

From: Debra Kazden <dkazden@igpp.ucla.edu>

Sent: Monday, May 09, 2016 4:48 PM

To: pds4ddwg

Subject:PDS DDWG notes from 2016 05 05

title: DDWG Notes 2016-05-05

layout: default

date: 2016-05-05

May 5, 2016

Notes by Debra Kazden

Known Attendees:

R. Chen, M. Gordon, E. Guinness, S. Hardman, L. Huber, S. Hughes, R. Joyner, D. Kazden, J. Mafi, T.

Morgan, L. Nagdimunov, L. Neakrase, J. Padams, C. Phillips, A. Raugh, E. Shaya, R. Simpson and J. Stone

Meeting Agenda and Summary

1) CCB/SCR Statuses

-- CCB-100: Remove Array 2D and Array 3D from File Area. (T.King)

-- status of CCB vote

(Not Discussed)

-- CCB-132: Units of Map Scale Improperly includes pixel/deg as a unit (J. Padams)

-- scheduled to be reviewed by CCB next week

(Not Discussed)

-- CCB-158: Restore J2C as valid format for supplemental data. (S.Slavney)

-- scheduled to be reviewed by CCB next week

(Not Discussed)

-- CCB-143: Validate field format via regex (Lev Nagdimunov)

-- scheduled to be reviewed by CCB next week

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

2) Task Statuses - 5 Minutes each

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

3) SCRs and Issues to Discuss:

-- CCB-65: Need additional Target Identification/type values (A.Raugh)

-- URGENT - enhancement / improvement

-- Open:

(1) Needs Proposed Solution

(2) Needs Requested Changes

-- 20150730: DDWG -- Anne to think about working the solution;

-- 20150813: formed WG: J.Mafi, Ed.G, A.Raugh, RJ

-- 20160225: Anne presented 4 questions to DDWG; will update JIRA with consensus

-- 20160324: Anne posted solution to JIRA; EN to review - done

-- 20160329: email to Steve to TA; then back to DDWG for review / discussion ?

****(Discussed)****

-- CCB-77: Augment Product Update with File Area Update - S.Hughes

-- Open: under DDWG discussion

-- has been TA'd

-- 20141002: There is now a tiger to work Update in general that will start in a few months

-- 20150519: Waiting for M.Gordon ?

-- 20150922: DDWG discussion topic; SCR needs to be updated by Mitch

-- 20160324: Mitch prefers to supersede this SCR and add new SCR

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

-- CCB-125: The bit mask attribute seems to be misplaced and possibly missing where needed (A.Raugh)

-- 20150915: Open; needs DDWG discussion

-- 20151008: Jordan to provide example label that uses bit mask

-- 20160323: WG: J.Padams, R.Simpson, A.Raugh, R.Joyner

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

-- CCB-131: Missing constraint on Special Constants attributes (A.Raugh)

- 20150922: Open
- 20160223: under DDWG discussion
- 20160322: EN governance; will take lead
- ** (Not Discussed) **
- CCB-132: Units of Map Scale Improperly includes pixel/deg as a unit (J. Padams)
 - 20151007: Open
 - 20151007: Email to Jordan to provide explicit changes to IM
 - 20151008: I updated SCR to include specific changes required; ready for Steve to TA
 - 20151012: TA'd; email to Emily and Dick to review
 - 20151013: Email from Jordan to pull back for further discussion / work
 - 20151022: Jordan to finalize new & improved proposal before next DDWG
 - 20151104: Jordan updated SCR as comment in JIRA
 - 20151105: DDWG agreed to send to CCB
 - 20151116: Needs TA; then ready for CCB per DDWG
 - 20151117: TA'd by Steve; Emily reviewed; Dick sent email with concerns
 - 20151118: Set up telecon to discuss Dick's concerns
 - 20151119: DDWG discussion; send to CCB if no comments
 - 20151123: Needs TA; then ready for CCB per DDWG
 - 20151202: Emily and Dick reviewed -- Ready
 - 20151208: CCB e-vote; rejected
 - M.Showalter proposed name changes:
 - Units of Map Pixel Resolution
 - Units of Map Pixel Scale
 - CCB will have telecon to discuss
 - 20151222: CCB sent back to WG / DDWG to re-work
 - 20160204: MC on 2016-02-04, Jordan et al violently agreed on a workable solution
 - 20160428: updates discussed by DDWG; sent to CCB
 - ** (Not Discussed) **
- CCB-133: Special Constants class precludes the ability to specify multiple invalid/missing constants (J.

