comments, do we want to make it units of mass density? Then we could add units of number density later.

Answer: Units of surface density is used in rings.

~ If there is more than one kind of density we should make this more specific.

~ Reporter likes that idea. Will change it.

Steve will do the TA and get it ready for a vote next time ****Action Item****

Question: Could it be units of volume mass density? If that's not too cumbersome?

Answer: Long name. Most people would think of volume mass.

~ Someone is not worried about the length of the name. We don't want ambiguity down the line.

~ Reporter is happy to do whatever.

~ Maybe the suggestion should be added in jira.

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

We don't have Mark anymore, but someone else could be the flag bearer for this.

~ Tanya will do it.

~ Someone thinks the issue comes down to where it fits.

~ Someone else had some strong opinions on this. Not sure if people want to discuss it now. The issue is where it goes in the label or if it actually goes in the label. Anne was concerned about putting it in the label. Might be better to have a pointer.

~ It might be good to allow for both in case PDS and PSA want to do different things. A lot of PSA users just download a few files, so it's easier for them if it is in the data files. Also, this was an optional proposal.

Question: Who all is interested in making this happen?

Answer: Mike, Tanya, Trent, Jordan, Anne.

~ **Action Item - Mike** will spin up a conversation off line to come to a consensus.

~ If it's an optional class with optional attributes then it could be up to individuals which granularity to use.

~ This is a case where enumerations scare me.

Another Question: Different types of licenses?

Answer: Yes. Maybe this should be free text.

~ Could be an issue when people at places like JAXA need to add things to an enumerated list.

So, Tanya is the stuckie.

~ The original class was put in jira. Just want to extend it to add a pointer.

~ Anne really extended this with provenance, DOIs, etc.

~ Someone is seeing some of this as separate.

~ Good to keep it simple and optional.

~ Someone talked to Anne about DOI metadata, but agrees we can consider that separately. ~ That deserves it's own issue.

Action Items - Tanya will add a comment about pointers in jira. **Tiger Team** will discuss this.

Question: So, in label or a pointer or both?

Answer: Will see if there are more ideas in Tiger Team too.

~ It could work to say a pointer points to the preferred version.

~ There is usually a latest version. Unclear on licenses.

Another Question: The pointer would go to another PDS product?

Answer: Not sure. It could be a file in the bundle.

~ Maybe a product document, so we could manage changes over time.

~ That would probably work fine.

~ Unclear on using a pointer and internal reference. Internal reference already is in the class, so unsure what we would add.

~ Someone is concerned about name and type, and changes over time. With pointer could make it a LID.

~ Someone is confused. If it's a product in a bundle, don't understand how we would use the internal reference.

~ Maybe we would, but have all the attributes optional. Will discuss it with Anne.

~ This implies that EN will have a bunch of license documents to maintain.

~ There could be one that a DP includes in their bundle.

~ Think about this for software too.

Someone is starting to think Anne's comments and licensing are a different scope outside the archive. Things could be licensed a different way and license could be outside the archive. Scary to be very specific about licenses inside the labels. Will organize thoughts.

~ Someone still thinks the proposed class covers all this.

~ Someone else likes the proposal that can use either, but not both. Likes the idea of access rights in documents and referenced.

PSA is under some pressure from ESA lawyers to attach a license to all their data.

~ It might be worth asking if can add a license file every time they download a file.

~ There are cases where a journalist only downloads an image. Thinks it would be good to just have a pointer to the license. Full license text isn't the plan for this class.

Question: Are there any examples?

Answer: The licensing information is part of the archive, but it can be referenced. Has to be part of archive and go in the deep archive.

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

Jesse has been asked to present this next time. It's a proposal for anything schema based to be archived and to supply the schema.

CCB-349 - Rename Dictionary to Submodel See <https://pds-jira.jpl.nasa.gov/browse/CCB-349>

This was in the DMSP conversation. A lot of implications for the re-name.

Action Item - Everyone Please add comments in jira.

CCB-347 - Importance of Reserved File Name Extensions vis a vis Other Requirements
See <https://pds-jira.jpl.nasa.gov/browse/CCB-347>

The problem was in reviewing an inventory file, noticed it came with a .TAB extension, which has to be fixed width. Validate didn't report errors, so put the issue in GitHub. EN didn't know how to deal with the issue and Jordan suggested the DDWG discuss this. There's a general question about what we mean by reserved file extensions. Think CSV is the way to go for inventory files.

Action Item - Everyone Please put comments in jira. We will discuss this next time.

Tier 2

Higher order products (CCBs to be created)

There's a new issue. There is a long term IMG node annex where products that aren't PDS compliant are put. PPI has one too, and ATMOS is in the process of standing up an annex. The issue is that we should have a way to extend the IM so annex data can be found. Need to be careful that it's known it is annex data. If in the annex it is not in the archive. We are talking about new products for the annex, like product simulation or product GIS. Want to respond to PDE recommendations. Want people to start thinking about this. Want people to come up with some SCRs in the future.

~ A comment. PDS needs to define science ready and derived. Highly related to where a class is and what users should use first. We need to be able to describe that.

~ Agreement. In primary results summery we probably need a new enumerated value like model output or something.

~ If we can solve the definition of derived that would be good. Annex has higher order products then PDS allows.

Eric's favorite keyword is derived. SBN uses it a lot. Planning a proposal on this.

~ IMG, PPI and ATMOS discussed this. Lynn is working on five possible SCRs. 1) Target types for context types for simulated. 2) Processing levels beyond derived. 3) Product that is not product observational, like product simulation or model. 4) Fix primary results summery. 5) Possible submodel to handle additional classes. IMG, PPI and ATMOS discussed this earlier this week.

~ Mike will let Eric know. Maybe he can help.

Someone is wondering if this is really different products or if different formats.

~ Both. Some of the products are in formats we don't recognize. A second issue is some are simulations and don't belong in the archive because they are always being improved.

~ So this is a way to point to things outside the archive.

~ We want them found.

~ We want them find-able.

The IRB and PDE pointed out that PDS isn't the end all for planetary. Need these to be discovered.

~ Someone is confused. If talking about products that don't have a hook, that can be worked on. If not PDS4 format that's a different issue. So how to connect to PDS is a PDE problem to discuss with Moses. ~ Not always.

~ Implications across other disciplines.

~ The important thing is to get a hook in so things can be registered and people can find them. That would be a win for all of us.

We are ten minutes over.

~ There is an MC F2F next week. Trent will make the DDWG report.

We will meet again in two weeks - on July 21.

DDWG Notes 2022-07-21

title: DDWG Notes 2022-07-21

layout: default

date: 2022-07-21

July 21, 2022

Notes by Debra Kazden

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

DDWG Agenda

(Included in email sent July 19, from M. Drum, DDWG Agenda for Thursday, July 21 2022)

Front-Matter

- 1) New issue: CCB-350 (Tier 1)
 - 2) New addition: DOI-rich metadata class (Tier 2)
 - 3) Approved SCRs by 8/26/22
- ** (Discussed) **

Ready for Vote

- CCB-348 – Add Units of Mass Density as a unit of measure – Trent Hare
- ** (No vote) **

Tier 1 issues

- CCB-336 – Add a License Information class to the Identification Area – Tanya Lim
- CCB-347 – Importance of Reserved File Name Extensions vis a vis Other Requirements – Dick Simpson
- CCB-211 – XML/schema based files as archival data – Jesse Stone, Mike Drum
- CCB-349 – Rename Dictionary to Submodel – Mike Drum
- CCB-350 – Adopt more rigorous / stringent rules for leap seconds and use of "24:00:00" in datetimes. – Ron Joyner, Dick Simpson ** (Some Discussions - Presentation on 211) **

Tier 2 issues

- Higher order products (tickets to be created) – Lyle Huber
- CCB-345 – Target Identification attributes name and type allow nullable – Trent Hare
- DOI-rich metadata into IM (ticket to be created), replacing CCB-222 and CCB-261 – Anne Raugh and any volunteers ** (Brief Discussion) **

Parking lot – need a driver!

- CCB-209 – Correct Definitions of Tagged-*Object
- CCB-326 - Inconsistent Direction on How to Use Discipline Facets
- CCB-324 – Create a data type for Type attributes that require a permissible value list.
- CCB-164 - Display Settings not required for images -- (IMG) - Need to think more about potential updates to Imaging and how Display might fit.
- RS: 259/281 (see new attachment to CCB-281) (on hold)
- CCB-261 - Add an optional method for indicating that a product is a member of a larger collection that has a DOI -- Last chance before recommending withdraw.

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

Next meeting: 8/4/22

DDWG

The meeting started late. Had to use Zoom instead of WebEx. People were unable to join the meeting with WebEx.

Front Matter

There have been a few additions since last time.

CCB-350 - Adopt more rigorous / stringent rules for leap seconds and use of "24:00:00" in datetimes.

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

Someone has already commented.

~ The problem is that datetime patterns are not picking up what needs to be picked up. Want us to look at the rules and get the pattern. Bug fix. One issue is leap seconds, another is allowing 24 hours, 0 minutes and 0 seconds. May be a few others. We need patterns that match what we think should be in there.

~ There have only been a couple leap seconds in the history of the world. PDS has an authoritative list of them. Anything not on that shouldn't be allowed. 24:00:00 is an abomination and should never happen.

~ NAIF can't speak to 24 hours, never accepted that, but ISO standard did. Unclear if we need to accept and change.

Question: Didn't SPICE just implement a new time?

Answer: No. Never accept 24 hours. This has come up in the past in PDS4 development. No decision was made. It's unclear what was implemented.

Another Question: If we don't accept 24 hours and someone submits it, do we just change the date?

Answer: If it's in already under the old standard. This is going forward.

~ Validate tool version 1.15 already won't accept things that were okay back in version 1.11.

~ Not perfect.

~ Validate tool can be updated for all that is described for future leap seconds, just not sure if easily implemented.

~ Someone didn't realize that 24:00:00 was a historical option, but prefers not to allow it. Migration cases could be fixed.

****Action Item - Dick and Ron**** will work to clear up this SCR.

~ The description of the SCR should include other times.

~ Totally ridiculous. Leap seconds should only be at the end of June or December.

~ Should only have leap seconds on dates defined in the leap second kernel.

CCB-348 - Add Units of Mass Density as a unit of measure See <https://pds-jira.jpl.nasa.gov/browse/CCB-348>

We will skip voting on this today because people are missing.

This got a TA.

~ The only thing to discuss is if the name can go forward.

~ Someone suggests volume mass density, but thinks mass density is fine.

~ Also, can call out surface or number density.

Question: In the description it was for density for Bennu, then in the assessment something else.

Curious about what kind of data products usage is for?

Answer: Reporter can't remember, but will put a better description in. Basically for derived products.

Generally seems like it would be in a document or paper.

~ Someone else had that confusion too.

Action Item - Trent will find out.

New addition: DOI-rich metadata class (Tier 2)

This is something new from Anne. Provenance related metadata. New SCR to replace CCB-222 (Add citation text to Citation Information) and CCB-261 (Add an optional method for indicating that a product is a member of a larger collection that has a DOI).

~ Haven't reviewed all the tickets yet to see if others mention DOIs.

~ Might want to make sure Anne and Bea are aware the DOI work group is working on this. Should probably work together.

~ **Action Item - Pat** will make sure Anne knows to talk to Bea.

Tier 1 issues

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

A conversation was kicked off with a few people. Outlined the issues, refined initial suggestion. Will stay off-line until there's more consensus. A comment was added.

~ We need to be careful about using the official license name. Can't use PDS license as a literal term.

~ This refers to the example in jira. Still some debate to have on this one. If licenses never change then it's helpful to have in the label.

~ Doesn't seem like a problem if license changes, could have a link to an update page. Don't see license being updated that often.

~ the concern is license being picked up in the wild without pointer to update. Might not be a real concern. If anyone else wants to be involved in the conversation let Mike know.

~ License information for any bundle needs to be part of the archive and goes in the deep archive.

Question: Do you mean full license legal text as a product or distributed with the archive?

Answer: Think of it like a context product. With a context product there has to be something someone can read in 25 years. Need to decide, but has to be kept forever.

~ Someone also had an idea to have products with LIDs.

~ Or it could be documents in a collection.

~ There are licenses that are required to be included with stuff.

~ We are expecting a license for PDS, another for ESA, not expecting it to be like in the software world. Hopefully.

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

Jesse Stone is here to present on this SCR today. He has slides.

~ Presentation is Archiving XML Data.

Started with use cases/background. Things like XML or JSON might need to be archived. Question is if can be described with the current IM. The answer is not really without hacking it up. The other option is to archive as text files. That preserves the data, but doesn't tell us much about it, so that is not a good solution.

The next question is how can we describe XML data. XML has schema to describe it. We relate data to schema with a label that indicates it's XML and has certain classes. Schemas include certain information. To describe schema we would need more information.

Question: Can you archive a style sheet with it?

Answer: Could probably describe it with some level of standardization. Interesting question. Can describe JSON, YAML and others. From EN, we would need capability to run the schemas and validate them. Can already do it for XML.

Some use cases were included on the slides (slide 12).

~ Another example is that the InSight seismometer has 2 types of XML files. And for the catalog, some schema included in the archive in the schema collection.

~ People will look at that to see if there's anything to include or adjust in this SCR.

~ The XML files are modelled as text objects for now. Labels have a kludge. Glad this SCR is being done.

Question: A fan of this, but why is XML better than CSV or TAB?

Answer: XML will put in metadata or values in a single field in a rigorous way. A CSV would need to kludge that in.

~ Array strings with field 1, field 2, etc. Likes the explanation that arrays don't work well.

~ This is covered in slide 3, plus some people don't want to convert.

~ This goes to the question if PDS should allow other ISO formats.

~ We've carved out exceptions for SPICE and GeoTiff.

~ GeoTiff has the building blocks to be PDS4 compliant.

There is a question in chat if schema could be proprietary and not allowed to be archived. If someone has an XML file from a propitiatory schema.

~ Interesting problem.

~ The latest version of Excel is XML. That could be horrifying in PDS.

~ Or powerful. Depends on your point of view.

~ Good question. People could shove Excel files in and claim that they are XML. We may need some limits.

~ We could have examples.

~ Someone never considered describing an XML file in a label.

~ There are XML Word files too.

~ We may need to consider if the format is open or not.

~ Not sure if PowerPoint is allowed by PDS.

~ Not sure how we would put bounds around this.

~ We could say they are just text files, and could say schema needs to describe the data. Unclear.

Question: Do things like Word have external dependencies?

Answer: Good question.

~ Some of the more proprietary formats might not be pure XML.

~ Just unzipped this PowerPoint presentation. It gave 13 XML files and a big structure.

Thank you to Jesse for the presentation.

Maybe some of this is a red herring.

Question: Any comments? Anyone want to cry heresy?

Answer: Someone wants to cry heresy, but it's time to wrap up.

~ Someone else wants to quickly respond. This is a long time coming. Likes this, but would like to point out that we are talking about complex structures and we have composite structures. That's a possible solution to this. Composite structure was implemented in a very constrained way, but it's a way to include schema. It describes the components and how they relate. For JSON, just describe what you have in the composite structure. The devil would be in the details to define the components and get that in the model. The point is that there is already a native way to do this. Would be happy to help provide examples.

~ Someone agrees that's a solution, but maybe not the best solution because it would have to be very tailored to every data product. We want something we can re-use.

~ Always the concept of giving DPs the ability to create schema. PDS would need to formalize that.

~ So, suggest a structure to associate a structure to an XML file.

~ We did look at this initially, but there's an extra dimension to XML files. Can see if we want to discuss how to do this in a native way.

~ Someone is willing to do prototypes.

Next meeting August 4.