Padams)

-- 20151012: Open

-- 20151021: Under DDWG review

-- 20151022: WG -- Jordan, Steve and RJ; sent email to WG with proposed changes

-- 20151105: Jordan -- special constants needs to be specified per "band" not per "axes"

** (Not Discussed) **

-- CCB-138: Mismatch between context object types and values of type in Observing System Component class (A.Raugh)

-- 20151202: Open; under DDWG review

-- 20151203: WG: Anne, Steve, Dick, Jordan, and RJ

-- 20160310: until someone volunteers to lead the effort -- on hold

** (Not Discussed) **

-- CCB-142: Create Data Quality Flags to hold metadata on Quality Flags (E.Shaya)

-- 20151229: Open;

-- 20160126: Under DDWG review

-- 20160322: Ed didn't like Simpson's CCB-142 implementation

-- Ed wants a lot of specifics embedded into XML

-- Simpson trying to figure out how to make it 'simpler'

-- 20160323: Simpson generated presentation for DDWG review / comment

-- tabled until next session (20160410)

-- 20160428: updates discussed by DDWG

-- requires IMG & others expertise to carry forward

-- 20160505: discussed by DDWG

** (Discussed - work group formed) **

-- CCB-143: Validate field format via regex (Lev Nagdimunov)

-- 20160210: Open & Under DDWG review

-- 20160322: EN governance; will take lead

-- 20160428: updates discussed by DDWG; sent to CCB

-- additional constraints in CCB-142; outside of this SCR

****(Discussed CCB-152)****

- CCB-144: Some examples in Examples collection are incorrect or out-of-date (Lev Nagdimunov)

-- 20160210: Open

-- 20160322: EN governance; will take lead

-- 20160323: may be augmented by CCB-155

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

-- CCB-149: Should PDS4 allow packed data? (E.Shaya)

-- 20160309: Open & Under DDWG review

-- 20160310: Sent email to E.Shaya asking that he upload his version of the IM for packet data to JIRA

-- DDWG will review and provide comments

-- PPI has volunteered to attempt to convert a PDS3 product using the Packed Data class

-- 20160322: dependency on CCB-153; and vice-versa

-- 20160428: sent email to J.Mafi for status on providing examples

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

-- CCB-151: Bundle Member Entry and Internal Reference do not require either LID or LIDVID. (A.Raugh)

-- 20160309: Open & Under DDWG review

-- 20160322: EN governance; will take lead

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

-- CCB-152: field format definition mismatch between IM and SR. (L.Nagdimunov)

-- 20160309: Open & Under DDWG review

-- 20160322: EN governance; will take lead

-- 20160428: dependency on CCB-143

****(Discussed)****

-- CCB-153: SR Needs Additional Description of Packed Data Fields. (E.Shaya)

-- 20160321: Open

-- 20160322: dependency on CCB-149; and vice-versa

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

-- CCB-154: Promote a Mission Information class to Discipline Governance Level. (S.Hughes)

-- 20160321: Open & Under DDWG review

-- 20160428: updates discussed by DDWG

-- formed WG: Steve, Anne, Jordan, Joe

(Not Discussed)

-- CCB-155: Need "Example Set" to include program test data. (A.Raugh)

-- 20160323: came from discussion of CCB-144

-- 20160323: Open; request to provide additional examples; to include 'test data'

(Not Discussed)

-- CCB-156: Inconsistent Discipline Dictionary Technique for Local Internal Reference, et al. (A.Raugh)

-- 20160418: Open

(Not Discussed)

-- CCB-157: Remove sampling parameters attribute from Uniformly Sampled. (L. Nagdimunov)

-- issue from CCB-128:

-- email from Lev requesting IM be changed before v1600 released

-- to make the following change:

-- propose everything be left alone, except sampling_parameters be removed.