DDWG Notes 2022-08-04

title: DDWG Notes 2022-08-04

layout: default

date: 2022-08-04

August 4, 2022

Notes by Debra Kazden

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

(Note: M. Tiscareno attended the July 7, 2022 meeting but was not listed as a known observer)

DDWG Agenda

(Included in email sent Aug 2, from M. Drum, DDWG Agenda for Thursday, August 4th 2022)

Front-matter/What's new

1) Dictionary/Submodel rename has been proposed for withdrawal

2) Approved SCRs by 8/26/22

(Discussed)

Ready for Vote (again)

- CCB-348 – Add Units of Mass Density as a unit of measure – Trent Hare ****(Voted to Pass)****

Tier 1 issues

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

- CCB-347 – Importance of Reserved File Name Extensions vis a vis Other Requirements – Dick Simpson

- CCB-211 – XML/schema based files as archival data – Jesse Stone, Mike Drum - Discuss data vs. document, edge cases

- CCB-349 – Rename Dictionary to Submodel – Mike Drum - discuss withdrawal or refinement

- CCB-350 – Adopt more rigorous / stringent rules for leap seconds and use of "24:00:00" in datetimes. – Ron Joyner, Dick Simpson ****(Some Discussion)****

Tier 2 issues

- Higher order products (tickets to be created) – Lyle Huber

- CCB-345 – Target Identification attributes name and type allow nillable – Trent Hare

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

Parking lot – need a driver!

- CCB-209 – Correct Definitions of Tagged-*Object

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

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

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

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

- CCB-261 - Add an optional method for indicating that a product is a member of a larger collection that has a DOI -- Last chance before recommending withdraw.
- ****(Discussed CCB-209)****

Next meeting: 8/11/22

DDWG

Front Matter/What's New

Very few meetings after this one for us left to approve new SCRs. Today will be a quick run down of what we think we can do.

From the top of the agenda, we need to decide what we can vote on next meeting and what we need to punt on. Hope to get through the bug fixes first - unlikely to have any roadblocks for those.

Ready for a Vote

CCB-348 - Add Units of Mass Density as a unit of measure See <https://pds-jira.jpl.nasa.gov/browse/CCB-348>

Question: Are we ready to vote? Anyone opposed to voting?

Answer (Silence)

(Debra called out the nodes for the vote.)

****The Vote for CCB-348:**

ATMOS - Yes

IMG - Yes

EN - Yes

GEO - Yes

IPDA - Yes

NAIF - Abstain

PPI - Yes

RMS - Yes

RS - Yes

SBN - Yes******

(PPI and RS had difficulty with audio)

Voted to Pass.

Dictionary/Submodel rename has been proposed for withdrawal (CCB-349)

A comment has been added in JIRA that this is probably more work than it's worth. It would be great if we had a unique term to use, but we don't. People has issues with the term sub-model. It's a lot of effort to change "dictionary." We need to document that this is more than a dictionary. Could use the term "class dictionary" or "model dictionary." Willing to withdraw unless someone feels very strongly.

~ Someone reported that every time they try to explain this to mission people it's a nightmare, but does like the idea of using class or model dictionary. A slow evolution, the name LDD is everywhere.

Question: It can be withdrawn without a vote. Anyone against this?

Another Question: Where would we enhance the documentation?

Answer: Can start with every public talk we give and the SR. Probably unnecessary in the IM.

~ That can be determined later.

~ We would need a different SCR for the documents.

~ Someone hasn't looked at the Glossary or Concepts Document in a while, but maybe the definition of dictionary could be enhanced if it's in the Glossary.

~ Someone else doesn't think it's there.

~ We could add it to the Glossary. Might need an SCR to update the Concepts Document.

~ Sounds like most people are on board with withdrawing this. We can do an SCR for a more nuanced change later.

~ Someone is okay with the withdrawal.

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

This is a big topic. There has been an ongoing discussion with Anne on this. A few people have been copied on it. Discussing the proper way to do this. After some research, Anne came to the conclusion that should include license information in the labels, so she wrote a more in depth proposal. Want everyone to read it. Not sure what we want to do for this build. One of the key things is that there should be management level policy discussions at MC and IPDA.

****Action Item - Everyone**** Read the updates in JIRA

Someone suggests we can't pass this without a complete enumerated list.

~ This was discussed with the reporter at IPDA. They are concerned about the list, but if new values can be added through the CCB for what they need in a fairly straight forward way then it would be okay.

~ Permanent value list wouldn't change often and would go through the DDWG.

Question: Making it required and in every label?

Answer: The decision has to be made at a higher level.

~ Someone thinks we need to keep it optional.

~ We'll need some sort of control to make the associations.

~ It would have to be heavily validated.

~ Someone is concerned about the mechanics for associations. Don't want an ESA product to use a NASA license or vice versa.

~ ESA might have to write their own schematron.

~ If NASA says all data has to be under a certain license, we have to check for that.

~ Someone is concerned. Thinks it needs to be optional and we roll with it.

~ It might not be our choice.

~ We will need to develop best practices. Presumably that's do-able.

~ Someone is assuming this would go in the identification area. We probably need Mike to do a presentation to the MC on this.

For today, the question is if this will be ready for this build.

- ~ We need the policy before the implementation.
- ~ Someone agrees, but doesn't see any harm in keeping it optional.

Question: Is there any concern with waiting for IPDA?

Answer: Can wait, but only asking for this to be optional.

- ~ We could maybe do this in a point build. We don't do them often, but we can do it if need be.

Question: Someone wants to make sure they understand - IPDA and ESA need this as optional and people are okay with that, but we are unclear if NASA wants it as a requirement?

Answer: Yes. And the details.

- ~ So we need to see if it's a requirement.
- ~ We haven't heard from the MC, but we know it's a topic from further up at SMD.

****Action Item - Mike**** - will put together a PowerPoint.

CCB-209 - Correct Definitions of Tagged-*Object See <https://pds-jira.jpl.nasa.gov/browse/CCB-209>

(From the Parking Lot)

This has been on the docket for many years. At this point it's about 75% there. The people working on it can't tell if it's done. Would like to implement it and see if there's any fall out. It doesn't affect operations in any way. It's more philosophical than practical.

CCB-351 - Term Mapping

See <https://pds-jira.jpl.nasa.gov/browse/CCB-351>

Because we have a model driven architecture, there has been a request from Jordan to start collecting mapping. There are four mappings included in the SCR. The approach will be to create a DD attribute external that would refer the x-path to something already defined. It works nicely to do DD attribute external and DD class external and use x-path and terminological entry. Would do referencing with logical identifiers. No new relationships, just a different type of referencing.

Someone needs to see some examples.

- ~ The use cases are prepared. Will be added to the SCR ****Action Item - Steve****

Question: Don't we already have a DD class external?

Answer: Very similar concept. Need it to literally map. Tried a few options. This seems to work best.

Another Question: If building an LDD, would use this to make associations clear?

Answer: Exactly.

- ~ Most of this will be behind the scenes, but we could do it in line while developing LDDs.
- ~ Will need to see the class.
- ~ There are examples. Mainly for attributes. No alias for class. Unclear if we need that.
- ~ Need something to put in the IM in Proposed Solution.
- ~ Will add that to the SCR ****Action -Item - Steve****

Question: Is this a way to create a meta-model to create relationships to semantically identical attributes and classes?

Answer: Yes, but you refer to existing one and then map to an extension, so it is a new attribute even if it is an exact match. Still a new version of that.

Another Question: Regarding the second option in jira, would this be allowing things outside PDS4 in?

Answer: Will need internal status. Is this PDS4 or a mapping to provide additional information. (See Jira)

There are some policy issues behind the scenes.

Someone sees a lot of power here, and the potential for getting it wrong. This close to the cut off we often move the DDWG meetings to a weekly meeting. If EN needs this we should meet August 11.

~ That makes sense. Also could restrict this to OPS and say no one can use this outside EN.

~ Might do that anyway. Can talk about more general usage at a later date.

Question: Is this at least partially to get registry to include products not archived in PDS4?

Answer: Not exactly.

~ Not sure why we would want to map PDS3 to PDS4 attributes.

~ Could say what's equivalent. Need more discussion before we can vote.

~ Powerful for referencing things that are identical. Versioning is an issue that would need to be tracked. (See comments in Jira) ~ Use cases will be added to Jira.

We will meet and discuss this on August 11. Maybe another issue too, and then we can vote in two weeks.

~ Might be able to do license too. The MC is on the 10th.

~ If we could vote on CCB-209 (Correct Definitions of Tagged-*Object, see <https://pds-jira.jpl.nasa.gov/browse/CCB-209>) that would be nice.

~ CCB-352 (Standards Reference Footnote Error, see <https://pds-jira.jpl.nasa.gov/browse/CCB-352>) too.

~ That's basically a typo.

~ Let's vote now.

~ No TA.

~ There will be later today.

Question: Okay with not wasting time. Anyone opposed?

Answer: People need to look at it.

~ No big rush. No problem waiting.

~ Will do a quick vote in a few weeks in case people can't come next week.

CCB-350 - Adopt more rigorous / stringent rules for leap seconds and use of "24:00:00" in datetimes.

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

People have been discussing this. The patterns aren't perfect, but trying to have some conformity. 24 hours is no longer acceptable to ISO. Leap seconds should only be in June and December. Trying to decide how to make the patterns better. Think can cover a lot of this, but no good way of checking these. Someone with tools can check. People working on this agree that this is a bug fix. Want everyone on the same page for the fixes.

~ Sees this as a duel phase - first, get consensus on leap year and 60, then implement, which includes changes to the test cases.

Question: Did the DDWG agree to the two questions in the SCR?

Answer: At least that 24 hours is not acceptable.

Another Question: What about 60?

Answer: Need it for leap seconds.

~ Absolutely.

Listed all the leap seconds.

~ Through 60. Invalid.

~ Yes, but in that example it was February 1.

~ The example had additional issues. Can only have leap seconds on specific dates.

~ It will be very hard...

~ Already in the Requested Changes.

We can go forward and re-do the patterns and run against the test cases.

~ Will confirm it's correct with NAIF.

Question: Sounds like a bug fix. Are we ready to vote?

Answer: Not sure we have to vote, but need some input for SBN on dates before 1582 including years designated with negative signs.

~ Someone will check in with someone at SBN on that.

Another Question: This is just going to be about leap seconds and 24 hours. Can we not bundle that in? This is cut and dry.

Answer: Someone agrees that should be a separate SCR.

~ It can be broken out.

~ Address that as a separate SCR. It's only coming up because working on the pattern. There are a finite number of leap seconds. Not sure why validate tool can't use the accepted list for NAIF. Seems like less work.

~ Someone agrees. We don't need to spend a lot of time on discussing how validate works.

We will check in on this again.

CCB-347 - Importance of Reserved File Name Extensions vis a vis Other Requirements

See <https://pds-jira.jpl.nasa.gov/browse/CCB-347>

Question: Do we want to vote on this in two weeks?

Answer: Let's discuss it next week. There are several options on how we can go forward.

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

Someone assumes this is too complicated with lots of edge cases to work out. Not aware of any urgent need to approve this. Will move it to Tier 2.

We will meet for a discussion next week - August 11

DDWG Notes 2022-08-11

title: DDWG Notes 2022-08-11

layout: default

date: 2022-08-11

August 11, 2022

Notes by Debra Kazden

Known Attendees: M. Drum, M. Gordon, E. Guinness, T. Hare, L. Huber, S. Hughes, R. Joyner, D. Kazden, J. Mafi, S. McLaughlin, B. Semenov and R. Simpson
Known Observers: P. Lawton, T. McClanahan and M. Tiscareno

DDWG Agenda

(Included in email sent Aug 9, from M. Drum, DDWG Agenda for Thursday, August 11 2022)

Front-matter/What's new

- Discussing PDS Licensing Information at MC 8/10/22
 - Do we need a slot during MC F2F 11/15 to discuss any issues?
 - Approved SCRs by 8/26/22
- **(Discussed)**

Ready for Vote

(none this week)

Tier 1 issues

- CCB-347 – Importance of Reserved File Name Extensions vis a vis Other Requirements – Dick Simpson
 - CCB-350 – Adopt more rigorous / stringent rules for leap seconds and use of "24:00:00" in datetimes. – Ron Joyner, Dick Simpson
 - CCB-351 – Term Mapping – Steve Hughes
 - CCB-209 – Correct Definitions of Tagged-*Object - TA complete, vote next week
 - CCB-336 – Add a License Information class to the Identification Area – Tanya Lim
- **(Discussed)**

Tier 2 issues

- CCB-211 – XML/schema based files as archival data – Jesse Stone, Mike Drum - Discuss data vs. document, edge cases
 - Higher order products (tickets to be created) – Lyle Huber
 - CCB-345 – Target Identification attributes name and type allow nillable – Trent Hare
 - DOI-rich metadata into IM (ticket to be created), replacing CCB-222 and CCB-261 – Anne Raugh and any volunteers
- **(Not Discussed)**

Parking lot – need a driver!

- CCB-326 - Inconsistent Direction on How to Use Discipline Facets
- CCB-324 – Create a data type for Type attributes that require a permissible value list.
- CCB-164 - Display Settings not required for images -- (IMG) - Need to think more about potential updates to Imaging and how Display might fit.
- RS: 259/281 (see new attachment to CCB-281) (on hold)

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

(Moved 164 to Tier 2)

Next meeting: 8/18/22

DDWG

At the beginning of the meeting there was some discussion about the Tuesday in St. Louis for the F2F. The thought was that if we are committed to the facility and can't cancel then the DDWG might be able to make use of the time if EN is not doing another town hall on the web.

~ The P.O. is still waiting to hear from the node leads.

~ People were asked to do a show of hands for who would be at the F2F - note-taker saw Mike, Mitch, Trent, Ed, Joe and Debra. (There may have been others)

Tim said that he will sort this out with Jordan and the MC and get a decision out.

There is a lot to discuss today. We want to dedicate most of the meeting to that today. Tim was invited to help discuss the next steps on the PDS license issue.

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

At the MC we decided we still need further clarification on the issue.

~ Someone is curious why we would consider anything other than CC-BY and optional. It supports open science, acknowledges sources. We should encourage the use of that to the providers.

Question: Comments?

Answer: Yeah, think we would all wish that data would be available under CC-BY. The question is if there are cases that would need more restrictive licenses.

~ We need to allow for ESA to have other values.

~ We would need to track those.

~ Not our responsibility, but we would need to know the values.

~ Someone is trying to understand.

If we put the license class in the labels we would need an enumerated list. It would include CC0, CC-BY and others. We would use a schematron rule to say if this is a NASA value it has to be this, and if it's an ESA value it has to be that. Can put a link to an FAQ that helps provides understand what licenses should be used.

~ People aren't sure the FAQ is correct. CC0 seems to destroy open science.

~ It just implies the copyright is gone, but we would still include author, DOI, citation information, etc.

~ Concerned about the downstream use of the data.

Tim will spin up a conversation with Steven Crawford to get some clarification.

~ It is reasonable for Tim to be the one to start the conversation.

~ We won't be voting on this next week.

(Tim might leave the meeting early.)

CCB-347 - Restrict Content and Naming of Inventory Files See <https://pds-jira.jpl.nasa.gov/browse/CCB-347>

Hasn't been much discussion on this yet. It will be re-introduced. There is a PowerPoint in Jira.

~ The problem came up when someone found an inventory file with a .TAB extension, which is reserved for fixed field tabular data. Validate didn't report that. Put it in as a Github issues, but Jordan said he wasn't sure what to do with it. There are additional concerns with other file extensions that also may need clarification. The question is if you see a .TAB if it needs to be checked early. The requirement is that tables be tabular data in fixed length records, but validate skips over that for inventory files. Need to consider what it means to have restricted file extensions and what requirements they need to meet. (See the PowerPoint in Jira) The issues for inventory files are if you can use .TAB. Most have variable length records. Could require they are always .CSV files. Not a requirement now, but most are. Another issue is that we prohibit having an offset before the inventory object - need to answer the question of if we could have something after that. Unclear if we could have a header after that. There are additional questions if a file contains fixed width and comma delimited objects. Need to decide if HTML or XML files can be .TXT files and if we want to distinguish between .DOC and .DOCX. There are no requirements. There are lots of issues in table 6C-1. Many with NAIF files.

The last slide includes possible solutions, for example, could make inventory file only contain one inventory and end with .CSV. Also recommends that if we use .TAB to include one or more table objects and nothing else and that digital objects with comma separated tabular data - no other digital objects. Docx for Microsoft Word XML objects and .DOC for Microsoft Word non-XML documents. Would like to be able to validate inventory files, but worried about the Pandora's Box here.

One small correction - 2.B.2.2 of the SR indicates collection files have inventory with .CSV.

~ Think that's the default.

~ Someone Likes the first two recommendations, but would add that when you get to field delimiter for inventory it should be restricted to just commas as the delimiter. Issue with objects with headers and more. File extension refers to the main thing in the file. We always allowed multiple digital objects in a file.

~ Sounds like asking for trouble.

~ Can't have headers and other objects separate from each other.

~ Not saying you can't have mixed files. Concerned over the file extensions here.

~ Someone said that mixed to me implies more than one format in a defined space. Recommendations seem like a common sense solution.

~ Someone agrees that the first two are easy, but there are probably billions of products in PDS with more than one object in them. Images with tables, etc. Limiting that could cause problems. Also, chapter 2 has always been problematic because it is just suggestions. We need to fix that at some point.

CCB-335 (Inventory Specification Allows Too Many Delimiters, see <https://pds-jira.jpl.nasa.gov/browse/CCB-335>) was implemented.

~ But validate didn't complain.

~ It's queued.

~ That was too many delimiters.

~ There was another issue awhile ago, SR said inventory files needed to have a .TAB extension, but that was fixed.

This sounds too complex for a vote, unless we just break out the inventory part.

~ Someone agrees. We should break out the first two bullets.
~ Reporter can break something out ****(Action Item - Dick)**** ~ We can put that to a vote.
~ CCB-117 (Change collection inventory file extension from ".tab" to ".csv" in PDS4 documents, see <https://pds-jira.jpl.nasa.gov/browse/CCB-117>) was implemented and released.

Question: Is this a matter of validate not checking?

Answer: Maybe the documentation was updated, not validate.

~ It was changed to allow CSV, not disallow TAB.

Let's vote next week on just .CSV. It will be on the agenda.

CCB-350 - Adopt more rigorous / stringent rules for leap seconds and use of "24:00:00" in datetimes.

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

Question: Is this ready?

Answer: People are still hashing it over. Hope to reach consensus and then it will be a bug fix.

CCB-353 - Standards Reference Footnote Error See <https://pds-jira.jpl.nasa.gov/browse/CCB-352>

This is ready for next week. We don't need to discuss it. Bug fix.

CCB-351 - Term Mapping

See <https://pds-jira.jpl.nasa.gov/browse/CCB-351>

We discussed this last week. It was introduced.

~ There was a request for use-cases. They were made available. This is OPS. The initial purpose is for mapping to be in LDDTool and maps to put in search. There was discussion that terminological entries could be used by designers. This implementation doesn't allow that. This is for system to do mapping. Experts will get involved in exactly what is needed.

Question: So, this is not in a typical dictionary?

Answer: Right, probably in a class. Search is mapping class in LDD ingest file for OPS.

Another Question: Who does the vetting to make sure it's scientifically correct?

Answer: We can get Jordan to talk to that.

~ We would have problems based on the examples.

~ Don't take them literally, they are just examples. Garbage in, garbage out.

~ There are two broad categories, PDS4 to PDS3 and PDS4 to PDS4 mappings. It would be nice if output was a simple table with the associations easy to see. Easy for science node reviews. Maybe could be a CSV.

~ Original requirement was for JSON because the system uses that, but CSV is no problem if the discipline nodes design the mappings. Unsure how Jordan's team could do that.

Another Question: In a separate box, dictionary?

Answer: Not a data dictionary document, just tables created for use by OPS.

~ But using an ingest file.

~ Yes.

~ So a dictionary of some sort would need review.

~ Just software with a bunch of flags.

- ~ So not doing terminological entry to accommodate that.
- ~ Solving the first problem first. Then as many use cases as we want.

Too many things to think about beyond OPS for now. One important point for model management is that including versioning would be very important. Show x-path migrations over history. A way to reach back in time to show this attribute relates to that one.

- ~ Could add version ID. Will need to figure out how to validate that. Still all in discussion.
- ~ Someone is unclear how we bundle useful search if we can't search old versions of the IM.
- ~ We might need to make version multi-valued.
- ~ Good discussion, but going way down the road. For now, we are talking about the current version. Multi-valued version is problematic. One to one mapping in current design. Assumes current version of IM for now.
- ~ If the current tools only work in one version then adding multiple versions might not be helpful. Could be valuable in the future.
- ~ Could make version IDs optional for now and come back to it later.
- ~ That raises the issue of purpose of how things are being used. For now, this is just trying to capture information.

We could vote next week. The current lien is to add version.

Question: For the preferred name - do we want extended versus external?

Answer: Either comes with unanticipated baggage.

- ~ Extended is less problematic. Will still have both in ingest.

CCB-209 - Correct Definitions of Tagged-*Object See <https://pds-jira.jpl.nasa.gov/browse/CCB-209>

This has a TA. We will vote next week.

****Action Item - Everyone**** take a look at it.

Question: Do we need to discuss this?

Answer: No, because we want to do it. If anyone wants to discuss it they can take it over.

- ~ Sounds like a veiled threat. If Steve and Dick are in agreement, won't object.
- ~ Someone says this is a good proposal to move this forward. Will vote for it.
- ~ Someone else won't argue.

We will vote on this next time.

Question: Anything else?

Answer: Have a feeling this will bring up some ire, but we have been struggling with video for the last few years. Compressed. Hard to read. Wonder if this is an opportunity to do it for derived only, for the modern era. Came up for HiRISE because JPEG2000 is going away. Unsure if should push on this.

- ~ The question is if this is archivable science data, if not PDS, then who.
- ~ That's not an answer.
- ~ Someone said yes, we should move forward with this. It's controversial, but yes.

~ We can discuss it in November. It came up in the Ecosystem. Want to limit to derived only. Just a thought. Things that can be reproduced.

~ The plan all along - partially driven for things like trying to recover Apollo data. Wanted to develop tools that could convert the archive to contemporary formats down the line.

~ Problem with that is a data glut. Can't afford raw and everything, but there are alternatives. Could say people have to pay for the raw. Will think about it more.

Someone would like to move CCB-164 (Display Settings not required for images, see <https://pds-jira.jpl.nasa.gov/browse/CCB-164>) to tier 2 to be in tier 1 after the build.

Next meeting August 18.

DDWG Notes 2022-08-18

title: DDWG Notes 2022-08-18

layout: default

date: 2022-08-18

August 18, 2022

Notes by Debra Kazden

Known Attendees: M. Drum, M. Gordon, E. Guinness, T. Hare, L. Huber, S. Hughes, D. Kazden, J. Mafi, S. McLaughlin, B. Semenov and R. Simpson
Known Observers: P. Lawton and M. Tiscareno

DDWG Agenda

(Included in email sent Aug 16, from M. Drum, DDWG Agenda for Thursday, August 18th 2022)

Front-matter/What's new

- New response from Steve Crawford re: licensing
 - MC F2F Tuesday slot is available for DMSF (or DDWG if needed)
 - Approved SCRs by 8/26/22
- **(Discussed)**

Ready for Vote

- CCB-209 – Correct Definitions of Tagged-*Object
- CCB-351 – Term Mapping – Steve Hughes
- CCB-347 – Restrict Content and Naming of Inventory Files (Original SCR title: Importance of Reserved File Name Extensions vis a vis Other Requirements) – Dick Simpson ****(Voted to Pass)****

Tier 1 issues

- CCB-336 – Add a License Information class to the Identification Area – Tanya Lim ****(Some Discussion)****

Bug Fixes

- CCB-350 – Adopt more rigorous / stringent rules for leap seconds and use of "24:00:00" in datetimes. – Ron Joyner, Dick Simpson
- CCB-352 – Standards Reference Footnote Error ****(No One Opposed)****

Tier 2 issues

- CCB-211 – XML/schema based files as archival data – Jesse Stone, Mike Drum - Discuss data vs. document, edge cases
 - Higher order products (tickets to be created) – Lyle Huber
 - CCB-345 – Target Identification attributes name and type allow nillable – Trent Hare
 - DOI-rich metadata into IM (ticket to be created), replacing CCB-222 and CCB-261 – Anne Raugh and any volunteers
 - CCB-164 - Display Settings not required for images -- (IMG) - Need to think more about potential updates to Imaging and how Display might fit.
- **(Not Discussed)**

Parking lot – need a driver!

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

Next meeting Sept 1

DDWG

Front-matter/What's new

There might be some time scheduled on Tuesday of the November F2F for the DSMP, but we can have an option to have DDWG discussion too.

- ~ Unsure at this point.
- ~ ATMOS will be there.
- ~ Haven't heard from SBN or McAuley yet.
- ~ Gerbs volunteered Ben to be there.
- ~ Anne and Ben were asked to go.
- ~ Someone is pretty sure Anne will be there.
- ~ Mike for PSI, Bea, maybe Jesse.
- ~ Trying to get numbers to Tim for how big a room the DSMP will need.
- ~ Again, that day is available for DSMP or DDWG to go over open issues and get a status on where we are for things like LDD Ingest, cross references, etc. Not to try to solve things, to decide status. Also hoping to have a splinter group work on revising the DPH while others attend the MC. Ben might be one possibility if he's willing. Hopes also for someone from IMG.
- ~ At least virtually.
- ~ Hoping to have people in a room, locked in. Can ask people at the MC for clarification as needed. Prefer not virtual.
- ~ Good plan. For DDWG topics, it's unclear if there will be any pressing issues, but eventually we do want to tackle bigger topics, like the 2.0 transition.

There was a response from Steve about NASA licenses. Added it to Jira. Basically, NASA funded data classified as civil servant data, should be CC0, but we are allowed to post data that is more restrictive than that. Still need more discussion with MC. See Jira. If there are no comments we can wait on this until after the build.

- ~ A two part question: We need a way to put information in a label and then what licenses are available.
- ~ And if required. Unclear who has strong opinions.
- ~ Someone thinks we want CC-BY.
- ~ There are important differences between CC0 and CC-BY. CC-BY implies a copyright so it wouldn't be much benefit.
- ~ Fair enough. Then it has to be based on LID. PSA wants CC-BY.
- ~ Science data developed solely by civil servants is a practically non-existent set.
- ~ Yes, a very tight definition. Need some clarification. Doesn't sound like this will be solved by this build.

Question: How do people feel about building the class? Not sure we can assume for NASA that it will always be CCO and entirely NASA funded.

Answer: Looking at CCB-336, it's to put in an optional class. No requirement on what the values are. Could restrict it later, but not sure it has all the information we need.

~ Danger if we put it in now and someone puts in a license that breaks the rules. The original SCR doesn't list possibilities. If we are not doing a tight enumerated list we can pass the original proposal, but not confident to decide about restrictions.

~ Someone thinks CCO is baked. Think we do need the option, but will be confusing when some have it and some don't. Could add a description.

~ Could provide encouragement for CCO and say you can get permission for others. Could be a software restriction.

~ The SCR says to provide a link. Won't work.

~ We could amend this, but it is owned by ESA. Mark put it in. He and Tanya are not here.

~ Could amend it to have an enumerated list and say for NASA what the options are.

****Action Item - Mike**** will reach out to Mark and Tanya to see if we can get their values and see if we can put a restriction on it.

Question: Where does the complete text of the license go if enumerated list? Has to be part of the archive.

Answer: Could be a URL or a link.

~ As long as some sort of document goes into the archive that is available 25 years from now.

~ Someone would hope that Creative Commons would still be available then.

Ready for Vote

CCB-209 - Correct Definitions of Tagged-*--Object See <https://pds-jira.jpl.nasa.gov/browse/CCB-209>

Question: There are no new comments in Jira. Are there any now?

Answered with Another Question: Why is there a bunch of cartography stuff in there?

Answer: Shouldn't be there. Ignore it. Probably from earlier. Don't remember, but the Tiger Team was working on the Common Model.

~ Included because they are in the tree from classes to sub-classes.

~ Not part of the core.

~ Someone thinks it is. It was a year or so ago.

~ Someone else didn't see it there.

~ TNDO (tagged non-digital object) context, not cartography. Looking at the specification now, not seeing cartography.

~ Not context, under TNDO.

~ Looking and not seeing it. The team has been working on common. Cartography is not part of that.

~ Then someone took it out.

~ Years ago - decided to keep the discipline dictionaries separate from common. Unclear if you were looking at the specification or the dictionaries.

~ Was using specification. It changed. Surprised the DDWG didn't hear, but it's okay. We don't need to worry about cartography.

~ Was originally in the same database as for common. A strange issue from the past that is resolved.

****The Vote for CCB-209:**

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

CCB-351 - Term Mapping

See <https://pds-jira.jpl.nasa.gov/browse/CCB-351>

This was amended to include the last comments.

~ Someone is curious how number 5 would be enforced.

~ In the Requested Changes, could make sure LDDTool wouldn't recognize them. Would use an OPS flag. For documenting, could control what goes in there.

~ Okay. Worried about property maps. Don't want problems.

~ Good point. This issue has been around for many years. Property maps were used for a while. Caused problems. That's why we need this focused through OPS.

Question: Any more comments?

Answer: (Silence)

**The Vote for CCB-351:

ATMOS - ("I guess. Interested to see how this will work in the wild.") - Yes EN - Yes GEO - Yes IMG - ("About as optimistic as ATMOS") - Yes IPDA - Absent NAIF - Abstain PPI - Yes RMS - Yes RS - Abstain SBN - Yes**

CCB-347 - Restrict Content and Naming of Inventory Files See <https://pds-jira.jpl.nasa.gov/browse/CCB-347>

We discussed this last week. There was a comment added to Jira last week.

~ The whole reason we have this SCR is that Jordan wanted clarification and we are basically side-stepping that. Will need another SCR,

Question: Is this strong enough to make validate tool do what we want for inventory?

Answer: Not sure, but it solves the original issue.

~ Someone has been harvesting data lately and seeing inventories that are archived without a CSV file extension. Can migrate those.

~ The question is if we give it a CSV file extension if it has to be a CSV file. That's part of what's tripping Jordan up.

~ Yes. Made recommendations last time, but not in this SCR.

~ Will be in the next SCR. Includes a lot of NAIF issues. Will want their help.

~ NAIF is unsure what issues.

~ Someone doesn't think Jordan is going to be happy.

~ Jordan's larger problem is not solved. No reserved file extension for delimited in general.

~ Someone thought we made a change on delimited tables for inventory tables last build.
~ Yes, we implemented one. A schematron rule. Constrained the delimited table. CCB-335 (Inventory Specification Allows Too Many Delimiters, see <https://pds-jira.jpl.nasa.gov/browse/CCB-335>).
~ It's not implemented yet. We already passed it. That requires inventories to use commas. That with this and CCB-346 (Additional Updates for Clarifying Inventory Specification Delimiters, see <https://pds-jira.jpl.nasa.gov/browse/CCB-346>) would make inventories be delimited tables, with commas and be CSV.
~ Fair.

So, we need CCB-346 in too.

~ Someone thought we discussed it.
~ There was a problem.
~ CCB-346 had issues. Was approved at the last minute. Implemented number 1, but couldn't modify the class and then ran out of time. The issues aren't trivial. Need more discussion. How to implement it wasn't clear.
~ We are approaching the last minute again.
~ Yes.

Question: Are they issues that the DDWG can discuss?

Answer: More of a tech issue. Could be discussed off-line.

Another Question: If it's already been voted on, would EN like us to vote and also try to discuss the issues?

Answer: Should continue. Concerned about getting it into this build.

~ We could hold CCB-346 for the next build and get a Tiger Team to work on it.
~ Should have brought it up earlier. Interesting to have three SCRs tied together.
~ Without numbers 2 to 5 the requirements would stand at must use commas and be CSV. Only a malicious actor would make an inventory that isn't.