-- how to document that EN made the change (outside of the scope of the original SCR)?

-- 20160418: Ready; sent to CCB

-- 20160419: Queued for Implementation

-- E-vote PASSED: 6 Yes (ATM, GEO, IMG, IPDA, PPI, RMS), 1 Failed to vote (SBN)

(Not Discussed)

-- CCB-158: Restore J2C as valid format for supplemental data. (S.Slavney)

-- 20160426: Open & Under DDWG review

-- 20160428: updates discussed by DDWG; sent to CCB

(Not Discussed)

-- CCB-159: Bug fixes for Version 1.7.0.0. (J.Hughes)

-- 20160426: Open & Under DDWG review

(Not Discussed)

4) Topics for Discussion

-- Proposal: CCB-1xx: Remove Enumerated List from Instrument.type (L.Huber)

-- Status & develop implementation plan

(Not Discussed)

-- SETI Issues (R.Simpson et al)

-- Status

-- Issues in XLS have been vetted by SETI notes

-- Issues to be "consolidated" & prioritized

(Discussed)

-- IPDA PDS4 Project: 2014-2015 Final Report (S.Martinez, S.Hughes)

-- Status & develop implementation plan

(Not Discussed)

Notice sent before the telecon in email from R. Joyner - May 3, 2016

There are three enclosures:

1. a list of the full topics under discussion by the DDWG
2. SETI Issue -- Reconciling Local Data Dictionary Versions with New System Builds
3. SETI Issue - Use of XML Catalogs in Archived Bundles

This week's agenda will focus on the following topics:

-- CCB-65: Need additional Target_Identification/type values (A.Raugh)

-- URGENT - enhancement / improvement

-- Open:

(1) Needs Proposed Solution

(2) Needs Requested Changes

-- 20140930 from A.Raugh: I thought Steve and I were supposed to work out how to fix this by eliminating inconsistent

taxonomies and modifying the model to allow multiple types (2 or 3 DDWG telecons ago).

But I'm not available until January, so I wasn't expecting to do anything about this

prior to that.

Have priorities changed? My available time has dramatically decreased.

- 20150730: DDWG -- Anne to think about working the solution;
- 20150813: formed WG: J.Mafi, Ed.G, A.Raugh, RJ
- 20160225: Anne presented 4 questions to DDWG; will update JIRA with consensus
- 20160324: Anne posted solution to JIRA; EN to review - done

- CCB-142: Create Data Quality Flags to hold metadata on Quality Flags (E.Shaya)
 - 20151229: Open;
 - 20160126: Under DDWG review
 - 20160322: Ed didn't like Simpson's CCB-142 implementation
 - Ed wants a lot of specifics embedded into XML
 - Simpson trying to figure out how to make it 'simpler'
 - 20160323: Simpson generated presentation for DDWG review / comment
 - tabled until next session (20160410)
 - 20160428: updates discussed by DDWG
 - requires IMG & others expertise to carry forward
 - 20160505: discussed by DDWG; RSimpson presentation

- CCB-152: field format definition mismatch between IM and SR. (L.Nagdimunov)
 - 20160309: Open & Under DDWG review
 - 20160322: EN governance; will take lead
 - 20160428: dependency on CCB-143

- SETI Issues (R.Simpson, etal)
 - Reconciling Local Data Dictionary Versions with New System Builds
 - Use of XML Catalogs in Archived Bundles

DDWG Telecon

This could be a very short telecon. Steve has been called away at 10:00a.m.