**The Vote for 347 -

ATMOS - Yes

EN - Yes

GEO - Yes

IMG - Abstain

IPDA - Absent

NAIF - Yes

PPI - Yes

RMS - Yes

RS - Yes

SBN - Yes**

It sounds like there will be future discussions, but that's through.

Tier 1 issues

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

This will slip to the next build. Tanya said it wasn't urgent. Don't think we have enough information to try to squeeze this in.

Question: Any objections?

Answer: (Silence)

~ Close, but not there yet.

~ Someone thinks there's a new version of SPD41 coming out, so it might benefit us to wait.

Bug Fixes

CCB-350 - Adopt more rigorous / stringent rules for leap seconds and use of "24:00:00" in datetimes.

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

No votes are required for bug fixes, but Ron wants official guidance on this. There are still four issues to resolve. See Jira. The SCR was cleaned up last night, now it has comments added from Anne. If what is in there is okay we can tell Ron to start working on it.

Question: Objections?

Answer: Not an objection, but item 3, think validate already checks that. Could be wrong.

In regard to voting - a previous DDWG Chair got beat up by the CCB on items the DDWG thought were bug fixes. Anytime we are changing the model we need to vote and send it to the CCB with a note that says we think something is a bug fix.

Question: Since no one is clamouring over this, does anybody... All opposed?

Answer: (Silence)

Action Item - Mike will tell the CCB what we think.

Another Question: For the additional bug fixes, do we need to wait for additional edits first? Can we submit them to the CCB or do we need to wait?

~ Answer:(person who would answer is having audio issues.)

CCB-352 - Standards Reference Footnote Error See <https://pds-jira.jpl.nasa.gov/browse/CCB-352>

This is the other bug fix.

Question: Anyone opposed?

Answer: (Silence)

Action Item - Mike will get clarification from Dick before he send the bug fixes out. Will email him since he is having audio issues.

We will meet again in a couple weeks.

Next meeting September 1

DDWG Notes 2022-09-01

title: DDWG Notes 2022-09-01

layout: default

date: 2022-09-01

September 1, 2022

Notes by Debra Kazden

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

DDWG Agenda

(Included in email sent Aug 30, from M. Drum, DDWG Agenda for Thursday, Sept.1 2022)

Front-matter/What's new

- New issue: CCB-353 – Fully label all bytes in data objects
 - CCB passed CCB-348, in process of reviewing 209, 347, 351
 - Approved SCRs by February 2023
- **{Discussed}**

Ready for Vote

(none this week)

Tier 1 issues

- CCB-336 – Add a License Information class to the Identification Area – Tanya Lim
 - CCB-211 – XML/schema based files as archival data – Jesse Stone, Mike Drum - Discuss data vs. document, edge cases
- **{Discussed}**

Tier 2 issues

- CCB-335 - Support for video as product observational – Trent Hare
 - Higher order products (tickets to be created) – Lyle Huber
 - CCB-345 – Target Identification attributes name and type allow nillable – Trent Hare
 - DOI-rich metadata into IM (ticket to be created), replacing CCB-222 and CCB-261 – Anne Raugh and any volunteers
 - CCB-164 - Display Settings not required for images -- (IMG) - Need to think more about potential updates to Imaging and how Display might fit.
- **{Brief Discussion}**

Parking lot – need a driver!

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

Next meeting: 9/15/22

DDWG

Front-matter/What's new

We are missing some regular attendees today.

New build cycle.

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

We have a new SCR, CCB-353 from Anne. It is requesting something for version 2.0. It's to require all bytes in data objects to be documented. This will be a long term SCR to work on.
~ Someone made sure that Anne saw that there are already comments in Jira. Anne replied.

Question: Anyone have anything to say about this yet?

Answer: Anne's comment added some clarity. IPDA has been discussing this.

~ It's a scary one. Spoke to Anne off-line about this. Still need to learn more about this.

The CCB approved the SCRs that we passed. They will be in this build, along with the bug fixes. Our new cut off date will be in February. License information and XML data were left over. We will work on those. Also want to give an opportunity to get others into Tier 1 too. Trent isn't here, but will want to talk about video and there's one for higher order products with tickets to be created.
~ That won't be ready for about a month.

Tier 1 issues

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

This SCR is from IPDA. From Mark. He is their archive lead. His boss and he discussed this and now the ESA lawyers are discussing it. The problem is that ESA is not a country. Many of the licenses seem to be country specific. This issue shouldn't hold everything up for PDS, especially since this is optional. When IPDA has a list of licenses for the enumerated list that can be added.

~ Someone is unclear on where that leaves this issue. Believes the DDWG still needs a policy decision. Steve Crawford said we can archive with restricted licenses. A more detailed proposal is probably what we need to go with. Can work with Steve and take a stab at describing the class. ***(Action Item - Mike and Steve)*** ~ There are still open questions about how required we want this, validation and which licenses are allowed.

Question: Does anyone have anything to add?

Another Question: Do you mean if we should wait or implement it and work out the license details later?

Answer: The open question is if we should make this required, and if not, maybe we don't need a policy recommendation. We could just say we allow certain licenses.

~ Could say if no license is specified then there's a default license. NASA could change it as they want.
~ That might be a good idea. Once a license is applied can't really be further restricted.

Another Question: Can it be loosened?

Answer: Yes. The page linked in Jira to CC licenses is very good. Everyone should go through that.

****(Action Item - Everyone)****

Question: What if we implement with only CC-0 and CC-BY, and then if anyone wants to add something else we can discuss that later? Or we could just implement them all?

Answer: CC-BY ND means you can't use this to create your own thing. Nothing in PDS should be that.

Not sure we should implement them all. If there are policy concerns, maybe we don't want to be tracking copyright.

~ With CC-BY we would be giving credit.

Question: So, PSA wanted this, and they are working on it now. Will they want anything more restrictive than CC-BY or CC-0?

Answer: Someone doesn't think so.

~ Someone else thought the MC mentioned a mission coming up with commercial concerns.

~ Maybe CLPS.

~ We could ask Maria Banks. She brought up the concept that NASA might have to buy data from some of these missions.

****Action Item - Mike**** will reach out to Maria to ask. Others can too.

Hopefully, we can make something concrete for this proposal.

~ Someone is not sure we would want to accept data under more restrictive licenses. We will need to have some conversations, but it is up to NASA. Thinks we are okay with CC-0 and CC-BY until that time.

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

Nothing new on this. It was introduced a few weeks ago. Wanted to know if anyone has had any thoughts on edge cases - compressed, license restricted data, etc.

~ (Silence)

~ Expecting input for some of the people who aren't here today. We can wait for that.

Tier 2 issues

Some of these will wait until Trent is here to discuss.

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

Question: Anyone interested in being involved? Helping with a bunch more providence metadata.

Answer: Steve has worked on it. Will be involved. Anne is talking to the Datacite people.

~ Good if Steve and Anne are already in contact.

Question: Does anyone have anything else to bring up today? This might be a short meeting...

Answer: (Silence)

Let's call it for today. Nice and quick.

DDWG Notes 2022-09-15

title: DDWG Notes 2022-09-15

layout: default

date: 2022-09-15

September 15, 2022

Notes by Debra Kazden

Known Attendees: M. Drum, T. Hare, L. Huber, S. Hughes, R. Joyner, D. Kazden, J. Mafi, S. McLaughlin, B. Semenov and R. Simpson Known Observers: D. Kahan, P. Lawton, J. Stone and M. Tiscareno

DDWG Agenda

(Included in email sent Sept 13, from M. Drum, DDWG Agenda for Thursday, Sept 15 2022)

Front-matter/What's new

- A quiet week!
- Approved SCRs by February 2023

Ready for Vote

(none this week)

Tier 1 issues

- CCB-336 – Add a License Information class to the Identification Area – Tanya Lim - Still awaiting updates about CLPS
 - CCB-211 – XML/schema based files as archival data – Jesse Stone, Mike Drum - Discuss data vs. document, edge cases
- **(Discussed)**

Tier 2 issues

- CCB-325 - Support for video as product observational – Trent Hare
- Higher order products (ticket ETA is October) – Lyle Huber
- CCB-345 – Target Identification attributes name and type allow nillable – Trent Hare
- DOI-rich metadata into IM (ticket to be created), replacing CCB-222 and CCB-261 – Anne Raugh and any volunteers
- CCB-164 - Display Settings not required for images -- (IMG) - Need to think more about potential updates to Imaging and how Display might fit.
- CCB-353 – Fully label all bytes in data objects ****(Briefly Discussed as part of 211)****

Parking lot – need a driver!

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

Next meeting: 9/29/22

DDWG

Front matter/What's new

Not much new this week to discuss. The two big issues haven't had much movement. Jesse is here today to discuss CCB-211. Maybe we can figure out what the next steps are for XML data. Been discussing it off line. Still some research to do.

Tier 1 Issues

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

Email was sent to Maria and the topic was brought up at the MC. Can't get a handle on CLPS. Seems like the most complicated license issue. Not sure what to do. Maybe we can implement something and amend it later if need be.

~ We can implement with an enumerated list that only lists CC-0 and CC-BY and if someone wants something else we can discuss it then.

~ This is being discussed at the IPDA meeting now. Anne did a good presentation and they are discussing CLPS too. Maybe we could go to Anne or Gerbs to see if they have any formal input.

~ Someone is already working with Anne. Her perspective keeps evolving. Will check in with her.

~ She did a very thorough presentation.

We will have updates on this for next time.

Someone was hoping that we could discuss video, but not everyone is here yet.

Question: Does anyone have anything to discuss?

Answer: We can start talking about higher order products (Tier 2), but would like everyone here for that too.

~ People are attending the IPDA meeting.

~ We can talk about XML and see if people show up.

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

This has been discussed. Trying to figure out what sorts of issues we would be signing up for. A few concrete XML files exist in the wild - Catalina Sky Survey has VO tables, Cassini Microsoft products.

~ Someone thinks CK files in XLSX, but doesn't think they are primary data. Not observational for a planetary target. Useful for analysing and interpreting.

~ That should be a CK file.

~ It was found. It was in Excel XML format in supplemental documents.

~ Maybe we'll need to dig into that. We need as many examples as possible. One idea is to have a set of schema that are allowed to be referenced, maybe owned by EN. They would have to be heavily vetted.

Seems like a way to offload. The IM is schema, so if archiving XML products, we are offloading on what can go into the archive. Might be a way to get some XML products in the archive.

~ Specifics about which collections is all someone can think of.

~ For native, we require people to provide a translated version. Understand there are readers to incorporate, for example XML into Python. Wondering if we can require a more standard format in parallel. Might be difficult, but maybe we could require that.

~ Might not be possible in every case. XML is better at handling arrays and other things. It's creating data that can't exist in existing formats.

~ Those products wouldn't get the same rigorous testing as the formats the DPs use. Problems discovered in original formats might not be fixed in a secondary version.

~ Someone is not so worried about that.

~ The promise we've had is that there will be a way to transform archive formats on the fly. Wonder if this will be a candidate for that, rather than having extra files sitting around.

~ EN would have to do that. Okay with that.

~ Someone is not excited about transformations. The IM expects certain schemas. It's more than arbitrarily archiving XML products, it's a way to define it in a complex structure. Transformations are good for human readers, but that's not the usual audience/reader for such products.

More people have joined the meeting.

Schema could validate the XML file, it could be part of a validate run, and it's correct that Python has an XML reader, but it might not dump out understandable data. The point is maybe we can come up with a set of recommendations and limit it to that. Say only allowed to do certain things.

~ The proposal is to limit the set of schema.

~ That would limit you to a few data formats.

~ Yes. Assume they are written in good faith, don't want to have to object to them.

The approach taken for CDF was that certain things allowed in the CDF specification are not allowed in CDF/A. People can archive what they want, but we don't allow things like internal compression. If possible, with XML we could restrict what's allowed and not the ability to use the format. We can limit, but not exclude it.

~ That gives someone an idea. We could have a supplemental set of schema files to do the restrictions.

~ So, if the goal is to archive XML documents, sounds like restricting. We're restricting what can go in.

That makes sense, but confused by the use of XML schema to do this. The IM is an ontology. We use XML schema to limit what people can do. A cascade level of constraints. Concerned if we allow XML schema in now that we are opening a hole in the model. Not sure why we are not constraining and writing a new class, and why we are jumping into the swamp with XML schema. About two months ago was introduced to two guys who have been using XML and now want to go to JSON. We can produce our IM in any language, So, concerned about the path we are going down.

~ This is a level deeper than the model. Proposing a collection of products, like context products. Would be in the model, described by the model and referenced by files that are then archived elsewhere. Agree that importing arbitrary XML schema is the wrong approach.

~ Not importing arbitrary schema, but you are extending the model.

~ The model is a set of schema generated by the ontology.

~ The opposite.

~ Want to separate the labels. Confusing the fact that we are importing schemas. They could be referenced by the data objects, not the labels. The labels would still be using schema in IM, only the schema we bless. Two completely separate models, side by side, one containing the other.

~ Someone understands, but there's a line in the sand. Same debate as we had for CDF. The data object no longer being described by the model. CDF had to come up with several LDDs specific to that format. The point is we should have a generic model to describe composite things. We keep dodging that. Very clearly not consistent with our original intent.

~ PDS4 labels and products are a new format. By injecting composite structures we make it very complicated to read or validate. Implementing complex structures makes it harder to have PDS4 drivers for all these applications. It is a line in the sand - allowing foreign formats to be in PDS4, like JPEG2000.

~ The original intent was that there would be a generic PDS4 reader for all of the archive. We started with simple formats, simple table etc. That was 85% of the data products. Now the 15% needs to be described with something that understands the composite structures. If we are not going to do that, fine, but we're crossing the line in the sand again. There's one composite structure you can describe everything with, but we constrained it.

~ Unfortunate to need readers for all these composite structures when readers exist for the other formats.

~ Regarding CDF - someone doesn't see this as being an analogy to this. Don't need an external reader to read the CDF files. The LDDs capture metadata. A good example is PDS4 Viewer - it can read every byte in there, but don't know if it could do that for the XML files.

~ Someone else agrees with that. Different subject. Went on for months. The point is that instead of using the composite structure we had to create LDDs. If had used composite structure we wouldn't have needed the extra LDDs. Similar analogy. How you describe it in PDS4 is the question.

~ Another person is not sure how practical it would be to redefine all the structures.

~ Someone thought we were talking about a limited set. Has to be constrained to very simple structures. Maybe we need some real examples and use cases.

Someone is confused. Trying to figure out who agrees with who. Likes having a limited white list. We don't want just anything. Maybe a rule that if there's an existing PDS4 format it could be archived in then you have to use that. Unclear who is agreeing with what.

~ If a constrained set of XML schema, someone is for that. Wants to see the white list. A cool idea - if you need a schema so software can read the data, then it's like a header, but we need to describe the relationship. Might be requiring Jordan to parse schema.

Question: Has anybody tried to to conceive of a complex PDS4 label to describe an XML file?

Answer: (Something about writing an XML ingest file - and more.) ~ Would automate it.

~ Someone created something for Bob Deen. A virtual structure. Described a very complex structure. Could do it for other use cases.

~ So, we are approaching a philosophical break - acknowledging some parsing is too complex. Thinks this will also come up for video. Hitting a practical limit. Think these are tied together if we are going to offload parsing from PDS4 to another standard.

~ If we do start crossing the line, what bounds. For me, video should only be a derived product. Maybe XML too.

~ Someone is not sure what that means.

~ The idea is to constrain what crosses the line.

~ We already allow FITS, but constrained by PDS and also requires an image label. Maybe. Also have to describe FITS with an array label. The question is if we can allow anything if we can describe it with a simple format. There are two issues here. Deciding what is not a simple data object. Issue is the point that we are already doing it for products with these headers. The other issue is multiples.

~ But the point of XML is that we can't describe it in the label.

~ We can with composite structure.

- ~ It might require us to tailor labels to data products.
- ~ A heavy lift. One exercise we need is to see if possible to describe an XML product in PDS4. Need actual examples and people to try to label them. If it's torture for us it will be worse for the DPs. For video too.
- ~ The question is how complex we are going to allow.