CCB-65 - Need additional Target Identification/type values See <https://pds-jira.jpl.nasa.gov/browse/CCB-65>

Unsure where we are on this. In March, a solution was posted to jira. It hasn't been TA'ed yet.

Question: Are people happy with this?

Answer: As far as we know - done.

Question: Anyone have issues with the proposed solution?

Answer: Crickets.

EN will TA this and hopefully the DDWG can vote next time.

CCB-142 - Create Data Quality Flags to hold metadata on Quality Flags See <https://pds-jira.jpl.nasa.gov/browse/CCB-142>

Question: Is Ed Shaya on? He was invited.

Answer: He knows about the call.

~ Someone told him twice.

We will come back to this when he is on the call.

CCB-152 - field format definition mismatch between IM and SR See <https://pds-jira.jpl.nasa.gov/browse/CCB-152>

This follows on from last week.

There are four things - two trivial and two not trivial. The first trivial thing is that the IM implies all POSIX string formats are allowed - proposed just a subset - specified by the SR. The second trivial thing is that SR has its limitations on what's valid - upper-case E is not valid, lower-case e is valid - there's a separate section for ASCII real in SR - says both are okay. So, the solution would be to add upper-case E to field format. The bigger changes are that it's not clear what the purpose of field format is and how it is to be used - has real world effect - can raise errors due to conversions of strongly typed languages (e.g. Java, Python, FORTRAN). Need to decide if we want to allow that. There's text in the SR that suggests field format is for validation- not to print a different way from in the file- need to decide one way or the other. It implies validation, but isn't explicit. Need to spell it out.

Question: Any questions?

Answered with a question: Was that two big things?

Answer: Yes. 1 - specifiers must match and 2 - length and precision must match for validation.

Question: Thought that format would specify what's in the data file, but that doesn't preclude leading blanks, right? They are still permitted?

Answer: Yes. Regexp isn't a data validator. It does not preclude leading blanks.

Question: Concerns? Comments?

Answer: (Silence)

The SCR was heavily modified last week

The plan going forward is that Steve will TA ****(Action Item)**** it and then we will vote next week.

Back to CCB-142

Ed Shaya is here now.

The idea here is to put in data quality flag somewhere as metadata. Data quality arrays have been around quite some time - a lot in PDS3, but added as second data array and then users just had to know what to do with them. We could do better in PDS4 with describing them. Every array (image) has lots of pixels - things go wrong - projects make arrays the same size to indicate what pixels have problems - so every image has a data quality array. If 64 things can go wrong then create 64 possible flags of what can go wrong - suggests a little table to tell what every flag means. The idea was added that it should have a special name so we know it's a data quality array, so we know it's flags, in XML description would be a description of what each bit means. A comment was added to CCB-142 this morning of how this could work in the XML. Data array, array data quality and then a little section with the flags - what the flag values equal and descriptions. Thinks this is all that needs to change in the specification. For IDL code, this would make it easy to add to a reader - what things are bad according to what flags are bad - or could be separate stand alone command to apply data quality array.

Someone recommends that people look at the CCB. Has another - different - suggested solution.

We need a work group to decide how best to implement. Work group should include a person from IMG. Not sure what implementation is better.

Question: This doesn't just affect images - where are the draft implementations?

Answer: Jira.

~ Can't find them...

~ Bottom of comments - code samples.

Someone agrees that we need a work group to discuss the two implementations.

~ The two implementations are not that different. The work group also needs to focus on if we want this.

~ Thought last week we were leaning towards this being necessary.

~ People were interested, but not sure if it is worth the work. People were going to decide if we need this.

~ That's right.

Question: Work group - who wants to play?

Answer: Steve, Anne (if there's going to be a quick wrap up), Lev, Jordan and Chris.