We want people to put example XML products in Jira and then we need a Tiger Torture Team to try to describe them in PDS4. Seems like a necessary step. Then it can be done for video.

- ~ That's the issue. Trying to decompose it to arrays and tables. Think we need something more than the simple products.
- ~ Someone thinks they get it. You are okay with XML, the issue is having the schema to organize it.
- ~ Yes. Need something for the not simple 15% of formats.
- ~ We have to allow complex formats.
- ~ If we are going to put so much constraint on these that they are reduced to simple tables and arrays, then we lose the added capability of XML.
- ~ It could be done for GIS.
- ~ IMG put a standard in CSV - a cheat that allows variable length. Did it by quoting it. Could quote an XML block in a table.
- ~ Good point. More people will find ways to cheat the system.
- ~ If DPs cheat the system, not sure if helps the users.
- ~ A lot of these formats help the users. Simple tables are hard to use to produce the plot.
- ~ The cheat for CSV is for users. Maybe that's a way to look at this. It was a well defined ISO standard.
- ~ For long term preservation, need to migrate output in simple formats. Wondering about that. Maybe we need a new class for user data.
- ~ That's what someone thinks of as derived. Derived for the user.
- ~ Issue is if it will be usable in 25 years.
- ~ If we can call it user data maybe we have a solution.

This ties into the annex products too.

- ~ We have a theme for the next six months.
- ~ Someone will do a session on this in November.
- ~ Good discussion.

Next meeting in two weeks.

DDWG Notes 2022-09-29

title: DDWG Notes 2022-09-29

layout: default

date: 2022-09-29

September 29, 2022

Notes by Debra Kazden

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

DDWG Agenda

(Included in email sent Sept 27, from M. Drum, DDWG Agenda for Thursday, September 29th)

Front-matter/What's new

- New issue: CCB-354 Permissible Value addition (Added to Tier 1)
- Approved SCRs by February 2023
- ****(Brief Discussion)****

Ready for Vote

- (none this week)

Tier 1 issues

- CCB-336 – Add a License Information class to the Identification Area – Tanya Lim - We may officially be in over our heads
- CCB-211 – XML/schema based files as archival data – Jesse Stone, Mike Drum - Good discussion last week
- CCB-354 – Add new permissible values to Header/parsing_standard_id to support new versions of CDF ****(Discussed)****

Tier 2 issues

- CCB-325 - Support for video as product observational – Trent Hare
- Higher order products (ticket ETA is October) – Lyle Huber - Ongoing discussion
- CCB-345 – Target Identification attributes name and type allow nillable – Trent Hare
- DOI-rich metadata into IM (ticket to be created), replacing CCB-222 and CCB-261 – Anne Raugh and any volunteers
- CCB-164 - Display Settings not required for images -- (IMG) - Need to think more about potential updates to Imaging and how Display might fit.
- CCB-353 – Fully label all bytes in data files ****(Some Discussion)****

Parking lot – need a driver!

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

Next meeting: 10/13/22

DDWG

Front Matter/what's new

We are late on attendance today.

A new issue has been created.

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

This is a new enumerated value one. It's from Joe. Shouldn't be too controversial.
~ CDF was contentious when it was added. It might not be uncontroversial.

Tier 1

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

Comments were added to Jira, including Anne's presentation to IPDA.

~ It's becoming clear that we may be in over our heads. The DDWG doesn't have the power to decide what licenses to use. Think the best we can do is to provide a generic structure.

~ The least we can do is provide the generic structure. Enumerations are problematic, but it's part of FAIR. We have to step up and do something.

~ Someone agrees. The downside to providing a generic structure is that we can be providing a way for people to get into trouble. Could cause problems that require lawyers and such. We need to be very careful and watch that the structure is used appropriately.

~ Someone else sees that point.

~ Another person agrees.

Question: Is FAIR required now? Thought it was guidelines. What's our responsibility to making sure we are following FAIR?

Answer: Other metadata standards have used licenses. There's some potential problems, but we can't hide it.

~ Anne thinks this will go all the way up to congress to decide. It won't be used for awhile. Part of why this is so hard. If things are not in the public domain we need copyright info too. Very complicated. Think we need consultation on what/how to gather the metadata. We need to distribute it, but it's not our decision.

~ At OIS level need access rights. The issue is copyright. In the meantime, maybe we should implement something like access rights, to allow us to do due diligence for FAIR. But it's not a copyright.

~ Yeah. Access constraints and use constraints. Access would be public and use, we could say, is constrained, and put a license or ISO terminology. A different term from license.

~ Yes, access rights. Some type of legal document that you have a right to hold the information.

Question: Can we do that? Are we allowed to determine that?

Answer: For USGS, access rights is allowed and no usage constraints.

~ So, maybe we should constrain this down to a structure and consult people on what we should do.

Question: What's bringing this up? Not CC0?

Answer: Neither we, nor the DPs, have the authority to decide what license is applied.

~ That's why we would use access rights. To show that we are allowed to have the data.

~ FAIR will be enforced at some point. Becky said a revision is coming.

~ Someone is interested in seeing what they come up with and how far they can take it.

~ As per SPD41 metadata...(read the policy) Maybe we can skirt the issue if we say it's a usage policy rather than license.

Question: What if we say we only accept FAIR compliant data?

Answer: Good point, but that's where we can get into trouble because if NASA funded something, but it's not FAIR compliant, lawyers will need to get involved.

We are not getting anywhere and spending a lot of time on this, yet this will be optional. And we are kicking the can.

~ Proposing the opposite. Get access rights in there - it's an ISO term. We can maintain, archive and document it. We could put it in. It shows good intentions.

~ We should discuss this with Anne.

~ Someone agrees. She has gone deep on this.

~ Beyond that, we need to bring it back to the MC.

~ Someone agrees that we are spending too much time on this. We need a blanket statement and maybe this doesn't belong in our labels. Maybe there should be a policy statement. We need guidance from above. We need to wait until this is actually settled.

~ There are lots of options. The original plan is probably not going to work. We will discuss this with the MC. We will move on for now.

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

Had a good discussion on this last time. Multiple philosophies going forward.

Question: Has anyone thought about this going forward? User products using more contemporary products?

Answer: Put in a discussion topic for annex data. More streamable data. Can go down a rabbit hole with this. Should keep it simple. We need examples.

~ None yet. Maybe we could move this to Tier 2 until we have some examples. Start with what we have.

~ Someone asked about this at a node meeting. Spreadsheet was mentioned, Excel format.

~ Yes, an XML format. Good example of something that can be converted.

~ Someone else has lots of examples. Seismic data from InSight and Mars 2020. Maybe MSL rover locations over time. There are XML cases out there. Straight forward XML.

Question: Are these currently archived but called text?

Answer: Yes.

Another Question: Could you put that in Jira for people to look at?

Answer: Yes. Think seismic data has an official schema. ****(Action Item - Ed)**** ~ People will look at that.

Someone else has an example being proposed for Europa Clipper, but the file type is under ITAR control, so can't share it with the DDWG.

Question: Could it be redacted?

Answer: Not proprietary, but stuck because of ITAR.

~ Maybe once we have some concrete examples we can make the proposal around them. We can try it, show them and see if we all think it will work.

~ There are a lot of hurdles to get over.

Tier 2

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

No update.

~ IMG wants to bring this up with higher order products. And annex products too. Feels all wrapped up, but no update. Don't know how to describe a video in a label. Will need to allow compression.

~ Unclear how we would describe it. Maybe some higher level way to describe. Don't want to invent our own video compression format.

~ The idea of creating a label that names every key frame, but that's what the video label is, so we have a format that redefines what the label does. Don't want to re-invent this like we did for CDF.

Question: Discuss this in November?

Answer: Yes.

CCB-345 - Target Identification attributes name and type allow nillable See <https://pds-jira.jpl.nasa.gov/browse/CCB-345>

There is a renewed need for this at SBN. Would like to get this through. The last comment was a while ago.

Question: Does the original proposer still have a need for this?

Answer: Took us down another rabbit hole with allowing more to be nillable in the system. Still needed, but it took a back seat. Some people are against it.

~ Someone strongly objects to having a nillable target. Just say what you know. Nil can be dangerous.

~ That's why it's on the back burner. The team that wanted to nil target is using solar system now.

The case at SBN would need more types if couldn't use nil. For ground based if the target isn't in the frame would want to add sky.

~ Someone wouldn't object to that. Okay with sky survey.

Question: Does that imply that it's on Earth? Could it be used as a Voyager target?

Answer: Okay with someone. Could do it from spacecraft.

~ Someone really dislikes the use of solar system. Sky survey would be better.

We need some new enumerations. Could do sky survey, sky, sky background, unknown, untargeted.

~ Could use name or description for unknown. Name is open. Could do multiple, if not targeting a specific target.

~ Would use sky survey for that.

~ Sky survey has positive connotations. It says just looking. This whole class is about the intended target, so multiple doesn't make sense. Would have sky survey. Wording that way guards against not just being lazy and using multiple.

~ Will run it by SBN.

Question: Can we vote on sky survey?

Another Question: Is multiple in the name?

Answer: Someone has used multiple asteroid.

~ We have a calibration field.

~ Not deprecated.

~ Sky or sky survey is okay.

~ Will get input. Can discuss this next time and maybe vote soon.

~ This could possibly open a can of worms. Link to the context product is optional.

Question: Want to further lock down name?

Answer: Don't know. Thought everything in target identifiers would have a context product.

~ We don't want to go down that road. Don't want a context product for every lunar sample, etc.

~ Could be computer generated.

~ Someone understands, but name is not very useful for search.

~ The concept of target, the way we are using it is way overloaded. It would be hard to have a list.

~ And since it's the observers intentions.

~ We can revisit this when we have a concrete search.

Someone withdraws the walk down that rabbit hole.

Higher Order Products

ATMOS, IMG and others have been discussing higher order products.

~ Some of what people want is not the same discussion. ATMOS, IMG and PPI are having annexes for things that cannot or will not be archived. Things like models that can't be in PDS4 format, but we want them to be able to be found. Want to clearly identify as not in the archive, but can be labeled with something like product annex. It would be like a product browse. Would be things that aren't tables or arrays, but encoded objects.

~ Great start. Likes where this is going. Could put video here, not so much with the name product annex. We will discuss that in November.

Question: Is this related to what was wanted by Rings for product ephemeris or supplemental?

Answer: Not exactly, but similar.

~ But that could be labeled as product whatever and point to the data.

~ People see similarities in the intent, but more usable and archivable.

~ Someone else was talking about simple higher order products. Those are product observational and do belong in the archive unless they are in a format that can't be archived.

~ Fuzzy. Wants higher order to be used for formats that aren't archivable.

~ Moving target.

Maybe the word product is the issue.

~ Product allows for things to be identified and registered.

~ That clarifies things.

~ Someone is unclear on what was wanted for the Rings product metadata supplemental or product ephemeris.

~ It was for updates.

~ Wanted metadata archived separate so could update metadata for archive products. It's part of the archive that gets updated. Wouldn't be in an annex.

Someone is struggling with the concept of an annex. If it's impossible for something to be archived, then there's something wrong in the standards. Seems inconsistent with the idea of an archive to have annex data.

~ Someone else has been thinking about our priorities, like longevity, usability, etc. It could help us decide what to do for video, XML data etc.

~ Something that is not archivable should go to PDE.

~ That didn't exist a year ago.

~ We don't need to build something for products that should go in PDE. NASA needs to decide how to attach PDS to PDE.

~ Another person thought there was a policy that products that can't be migrated still need to be archivable.

Question: Is there really no video format that would be usable in 50 years?

Answer: The issue is compression. Would argue that MP4 could be archived. Code is very important.

~ The archive was only supposed to include a few simple formats. There was a proposal from Bob Deen to archive product observational tiles. Difficult to archive. The components of his product were already archived.

~ Composite structure.

~ Yes, but voted down. Had to write an LDD.

~ Someone would like to see that SCR.

~ It was to do with nillable file products. (CCB-259 - Need a "Null Result" option for Product Observational, see <https://pds-jira.jpl.nasa.gov/browse/CCB-259>).

Question: So, it tried to do something to do a product and ended up with a null result?

Answer: He had several tiles. Used another component to try to bring them together.

~ People need to review that.

~ Label without a data product.

This is a big heavy topic. We need to stick to the proposals. We are making some progress. Good discussion.

DDWG Notes 2022-10-13

title: DDWG Notes 2022-10-13

layout: default

date: 2022-10-13

October 13, 2022

Notes by Debra Kazden

Known Attendees: M. Drum, M. Gordon, E. Guinness, T. Hare, L. Huber, R. Joyner, D. Kazden, J. Mafi, S. McLaughlin, B. Semenov and R. Simpson ****CHECK**** Known Observers: M. Crombie, B. Hirsch, D. Kahan, M. McAuley, C. Seyb and M. Tiscareno

DDWG Agenda

(Included in email sent October 11, from M. Drum, DDWG Agenda for October 13, 2022)

Front-matter/What's new

- New issues: CCB-355 (Add Funding Acknowledgement) and CCB-356 (Update Constraints on Author/Editor Lists)
- Topics for November F2F
- Approved SCRs by February 2023
- **** (Discussed - also CCB-357) ****

Ready for Vote

(none this week)

Tier 1 issues

- CCB-345 – Target Identification add Sky Survey target type (formerly: nillable) – Trent Hare - Vote next session
- CCB-211 – XML/schema based files as archival data – Jesse Stone, Mike Drum
- CCB-354 – Add new permissible values to Header/parsing standard id to support new versions of CDF – Joe Mafi
- CCB-355 – Add Funding Acknowledgement class to Citation Information Class – Myche McAuley - new
- **** (Discussed) ****

Tier 2 issues

- CCB-336 – Add a License Information class to the Identification Area – Tanya Lim - We may officially be in over our heads
- CCB-325 - Support for video as product observational – Trent Hare - Revisit in November
- Higher order products (ticket ETA is October) – Lyle Huber
- DOI-rich metadata into IM (ticket to be created), replacing CCB-222 and CCB-261 – Anne Raugh and any volunteers
- CCB-164 - Display Settings not required for images -- (IMG) - Need to think more about potential updates to Imaging and how Display might fit.
- CCB-353 – Fully label all bytes in data files – Anne
- CCB-356 – Update Constraints on Author/Editor Lists – Dick Simpson - new **** (Brief Discussion) ****

Parking lot – need a driver!

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

Next meeting: 10/27/22

DDWG

Front Matter/What's new

There are a lot of new people attending the meeting today.

There are three new issues this week to introduce.

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

New issue from Mike McAuley.

~ Noticed that there is a hole in trying to relate funding results to the archive. Becky thought it was important too. Funding information is included in papers. Thought we could fix this with an SCR.

Question: Fix acknowledgement text?

Another Question: In the bundle or everywhere?

Answer: Was thinking collection.

~ We need to allow for more than one funding source. Some data is funded by multiple sources.

~ It would be optional.

~ It might be better to have it in every label if there are different funding sources within one bundle.

~ This is the same problem we have with licenses. Need to decide if it's an umbrella.

Question: Is there a penalty for putting it in bundle, collection and products and letting DPs decide where it should go? Could be optional everywhere.

Answer: Could have more specificity if it's needed.

~ It could also just go into documentation.

~ True, but have heard PDS HQ say twice that they would like to track all the things that money went to. Documents might not be so trackable.

~ Contract numbers and awards are unique and trackable. The example in the SCR is real, but only an example.

Question: Would this be an appropriate thing to have in a separate dictionary? Unclear if the shape of these is always the same.

Answer: Good point. Maybe a dictionary could include the things people might want to search for.

~ We don't have to solve this today.

Someone doesn't think a lot of users will care about this.

~ Someone else considers HQ a user.

~ There is a use case. If wanting to look for data from a certain award, could look up the award number.

~ This is something we can explore.

~ We can keep it very open, plain and optional. Don't see a problem with this.

Action Item-Everyone put comments in Jira.