Someone wants to comment - kinda don't like the approach - separate array. Think it would be better to describe relationships of arrays separate from the physical description.

~ Someone said this is to add metadata, but this is in a sense data. So, agrees with folks who say we can do this with tools we already have.

~ Someone agrees. Useful.

To someone, metadata is the flags - data quality array is sort of data - not really changing that, values are in data file - just saying description of the meaning of the flag is metadata.

~ True.

~ We have other bit flags in other contexts.

Question to Joe: Could you write up your concerns for the work group if you aren't going to be on it? To make your issues clear?

Answer: Yes - will write a paragraph. Not on work group. ****(Action Item)****

Question: I'm on the CCB - so what's the role of the CCB here? Thought they would decide if we do or do not do this. Thought this group just came up with possible solutions?

Answer: The CCB has the final decision, but this group makes suggestions, possible implementations etc.

~ Someone is concerned that a CCB member doesn't know what we do.

~ There's a CCB document - in limbo, but it lays it all out.

(Steve has left the call)

SETI Issues - Use of XML Catalogs in Archived Bundles See attachment to telecon announcement email - May 3, 2016, from Ron Joyner, Use_of_XML_Cata...hives_20160330.pdf

Someone hesitates to go over the SETI issues without Steve.

When someone read the XML catalog attachment was shocked. Thought we can't do this. We aren't giving data providers a complete environment to make their labels. XML catalog files do not belong in the archive. Thought we killed this at SETI. We are just able to suggest best practices.

~ You read this different from me - Thought it said when a data provider submits data - all has to resolve without referring to an XML catalog.

Question: Can we have someone re-write these? I can't understand what this says...

Another Question: Are these raw notes from the meeting at SETI?

Answer: No, this was produced at EN.

Question: In the attachment, there are five bullets - was the intent to specify requirements? What's the intent? That can't resolve to catalog file...

Answer: All five bullets are there as best practices - and put a shall by each, but use of catalog file could negate them all.

~ Catalog file is for local environment.

~ Yes, if under development use a catalog file, but when ready for NSSDC don't think there's room for a catalog file.

This was the only discussion that came to a conclusion at SETI.

Confusion.

Concern that catalog file users would validate and then change things.

Someone lost the plot here. Understands that schemaLocation is not an XML requirement.

~ Need to resolve by full namespace URI declaration. That always has to be there. Have to validate against that conical namespace regardless of anything else in the label.

Question: Eclipse doesn't use catalog file? If you want a local version of a file you specify that under schemaLocation - give that when archiving - but it's not useful then?

Answer: It's a hint.

~ Someone takes issue with that. In our implementation, taken to the extreme, it's a restricted value the way we use it.

~ We can't provide data providers with an environment so can't make requirements for processing hints that are not conformative with the standard.

~ We are exceeding the standard.

~ Can't require something that's not required in the standard.

~ This is a best practices implementation. It doesn't invalidate anything.

~ It's telling users to change validated data and then submit it.

Question: What if they data providers are using a processing environment that doesn't implement XML catalog?

Answer: Has been discussed with Sean. We're not in favor - can get in a lot of trouble using catalog files.

Another Question: Suggesting we use schemaLocation?

Answered with another question: So what if data providers reference a local thing?

Answer: From archive, should only be looking at the namespace.

We're not even on the same planet on this.

~ Right. This is trying to put a requirement on something that should just be a hint. SchemaLocation has nothing to do with the permanent archive. All you need is IM number and namespace.

Question: Anyone else?

Answer: We need a work group.

Question: Can I say what I think was the agreement - resolution of archive product should be irrelevant to the catalog file, yes? Someone says schemaLocation should be ignored by anyone looking at a label that's been archived ...?

Another Question: The value should resolve to EN. We should accept a note that says schemaLocation was for development only. It doesn't invalidate it, does it?