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

Dick is the reporter, but this is really from Ron. Validate is refusing to accept author lists that don't have a person. If it's an agency or company validate doesn't accept it or correctly validate it. This is an attempt to cover a broad range of authors and answers. Anne may have ideas on this. Ron thinks if we get it looking like Datacite it will be easy to extract information. Looking for a good solution.

Someone remembers when we were originally designing this. We tried to design it for the ADS database. It will still need to work with ADS. Having Anne on the work group will be good.

~ She's also still wanting to create an SCR on DOI rich metadata in the IM. This seems very related. Kind of complicated and scary.

Question: Any other comments?

Answer: (Silence)

CCB-357 - Create Product Annex

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

New SCR from Lyle.

~ We talked a lot about this last time. Wanted to have something that we could throw darts at. Ran it by Joe and Trent. Product annex would be similar to product browse. Encoding would need to be populated with things in an annex. Has start and stop time. Will also want bundle and collection annex types too so validation will know it all requires different validation. Will also need to update the SR to say this is not part of the archive, it's so things can be found and cited. Kind of like PDS3 Extras, but kind of not. A tier 1 issue to discuss and flesh out.

~ Great start. Unclear on one thing. Not sure if we are going to include data objects or if this is just references.

~ Yes, include data objects. File formats that can't be included in the PDS4 archive.

Question: Where would they be held?

Answer: In the annex. Annex will hold files that don't meet the standard. They could be short lived or not passed peer review. Things we want to be able to find and cite.

~ There are DDWG and MC issues here to work out. At the MC yesterday this was discussed. We don't want lazy DPs to just put their data in an annex instead of doing the work.

~ Absolutely. That's why item 3 is in the SCR. (Item 3 says "Create new Standards Reference subsection to explicitly state that Annex Products are only to be used for non-archival products.") ~ This is tricky. We need to have categories. We discussed it some. Troubled by products that are of long term interest that we can't archive, like streaming video.

~ Understands, but some things, like models, are short term.

~ Different category.

~ Yes, we need to be careful when we write the enumerated value list.

~ That makes someone feel better.

Question: We could have product extras. Can something be in an official archive and not in an annex?

Answer: To be discussed. Makes sense to try to keep annex products together, but could have an annex product in an archive bundle.

~ Someone would like to discuss that sort of thing in November. And edge cases.

~ We will talk to the MC then too.

Question: Said doing this so things can be findable and citable, but also temporary. Conflict. Need to consider that. Why would we want to make something that goes away citable?

Answer: DDWG will have to discuss that.

~ Also, in yesterday's conversation. Would like to think of this as PDE stuff, but that's not happening soon and nodes that are funded to do this need to step up. Maybe we need an enumerated list with why something is in the annex - model, incompatible, safed, not of long term interest, not passed peer review. A way to say why a data set is in the annex.

~ Great idea! Wrote it down. Will try to add that to the SCR.

~ It has to be nillable.

~ Could build this into the model. Not in love with the name product annex because it's destinational. Could group products into product non-archival.

~ Product model, product software, etc.

~ Someone is not sure we want to go that way. Make a bin for all the non-archival and then divide from there.

~ Some could become more archival over time.

We are into the weeds. We need a Tiger Team. We should move on.

~ Ron volunteers to be on the team.

Question to Ron: Will you be at the F2F?

Answer: No.

~ It will be discussed and then we can build a Tiger Team.

~ There should be a virtual option for the F2F.

~ Lyle, Trent and Joe are on the Tiger Team. Should have Ron and Steve too.

~ This is an important issue that should have been addressed long ago.

The Tiger Team can kick off the discussion and present at the F2F and MC.

F2F workshop

For Tuesday of the MC, time will be split between the DDWG and DSMP. Mike asked for a tally of the big issues people want to discuss. Knows about video already.

~ If there is time, we should start mapping out product annex. Will give us an idea where the Tiger Team will head.

~ People will gather their thoughts on this and consider the issues. Probably this will be the theme of the meeting.

Question: Is anyone strongly opposed to product annex?

Answer: Someone is not strongly opposed, but wants to understand what it is and how it fits. We already have ways to reference things inside and outside PDS. Unclear we need this.

Another Question: Would specific use cases help?

Answer: We could make some.

~ This goes back to showing why some things aren't archival.

~ This can be a Pandora's Box, but having a product puts them into the system. Hooks to allow for DOIs, references, etc.

~ Someone thinks how he feels depends on the constraints. This is more of a PDE issue, and NASA needs to come up with a generic alternative repository. Not sure what constraints they would put on it, but it's up to us because they haven't done it. It all comes down to constraints. Sees there is a need for something.

~ There's a worry that people will choose the alternative repository for things that really do belong in PDS. A real concern. Need to assure that doesn't happen.

So, the topics we are planning to discuss at the DDWG F2F are product annex and video. Not sure there's time for more.

~ Someone agrees. A good opportunity for a long discussion.

Question: Don't we also have DSMP wrap up?

Answer: Yes.

~ And there's an issue, something we may want to discuss there. Intending to submit an SCR to relax prohibitions on IEEE754, nans and inf - concerned it may cause fights. Will put together a presentation.

~ Someone had to speak to that at a lunar meeting yesterday. Told them that it wasn't allowed in our archive. Definitely a concern.

~ Definitely need some discussion. Could put a presentation together for a DDWG meeting or the F2F.

~ Someone doesn't know enough about the issue to decide.

****Action Item - Joe**** will put together a presentation for next week. Then we can decide if we want to discuss it at the F2F.

SCRs

CCB-345 - Target Identification attribute target, add Sky Survey enumeration See <https://pds-jira.jpl.nasa.gov/browse/CCB-345>

SCR was fixed to add sky survey. Ran it by the folks at SBN. It solves the issues there. Better than nillable.

~ Someone didn't see a definition for sky survey in the SCR. Not sure it will solve all the issues. Suggests for legacy data that we should have unknown and not dump everything into sky survey.

~ Someone doesn't like sky survey.

~ It's more descriptive of the investigation.

Question: Can't we change it to none? Not everything is looking at the sky. There's not always a defined target.

Another Question: How is none different from unknown or nil?

Answer: In the Voyager case they are not looking at nothing.

~ A target is something you are aiming at.

Another Question: What about unspecified?

Answer: This is not restricted to legacy data. Some are sky surveys, but some things point into the middle of no where. Those don't belong in the same bucket.

~ Sky survey could be defined in other ways too.

- ~ In PDS3 we had dark sky and star field.
- ~ We have star field for calibration, but unsure when things are pointing into random directions.
- ~ Someone still likes going with sky.
- ~ Some one else needs to see definitions.
- ~ Sky is very image centric. Unsure what would be better for radio science or magnetometers.
- ~ We just need to solve this for identified cases for now.

Question: So, change this to sky and add a definition?

Answer: Yes.

Action Item - Trent

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

No progress yet. Got examples to use to create use cases.

- ~ XML could be a product annex.
- ~ These are archive products.
- ~ Someone sees that.
- ~ We need to think about the use cases first.

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

Reporter saw the comment and needs to respond. The intent was to use version 3.4 or later. Need to make sure modifications to later versions won't make PDS4 compliant CDF impossible. Need to verify that's true.

- ~ Someone will watch the comments.

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

Chair emailed the project office and put this in tier 2. It might go to the parking lot while we wait to hear.

CCB-350 - Adopt more rigorous / stringent rules for leap seconds and use of "24:00:00" in datetimes. See <https://pds-jira.jpl.nasa.gov/browse/CCB-350>

Someone asked the status of the date format, leap seconds SCR. Wanted to know if it has passed.

- ~ It was originally thought to be a bug fix, but some are possible changes for the SR. It was passed back to Ron. Unsure where it is now.
- ~ Seems like we said to tie it off and move on, not figure out all possible validation issues with leap seconds and years.
- ~ Correct, but what we can do is better than what we have now.
- ~ This will be put back on the agenda.
- ~ Leap years won't be too hard to define.
- ~ But there are some exceptions.

Thanks for the reminder on this one.

DSMP

We will have some time to talk on Tuesday of the F2F. The DDWG will be part of the day and the DSMP will be part. The focus of the DSMP will be to have a small Tiger Team update the DPH, SR and do other editing. Want to use the time Tuesday to decide what is required documentation, what are the required classes and regression testing. If someone else thinks there's more to discuss let Mitch know. The team will be one person for SBN and would like someone from GEO and IMG too. Need to know who will participate.

~ Mike and Mitch will decide how to divide the time up on Tuesday.

DDWG Notes 2022-10-27

title: DDWG Notes 2022-10-27

layout: default

date: 2022-10-27

October 27, 2022

Notes by Debra Kazden

Known Attendees: M. Drum, M. Gordon, E. Guinness, T. Hare, L. Huber, R. Joyner, D. Kazden, J. Mafi, S. McLaughlin and R. Simpson
Known Observers: P. Adhikar, D. Kahan, P. Lawton, M. McAuley, L. Neakrase and M. Tiscareno

DDWG Agenda

(Included in email sent October 25, from M. Drum, DDWG Agenda for October 27, 2022)

Front-matter/What's new

1) Topics for Tuesday afternoon workshop @ November F2F a. Annex b. Video c. Possibly IEEE754

2) Approved SCRs by February 2023

(Discussed)

Ready for Vote

- CCB-345 – Target Identification add Sky target type (formerly: nillable) – Trent Hare ****(Voted to Pass)****

Tier 1 issues

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

- CCB-354 – Add new permissible values to Header/parsing standard id to support new versions of CDF – Joe Mafi - added definitions and change logs

- CCB-355 – Add Funding Acknowledgement class to Citation Information Class – Myche McAuley

(Discussed)

Tier 2 issues

- CCB-336 – Add a License Information class to the Identification Area – Tanya Lim - Good progress made with HQ/PO

- CCB-325 - Support for video as product observational – Trent Hare - Revisit in November

- CCB-357 – Create Product Annex – Lyle Huber

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

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

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

- CCB-356 – Update Constraints on Author/Editor Lists – Dick Simpson - new ****(Some Discussion)****

Parking lot – need a driver!

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

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

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

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

** (Not Discussed) **

Next meeting: 11/10/22

DDWG

Front Matter/What's new

We talked about the F2F last time. It will be Tuesday, November 15. The DSMP discussion will be in the morning, and the DDWG will meet in the afternoon. We will discuss annex, video, and maybe IEEE754.

Question: Has Laura given us information about the room yet?

Answer: Not yet, but someone understands it will be in the Knight Center.

~ Asking because of splinter groups and the meeting on the second and third days. Wondering how that will be.

Another Question: Are there lounges or something?

Answer: There are some tables and chairs. It depends on where the meeting room is.

~ Someone thought Ben and others working on the documentation could be somewhere at a table with access to people in the meeting.

~ That depends on where we are. Should contact Laura to see if she can contact the Knight Center. Ed will be out of town the two weeks before the meeting.

~ It would be convenient to know the room and floor.

Ed will be trying to bring in another GEO person in case he ever retires. Would like to get them on the mailing list. It's a grad student from U of A.

~ Send Vivian email. CC Mike.

Ready for Vote

CCB-345 - Target Identification attribute target, add Sky enumeration See <https://pds-jira.jpl.nasa.gov/browse/CCB-345>

There were some comments in the last few days and a definition change, just to finish up the issue of what we are voting on.

Question: Are there any last minute comments?

Answer: (Silence)

**The Vote for CCB-345

ATMOS - Yes

EN - Yes

GEO - Yes

IMG - Yes

IPDA - Absent

NAIF - Absent

PPI - Yes
RMS - Yes
RS - Yes
SBN - Yes**

That will help out.

Tier 1 and 2 Issues

CCB-257 - Create Product Annex

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

Question: Should we discuss product annex today or wait for the F2F?

Answer: Someone wondered if there was any feedback yet. Haven't seen the SCR.

~ The hardest thing is the name. If people could start thinking about that.

~ Someone has some ideas. Will add them to Jira.

There are also some TBDs, like if products and collections can be interspersed with the archive or if they need to all be stand alones and the wording for the SR.

~ We also might want to think about the process for things going away from the annex someday. If they have a DOI we should send it to NSSDCA so it doesn't go away.

~ Not sure the NSSDCA would want it.

~ If it has a DOI, a persistent identifier, it should be around awhile. NSSDCA might be a place to move it to to keep around.

~ The NSSDCA MOU says they accept valid PDS3 and PDS4 products, but they could work with us outside the PDS interface.

~ They are not obligated to accept non-archival products.

~ Correct, but they can work with you to ingest and make it searchable in their catalog.

~ We might need to discuss this with the DOI work group. DOIs are in their infancy. We might need a way to sunset DOIs.

~ Might be a registry issue too.

~ That's true for superseded data too.

~ The SCR needs to consider the entire process.

Question for NSSDCA: If the products do end up at NSSDCA could they have a direct link?

Answer: Yes, but they wouldn't keep the DOI. Would assign a PDS identifier. We don't track DOIs at the NSSDCA.

~ Would need a persistent link.

~ At best it would be a collection link ID at NSSDCA. Can search the catalog there.

~ DOIs are supposed to be more direct, but okay if they can get there with NSSDCA ID.

~ Someone is not sure those would be compliant landing pages.

~ We don't assign DOIs to products, but we do to collections. Would think annex would maintain the landing page and put where it's at at NSSDCA. Not sure DOI data has to be online, just have to be able to get to it.

~ It seems strange to put this at NSSDCA.

~ Will need more discussion between annex and NSSDCA because they don't maintain landing pages.

~ We will figure it out.

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

The examples still need to be worked on. Will do that before the F2F so we can discuss them, but not making it a F2F topic.

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

There were some changes.

Question: How close are we to finalizing this for a vote?

Answer: Responded to comments in Jira. Still going through changes in versions to see if there's anything problematic for CDF/A. Don't expect that.

~ Added a change log.

Question: How challenging is it to push to a different version?

Answer: The main changes are to the tools. Main issue is not to affect the ability to produce valid CDF/A. We already have requirements. The goal is to make sure newer versions don't make it impossible.

~ Someone sees this as an enumeration change.

~ Valid point, but some versions would make it non-PDS compliant. Some versions before 3.4 can't create a PDS4 compliant format.

~ Probably too late to separate this from the version.

~ 3.4 is the minimum version we want to accept. Other enumerations for the attribute could be a separate SCR.

~ When we did TIFF 6.0 it was difficult.

~ Being tied to a specific version is not specifically bad, prevents us hurting ourselves.

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

We discussed last time that this could be in an LDD.

~ It should be applicable for everything. We can try to do a small sub-model for citation augmentation. Could do it as an LDD change.

~ If in an LDD it can't live in the the citation information.

~ Seems like it should live there.

~ Someone doesn't agree with the comment in Jira.

~ Citation description is being discussed now - making it broader for DOIs. No movement lately. Can still potentially go forward with this. The question is if this is relevant to citation.

~ That's out of someone's wheel house, but it seems married to a citation thing.

~ Alternative is to put in a class under the identification area and make it optional. Seems like if I'm citing data I used I don't cite the funding source. Could be a class in the identification area.

~ This is out of someone's wheel house.

~ Someone had said that this would be good for search. Unclear that people would do that. Could put it in the documentation.

~ Want to associate funding organization to results in the archive.

~ Problems if it's optional. Every bundle has a documentation.

~ Was looking for a more readily available way to search.

We could ask at the MC if NASA wants this.

~ Bean counters will want to track stuff. Providers could be stuck doing it.

~ Believe LDDs came up because we weren't sure if the structure would be the same for different types of awards.

~ The question is if we need to do this at the product level for all stuff or if we can do it at the collection or bundle level.

~ That was brought up.

~ We require input from the bosses. Will need to bring this up at the MC. We need HQ to weigh in.

****Action Item - Mike**** will bring this up at the MC.

~ It's a good idea to find out what HQ thinks.

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

A brief update - there was a productive meeting with Steve Crawford and the PO on how to approach this. The outcome was that there's reason to believe that most missions are funded under terms that give up data rights and can be CC0, but there are plenty of exceptions - international, individual data providers, CLPS, etc., so we probably need to build something for more restrictive licenses. PO and nodes will need to reach out and make sure they are okay with us applying that license. No action items yet.