Answer: That's an option. Can force schemaLocation or override - there's minimum implementation for using the catalog file. Does not mix and match. Two options ignore schemaLocation, one uses it - uses XML catalog to specify schemas in labels.

~ So validator doesn't validate the URIs.

~ Namespaces can be anything.

~ URIs should resolve.

If you pass in the schema - validator will check schema to namespace or only using schema in schemaLocation with the namespace specified in the label. We are checking that schema matches namespace in the label.

~ Someone thinks - the thing it doesn't do - it should go to the web, pull out the current schema for that namespace and try to verify that.

~ That won't work for most labels cause most are not at the current version.

~ That's another issue.

Someone needs time to consider all of this. Moves we delay this discussion.

~ That works for someone else.

~ Someone would like to move on to the next topic - LDDs.

SETI Issues - Reconciling Local Data Dictionary Versions with New System Builds See attachment to telecon announcement email - May 3, 2016, from Ron Joyner, Versioning_of_D...aries_20160330.pdf

The next issue is LDDs - about regenerating local data dictionaries with each new release of the IM. We originally just said that we needed to validate that they were still valid. Don't want us to drop the discussion. Also, (something about file names) - internal file version. Need file name to include IM and LDD version. Not sure if we should discuss this now or not, but doesn't want this to go away. Would like to discuss next week if not now.

Comment - Someone is wondering if anyone at EN is considering the scenario about mission DDs - they archive to a particular IM, but the LDD may have to update.

~ This is about mission DDs.

Question: Isn't this driven by a problem driven by a mission DD? But not changing core version..?

Answer: If they signed up for 1.5.0.0, they can build to 1.5.0.0 - committed to that. But if building a DD update that needs an update from the IM - would make two versions so they could have support at the discipline level without more work at the mission level.

Question: Stewards were supposed to be responsible, but should they all be ingested in the IM, then when push the button it would get spit out at a new version - all in sync to the next build. So, what if EN took over the responsibility of keeping it all in sync? Stewards would continue to be in charge of content. There are pros and cons to both, but if everything was ingested, then everything would be plopped out. It would be a change to the process.

Answer: Someone just doesn't understand why EN would need to be so involved with node DDs. Could be a nightmare.

Another Question: Would it involve higher requirements on vetting?

Answer: Yeah, if we would have to go to the CCB - don't want that.

~ CCB doesn't care about LDDs.

~ But if they are incorporated into the IM that could be an issue for discipline DDs, but we all appreciate ENs offer to take the administrative work.

One comment on LDDs - There was a recent change to an LDD - not sure the stewards knew it was

changed. Even with stable ones - things break cause there's no oversight.

~ Need to report the problem to the steward.

~ Don't know who that is. The problem is in the Display DD.

We need a consistent policy.

~ Just wanted to point out that stable things still sometimes change and break cause of no oversight.

Question: Comments?

Answer: It's confusing to users if version numbers change, but there is no other change.

~ Agreement.

When we were designing this years ago, it was someone's understanding that versions of DDs were going to stay independent of the version of the I.M. version.

~ We didn't fully understand the full implementations of schema and XML references. That surfaced with issues in the MAVEN DD. That was unpredicted.

~ The solution was that the LDD shouldn't refer to the IM , so no namespace collision.

Someone has been concerned about all this. Has been working on possible solutions. It's a complicated application of a complicated standard. Working on mock ups. Thinks there's a good solution - not a quick, easy change. But thinks it's a better way - needs to make sure it can handle this.

~ So, we'll wait to see the mock ups.

~ It's complicated. Will take a few weeks.

~

So, next week we have two SCRs maybe that mission thing and the ever popular enumerated list for targets stuff.

~ Final changes are hopefully trivial for the enumerated values for the target list - CCB-65.

~ That's going to be TAed. There are four things for next week.

Question: What about enumerated list for instrument type?

Answer: Meant that - not target type.

Meeting again next week.