~ Someone who was at that meeting wants to thank Mike for taking the lead on this.

Question: Will language be in the initial award letter?

Answer: Steve said there is language in the contracts about giving up data rights which would equate to us being able to use a license, but no one's thought about it.

~ There are two categories - missions and everyone else. Unclear where the cut off is for missions to make everything publically available. Bigger problem for R&A programs and complications from journals. Haven't thought about this ever. We may need to contact journals.

~ Someone thinks they have a copyright on the publication, not the content. Also, IPDA first submitted this request. Third category.

~ Yes, hopefully can still use CC licenses.

~ Our problem is that we'll put a class up, most US data will be CC0, not sure we want to be the gatekeeper. More concerned about migrations. For new data they can just tell us what to use. Asked if something is submitted to PDS if it is automatically public domain and the answer was no.

~ Not a lot of escape hatches.

Compressed Video

Someone will be shooting for compressed video as a full up product observational. Would like to know if there are any objections.

~ Someone has reservations. Concerned about the long term.

~ Someone else is concerned with product native.

~ Product observational is supposed to be simple formats.

~ Another person agrees. And has concerns.

~ We will have to acknowledge that this is something different. Unprecedented.

~ This is where someone wants to throw it into product annex, but not if it can't be in product collection

and bundle. We can discuss this in November.

~ Someone has less reservations with that.

~ We need to capture the data.

CCB-350 - Adopt more rigorous / stringent rules for leap seconds and use of "24:00:00" in datetimes.

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

This was added back to our list. Still being worked on. Will need IM change for earlier dates. Nothing to report today.

IEEE754

Regarding nans and infs, been doing some research on this. The only place this is mentioned in the SR is in the 19.28 ASCII Real section. The question is if it was intended to apply everywhere. If yes, we need an SCR, but if it really is only for ASCII real then we just need an issue on validate not to report nan and infs in binary types.

Question: Why is nan a problem?

Answer: (Very hard to hear)... one too many...

~ And quiet nans.

~ Problem for Mars 2020 - they had to reprocess to get rid of nans. Re-ran the calibration.

~ So, in the end it was a good thing.

Question: Will we define a valid nan?

Answer: IEEE defines them. We could have an LDD.

~ Different hardware has different interpretations. Not all the same thing.

~ There's more work to do.

~ If we want to clarify the SR that that restriction is only for ASCII real we might need an SCR for better documentation.

DDWG Notes 2022-11-15

title: DDWG Notes 2022-11-15

layout: default

date: 2022-11-15

November 15, 2022

Notes by Debra Kazden

*There are also notes by Laura Givens.

DDWG F2F

Annex

Some nodes are housing data that needs to be discoverable with LIDs and be registered. A big issue is that there are lots of products that are not appropriate for an archive as-is, but would be nice to have. PDS4 is for long term usability. The specific issue now is the annex.

The concept behind this is that there are items that belong in a planetary repository which are not archive worthy for various reasons. Some of these are easily superseded, temporary or transitory. Rings supplemental products may or may not belong in this category. Things might have been included in PDS3 as extras. Some are not PDS4 compliant.

There are also some that need to be available pre-peer review for publications with DOIs and LIDs. DOIs can be reserved ahead of time.

Have to have it done when manuscripts are submitted.

A chicken and egg game.

These are some of the reasons. There are a few tall poles to hash out:

- Do these need to be separate from the archive products or can they be in a bundle? (Someone thinks they should be separate.)
- We need to write a SR section that says these are specifically not part of the archive, but will follow the other rules for PDS4.
- Need to populate an enumerated list for encoding standards for annex products.

Transitory products being things people might not want long term. Other products are safed. There's a fuzzy in-between part too - not PDS archival, not exactly transitory or contemporary.

Some things shouldn't have archive LIDs because they will later be in the archive.

A separate category is non-compliant formats. PDS problem - should be a goal to put them in a compliant format. Transitory PDS formats that PDS4 isn't ready for.

All part of the PDE, not the archive.

Question: Getting data published for a journal - data repository - is that a PDS or PDE problem?

Answer: PDE.

~ Some nodes are doing PDE stuff.

~ Everything PDS is PDE.

~ Rings does things that are archival.

~ All data in the PPI annex are things that we would have wanted. The most recent products are polar maps of the moon.

One thing to note - every four or five years record retention schedules are brought up by HQ. PDS is a permanent record. There is a document that says the National Archives has an agreement with the space agencies that this is a permanent record. Annex is harder to defend as a permanent record.

Regarding the PPI latest product, sounds like the PI got flat footed. Had an unfunded mandate from a journal. PDS shouldn't be doing work arounds because a PI didn't plan ahead.

FAIR hasn't been around that long. Proposals from before that and how journals interpret FAIR are issues too.

PDE may be able to influence the journals.

It's becoming more clear what is expected. Lots of potential gotchas. SPD41 is still in comment mode.

A lot of philosophical discussions. We need to talk nitty-gritty.

Journals aren't saying data has to be in PDS. PIs are aware PDS is the right place. We aren't obligated to do it.

Someone agrees. Encouraged by the idea of us getting ahead of it and defining this. Can use that and integrate with our own system, rather than having to integrate with several other possible systems.

That's the point. Want to use the registry and make use of PDS4 and have this be citeable. So, the question is if we give a DOI to something non-permanent, does that imply permanence?

We can have tombstone pages.

Question: So a journal will accept and point to a page that's no longer there?

Answer: Last resort for things that get a permanent identifier.

~ Will be expecting that by the time of the gravestone that there will be an update.

Regarding the original questions - that assumes we are doing product annex. We still haven't answered that.

We should define what this looks like. It can live elsewhere.

That's NASA's decision. If they decide PDS4 is going to be the structure for the whole PDE, that's fine.

Really think SMD needs a plan for NASA data, including a repository for things that can be curated up to the archive, but doesn't have to meet all the standards to be in the repository. Could be transitory,

maybe archival later. Should be SMD wide.

When we say PDS4 we are thinking archive, but that's a metadata definition.

Yes, so if PDS could help with extending the IM for repositories that would be very helpful.

If we extend the IM to archive and repository there's no impact on the permanent record issue.

Chief archivist at the National Archives sets the standards. Have staved them off so far, but as bureaucrats they would want to see record retention plans.

We can cross that bridge when we get there.

If we get to that point, having product annex for repository, it doesn't hurt us. Just a definition. If we supply the repository it needs to be separate so the archive still complies. Consensus seems to be the model should be extended.

We are taking data in an annex. At the last DDWG there was no strong opposition.

Regarding the questions - Yes to number 1 if NASA wants us to extend beyond the IM.

There needs to be a clear line between the archive and other activities, software, trainings, annex. Extending the IM is consistent with the PDE IRB. A caveat is that we wouldn't want to see such strict requirements on an annex that archive stuff becomes less valuable.

We need to write the SR section very carefully.

At the SMD level, as it gets it's feet, we can say in planetary we have this and see if we can integrate. Sees an SMD repository in the future. Some products could go to the deep archive. Need to be able to easily pull things out.

If that's the case there's a nomenclature problem. We use PDS4 to mean IM and archive. Things for PDE should be pretty minimal. Need a bundle product and whatever Moses wants.

If we want to extend the IM there are aspects of the IM that summarize knowledge. Can be any context. Fine. Key will be to completely separate documentation. Bundle is an archiving concept. Only need it for the archive. Can have lightweight IM model and let go of some of the archive metadata.

Need a bundle for LID.

Question: Why can't it be a bundle?

Answer: You lose people when you talk about bundles. They understand the concept of collection.

~ Let's table this awhile.

Yes is the answer to should we extend the IM for annex data.

For question 2, should annex products be interleaved with other products the answer is no.

Question 3 is which standards should annex products comply with.

Question: Should we provide full support for annex products?

Answer: We won't be able to support some of this.

~ What about when someone wants an explanation about products? Should be in the policy.

~ This is here as a courtesy.

~ Tell people to contact the PI for explanations. PDE wants to participate in the findability whether it's in the PDS archive or repository. Would be great if search could provide all answers.

Question 4 is do annex products imply the full support of PDS. The answer is TBD.

It may behoove PDS to get ahead of defining this.

Nitty-gritty. Steve Hughes has a deep dive available.

Number 2 is no.

People are not sure what that means.

We used to throw things into PDS3 Extras.

Someone supports the idea of them being findable.

Others agree. A bundle has all it's documents, data collections, etc.

Question: And it can all be found in search for the archive or repository?

Answer: Yes.

~ Could have external references, constrained somehow.

~ Good idea to have front of LID not be NASA PDS.

~ Co-locating might be helpful. There's useful stuff in the PDS3 extras.

Question 3 - PDS4 SR must be appended. It has to be in the SR if we are going to extend the IM. Can be an appendix of the SR and say these exist and are not part of the archive.

~ Repository, non-archival data.

Need to write a separate document so we are not confusing PDS4 with archive. It should be how to label your data.

Should say repository instead of annex.

~ Someone is not sure repository is better.

Of 4 questions, 2.5 are answered so far.

For question 4, we need a policy to establish limits.

Yes, but we can't do that until we know what the products are.

Support of annexes should not come at the expense of archives and not take away resources to support the primary archive.

Question 4 - Not for products outside of PDS. PDS is pointing users to these products as a courtesy to the community. If the nodes want to host an annex, it's up to the nodes.

NASA support for the annex comes from HQ.

Stuff in Zenodo could be pointed to.

A lot of problems go back to the name annex. Using annex here implies a PDS relationship. IMG, ATMOS and PPI using repositories. Might want a different word.

NASA PDE should define the standards. And we should have an enumerated list for why a data set is not in the archive.

Something besides the DP didn't feel like meeting the archiving obligation.

Need to be clear.

That's offensive to DPs who are doing more work than just dumping it into Zenodo.

That's more at the proposal review stage - if a DP says they will be archiving in a repository.

PDS isn't where all R&A data goes.

Someone agrees. For missions, yes.

Right.

The main products from missions should be in the archive.

Concerned about smaller DPs who don't have funding for the full archiving job. Don't want to dissuade them from doing a good job.

There are huge European products that aren't funded, but DPs want to store their data long term. Really supports the repository idea, but ESA doesn't have power to impose repository standards. Having this to point to would be very helpful.

Presentation by Steve Hughes

The discussion has helped us understand where we are going.

Slide 1 - Extending the model, product annex, etc.

Slide 2 - CCB-357 is to create product annex.

Slide 3 - There are three topics to discuss

Slide 4 - What is an archive?

archive
new concept - "trustworthy archive"

The terms archive and repository are often used interchangeably.

PDS does a great job of meeting long term requirements of their community.

Slide 5 - Aspects of a trustworthy archive

Slide 6 - Describing the data object

Slide 7 - Representation/Format

PDS3 in the archive - scientifically useful.

PDS4 - readily useful - right out of the archive with simple formats.

The PDE documentation defines readily useful and analysis ready different. A higher tier of data requirements.

Defined by a specific scientist who wants to do something specific for analysis ready, but we also want that lower tier of readily useful.

We archive what missions send us. Hard to promise analysis ready.

PDS3 accepted data in any format and it became obsolete. Made PDS4 simple formats suitable for archive. Planned to have tools to convert to contemporary formats. Flaw in that was a lack of funding for EN to make the conversion tools.

Someone agrees completely. We promised the community that, and we didn't do it. Promised when we developed PDS4.

Analysis ready is why the IMG annex started.

Means it's a particular way - different from readily usable. Some processing is already done.

Question: Why are we discussing analysis ready? Others are talking about conversion tools.

Answer: The term analysis ready is already getting muddy. People who use FITS consider that analysis ready.

~ Analysis ready is a higher bar than what we try to address.

Slide 8 - Readily Usable

Basic requirements for your community.

Slide 9 - Topic.

Slide 10 - Current modeling paradigm

Details are tagged digital objects.

Slide 11 - Formal structures, tools and languages.

Slide 12 - Application of modeling paradigm to PDS4 Options - option one - product observational

Slide 13 - Option 1.5 - product native

Slide 14 - Option 2 - product context

Slide 15 - Option 3 - product XXX (browse, bundle, collection, document, etc)

Some things here are similar to product-observational. Need to go into the file area to know what they are.

Slide 16 - Option 4 - LDD parent and child classes

Slide 17 - Take Away

Slide 18 - Topic

Slide 19 - Possible paths forward

Slide 20 - Augment the existing modeling paradigm

Slide 21 - Merge option 1.5 and 2. (Written before knew what the discussions today would be. We should be thinking of this as archive usability, not anything else.)

Enumerated list would be applicable for product annex.

Yes, same IM, with different requirements.

Question: Why restrict this not to include array?

Answer: Could.

That's the end of the slides.

The point is to show that the model is easy to extend, just need to put in correct constraints.

Seems complicated, but when people put data in Zenodo they just fill in a form. Could do that for planetary.

And add aspects for rich DOI metadata.

The ATMOS annex will be presented at the MC tomorrow. Trent and Ray will be presenting on Thursday. Things to think about and be ready to discuss at our next DDWG meeting.

Last statement was data object, but there are services too.

Someone is not sure they see this as watered down, think we could have all the rich metadata in the

product annex products, unless all we want is findability. Labels were sounding very watered down. Two competing goods. Need to find balance. Can suggest the rich metadata, but NASA needs to decide. Steve is working on an IM for all JPL repositories - using an LDD for if want more metadata.

A lot of this has non-PDS metadata, USGS maps, etc. Would be nice to figure that out.

People would like to put things like net-CDF in repositories. Very hard to shoehorn them into PDS data types, so even with new products there will still be things we won't be able to describe.

That's the point of the encoding standard list - enumerated list for product annex. Could put net-CDF in the list.

It may still be moderately difficult to fit into the standard and file area.

File area just says go read the standard.

This can be solved by software. The registry can be extended too.

Most of what's going to be in the ATMOS annex will be in net-CDF.

Question: The registry will know who is doing that?

Answer: Can wrap in standard libraries.

~ NAIF wants a policy to say there is no commitment to support it.

~ Nodes can do their best effort based on user priorities. Best effort based on needs of the community.

PDS is designed with principals. If talking about re-usability and interoperability should look at what euro-planet and IVOA are doing. Don't think the best solution is to take the archiving solution. Need to solve problems in a way that supports our community.

Someone agrees.

That's a four year solution. Need a one year solution.

It's easier than you think. Just need a DOI. Don't need a LID.

Want LID for search.

You don't search on LIDs. You don't need all that overhead.

Could all be product nasa:pds:product-annex. Don't need so many layers.

LIDs are for the user friendliness. People search on metadata.

Can assign random identifiers.

Had LIDs for uniqueness.

In the archive.

But that's been broken. There are already products with two different version 1 LIDs.

We are getting into implementation. We can come back to this later.

Video as Product Observational by Trent Hare (for annex)

Slide 1 - Issue for the last year or so. Trying to understand if we can move forward. Video breaks the policy on formats.

Slide 2 - Policy on formats for PDS4 data and documentation. Video is not allowed as a first class citizen in an archive. We did accept MPEG-4 as a document. Important because video is becoming a primary product on missions (Mars 2020, Artemis, etc.) also, there are proposals where the primary product is video.

Slide(s) with pictures that are data.

Video can't be uncompressed or it's no longer usable. Even loss-less is too slow to be used on a computer. Question is if we can move past the idea that PDS4 will not allow compression.

Might be a question for more than just video, but the question now is how to get video in PDS4. Pandora's Box.

Question: Is it okay to require a compliant version of MPEG-4 for PDS4 and the original as a product native?

Answer: Okay to require a compliant version. Could require product native, but a lot of times it needs to be converted anyway. Tried with the first SCR to use product native for video.

Question: If we agree MPEG-4 is acceptable, should we also preserve the original format?

Answer: Think that should be up to the DP. They should have the option to do product native.

Assuming we accept MPEG-4 for product observational, then we would want significant restrictions, especially on compression.

If we are not trying to further compress data...

There are so many video formats that we need a standardized video format. Would use the video as a product. Paper with a specific standard product.

The paper describes video and audio. Audio is the next step. Uncompressed audio.

This is a PDS problem. When we designed PDS4 we wanted simple formats. We're now at the point where we have to look at those again. Things that were a nightmare, aren't a nightmare now. So, not ready to say no, but we will have to go to the MC.

Pandora's Box is a concern. Need to get past that. Take it one step at a time.

Need to avoid abuse. Need a solution that doesn't strain our system.

Question: Wave is PDS4 compliant?

Answer: An array of sound bytes. Binary table.

Question: JPEG 2000?

Answer: Document or browse for PDS4.

Question: Are you making a proposal?

Answer: It stalled in the DDWG. Have to go to the MC.

~ What if we just have another bullet point - video and audio files with certain caviots?

~ Also product native with the exact product that came off the spacecraft.

Question: What are the concerns about loss in the compressed format?

Answer: Confident can set quality flags to make it a loss-less conversion.

~ And include product native.

~ That will be harder than it sounds.

We will need to expand product native with a LDD.

Thought it could be anything. Worried that if we force everyone to have product observational and native it will be massive.

~ Less than if we translate every byte. Can buy more hard drives.

Not all video being taken should be product observational. Maybe some is documents. But we have implied that observational data is first class.

Question: What comes first? A new SCR?

Answer: SCR to allow video as compressed formats and audio.

We already have a solution for audio.

~ Uncompressed.

PDS is supposed to make sure the data is here for the future. Sounds like we are saying the format we want for future generations and otherwise would be orders of magnitude of resources.

The other issue is usability.

Issue of the .1% of it that might be lost and future usability. Unclear how to make sure it can be used in 20 years.

Migrate it in 20 years.

Need to consider the economics of the trade off. A lot of disk space now versus ten years versus migrating in 20 years. With MPEG-4, someone will write code to migrate that. It's a trade off to think about.

We have migrated data from PDS3 to PDS4 that weren't PDS4 compliant. Migration was complicated.

Confident MPEG-4 standard will be around a long time. Capability won't get lost.

Same trade off we made deciding to accept PDF.

So, the answer is yes. Raise the SCR. Take it to the DDWG, then if it passes the CCB it will go to the MC for an update to the policy.

Everything is in line. Just wanted people to know.

Just open the corner of Pandora's Box.

That's what we need to be careful of.

Another issue is if validate tool can validate it.

If this is encoded data than it needs a different product type because there are things that need to be included for validate.

Question: Horrified at the idea of migrating everything again in the future. Sounds like we should have product video. Is there anything comparable to wave?

Answer: (technical talk - factor of 10, 100 , compression - something...) ~ Bottom line is don't think we'll be migrating these in 20 years. Can review them every decade.

Question: What about product observational compressed?

Answer: Or product audio video so we don't have to do it for everything.

~ Could leave as document, but it needs to be a first class citizen.

Documents aren't second class citizens. If we want to use product observational we need to make sure there's metadata that describes the compression.

That's why we should have a class of compressed.

Comments in chat:

~ From Mark B. - Very interesting discussion. Likes the idea of some product video thing.

~ From John B. - Talking about audio/video, but also shape files. Might want to archive them in the future. Maybe product ancillary as a catch all to indicate that it might not be in a format that will be around in 20 years.

~ We already archive shape models. We do have product ancillary, but it is used for things for support information.

~ Naming debates are our favorite distraction in PDS4.

~ Just meant to have a bin we could chuck things into and then look at later to make sure they are still relevant.

One risk is that it does make it easier to shove more in without traditional requirements.

Someone agrees. Product observational makes it easier to find the data. We need a way to identify video data. The primary observation should be added as a specific type for product observational or we should make a product observational audio video that looks just like product observational.

Of those, the DDWG can work out, but agrees. Audio and video are unique in needing specialized attention right now. Especially since SBN is already archiving shape models.

We really need to constrain this.

We need a proposal with requirements in Jira.

Someone doesn't understand the implications of adding a new product observational.

Revisiting in 20 years is an issue.

~ VAX reads in vanilla in PDS format. No longer doing that. Not sure. Think we will still get video from spacecrafts in 20 years.

Might be superseded by 3D MPEG-4 in 20 years.

Hard to discuss in the DDWG.

Can review and come up with action items.

Laura will send action items by email.

Question: Anything else today?

Answer: (Silence)

DDWG Notes 2022-12-08

title: DDWG Notes 2022-12-08

layout: default

date: 2022-12-08

December 8, 2022

Notes by Debra Kazden

(Version 2- January 4, 2020 - Includes the following text edit:

Version 1: "to loose with saying PDS4. It's an adjective. We are talking about archiving standards now."

Version 2: "too loose with saying PDS4. It's an adjective. There is the PDS4 information model, and there is the PDS4 archiving standard. We are talking about archiving standards now.")

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

DDWG Agenda

(Included in email sent December 6, from M. Drum, DDWG Agenda for December 8, 2022)

Front-matter/What's new

- Welcome Sheri Loftin?
- Product Metadata Errata?
- Create a supplemental label to live next to original label, that can add aliases and supersede metadata
- Improve SR around bundle/collection names?
[https://urldefense.us/v3/https://github.com/NASA-PDS/validate/issues/561_!!PvBDto6Hs4WbVuu7!aLYEeXh8Ls1vWFnckIFHK5MtCENpK5eVxpzQefXRHx9z7_kwWJxn6Ww8diGbo5nvw1jUf891akA\\$](https://urldefense.us/v3/https://github.com/NASA-PDS/validate/issues/561_!!PvBDto6Hs4WbVuu7!aLYEeXh8Ls1vWFnckIFHK5MtCENpK5eVxpzQefXRHx9z7_kwWJxn6Ww8diGbo5nvw1jUf891akA$)
- Approved SCRs by February 2023
- ****(Discussed)****

Ready for Vote

- None this week

Tier 1 issues

- CCB-357 – Create Product Annex – Lyle Huber
- CCB-325 - Support for video as product observational – Trent Hare
- CCB-211 – XML/schema based files as archival data – Jesse Stone, Mike Drum - Have some examples, will post in JIRA over December break
- CCB-354 – Add new permissible values to Header/parsing standard id to support new versions of CDF – Joe Mafi - added definitions and changelogs
- CCB-355 – Add Funding Acknowledgement class to Citation Information Class – Myche McAuley - Talk about in MC F2F (?)
- CCB-358 – Context object for Instrument Package – Anne Raugh - Up for debate
- CCB-359 – Allow Product Ancillary to include dictionaries – Anne Raugh
- ****(Discussed)****

Tier 2 issues

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

Parking lot – need a driver!

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

Next meeting: 1/5/23

#DDWG

First meeting since the F2F where product annex and observational video were the main topics. They are the main issues for this next build.

SCR freeze is February 3. DDWG deadline.

~ Not a ton more time in terms of meetings. One in December.

~ Sheri Loftin is now sitting in in an effort to implement changing IM with training efforts.

Regarding the freeze date - the date is for the DDWG. We too often wait for the last minute. Trying not to do that. It's hard to get a bunch through the CCB at the last minute.

~ Critical one, annex, will probably take until the end.

Another new topic from the F2F is product errata. There's a need for tracking things like bug fixes where we don't want to update the entire label version. Hopefully will be a supplemental label similar to product supplemental. Can be turned into a more concrete proposal in the future.

~ We crate products in two ways - concrete where we are not proliferating product types and where we do have new product types. That's probably where this is going.

~ We need to think about that.

Something that came up is issue 561. Jordan was encountering an issue with validate tool where validate incorrectly enforces file naming requirements on bundles and collections. Bundle must be called bundle blank blank, etc. Not in the SR, so he thought that requirement did not exist, but it does. Wants us to improve the SR to clarify or remove that. If anyone has more information on this - let Mike know.

(S. McLaughlin emailed Mike and Debra after the meeting. She wrote:

"Regarding <https://urldefense.us/v3/>

[PDS/validate/issues/561_!!PvBDto6Hs4WbVuu7!aLYEeXh8Ls1vWFncKIFHK5MtCENpK5eVxpzQefXRHx9z7_kwWJxn6Ww8diGbo5nvw1jUf891akA\\$, bundle/collection names, I confirmed Chapter 2B "Archive Physical Org" of the current SR specifies the naming conventions for bundle and collection products. I probably should have put this in Chat. Perhaps Deb can add it to the minutes? Thanks!!"\)](https://pds-jira.jpl.nasa.gov/browse/CCB-357)

#Tier 1 issues

CCB-357 - Create Product Annex

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

Good discussion at the F2F. Updated notes are in jira.

Question to Lyle: Any thoughts on next steps or a time line? We want it ASAP. Lyle is the point person. Want to make sure we are all on the same page.

Answer: Want a decision on what it's going to be called today. Everything going forward follows that. Haven't heard a better name than annex yet. Will incorporate the comments. One is that the whole point of doing this is to expand the IM for either the PDS or PDE. If we leave parts for others to do we lose control over it.

~ We maybe wouldn't have governance.

~ Someone really likes urn:nasa:pde.

~ We need to outline what it will start with. For the ATMOS model, was thinking ama. There are a number of ways we can go about it.

~ The reason to embrace this is our chief scientist is a PDE scientist.

~ Sure. At this point curious to hear from others.

~ Someone put some comments in jira. Wasn't at the F2F. Unclear why PDS is putting effort into this if it's for the PDE. Can let them figure it out. Don't want product annex to imply it belongs in the PDS archive. Have product native. Thought the reason to do this is to reference from PDS. Seems like urn:nasa:pde would do it.

~ There is no available class structure for somebody else to operate from. We want it to work with the PDS4 registry.

~ All we need is a LID or something equivalent.

~ Decided PDS4 can define this product, but not put it in the archive.

~ Product observational to meet our standard.

The real question is about interoperability. Jordan was saying we should be compatible. We can design the bare bones product and then use an LDD. Wasn't clear that PDS4 is really the way to go. Management needs to clear it up. We're stuck trying to define it.

~ Interoperability is just a reference. Bringing it into PDS and using our tools is a bigger order.

~ We need to know what interoperability means to them. Basic principal.

~ We have lots of different class structures. This is beyond having an XML file with a LID. Ultimate idea is no one knows what they want yet. We can create what we want for our annexes and offer it up to the PDE.

~ Someone wants to create product annex labels for the ATMOS annex. Prefers to have this group's oversight to doing it on his own.

~ Someone else didn't think it was that unclear at the F2F. They want a way to expand discoverability beyond PDS. We have already created infrastructure. They would like to piggy back on that. Been thinking we are too loose with saying PDS4. It's an adjective. There is the PDS4 information model, and there is the PDS4 archiving standard. We are talking about archiving standards now.

- ~ Good point. Lots of ambiguity with the term PDS4.
- ~ Another person completely agrees. Thinks we are dealing with search, so products don't disappear. Finding products is critical.

To someone, regarding annex, doesn't matter if outside PDS, like Zenodo or ATMOS annex. Same thing. If we want annex like PDS we can do that, but it's not part of the archive even if we are using archiving standards.

- ~ For discoverability need a registry.
- ~ Discussed minimum of what we want at the F2F.
- ~ Sounds like Dublin Core.
- ~ But we want it available to harvesting tools.
- ~ So streamlined PDS4 label, but if not an array or table, same problem.
- ~ We have it with browse, etc.
- ~ Intention was to use something similar to product browse or document which require less infrastructure. Picture an enumerated list of formats. If trying to push outside PDS standard than maybe don't need that.
- ~ Someone thinks enumerated list is outside our governance.

We need to come up with a name and new class structure adjacent to the product observational tree. Hoping to find a way. Maybe do some refactoring of the tree. Need a way to keep things separate so it's clear where the tree is meeting the archiving standard. That would help.

Question: How do we need to modify this SCR to get it done by February?

Answer: Need an example, name and proposal for governance.

- ~ Need to understand the interoperable boundaries.
- ~ That can grow over time. Start with a LID for now.
- ~ Can say product x - one attribute - a LID and can't use it in any PDS4 archives. Can grow from there. Allows us to grow. Keeps most governance outside PDS4.
- ~ Someone thinks it needs more.
- ~ Need more than just a LID. Think what's in the identification area is a good list of what should be there. People can shoot arrows, trim it down.
- ~ We will treat the description as the proposal. Hash out issues. Can go with the name annex until we hear better.
- ~ Someone likes product ecosystem.
- ~ Someone else says it's weird, but let it sit. Has it's points. Known in the community.
- ~ The problem is that PDS is part of the ecosystem.
- ~ Could have annex be a subset of product observation and part of the ecosystem.
- ~ We are trying to say product not PDS.
- ~ Product external.
- ~ That has meaning in PDS. Defined meaning.
- ~ Someone likes it.

Question: What if it's in PSA? Only external to PDS? There are internal products in every system.

Answer: But outside the PDS4 standard. Our discoverability system.

- ~ Someone can see a future where PDS is a minority user of this IM.
- ~ Someone else almost likes PDE and PDS4 is a member, and ecosystem ATMOS is it's own tree.
- ~ Another person likes ecosystem best despite PDS not being outside it. We need to pick something.

Question to Lyle: Okay with ecosystem?

Answer: Like external better.

~ Let's do a quick straw poll. Can put opinions in chat.

~ Not everyone remembers how to use chat. Someone doesn't like either one. Ecosystem is too broad, and external doesn't make sense.

~ We can change it in the future.

~ Before the SCR is passed.

From the votes in chat, it looks like the majority likes ecosystem. We will go forward with that. The ticket will be updated and we will work on this next month.

~ ****Action Item - Everyone**** - go after Item 1, structure, in jira. Make comments.

We can pull Moses in once we have aligned behind an idea.

~ We can invite Moses two meetings from now.

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

Comments added in jira.

(See Jira for F2F notes added.)

No strong objection. There was a strawman vote. Becky was concerned, but for it. Wanted reassurances. Moses asked about loss-less codec available - hard to use - requires a specific color domain. Not normal, not viable. Did the research. In the white paper, Becky wanted proof about the loss. Show the data wasn't crazy crushed. Will be a lot of spacecrafts using compressed with product native. Will try to update the white paper, then we will be ready to vote. If it doesn't pass we can loop back.

We are voting on a standard not the white paper.

~ Right.

~ We need specific examples of changes to the IM. Need to add Requested Changes.

~ Reporter will try to make it all more clear. Doing what's under document for product observational.

~The Proposed Solution section needs the actual changes - that product observational will have a video option.

~ Reporter will fill it all out.

~ Someone thinks we will also need changes to the SR to clarify this is only format we allow. Show difference between compressed and uncompressed.

We can review this next time. Maybe vote after.

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

We might need to punt this to the next cycle. Still a big discussion to have. Still no dire need for this. Will move to tier 2. Will keep moving it along, working on examples.

The more someone thinks about using LBLX, the more benefit they see. Interested to see when the first person uses it.

~ PPI has some. Not sure if it's released yet.
~ Excellent.

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

No updates on this. Still need to finish research on the changes to verify that none will make the ability to create CDF/A go away.

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

We did not discuss this at the F2F.

~ Simple request.

~ It comes down to formalizing it, at the end of in a document or a flag. May come down to searchability of award number.

~ We need to consider if there are other things that would fit into this.

Action Item - Trent - will touch base with Mike on this. We can probably vote soon.

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

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

From Anne. Not everyone agrees this is appropriate. Somewhat contentious, but maybe we should just vote.

Question: Comments?

Answer: Someone is worried about the impact. Would like to hear from Jordan and how it effects search. Same issue as with composite.

~ Doesn't fit real well with new website either.

~ Someone would ask Jordan how it compares to APSS he's doing. If that's working there's no need to do it.

~ APSS products have it in the LID.

~ If the whole point is search, if Jordan's search incorporates the aliases then it doesn't matter. Should return the same answer.

~ Search still doesn't exist, but we will run this by him.

CCB-359 - Product Ancillary does not allow inclusion of either a Discipline Area or a Mission Area

See <https://pds-jira.jpl.nasa.gov/browse/CCB-359>

From Anne.

No real discussion. Sounds relatively uncontroversial.

~ Bug fix.